SERIE SCHAUM

Simon Plouffe
29 avril 1979

SEYMOUR LIPSCHUTZ
Temple University

ALGEBRE LINEAIRE

COURS ET PROBLEMES

Nouveau tirage

EDISCIENCE

Groupe McGraw-Hill :
Paris, New York, Londres, Saint Louis, San Francisco, Dusseldorf
Johannesbourg, Madrid, Mexico, Montréal, New Delhi, Panama,
Rio de Janeiro, Singapour, Sydney, Tokyo, Toronto
1977

McGraw-Hill

ALGEBRE LINEAIRE, Cours et problèmes, est traduit de :

Theory and problems of Linear Algebra, by Seymour Lipschutz

ISBN France : 2.7042.0001.7
ISBN Canada : 007.084.394.5

· McGraw-Hill Inc. 28 Rue Beaunier. 75014 Paris

ALGEBRE LINEAIRE

Table des matières

Remarque

On trouvera dans ce livre l'abréviation suivante : "ssi" qui signifie "si et seulement si".

Préface

Depuis quelques années, l'algèbre linéaire est devenue une partie essentielle du bagage mathématique nécessaire aux ingénieurs, physiciens et autres scientifiques. Ce besoin reflète l'importance et les applications étendues du sujet.

Ce livre est destiné à être utilisé comme manuel pour un cours d'algèbre linéaire ou comme supplément à d'autres ouvrages. Il vise à présenter une introduction à l'algèbre linéaire qui sera utile à tous les lecteurs quelle que soit leur spécialisation. Il est inclus plus de matière que l'on en peut insérer dans la plupart des premiers cours d'algèbre linéaire. Ceci a été fait dans le but de rendre l'ouvrage plus souple, de fournir un livre de référence utile et stimuler l'intérêt porté à cette matière.

Chaque chapitre comprend des énoncés clairs de définitions de principes et de théorèmes, des éléments d'illustration et de description et des exercices progressifs résolus. Ces derniers servent à illustrer et à amplifier la théorie, à mettre au point de façon précise les passages délicats, sans lesquels l'étudiant se sent constamment sur un terrain incertain, et à permettre la répétition des principes fondamentaux, si essentiels à une étude efficace. De nombreuses démonstrations de théorèmes illustrent les problèmes résolus. Les problèmes supplémentaires permettent une révision complète de chaque chapitre.

Les trois premiers chapitres traitent des vecteurs dans l'espace euclidien, des équations linéaires et des matrices. Ceux-ci fournissent la motivation et les instruments de calcul fondamentaux pour l'étude abstraite des espaces vectoriels et des applications linéaires qui suivent. Un chapitre sur les vecteurs propres et les valeurs propres, précédé par les déterminants, donne les conditions permettant de représenter un opérateur linéaire par une matrice diagonale. Ceci conduit naturellement à l'étude de formes canoniques et spécialement des formes canoniques triangulaires, de Jordan, et des formes canoniques rationnelles. Dans le dernier chapitre sur les espaces hermitiens, le théorème spécial pour les opérateurs symétriques est obtenu et appliqué à la diagonalisation des formes réelles quadratiques.

On termine avec des appendices comprenant des paragraphes sur les ensembles et les relations, les structures algébriques et les polynômes sur un corps.

Je voudrais remercier de nombreux amis et collègues, en particulier le Dr. Martin Silverstein et Dr. Hwa-Tsang pour leurs suggestions et la révision du manuscrit. Je veux également exprimer ma gratitude à Daniel Schaum et Nicola Monti pour leur précieuse collaboration.

Seymour Lipschutz
Temple University

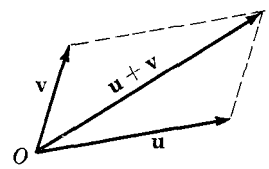
CHAPITRE 1

Vecteurs de $\mathbf{R}^n$ et $\mathbf{C}^n$

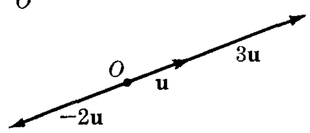
INTRODUCTION

Dans diverses applications physiques apparaissent certaines quantités, comme la température et la valeur absolue de la vitesse qui possèdent seulement "une amplitude". Ces quantités peuvent être représentées par des nombres réels et sont appelées *scalaires.* D'autre part, il existe d'autres quantités comme la force et l'accélération qui possèdent à la fois une amplitude et une direction. Ces quantités peuvent être représentées par des segments orientés (ayant une longueur et une direction données et ayant pour origine un point de référence et sont appelées des *vecteurs.* Dans ce chapitre nous étudierons en détail les propriétés de ces vecteurs.

Commençons par les opérations élémentaires sur les vecteurs.



1) *Addition* : La résultante $\mathbf{u} + \mathbf{v}$ de deux vecteurs $\mathbf{u}$ et $\mathbf{v}$ est obtenue d'après la règle suivante appelée règle du parallélogramme, $\mathbf{u} + \mathbf{v}$ est la diagonale du parallélogramme construit sur $\mathbf{u}$ et $\mathbf{v}$.



2) *Multiplication scalaire* : Le produit $k\mathbf{u}$ d'un nombre réel k par un vecteur $\mathbf{u}$ est obtenu en multipliant la longueur du vecteur $\mathbf{u}$ par la valeur absolue de k, et en conservant le même sens que celui de $\mathbf{u}$ si $k \geqslant 0$, ou le sens opposé si $k < 0$.

Nous supposons que le lecteur est familiarisé avec la représentation des points du plan par des couples ordonnés de nombres réels. Si l'origine des axes est choisie comme point de référence O, comme précédemment, chaque vecteur est déterminé de manière unique par les coordonnées de son extrémité. Il s'ensuit les relations suivantes entre les opérations précédentes et les coordonnées des extrémités de vecteurs.

i) *Addition :* Si (a, b) et (c, d) sont les coordonnées des extrémités des vecteurs $\mathbf{u}$ et $\mathbf{v}$, alors $(a + c,\ b + d)$ sont les coordonnées de l'extrémité du vecteur $\mathbf{u} + \mathbf{v}$, comme l'indique la figure (a).

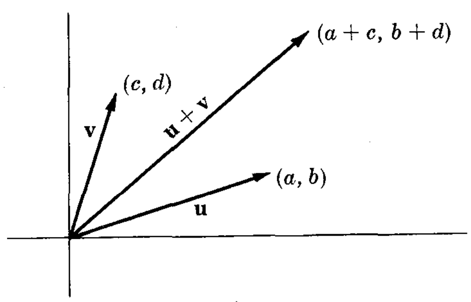


Fig. (a)

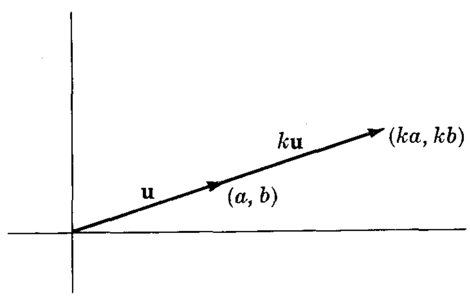


Fig. (b)

ii) *Multiplication scalaire* : (a, b) étant les coordonnées de l'extrémité du vecteur $\mathbf{u}$, (ka, kb) sont les coordonnées de l'extrémité du vecteur $k\mathbf{u}$, voir (Fig. b) ci-dessus.

De façon générale, nous identifierons un vecteur par son extrémité ; c'est-à-dire que nous appellerons vecteur le couple ordonné (a, b) de nombres réels, En généralisant cette notion nous appellerons vecteur un n-tuple $(a_1, a_2, \ldots a_n)$ de nombres réels. En généralisant de nouveau nous supposerons que les coordonnées du n-tuple sont des nombres réels ou complexes. Plus loin, dans le Chapitre IV, nous ferons abstraction des propriétés de ces n-tuples et définirons de façon formelle l'ensemble mathématique appelé *espace vectoriel.*

Nous supposons que le lecteur est familiarisé avec les propriétés élémentaires du corps des nombres réels noté R.

VECTEURS DANS $\mathbf{R}^n$

L'ensemble de tous les n-tuples de nombres réels, noté $\mathbf{R}^n$, est appelé *n-espace* ou espace à n dimensions. En particulier un n-tuple dans $\mathbf{R}^n$,

$$u = (u_1, u_2, \ldots, u_n)$$

est appelé un point ou un vecteur, les nombres réels u_i sont les *composantes* (ou *coordonnées*) du vecteur u. En outre, lorsque nous serons dans l'espace $\mathbf{R}^n$, nous emploierons le terme scalaire pour les éléments de $\mathbf{R}$, c'est-à-dire pour les nombres réels.

Exemple 1.1 : Considérons les vecteurs suivants,

$$(0, 1), \quad (1, -3), \quad (1, 2, \sqrt{3}, 4), \quad (-5, \tfrac{1}{2}, 0, \pi)$$

Les deux premiers ont deux composantes et sont ainsi des points de $\mathbf{R}^2$; les deux derniers ont quatre composantes et sont donc des points de $\mathbf{R}^4$.

Deux vecteurs u et v sont égaux (on écrit $\mathbf{u} = \mathbf{v}$) s'ils ont les mêmes composantes, donc appartiennent au même espace, et si les composantes correspondantes sont égales. Les vecteurs (1, 2, 3) et (2, 3, 1) ne sont pas égaux, puisque leurs composantes correspondantes ne sont pas égales.

Exemple 1.2 : Soit $(x - y, x + y, z - 1) = (4, 2, 3)$. Alors par définition de l'égalité des vecteurs,

$$x - y = 4$$
$$x + y = 2$$
$$z - 1 = 3$$

Le système précédent résolu donne $x = 3, y = -1$ et $z = 4$.

ADDITION VECTORIELLE ET MULTIPLICATION SCALAIRE

Soit u et v deux vecteurs de $\mathbf{R}^n$

$$u = (u_1, u_2, \ldots, u_n) \qquad \text{et} \qquad v = (v_1, v_2, \ldots, v_n)$$

La *somme* de u et v, que l'on écrit $u + v$. est le vecteur obtenu en ajoutant les composantes correspondantes de u et de v:

$$u + v = (u_1 + v_1, u_2 + v_2, \ldots; u_n + v_n)$$

Le *produit* d'un nombre réel k par un vecteur u, écrit ku, est le vecteur obtenu en multipliant chaque composante de u par k:

$$ku = (ku_1, ku_2, \ldots, ku_n)$$

Remarquons que $u + v$ et ku sont aussi des vecteurs de $\mathbf{R}^n$. On définit aussi

$$-u = -1u \qquad \text{et} \qquad u - v = u + (-v)$$

La somme de vecteurs ayant un nombre distinct de composantes n'est pas définie.

Exemple 1.3 : Soit $u = (1, -3, 2, 4)$ et $v = (3, 5, -1, -2)$. Alors

$$\begin{aligned} u + v &= (1+3, -3+5, 2-1, 4-2) = (4, 2, 1, 2) \\ 5u &= (5 \cdot 1, 5 \cdot (-3), 5 \cdot 2, 5 \cdot 4) = (5, -15, 10, 20) \\ 2u - 3v &= (2, -6, 4, 8) + (-9, -15, 3, 6) = (-7, -21, 7, 14) \end{aligned}$$

Exemple 1.4 : Le vecteur $(0, 0, 0 \ldots, 0)$ de $\mathbf{R}^n$, noté 0, est appelé vecteur nul. Il est semblable au scalaire 0 en ce sens que pour chaque vecteur $u = (u_1, u_2, u_3 \ldots, u_n)$,

$$u + 0 = (u_1 + 0, u_2 + 0, \ldots, u_n + 0) = (u_1, u_2, \ldots, u_n) = u$$

Les propriétés essentielles des vecteurs de $\mathbf{R}^n$ par rapport à l'addition vectorielle et à la multiplication par un scalaire sont résumées dans le théorème suivant.

Théorème 1.1 : Quels que soient les vecteurs $u, v, w \in \mathbf{R}^n$ et quels que soient les scalaires $k, k' \in \mathbf{R}$,

(i) $(u+v)+w = u+(v+w)$

(ii) $u+0 = u$

(iii) $u+(-u) = 0$

(iv) $u+v = v+u$

(v) $k(u+v) = ku+kv$

(vi) $(k+k')u = ku+k'u$

(vii) $(kk')u = k(k'u)$

(viii) $1u = u$

Remarque : Supposons que u et v sont des vecteurs de $\mathbf{R}^n$ pour lesquels on a $u = kv$, k étant un scalaire non nul $k \in \mathbf{R}$. On dit alors que u a le *même sens* que v si $k > 0$ et le *sens contraire* si $k < 0$.

PRODUIT SCALAIRE

Soient u et v deux vecteurs de $\mathbf{R}^n$:

$$u = (u_1, u_2, \ldots, u_n) \quad \text{et} \quad v = (v_1, v_2, \ldots, v_n)$$

Le produit scalaire de u et v, noté $u \cdot v$, est le scalaire obtenu en multipliant les composantes correspondantes et en ajoutant les produits obtenus:

$$u \cdot v = u_1v_1 + u_2v_2 + \cdots + u_nv_n$$

Les vecteurs u et v sont *orthogonaux* (ou *perpendiculaires*) si leur produit scalaire est nul: $u \cdot v = 0$

Exemple 1.5 : Soit $u = (1, -2, 3, -4)$, $v = (6, 7, 1, -2)$ et $w = (5, -4, 5, 7)$. Alors

$$\begin{aligned} u \cdot v &= 1 \cdot 6 + (-2) \cdot 7 + 3 \cdot 1 + (-4) \cdot (-2) = 6 - 14 + 3 + 8 = 3 \\ u \cdot w &= 1 \cdot 5 + (-2) \cdot (-4) + 3 \cdot 5 + (-4) \cdot 7 = 5 + 8 + 15 - 28 = 0 \end{aligned}$$

Donc u et w sont orthogonaux.

Les propriétés remarquables du produit scalaire dans $\mathbf{R}^n$ sont les suivantes:

Théorème 1.2 : Quels que soient les vecteurs $u, v, w \in \mathbf{R}^n$ et quel que soit le scalaire $k \in \mathbf{R}$,

(i) $(u+v) \cdot w = u \cdot w + v \cdot w$

(ii) $(ku) \cdot v = k(u \cdot v)$

(iii) $u \cdot v = v \cdot u$

(iv) $u \cdot u \geq 0$, et $u \cdot u = 0$ ssi $u = 0$

Remarque : L'espace $\mathbf{R}^n$, avec les propriétés précédentes d'addition vectorielle et de multiplication *par un réel et le produit scalaire*, est appelé communément *espace euclidien à n-dimensions.*

NORME ET DISTANCE DANS $\mathbf{R}^n$

Soient u et v des vecteurs de $\mathbf{R}^n$: $u = (u_1, u_2, \ldots, u_n)$ et $v = (v_1, v_2 \ldots v_n)$. La distance des deux points u et v, notée $d(u, v)$, est définie par,

$$d(u, v) = \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + \cdots + (u_n - v_n)^2}$$

La norme (ou longueur) du vecteur u est notée $||u||$, et est définie comme la racine carrée positive de $u \, . \, u$:

$$||u|| = \sqrt{u \cdot u} = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2}$$

D'après le théorème 1.2, $u \, . \, u \geqslant 0$ et ainsi la racine carrée existe. Observons que :

$$d(u, v) = ||u - v||$$

Exemple 1.6 : Soient $u = (1, -2, 4, 1)$ et $v = (3, 1, -5, 0)$. Alors

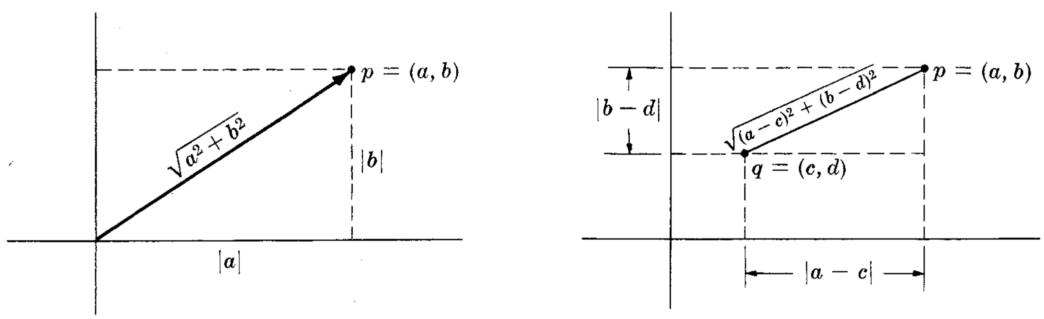
$$d(u, v) = \sqrt{(1-3)^2 + (-2-1)^2 + (4+5)^2 + (1-0)^2} = \sqrt{95}$$

$$||v|| = \sqrt{3^2 + 1^2 + (-5)^2 + 0^2} = \sqrt{35}$$

Si nous considérons maintenant deux points $p = (a, b)$ et $q = (c, d)$, dans le plan $\mathbf{R}^2$, alors,

$$||p|| = \sqrt{a^2 + b^2} \qquad \text{et} \qquad d(p, q) = \sqrt{(a-c)^2 + (b-d)^2}$$

Donc $||p||$ correspond à l'habituelle distance euclidienne de l'origine au point p, et $d(p, q)$ correspond à la distance euclidienne séparant les points p et q, comme ci-dessous:



Des résultats semblables seraient obtenus sur la droite réelle $\mathbf{R}$, ou dans $\mathbf{R}^3$.

Remarque : Un vecteur e est appelé *vecteur unitaire* si sa norme est 1 : $||e|| = 1$. Remarquons que, pour chaque vecteur non nul $u \in \mathbf{R}^n$, le vecteur $e_u = u/||u||$ est un vecteur unitaire de même direction que u.

Etablissons maintenant une relation fondamentale, connue sous le nom d'inégalité de Cauchy-Schwarz.

Théorème 1.3 (Cauchy-Schwarz) : Quels que soient les vecteurs $u, v \in \mathbf{R}^n$, $|u \, . \, v| \leqslant ||u|| \; ||v||$.

En utilisant l'inégalité précédente, nous pouvons définir maintenant l'angle θ de deux vecteurs non nuls $u, v \in \mathbf{R}^n$ par :

$$\cos \theta = \frac{u \cdot v}{||u|| \; ||v||}$$

Remarquons que si $u \cdot v = 0$ alors $\theta = 90°$ ou $\theta = \pi/2$. Ce qui concorde bien avec notre définition de l'orthogonalité.

NOMBRES COMPLEXES

L'ensemble des nombres complexes est noté $\mathbf{C}$. Formellement, un nombre complexe est un couple ordonné de nombres réels ; l'égalité, l'addition et la multiplication de nombres complexes sont définies de la manière suivante :

$$(a, b) = (c, d) \quad \text{ssi} \quad a = c \quad \text{et} \quad b = d$$

$$(a, b) + (c, d) = (a + c, b + d)$$

$$(a, b)(c, d) = (ac - bd, ad + bc)$$

Identifions le nombre réel a avec le nombre complexe $(a, 0)$:

$$a \leftrightarrow (a, 0)$$

Ceci est possible car les opérations, addition et multiplication, entre nombres réels sont conservées par la correspondance précédente :

$$(a, 0) + (b, 0) = (a+b, 0) \qquad \text{et} \qquad (a, 0)(b, 0) = (ab, 0)$$

Donc R est un sous-ensemble de C et nous pouvons remplacer $(a, 0)$ par a toutes les fois que cela est commode et possible.

Le nombre complexe $(0, 1)$, noté i, a la propriété importante suivante :

$$i^2 = ii = (0, 1)(0, 1) = (-1, 0) = -1 \qquad \text{ou} \qquad i = \sqrt{-1}$$

De plus, en utilisant le fait que :

$$(a, b) = (a, 0) + (0, b) \qquad \text{et} \qquad (0, b) = (b, 0)(0, 1)$$

$$(a, b) = (a, 0) + (b, 0)(0, 1) = a + bi$$

La notation $a + bi$ est plus commode que (a, b). Ainsi la somme et le produit de deux nombres complexes peuvent être obtenus en utilisant simplement la commutativité et la distributivité des deux lois et le fait que $i^2 = -1$:

$$(a+bi) + (c+di) = a + c + bi + di = (a+c) + (b+d)i$$

$$(a+bi)(c+di) = ac + bci + adi + bdi^2 = (ac-bd) + (bc+ad)i$$

Le nombre complexe conjugué de $z = (a, b) = a + bi$ est noté et défini par

$$\bar{z} = a - bi$$

(Notons que $z\bar{z} = a^2 + b^2$). Si finalement, $z \neq 0$, alors l'inverse z^{-1} de z et la division par z sont donnés par

$$z^{-1} = \frac{\bar{z}}{z\bar{z}} = \frac{a}{a^2+b^2} + \frac{-b}{a^2+b^2}i \qquad \text{et} \qquad \frac{w}{z} = wz^{-1}$$

où $w \in C$. Nous définissons aussi

$$-z = -1z \qquad \text{et} \qquad w - z = w + (-z)$$

Exemple 1.7 : Supposons $z = 2 + 3i$ et $w = 5 - 2i$. Alors

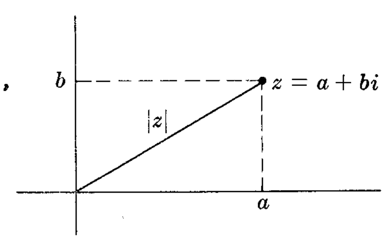
$$z + w = (2+3i) + (5-2i) = 2 + 5 + 3i - 2i = 7 + i$$

$$zw = (2+3i)(5-2i) = 10 + 15i - 4i - 6i^2 = 16 + 11i$$

$$\bar{z} = \overline{2+3i} = 2 - 3i \qquad \text{et} \qquad \bar{w} = \overline{5-2i} = 5 + 2i$$

$$\frac{w}{z} = \frac{5-2i}{2+3i} = \frac{(5-2i)(2-3i)}{(2+3i)(2-3i)} = \frac{4-19i}{13} = \frac{4}{13} - \frac{19}{13}i$$

Comme les nombres réels qui peuvent être représentés par les points d'une droite, les nombres complexes peuvent être représentés par les points d'un plan. En particulier le point (a, b) du plan, représente le nombre complexe $z = a + bi$ dont la *partie réelle* est a, et dont la *partie imaginaire* est b. La valeur absolue de z, notée $|z|$, est définie comme la distance de z à l'origine:



$$|z| = \sqrt{a^2 + b^2}$$

Remarquons que $|z|$ est égal à la norme du vecteur (a, b). Ainsi $|z| = \sqrt{z\bar{z}}$

Exemple 1.8 : Supposons $z = 2 + 3i$ et $w = 12 - 5i$. Alors,

$$|z| = \sqrt{4+9} = \sqrt{13} \qquad \text{et} \qquad |w| = \sqrt{144+25} = 13$$

Remarque : Dans l'Appendice B, nous définissons la structure algébrique appelée *corps*. Insistons sur le fait que l'ensemble **C** des nombres complexes muni des opérations d'addition et de multiplication définies ci-dessus est un corps.

VECTEURS DANS $\mathbf{C}^n$

L'ensemble de tous les n-tuples de nombres complexes, noté $\mathbf{C}^n$, est appelé *espace complexe à n dimensions*. Comme dans le cas réel, les éléments de $\mathbf{C}^n$ sont appelés des *points* ou des *vecteurs*, les éléments de **C** sont appelés *scalaires*, et *l'addition vectorielle* dans $\mathbf{C}^n$, *et la multiplication* par *un scalaire* dans $\mathbf{C}^n$ sont données par :

$$(z_1, z_2, \ldots, z_n) + (w_1, w_2, \ldots, w_n) = (z_1 + w_1, z_2 + w_2, \ldots, z_n + w_n)$$
$$z(z_1, z_2, \ldots, z_n) = (zz_1, zz_2, \ldots, zz_n)$$

où $z_i, w_i, z \in \mathbf{C}$.

Exemple 1.9 : $(2+3i, 4-i, 3) + (3-2i, 5i, 4-6i) = (5+i, 4+4i, 7-6i)$

$2i(2+3i, 4-i, 3) = (-6+4i, 2+8i, 6i)$

Soient u et v deux vecteurs quelconques de $\mathbf{C}^n$:

$$u = (z_1, z_2, \ldots, z_n), \qquad v = (w_1, w_2, \ldots, w_n), \qquad z_i, w_i \in \mathbf{C}$$

Le produit scalaire de u et v est défini par :

$$u \cdot v = z_1\bar{w}_1 + z_2\bar{w}_2 + \cdots + z_n\bar{w}_n$$

Remarquons que cette définition se réduit à la définition du produit scalaire dans le cas réel, puisque $w_i = \bar{w}_i$, lorsque w_i est réel. La norme de u est définie par

$$\|u\| = \sqrt{u \cdot u} = \sqrt{z_1\bar{z}_1 + z_2\bar{z}_2 + \cdots + z_n\bar{z}_n} = \sqrt{|z_1|^2 + |z_2|^2 + \cdots + |z_n|^2}$$

Observons que $u \,.\, u$ et donc $\|u\|$ sont des nombres réels positifs, lorsque $u \neq 0$, et nuls lorsque $u = 0$.

Exemple 1.10 : Soient $u = (2+3i, 4-i, 2i)$ et $v = (3-2i, 5, 4-6i)$. Alors

$$\begin{aligned}
u \cdot v &= (2+3i)\overline{(3-2i)} + (4-i)\overline{(5)} + (2i)\overline{(4-6i)} \\
&= (2+3i)(3+2i) + (4-i)(5) + (2i)(4+6i) \\
&= 13i + 20 - 5i - 12 + 8i = 8 + 16i
\end{aligned}$$

$$\begin{aligned}
u \cdot u &= (2+3i)\overline{(2+3i)} + (4-i)\overline{(4-i)} + (2i)\overline{(2i)} \\
&= (2+3i)(2-3i) + (4-i)(4+i) + (2i)(-2i) \\
&= 13 + 17 + 4 = 34
\end{aligned}$$

$$\|u\| = \sqrt{u \cdot u} = \sqrt{34}$$

L'espace $\mathbf{C}^n$, avec les opérations précédemment définies, addition vectorielle, multiplication par un scalaire et produit scalaire, est appelé *espace euclidien complexe à n dimensions.*

Remarque : Si $u \,.\, v$ était défini par $u \,.\, v = z_1 w_1 + \cdots + z_n w_n$, alors il serait possible pour $u \,.\, u = 0$, même si $u \neq 0$ par exemple avec $u = (1, i, 0)$. En fait $u \,.\, u$ pourrait même ne pas être réel.

PROBLEMES RESOLUS

VECTEURS DE R^n

1.1. Calculer *1*) $(3, -4, 5) + (1, 1, -2)$; *2*) $(1, 2, -3) + (4, -5)$; *3*) $-3(4, -5, -6)$; *4*) $-(-6, 7, -8)$

(1) Additionnons les composantes correspondantes $(3, -4, 5) + (1, 1, -2) = (3 + 1, -4 + 1, 5 - 2) = (4, -3, 3)$

(2) La somme n'est pas définie car les vecteurs ont des nombres différents de composantes.

(3) Multiplions chaque composante par le scalaire : $-3(4, -5, -6) = (-12, 15, 18)$.

(4) Multiplions chaque composante par -1, $-(-6, 7, -8) = (6, -7, 8)$

1.2. Soient $u = (2, -7, 1)$, $v = (-3, 0, 4)$, $w = (0, 5, -8)$. Trouver 1) $3u - 4v$, 2) $2u + 3v - 5w$.

Effectuer d'abord la multiplication scalaire, puis l'addition vectorielle.

(1) $3u - 4v = 3(2, -7, 1) - 4(-3, 0, 4) = (6, -21, 3) + (12, 0, -16) = (18, -21, -13)$

(2)
$$\begin{aligned} 2u + 3v - 5w &= 2(2, -7, 1) + 3(-3, 0, 4) - 5(0, 5, -8) \\ &= (4, -14, 2) + (-9, 0, 12) + (0, -25, 40) \\ &= (4 - 9 + 0, -14 + 0 - 25, 2 + 12 + 40) = (-5, -39, 54) \end{aligned}$$

1.3. Trouver x et y si $(x, 3) = (2, x + y)$.

Les deux vecteurs étant égaux, leurs composantes correspondantes sont égales entre elles:

$$x = 2, \qquad 3 = x + y$$

En remplaçant x par 2 dans la seconde équation on obtient $y = 1$. Ainsi $x = 2$ et $y = 1$.

1.4. Trouver x et y si $(4, y) = x(2, 3)$.

En multipliant par le scalaire x on obtient $(4, y) = x(2, 3) = (2x, 3x)$.

En égalant les composantes correspondantes : $4 = 2x$, $y = 3x$.

Résolvons enfin les équations linéaires en x et y : $x = 2$ et $y = 6$.

1.5. Trouver x, y et z si $(2, -3, 4) = x(1, 1, 1) + y(1, 1, 0) + z(1, 0, 0)$.

Multiplions d'abord par les scalaires x, y et z et additionnons:

$$\begin{aligned} (2, -3, 4) &= x(1, 1, 1) + y(1, 1, 0) + z(1, 0, 0) \\ &= (x, x, x) + (y, y, 0) + (z, 0, 0) \\ &= (x + y + z, x + y, x) \end{aligned}$$

Egalons les composantes correspondantes les unes aux autres, on a :

$$x + y + z = 2, \qquad x + y = -3, \qquad x = 4$$

Pour résoudre ce système d'équations, nous remplaçons x par 4 dans la seconde équation, on obtient $4 + y = -3$ ou $y = -7$. Alors en remplaçant x et y par leur valeur dans la première équation, on trouve $z = 5$. D'où $x = 4$, $y = -7$, $z = 5$.

1.6. Démontrer le théorème 1.1 : Quels que soient les vecteurs $u, v, w \in R^n$ et quels que soient les scalaires $k, k' \in R$:

(1)	$(u + v) + w = u + (v + w)$	(5)	$k(u + v) = ku + kv$
(2)	$u + 0 = u$	(6)	$(k + k')u = ku + k'u$
(3)	$u + (-u) = 0$	(7)	$(kk')u = k(k'u)$
(4)	$u + v = v + u$	(8)	$1u = u$

Soient u_i, v_i, w_i les i^{emes} composantes de u, v et w respectivement.

(1) Par définition, $u_i + v_i$ est la $i^{\text{ème}}$ composante de $u + v$ et donc $(u_i + v_i) + w_i$ est la $i^{\text{ème}}$ composante de $(u + v) + av$. D'autre part, $v_i + w_i$ est la $i^{\text{ème}}$ composante de $v + w$ et donc $u_i + (v_i + w_i)$ est la $i^{\text{ème}}$ composante de $u + (v + w)$. Mais u_i, v_i et w_i sont des nombres réels pour lesquels l'association de l'addition est vérifiée, donc

$$(u_i + v_i) + w_i = u_i + (v_i + w_i) \quad \text{pour} \quad i = 1, \ldots, n$$

D'où $(u + v) + w = u + (v + w)$ car leurs composantes correspondantes sont égales.

(2) Avec $0 = (0, 0, \ldots, 0)$, on a

$$\begin{aligned} u + 0 &= (u_1, u_2, \ldots, u_n) + (0, 0, \ldots, 0) \\ &= (u_1 + 0, u_2 + 0, \ldots, u_n + 0) = (u_1, u_2, \ldots, u_n) = u \end{aligned}$$

(3) Puisque $-u = -1(u_1, u_2, \ldots, u_n) = (-u_1, -u_2, \ldots, -u_n)$,

$$\begin{aligned} u + (-u) &= (u_1, u_2, \ldots, u_n) + (-u_1, -u_2, \ldots, -u_n) \\ &= (u_1 - u_1, u_2 - u_2, \ldots, u_n - u_n) = (0, 0, \ldots, 0) = 0 \end{aligned}$$

(4) Par définition $u_i + v_i$ est la $i^{\text{ième}}$ composante de $u + v$ et $v_i + u_i$ est la $i^{\text{ème}}$ composante de $v + u$. Or u_i et v_i sont des nombres réels pour lesquels la commutativité de l'addition est vérifiée, donc

$$u_i + v_i = v_i + u_i, \qquad i = 1, \ldots, n$$

D'où $u + v = v + u$ car leurs composantes correspondantes sont égales.

(5) Puisque $u_i + v_i$ est la $i^{\text{ème}}$ composante de $u + v$, $k(u_i + v_i)$ est la $i^{\text{ème}}$ composante de $k(u + v)$. De même ku_i et kv_i sont les $i^{\text{èmes}}$ composantes de ku et kv respectivement, $ku_i + kv_i$ est donc la $i^{\text{ème}}$ composante de $ku + kv$. Or k, u_i et v_i sont des nombres réels, d'où

$$k(u_i + v_i) = ku_i + kv_i, \qquad i = 1, \ldots, n$$

Ainsi $k(u + v) = ku + kv$, les composantes correspondantes étant égales.

(6) Remarquons que le premier signe plus se rapporte à l'addition de deux scalaires k et k', alors que le second signe plus, se rapporte à une addition vectorielle entre les deux vecteurs ku et kv.

Par définition, $(k + k')u_i$ est la $i^{\text{ème}}$ composante du vecteur $(k + k')u$. Puisque ku_i et $k'u_i$ sont les $i^{\text{èmes}}$ composantes de ku et ku' respectivement, $ku_i + k'u_i$ est la $i^{\text{ème}}$ composante de $ku + ku'$. Or k, k' et u_i sont des nombres réels, donc

$$(k + k')u_i = ku_i + k'u_i, \qquad i = 1, \ldots, n$$

Ainsi $(k + k')u = ku + k'u$, les composantes correspondantes étant égales.

(7) Puisque $k'u_i$ est la $i^{\text{ème}}$ composante de $k'u$, $k(k'u_i)$ est la $i^{\text{ème}}$ composante de $k(k'u)$. Or $(kk')u_i$ est la $i^{\text{ème}}$ composante de $(kk')u$ et k, k', u_i étant des nombres réels, on a

$$(kk')u_i = k(k'u_i), \qquad i = 1, \ldots, n$$

Donc $(kk')u = k(k'u)$, les composantes correspondantes étant égales.

(8) $1 \cdot u = 1(u_1, u_2, \ldots, u_n) = (1u_1, 1u_2, \ldots, 1u_n) = (u_1, u_2, \ldots, u_n) = u.$

1.7. Montrer que $0u = 0$ quel que soit le vecteur u, où le premier 0 est un scalaire et le second un vecteur.

Méthode 1 : $0u = 0(u_1, u_2, \ldots, u_n) = (0u_1, 0u_2, \ldots, 0u_n) = (0, 0, \ldots, 0) = 0$

Méthode 2 : D'après le théorème 1.1, $0u = (0 + 0)u = 0u + 0u$

En ajoutant $-0u$ aux deux membres, on trouve le résultat demandé.

PRODUIT SCALAIRE

1.8. Calculer $u \cdot v$ avec : 1) $u = (2, -3, 6)$, $v = (8, 2, -3)$; 2) $u = (1, -8, 0, 5)$, $v = (3, 6, 4)$; 3) $u = (3, -5, 2, 1)$, $v = (4, 1, -2, 5)$.

(1) Multiplions les composantes correspondantes et ajoutons: $u \cdot v = 2 \cdot 8 + (-3) \cdot 2 + 6 \cdot (-3) = -8$.

(2) Le produit scalaire n'est pas défini, les vecteurs ayant un nombre distinct de composantes.

(3) En multipliant les composantes correspondantes et en ajoutant: $u \cdot v = 3 \cdot 4 + (-5) \cdot 1 + 2 \cdot (-2) + 1 \cdot 5 = 8$.

1.9. Déterminer k de manière que les vecteurs u et v soient orthogonaux où

(i) $u = (1, k, -3)$ et $v = (2, -5, 4)$

(ii) $u = (2, 3k, -4, 1, 5)$ et $v = (6, -1, 3, 7, 2k)$

Dans chaque cas, calculons $u \cdot v$, et égalons à 0, en résolvant en k.

(i) $u \cdot v = 1 \cdot 2 + k \cdot (-5) + (-3) \cdot 4 = 2 - 5k - 12 = 0, \quad -5k - 10 = 0, \quad k = -2$

(ii) $u \cdot v = 2 \cdot 6 + 3k \cdot (-1) + (-4) \cdot 3 + 1 \cdot 7 + 5 \cdot 2k$
$= 12 - 3k - 12 + 7 + 10k = 0, \quad k = -1$

1.10. Démontrer le théorème 1.2 : quels que soient les vecteurs $u, v, w \in \mathbf{R}^n$ et quel que soit le scalaire $k \in \mathbf{R}$,

(i) $(u+v) \cdot w = u \cdot w + v \cdot w$ (iii) $u \cdot v = v \cdot u$

(ii) $(ku) \cdot v = k(u \cdot v)$ (iv) $u \cdot u \geqq 0$, et $u \cdot u = 0$ ssi $u = 0$

Soient $u = (u_1, u_2, \ldots, u_n)$, $v = (v_1, v_2, \ldots, v_n)$, $w = (w_1, w_2, \ldots, w_n)$.

(i) Puisque $u + v = (u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n)$,

$$\begin{aligned}(u+v) \cdot w &= (u_1 + v_1)w_1 + (u_2 + v_2)w_2 + \cdots + (u_n + v_n)w_n \\ &= u_1w_1 + v_1w_1 + u_2w_2 + v_2w_2 + \cdots + u_nw_n + v_nw_n \\ &= (u_1w_1 + u_2w_2 + \cdots + u_nw_n) + (v_1w_1 + v_2w_2 + \cdots + v_nw_n) \\ &= u \cdot w + v \cdot w\end{aligned}$$

(ii) Puisque $ku = (ku_1, ku_2, \ldots, ku_n)$,

$$(ku) \cdot v = ku_1v_1 + ku_2v_2 + \cdots + ku_nv_n = k(u_1v_1 + u_2v_2 + \cdots + u_nv_n) = k(u \cdot v)$$

(iii) $u \cdot v = u_1v_1 + u_2v_2 + \cdots + u_nv_n = v_1u_1 + v_2u_2 + \cdots + v_nu_n = v \cdot u$

(iv) u_i^2 étant positif pour chaque i, et la somme de nombres réels positifs étant positive,

$$u \cdot u = u_1^2 + u_2^2 + \cdots + u_n^2 \geqq 0$$

De plus $u \cdot u = 0$, ssi $u_i = 0$ pour chaque i, c'est-à-dire ssi $u = 0$.

DISTANCE ET NORME DANS $\mathbf{R}^n$

1.11. Trouver la distance $d(u, v)$ des deux points u et v où (i) $u = (1,7), v = (6, -5)$; (ii) $u = (3, -5, 4), v = (6, 2, -1)$; (iii) $u = (5, 3, -2, -4, -1)$, $v = (2, -1, 0, -7, 2)$.

Utilisons dans chaque cas la formule $d(u, v) = \sqrt{(u_1 - v_1)^2 + \cdots + (u_n - v_n)^2}$

(i) $d(u,v) = \sqrt{(1-6)^2 + (7+5)^2} = \sqrt{25 + 144} = \sqrt{169} = 13$

(ii) $d(u,v) = \sqrt{(3-6)^2 + (-5-2)^2 + (4+1)^2} = \sqrt{9 + 49 + 25} = \sqrt{83}$

(iii) $d(u,v) = \sqrt{(5-2)^2 + (3+1)^2 + (-2+0)^2 + (-4+7)^2 + (-1-2)^2} = \sqrt{47}$

1.12. Trouver k de telle sorte que $d(u, v) = 6$ où $u = (2, k, 1, -4)$ et $v = (3, -1, 6, -3)$.

$$(d(u,v))^2 = (2-3)^2 + (k+1)^2 + (1-6)^2 + (-4+3)^2 = k^2 + 2k + 28$$

Résolvons maintenant $k^2 + 2k + 28 = 6^2$ on obtient $k = 2, -4$.

1.13. Trouver la norme du vecteur u si (i) $u = (2, -7)$, (ii) $u = (3, -12, -4)$

Utilisons dans chaque cas la formule $\|u\| = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2}$.

(i) $\|u\| = \sqrt{2^2 + (-7)^2} = \sqrt{4 + 49} = \sqrt{53}$

(ii) $\|u\| = \sqrt{3^2 + (-12)^2 + (-4)^2} = \sqrt{9 + 144 + 16} = \sqrt{169} = 13$

1.14. Déterminer k de telle manière que $||u|| = \sqrt{39}$ où $u = (1, k, -2, 5)$.

$$||u||^2 = 1^2 + k^2 + (-2)^2 + 5^2 = k^2 + 30$$

Résolvons $k^2 + 30 = 39$ on obtient $k = 3, -3$.

1.15. Montrer que $||u|| \geqslant 0$ et $||u|| = 0$ ssi $u = 0$.

D'après le théorème 1.2 $uu \geqslant 0$ et $u \,.\, u = 0$, ssi $u = 0$. Puisque $||u|| = \sqrt{u \,.\, u}$, le résultat en découle.

1.16. Démontrer le théorème 1.3 (Cauchy-Schwarz):

Quels que soient les vecteurs $u = (u_1, \ldots, u_n)$ et $v = (v_1, \ldots, v_n)$ de $\mathbf{R}^n$, $|u \,.\, v| \leqslant ||u||\, ||v||$.

Nous démontrerons d'abord la double inégalité $|u \cdot v| \leq \sum_{i=1}^{n} |u_i v_i| \leq ||u||\, ||v||$.

Si $u = 0$ ou $v = 0$, l'inégalité se réduit à $0 \leqslant 0 \leqslant 0$ qui est évidemment vérifiée. Il reste donc à considérer le cas où $u \neq 0$ et $v \neq 0$ d'où $||u|| \neq 0$ et $||v|| \neq 0$. De plus

$$|u \cdot v| = |u_1 v_1 + \cdots + u_n v_n| \leq |u_1 v_1| + \cdots + |u_n v_n| = \sum |u_i v_i|$$

Démontrons maintenant la seconde inégalité.

Pour tous nombres réels $x, y \in \mathbf{R}$, $0 \leqslant (x - y)^2 = x^2 - 2xy + y^2$, ce qui est équivalent à

$$2xy \leq x^2 + y^2 \tag{1}$$

Posons $x = |u_i|/||u||$ et $y = |v_i|/||v||$ dans (1), nous obtenons, quel que soit i,

$$2 \frac{|u_i|}{||u||} \frac{|v_i|}{||v||} \leq \frac{|u_i|^2}{||u||^2} + \frac{|v_i|^2}{||v||^2} \tag{2}$$

Mais, par définition de la norme d'un vecteur on a $||u|| = \sum u_i^2 = \sum |u_i|^2$ et $||v|| = \sum v_i^2 = \sum |v_i|^2$.

En sommant par rapport à i et en utilisant $|u_i v_i| = |u_i|\, |v_i|$, on a

$$2 \frac{\sum |u_i v_i|}{||u||\, ||v||} \leq \frac{\sum |u_i|^2}{||u||^2} + \frac{\sum |v_i|^2}{||v||^2} = \frac{||u||^2}{||u||^2} + \frac{||v||^2}{||v||^2} = 2$$

c'est-à-dire

$$\frac{\sum |u_i v_i|}{||u||\, ||v||} \leq 1$$

Enfin en multipliant les deux membres par $||u||\, ||v||$, nous obtenons l'inégalité demandée.

1.17. Démontrer l'inégalité de Minkowski:

Quels que soient les vecteurs $u = (u_1, \ldots, u_n)$ et $v = (v_1, \ldots, v_n)$ de $\mathbf{R}^n$, $||u + v|| \leqslant ||u|| + ||v||$.

Si $||u + v|| = 0$, l'inégalité est vérifiée. Il reste alors à considérer le cas $||u + v|| \neq 0$.

Or $|u_i + v_i| \leqslant |u_i| + |v_i|$, quels que soient les nombres réels $u_i, v_i \in \mathbf{R}$. Donc

$$\begin{aligned} ||u + v||^2 &= \sum (u_i + v_i)^2 = \sum |u_i + v_i|^2 \\ &= \sum |u_i + v_i|\, |u_i + v_i| \leq \sum |u_i + v_i|\, (|u_i| + |v_i|) \\ &= \sum |u_i + v_i|\, |u_i| + \sum |u_i + v_i|\, |v_i| \end{aligned}$$

D'après l'inégalité de Cauchy-Schwarz (voir le précédent problème)

$$\sum |u_i + v_i|\, |u_i| \leq ||u + v||\, ||u|| \quad \text{et} \quad \sum |u_i + v_i|\, |v_i| \leq ||u + v||\, ||v||$$

Donc

$$||u + v||^2 \leq ||u + v||\, ||u|| + ||u + v||\, ||v|| = ||u + v||\, (||u|| + ||v||)$$

En divisant par $||u + v||$, nous obtenons l'inégalité demandée.

1.18. Démontrer que la norme dans $\mathbf{R}^n$ satisfait les lois suivantes:

$[N_1]$: Quel que soit le vecteur u, $||u|| \geqslant 0$ et $||u|| = 0$ ssi $u = 0$.

$[N_2]$: Quel que soit le vecteur u, et quel que soit le scalaire k, $||ku|| = |k|\,.\,||u||$.

$[N_3]$: Quels que soient les vecteurs u et v, $||u + v|| \leqslant ||u|| + ||v||$.

$[N_1]$ est démontré dans le problème 1.15 et $[N_3]$ dans le problème 1.17. Il reste uniquement à démontrer $[N_2]$.

Supposons $u = (u_1, u_2, \ldots, u_n)$ et donc $ku = (ku_1, ku_2, \ldots, ku_n)$. Alors

$$\begin{aligned} ||ku||^2 &= (ku_1)^2 + (ku_2)^2 + \cdots + (ku_n)^2 = k^2u_1^2 + k^2u_2^2 + \cdots + k^2u_n^2 \\ &= k^2(u_1^2 + u_2^2 + \cdots + u_n^2) = k^2\,||u||^2 \end{aligned}$$

En prenant la racine carrée des deux membres, on obtient l'égalité demandée.

NOMBRES COMPLEXES

1.19. Simplifier : 1) $(5+3i)(2-7i)$; 2) $(4-3i)^2$; 3) $\dfrac{1}{3-4i}$; 4) $\dfrac{2-7i}{5+3i}$; 5) i^3, i^4, i^{31}; 6) $(1+2i)^3$; 7) $\left(\dfrac{1}{2-3i}\right)^2$.

1) $(5+3i)(2-7i) = 10 + 6i - 35i - 21i^2 = 31 - 29i$

2) $(4-3i)^2 = 16 - 24i + 9i^2 = 7 - 24i$

3) $\dfrac{1}{3-4i} = \dfrac{(3+4i)}{(3-4i)(3+4i)} = \dfrac{3+4i}{25} = \dfrac{3}{25} + \dfrac{4}{25}i$

4) $\dfrac{2-7i}{5+3i} = \dfrac{(2-7i)(5-3i)}{(5+3i)(5-3i)} = \dfrac{-11-41i}{34} = -\dfrac{11}{34} - \dfrac{41}{34}i$

5) $i^3 = i^2 \cdot i = (-1)i = -i$; $i^4 = i^2 \cdot i^2 = 1$; $i^{31} = (i^4)^7 \cdot i^3 = 1^7 \cdot (-i) = -i$

6) $(1+2i)^3 = 1 + 6i + 12i^2 + 8i^3 = 1 + 6i - 12 - 8i = -11 - 2i$

7) $\left(\dfrac{1}{2-3i}\right)^2 = \dfrac{1}{-5-12i} = \dfrac{(-5+12i)}{(-5-12i)(-5+12i)} = \dfrac{-5+12i}{169} = -\dfrac{5}{169} + \dfrac{12}{169}i$

1.20. Soient $z = 2 - 3i$ et $w = 4 + 5i$. Trouver

(1) $z + w$ et zw; (2) z/w; (3) $\bar{z}$ et $\bar{w}$; (4) $|z|$ et $|w|$.

1) $z + w = 2 - 3i + 4 + 5i = 6 + 2i$

$zw = (2-3i)(4+5i) = 8 - 12i + 10i - 15i^2 = 23 - 2i$

2) $\dfrac{z}{w} = \dfrac{2-3i}{4+5i} = \dfrac{(2-3i)(4-5i)}{(4+5i)(4-5i)} = \dfrac{-7-22i}{41} = -\dfrac{7}{41} - \dfrac{22}{41}i$

3) Utilisons $\overline{a+bi} = a - bi$: $\bar{z} = \overline{2-3i} = 2 + 3i$; $\bar{w} = \overline{4+5i} = 4 - 5i$.

4) Utilisons $|a+bi| = \sqrt{a^2+b^2}$: $|z| = |2-3i| = \sqrt{4+9} = \sqrt{13}$; $|w| = |4+5i| = \sqrt{16+25} = \sqrt{41}$.

1.21. Démontrer que quels que soient les nombres complexes $z, w \in \mathbf{C}$,

(i) $\overline{z+w} = \bar{z} + \bar{w}$, (ii) $\overline{zw} = \bar{z}\,\bar{w}$, (iii) $\bar{\bar{z}} = z$.

Posons $z = a + bi$ et $w = c + di$ où $a, b, c, d \in \mathbf{R}$.

(i) $$\begin{aligned} \overline{z+w} &= \overline{(a+bi) + (c+di)} = \overline{(a+c) + (b+d)i} \\ &= (a+c) - (b+d)i = a + c - bi - di \\ &= (a-bi) + (c-di) = \bar{z} + \bar{w} \end{aligned}$$

(ii) $$\begin{aligned} \overline{zw} &= \overline{(a+bi)(c+di)} = \overline{(ac-bd) + (ad+bc)i} \\ &= (ac-bd) - (ad+bc)i = (a-bi)(c-di) = \bar{z}\,\bar{w} \end{aligned}$$

(iii) $\bar{\bar{z}} = \overline{\overline{a+bi}} = \overline{a-bi} = a - (-b)i = a + bi = z$

1.22. Démontrer que quels que soient les nombres complexes $z, w \in \mathbf{C}, |zw| = |z|\ |w|$.

Posons $z = a + bi$ et $w = c + di$, où $a, b, c, d \in \mathbf{R}$. Alors

$$|z|^2 = a^2 + b^2, \quad |w|^2 = c^2 + d^2, \quad \text{et} \quad zw = (ac - bd) + (ad + bc)i$$

Alors
$$\begin{aligned} |zw|^2 &= (ac - bd)^2 + (ad + bc)^2 \\ &= a^2c^2 - 2abcd + b^2d^2 + a^2d^2 + 2abcd + b^2c^2 \\ &= a^2(c^2 + d^2) + b^2(c^2 + d^2) = (a^2 + b^2)(c^2 + d^2) = |z|^2\,|w|^2 \end{aligned}$$

En prenant la racine carrée des deux membres on obtient le résultat demandé.

1.23. Démontrer que quels que soient les nombres complexes $z, w \in \mathbf{C}$, $|z + w| \leqslant |z| + |w|$.

Posons $z = a + bi$ et $w = c + di$ où $a, b, c, d \in \mathbf{R}$. Considérons les vecteurs $u = (a, b)$ et $v = (c, d)$ de $\mathbf{R}^2$, remarquons que

$$|z| = \sqrt{a^2 + b^2} = ||u||, \quad |w| = \sqrt{c^2 + d^2} = ||v||$$

et $$|z + w| = |(a + c) + (b + d)i| = \sqrt{(a + c)^2 + (b + d)^2} = ||(a + c, b + d)|| = ||u + v||$$

D'après l'inégalité de Minkowski (Problème 1.17) $||u + v|| \leqslant ||u|| + ||v||$. On a donc

$$|z + w| = ||u + v|| \leq ||u|| + ||v|| = |z| + |w|$$

VECTEURS DE $\mathbf{C}^n$

1.24. Soient $u = (3 - 2i, 4i, 1 + 6i)$ et $v = (5 + i, 2 - 3i, 5)$. Trouver (1) $u + v$, (2) $4iu$, (3) $(1 + i)v$, (4) $(1 - 2i)u + (3 + i)v$.

1) Additionnons les composantes correspondantes : $u + v = (8 - i, 2 + i, 6 + 6i)$.

2) Multiplions chaque composante de u par le scalaire $4i$: $4iu = (8 + 12i, - 16, - 24 + 4i)$.

3) Multiplions chaque composante de v par le scalaire $1 + i$:

$$(1 + i)v = (5 + 6i + i^2, 2 - i - 3i^2, 5 + 5i) = (4 + 6i, 5 - i, 5 + 5i)$$

4) Effectuons d'abord la multiplication scalaire puis l'addition vectorielle:

$$\begin{aligned} (1 - 2i)u + (3 + i)v &= (-1 - 8i, 8 + 4i, 13 + 4i) + (14 + 8i, 9 - 7i, 15 + 5i) \\ &= (13, 17 - 3i, 28 + 9i) \end{aligned}$$

1.25. Trouver $u \cdot v$ et $v \cdot u$ avec 1) $u = (1 - 2i, 3 + i), v = (4 + 2i, 5 - 6i)$; 2) $u = (3 - 2i, 4i, 1 + 6i)$, $v = (5 + i, 2 - 3i, 7 + 2i)$.

Rappelons que les composantes conjuguées du second vecteur apparaissent dans le produit scalaire:

$$(z_1, \ldots, z_n) \cdot (w_1, \ldots, w_n) = z_1\overline{w}_1 + \cdots + z_n\overline{w}_n$$

1)
$$\begin{aligned} u \cdot v &= (1 - 2i)\overline{(4 + 2i)} + (3 + i)\overline{(5 - 6i)} \\ &= (1 - 2i)(4 - 2i) + (3 + i)(5 + 6i) = -10i + 9 + 23i = 9 + 13i \end{aligned}$$

$$\begin{aligned} v \cdot u &= (4 + 2i)\overline{(1 - 2i)} + (5 - 6i)\overline{(3 + i)} \\ &= (4 + 2i)(1 + 2i) + (5 - 6i)(3 - i) = 10i + 9 - 23i = 9 - 13i \end{aligned}$$

2)
$$\begin{aligned} u \cdot v &= (3 - 2i)\overline{(5 + i)} + (4i)\overline{(2 - 3i)} + (1 + 6i)\overline{(7 + 2i)} \\ &= (3 - 2i)(5 - i) + (4i)(2 + 3i) + (1 + 6i)(7 - 2i) = 20 + 35i \end{aligned}$$

$$\begin{aligned} v \cdot u &= (5 + i)\overline{(3 - 2i)} + (2 - 3i)\overline{(4i)} + (7 + 2i)\overline{(1 + 6i)} \\ &= (5 + i)(3 + 2i) + (2 - 3i)(-4i) + (7 + 2i)(1 - 6i) = 20 - 35i \end{aligned}$$

Dans les deux exemples $v \cdot u = \overline{u \cdot v}$. Ce qui est vrai en général, comme on le montre dans le problème 1.27.

1.26. Trouver $||u||$ où (i) $u = (3+4i, 5-2i, 1-3i)$; (ii) $u = (4-i, 2i, 3+2i, 1-5i)$.

Rappelons que $z\bar{z} = a^2 + b^2$ lorsque $z = a + bi$. Employons

$$||u||^2 = u \cdot u = z_1\bar{z}_1 + z_2\bar{z}_2 + \cdots + z_n\bar{z}_n \quad \text{où} \quad z = (z_1, z_2, \ldots, z_n)$$

(i) $||u||^2 = (3)^2 + (4)^2 + (5)^2 + (-2)^2 + (1)^2 + (-3)^2 = 64$, ou $||u|| = 8$

(ii) $||u||^2 = 4^2 + (-1)^2 + 2^2 + 3^2 + 2^2 + 1^2 + (-5)^2 = 60$, ou $||u|| = \sqrt{60} = 2\sqrt{15}$

1.27. Montrer que quels que soient les vecteurs $u, v \in C^n$ et quel que soit le scalaire $z \in C$, 1) $u \,.\, v = \overline{v \,.\, u}$, 2) $(zu).v = z.(uv)$ 3) $u \,.\, (zv) = \bar{z}\,(u \cdot v)$. (Comparer avec le théorème 1.2).

Supposons que $u = (z_1, z_2, \ldots, z_n)$ et $v = (w_1, w_2, \ldots, w_n)$.

1) En utilisant les propriétés du conjugué établies dans le Problème 1.21,

$$\begin{aligned}\overline{v \cdot u} &= \overline{w_1\bar{z}_1 + w_2\bar{z}_2 + \cdots + w_n\bar{z}_n} = \overline{w_1\bar{z}_1} + \overline{w_2\bar{z}_2} + \cdots + \overline{w_n\bar{z}_n} \\ &= \bar{w}_1 z_1 + \bar{w}_2 z_2 + \cdots + \bar{w}_n z_n = z_1\bar{w}_1 + z_2\bar{w}_2 + \cdots + z_n\bar{w}_n = u \cdot v\end{aligned}$$

2) Puisque $zu = (zz_1, zz_2, \ldots, zz_n)$,

$$(zu) \cdot v = zz_1\bar{w}_1 + zz_2\bar{w}_2 + \cdots + zz_n\bar{w}_n = z(z_1\bar{w}_1 + z_2\bar{w}_2 + \cdots + z_n\bar{w}_n) = z(u \cdot v)$$

3) **Méthode 1 :** On a $zv = (zw_1, zw_2, \ldots, zw_n)$,

$$\begin{aligned}u \cdot (zv) &= z_1\overline{zw_1} + z_2\overline{zw_2} + \cdots + z_n\overline{zw_n} = z_1\bar{z}\bar{w}_1 + z_2\bar{z}\bar{w}_2 + \cdots + z_n\bar{z}\bar{w}_n \\ &= \bar{z}(z_1\bar{w}_1 + z_2\bar{w}_2 + \cdots + z_n\bar{w}_n) = \bar{z}(u \cdot v)\end{aligned}$$

Méthode 2 : Utilisons 1) et 2)

$$u \cdot (zv) = \overline{(zv) \cdot u} = \overline{z(v \cdot u)} = \bar{z}\,(\overline{v \cdot u}) = \bar{z}\,(u \cdot v)$$

PROBLEMES DIVERS

1.28. Soit $u = (3, -2, 1, 4)$ et $v = (7, 1, -3, 6)$. Trouver:

(i) $u+v$; (ii) $4u$; (iii) $2u-3v$; (iv) $u \cdot v$; (v) $||u||$ et $||v||$; (vi) $d(u, v)$.

(i) $u + v = (3+7, -2+1, 1-3, 4+6) = (10, -1, -2, 10)$

(ii) $4u = (4 \cdot 3, 4 \cdot (-2), 4 \cdot 1, 4 \cdot 4) = (12, -8, 4, 16)$

(iii) $2u - 3v = (6, -4, 2, 8) + (-21, -3, 9, -18) = (-15, -7, 11, -10)$

(iv) $u \cdot v = 21 - 2 - 3 + 24 = 40$

(v) $||u|| = \sqrt{9+4+1+16} = \sqrt{30}$, $||v|| = \sqrt{49+1+9+36} = \sqrt{95}$

(vi) $d(u,v) = \sqrt{(3-7)^2 + (-2-1)^2 + (1+3)^2 + (4-6)^2} = \sqrt{45} = 3\sqrt{5}$

1.29. Soit $u = (7 - 2i, 2 + 5i)$ et $v = (1 + i, -3 - 6i)$. Trouver :

(i) $u+v$; (ii) $2iu$; (iii) $(3-i)v$; (iv) $u \cdot v$; (v) $||u||$ et $||v||$.

(i) $u + v = (7 - 2i + 1 + i, 2 + 5i - 3 - 6i) = (8 - i, -1 - i)$

(ii) $2iu = (14i - 4i^2, 4i + 10i^2) = (4 + 14i, -10 + 4i)$

(iii) $(3-i)v = (3 + 3i - i - i^2, -9 - 18i + 3i + 6i^2) = (4 + 2i, -15 - 15i)$

(iv) $$\begin{aligned}u \cdot v &= (7-2i)(\overline{1+i}) + (2+5i)(\overline{-3-6i}) \\ &= (7-2i)(1-i) + (2+5i)(-3+6i) = 5 - 9i - 36 - 3i = -31 - 12i\end{aligned}$$

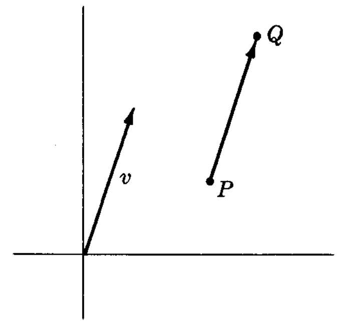
(v) $||u|| = \sqrt{7^2 + (-2)^2 + 2^2 + 5^2} = \sqrt{82}$, $||v|| = \sqrt{1^2 + 1^2 + (-3)^2 + (-6)^2} = \sqrt{47}$

1.30. Tout couple de points $P = (a_i)$ et $Q = (b_i)$ de $\mathbf{R}^n$ définit un *segment orienté* $\overrightarrow{PQ}$. Identifions $\overrightarrow{PQ}$ avec le vecteur $v = Q - P$:

$$\overrightarrow{PQ} = v = (b_1 - a_1, b_2 - a_2, \ldots, b_n - a_n)$$

Trouver le vecteur v identifié à $\overrightarrow{PQ}$ où

(i) $P = (2, 5)$, $Q = (-3, 4)$
(ii) $P = (1, -2, 4)$, $Q = (6, 0, -3)$



(i) $v = Q - P = (-3-2, 4-5) = (-5, -1)$
(ii) $v = Q - P = (6-1, 0+2, -3-4) = (5, 2, -7)$

1.31. L'ensemble H des éléments de $\mathbf{R}^n$ qui sont les solutions d'une équation linéaire à n inconnues $x_1, x_2, \ldots x_n$ de la forme

$$c_1x_1 + c_2x_2 + \cdots + c_nx_n = b \qquad (*)$$

avec $u = (c_1, c_2, \ldots, c_n) \neq 0$ dans $\mathbf{R}^n$, est appelé un *hyperplan* de $\mathbf{R}^n$, et (*) est l'équation de l'hyperplan H.(Nous identifierons souvent H avec (*)). Montrons que le segment orienté $\overrightarrow{PQ}$ pour tout couple de points $P, Q \in H$ est orthogonal au vecteur u de composantes $(c_1, c_2, \ldots, c_n)$; le vecteur u est appelé vecteur *normal* à l'hyperplan H.

Supposons $P = (a_1, \ldots, a_n)$ et $Q = (b_1, \ldots, b_n)$. Les a_i et b_i sont alors les solutions de l'équation donnée

$$c_1a_1 + c_2a_2 + \cdots + c_na_n = b, \qquad c_1b_1 + c_2b_2 + \cdots + c_nb_n = b$$

Soit $$v = \overrightarrow{PQ} = Q - P = (b_1 - a_1, b_2 - a_2, \ldots, b_n - a_n)$$

Alors
$$\begin{aligned} u \cdot v &= c_1(b_1 - a_1) + c_2(b_2 - a_2) + \cdots + c_n(b_n - a_n) \\ &= c_1b_1 - c_1a_1 + c_2b_2 - c_2a_2 + \cdots + c_nb_n - c_na_n \\ &= (c_1b_1 + c_2b_2 + \cdots + c_nb_n) - (c_1a_1 + c_2a_2 + \cdots + c_na_n) = b - b = 0 \end{aligned}$$

D'où v, donc $\overrightarrow{PQ}$, est orthogonal à u.

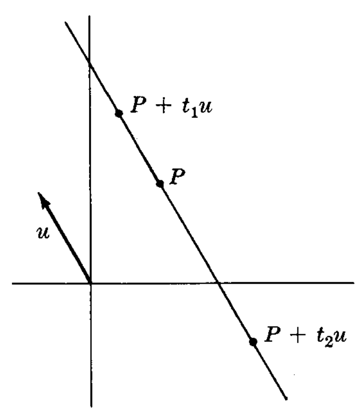
1.32. Trouver l'équation de l'hyperplan H de $\mathbf{R}^4$ si 1) H passe par $P = (3, -2, 1, -4)$ et est normal au vecteur $u = (2, 5, -6, -2)$; 2) H passe le point $P(1, -2, 3, 5)$ et est parallèle à l'hyperplan H' déterminé par $4x - 5y + 2z + w = 11$.

1) L'équation de H est de la forme $2x + 5y - 6z - 2w = k$ puisqu'il est normal à u. Remplaçons dans l'équation de H (x, y, z, w) par les coordonnées de P; on obtient $k = -2$. D'où l'équation de H est $2x + 5y - 6z - 2w = -2)$.

2) H et H' sont parallèles ssi les vecteurs normaux ont la même direction ou des directions opposées. Ainsi l'équation de H est de la forme $4x - 5y + 2z + w = k$. Remplaçons (x, y, z, w) par les coordonnées de P dans l'équation précédente, on trouve $k = 25$. L'équation de H est donc $4x - 5y + 2z + w = 25$.

1.33. La droite l de $\mathbf{R}^n$ passant par le point $P = (a_i)$ et parallèle à $u = (u_i) \neq 0$ est l'ensemble des points X tels que $X = P + tu$, $t \in \mathbf{R}$, donc est l'ensemble des points $X = (x_i)$ tels que

$$(*) \begin{cases} x_1 = a_1 + u_1t \\ x_2 = a_2 + u_2t \\ \cdots\cdots\cdots\cdots \\ x_n = a_n + u_nt \end{cases}$$



où t prend toutes les valeurs réelles. La variable t est appelée *paramètre* et (*) est appelée la *représentation paramétrique* de la droite l.

(1) Trouver une représentation paramétrique de la droite passant par P et de vecteur directeur u où (a) $P = (2, 5)$ et $u = (-3, 4)$;(b) $P = (4, -2, 3, 1)$ et $u = (2, 5, -7, 11)$.

(2) Trouver une représentation paramétrique de la droite passant par les points P et Q où (a) $P = (7, -2)$ et $Q = (9, 3)$;(b) $P = (5, 4, -3)$ et $Q = (1, -3, 2)$.

(1) En utilisant la formule (*),

$$(a)\quad \begin{cases} x = 2 - 3t \\ y = 5 + 4t \end{cases} \qquad (b)\quad \begin{cases} x = 4 + 2t \\ y = -2 + 5t \\ z = 3 - 7t \\ w = 1 + 11t \end{cases}$$

Dans $\mathbf{R}^2$, on élimine habituellement t entre les deux équations et la droite est alors déterminée par la seule équation $4x + 3y = 23$.

(2) Calculons d'abord $u = \overrightarrow{PQ} = Q - P$. Utilisons ensuite la formule (*).

$$(a)\quad u = Q - P = (2, 5) \qquad (b)\quad u = Q - P = (-4, -7, 5)$$

$$\begin{cases} x = 7 + 2t \\ y = -2 + 5t \end{cases} \qquad \begin{cases} x = 5 - 4t \\ y = 4 - 7t \\ z = -3 + 5t \end{cases}$$

(Remarquons que dans chaque cas on peut aussi écrire $u = \overrightarrow{QP} = P - Q$.)

PROBLEMES SUPPLEMENTAIRES

VECTEURS DE $\mathbf{R}^n$

1.34. Soit $u = (1, -2, 5)$, $v = (3, 1, -2)$. Trouver (i) $u + v$; (ii) $-6u$; (iii) $2u - 5v$; (iv) $u \cdot v$; (v) $\|u\|$ et $\|v\|$; (vi) $d(u, v)$.

1.35. Soit $u = (2, -1, 0, -3)$, $v = (1, -1, -1, 3)$, $w = (1, 3, -2, 2)$. Trouver : (i) $2u - 3v$; (ii) $5u - 3v - 4w$; (iii) $-u + 2v - 2w$; (iv) $u \cdot v$, $u \cdot w$ et $v \cdot w$; (v) $d(u, v)$ et $d(v, w)$.

1.36. Soit $u = (2, 1, -3, 0, 4)$, $v = (5, -3, -1, 2, 7)$. Trouver : (i) $u + v$; (ii) $3u - 2v$; (iii) $u \cdot v$; (iv) $\|u\|$ et $\|v\|$; (v) $d(u, v)$.

1.37. Déterminer k de telle sorte que les vecteurs u et v soient orthogonaux. (i) $u = (3, k, -2)$, $v = (6, -4, -3)$. (ii) $u = (5, k, -4, 2)$, $v = (1, -3, 2, 2k)$. (iii) $u = (1, 7, k + 2, -2)$, $v = (3, k, -3, k)$.

1.38. Déterminer x et y si : (i) $(x, x + y) = (y - 2, 6)$; (ii) $x(1, 2) = -4(y, 3)$.

1.39. Déterminer x et y si : (i) $x(3, 2) = 2(y, -1)$; (ii) $x(2, y) = y(1, -2)$.

1.40. Trouver x et y tels que

(i) $(3, -1, 2) = x(1, 1, 1) + y(1, -1, 0) + z(1, 0, 0)$

(ii) $(-1, 3, 3) = x(1, 1, 0) + y(0, 0, -1) + z(0, 1, 1)$

1.41. Soit $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$ $e_z = (0, 0, 1)$. Montrer que pour tout vecteur $u = (a, b, c)$ de $\mathbf{R}^3$: (i) $u = ae_1 + be_2 + ce_3$; (ii) $u \cdot e_1 = a$, $u \cdot e_2 = b$, $u \cdot e_3 = c$.

1.42. Généralisons le résultat du précédent problème. Soit $e_i \in \mathbf{R}^n$ le vecteur admettant 1 comme $i^{\text{ème}}$ coordonnée et 0 pour toutes les autres:

$$e_1 = (1, 0, 0, \ldots, 0, 0), \quad e_2 = (0, 1, 0, \ldots, 0, 0), \quad \ldots, \quad e_n = (0, 0, 0, \ldots, 0, 1)$$

Montrer que pour tout vecteur $u = (a_1, a_2, \ldots, a_n)$ on a

(i) $u = a_1e_1 + a_2e_2 + \cdots + a_ne_n$, (ii) $u \cdot e_i = a_i$ pour $i = 1, \ldots, n$

1.43. Soit $u \in \mathbf{R}^n$; supposons que $u \cdot v = 0$ quel que soit $v \in \mathbf{R}^n$. Montrer alors que $u = 0$.

1.44. En utilisant la formule $d(u \cdot v) = \|u - v\|$ et les propriétés $[N_1]$, $[N_2]$, $[N_3]$ des normes dans le problème 1.18, montrer que la fonction distance vérifie les propriétés suivantes quels que soient les vecteurs $u, v, w \in \mathbf{R}^n$:

(i) $d(u, v) \geqq 0$, et $d(u, v) = 0$ ssi $u = v$; (ii) $d(u, v) = d(v, u)$; (iii) $d(u, w) \leqq d(u, v) + d(v, w)$.

NOMBRES COMPLEXES

1.45. Simplifier (i) $(4 - 7i)(9 + 2i)$; (ii) $(3 - 5i)^2$; (iii) $\dfrac{1}{4 - 7i}$; (iv) $\dfrac{9 + 2i}{3 - 5i}$; (v) $(1 - i)^3$.

1.46. Simplifier (i) $\dfrac{1}{2i}$; (ii) $\dfrac{2 + 3i}{7 - 3i}$; (iii) i^{15}, i^{25}, i^{34}; (iv) $\left(\dfrac{1}{3 - i}\right)^2$.

1.47. Soit $z = 2 - 5i$ et $w = 7 + 3i$. Trouver (i) $z + w$; (ii) zw; (iii) z/w; (iv) $\bar{z}, \bar{w}$; (v) $|z|, |w|$.

1.48. Soit $z = 2 + i$ et $w = 6 - 5i$. Trouver (i) z/w; (ii) $\bar{z}, \bar{w}$; (iii) $|z|, |w|$.

1.49. Montrer que 1) $zz^{-1} = 1$; 2) $|z| = |\bar{z}|$; 3) Montrer que la partie réelle de z $\mathrm{Re}(z) = \dfrac{1}{2}(z + \bar{z})\,|z|$
4) Montrer que la partie imaginaire de z, $I(z) = (z - \bar{z})/2i$.

1.50. Montrer que $zw = 0$ implique $z = 0$ ou $w = 0$.

VECTEURS DE $\mathbf{C}^n$

1.51. Soit $u = (1 + 7i, 2 - 6i)$ et $v = (5 - 2i, 3 - 4i)$. Trouver 1) $u + v$; 2) $(3 + i)u$; 3) $2iu + (4 - 7i)v$; 4) $u \cdot v$ et $v \cdot u$; 5) $\|u\|$ et $\|v\|$.

1.52. Soit $u = (3 - 7i, 2i, -1 + i)$ et $v = (4 - i, 11 + 2i, 8 - 3i)$. Trouver 1) $u - v$; 2) $(3 + i)v$; 3) $u \cdot v$ et $v \cdot u$; 4) $\|u\|$ et $\|v\|$.

1.53. Démontrer que quels que soient les vecteurs $u, v, w \in \mathbf{C}^n$:

1) $(u + v) \cdot w = u \cdot w + v \cdot w$; 2) $w \cdot (u + v) = w \cdot u + w \cdot v$. (Comparer avec le théorème 1.2).

1.54. Démontrer que la norme dans $\mathbf{C}^n$ satisfait aux lois suivantes:

$[N_1]$ quel que soit le vecteur u, $\|u\| \geqslant 0$ et $\|u\| = 0$ ssi $u = 0$.

$[N_2]$ quel que soit le vecteur u, et quel que soit le nombre complexe z, $\|zu\| = |z|\,\|u\|$,

$[N_3]$ quels que soient les vecteurs u et v, $\|u + v\| \leqslant \|u\| + \|v\|$.

(Comparer avec le théorème 1.18).

PROBLEMES DIVERS

1.55. Trouver une équation de l'hyperplan de $\mathbf{R}^3$ qui :

1) passe par le point $(2, -7, 1)$ et est normal au vecteur $(3, 1, -11)$;

2) contient les points $(1, -2, 2)$, $(0, 1, 3)$ et $(0, 2, -1)$;

3) contient le point $(1, -5, 2)$ et est parallèle au plan $3x - 7y + 4z = 5$.

1.56. Trouver la valeur de k de telle sorte que $2x - ky + 4z - 5w = 11$ soit perpendiculaire à $7x + 2y - z + 2w = 8$ (Deux hyperplans sont perpendiculaires ssi leurs vecteurs normaux respectifs sont orthogonaux).

1.57. Trouver une représentation paramétrique de la droite qui :

1) passe par $(7, -1, 8)$ et est parallèle au vecteur $(1, 3, -5)$
2) passe par les deux points $(1, 9, -4, 5)$ et $(2, -3, 0, 4)$
3) passe par $(4, -1, 9)$ et est perpendiculaire au plan $3x - 2y + z = 18$.

1.58. Soit P, Q et R des points d'une droite déterminée par

$$x_1 = a_1 + u_1 t, \quad x_2 = a_2 + u_2 t, \quad \ldots, \quad x_n = a_n + u_n t$$

qui correspondent aux valeurs t_1, t_2, t_3 du paramètre t. Montrer que si $t_1 < t_2 < t_3$ alors $d(P, Q) + d(Q, R) = d(P, R)$.

REPONSES AUX PROBLEMES SUPPLEMENTAIRES

1.34. (i) $u + v = (4, -1, 3)$; (ii) $-6u = (-6, 12, -30)$; (iii) $2u - 5v = (-13, -9, 20)$; (iv) $u \cdot v = -9$; (v) $\|u\| = \sqrt{30}$, $\|v\| = \sqrt{14}$; (vi) $d(u, v) = \sqrt{62}$

1.35. (i) $2u - 3v = (1, 1, 3, -15)$; (ii) $5u - 3v - 4w = (3, -14, 11, -32)$; (iii) $-u + 2v - 2w = (-2, -7, 2, 5)$; (iv) $u \cdot v = -6$, $u \cdot w = -7$, $v \cdot w = 6$; (v) $d(u, v) = \sqrt{38}$, $d(v, w) = 3\sqrt{2}$

1.36. (i) $u + v = (7, -2, -4, 2, 11)$; (ii) $3u - 2v = (-4, 9, -7, -4, -2)$; (iii) $u \cdot v = 38$; (iv) $\|u\| = \sqrt{30}$, $\|v\| = 2\sqrt{22}$; (v) $d(u, v) = \sqrt{42}$

1.37. (i) $k = 6$; (ii) $k = 3$; (iii) $k = 3/2$

1.38. (i) $x = 2,\ y = 4$; (ii) $x = -6,\ y = 3/2$

1.39. (i) $x = -1,\ y = -3/2$; (ii) $x = 0,\ y = 0$; ou $x = -2,\ y = -4$

1.40 (i) $x = 2,\ y = 3,\ z = -2$; (ii) $x = -1,\ y = 1,\ z = 4$

1.43. Nous obtenons $u \,.\, u = 0$, ce qui implique $u = 0$.

1.45. (i) $50 - 55i$; (ii) $-16 - 30i$; (iii) $(4 + 7i)/65$; (iv) $(1 + 3i)/2$; (v) $-2 - 2i$.

1.46. (i) $-\frac{1}{2}i$; (ii) $(5 + 27i)/58$; (iii) $-i, i, -1$; (iv) $(4 + 3i)/50$.

1.47. (i) $z + w = 9 - 2i$; (ii) $zw = 29 - 29i$; (iii) $z/w = (-1 - 41i)/58$; (iv) $\bar{z} = 2 + 5i$, $\bar{w} = 7 - 3i$; (v) $|z| = \sqrt{29}$, $|w| = \sqrt{58}$.

1.48. (i) $z/w = (7 + 16i)/61$; (ii) $\bar{z} = 2 - i$, $\bar{w} = 6 + 5i$; (iii) $|z| = \sqrt{5}$, $|w| = \sqrt{61}$.

1.50. Si $zw = 0$ alors $|zw| = |z|\,|w| = |0| = 0$. D'où $|z| = 0$ ou $|w| = 0$ et donc $z = 0$ ou $w = 0$.

1.51. (i) $u + v = (6 + 5i,\ 5 - 10i)$
(ii) $(3 + i)u = (-4 + 22i,\ 12 - 16i)$
(iii) $2iu + (4 - 7i)v = (-8 - 41i,\ -4 - 33i)$
(iv) $u \cdot v = 21 + 27i$, $v \cdot u = 21 - 27i$
(v) $\|u\| = 3\sqrt{10}$, $\|v\| = 3\sqrt{6}$

1.52. (i) $u - v = (-1 - 6i, -11, -9 + 4i)$
(ii) $(3 + i)v = (13 + i, 31 + 17i, 27 - i)$
(iii) $u \cdot v = 12 + 2i$, $v \cdot u = 12 - 2i$
(iv) $\|u\| = 8$, $\|v\| = \sqrt{215}$

1.55. (i) $3x + y - 11z = -12$; (ii) $13x + 4y + z = 7$; (iii) $3x - 7y + 4z = 46$.

1.56. $k = 0$

1.57. (i) $\begin{cases} x = 7 + t \\ y = -1 + 3t \\ z = 8 - 5t \end{cases}$ (ii) $\begin{cases} x = 1 + t \\ y = 9 - 12t \\ z = -4 + 4t \\ w = 5 - t \end{cases}$ (iii) $\begin{cases} x = 4 + 3t \\ y = -1 - 2t \\ z = 9 + t \end{cases}$

CHAPITRE 2

Equations linéaires

INTRODUCTION

La théorie des équations linéaires joue un rôle important en algèbre linéaire. En fait de très nombreux problèmes de cette matière se résument à l'étude d'un système d'équations linéaires, par exemple la recherche du noyau d'un homomorphisme d'espace vectoriel et la détermination d'un sous-espace vectoriel donné par un ensemble de vecteurs. En conséquence les techniques introduites dans ce chapitre seront applicables aux méthodes plus abstraites données par la suite. D'autre part quelques-uns des résultats de ces méthodes abstraites nous donneront de nouveaux aperçus sur la structure des systèmes concrets d'équations linéaires.

Pour simplifier, nous supposerons que toutes les équations de ce chapitre sont données sur le corps **R**. Les résultats et les méthodes employées resteront valables sur le corps complexe **C**, ou dans n'importe quel corps K.

EQUATION LINEAIRE

Par équation linéaire sur le corps **R**, nous entendons une expression de la forme

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b \tag{1}$$

où les $a_i, b_i \in \mathbf{R}$ et les x_i sont les *indéterminées (ou inconnues ou variables)*. Les scalaires a_i sont appelés les *coefficients* des x_i, et b est le *terme constant*, ou simplement la *constante* de l'équation. Un ensemble de valeurs des inconnues, par exemple

$$x_1 = k_1,\ x_2 = k_2,\ \ldots,\ x_n = k_n$$

est une *solution* de (*1*) si l'égalité obtenue en remplaçant x_i par k_i,

$$a_1k_1 + a_2k_2 + \cdots + a_nk_n = b$$

est vérifiée. Cet ensemble de valeurs satisfait alors l'équation. S'il n'y a aucune ambiguïté sur la position des inconnues dans l'équation, cette solution pourra alors s'écrire sous forme de n-tuple

$$u = (k_1, k_2, \ldots, k_n)$$

Exemple 2.1 : Considérons l'équation $x + 2y - 4z + w = 3$.

Le 4-tuple $u = (3, 2, 1, 0)$ est une solution de l'équation puisque

$$3 + 2\cdot 2 - 4\cdot 1 + 0 = 3 \quad \text{ou} \quad 3 = 3$$

est une égalité vraie. Cependant le 4-tuple $v = (1, 2, 4, 5)$ n'est pas une solution de l'équation puisque

$$1 + 2\cdot 2 - 4\cdot 4 + 5 = 3 \quad \text{ou} \quad -6 = 3$$

n'est pas vérifié.

Les solutions de l'équation (1) peuvent être aisément obtenues. Il y a 3 cas:

Cas 1 : Un des coefficients de (*1*) n'est pas nul, par exemple $a_1 \neq 0$. Alors nous pouvons écrire l'équation sous la forme

$$a_1x_1 = b - a_2x_2 - \cdots - a_nx_n \quad \text{ou} \quad x_1 = a_1^{-1}b - a_1^{-1}a_2x_2 - \cdots - a_1^{-1}a_nx_n$$

En donnant arbitrairement des valeurs déterminées aux inconnues $x_2, \ldots, x_n$, nous obtenons une valeur pour x_1 ; ces valeurs forment une solution de l'équation. De plus, toute solution de l'équation peut être obtenue de cette manière. Remarquons en particulier que l'équation linéaire à une inconnue

$$ax = b \quad \text{avec} \quad a \neq 0$$

a une solution unique $x = a^{-1}. b$.

Exemple 2.2 : Considérons l'équation $2x - 4y + z = 8$.

Ecrivons l'équation sous la forme

$$2x = 8 + 4y - z \quad \text{ou} \quad x = 4 + 2y - \frac{1}{2} z$$

A chaque valeur de y et z correspondra une valeur de x, et ces trois valeurs constitueront une solution de l'équation. Par exemple, soient $y = 3$ et $z = 2$ d'où $x = 4 + 2 \cdot 3 - \frac{1}{2} \cdot 2 = 9$ On peut donc dire que le 3-tuple (triplet) $u = (9, 3, 2)$ est une solution de l'équation.

Cas 2 : Tous les coefficients de l'équation (*1*) sont nuls, mais le terme constant est différent de zéro. L'équation est alors de la forme

$$0x_1 + 0x_2 + \cdots + 0x_n = b, \quad \text{avec} \quad b \neq 0$$

L'équation n'a alors pas de solution.

Cas 3 : Tous les coefficients de l'équation (*1*) sont nuls, et le terme constant est nul également. L'équation est de la forme

$$0x_1 + 0x_2 + \cdots + 0x_n = 0$$

Quel que soit le n-tuple constitué de scalaires de **R**, il s'agira d'une solution de l'équation.

SYSTEME D'EQUATIONS LINEAIRES

Considérons maintenant un système de m équations linéaires à n inconnues $x_1, x_2, \ldots, x_n$:

$$\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_2 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{array} \qquad (*)$$

où les a_{ij}, b_i appartiennent au corps des réels **R**. On dit que le système est *homogène* si les constantes $b_1, \ldots, b_m$ sont nulles. Un n-tuple $u = (k_1, k_2 \ldots k_n)$ de nombres réels est une *solution* (ou une *solution particulière*) s'il satisfait chacune des équations. L'ensemble des solutions est souvent appelé *solution générale du système.*

Le système d'équations linéaires

$$\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0 \end{array} \qquad (**)$$

est appelé système homogène associé à (*). Le système précédent a toujours une solution, qui est le n-tuple nul $0 = (0, 0, \ldots, 0)$, appelé solution *triviale* ou *nulle.* N'importe quelle autre solution, si elle existe, est appelée solution *non triviale* ou *non nulle.*

Il existe la relation fondamentale suivante entre les système (*) et (**).

Théorème 2.1 : Soit u une solution particulière du système non homogène (*) et soit W la solution générale du système homogène associé (**). Alors

$$u + W = \{u + w : w \in W\}$$

est la solution générale du système non homogène (*).

Cependant le théorème précédent n'a qu'un intérêt théorique et ne peut en aucune manière nous permettre d'obtenir les solutions explicites du système (*). Celles-ci seront fournies par la méthode habituelle d'élimination exposée ci-après.

SOLUTIONS D'UN SYSTEME D'EQUATIONS LINEAIRES

Considérons le système (*) d'équations linéaires. Réduisons-le à un système plus simple de la manière suivante.

Premièrement. Echanger les équations de telle sorte que la première inconnue x_1 ait un coefficient non nul dans la première équation, ainsi $a_{11} \neq 0$.

Deuxièmement. Pour chaque $i > 1$, appliquer

$$L_i \to -a_{i1}L_1 + a_{11}L_i$$

c'est-à-dire remplacer la $i$ème équation linéaire L_i par l'équation obtenue en multipliant la première équation L_1 par $-a_{i1}$, et en ajoutant la $i$ème équation L_i multipliée par a_{11}.

Nous obtenons alors le système suivant qui d'après le Problème 2.13 est équivalent à (*), c'est-à-dire a le même ensemble de solutions que (*):

$$\begin{aligned} a_{11}x_1 + a'_{12}x_2 + a'_{13}x_3 + \cdots + a'_{1n}x_n &= b'_1 \\ a'_{2j_2}x_{j_2} + \cdots\cdots + a'_{2n}x_n &= b'_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\ a'_{mj_2}x_{j_2} + \cdots\cdots + a'_{mn}x_n &= b'_m \end{aligned}$$

où $a_{11} \neq 0$. Ici x_{j_2} représente la première inconnue dont le coefficient est non nul dans une équation autre que la première ; d'après le deuxièmement $x_{j_2} \neq x_1$. Ce procédé consistant à éliminer une inconnue à partir des diverses équations successives est appelé procédé d'élimination ou de Gauss.

Exemple 2.3 : Soit le système suivant d'équations linéaires:

$$\begin{aligned} 2x + 4y - z + 2v + 2w &= 1 \\ 3x + 6y + z - v + 4w &= -7 \\ 4x + 8y + z + 5v - w &= 3 \end{aligned}$$

Eliminons l'inconnue x à partir de la seconde et de la troisième équation en appliquant les combinaisons suivantes:

$$L_2 \to -3L_1 + 2L_2 \quad \text{et} \quad L_3 \to -2L_1 + L_3$$

Calculons

$$\begin{array}{rr} -3L_1: & -6x - 12y + 3z - 6v - 6w = -3 \\ 2L_2: & 6x + 12y + 2z - 2v + 8w = -14 \\ \hline -3L_1 + 2L_2: & 5z - 8v + 2w = -17 \end{array}$$

et

$$\begin{array}{rr} -2L_1: & -4x - 8y + 2z - 4v - 4w = -2 \\ L_3: & 4x + 8y + z + 5v - w = 3 \\ \hline -2L_1 + L_3: & 3z + v - 5w = 1 \end{array}$$

Ainsi le système initial a été réduit au système équivalent :

$$\begin{aligned} 2x + 4y - z + 2v + 2w &= 1 \\ 5z - 8v + 2w &= -17 \\ 3z + v - 5w &= 1 \end{aligned}$$

Remarquons que y a aussi été éliminé à partir de la seconde et de la troisième équation. Ici l'inconnue z joue le rôle de l'inconnue précédente x_{j_2}

Les équations précédentes, sauf la première, forment un système partiel qui a moins d'équations et moins d'inconnues que le système initial. De plus

1) s'il se présente une équation de la forme $0x_1 + 0x_2 + \ldots + 0x_n = b$, $b \neq 0$, le système est alors *impossible* et n'a pas de solution;

2) s'il se présente une équation de la forme, $0x_1 + 0x_2 + \ldots + 0x_n = 0$, cette équation peut être supprimée et ceci sans affecter la solution.

En continuant le procédé précédent de Gauss, avec chaque nouveau sous-système, nous obtenons par récurrence soit un système impossible soit un système réduit à la forme équivalente suivante

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots\cdots\cdots\cdots + a_{1n}x_n &= b_1 \\ a_{2j_2}x_{j_2} + a_{2,j_2+1}x_{j_2+1} + \cdots\cdots + a_{2n}x_n &= b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{rj_r}x_{j_r} + a_{r,j_r+1}x_{j_r+1} + \cdots + a_{rn}x_n &= b_r \end{aligned} \qquad (***)$$

où $1 < j_2 < \ldots < j_r$ et où les coefficients restants sont non nuls :

$$a_{11} \neq 0, \; a_{2j_2} \neq 0, \; \ldots, \; a_{rj_r} \neq 0$$

(Nous utiliserons pour des commodités de notations les mêmes symboles a_{ij}, b_k dans le système (***) que ceux utilisés dans le système (*) mais ils désigneront des scalaires différents).

Définition : Le système (***) est sous une forme dite *échelonnée* ; les inconnues x_i qui n'apparaissent pas au commencement de chaque équation ($i \neq 1, j_2, \ldots, j_r$) sont appelées des inconnues libres ou *variables libres.*

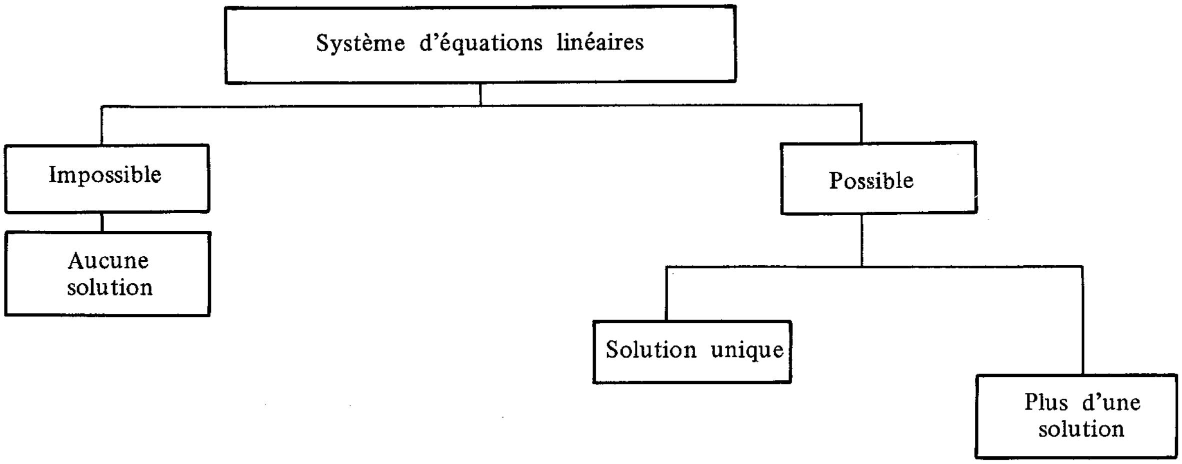
Le théorème suivant s'applique alors.

Théorème 2.2 : Dans un système réduit à une forme échelonnée, on peut distinguer deux cas :

1) $r = n$; il y a autant d'équations que d'inconnues. Le système admet alors une solution unique.

2) $r < n$; il y a moins d'équations que d'inconnues. Nous attribuerons alors arbitrairement des valeurs aux $n - r$ inconnues libres et nous obtiendrons une solution du système.

Le théorème précédent implique évidemment que le système (***) et n'importe quels systèmes équivalents sont possibles. Ainsi si le système (*) est possible et ramené au précédent cas 2), nous pourrons alors donner différentes valeurs aux inconnues libres et obtenir ainsi plusieurs solutions du système, comme le schéma suivant le montre.



D'après le théorème 2.1, la solution unique précédente peut se présenter seulement lorsque le système homogène associé admet seulement la solution triviale nulle.

Exemple 2.4 : Réduisons le système suivant en appliquant les combinaisons $L_2 \to -3L_1 + 2L_2$ et $L_3 \to -3L_1 + 2L_3$ et $L_3 \to -3L_2 + L_3$:

$$\begin{aligned} 2x + y - 2z + 3w &= 1 \\ 3x + 2y - z + 2w &= 4 \\ 3x + 3y + 3z - 3w &= 5 \end{aligned} \qquad \begin{aligned} 2x + y - 2z + 3w &= 1 \\ y + 4z - 5w &= 5 \\ 3y + 12z - 15w &= 7 \end{aligned} \qquad \begin{aligned} 2x + y - 2z + 3w &= 1 \\ y + 4z - 5w &= 5 \\ 0 &= -8 \end{aligned}$$

L'équation $0 = -8$, c'est-à-dire $0x + 0y + 0w = -8$, nous montre que le système initial est impossible et donc n'a aucune solution.

Exemple 2.5 : Réduisons le système suivant en appliquant les combinaisons $L_2 \to -L_1 + L_2$, $L_3 \to -2L_1 + L_3$ et $L_4 \to -2L_1 + L_4$ ainsi que les combinaisons $L_3 \to L_2 - L_3$ et $L_4 \to -2L_2 + L_4$:

$$\begin{aligned} x + 2y - 3z &= 4 \\ x + 3y + z &= 11 \\ 2x + 5y - 4z &= 13 \\ 2x + 6y + 2z &= 22 \end{aligned} \qquad \begin{aligned} x + 2y - 3z &= 4 \\ y + 4z &= 7 \\ y + 2z &= 5 \\ 2y + 8z &= 14 \end{aligned} \qquad \begin{aligned} x + 2y - 3z &= 4 \\ y + 4z &= 7 \\ 2z &= 2 \\ 0 &= 0 \end{aligned}$$

$$\begin{aligned} x + 2y - 3z &= 4 \\ y + 4z &= 7 \\ 2z &= 2 \end{aligned}$$

Remarquons d'abord que le système est possible puisqu'il n'y a aucune équation de la forme $0 = b$ avec $b \neq 0$. De plus, puisque dans la forme réduite échelonnée il y a trois équations à trois inconnues, le système admet une solution unique. Dans la troisième on a $z = 1$. En remplaçant z par 1 dans la deuxième équation on obtient $y = 3$. En remplaçant z par 1 et y par 3 dans la première équation on trouve $x = 1$. Ainsi $x = 1, y = 3$ et $z = 1$ ou le triplet $(1, 3, 1)$ est la solution unique du système.

Exemple 2.6 : Réduisons le système suivant à l'aide des combinaisons $L_2 \to -2L_1 + L_2$ et $L_3 \to -5L_1 + l$ et $L_3 \to -2L_2 + L_3$:

$$\begin{aligned} x + 2y - 2z + 3w &= 2 \\ 2x + 4y - 3z + 4w &= 5 \\ 5x + 10y - 8z + 11w &= 12 \end{aligned} \qquad \begin{aligned} x + 2y - 2z + 3w &= 2 \\ z - 2w &= 1 \\ 2z - 4w &= 2 \end{aligned} \qquad \begin{aligned} x + 2y - 2z + 3w &= 2 \\ z - 2w &= 1 \\ 0 &= 0 \end{aligned}$$

$$\begin{aligned} x + 2y - 2z + 3w &= 2 \\ z - 2w &= 1 \end{aligned}$$

Le système est possible, et puisqu'il y a plus d'inconnues que d'équations dans la forme réduite échelonnée, le système a un nombre infini de solutions. En fait il y a deux inconnues libres y et w, et donc une solution particulière peut être obtenue en donnant des valeurs à y et w. Par exemple prenons $w = 1$ et $y = -2$. Remplaçons w par 1 dans la seconde équation, on obtient $z = 3$. Remplaçons w par 1, z par 3 et y par -2 dans la première équation, on trouve $x = 9$. Ainsi $x = 9$, $y = -2$, $z = 3$ et $w = 1$, ou le 4-tuple $(9, -2, 3, 1)$ est une solution particulière du système.

Remarque : Dans l'exemple précédent, nous pouvons trouver la solution générale du système de la manière suivante. Donnons aux inconnues libres des valeurs bien déterminées ; par exemple $y = a$ et $w = b$. Remplaçons w par b, dans la seconde équation, nous obtenons $z = 1 + 2b$. Remplaçons y par a, z par $1 + 2b$ et w par b dans la première équation, nous trouvons $x = 4 - 2a + b$. Donc la solution générale du système est

$$x = 4 - 2a + b, \quad y = a, \quad z = 1 + 2b, \quad w = b$$

ou $(4 - 2a + b, a, 1 + 2b, b)$, où a et b sont des constantes arbitraires. Fréquemment on laisse dans la solution générale les inconnus libres y et w (au lieu de a et b), comme suit :

$$x = 4 - 2y + w, \quad z = 1 + 2w \qquad \text{ou} \qquad (4 - 2y + w, y, 1 + 2w, w)$$

Nous examinerons de nouveau la représentation de la solution générale d'un système d'équations linéaires dans un prochain chapitre.

Exemple 2-7 : Soit un système de deux équations à deux inconnues :

$$\begin{aligned} a_1x + b_1y &= c_1 \\ a_2x + b_2y &= c_2 \end{aligned}$$

D'après notre théorie, il se présente seulement un des trois cas suivants :

1) Le système est impossible.

2) Le système peut être réduit à un système équivalent de forme échelonnée à 2 équations.

3) Le système est équivalent à un système de forme échelonnée à une équation.

Lorsque les équations linéaires à deux inconnues à coefficients réels peuvent être représentées comme des droites de $\mathbf{R}^2$, les cas précédents peuvent alors être interprétés géométriquement de la manière suivante :

1) Les deux droites sont parallèles.

2) Les deux droites se coupent en un point unique.

3) Les deux droites coïncident.

SOLUTION D'UN SYSTEME HOMOGENE D'EQUATIONS LINEAIRES

Si nous commençons par un système homogène d'équations linéaires, le système est alors possible, puisque, par exemple, il admet comme solution la solution nulle $0 = (0, 0, \ldots, 0)$. Il peut donc toujours être réduit à un système équivalent homogène mis sous forme échelonnée:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots\cdots\cdots\cdots + a_{1n}x_n &= 0 \\ a_{2j_2}x_{j_2} + a_{2,j_2+1}x_{j_2+1} + \cdots\cdots + a_{2n}x_n &= 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\ a_{rj_r}x_{j_r} + a_{r,j_r+1}x_{j_r+1} + \cdots + a_{rn}x_n &= 0 \end{aligned}$$

Nous avons alors deux possibilités.

1) $r = n$. Alors le système a seulement la solution nulle.

2) $r < n$. Alors le système a une solution non nulle.

Si nous commençons avec moins d'équations que d'inconnues, alors dans la forme échelonnée, $r < n$ et le système admet une solution non nulle. D'où

Théorème 2.3 : Un système d'équations linéaires homogène avec plus d'inconnues que d'équations admet une solution non nulle.

Exemple 2-8 : Le système homogène

$$\begin{aligned} x + 2y - 3z + w &= 0 \\ x - 3y + z - 2w &= 0 \\ 2x + y - 3z + 5w &= 0 \end{aligned}$$

a une solution non nulle puisqu'il y a quatre inconnues et seulement trois équations.

Exemple 2-9 : Réduisons le système suivant à sa forme échelonnée :

$$\begin{aligned} x + y - z &= 0 \\ 2x - 3y + z &= 0 \\ x - 4y + 2z &= 0 \end{aligned} \qquad \begin{aligned} x + y - z &= 0 \\ -5y + 3z &= 0 \\ -5y + 3z &= 0 \end{aligned} \qquad \begin{aligned} x + y - z &= 0 \\ -5y + 3z &= 0 \end{aligned}$$

Le système a une solution non nulle, puisque nous obtenons seulement deux équations à trois inconnues dans la forme échelonnée. Par exemple, soit $z = 5$; alors $y = 3$ et $x = 2$. En d'autres termes le triplet (2, 3, 5) est une solution particulière non nulle du système.

Exemple 2-10 : Réduisons le système suivant à sa forme échelonnée :

$$\begin{aligned} x + y - z &= 0 \\ 2x + 4y - z &= 0 \\ 3x + 2y + 2z &= 0 \end{aligned} \qquad \begin{aligned} x + y - z &= 0 \\ 2y + z &= 0 \\ -y + 5z &= 0 \end{aligned} \qquad \begin{aligned} x + y - z &= 0 \\ 2y + z &= 0 \\ 11z &= 0 \end{aligned}$$

Dans la forme réduite échelonnée, il y a trois équations à trois inconnues, le système admet seulement la solution nulle (0, 0, 0).

PROBLEMES RESOLUS

SOLUTION D'EQUATIONS LINEAIRES

2.1. Résoudre le système :

$$\begin{aligned} 2x - 3y + 6z + 2v - 5w &= 3 \\ y - 4z + v &= 1\,. \\ v - 3w &= 2 \end{aligned}$$

Le système est déjà mis sous sa forme réduite échelonnée, les équations commençant par les inconnues x, y et v respectivement, les autres inconnues z et w sont des inconnues libres.

Pour trouver la solution générale, posons $z = a$ et $w = b$. En remplaçant z et w par leurs valeurs dans la troisième équation on obtient

$$v - 3b = 2 \quad \text{ou} \quad v = 2 + 3b$$

En remplaçant dans la seconde équation, on a

$$y - 4a + 2 + 3b = 1 \quad \text{ou} \quad y = 4a - 3b - 1$$

et dans la première, on obtient

$$2x - 3(4a - 3b - 1) + 6a + 2(2 + 3b) - 5b = 3 \quad \text{ou} \quad x = 3a - 5b - 2$$

Ainsi la solution générale de ce système est

$$x = 3a - 5b - 2, \quad y = 4a - 3b - 1, \quad z = a, \quad v = 2 + 3b, \quad w = b$$

ou $(3a - 5b - 2, 4a - 3b - 1, a, 2 + 3b, b)$, où a et b sont des nombres réels arbitraires. Certains préfèrent écrire la solution générale en conservant les inconnues libres z et w au lieu de a et b, comme suit :

$$\begin{aligned} x &= 3z - 5w - 2 \\ y &= 4z - 3w - 1 \qquad \text{ou} \qquad (3z-5w-2,\ 4z-3w-1,\ z,\ 2+3w,\ w) \\ v &= 2 + 3w \end{aligned}$$

Après avoir trouvé la solution générale, nous pouvons trouver une solution particulière, en prenant par exemple $a = 2$ et $b = 1$; alors

$$x=-1,\ y=4,\ z=2,\ v=5,\ w=1 \qquad \text{ou} \qquad (-1, 4, 2, 5, 1)$$

est une solution particulière du système donné.

2.2. Résoudre le système
$$\begin{aligned} x + 2y - 3z &= -1 \\ 3x - y + 2z &= 7\,. \\ 5x + 3y - 4z &= 2 \end{aligned}$$

Réduisons à la forme échelonnée. Eliminons x à partir de la seconde et de la troisième équation par les combinaisons $L_2 \to -3L_1 + L_2$ et $L_3 \to -5L_1 + L_3$:

$$\begin{array}{rrcl} -3L_1: & -3x - 6y + 9z &=& 3 \\ L_2: & 3x - y + 2z &=& 7 \\ \hline -3L_1 + L_2: & -7y + 11z &=& 10 \end{array} \qquad \begin{array}{rrcl} -5L_1: & -5x - 10y + 15z &=& 5 \\ L_3: & 5x + 3y - 4z &=& 2 \\ \hline -5L_1 + L_3: & -7y + 11z &=& 7 \end{array}$$

Nous obtenons ainsi le système équivalent

$$\begin{aligned} x + 2y - 3z &= -1 \\ -7y + 11z &= 10 \\ -7y + 11z &= 7 \end{aligned}$$

La seconde et la troisième équation montrent que le système est impossible, car en soustrayant, nous obtenons $0x + 0y + 0z = 3$ ou $0 = 3$.

2.3. Résoudre le système
$$\begin{aligned} 2x + y - 2z &= 10 \\ 3x + 2y + 2z &= 1\,. \\ 5x + 4y + 3z &= 4 \end{aligned}$$

Réduisons le système à sa forme échelonnée. Eliminons x à l'aide de la seconde et de la troisième équation par les combinaisons $L_2 \to -3L_1 + 2L_2$ et $L_3 \to -5L_1 + 2L_3$:

$$\begin{array}{rrcl} -3L_1: & -6x - 3y + 6z &=& -30 \\ 2L_2: & 6x + 4y + 4z &=& 2 \\ \hline -3L_1 + 2L_2: & y + 10z &=& -28 \end{array} \qquad \begin{array}{rrcl} -5L_1: & -10x - 5y + 10z &=& -50 \\ 2L_3: & 10x + 8y + 6z &=& 8 \\ \hline -5L_1 + 2L_3: & 3y + 16z &=& -42 \end{array}$$

Nous obtenons ainsi le système suivant, à partir duquel nous éliminons y à partir de la troisième équation par la combinaison $L_3 \to -3L_2 + L_3$

$$\begin{aligned} 2x + y - 2z &= 10 \\ y + 10z &= -28 \\ 3y + 16z &= -42 \end{aligned} \qquad \text{d'où} \qquad \begin{aligned} 2x + y - 2z &= 10 \\ y + 10z &= -28 \\ -14z &= 42 \end{aligned}$$

Dans la forme réduite échelonnée il y a trois équations à trois inconnues ; donc le système a une solution unique. D'après la troisième équation, $z = -3$. En remplaçant z par cette valeur dans la seconde équation nous trouvons $y = 2$. En remplaçant y et z par leurs valeurs dans la première équation, on obtient $x = 1$. Ainsi $x = 1$, $y = 2$ et $z = -3$, c'est-à-dire que le triplet $(1, 2, -3)$ est la solution unique du système précédent.

2.4. Résoudre le système
$$\begin{aligned} x + 2y - 3z &= 6 \\ 2x - y + 4z &= 2 \\ 4x + 3y - 2z &= 14 \end{aligned}.$$

Réduisons ce système à sa forme échelonnée. Eliminons x entre la seconde et la troisième équation par les combinaisons $L_2 \to -2L_1 + L_2$ et $L_3 \to -4L_1 + L_3$:

$$\begin{array}{rrcr} -2L_1: & -2x - 4y + 6z & = & -12 \\ L_2: & 2x - y + 4z & = & 2 \\ \hline & -5y + 10z & = & -10 \\ \text{ou} & y - 2z & = & 2 \end{array} \qquad \begin{array}{rrcr} -4L_1: & -4x - 8y + 12z & = & -24 \\ L_3: & 4x + 3y - 2z & = & 14 \\ \hline & -5y + 10z & = & -10 \\ \text{ou} & y - 2z & = & 2 \end{array}$$

Ainsi on obtient le système équivalent

$$\begin{aligned} x + 2y - 3z &= 6 \\ y - 2z &= 2 \\ y - 2z &= 2 \end{aligned} \quad \text{ou en simplifiant} \quad \begin{aligned} x + 2y - 3z &= 6 \\ y - 2z &= 2 \end{aligned}$$

(Puisque la seconde et la troisième équation sont identiques, nous pouvons supprimer une d'entre elles).

Dans la forme réduite échelonnée, il y a seulement deux équations à trois inconnues ; le système a donc une infinité de solutions et en particulier $3 - 2 = 1$ inconnue libre qui est z.

Pour obtenir la solution générale, posons $z = a$. En remplaçant z par a dans la seconde équation, on obtient $y = 2 + 2a$. En remplaçant de nouveau z et y par leurs valeurs dans la première équation on obtient $x + 2(2 + 2a) - 3a = 6$ ou $x = 2 - a$. On obtient ainsi la solution générale

$$x = 2 - a, \quad y = 2 + 2a, \quad z = a \qquad \text{ou} \qquad (2 - a, 2 + 2a, a)$$

où a est un nombre réel quelconque.

Donnons à a une valeur $a = 1$, nous obtenons une solution particulière $x = 1$, $y = 4$, $z = 1$ ou $(1, 4, 1)$.

2.5. Résoudre le système
$$\begin{aligned} x - 3y + 4z - 2w &= 5 \\ 2y + 5z + w &= 2 \\ y - 3z \quad\quad &= 4 \end{aligned}.$$

Ce système n'est pas mis sous sa forme échelonnée, car y est la première inconnue dans la seconde et la troisième équation. Cependant, si nous réécrivons le système de telle sorte que w soit la seconde inconnue, on obtient alors le système suivant qui est écrit sous sa forme réduite échelonnée:

$$\begin{aligned} x - 2w - 3y + 4z &= 5 \\ w + 2y + 5z &= 2 \\ y - 3z &= 4 \end{aligned}$$

Si un 4-tuple (a, b, c, d) est donné comme solution, on ne sait s'il faut remplacer w ou y par b ; donc pour des raisons théoriques nous considérerons les deux systèmes comme distincts. Cependant ceci ne nous empêche pas d'utiliser le nouveau système pour obtenir la solution du système initial.

Soit $z = a$. En remplaçant z par a dans la troisième équation, nous trouvons $y = 4 + 3a$. En remplaçant z et y par leurs valeurs respectives dans la seconde équation, on obtient $w + 2(4 + 3a) + 5a = 2$ ou $w = -6 - 11a$. En remplaçant dans la première équation on a

$$x - 2(-6 - 11a) - 3(4 + 3a) + 4a = 5 \qquad \text{ou} \qquad x = 5 - 17a$$

Ainsi la solution générale du système initial est

$$x = 5 - 17a, \quad y = 4 + 3a, \quad z = a, \quad w = -6 - 11a$$

où a est un nombre réel quelconque.

2.6. Déterminer les valeurs de a de telle sorte que le système suivant dont les inconnues sont x, y et z ait : 1) aucune solution, 2) plus d'une solution, 3) une solution unique:

$$\begin{aligned} x + y - z &= 1 \\ 2x + 3y + az &= 3 \\ x + ay + 3z &= 2 \end{aligned}$$

Réduisons le système à sa forme échelonnée. Eliminons x à partir de la seconde et de la troisième équation par les combinaisons. $L_2 \to -2L_1 + L_3$ et $L_3 \to -L_1 + L_3$.

$$\begin{array}{rrcl} -2L_1: & -2x - 2y + 2z &=& -2 \\ L_2: & 2x + 3y + az &=& 3 \\ \hline & y + (a+2)z &=& 1 \end{array} \qquad \begin{array}{rrcl} -L_1: & -x - y + z &=& -1 \\ L_3: & x + ay + 3z &=& 2 \\ \hline & (a-1)y + 4z &=& 1 \end{array}$$

Le système équivalent est ainsi

$$\begin{aligned} x + y - z &= 1 \\ y + (a+2)z &= 1 \\ (a-1)y + 4z &= 1 \end{aligned}$$

Eliminons maintenant y à partir de la troisième équation à l'aide de la combinaison $L_3 \to -(a-1)L_2 + L_3$:

$$\begin{array}{rrcl} -(a-1)L_2: & -(a-1)y + (2-a-a^2)z &=& 1-a \\ L_3: & (a-1)y + 4z &=& 1 \\ \hline & (6-a-a^2)z &=& 2-a \\ \text{ou} & (3+a)(2-a)z &=& 2-a \end{array}$$

nous obtenons le système équivalent

$$\begin{aligned} x + y - z &= 1 \\ y + (a+2)z &= 1 \\ (3+a)(2-a)z &= 2-a \end{aligned}$$

qui a une solution unique si le coefficient de z dans la troisième équation n'est pas nul , c'est-à-dire si $a \neq 2$ et $a \neq -3$. Dans le cas où $a = 2$ la troisième équation devient $0 = 0$ et le système a plus d'une solution. Dans le cas où $a = -3$, la troisième équation est $0 = 5$ et le système n'a pas de solution.

En résumé nous avons 1) $a = -3$ 2) $a = 2$ 3) $a \neq 2$ et $a \neq -3$.

2.7. Quelle condition doit-on avoir sur a, b et c, de telle sorte que le système suivant en x, y et z admette une solution ?

$$\begin{aligned} x + 2y - 3z &= a \\ 2x + 6y - 11z &= b \\ x - 2y + 7z &= c \end{aligned}$$

En réduisant à la forme échelonnée, et en éliminant x à l'aide de la seconde et de la troisième équation par les combinaisons $L_2 \to -2L_1 + L_2$ et $L_3 \to -L_1 + L_3$, nous obtenons le système équivalent

$$\begin{aligned} x + 2y - 3z &= a \\ 2y - 5z &= b - 2a \\ -4y + 10z &= c - a \end{aligned}$$

Eliminons y à partir de la troisième équation par la combinaison $L_3 \to 2L_2 + L_3$, nous obtenons finalement le système équivalent

$$\begin{aligned} x + 2y - 3z &= a \\ 2y - 5z &= b - 2a \\ 0 &= c + 2b - 5a \end{aligned}$$

Ce système n'a pas de solution si la troisième équation est de la forme $0 = k$ avec $k \neq 0$, c'est-à-dire si $c + 2b - 5a \neq 0$. En conclusion le système aura donc une solution si

$$c + 2b - 5a = 0 \quad \text{ou} \quad 5a = 2b + c$$

Remarquons, dans ce cas, que le système aura plus d'une solution. En d'autres termes, le système ne peut pas avoir une solution unique.

SYSTEMES HOMOGENES D'EQUATIONS LINEAIRES

2.8. Déterminer si chacun des systèmes suivants a une solution non nulle:

$$\begin{aligned} x - 2y + 3z - 2w &= 0 \\ 3x - 7y - 2z + 4w &= 0 \\ 4x + 3y + 5z + 2w &= 0 \end{aligned} \qquad \begin{aligned} x + 2y - 3z &= 0 \\ 2x + 5y + 2z &= 0 \\ 3x - y - 4z &= 0 \end{aligned} \qquad \begin{aligned} x + 2y - z &= 0 \\ 2x + 5y + 2z &= 0 \\ x + 4y + 7z &= 0 \\ x + 3y + 3z &= 0 \end{aligned}$$

(i) (ii) (iii)

(i) Le système doit avoir une solution non nulle puisqu'il y a plus d'inconnues que d'équations.

(ii) Réduisons à la forme échelonnée :

$$\begin{aligned} x + 2y - 3z &= 0 \\ 2x + 5y + 2z &= 0 \\ 3x - y - 4z &= 0 \end{aligned} \quad \text{et} \quad \begin{aligned} x + 2y - 3z &= 0 \\ y + 8z &= 0 \\ -7y + 5z &= 0 \end{aligned} \quad \text{et} \quad \begin{aligned} x + 2y - 3z &= 0 \\ y + 8z &= 0 \\ 61z &= 0 \end{aligned}$$

Dans la forme réduite échelonnée il y a trois équations à trois inconnues ; ainsi le système admet une solution unique, la solution nulle.

(iii) Réduisons à la forme échelonnée:

$$\begin{aligned} x + 2y - z &= 0 \\ 2x + 5y + 2z &= 0 \\ x + 4y + 7z &= 0 \\ x + 3y + 3z &= 0 \end{aligned} \qquad \begin{aligned} x + 2y - z &= 0 \\ y + 4z &= 0 \\ 2y + 8z &= 0 \\ y + 4z &= 0 \end{aligned} \qquad \begin{aligned} x + 2y - z &= 0 \\ y + 4z &= 0 \end{aligned}$$

Dans la forme échelonnée il y a deux équations à trois inconnues ; ainsi le système admet une solution non nulle.

2.9. Les vecteurs $u_1, u_2, \ldots, u_m$ de $\mathbf{R}^n$ sont dits *linéairement dépendants*, ou simplement *dépendants*, s'il existe des scalaires $k_1, k_2, \ldots, k_m$ non tous nuls de telle sorte que $k_1 u_1 + k_2 u_2 + \cdots + k_m u_m = 0$. Si ces conditions ne sont pas remplies on dit qu'ils sont *indépendants*. Déterminer si les vecteurs u, v, w sont dépendants ou indépendants avec

(i) $u = (1, 1, -1)$, $v = (2, -3, 1)$, $w = (8, -7, 1)$

(ii) $u = (1, -2, -3)$, $v = (2, 3, -1)$, $w = (3, 2, 1)$

(iii) $u = (a_1, a_2)$, $v = (b_1, b_2)$, $w = (c_1, c_2)$

Dans chaque cas

(*a*) Soit $xu + yv + zw = 0$ où x, y et z sont des scalaires inconnus.

(*b*) Il s'agit de trouver le système d'équations linéaires homogène équivalent.

(*c*) Enfin il faut déterminer si le système a une solution non nulle ; s'il en est ainsi, alors les vecteurs sont dépendants ; sinon les vecteurs sont indépendants.

(i) Soit $xu + yv + zw = 0$

$$x(1, 1, -1) + y(2, -3, 1) + z(8, -7, 1) = (0, 0, 0)$$

ou $$(x, x, -x) + (2y, -3y, y) + (8z, -7z, z) = (0, 0, 0)$$

ou $$(x + 2y + 8z,\ x - 3y - 7z,\ -x + y + z) = (0, 0, 0)$$

En égalant les composantes correspondantes les unes aux autres, et en réduisant le système à sa forme échelonnée :

$$\begin{array}{rcl} x+2y+8z&=&0\\ x-3y-7z&=&0\\ -x+y+z&=&0 \end{array} \qquad \begin{array}{rcl} x+2y+8z&=&0\\ -5y-15z&=&0\\ 3y+9z&=&0 \end{array} \qquad \begin{array}{rcl} x+2y+8z&=&0\\ y+3z&=&0\\ y+3z&=&0 \end{array} \qquad \begin{array}{rcl} x+2y+8z&=&0\\ y+3z&=&0 \end{array}$$

Dans la forme échelonnée, il y a seulement deux équations à trois inconnues ; ainsi le système a une solution non nulle. En conséquence les vecteurs sont dépendants.

Remarque : Nous n'avons pas besoin de résoudre le système pour déterminer la dépendance ou l'indépendance des vecteurs ; nous avons besoin uniquement de connaître s'il existe une solution non nulle.

(ii)
$$\begin{aligned} x(1,-2,-3)+y(2,3,-1)+z(3,2,1) &= (0,0,0)\\ (x,-2x,-3x)+(2y,3y,-y)+(3z,2z,z) &= (0,0,0)\\ (x+2y+3z,\ -2x+3y+2z,\ -3x-y+z) &= (0,0,0) \end{aligned}$$

$$\begin{array}{rcl} x+2y+3z&=&0\\ -2x+3y+2z&=&0\\ -3x-y+z&=&0 \end{array} \qquad \begin{array}{rcl} x+2y+3z&=&0\\ 7y+8z&=&0\\ 5y+10z&=&0 \end{array} \qquad \begin{array}{rcl} x+2y+3z&=&0\\ 7y+8z&=&0\\ 30z&=&0 \end{array}$$

Dans la forme échelonnée, il y a exactement trois équations à trois inconnues ; ainsi le système a seulement une solution nulle. Les vecteurs sont donc indépendants.

(iii)
$$\begin{aligned} x(a_1,a_2)+y(b_1,b_2)+z(c_1,c_2) &= (0,0)\\ (a_1x,a_2x)+(b_1y,b_2y)+(c_1z,c_2z) &= (0,0)\\ (a_1x+b_1y+c_1z,\ a_2x+b_2y+c_2z) &= (0,0) \end{aligned} \quad \text{et ainsi} \quad \begin{array}{rcl} a_1x+b_1y+c_1z&=&0\\ a_2x+b_2y+c_2z&=&0 \end{array}$$

Le système a une solution non nulle d'après le théorème 2-3, car il y a plus d'inconnues que d'équations, ainsi les vecteurs sont dépendants. En d'autres termes, nous avons prouvé que trois vecteurs quelconque de $\mathbf{R}^2$ *sont dépendants.*

2.10. Supposons que dans un système homogène d'équations linéaires les coefficients de l'une des inconnues soient tous nuls. Montrer que le système a une solution non nulle.

Supposons que $x_1, \ldots, x_n$ soient les inconnues d'un système et x_j est l'inconnue dont les coefficients sont tous nuls. Alors chaque équation de ce système est de la forme.

$$a_1x_1+\cdots+a_{j-1}x_{j-1}+0x_j+a_{j+1}x_{j+1}+\cdots+a_nx_n = 0$$

Par exemple $(0,\ldots,0,1,0,\ldots,0)$ où 1 est la $j^{\text{ième}}$ composante est une solution non nulle de chaque équation et donc de ce système.

PROBLEMES DIVERS

2.11. Démontrer le théorème 2.1 : Supposons que u soit une solution particulière du système homogène (*) et supposons que W soit une solution générale du système homogène associé (**). Alors

$$u+W = \{u+w:\ w\in W\}$$

est la solution générale du système non homogène (*).

Soit U la solution générale du système non homogène (*). Supposons que $u \in U$ et que $u = (u_1, \ldots u_n)$. Puisque u est une solution de (*) nous avons pour $i = 1, \ldots, m$,

$$a_{i1}u_1+a_{i2}u_2+\cdots+a_{in}u_n = b_i$$

Supposons maintenant que $w \in W$ et que $w = (w_1, \ldots, w_n)$. Puisque w est une solution du système homogène (**) nous avons pour $i = 1, \ldots, m$,

$$a_{i1}w_1+a_{i2}w_2+\cdots+a_{in}w_n = 0$$

Donc, pour $i = 1, \ldots, m$,

$$\begin{aligned} a_{i1}(u_1+w_1) + a_{i2}(u_2+w_2) + \cdots + a_{in}(u_n+w_n) \\ &= a_{i1}u_1 + a_{i1}w_1 + a_{i2}u_2 + a_{i2}w_2 + \cdots + a_{in}u_n + a_{in}w_n \\ &= (a_{i1}u_1 + a_{i2}u_2 + \cdots + a_{in}u_n) + (a_{i1}w_1 + a_{i2}w_2 + \cdots + a_{in}w_n) \\ &= b_i + 0 = b_i \end{aligned}$$

C'est-à-dire, $u + w$ est une solution de (*). Ainsi $u + w \in U$, et donc

$$u + W \subset U$$

Supposons maintenant que $v = (v_1, \ldots, v_n)$ est un élément arbitraire de U ; par exemple une solution de (*). Alors, pour $i = 1, \ldots, m$,

$$a_{i1}v_1 + a_{i2}v_2 + \cdots + a_{in}v_n = b_i$$

Remarquons que $v = u + (v - u)$; montrons que $v - u \in W$. Pour $i = 1, \ldots, m$,

$$\begin{aligned} a_{i1}(v_1-u_1) + a_{i2}(v_2-u_2) + \cdots + a_{in}(v_n-u_n) \\ &= (a_{i1}v_1 + a_{i2}v_2 + \cdots + a_{in}v_n) - (a_{i1}u_1 + a_{i2}u_2 + \cdots + a_{in}u_n) \\ &= b_i - b_i = 0 \end{aligned}$$

Ainsi $v - u$ est une solution du système homogène (*), c'est-à-dire $v - u \in W$. Alors $v \in u + W$ et donc

$$U \subset u + W$$

Les deux relations d'inclusion donnent $U = u + W$; c'est-à-dire $u + W$ est une solution générale du système non homogène (**).

2.12. Considérons le système (*) d'équations linéaires (page 18). Multiplions la $i^{\text{ème}}$ équation par c_i, et additionnons ; nous obtenons l'équation

$$(c_1a_{11} + \cdots + c_ma_{m1})x_1 + \cdots + (c_1a_{1n} + \cdots + c_ma_{mn})x_n = c_1b_1 + \cdots + c_mb_m \quad (1)$$

Une telle équation est appelée ***combinaison linéaire*** des équations de (*). Montrer que toute solution de (*) est aussi une solution de la combinaison linéaire (1).

Supposons que $u = (k_1, \ldots, k_n)$ soit une solution de (*). Alors on a

$$a_{i1}k_1 + a_{i2}k_2 + \cdots + a_{in}k_n = b_i, \qquad i = 1, \ldots, m \qquad (2)$$

Pour montrer que u est une solution de (*1*), nous devons vérifier l'équation

$$(c_1a_{11} + \cdots + c_ma_{m1})k_1 + \cdots + (c_1a_{1n} + \cdots + c_ma_{mn})k_n = c_1b_1 + \cdots + c_mb_m$$

Mais ceci peut être mis sous la forme

$$c_1(a_{11}k_1 + \cdots + a_{1n}k_n) + \cdots + c_m(a_{m1} + \cdots + a_{mn}k_n) = c_1b_1 + \cdots + c_mb_m$$

ou d'après (*2*), $\qquad c_1b_1 + \cdots + c_mb_m = c_1b_1 + \cdots + c_mb_m$

qui est une égalité vraie.

2.13. Dans le système (*) d'équations linéaires, supposons $a_{11} \neq 0$. Soit (#) le système obtenu à partir de (*) par la combinaison $L_i \to -a_{i1}L_1 + a_{11}L_i$, $i \neq 1$. Montrer que (*) et (#) sont des systèmes équivalents, c'est-à-dire ont le même ensemble de solutions.

D'après la combinaison précédente sur (*), chaque équation de (#) est une combinaison linéaire des équations de (*) ; ainsi d'après le problème précédent toute solution de (*) est aussi une solution de (#).

D'autre part, en appliquant la combinaison $L_1 \to \frac{1}{a_{11}}(-a_{i1}L_1 + L_i)$ à (#), nous obtenons le système initial (*). c'est-à-dire que chaque équation de (*) est une combinaison linéaire des équations de (#) ; ainsi chaque solution de (#) est aussi une solution de (*).

Les deux conditions montrent que (*) et (#) ont le même ensemble de solutions.

2.14. Démontrer le théorème 2.2 : Considérons un système réduit à sa forme échelonnée :

$$\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots\cdots\cdots\cdots + a_{1n}x_n &=& b_1 \\ a_{2j_2}x_{j_2} + a_{2,j_2+1}x_{j_2+1} + \cdots\cdots + a_{2n}x_n &=& b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots && \\ a_{rj_r}x_{j_r} + a_{r,j_r+1}x_{j_r+1} + \cdots + a_{rn}x_n &=& b_r \end{array}$$

où $1 < j_2 < \cdots < j_r$ et où $a_{11} \neq 0$, $a_{2j_2} \neq 0, \ldots, a_{rj_r} \neq 0$. La solution est donnée de la manière suivante. Il y a deux cas :

(i) $r = n$. Alors le système a une solution unique.

(ii) $r < n$. Alors nous pouvons donner aux $n - r$ inconnues libres des valeurs déterminées et nous obtenons une solution du système.

La démonstration se fait par récurrence sur le nombre r d'équations du système. Si $r = 1$, nous avons une seule équation linéaire

$$a_1x_1 + a_2x_2 + a_3x_3 + \cdots + a_nx_n = b, \qquad \text{où} \qquad a_1 \neq 0$$

Les inconnues libres sont $x_2, \ldots, x_r$. Donnons arbitrairement des valeurs aux inconnues libres ; par exemple $x_2 = k_2$, $x_3 = k_3, \ldots, x_n = k_n$. En remplaçant $x_2, x_3, \ldots, x_n$ par leurs valeurs dans l'équation et en résolvant en x_1, on a

$$x_1 = \frac{1}{a_1}(b - a_2k_2 - a_3k_3 - \cdots - a_nk_n)$$

Ces valeurs constituent une solution de l'équation; en remplaçant x_1 par la valeur précédente, nous obtenons

$$a_1\left[\frac{1}{a_1}(b - a_2k_2 - \cdots - a_nk_n)\right] + a_2k_2 + \cdots + a_nk_n = b \qquad \text{ou} \qquad b = b$$

qui nous donne une relation vérifiée.

De plus si $r = n = 1$, nous avons $ax = b$ où $a \neq 0$. Remarquons que $x = b/a$ est une solution puisque $a(b/a) = b$ est vraie. De plus si $x = k$ est une solution, par exemple $ak = b$, alors $k = b/a$. Ainsi l'équation a une solution unique.

Maintenant supposons que $r > 1$ et que le théorème soit vrai pour un système de $r - 1$ équations. Nous considérons les $r - 1$ équations

$$\begin{array}{rcl} a_{2j_2}x_{j_2} + a_{2,j_2+1}x_{j_2+1} + \cdots\cdots\cdots\cdots + a_{2n}x_n &=& b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots && \\ a_{rj_r}x_{j_r} + a_{r,j_r+1}x_{j_r+1} + \cdots + a_{rn}x_n &=& b_r \end{array}$$

comme un système d'inconnues $x_{j_2}, \ldots, x_n$. Remarquons que ce système est mis sous la forme réduite échelonnée. Par récurrence, nous pouvons donner arbitrairement aux $(n - j_2 + 1) - (r - 1)$ inconnues libres des valeurs dans le système réduit pour obtenir une solution (par exemple $x_{j_2} = k_{j_2}, \ldots, x_n = k_n$). Comme dans le cas $r = 1$ ces valeurs, ainsi que les valeurs arbitraires données aux $j_2 - 2$ inconnues libres (c'est-à-dire $x_2 = k_2, \ldots, x_{j_2} - 1 = k_{j_2-1}$) forment une solution de la première équation avec

$$x_1 = \frac{1}{a_{11}}(b_1 - a_{12}k_2 - \cdots - a_{1n}k_n)$$

(Remarquons qu'il y a $(n - j_2 + 1) - (r - 1) - (j_2 - 2) = n - r$ inconnues libres). De plus ces valeurs pour $x_1, \ldots, x_n$ satisfont aussi les autres équations puisque dans ces équations les coefficients de $x_1, \ldots, x_{j_2-1}$ sont nuls.

Enfin si $r = n$, alors $j_2 = 2$. Ainsi par récurrence nous obtenons une solution unique du système partiel et donc une solution unique du système complet. En conséquence le théorème est démontré.

2.15. Un système (*) d'équations linéaires est dit *possible*, si aucune combinaison linéaire de ses équations ne donne l'équation

$$0x_1 + 0x_2 + \cdots + 0x_n = b, \qquad \text{où} \qquad b \neq 0 \tag{1}$$

Montrer que le système (*) est possible si et seulement si il est réductible à une forme échelonnée.

Supposons (*) réductible à une forme échelonnée. Alors il a une solution qui, d'après le problème 2-12, est une solution de chaque combinaison linéaire de ses équations. Puisque (1) n'a pas de solution, il ne peut être une combinaison linéaire des équations de (*). Donc (*) est possible.

D'autre part, supposons (*) non réductible à une forme échelonnée. Alors dans le procédé de réduction il doit donner une équation de la forme (1). C'est-à-dire que (1) est une combinaison linéaire des équations de (*). Donc (*) n'est pas possible, d'où impossible.

PROBLEMES SUPPLEMENTAIRES

SOLUTION D'EQUATIONS LINEAIRES

2.16. Résoudre (i) $\begin{aligned} 2x + 3y &= 1 \\ 5x + 7y &= 3 \end{aligned}$ (ii) $\begin{aligned} 2x + 4y &= 10 \\ 3x + 6y &= 15 \end{aligned}$ (iii) $\begin{aligned} 4x - 2y &= 5 \\ -6x + 3y &= 1 \end{aligned}$

2.17. Résoudre

(i) $\begin{aligned} 2x + y - 3z &= 5 \\ 3x - 2y + 2z &= 5 \\ 5x - 3y - z &= 16 \end{aligned}$ (ii) $\begin{aligned} 2x + 3y - 2z &= 5 \\ x - 2y + 3z &= 2 \\ 4x - y + 4z &= 1 \end{aligned}$ (iii) $\begin{aligned} x + 2y + 3z &= 3 \\ 2x + 3y + 8z &= 4 \\ 3x + 2y + 17z &= 1 \end{aligned}$

2.18 Résoudre

(i) $\begin{aligned} 2x + 3y &= 3 \\ x - 2y &= 5 \\ 3x + 2y &= 7 \end{aligned}$ (ii) $\begin{aligned} x + 2y - 3z + 2w &= 2 \\ 2x + 5y - 8z + 6w &= 5 \\ 3x + 4y - 5z + 2w &= 4 \end{aligned}$ (iii) $\begin{aligned} x + 2y - z + 3w &= 3 \\ 2x + 4y + 4z + 3w &= 9 \\ 3x + 6y - z + 8w &= 10 \end{aligned}$

2.19. Résoudre (i) $\begin{aligned} x + 2y + 2z &= 2 \\ 3x - 2y - z &= 5 \\ 2x - 5y + 3z &= -4 \\ x + 4y + 6z &= 0 \end{aligned}$ (ii) $\begin{aligned} x + 5y + 4z - 13w &= 3 \\ 3x - y + 2z + 5w &= 2 \\ 2x + 2y + 3z - 4w &= 1 \end{aligned}$

2.20. Déterminer les valeurs de k de telle sorte que le système d'inconnues x, y, z ait (i) une solution unique (ii) aucune solution (iii) plus d'une solution.

(a) $\begin{aligned} kx + y + z &= 1 \\ x + ky + z &= 1 \\ x + y + kz &= 1 \end{aligned}$ (b) $\begin{aligned} x + 2y + kz &= 1 \\ 2x + ky + 8z &= 3 \end{aligned}$

2.21. Déterminer les valeurs de k de telle sorte que le système d'inconnues x, y et z ait (i) une solution unique. (ii) aucune solution,(iii) plus d'une solution:

(a) $\begin{aligned} x + y + kz &= 2 \\ 3x + 4y + 2z &= k \\ 2x + 3y - z &= 1 \end{aligned}$ (b) $\begin{aligned} x \quad - 3z &= -3 \\ 2x + ky - z &= -2 \\ x + 2y + kz &= 1 \end{aligned}$

2.22. Déterminer les conditions sur a, b, c de telle sorte que le système ayant pour inconnues x, y et z ait une solution:

(i) $\begin{aligned} x + 2y - 3z &= a \\ 3x - y + 2z &= b \\ x - 5y + 8z &= c \end{aligned}$ (ii) $\begin{aligned} x - 2y + 4z &= a \\ 2x + 3y - z &= b \\ 3x + y + 2z &= c \end{aligned}$

SYSTEMES HOMOGENES

2.23. Déterminer si chacun des systèmes suivants a une solution non nulle:

(i) $\begin{aligned} x + 3y - 2z &= 0 \\ x - 8y + 8z &= 0 \\ 3x - 2y + 4z &= 0 \end{aligned}$ (ii) $\begin{aligned} x + 3y - 2z &= 0 \\ 2x - 3y + z &= 0 \\ 3x - 2y + 2z &= 0 \end{aligned}$ (iii) $\begin{aligned} x + 2y - 5z + 4w &= 0 \\ 2x - 3y + 2z + 3w &= 0 \\ 4x - 7y + z - 6w &= 0 \end{aligned}$

2.24. Déterminer si chacun des systèmes suivants a une solution non nulle:

$$\text{(i)}\quad \begin{aligned} x - 2y + 2z &= 0 \\ 2x + y - 2z &= 0 \\ 3x + 4y - 6z &= 0 \\ 3x - 11y + 12z &= 0 \end{aligned} \qquad \text{(ii)}\quad \begin{aligned} 2x - 4y + 7z + 4v - 5w &= 0 \\ 9x + 3y + 2z - 7v + w &= 0 \\ 5x + 2y - 3z + v + 3w &= 0 \\ 6x - 5y + 4z - 3v - 2w &= 0 \end{aligned}$$

2.25. Déterminer ceux des vecteurs u, v, w qui sont dépendants ou indépendants [voir le problème 2-9) où :

(i) $u = (1, 3, -1)$, $v = (2, 0, 1)$, $w = (1, -1, 1)$

(ii) $u = (1, 1, -1)$, $v = (2, 1, 0)$, $w = (-1, 1, 2)$

(iii) $u = (1, -2, 3, 1)$, $v = (3, 2, 1, -2)$, $w = (1, 6, -5, -4)$

PROBLEMES DIVERS

2.26. Soient deux équations linéaires générales à deux inconnues x et y appartenant au corps $\mathbf{R}$:

$$\begin{aligned} ax + by &= e \\ cx + dy &= f \end{aligned}$$

Montrer que

(i) si $\dfrac{a}{c} \neq \dfrac{b}{d}$ c'est-à-dire $ad - bc \neq 0$, alors le système admet une solution unique $x = \dfrac{de - bf}{ad - bc}$, $y = \dfrac{af - ce}{ad - bc}$;

(ii) si $\dfrac{a}{c} = \dfrac{b}{d} \neq \dfrac{e}{f}$, alors le système n'a pas de solution;

(iii)· si $\dfrac{a}{c} = \dfrac{b}{d} = \dfrac{e}{f}$, alors le système a plus d'une solution.

2.27. Soit le système

$$\begin{aligned} ax + by &= 1 \\ cx + dy &= 0 \end{aligned}$$

Montrer que si $ad - bc \neq 0$, alors le système a une solution unique $x = d/(ad - bc)$, $y = - c/(ad - bc)$. Montrer aussi que si $ad - bc = 0$, $c \neq 0$ ou $d \neq 0$,alors le système n'a pas de solution.

2.28. Montrer qu'une équation de la forme $0x_1 + 0x_2 + \cdots + 0x_n = 0$ peut être ajoutée ou supprimée d'un système sans affecter l'ensemble des solutions.

2.29. Soit un système d'équations linéaires avec le même nombre d'équations que d'inconnues ;

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots& \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n \end{aligned} \qquad (1)$$

(i) On suppose que le système homogène associé admet seulement la solution nulle. Montrer que (1) a une solution unique pour chaque choix des constantes b_i.

(ii) On suppose que le système homogène associé à une solution nulle. Montrer qu'il existe des constantes b_i pour lesquelles (1) n'a pas de solution. Montrer aussi que si (1) a une solution, alors il en a plus d'une.

REPONSES AUX PROBLEMES SUPPLEMENTAIRES

2.16. (i) $x = 2,\ y = -1$; (ii) $x = 5 - 2a,\ y = a$; (iii) aucune solution

2.17. (i) $(1, -3, -2)$; (ii) aucune solution; (iii) $(-1 - 7a,\ 2 + 2a,\ a)$ ou $\begin{cases} x = -1 - 7z \\ y = 2 + 2z \end{cases}$

2.18. (i) $x = 3,\ y = -1$

(ii) $(-a + 2b,\ 1 + 2a - 2b,\ a,\ b)$ ou $\begin{cases} x = -z + 2w \\ y = 1 + 2z - 2w \end{cases}$

(iii) $(7/2 - 5b/2 - 2a,\ a,\ 1/2 + b/2,\ b)$ ou $\begin{cases} x = 7/2 - 5w/2 - 2y \\ z = 1/2 + w/2 \end{cases}$

2.19. (i) $(2,\ 1,\ -1)$; (ii) aucune solution.

2.20. (*a*) (i) $k \neq 1$ et $k \neq -2$; (ii) $k = -2$; (iii) $k = 1$.

(*b*) (i) n'a jamais une solution unique ; (ii) $k = 4$; (iii) $k \neq 4$.

2.21. (*a*) (i) $k \neq 3$; (ii) a toujours une solution ; (iii) $k = 3$.

(*b*) (i) $k \neq 2$ et $k \neq -5$; (ii) $k = -5$; (iii) $k = 2$.

2.22. (i) $2a - b + c = 0$. (ii) N'importe quelle valeur de a, b et c forme une solution

2.23. (i) oui; (ii) non; (iii) oui d'après le théorème 2.3.

2.24. (i) oui; (ii) oui d'après le théorème 2.3.

2.25. (i) dépendants; (ii) indépendants; (iii) dépendants.

CHAPITRE 3

Matrices

INTRODUCTION

Nous avons remarqué en travaillant sur les systèmes d'équations linéaires que seulement les coefficients et leurs positions respectives sont importants. De même, en réduisant les systèmes à leurs formes échelonnées, il est essentiel de garder les équations strictement alignées. Les coefficients peuvent donc être rassemblés dans un tableau rectangulaire appelé une *matrice*. De plus, certaines notions abstraites introduites dans les chapitres suivants, comme les "changements de bases", les "opérateurs linéaires" et les "formes bilinéaires", peuvent être aussi représentées par ces tableaux rectangulaires ou matrices.

Dans ce chapitre, nous étudierons ces matrices et certaines opérations algébriques définies entre elles. Il s'agira principalement dans ce chapitre de calculs. Cependant, comme pour les équations linéaires, les méthodes abstraites données plus loin nous permettront d'avoir un nouvel aperçu de la structure de ces matrices.

A moins qu'il n'en soit spécifié autrement, les matrices que nous considérerons appartiendront à un corps fixé arbitraire K. (Voir l'Appendice B). Les éléments de K sont appelés des *scalaires*. Le lecteur pourra supposer que K est, soit le corps réel $\mathbf{R}$ soit le corps complexe $\mathbf{C}$.

Plus tard, nous remarquerons que les "vecteurs lignes" ou les "vecteurs colonnes", qui sont des matrices particulières, sont des éléments de $\mathbf{R}^n$ ou de $\mathbf{C}^n$.

MATRICES

Soit K un corps arbitraire. Un tableau rectangulaire de la forme

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

où les a_{ij} sont des scalaires du corps K, est appelé *une matrice sur K*, ou simplement une *matrice* si K est implicite. La matrice précédente est aussi notée par (a_{ij}), $i = 1, \dots, m$, $j = 1, \dots, n$, ou simplement par (a_{ij}). Les m, n-tuples horizontaux

$$(a_{11}, a_{12}, \dots, a_{1n}), \ (a_{21}, a_{22}, \dots, a_{2n}), \ \dots, \ (a_{m1}, a_{m2}, \dots, a_{mn})$$

sont les *lignes* de la matrice, et les n, m-tuples verticaux

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \cdot\,\cdot \\ a_{m1} \end{pmatrix}, \quad \begin{pmatrix} a_{12} \\ a_{22} \\ \cdot\,\cdot \\ a_{m2} \end{pmatrix}, \quad \dots, \quad \begin{pmatrix} a_{1n} \\ a_{2n} \\ \cdot\,\cdot \\ a_{mn} \end{pmatrix}$$

sont ses *colonnes*. Remarquons que l'élément a_{ij}, appelé le *ij*-élément ou la *ij*$^{\text{ème}}$ *composante*, se trouve à l'intersection de la $i^{\text{ième}}$ ligne et de la $j^{\text{ième}}$ colonne. Une matrice ayant m lignes et n colonnes est appelée une (m, n) matrice, ou $m \times n$ matrice ; le couple de nombres (m, n) est appelé la *dimension* de la matrice ; n est sa largeur, m sa hauteur.

Exemple 3.1 : Soit la matrice 2×3 : $\begin{pmatrix} 1 & -3 & 4 \\ 0 & 5 & -2 \end{pmatrix}$.

Ses lignes sont $(1, -3, 4)$ et $(0, 5, -2)$; ses colonnes sont $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\begin{pmatrix} -3 \\ 5 \end{pmatrix}$ et $\begin{pmatrix} 4 \\ -2 \end{pmatrix}$.

Les matrices seront notées habituellement par des lettres capitales A, B . . . , et les éléments du corps K par des lettres minuscules a, b,. . . Deux matrices A et B sont *égales*, et on écrit $A = B$, si elles ont même dimension et si leurs éléments correspondants sont égaux. Ainsi l'égalité de deux $m \times n$ matrices est équivalente à un système de mn égalités, une pour chaque paire d'indices.

Exemple 3.2 : L'égalité $\begin{pmatrix} x+y & 2z+w \\ x-y & z-w \end{pmatrix} = \begin{pmatrix} 3 & 5 \\ 1 & 4 \end{pmatrix}$ est équivalente au système suivant d'équations :

$$\begin{cases} x + y = 3 \\ x - y = 1 \\ 2z + w = 5 \\ z - w = 4 \end{cases}$$

La solution de ce système est $x = 2$, $y = 1$, $z = 3$, $w = -1$.

Remarque : Une matrice à une ligne peut être considérée comme un *vecteur ligne*, et une matrice à une colonne comme un *vecteur colonne.* En particulier un élément du corps K peut être interprété comme une matrice 1×1.

ADDITION DES MATRICES ET MULTIPLICATION PAR UN SCALAIRE

Soit A et B deux matrices de même dimension, c'est-à-dire ayant le même nombre de lignes et de colonnes, c.à.d. deux $m \times n$ matrices :

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \dots & \dots & \dots & \dots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{pmatrix}$$

La *somme* de A et B, écrite $A + B$, est la matrice obtenue en ajoutant les éléments correspondants des deux matrices :

$$A + B = \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \dots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \dots & a_{2n}+b_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \dots & a_{mn}+b_{mn} \end{pmatrix}$$

Le *produit* d'une matrice A par un scalaire k, noté $k \, . \, A$, ou kA, est la matrice obtenue en multipliant chaque élément de la matrice A par k :

$$kA = \begin{pmatrix} ka_{11} & ka_{12} & \dots & ka_{1n} \\ ka_{21} & ka_{22} & \dots & ka_{2n} \\ \dots & \dots & \dots & \dots \\ ka_{m1} & ka_{m2} & \dots & ka_{mn} \end{pmatrix}$$

Remarquons que $A + B$ et kA sont aussi des $m \times n$ matrices. On définit aussi

$$-A = -1 \cdot A \qquad \text{et} \qquad A - B = A + (-B)$$

La somme de deux matrices de dimensions différentes n'est pas définie.

Exemple 3.3 : Soient $A = \begin{pmatrix} 1 & -2 & 3 \\ 4 & 5 & -6 \end{pmatrix}$ et $B = \begin{pmatrix} 3 & 0 & 2 \\ -7 & 1 & 8 \end{pmatrix}$. Alors

$$A + B = \begin{pmatrix} 1+3 & -2+0 & 3+2 \\ 4-7 & 5+1 & -6+8 \end{pmatrix} = \begin{pmatrix} 4 & -2 & 5 \\ -3 & 6 & 2 \end{pmatrix}$$

$$3A = \begin{pmatrix} 3\cdot 1 & 3\cdot(-2) & 3\cdot 3 \\ 3\cdot 4 & 3\cdot 5 & 3\cdot(-6) \end{pmatrix} = \begin{pmatrix} 3 & -6 & 9 \\ 12 & 15 & -18 \end{pmatrix}$$

$$2A - 3B = \begin{pmatrix} 2 & -4 & 6 \\ 8 & 10 & -12 \end{pmatrix} + \begin{pmatrix} -9 & 0 & -6 \\ 21 & -3 & -24 \end{pmatrix} = \begin{pmatrix} -7 & -4 & 0 \\ 29 & 7 & -36 \end{pmatrix}$$

Exemple 3.4. : La $m \times n$ matrice dont les éléments sont tous nuls,

$$\begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \end{pmatrix}$$

est appelée la *matrice nulle*, et sera notée 0. Elle est semblable au scalaire 0 en ce sens que, pour toute $m \times n$ matrice $A = (a_{ij})$, $A + 0 = (a_{ij} + 0) = (a_{ij}) = A$.

Les principales propriétés des matrices se déduisant de l'addition et de la multiplication par un scalaire sont :

Théorème 3.1. : Soit V l'ensemble de toutes les $m \times n$ matrices sur un corps K. Quelles que soient les matrices $A, B, C \in V$ et quels que soient les scalaires $k_1, k_2 \in K$.

(i) $(A+B)+C = A+(B+C)$	(v) $k_1(A+B) = k_1A + k_1B$
(ii) $A+0 = A$	(vi) $(k_1+k_2)A = k_1A + k_2A$
(iii) $A+(-A) = 0$	(vii) $(k_1k_2)A = k_1(k_2A)$
(iv) $A+B = B+A$	(viii) $1\cdot A = A$ et $0A = 0$

En utilisant (vi) et (vii) on a aussi $A + A = 2A$, $A + A + A = 3A$, ...

Remarque : Supposons que les vecteurs de R^n soient représentés par des vecteurs lignes (ou des vecteurs colonnes), c'est-à-dire

$$u = (a_1, a_2, \dots, a_n) \quad \text{et} \quad v = (b_1, b_2, \dots, b_n)$$

En considérant ces vecteurs comme des matrices, la somme $u + v$ et le produit par un scalaire ku sont :

$$u + v = (a_1 + b_1, a_2 + b_2, \dots, a_n + b_n) \quad \text{et} \quad ku = (ka_1, ka_2, \dots, ka_n)$$

Ces résultats correspondent précisément à la somme vectorielle et au produit par un scalaire tels qu'ils sont définis dans le chapitre I. En d'autres termes les opérations précédentes sur les matrices peuvent être considérées comme une généralisation des opérations correspondantes définies dans le chapitre I.

MULTIPLICATION DES MATRICES

Le produit de matrices A et B, écrit AB, est quelque peu compliqué. Pour cette raison, nous ferons les remarques suivantes.

(i) Soit $A = (a_i)$ et $B = (b_i)$ appartenant à $\mathbf{R}^n$, A étant représenté par un vecteur ligne et B par un vecteur colonne. Leur produit scalaire $A \cdot B$ peut être trouvé en combinant les matrices de la façon suivante :

$$A \cdot B = (a_1, a_2, \dots, a_n)\begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_n \end{pmatrix} = a_1b_1 + a_2b_2 + \cdots + a_nb_n$$

Nous définirons donc la matrice produit d'un vecteur ligne A par un vecteur colonne B comme précédemment.

(ii) Considérons les équations

$$\begin{aligned} b_{11}x_1 + b_{12}x_2 + b_{13}x_3 &= y_1 \\ b_{21}x_1 + b_{22}x_2 + b_{23}x_3 &= y_2 \end{aligned} \qquad (1)$$

Ce système est équivalent à l'équation matricielle

$$\begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \quad \text{ou simplement } BX = Y$$

où $B = (b_{ij})$, $X = (x_i)$ et $Y = (y_i)$, si nous multiplions la matrice B et le vecteur colonne X comme suit :

$$BX = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_{11}x_1 + b_{12}x_2 + b_{13}x_3 \\ b_{21}x_1 + b_{22}x_2 + b_{23}x_3 \end{pmatrix} = \begin{pmatrix} B_1 \cdot X \\ B_2 \cdot X \end{pmatrix}$$

où B_1 et B_2 sont les lignes de B; on peut remarquer que le produit d'une matrice et d'un vecteur colonne donne un autre vecteur colonne.

(iii) Considérons maintenant les équations

$$\begin{aligned} a_{11}y_1 + a_{12}y_2 &= z_1 \\ a_{21}y_1 + a_{22}y_2 &= z_2 \end{aligned} \qquad (2)$$

qui peuvent être représentées par l'équation matricielle

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \quad \text{ou simplement } AY = Z$$

où $A = (a_{ij})$, $Y = (y_i)$ comme plus haut et $Z = (z_i)$. En remplaçant y_1 et y_2 par leurs valeurs données par (1) dans les équations (2) on obtient

$$\begin{aligned} a_{11}(b_{11}x_1 + b_{12}x_2 + b_{13}x_3) + a_{12}(b_{21}x_1 + b_{22}x_2 + b_{23}x_3) &= z_1 \\ a_{21}(b_{11}x_1 + b_{12}x_2 + b_{13}x_3) + a_{22}(b_{21}x_1 + b_{22}x_2 + b_{23}x_3) &= z_2 \end{aligned}$$

où en ordonnant par rapport à x_1, x_2, x_3

$$\begin{aligned} (a_{11}b_{11} + a_{12}b_{21})x_1 + (a_{11}b_{12} + a_{12}b_{22})x_2 + (a_{11}b_{13} + a_{12}b_{23})x_3 &= z_1 \\ (a_{21}b_{11} + a_{22}b_{21})x_1 + (a_{21}b_{12} + a_{22}b_{22})x_2 + (a_{21}b_{13} + a_{22}b_{23})x_3 &= z_2 \end{aligned} \qquad (3)$$

D'une autre manière, en utilisant l'équation matricielle $BX = Y$ et en remplaçant Y par BX dans l'équation matricielle $AY = Z$ on obtient l'expression

$$ABX = Z$$

Cette équation matricielle représente le système (3), si nous définissons le produit de A et B de la manière suivante :

$$\begin{aligned} AB &= \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} & a_{11}b_{13} + a_{12}b_{23} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} & a_{21}b_{13} + a_{22}b_{23} \end{pmatrix} \\ &= \begin{pmatrix} A_1 \cdot B^1 & A_1 \cdot B^2 & A_1 \cdot B^3 \\ A_2 \cdot B^1 & A_2 \cdot B^2 & A_2 \cdot B^3 \end{pmatrix} \end{aligned}$$

où A_1 et A_2 sont les lignes de A et B^1 et B^2 et B^3 les colonnes de B. Nous pouvons donc dire que si ces calculs sont généraux, il est nécessaire que le nombre des y_i dans (1) et (2) reste le même. Ceci correspond au fait que le nombre de colonnes de la matrice A doit être égal au nombre de lignes de la matrice B.

A l'aide de la précédente introduction, nous pouvons définir formellement la multiplication des matrices.

Définition : Supposons que $A = (a_{ij})$ et $B = (b_{ij})$ soient des matrices dont le nombre de colonnes de A soit égal au nombre de lignes de B ; c'est-à-dire A est une $m \times p$ matrice et B est une $p \times n$ matrice. Alors, le produit AB est une $m \times n$ matrice où le ij élément est obtenu en multipliant la $i$ème ligne A_i de A par la $j$ème colonne B^j de B:

$$AB = \begin{pmatrix} A_1 \cdot B^1 & A_1 \cdot B^2 & \dots & A_1 \cdot B^n \\ A_2 \cdot B^1 & A_2 \cdot B^2 & \dots & A_2 \cdot B^n \\ \dots & \dots & \dots & \dots \\ A_m \cdot B^1 & A_m \cdot B^2 & \dots & A_m \cdot B^n \end{pmatrix}$$

D'où

$$\begin{pmatrix} a_{11} & \dots & a_{1p} \\ . & \dots & . \\ a_{i1} & \dots & a_{ip} \\ . & \dots & . \\ a_{m1} & \dots & a_{mp} \end{pmatrix} \begin{pmatrix} b_{11} & \dots & b_{1j} & \dots & b_{1n} \\ . & \dots & . & \dots & . \\ . & \dots & . & \dots & . \\ . & \dots & . & \dots & . \\ b_{p1} & \dots & b_{pj} & \dots & b_{pn} \end{pmatrix} = \begin{pmatrix} c_{11} & \dots & c_{1n} \\ . & \dots & . \\ . & c_{ij} & . \\ . & \dots & . \\ c_{m1} & \dots & c_{mn} \end{pmatrix}$$

où $\quad c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ip}b_{pj} = \sum_{k=1}^{p} a_{ik}b_{kj}.$

Il est important de remarquer que le produit AB n'est pas défini si A est une $m \times p$ matrice et B une $q \times n$ matrice, où $p \neq q$.

Exemple 3.5 :
$$\begin{pmatrix} r & s \\ t & u \end{pmatrix} \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix} = \begin{pmatrix} ra_1 + sb_1 & ra_2 + sb_2 & ra_3 + sb_3 \\ ta_1 + ub_1 & ta_2 + ub_2 & ta_3 + ub_3 \end{pmatrix}$$

Exemple 3.6 :
$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 \cdot 1 + 2 \cdot 0 & 1 \cdot 1 + 2 \cdot 2 \\ 3 \cdot 1 + 4 \cdot 0 & 3 \cdot 1 + 4 \cdot 2 \end{pmatrix} = \begin{pmatrix} 1 & 5 \\ 3 & 11 \end{pmatrix}$$
$$\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 \cdot 1 + 1 \cdot 3 & 1 \cdot 2 + 1 \cdot 4 \\ 0 \cdot 1 + 2 \cdot 3 & 0 \cdot 2 + 2 \cdot 4 \end{pmatrix} = \begin{pmatrix} 4 & 6 \\ 6 & 8 \end{pmatrix}$$

Les exemples précédents montrent que la multiplication des matrices n'est pas commutative ; les produits AB et BA des matrices ne sont pas égaux.

De plus la multiplication des matrices satisfait aux propriétés suivantes :

Théorème 3.2 : (i) $(AB)C = A(BC)$ (associativité)

(ii) $A(B + C) = AB + AC$ (distributivité à gauche)

(iii) $(B + C)A = BA + CA$ (distributivité à droite)

(iv) $k(AB) = (kA)B = A(kB)$ où k est un scalaire

Nous supposons que les sommes et produits précédents sont définis.

Remarquons que $0A = 0$ et $B0 = 0$ où 0 est la matrice nulle.

TRANSPOSITION

La matrice *transposée* d'une matrice A, écrite A^t, est la matrice obtenue en écrivant les lignes de A en colonnes :

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}^t = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \dots & \dots & \dots & \dots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{pmatrix}$$

Observons que si A est une matrice $m \times n$, alors A^t est une $n \times m$ matrice.

Exemple 3-7 : $$\begin{pmatrix} 1 & 2 & 3 \\ 4 & -5 & -6 \end{pmatrix}^t = \begin{pmatrix} 1 & 4 \\ 2 & -5 \\ 3 & -6 \end{pmatrix}$$

La transposition des matrices satisfait aux propriétés suivantes:

Théorème 3-3 :

(i) $(A+B)^t = A^t + B^t$

(ii) $(A^t)^t = A$

(iii) $(kA)^t = kA^t$, pour k scalaire

(iv) $(AB)^t = B^tA^t$

MATRICES ET SYSTEMES D'EQUATIONS LINEAIRES

Le système suivant d'équations linéaires

$$\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{array} \qquad (1)$$

est équivalent à l'équation matricielle

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \cdot\cdot \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \cdot\cdot \\ b_m \end{pmatrix} \quad \text{ou simplement } AX = B \qquad (2)$$

où $A = (a_{ij})$, $X = (x_i)$ et $B = (b_i)$. En somme, chaque solution du système (*1*) est une solution de l'équation matricielle (*2*) est vice versa. Observons que le système homogène associé au système (*1*) est alors équivalent à l'équation matricielle $AX = 0$.

La matrice A précédente est appelée la *matrice des coefficients* du système (*1*) et la matrice

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{pmatrix}$$

est appelée *matrice augmentée* de (*1*). Le système (*1*) est complètement défini par sa matrice augmentée.

Exemple 3-8 : La matrice des coefficients et la matrice augmentée du système

$$\begin{array}{r} 2x + 3y - 4z = 7 \\ x - 2y - 5z = 3 \end{array}$$

sont respectivement les matrices suivantes :

$$\begin{pmatrix} 2 & 3 & -4 \\ 1 & -2 & -5 \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} 2 & 3 & -4 & 7 \\ 1 & -2 & -5 & 3 \end{pmatrix}$$

Observons que le système est équivalent à l'équation matricielle

$$\begin{pmatrix} 2 & 3 & -4 \\ 1 & -2 & -5 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 7 \\ 3 \end{pmatrix}$$

En étudiant les équations linéaires, il est habituellement plus simple d'employer le langage et la théorie des matrices, comme le montrent les théorèmes suivants.

Théorème 3.4 : Supposons que $u_1, u_2, \ldots, u_n$ soient les solutions d'un système homogène d'équations linéaires $AX = 0$. Chaque combinaison linéaire des u_i de la forme $k_1u_1 + k_2u_2 + \cdots + k_nu_n$, où les k_i sont des scalaires, est aussi une solution de $AX = 0$. Ainsi, en particulier, chaque multiple ku d'une solution quelconque u de $AX = 0$ est aussi une solution de $AX = 0$

Preuve : $u_1, u_2, \ldots, u_n$ étant des solutions du système donc $Au_1 = 0, Au_2 = 0, \ldots, Au_n = 0$. Ainsi :

$$\begin{aligned} A(ku_1 + ku_2 + \cdots + ku_n) &= k_1Au_1 + k_2Au_2 + \cdots + k_nAu_n \\ &= k_10 + k_20 + \cdots + k_n0 = 0 \end{aligned}$$

En conséquence $k_1u_1 + \cdots + k_nu_n$ est une solution du système homogène $AX = 0$.

Théorème 3.5 : Supposons le corps K infini (ce qui a lieu si K est le corps réel **R** ou le corps complexe **C**). Le système $AX = B$, soit n'a pas de solution, soit a une solution unique, soit a un nombre infini de solutions.

Preuve : Il suffit de montrer que si $AX = B$ a plus d'une solution alors il en a une infinité. Supposons que u et v soient des solutions distinctes de $AX = B$; donc $Au = B$ et $Av = B$. Alors, quel que soit $k \in K$,

$$A(u + k(u-v)) = Au + k(Au - Av) = B + k(B - B) = B$$

En d'autres termes, quel que soit $k \in K$, $u + k(u - v)$ est une solution de $AX = B$. Puisque toutes ces solutions sont distinctes (Problème 3.31), $AX = B$ a un nombre infini de solutions comme il était prévu.

MATRICES ECHELONNEES

Une matrice $A = (a_{ij})$ est une *matrice échelonnée*, ou est dite mise sous *forme échelonnée* si le nombre de zéros précédant le premier élément non nul d'une ligne augmente de ligne en ligne jusqu'à ce qu'il n'y ait plus de ligne ; c'est-à-dire, s'il existe des éléments non nuls

$$a_{1j_1}, a_{2j_2}, \ldots, a_{rj_r}, \qquad \text{où} \qquad j_1 < j_2 < \cdots < j_r$$

avec la propriété que

$$a_{ij} = 0 \text{ pour } i \leq r,\ j < j_i, \text{ et pour } i > r$$

Nous appelons $a_{1j_1}, \ldots, a_{rj_r}$ les éléments *distingués* ou *remarquables* de la matrice échelonnée.

Exemple 3.9 : Les matrices suivantes sont des matrices échelonnées où les éléments distingués ont été encerclés:

$$\begin{pmatrix} 2 & 3 & 2 & 0 & 4 & 5 & -6 \\ 0 & 0 & 7 & 1 & -3 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 6 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 & 3 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 1 & 0 & -3 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

En particulier une matrice échelonnée, est appelée *matrice échelonnée réduite par les lignes* ou encore une matrice e.r.l. si les éléments distingués sont :

(i) les seuls éléments non nuls dans leurs colonnes respectives;

(ii) chacun égal à 1.

La troisième matrice précédente est un exemple de matrice e.r.l. Remarquons que la matrice 0, pour n'importe quel nombre de lignes ou de colonnes, est aussi une matrice e.r.l.

EQUIVALENCE LIGNE ET OPERATIONS ELEMENTAIRES SUR LES LIGNES

Une matrice A est dite *équivalente ligne* à une matrice B, si B peut être obtenue à partir de A par un nombre fini d'opérations appelées *opérations élémentaires sur les lignes* qui sont données plus bas.

$[E_1]$: On peut interchanger ou permuter la $i$ème ligne et la $j$ème ligne : $R_i \leftrightarrow R_j$.

$[E_2]$: On peut multiplier la $i$ème ligne par un scalaire non nul k: $R_i \to kR_i$, $k \neq 0$.

$[E_3]$: On peut remplacer la $i$ème ligne par k fois la $j$ème ligne plus la $i$ème ligne: $R_i \to kR_j + R_i$,

En pratique on applique $[E_2]$ puis $[E_3]$ dans un premier stade, d'où l'opération

$[E]$: On peut remplacer la $i$ème ligne par k' fois la $j$ème ligne plus k fois avec $k \neq 0$ la $i$ème ligne : $R_i \to k'R_j + kR_i$, $k \neq 0$.

Le lecteur reconnaîtra facilement la similitude des opérations précédentes et de celles utilisées dans la résolution des systèmes d'équations linéaires. En fait, deux systèmes avec des matrices augmentées à lignes équivalentes auront le même ensemble de solutions. L'algorithme suivant est aussi semblable à l'algorithme utilisé à propos des équations linéaires.

Algorithme **réduisant les lignes** *d'une matrice à une forme échelonnée :*

Premièrement : Supposons que la j_1 ème colonne soit la première colonne avec un élément non nul. Echangeons les lignes de telle sorte que ce premier élément non nul soit dans la première ligne, d'où $a_{1j_1} \neq 0$.

Deuxièmement : Pour chaque $i > 1$, appliquons la combinaison

$$R_i \to -a_{ij_1}R_1 + a_{1j_1}R_i$$

On répète le premièrement et le deuxièmement avec la sous-matrice formée par toutes les lignes sauf la première, et on continue cette méthode jusqu'à ce que la matrice soit sous forme échelonnée.

Remarque : La *réduction des lignes* d'une matrice indique la transformation des lignes par des combinaisons sur celles-ci.

Exemple 3.10 : La matrice suivante A est réduite ligne à une forme échelonnée en appliquant les combinaisons $R_2 \to -2R_1 + R_2$ et $R_3 \to -3R_1 + R_3$ et enfin $R_3 \to -5R_2 + 4R_3$:

$$A = \begin{pmatrix} 1 & 2 & -3 & 0 \\ 2 & 4 & -2 & 2 \\ 3 & 6 & -4 & 3 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & 2 & -3 & 0 \\ 0 & 0 & 4 & 2 \\ 0 & 0 & 5 & 3 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & 2 & -3 & 0 \\ 0 & 0 & 4 & 2 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

Supposons que $A = (a_{ij})$ soit une matrice de forme échelonnée ayant pour éléments distingués $a_{1j_1}, a_{2j_2} \dots a_{rj_r}$. Appliquons les combinaisons

$$R_k \to -a_{kj_i}R_i + a_{ij_i}R_k, \qquad k = 1, \dots, i-1$$

pour $i = 2$, $i = 3, \dots, i = r$. Ainsi A est remplacée par une matrice échelonnée dont les éléments distingués sont les seuls éléments non nuls dans leurs colonnes respectives. Multiplions ensuite R_i par $a_{ij_i}^{-1}$ $i, \leqslant r$. Ainsi, finalement, les éléments distingués sont chacun égaux à 1. En d'autres termes on a ainsi réduit par cette méthode la matrice donnée en une matrice e.l.r.

Exemple 3.11 : Sur la matrice échelonnée suivante A appliquons la combinaison $R_1 \to -4R_2 + 3R_1$ ainsi que les combinaisons $R_1 \to R_3 + R_1$ et $R_2 \to -5R_3 + 2R_2$:

$$A = \begin{pmatrix} 2 & 3 & 4 & 5 & 6 \\ 0 & 0 & 3 & 2 & 5 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix} \text{d'où} \begin{pmatrix} 6 & 9 & 0 & 7 & -2 \\ 0 & 0 & 3 & 2 & 5 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix} \text{ d'où } \begin{pmatrix} 6 & 9 & 0 & 7 & 0 \\ 0 & 0 & 6 & 4 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}$$

Multiplions ensuite R_1 par $1/6$, R_2 par $1/6$ et R_3 par $1/2$ pour obtenir la matrice e.l.r :

$$\begin{pmatrix} 1 & 3/2 & 0 & 7/6 & 0 \\ 0 & 0 & 1 & 2/3 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Les précédentes remarques montrent que n'importe quelle matrice A est équivalente ligne à une matrice e.l.r. Dans le prochain chapitre nous démontrerons, théorème 4-8, que A est équivalente ligne à une seule matrice e.l.r ; nous l'appellerons la *forme ligne canonique* de A.

MATRICES CARREES

Une matrice ayant le même nombre de lignes que de colonnes est appelée une *matrice carrée.* Une matrice carrée de n lignes et n colonnes est appelée *d'ordre* n, ou une n-matrice carrée. La *diagonale* (ou diagonale principale) de la matrice carrée d'ordre n. $A = (a_{ij})$ est constituée des éléments $a_{11}, a_{22}, \ldots, a_{nn}$.

Exemple 3.12 : La matrice suivante d'ordre 3: $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$.

Les éléments diagonaux sont 1, 5 et 9.

Une *matrice triangulaire supérieure* ou plus simplement une *matrice triangulaire* est une matrice carrée dont les éléments au-dessous de la diagonale principale sont tous nuls :

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ 0 & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & a_{nn} \end{pmatrix} \quad \text{ou} \quad \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ & a_{22} & \ldots & a_{2n} \\ & & \ldots & \ldots \\ & & & a_{nn} \end{pmatrix}$$

De façon semblable une *matrice triangulaire inférieure* est une matrice carrée dont les éléments au-dessus de la diagonale principale sont tous nuls.

Une *matrice diagonale* est une matrice carrée dont les éléments non diagonaux sont tous nuls :

$$\begin{pmatrix} a_1 & 0 & \ldots & 0 \\ 0 & a_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & a_n \end{pmatrix} \quad \text{ou} \quad \begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{pmatrix}$$

En particulier, les matrices carrées d'ordre n, ne contenant que des 1 sur la diagonale principale et des 0 pour tous les autres éléments, sont notées I_n ou simplement I, et sont appelées matrices *unités* ou *identités*; par exemple

$$I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Cette matrice I est semblable au scalaire 1, en ce sens que pour toute matrice carrée d'ordre n, A, on a

$$AI = IA = A$$

La matrice kI, pour un scalaire $k \in K$, est appelée une *matrice scalaire* ; c'est une matrice diagonale dont les éléments diagonaux sont égaux à k.

ALGEBRE DES MATRICES CARREES

Rappelons que l'on ne peut pas additionner ou multiplier deux matrices quelconques. Cependant si nous considérons seulement des matrices carrées d'ordre donné n, cet inconvénient disparaît. En particulier les opérations d'addition, de multiplication, de multiplication par un scalaire, ainsi que de transposition peuvent être appliquées aux matrices d'ordre n et le résultat est de nouveau une matrice $n \times n$.

En particulier, si A est une matrice carrée quelconque d'ordre n, nous pouvons former les diverses puissances de A :

$$A^2 = AA, \; A^3 = A^2A, \; \ldots \quad \text{et} \quad A^0 = I$$

Nous pouvons aussi former des polynômes de la matrice A ; à chaque polynôme

$$f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$$

où les a_i sont des scalaires, nous pouvons associer $f(A)$ qui est la matrice

$$f(A) = a_0 I + a_1 A + a_2 A^2 + \cdots + a_n A^n$$

Dans le cas où $f(A)$ est la matrice nulle, alors A est appelée ***zéro*** ou ***racine*** du polynôme $f(x)$.

Exemple 3.13 : Soit $A = \begin{pmatrix} 1 & 2 \\ 3 & -4 \end{pmatrix}$; alors $A^2 = \begin{pmatrix} 1 & 2 \\ 3 & -4 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & -4 \end{pmatrix} = \begin{pmatrix} 7 & -6 \\ -9 & 22 \end{pmatrix}$.

Si $f(x) = 2x^2 - 3x + 5$, alors

$$f(A) = 2\begin{pmatrix} 7 & -6 \\ -9 & 22 \end{pmatrix} - 3\begin{pmatrix} 1 & 2 \\ 3 & -4 \end{pmatrix} + 5\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 16 & -18 \\ -27 & 61 \end{pmatrix}$$

Si $g(x) = x^2 + 3x - 10$, alors

$$g(A) = \begin{pmatrix} 7 & -6 \\ -9 & 22 \end{pmatrix} + 3\begin{pmatrix} 1 & 2 \\ 3 & -4 \end{pmatrix} - 10\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

D'où A est un zéro du polynôme $g(x)$.

MATRICES INVERSIBLES

Une matrice carrée A est dite inversible s'il existe une matrice B telle que

$$AB = BA = I$$

où I est la matrice identité. Donc la matrice B est unique ; car

$$AB_1 = B_1A = I \quad \text{et} \quad AB_2 = B_2A = I \text{ implique} \quad B_1 = B_1I = B_1(AB_2) = (B_1A)B_2 = IB_2 = B_2$$

Une telle matrice B est appelée l'inverse de la matrice A et est notée par A^{-1}. Remarquons que la relation précédente est symétrique ; c'est-à-dire que si B est l'inverse de A, alors A est l'inverse de B.

Exemple 3.14 :

$$\begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 6-5 & -10+10 \\ 3-3 & -5+6 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 6-5 & 15-15 \\ -2+2 & -5+6 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Ainsi $\begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix}$ et $\begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix}$ sont inversibles et sont les inverses l'une de l'autre.

Nous montrerons (Problème 3.37) que pour les matrices carrées, $AB = I$ si et seulement si $BA = I$; ainsi il suffira de montrer qu'un seul produit est égal à I pour prouver que deux matrices sont inverses l'une de l'autre, comme dans l'exemple suivant.

Exemple 3.15 :

$$\begin{pmatrix} 1 & 0 & 2 \\ 2 & -1 & 3 \\ 4 & 1 & 8 \end{pmatrix}\begin{pmatrix} -11 & 2 & 2 \\ -4 & 0 & 1 \\ 6 & -1 & -1 \end{pmatrix} = \begin{pmatrix} -11+0+12 & 2+0-2 & 2+0-2 \\ -22+4+18 & 4+0-3 & 4-1-3 \\ -44-4+48 & 8+0-8 & 8+1-8 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Ainsi les deux matrices sont inversibles, et sont les inverses l'une de l'autre.

Calculons maintenant l'inverse d'une matrice générale 2×2, $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Déterminons les scalaires x, y, z, w tels que

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x & y \\ z & w \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{ou} \quad \begin{pmatrix} ax+bz & ay+bw \\ cx+dz & cy+dw \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

ce qui se résume à la résolution des deux systèmes suivants d'équations linéaires à deux inconnues:

$$\begin{cases} ax + bz = 1 \\ cx + dz = 0 \end{cases} \qquad \begin{cases} ay + bw = 0 \\ cy + dw = 1 \end{cases}$$

Si nous posons $|A| = ad - bc$, alors d'après le problème 2.27 page 33, les systèmes précédents ont des solutions si et seulement si $|A| \neq 0$; de telles solutions sont uniques et sont

$$x = \frac{d}{ad - bc} = \frac{d}{|A|}, \quad y = \frac{-b}{ad - bc} = \frac{-b}{|A|}, \quad z = \frac{-c}{ad - bc} = \frac{-c}{|A|}, \quad w = \frac{a}{ad - bc} = \frac{a}{|A|}$$

D'où
$$A^{-1} = \begin{pmatrix} d/|A| & -b/|A| \\ -c/|A| & a/|A| \end{pmatrix} = \frac{1}{|A|}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Remarque : Le lecteur reconnaîtra en $|A| = ad - bc$ le déterminant de la matrice A ; ainsi nous remarquerons qu'une 2×2 matrice a une inverse, si, et seulement si, son déterminant est différent de zéro. Cette propriété, qui reste vraie en général, sera démontrée plus loin dans le chapitre 9 sur les déterminants.

MATRICES DECOMPOSEES EN BLOCS

En utilisant un système de lignes horizontales et verticales, nous pouvons décomposer une matrice A en plusieurs matrices plus petites appelées *blocs* de A. La matrice A est alors appelée une matrice décomposée en blocs. Il est évident qu'une matrice donnée peut être décomposée de différentes manières ; par exemple,

$$\begin{pmatrix} 1 & -2 & 0 & 1 & 3 \\ 2 & 3 & 5 & 7 & -2 \\ 3 & 1 & 4 & 5 & 9 \end{pmatrix} = \left(\begin{array}{cc|cc|c} 1 & -2 & 0 & 1 & 3 \\ 2 & 3 & 5 & 7 & -2 \\ \hline 3 & 1 & 4 & 5 & 9 \end{array}\right) = \left(\begin{array}{ccc|cc} 1 & -2 & 0 & 1 & 3 \\ \hline 2 & 3 & 5 & 7 & -2 \\ \hline 3 & 1 & 4 & 5 & 9 \end{array}\right)$$

L'utilité de la décomposition en blocs est que le résultat des opérations dans les matrices décomposées en blocs peut être obtenu en calculant avec les blocs comme si ceux-ci étaient effectivement des éléments des matrices. Ce qui est illustré par les exemples ci-dessous.

Supposons que A soit décomposée en blocs ; c'est-à-dire

$$A = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \dots & \dots & \dots & \dots \\ A_{m1} & A_{m2} & \dots & A_{mn} \end{pmatrix}$$

En multipliant chaque bloc par un scalaire k, on multiplie chaque élément de A par k ; ainsi

$$kA = \begin{pmatrix} kA_{11} & kA_{12} & \dots & kA_{1n} \\ kA_{21} & kA_{22} & \dots & kA_{2n} \\ \dots & \dots & \dots & \dots \\ kA_{m1} & kA_{m2} & \dots & kA_{mn} \end{pmatrix}$$

Supposons maintenant que la matrice B soit décomposée en un même nombre de blocs que la matrice A ; c'est-à-dire

$$B = \begin{pmatrix} B_{11} & B_{12} & \dots & B_{1n} \\ B_{21} & B_{22} & \dots & B_{2n} \\ \dots & \dots & \dots & \dots \\ B_{m1} & B_{m2} & \dots & B_{mn} \end{pmatrix}$$

De plus nous supposons que les blocs correspondants de A et de B ont la même dimension. En additionnant les blocs correspondants, on additionne les éléments correspondants de A et de B. D'où

$$A + B = \begin{pmatrix} A_{11}+B_{11} & A_{12}+B_{12} & \dots & A_{1n}+B_{1n} \\ A_{21}+B_{21} & A_{22}+B_{22} & \dots & A_{2n}+B_{2n} \\ \dots & \dots & \dots & \dots \\ A_{m1}+B_{m1} & A_{m2}+B_{m2} & \dots & A_{mn}+B_{mn} \end{pmatrix}$$

Le cas de la multiplication est moins évident, mais reste néanmoins vrai. Supposons que A et B soient deux matrices décomposées en blocs comme suit :

$$U = \begin{pmatrix} U_{11} & U_{12} & \dots & U_{1p} \\ U_{21} & U_{22} & \dots & U_{2p} \\ \dots & \dots & \dots & \dots \\ U_{m1} & U_{m2} & \dots & U_{mp} \end{pmatrix} \quad \text{et} \quad V = \begin{pmatrix} V_{11} & V_{12} & \dots & V_{1n} \\ V_{21} & V_{22} & \dots & V_{2n} \\ \dots & \dots & \dots & \dots \\ V_{p1} & V_{22} & \dots & V_{pn} \end{pmatrix}$$

de telle sorte que le nombre de colonnes de chaque bloc U_{ik} est égal au nombre de lignes de chaque bloc V_{kj}. Alors

$$UV = \begin{pmatrix} W_{11} & W_{12} & \dots & W_{1n} \\ W_{21} & W_{22} & \dots & W_{2n} \\ \dots & \dots & \dots & \dots \\ W_{m1} & W_{m2} & \dots & W_{mn} \end{pmatrix}$$

$$W_{ij} = U_{i1}V_{1j} + U_{i2}V_{2j} + \cdots + U_{ip}V_{pj}$$

La démonstration de la formule ci-dessus pour UV se fait directement sans difficulté ; mais elle est longue et fastidieuse. Elle fait l'objet d'un problème supplémentaire (Problème 3.68).

PROBLEMES RESOLUS

ADDITION ET MULTIPLICATION SCALAIRE DE MATRICES

3.1. Calculer

(i) $\begin{pmatrix} 1 & 2 & -3 & 4 \\ 0 & -5 & 1 & -1 \end{pmatrix} + \begin{pmatrix} 3 & -5 & 6 & -1 \\ 2 & 0 & -2 & -3 \end{pmatrix}$

(ii) $\begin{pmatrix} 1 & 2 & -3 \\ 0 & -4 & 1 \end{pmatrix} + \begin{pmatrix} 3 & 5 \\ 1 & -2 \end{pmatrix}$ (iii) $-3\begin{pmatrix} 1 & 2 & -3 \\ 4 & -5 & 6 \end{pmatrix}$

(i) Additionnons les éléments correspondants :

$$\begin{pmatrix} 1 & 2 & -3 & 4 \\ 0 & -5 & 1 & -1 \end{pmatrix} + \begin{pmatrix} 3 & -5 & 6 & -1 \\ 2 & 0 & -2 & -3 \end{pmatrix}$$
$$= \begin{pmatrix} 1+3 & 2-5 & -3+6 & 4-1 \\ 0+2 & -5+0 & 1-2 & -1-3 \end{pmatrix} = \begin{pmatrix} 4 & -3 & 3 & 3 \\ 2 & -5 & -1 & -4 \end{pmatrix}$$

(ii) La somme n'est pas définie puisque les matrices n'ont pas les mêmes dimensions.

(iii) Multiplions chaque élément de la matrice par le scalaire -3 :

$$-3\begin{pmatrix} 1 & 2 & -3 \\ 4 & -5 & 6 \end{pmatrix} = \begin{pmatrix} -3 & -6 & 9 \\ -12 & 15 & -18 \end{pmatrix}$$

3.2. Soient $A = \begin{pmatrix} 2 & -5 & 1 \\ 3 & 0 & -4 \end{pmatrix}$, $B = \begin{pmatrix} 1 & -2 & -3 \\ 0 & -1 & 5 \end{pmatrix}$, $C = \begin{pmatrix} 0 & 1 & -2 \\ 1 & -1 & -1 \end{pmatrix}$. Trouver $3A + 4B - 2C$.

En faisant d'abord la multiplication scalaire puis l'addition:

$$3A + 4B - 2C = \begin{pmatrix} 6 & -15 & 3 \\ 9 & 0 & -12 \end{pmatrix} + \begin{pmatrix} 4 & -8 & -12 \\ 0 & -4 & 20 \end{pmatrix} + \begin{pmatrix} 0 & -2 & 4 \\ -2 & 2 & 2 \end{pmatrix} = \begin{pmatrix} 10 & -25 & -5 \\ 7 & -2 & 10 \end{pmatrix}$$

3.3. Trouver x, y, z et w si $3\begin{pmatrix} x & y \\ z & w \end{pmatrix} = \begin{pmatrix} x & 6 \\ -1 & 2w \end{pmatrix} + \begin{pmatrix} 4 & x+y \\ z+w & 3 \end{pmatrix}$.

Réduisons chaque membre à une seule matrice:

$$\begin{pmatrix} 3x & 3y \\ 3z & 3w \end{pmatrix} = \begin{pmatrix} x+4 & x+y+6 \\ z+w-1 & 2w+3 \end{pmatrix}$$

En égalant les éléments correspondants les uns aux autres,

$$\begin{aligned} 3x &= x+4 \\ 3y &= x+y+6 \\ 3z &= z+w-1 \\ 3w &= 2w+3 \end{aligned} \qquad \text{ou} \qquad \begin{aligned} 2x &= 4 \\ 2y &= 6+x \\ 2z &= w-1 \\ w &= 3 \end{aligned}$$

La solution est donnée par $x = 2$, $y = 4$, $z = 1$, $w = 3$.

3.4. Démontrer le théorème 3.1(v) : Soient A et B deux matrices $m \times n$ et k un scalaire. Alors $k(A+B) = kA + kB$.

Soient $A = (a_{ij})$ et $B = (b_{ij})$. D'où $a_{ij} + b_{ij}$ est le ij-élément de $A + B$, et ainsi $k(a_{ij} + b_{ij})$ est le ij-élément de $k(A+B)$. D'autre part, ka_{ij} et kb_{ij} sont les ij-éléments de kA et kB respectivement et ainsi $ka_{ij} + kb_{ij}$ est le ij-élément de $kA + kB$. Or k, a_{ij} et b_{ij} sont des scalaires appartenant à un corps, d'où

$$k(a_{ij} + b_{ij}) = ka_{ij} + kb_{ij} \quad \text{pour chaque } i, j$$

Ainsi $k(A+B) = kA + kB$, puisque les éléments correspondants sont égaux.

Remarque : Observons que cette démonstration est semblable à la démonstration du Théorème 1.1(v) dans le problème 1.6 page 7. En fait toutes les parties dans le théorème précédent sont démontrées de la même manière que les parties correspondantes du Théorème 1.1.

MULTIPLICATION MATRICIELLE

3.5. On note $(r \times s)$ une matrice de dimensions $r \times s$. Trouver les dimensions des produits suivants si le produit est défini par

(1) $(2\times 3)(3\times 4)$ (3) $(1\times 2)(3\times 1)$ (5) $(3\times 4)(3\times 4)$

(2) $(4\times 1)(1\times 2)$ (4) $(5\times 2)(2\times 3)$ (6) $(2\times 2)(2\times 4)$

Rappelons qu'une $m \times p$ matrice et une $q \times n$ matrice sont multipliables seulement si $p = q$, le produit étant alors une $m \times n$ matrice. Ainsi chacun des produits précédents est défini si le second élément du premier couple est égal au premier élément du second couple, le produit ayant pour dimensions les extrêmes des couples donnés dans l'ordre.

1) Le produit est une 2×4 matrice.

2) Le produit est une 4×2 matrice.

3) Le produit n'est pas défini puisque les moyens 2 et 3 ne sont pas égaux.

4) Le produit est une 5×3 matrice.

5) Le produit n'est pas défini, quoique les matrices aient les mêmes dimensions.

6) Le produit est une 2×4 matrice.

3.6. Soient $A = \begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix}$ et $B = \begin{pmatrix} 2 & 0 & -4 \\ 3 & -2 & 6 \end{pmatrix}$. Trouver 1) AB, 2) BA.

1) Puisque A est une matrice 2 x 2 et B une matrice 2 x 3, le produit AB est défini et est une 2 x 3 matrice. Pour obtenir les éléments de la première ligne de AB, on multiplie la première ligne (1, 3) de A par les colonnes $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$, $\begin{pmatrix} 0 \\ -2 \end{pmatrix}$ et $\begin{pmatrix} -4 \\ 6 \end{pmatrix}$ de B respectivement :

$$\begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix}\begin{pmatrix} 2 & 0 & -4 \\ 3 & -2 & 6 \end{pmatrix} = \begin{pmatrix} 1\cdot 2 + 3\cdot 3 & 1\cdot 0 + 3\cdot(-2) & 1\cdot(-4) + 3\cdot 6 \\ & & \end{pmatrix}$$

$$= \begin{pmatrix} 2+9 & 0-6 & -4+18 \\ & & \end{pmatrix} = \begin{pmatrix} 11 & -6 & 14 \\ & & \end{pmatrix}$$

Pour obtenir les éléments de la deuxième ligne de AB, on multiplie la deuxième ligne (2, – 1) de A par les colonnes de B, respectivement :

$$\begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix}\begin{pmatrix} 2 & 0 & -4 \\ 3 & -2 & 6 \end{pmatrix} = \begin{pmatrix} 11 & -6 & 14 \\ 2\cdot 2 + (-1)\cdot 3 & 2\cdot 0 + (-1)\cdot(-2) & 2\cdot(-4) + (-1)\cdot 6 \end{pmatrix}$$

Ainsi $$AB = \begin{pmatrix} 11 & -6 & 14 \\ 1 & 2 & -14 \end{pmatrix}$$

2) Remarquons que B est une matrice 2 x 3 et A une matrice 2 x 2. Les moyens 3 et 2 n'étant pas égaux, le produit BA n'est pas défini.

3.7. Soient $A = (2, 1)$ et $B = \begin{pmatrix} 1 & -2 & 0 \\ 4 & 5 & -3 \end{pmatrix}$, trouver (1) AB, (2) BA.

1) Puisque A est une matrice 1 x 2, et B une matrice 2 x 3, le produit AB est défini et est une matrice 1 x 3, c'est-à-dire à une ligne et trois colonnes donc un vecteur ligne à trois composantes. Pour obtenir les composantes de AB, on multiplie la ligne de A par chaque colonne de B :

$$AB = (2, 1)\begin{pmatrix} 1 & -2 & 0 \\ 4 & 5 & -3 \end{pmatrix} = (2\cdot 1 + 1\cdot 4,\ 2\cdot(-2) + 1\cdot 5,\ 2\cdot 0 + 1\cdot(-3)) = (6, 1, -3)$$

2) Remarquons que B est une matrice 2 x 3 et A une matrice 1 × 2. Les moyens 3 et 1 n'étant pas égaux, le produit BA n'est pas défini.

3.8. Soient $A = \begin{pmatrix} 2 & -1 \\ 1 & 0 \\ -3 & 4 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & -2 & -5 \\ 3 & 4 & 0 \end{pmatrix}$, trouver (1) AB, (2) BA.

1) Puisque A est une matrice 3 x 2 et B une matrice 2 x 3, le produit AB est défini et est une 3 x 3 matrice. Pour obtenir la première ligne de AB, on multiplie la première ligne de A par chaque colonne de B respectivement :

$$\begin{pmatrix} 2 & -1 \\ 1 & 0 \\ -3 & 4 \end{pmatrix}\begin{pmatrix} 1 & -2 & -5 \\ 3 & 4 & 0 \end{pmatrix} = \begin{pmatrix} 2-3 & -4-4 & -10+0 \\ & & \\ & & \end{pmatrix} = \begin{pmatrix} -1 & -8 & -10 \\ & & \\ & & \end{pmatrix}$$

Pour obtenir la seconde ligne de AB on multiplie la seconde ligne de A par chaque colonne de B respectivement :

$$\begin{pmatrix} 2 & -1 \\ 1 & 0 \\ -3 & 4 \end{pmatrix}\begin{pmatrix} 1 & -2 & -5 \\ 3 & 4 & 0 \end{pmatrix} = \begin{pmatrix} -1 & -8 & -10 \\ 1+0 & -2+0 & -5+0 \\ & & \end{pmatrix} = \begin{pmatrix} -1 & -8 & -10 \\ 1 & -2 & -5 \\ & & \end{pmatrix}$$

Pour obtenir la troisième ligne de AB on multiplie la troisième ligne de A par chaque colonne de B respectivement :

$$\begin{pmatrix} 2 & -1 \\ 1 & 0 \\ -3 & 4 \end{pmatrix}\begin{pmatrix} 1 & -2 & -5 \\ 3 & 4 & 0 \end{pmatrix} = \begin{pmatrix} -1 & -8 & -10 \\ 1 & -2 & -5 \\ -3+12 & 6+16 & 15+0 \end{pmatrix} = \begin{pmatrix} -1 & -8 & -10 \\ 1 & -2 & -5 \\ 9 & 22 & 15 \end{pmatrix}$$

Ainsi
$$AB = \begin{pmatrix} -1 & -8 & -10 \\ 1 & -2 & -5 \\ 9 & 22 & 15 \end{pmatrix}$$

2) Puisque B est une matrice 2 x 3 et A une matrice 3 x 2, le produit BA est défini et est une matrice 2 x 2. Pour obtenir la première ligne de BA on multiplie la première ligne de B par chaque colonne de A respectivement :

$$\begin{pmatrix} 1 & -2 & -5 \\ 3 & 4 & 0 \end{pmatrix}\begin{pmatrix} 2 & -1 \\ 1 & 0 \\ -3 & 4 \end{pmatrix} = \begin{pmatrix} 2-2+15 & -1+0-20 \\ & \end{pmatrix} = \begin{pmatrix} 15 & -21 \\ & \end{pmatrix}$$

Pour obtenir la deuxième ligne de BA on multiplie la deuxième ligne de B par chaque colonne de A respectivement:

$$\begin{pmatrix} 1 & -2 & -5 \\ 3 & 4 & 0 \end{pmatrix}\begin{pmatrix} 2 & -1 \\ 1 & 0 \\ -3 & 4 \end{pmatrix} = \begin{pmatrix} 15 & -21 \\ 6+4+0 & -3+0+0 \end{pmatrix} = \begin{pmatrix} 15 & -21 \\ 10 & -3 \end{pmatrix}$$

Ainsi
$$BA = \begin{pmatrix} 15 & -21 \\ 10 & -3 \end{pmatrix}$$

Remarque : Observons que dans ce cas à la fois AB et BA sont définis, mais ne sont pas égaux ; en fait les matrices résultats n'ont pas les mêmes dimensions.

3.9. Soient $A = \begin{pmatrix} 2 & -1 & 0 \\ 1 & 0 & -3 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & -4 & 0 & 1 \\ 2 & -1 & 3 & -1 \\ 4 & 0 & -2 & 0 \end{pmatrix}$.

1) Déterminer les dimensions de AB. 2) Soit c_{ij} l'élément de la $i^{\text{ème}}$ ligne et de la $j^{\text{ème}}$ colonne du produit AB, c'est-à-dire $AB = (c_{ij})$; trouver c_{23}, c_{14} et c_{21}.

1) Puisque A est une matrice 2 x 3 et B une matrice 3 x 4, le produit AB est une matrice 2 x 4.

2) c_{ij} est défini comme le produit de la $i^{\text{ème}}$ ligne de A par la $j^{\text{ème}}$ colonne de B, d'où

$$c_{23} = (1, 0, -3)\begin{pmatrix} 0 \\ 3 \\ -2 \end{pmatrix} = 1\cdot 0 + 0\cdot 3 + (-3)\cdot(-2) = 0 + 0 + 6 = 6$$

$$c_{14} = (2, -1, 0)\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = 2\cdot 1 + (-1)\cdot(-1) + 0\cdot 0 = 2 + 1 + 0 = 3$$

$$c_{21} = (1, 0, -3)\begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix} = 1\cdot 1 + 0\cdot 2 + (-3)\cdot 4 = 1 + 0 - 12 = -11$$

3.10. Calculer

(1) $\begin{pmatrix} 1 & 6 \\ -3 & 5 \end{pmatrix}\begin{pmatrix} 4 & 0 \\ 2 & -1 \end{pmatrix}$ (3) $\begin{pmatrix} 1 \\ -6 \end{pmatrix}\begin{pmatrix} 1 & 6 \\ -3 & 5 \end{pmatrix}$ (5) $(2, -1)\begin{pmatrix} 1 \\ -6 \end{pmatrix}$

(2) $\begin{pmatrix} 1 & 6 \\ -3 & 5 \end{pmatrix}\begin{pmatrix} 2 \\ -7 \end{pmatrix}$ (4) $\begin{pmatrix} 1 \\ 6 \end{pmatrix}(3, 2)$

1) Le premier facteur est une matrice 2 x 2 et le second une matrice 2 x 2 de sorte que le produit est défini et est une matrice 2 x 2:

$$\begin{pmatrix} 1 & 6 \\ -3 & 5 \end{pmatrix}\begin{pmatrix} 4 & 0 \\ 2 & -1 \end{pmatrix} = \begin{pmatrix} 1\cdot 4 + 6\cdot 2 & 1\cdot 0 + 6\cdot(-1) \\ (-3)\cdot 4 + 5\cdot 2 & (-3)\cdot 0 + 5\cdot(-1) \end{pmatrix} = \begin{pmatrix} 16 & -6 \\ -2 & -5 \end{pmatrix}$$

(2) Le premier facteur est une matrice 2 x 2 et le second une matrice 2 x 1 de sorte que le produit est défini et est une matrice 2 x 1:

$$\begin{pmatrix} 1 & 6 \\ -3 & 5 \end{pmatrix}\begin{pmatrix} 2 \\ -7 \end{pmatrix} = \begin{pmatrix} 1\cdot 2 + 6\cdot(-7) \\ (-3)\cdot 2 + 5\cdot(-7) \end{pmatrix} = \begin{pmatrix} -40 \\ -41 \end{pmatrix}$$

(3) Maintenant le premier facteur est une matrice 2 x 1 et le second une matrice 2 x 2. Les moyens 1 et 2 étant distincts, le produit n'est pas défini.

(4) Ici le premier facteur est une matrice 2 x 1 et le second une matrice 1 x 2, de sorte que le produit est défini et est une matrice 2 x 2 :

$$\begin{pmatrix} 1 \\ 6 \end{pmatrix}(3, 2) = \begin{pmatrix} 1\cdot 3 & 1\cdot 2 \\ 6\cdot 3 & 6\cdot 2 \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ 18 & 12 \end{pmatrix}$$

(5) Le premier facteur est une matrice 1 x 2 et le second une matrice 2 x 1, de sorte que le produit est défini et est une 1 x 1 matrice, que l'on note comme un scalaire.

$$(2, -1)\begin{pmatrix} 1 \\ -6 \end{pmatrix} = (2\cdot 1 + (-1)\cdot(-6)) = (8) = 8$$

3.11. Démontrer le théorème 3.2 (i) : $(AB)C = A(BC)$.

Soient $A = (a_{ij}), B = (b_{jk})$ et $C = (c_{kl})$. De plus soit $AB = S = (s_{ik})$ et $BC = T = (t_{jl})$. Alors

$$s_{ik} = a_{i1}b_{1k} + a_{i2}b_{2k} + \cdots + a_{im}b_{mk} = \sum_{j=1}^{m} a_{ij}b_{jk}$$

$$t_{jl} = b_{j1}c_{1l} + b_{j2}c_{2l} + \cdots + b_{jn}c_{nl} = \sum_{k=1}^{n} b_{jk}c_{kl}$$

Multiplions maintenant S par C c'est-à-dire (AB) par C; l'élément de la $i$ème ligne et de la $l$ième colonne de la matrice $(AB)C$ est

$$s_{i1}c_{1l} + s_{i2}c_{2l} + \cdots + s_{in}c_{nl} = \sum_{k=1}^{n} s_{ik}c_{kl} = \sum_{k=1}^{n}\sum_{j=1}^{m} (a_{ij}b_{jk})c_{kl}$$

D'autre part, multiplions A par T, c'est-à-dire A par BC; l'élément de la $i$ème ligne et de la $l$ème colonne de la matrice $A(BC)$ est

$$a_{i1}t_{1l} + a_{i2}t_{2l} + \cdots + a_{im}t_{ml} = \sum_{j=1}^{m} a_{ij}t_{jl} = \sum_{j=1}^{m}\sum_{k=1}^{n} a_{ij}(b_{jk}c_{kl})$$

Les sommes précédentes étant égales, on a ainsi démontré le théorème.

3.12. Démontrer le théorème 3.2 (ii) : $A(B + C) = AB + AC$.

Soient $A = (a_{ij})$, $B = (b_{jk})$ et $C = (c_{jk})$. De plus appelons $D = B + C = (d_{jk}), E = AB = (e_{ik})$ et $F = AC = (f_{ik})$; alors

$$d_{jk} = b_{jk} + c_{jk}$$

$$e_{ik} = a_{i1}b_{1k} + a_{i2}b_{2k} + \cdots + a_{im}b_{mk} = \sum_{j=1}^{m} a_{ij}b_{jk}$$

$$f_{ik} = a_{i1}c_{1k} + a_{i2}c_{2k} + \cdots + a_{im}c_{mk} = \sum_{j=1}^{m} a_{ij}c_{jk}$$

Ainsi l'élément de la $i$ème ligne et de la $k$ième colonne de la matrice $AB + AC$ est :

$$e_{ik} + f_{ik} = \sum_{j=1}^{m} a_{ij}b_{jk} + \sum_{j=1}^{m} a_{ij}c_{jk} = \sum_{j=1}^{m} a_{ij}(b_{jk} + c_{jk})$$

D'autre part, l'élément de la $i$ème ligne est de la $k$ième colonne de la matrice $AD = A(B + C)$ est

$$a_{i1}d_{1k} + a_{i2}d_{2k} + \cdots + a_{im}d_{mk} = \sum_{j=1}^{m} a_{ij}d_{jk} = \sum_{j=1}^{m} a_{ij}(b_{jk} + c_{jk})$$

D'où $A(B + C) = AB + AC$ puisque les éléments correspondants sont égaux.

TRANSPOSITION

3.13. Trouver la matrice transposée de $A = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 2 & 3 & 4 & 5 \\ 4 & 4 & 4 & 4 \end{pmatrix}$.

En écrivant les lignes de A comme colonnes de A^t : $A^t = \begin{pmatrix} 1 & 2 & 4 \\ 0 & 3 & 4 \\ 1 & 4 & 4 \\ 0 & 5 & 4 \end{pmatrix}$.

3.14. Soit A une matrice quelconque. Sous quelles conditions le produit AA^t est-il défini ?

Supposons que A soit une matrice $m \times n$, alors A^t est une matrice $n \times m$. Ainsi le produit AA^t est toujours défini. Observons que AA^t est aussi défini. Ainsi AA^t est une $m \times m$ matrice, tandis que A^tA est une $n \times n$ matrice.

3.15. Soit $A = \begin{pmatrix} 1 & 2 & 0 \\ 3 & -1 & 4 \end{pmatrix}$. Trouver (i) AA^t, (ii) A^tA.

Pour obtenir A^t on écrit les lignes de A en colonnes : $A^t = \begin{pmatrix} 1 & 3 \\ 2 & -1 \\ 0 & 4 \end{pmatrix}$ Alors

$$AA^t = \begin{pmatrix} 1 & 2 & 0 \\ 3 & -1 & 4 \end{pmatrix}\begin{pmatrix} 1 & 3 \\ 2 & -1 \\ 0 & 4 \end{pmatrix}$$

$$= \begin{pmatrix} 1\cdot1 + 2\cdot2 + 0\cdot0 & 1\cdot3 + 2\cdot(-1) + 0\cdot4 \\ 3\cdot1 + (-1)\cdot2 + 4\cdot0 & 3\cdot3 + (-1)\cdot(-1) + 4\cdot4 \end{pmatrix} = \begin{pmatrix} 5 & 1 \\ 1 & 26 \end{pmatrix}$$

$$A^tA = \begin{pmatrix} 1 & 3 \\ 2 & -1 \\ 0 & 4 \end{pmatrix}\begin{pmatrix} 1 & 2 & 0 \\ 3 & -1 & 4 \end{pmatrix}$$

$$= \begin{pmatrix} 1\cdot1 + 3\cdot3 & 1\cdot2 + 3\cdot(-1) & 1\cdot0 + 3\cdot4 \\ 2\cdot1 + (-1)\cdot3 & 2\cdot2 + (-1)\cdot(-1) & 2\cdot0 + (-1)\cdot4 \\ 0\cdot1 + 4\cdot3 & 0\cdot2 + 4\cdot(-1) & 0\cdot0 + 4\cdot4 \end{pmatrix} = \begin{pmatrix} 10 & -1 & 12 \\ -1 & 5 & -4 \\ 12 & -4 & 16 \end{pmatrix}$$

3.16. Démontrer le théorème 3.3 (iv) : $(AB)^t = B^t A^t$.

Soient $A = (a_{ij})$ et $B = (b_{jk})$. Alors l'élément de la $i$ème ligne et de la $j$ème colonne de la matrice AB est

$$a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{im}b_{mj} \tag{1}$$

Ainsi (*1*) est l'élément de la $j$ème ligne et de la $i$ème colonne de la matrice transposée $(AB)^t$.

D'autre part, la $j$ème ligne de B^t est la $j$ème colonne de B:

$$(b_{1j} \;\; b_{2j} \;\; \ldots \;\; b_{mj}) \tag{2}$$

De plus, la $i$ème colonne de A^t est la $i$ème ligne de A:

$$\begin{pmatrix} a_{i1} \\ a_{i2} \\ \cdot\cdot \\ \cdot\cdot \\ \cdot\cdot \\ a_{im} \end{pmatrix} \tag{3}$$

En conséquence, l'élément qui est à l'intersection de la $j$ème ligne et de la $i$ème colonne de la matrice B^tA^t est le produit de (*2*) et (*3*), ce qui donne (*1*). D'où $(AB)^t = B^tA^t$.

MATRICES ECHELONNEES ET OPERATIONS ELEMENTAIRES SUR LES LIGNES

3.17. Encerclez les éléments distingués dans chacune des matrices échelonnées suivantes. Lesquelles sont réduites par les lignes (e.r.l.) ?

$$\begin{pmatrix}1 & 2 & -3 & 0 & 1\\0 & 0 & 5 & 2 & -4\\0 & 0 & 0 & 7 & 3\end{pmatrix}, \quad \begin{pmatrix}0 & 1 & 7 & -5 & 0\\0 & 0 & 0 & 0 & 1\\0 & 0 & 0 & 0 & 0\end{pmatrix}, \quad \begin{pmatrix}1 & 0 & 5 & 0 & 2\\0 & 1 & 2 & 0 & 4\\0 & 0 & 0 & 1 & 7\end{pmatrix}$$

Les éléments distingués sont les premiers éléments non nuls dans les diverses lignes, d'où

$$\begin{pmatrix}1 & 2 & -3 & 0 & 1\\0 & 0 & 5 & 2 & -4\\0 & 0 & 0 & 7 & 3\end{pmatrix}, \quad \begin{pmatrix}0 & 1 & 7 & -5 & 0\\0 & 0 & 0 & 0 & 1\\0 & 0 & 0 & 0 & 0\end{pmatrix}, \quad \begin{pmatrix}1 & 0 & 5 & 0 & 2\\0 & 1 & 2 & 0 & 4\\0 & 0 & 0 & 1 & 7\end{pmatrix}$$

Une matrice échelonnée est réduite par les lignes si les éléments distingués sont chacun égaux à 1 et sont les seuls éléments non nuls dans leurs colonnes respectives. Ainsi la seconde et la troisième matrice sont réduites par les lignes, mais la première non.

3.18. Soit $A = \begin{pmatrix}1 & -2 & 3 & -1\\2 & -1 & 2 & 2\\3 & 1 & 2 & 3\end{pmatrix}$. (i) Réduire A à la forme échelonnée. (ii) Réduire A à sa forme canonique ligne, c'est-à-dire à la forme e.r.l.

(i) Appliquons les combinaisons suivantes: $R_2 \to -2R_1 + R_2$ et $R_3 \to -3R_1 + R_3$ et enfin $R_3 \to -7R_2 + 3R_3$ pour réduire A à sa forme échelonnée :

$$A \text{ donne } \begin{pmatrix}1 & -2 & 3 & -1\\0 & 3 & -4 & 4\\0 & 7 & -7 & 6\end{pmatrix} \text{ d'où } \begin{pmatrix}1 & -2 & 3 & -1\\0 & 3 & -4 & 4\\0 & 0 & 7 & -10\end{pmatrix}$$

(ii) **Première méthode.** Appliquons les combinaisons $R_1 \to 2R_2 + 3R_1$ et ensuite $R_1 \to -R_3 + 7R_1$ et $R_2 \to 4R_3 + 7R_2$ à la dernière matrice (i) pour réduire A :

$$\text{donc } \begin{pmatrix}3 & 0 & 1 & 5\\0 & 3 & -4 & 4\\0 & 0 & 7 & -10\end{pmatrix} \text{ d'où } \begin{pmatrix}21 & 0 & 0 & 45\\0 & 21 & 0 & -12\\0 & 0 & 7 & -10\end{pmatrix}$$

Finalement multiplions R_1 par 1/21, R_2 par 1/21 et R_3 par 1/7 pour obtenir la forme canonique ligne de A :

$$\begin{pmatrix}1 & 0 & 0 & 15/7\\0 & 1 & 0 & -4/7\\0 & 0 & 1 & -10/7\end{pmatrix}$$

Deuxième méthode. Dans la dernière matrice de (i) multiplions R_2 par 1/3 et R_3 par 1/7 pour obtenir une matrice échelonnée où les éléments distingués sont chacun égaux à 1 :

$$\begin{pmatrix}1 & -2 & 3 & -1\\0 & 1 & -4/3 & 4/3\\0 & 0 & 1 & -10/7\end{pmatrix}$$

Appliquons ensuite la combinaison $R_1 \to 2R_2 + R_1$ puis les combinaisons $R_2 \to (4/3)\,R_3 + R_2$ et $R_1 \to (-\,1/3)\,R_3 + R_1$ pour obtenir la forme canonique ligne de A précédente.

Remarque : Observons qu'un avantage de la première méthode est que les fractions n'apparaissent pas jusqu'au dernier stade.

3.19. Déterminer la forme canonique ligne de $A = \begin{pmatrix} 0 & 1 & 3 & -2 \\ 2 & 1 & -4 & 3 \\ 2 & 3 & 2 & -1 \end{pmatrix}$.

$$A \text{ d'où } \begin{pmatrix} 2 & 1 & -4 & 3 \\ 0 & 1 & 3 & -2 \\ 2 & 3 & 2 & -1 \end{pmatrix} \text{ d'où } \begin{pmatrix} 2 & 1 & -4 & 3 \\ 0 & 1 & 3 & -2 \\ 0 & 2 & 6 & -4 \end{pmatrix} \text{ d'où } \begin{pmatrix} 2 & 1 & -4 & 3 \\ 0 & 1 & 3 & -2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\text{d'où } \begin{pmatrix} 2 & 0 & -7 & 5 \\ 0 & 1 & 3 & -2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & 0 & -7/2 & 5/2 \\ 0 & 1 & 3 & -2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Remarquons que la troisième matrice est déjà sous forme échelonnée.

3.20. Réduire $A = \begin{pmatrix} 6 & 3 & -4 \\ -4 & 1 & -6 \\ 1 & 2 & -5 \end{pmatrix}$ à une forme échelonnée, et enfin à une forme e.r.l. c'est-à-dire à sa forme canonique ligne.

Les calculs sont habituellement simples si l'élément "pivot" est 1. C'est pourquoi on commence par un échange de la première et de la troisième ligne :

$$A \text{ d'où } \begin{pmatrix} 1 & 2 & -5 \\ -4 & 1 & -6 \\ 6 & 3 & -4 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & 2 & -5 \\ 0 & 9 & -26 \\ 0 & -9 & 26 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & 2 & -5 \\ 0 & 9 & -26 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\text{d'où } \begin{pmatrix} 1 & 2 & -5 \\ 0 & 1 & -26/9 \\ 0 & 0 & 0 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & 0 & 7/9 \\ 0 & 1 & -26/9 \\ 0 & 0 & 0 \end{pmatrix}$$

Remarquons que la *troisième matrice* est déjà sous forme échelonnée.

3.21. Montrer que chacune des opérations élémentaires précédentes sur les lignes a une opération inverse du même type :

$[E_1]$: Echange de la $i$ème ligne et de la $j$ème ligne : $R_i \leftrightarrow R_j$.

$[E_2]$: On peut multiplier la $i$ème ligne par un scalaire non nul k: $R_i \to kR_i$, $k \neq 0$.

$[E_3]$: On peut remplacer la $i$ème ligne par k fois la $j$ème ligne plus la $i$ème ligne: $R_i \to kR_j + R_i$.

(i) En échangeant deux fois les deux mêmes lignes, on obtient la matrice initiale ; c'est-à-dire que cette opération est sa propre inverse.

(ii) En multipliant la $i$ème ligne par k puis par k^{-1}, ou par k^{-1} puis par k, nous obtenons la matrice initiale. En d'autres termes, les opérations ou combinaisons $R_i \to kR_i$ et $R_i \to k^{-1}R_i$ sont inverses.

(iii) En appliquant l'opération $R_i \to kR_j + R_i$ puis l'opération $R_i \to -kR_j + R_i$ ou en appliquant l'opération $R_i \to -kR_j + R_i$ puis l'opération $R_i \to kR_j + R_i$, nous obtenons la matrice initiale. En d'autres termes les opérations $R_i \to kR_j + R_i$ et $R_i \to -kR_j + R_i$ sont inverses.

MATRICES CARREES

3.22. Soit $A = \begin{pmatrix} 1 & 2 \\ 4 & -3 \end{pmatrix}$. Trouver (1) A^2, (2) A^3 (3) $f(A)$ où $f(x) = 2x^3 - 4x + 5$.

1) $A^2 = AA = \begin{pmatrix} 1 & 2 \\ 4 & -3 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 4 & -3 \end{pmatrix}$

$$= \begin{pmatrix} 1\cdot 1 + 2\cdot 4 & 1\cdot 2 + 2\cdot(-3) \\ 4\cdot 1 + (-3)\cdot 4 & 4\cdot 2 + (-3)\cdot(-3) \end{pmatrix} = \begin{pmatrix} 9 & -4 \\ -8 & 17 \end{pmatrix}$$

2) $$A^3 = AA^2 = \begin{pmatrix}1 & 2\\ 4 & -3\end{pmatrix}\begin{pmatrix}9 & -4\\ -8 & 17\end{pmatrix}$$
$$= \begin{pmatrix}1\cdot 9 + 2\cdot(-8) & 1\cdot(-4) + 2\cdot 17\\ 4\cdot 9 + (-3)\cdot(-8) & 4\cdot(-4) + (-3)\cdot 17\end{pmatrix} = \begin{pmatrix}-7 & 30\\ 60 & -67\end{pmatrix}$$

3) Pour trouver $f(A)$ il faut d'abord substituer à x, A et à 5, $5I$ pour éliminer le terme constant dans le polynôme donné $f(x) = 2x^3 - 4x + 5$:

$$f(A) = 2A^3 - 4A + 5I = 2\begin{pmatrix}-7 & 30\\ 60 & -67\end{pmatrix} - 4\begin{pmatrix}1 & 2\\ 4 & -3\end{pmatrix} + 5\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$$

En multipliant chaque matrice par son scalaire respectif:

$$= \begin{pmatrix}-14 & 60\\ 120 & -134\end{pmatrix} + \begin{pmatrix}-4 & -8\\ -16 & 12\end{pmatrix} + \begin{pmatrix}5 & 0\\ 0 & 5\end{pmatrix}$$

Finalement, en additionnant des éléments correspondants dans les matrices, on trouve

$$= \begin{pmatrix}-14 - 4 + 5 & 60 - 8 + 0\\ 120 - 16 + 0 & -134 + 12 + 5\end{pmatrix} = \begin{pmatrix}-13 & 52\\ 104 & -117\end{pmatrix}$$

3.23. En se rapportant au problème 3.22 montrer que A est un zéro du polynôme $g(x) = x^2 + 2x - 11$

A est un zéro de $g(x)$ si la matrice $g(A)$ est la matrice nulle. Calculons $g(A)$ comme nous l'avons fait pour $f(A)$, c'est-à-dire remplaçons x par A et 11 par $11I$ pour le terme constant de $g(x) = x^2 + 2x - 11$:

$$g(A) = A^2 + 2A - 11I = \begin{pmatrix}9 & -4\\ -8 & 17\end{pmatrix} + 2\begin{pmatrix}1 & 2\\ 4 & -3\end{pmatrix} - 11\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$$

En mulitpliant chaque matrice par le scalaire qui la précède :

$$g(A) = \begin{pmatrix}9 & -4\\ -8 & 17\end{pmatrix} + \begin{pmatrix}2 & 4\\ 8 & -6\end{pmatrix} + \begin{pmatrix}-11 & 0\\ 0 & -11\end{pmatrix}$$

Enfin en additionnant les éléments correspondants dans les matrices:

$$g(A) = \begin{pmatrix}9 + 2 - 11 & -4 + 4 + 0\\ -8 + 8 + 0 & 17 - 6 - 11\end{pmatrix} = \begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix}$$

Puisque $g(A) = 0$, A est un zéro du polynôme $g(x)$.

3.24. Soit $A = \begin{pmatrix}1 & 3\\ 4 & -3\end{pmatrix}$. Trouver un vecteur colonne non nul $u = \begin{pmatrix}x\\ y\end{pmatrix}$ de telle sorte que $Au = 3u$.

Tout d'abord écrivons l'équation matricielle $Au = 3u$:

$$\begin{pmatrix}1 & 3\\ 4 & -3\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix} = 3\begin{pmatrix}x\\ y\end{pmatrix}$$

Réduisons chaque membre à une matrice unique:

$$\begin{pmatrix}x + 3y\\ 4x - 3y\end{pmatrix} = \begin{pmatrix}3x\\ 3y\end{pmatrix}$$

Egalons les éléments correspondants entre eux pour obtenir un système d'équations (et réduisons à la forme échelonnée) :

$$\begin{aligned} x + 3y &= 3x\\ 4x - 3y &= 3y\end{aligned} \quad \text{ou} \quad \begin{aligned} 2x - 3y &= 0\\ 4x - 6y &= 0\end{aligned} \quad \text{ou} \quad \begin{aligned} 2x - 3y &= 0\\ 0 &= 0\end{aligned} \quad \text{ou} \quad 2x - 3y = 0$$

Ce système se réduit à une équation homogène à deux inconnues et a donc un nombre infini de solutions. Pour obtenir une solution non nulle, fixons y, par exemple $y = 2$ d'où $x = 3$. Donc $x = 3$, $y = 2$ est une solution du système. Ainsi le vecteur $u = \begin{pmatrix}3\\ 2\end{pmatrix}$ est un vecteur colonne non nul et vérifie la propriété $Au = 3u$.

3.25. Trouver l'inverse de la matrice $\begin{pmatrix} 3 & 5 \\ 2 & 3 \end{pmatrix}$.

1$^{\text{ère}}$ **méthode.** Cherchons les scalaires x, y, z et w pour lesquels

$$\begin{pmatrix} 3 & 5 \\ 2 & 3 \end{pmatrix}\begin{pmatrix} x & y \\ z & w \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{ou} \quad \begin{pmatrix} 3x+5z & 3y+5w \\ 2x+3z & 2y+3w \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

où qui satisfont
$$\begin{cases} 3x+5z = 1 \\ 2x+3z = 0 \end{cases} \quad \text{et} \quad \begin{cases} 3y+5w = 0 \\ 2y+3w = 1 \end{cases}$$

La solution du premier système est $x = -3, z = 2$ et celle du second système est $y = 5$, $w = -3$. Ainsi la matrice inverse est donc $\begin{pmatrix} -3 & 5 \\ 2 & -3 \end{pmatrix}$.

2$^{\text{ème}}$ **méthode.** En appliquant la formule générale donnant la matrice inverse A^{-1} de la matrice 2×2 $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$:

$$A^{-1} = \frac{1}{|A|}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \qquad \text{où} \qquad |A| = ad - bc$$

Ainsi si $A = \begin{pmatrix} 3 & 5 \\ 2 & 3 \end{pmatrix}$, alors $|A| = 9 - 10 = -1$ et $A^{-1} = -1\begin{pmatrix} 3 & -5 \\ -2 & 3 \end{pmatrix} = \begin{pmatrix} -3 & 5 \\ 2 & -3 \end{pmatrix}$.

PROBLEMES DIVERS

3.26. Calculer AB en utilisant la multiplication par blocs ; où

$$A = \left(\begin{array}{cc|c} 1 & 2 & 1 \\ 3 & 4 & 0 \\ \hline 0 & 0 & 2 \end{array}\right) \qquad \text{et} \qquad B = \left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 4 & 5 & 6 & 1 \\ \hline 0 & 0 & 0 & 1 \end{array}\right)$$

Ici $A = \begin{pmatrix} E & F \\ 0 & G \end{pmatrix}$ et $B = \begin{pmatrix} R & S \\ 0 & T \end{pmatrix}$ où E, F, G, R, S et T sont les blocs donnés.

D'où

$$AB = \begin{pmatrix} ER & ES+FT \\ 0 & GT \end{pmatrix} = \begin{pmatrix} \begin{pmatrix} 9 & 12 & 15 \\ 19 & 26 & 33 \end{pmatrix} & \begin{pmatrix} 3 \\ 7 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ (0 \quad 0 \quad 0) & (2) \end{pmatrix} = \begin{pmatrix} 9 & 12 & 15 & 4 \\ 19 & 26 & 33 & 7 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

3.27. Supposons que $B = (R_1, R_2, \ldots, R_n)$ c'est-à-dire que R_i est la $i^{\text{ème}}$ ligne de B. On suppose BA défini. Montrer que $BA = (R_1A, R_2A, R_3A, \ldots, R_nA)$ c'est-à-dire que R_iA est la $i^{\text{ème}}$ ligne de BA.

Soient $A^1, A^2, \ldots, A^m$ les colonnes de A. Par définition de la multiplication des matrices, la $i^{\text{ème}}$ ligne de BA est $(R_iA^1, R_iA^2, \ldots, R_iA^m)$. Mais d'après la multiplication matricielle $R_iA = (R_iA^1, R_iA^2, \ldots, R_iA^m)$. Ainsi la $i^{\text{ème}}$ ligne de BA est R_iA.

3.28. Soit $e_i = (0, \ldots, 1, \ldots, 0)$ le vecteur ligne admettant 1 comme $i^{\text{ème}}$ composante, et 0 partout d'ailleurs. Montrer que $e_iA = R_i$ est la $i^{\text{ème}}$ ligne de A.

Remarquons que e_i est la $i^{\text{ème}}$ ligne de I, la matrice identité. D'après le précédent problème, la $i^{\text{ème}}$ ligne de IA est e_iA. Mais $IA = A$. D'où $e_iA = R_i$ est la $i^{\text{ème}}$ ligne de A.

3.29. Montrer que : 1) Si A a une ligne nulle, alors AB a une ligne nulle.
2) Si B a une colonne nulle, alors AB a une colonne nulle.
3) Une matrice quelconque ayant une ligne nulle ou une colonne nulle n'est pas inversible.

1) Soit R_i la ligne nulle de A et $B^1, \ldots, B^n$ les colonnes de B. Alors la $i^{\text{ème}}$ ligne de AB est :

$$(R_i \cdot B^1, R_i \cdot B^2, \ldots, R_i \cdot B^n) = (0, 0, \ldots, 0)$$

2) Soit C_j la colonne nulle de B et $A_1, \ldots, A_m$ les lignes de A. Alors la $j^{\text{ème}}$ colonne de AB est

$$\begin{pmatrix} A_1 \cdot C_j \\ A_2 \cdot C_j \\ \ldots \\ A_m \cdot C_j \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \ldots \\ 0 \end{pmatrix}$$

3) Une matrice A est inversible veut dire qu'il existe une matrice A^{-1} telle que $AA^{-1} = A^{-1}A = I$. Mais la matrice identité I n'a aucune ligne ni aucune colonne nulle ; ainsi d'après 1) et 2) A ne peut avoir une ligne nulle ou une colonne nulle. En d'autres termes, une matrice avec une ligne nulle ou une colonne nulle ne peut être inversible.

3.30. Soient A et B deux matrices inversibles (du même ordre). Montrer que le produit AB est aussi inversible et que $(AB)^{-1} = B^{-1}A^{-1}$. D'où par récurrence $(A_1 A_2 \ldots A_n)^{-1} = A_n^{-1} \ldots A_2^{-1} A_1^{-1}$ où les A_i sont inversibles.

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I$$

et

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}IB = B^{-1}B = I$$

D'où $(AB)^{-1} = B^{-1}A^{-1}$.

3.31. Soient u et v deux vecteurs distincts. Montrer que, pour chaque scalaire $k \in K$, les vecteurs $u + k(u-v)$ sont distincts.

Il suffit de montrer que si

$$u + k_1(u-v) = u + k_2(u-v) \tag{1}$$

alors $k_1 = k_2$. Supposons que (1) soit vraie, d'où

$$k_1(u-v) = k_2(u-v) \quad \text{ou} \quad (k_1-k_2)(u-v) = 0$$

Puisque u et v sont distincts, $u - v \neq 0$. D'où $k_1 - k_2 = 0$ et $k_1 = k_2$.

MATRICES ELEMENTAIRES ET APPLICATIONS*

3.32. Une matrice obtenue à partir de la matrice identique par une seule opération élémentaire sur les lignes est appelée une matrice élémentaire. Déterminer les matrices carrées élémentaires d'ordre 3 correspondant aux opérations $R_1 \leftrightarrow R_2$, $R_3 \to -7R_3$ et $R_2 \to -3R_1 + R_2$.

Appliquons les opérations précédentes à la matrice identique $I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ pour obtenir

$$E_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -7 \end{pmatrix}, \quad E_3 = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

3.33. Démontrer : Soit e une opération élémentaire sur les lignes et E la matrice carrée élémentaire d'ordre m correspondante, c'est-à-dire $E = e(I_m)$. Alors pour n'importe quelle matrice $m \times n$ A, $e(A) = EA$, c'est-à-dire, le résultat $e(A)$ de l'application de l'opération e à la matrice A peut être obtenu en multipliant A par la matrice élémentaire correspondante E.

Soit R_i la $i^{\text{ème}}$ ligne de A ; nous notons ceci en écrivant $A = (R_1, \ldots, R_m)$. D'après le problème 3.27, si B est une matrice pour laquelle le produit AB est défini, alors $AB = (R_1B, R_2B, \ldots, R_mB)$. Posons aussi

$$e_i = (0, \ldots, 0, \hat{1}, 0, \ldots, 0), \qquad \wedge = i$$

----

* Ce paragraphe, qui est assez détaillé, peut être omis en première lecture. Il n'est pas essentiel sauf pour quelques résultats dans le chapitre 9 sur les déterminants.

Ici $\wedge = i$ veut dire que 1 est la $i^{\text{ème}}$ composante. D'après le problème 3.28, $e_i A = R_i$. Nous remarquons aussi que $I = (e_1, \ldots, e_m)$ est la matrice identité.

1) Soit e l'opération élémentaire sur les lignes $R_i \leftrightarrow R_j$. Alors, pour $\wedge = i$ et $\hat{\wedge} = j$,

$$E = e(I) = (e_1, \ldots, \widehat{e_j}, \ldots, \widehat{\widehat{e_i}}, \ldots, e_m)$$

et
$$e(A) = (R_1, \ldots, \widehat{R_j}, \ldots, \widehat{\widehat{R_i}}, \ldots, R_m)$$

Ainsi
$$EA = (e_1A, \ldots, \widehat{e_jA}, \ldots, \widehat{\widehat{e_iA}}, \ldots, e_mA) = (R_1, \ldots, \widehat{R_j}, \ldots, \widehat{\widehat{R_i}}, \ldots, R_m) = e(A)$$

2) Soit e l'opération élémentaire sur les lignes $R_i \to k R_i, k \neq 0$. Alors pour $\wedge = i$,

$$E = e(I) = (e_1, \ldots, \widehat{ke_i}, \ldots, e_m) \quad \text{et} \quad e(A) = (R_1, \ldots, \widehat{kR_i}, \ldots, R_m)$$

Ainsi
$$EA = (e_1A, \ldots, \widehat{ke_iA}, \ldots, e_mA) = (R_1, \ldots, \widehat{kR_i}, \ldots, R_m) = e(A)$$

3) Enfin appelons maintenant e l'opération élémentaire sur les lignes suivantes $R_i \to k R_j + R_i$. D'où, pour $\wedge = i$,

$$E = e(I) = (e_1, \ldots, \widehat{ke_j + e_i}, \ldots, e_m) \quad \text{et} \quad e(A) = (R_1, \ldots, \widehat{kR_j + R_i}, \ldots, R_m)$$

En utilisant $(ke_j + e_i)A = k(e_jA) + e_iA = kR_j + R_i$, nous avons

$$EA = (e_1A, \ldots, \widehat{(ke_j + e_i)A}, \ldots, e_mA) = (R_1, \ldots, \widehat{kR_j + R_i}, \ldots, R_m) = e(A)$$

Nous avons ainsi démontré le théorème.

3.34. Montrer que A est équivalente ligne à B si et seulement si, il existe des matrices élémentaires $E_1, \ldots, E_s$ telles que $E_s \ldots E_2 E_1 A = B$.

D'après la définition, A est équivalente ligne à B s'il existe des opérations élémentaires sur les lignes, $e_1, e_2, \ldots, e_s$ pour lesquelles $e_s(\ldots(e_2(e_1(A)))\ldots) = B$. Mais d'après le précédent problème, les opérations précédentes sont vraies si et seulement si $E_s \ldots E_2 E_1 A = B$, où E_i est la matrice élémentaire correspondant à l'opération e_i.

3.35. Montrer que les matrices élémentaires sont inversibles et que leurs inverses sont aussi des matrices élémentaires.

Soit E la matrice élémentaire correspondant à l'opération élémentaire sur les lignes e: $e(I) = E$. Soit e' l'opération inverse de e (Voir Problème 3.21) et E' sa matrice élémentaire correspondante. D'après le Problème 3.33,

$$I = e'(e(I)) = e'E = E'E \quad \text{et} \quad I = e(e'(I)) = eE' = EE'$$

D'où E' est l'inverse de E.

3.36. Démontrer que les propriétés suivantes sont équivalentes:

1) A est inversible.

2) A est équivalente ligne à la matrice identité I.

3) A est un produit de matrices élémentaires.

Supposons que A soit inversible et supposons que A soit équivalente ligne à une matrice e.r.l. B. Il existe alors des matrices élémentaires $E_1, E_2, \ldots, E_s$ telles que $E_s, \ldots, E_2 E_1 A = B$. Puisque A est inversible et que chaque matrice élémentaire E_i est inversible, le produit est donc inversible. Mais si $B \neq I$, alors B a une ligne nulle (Problème 3.47) ; ainsi B n'est pas inversible (Problème 3.29). Ainsi $B = I$. En d'autres termes 1) implique 2).

Si 2) est vérifié, il existe donc des matrices élémentaires $E_1, E_2, \ldots, E_s$ de telle sorte que

$$E_s \cdots E_2 E_1 A = I, \quad \text{et aussi} \quad A = (E_s \cdots E_2 E_1)^{-1} = E_1^{-1} E_2^{-1} \cdots E_s^{-1}$$

D'après le problème précédent, les E_i^{-1} sont aussi des matrices élémentaires. Donc 2) implique 3).

Si 3) est vérifié ($A = E_1 E_2 \ldots . E_s$) d'où 1) en découle puisque le produit de matrices inversibles est inversible.

3.37. Soient A et B deux matrices carrées du même ordre. Montrer que si $AB = I$, alors $B = A^{-1}$. Ainsi $AB = I$ si et seulement si $BA = I$.

Supposons que A ne soit pas inversible. Alors A n'est pas équivalente ligne à une matrice identité I, et ainsi A est équivalente ligne à une matrice avec une ligne nulle. En d'autres termes, il existe des matrices élémentaires $E_1, \ldots, E_s$ de telle sorte que $E_s \ldots E_2 E_1 A$ a une ligne nulle. D'où $E_s, \ldots, E_2 E_1 AB$ a une ligne nulle. D'où AB est équivalente ligne à une matrice avec une ligne nulle et donc n'est pas équivalente ligne à I. Mais ceci est en contradiction avec le fait que $AB = I$. Ainsi A est inversible. En conséquence

$$B = IB = (A^{-1}A)B = A^{-1}(AB) = A^{-1}I = A^{-1}$$

3.38. Supposons que A soit inversible, et donc peut être réduite ligne à une matrice identité I par la suite des opérations élémentaires $e_1, e_2, \ldots, e_n$. 1) Montrer que cette suite d'opérations élémentaires sur les lignes appliquée à I donne A^{-1}. 2) Utiliser ce résultat pour obtenir l'inverse de la matrice $A = \begin{pmatrix} 1 & 0 & 2 \\ 2 & -1 & 3 \\ 4 & 1 & 8 \end{pmatrix}$.

1) Soit E_i la matrice élémentaire correspondant à l'opération e_i. Alors, d'après l'hypothèse et le problème 3.34, $E_n \ldots E_2 E_1 A = I$. Ainsi $(E_n \ldots E_2 E_1 I) A = I$ et donc, $A^{-1} = E_n, \ldots E_2 E_1 I$. En d'autres termes A^{-1} peut être obtenue à partir de I en appliquant les opérations élémentaires sur les lignes $e_1, \ldots, e_n$.

2) Former la matrice bloc (A, I) et réduire ses lignes de manière à obtenir sa forme canonique ligne :

$$(A, I) = \left(\begin{array}{rrr|rrr} 1 & 0 & 2 & 1 & 0 & 0 \\ 2 & -1 & 3 & 0 & 1 & 0 \\ 4 & 1 & 8 & 0 & 0 & 1 \end{array}\right) \text{ d'où } \left(\begin{array}{rrr|rrr} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & -1 & -1 & -2 & 1 & 0 \\ 0 & 1 & 0 & -4 & 0 & 1 \end{array}\right)$$

$$\text{d'où} \left(\begin{array}{rrr|rrr} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & -1 & -1 & -2 & 1 & 0 \\ 0 & 0 & -1 & -6 & 1 & 1 \end{array}\right) \text{ d'où } \left(\begin{array}{rrr|rrr} 1 & 0 & 0 & -11 & 2 & 2 \\ 0 & -1 & 0 & 4 & 0 & -1 \\ 0 & 0 & -1 & -6 & 1 & 1 \end{array}\right)$$

$$\text{d'où} \left(\begin{array}{rrr|rrr} 1 & 0 & 0 & -11 & 2 & 2 \\ 0 & 1 & 0 & -4 & 0 & 1 \\ 0 & 0 & 1 & 6 & -1 & -1 \end{array}\right)$$

Remarquons que la matrice finale est de la forme (I, B). Ainsi A est inversible et B est son inverse :

$$A^{-1} = \begin{pmatrix} -11 & 2 & 2 \\ -4 & 0 & 1 \\ 6 & -1 & -1 \end{pmatrix}$$

Remarque : Dans ce cas, la matrice bloc finale n'est pas de la forme (I, B), d'où la matrice donnée n'est pas équivalente ligne à I et donc n'est pas inversible.

PROBLEMES SUPPLEMENTAIRES

OPERATIONS SUR LES MATRICES

Dans les problèmes 3.39 – 3.41, soient :

$$A = \begin{pmatrix} 1 & -1 & 2 \\ 0 & 3 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 4 & 0 & -3 \\ -1 & -2 & 3 \end{pmatrix}, \quad C = \begin{pmatrix} 2 & -3 & 0 & 1 \\ 5 & -1 & -4 & 2 \\ -1 & 0 & 0 & 3 \end{pmatrix}, \quad D = \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}$$

3.39. (i) $A + B$, (ii) $A + C$, (iii) $3A - 4B$.

3.40. Trouver (i) AB, (ii) AC, (iii) AD, (iv) BC, (v) BD, (vi) CD.

3.41. Trouver (i) A^t, (ii) A^tC, (iii) D^tA^t, (iv) B^tA, (v) D^tD, (vi) DD^t.

3.42. Soient $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$ et $e_3 = (0, 0, 1)$. Etant donné $A = \begin{pmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \end{pmatrix}$, trouver 1) $e_1 A$, 2) $e_2 A$, 3) $e_3 A$,

3.43. Soient $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$ où 1 est la $i^{\text{ème}}$ composante. Démontrer les propriétés suivantes

1) $Be_j^t = C_j$, la $j^{\text{ème}}$ colonne de B (D'après le problème 3.28, $e_i A = R_i$)

2) Si $e_i A = e_i B$ pour chaque i, alors $A = B$.

3) Si $A e_i^t = Be_i^t$ pour chaque i, alors $A = B$.

MATRICES ECHELONNEES ET OPERATIONS ELEMENTAIRES SUR LES LIGNES

3.44. Réduire A à la forme échelonnée, et ensuite à sa forme canonique ligne, où

(i) $A = \begin{pmatrix} 1 & 2 & -1 & 2 & 1 \\ 2 & 4 & 1 & -2 & 3 \\ 3 & 6 & 2 & -6 & 5 \end{pmatrix}$, (ii) $A = \begin{pmatrix} 2 & 3 & -2 & 5 & 1 \\ 3 & -1 & 2 & 0 & 4 \\ 4 & -5 & 6 & -5 & 7 \end{pmatrix}$.

3.45. Réduire A à la forme échelonnée, et ensuite à sa forme canonique ligne, où

(i) $A = \begin{pmatrix} 1 & 3 & -1 & 2 \\ 0 & 11 & -5 & 3 \\ 2 & -5 & 3 & 1 \\ 4 & 1 & 1 & 5 \end{pmatrix}$, (ii) $A = \begin{pmatrix} 0 & 1 & 3 & -2 \\ 0 & 4 & -1 & 3 \\ 0 & 0 & 2 & 1 \\ 0 & 5 & -3 & 4 \end{pmatrix}$.

3.46. Donner toutes les matrices 2×2 possibles qui sont réduites ligne à leur forme échelonnée réduite par les lignes.

3.47. Soit A une matrice carrée réduite à sa forme e.r.l. Montrer que si $A \neq I$, la matrice identité, alors A a une ligne nulle.

3.48. Montrer que chaque matrice carrée échelonnée est triangulaire supérieure, mais non vice versa.

3.49. Montrer que l'équivalence-ligne est une relation d'équivalence :

1) A est équivalente-ligne à A.

2) A est équivalente-ligne à B implique B équivalente-ligne à A.

3) A équivalente-ligne à B et B équivalente-ligne à C implique A équivalente-ligne à C.

MATRICES CARREES

3.50. Soit $A = \begin{pmatrix} 2 & 2 \\ 3 & -1 \end{pmatrix}$. 1) Trouver A^2 et A^3; 2) si $f(x) = x^3 - 3x^2 - 2x + 4$, trouver $f(A)$. 3) si $g(x) = x^2 - x - 8$, trouver $g(A)$.

3.51. Soit $B = \begin{pmatrix} 1 & 3 \\ 5 & 3 \end{pmatrix}$. 1) Si $f(x) = 2x^2 - 4x + 3$, trouver $f(B)$; 2) si $g(x) = x^2 - 4x - 12$, trouver $g(B)$; 3) Trouver un vecteur colonne non nul $u = \begin{pmatrix} x \\ y \end{pmatrix}$ de telle sorte que $Bu = 6u$.

3.52. Les matrices A et B sont dites commutables si $AB = BA$. Trouver toutes les matrices $\begin{pmatrix} x & y \\ z & w \end{pmatrix}$ qui commutent avec $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

3.53. Soit $A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$. Trouver A^n.

3.54. Soit $A = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$ et $B = \begin{pmatrix} 7 & 0 \\ 0 & 11 \end{pmatrix}$.

Trouver : (i) $A + B$, (ii) $A \cdot B$, (iii) A^2 et A^3, (iv) A^n, (v) $f(A)$ pour le polynôme $f(x)$.

3.55. Soient $D = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}$, $A = \begin{pmatrix} a_1 & a_2 & \dots & a_n \\ b_1 & b_2 & \dots & b_n \end{pmatrix}$, $B = \begin{pmatrix} c_1 & d_1 \\ c_2 & d_2 \\ \dots & \dots \\ c_n & d_n \end{pmatrix}$. Trouver DA et BD.

3.56. On suppose que la matrice carrée 2×2 B commute avec toute matrice carrée d'ordre 2, A, c'est-à-dire $AB = BA$. Montrer que $B = \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}$ pour un scalaire k, c'est-à-dire que B est une matrice scalaire.

3.57. Soit D_k la matrice scalaire carrée d'ordre m, dont les éléments diagonaux sont k. Montrer que 1) pour toute matrice $m \times n$ A, $D_k A = kA_k$ 2) pour une matrice $n \times m$ B, $BD_k = kB$.

3.58. Montrer que la somme, le produit et le multiple scalaire des:

1) matrices triangulaires supérieures est triangulaire supérieure;
2) matrices triangulaires inférieures est triangulaire inférieure;
3) matrices diagonales est diagonale;
4) matrices scalaires est scalaire.

MATRICES INVERSIBLES

3.59. Trouver l'inverse de chacune des matrices : 1) $\begin{pmatrix} 3 & 2 \\ 7 & 5 \end{pmatrix}$, 2) $\begin{pmatrix} 2 & -3 \\ 1 & 3 \end{pmatrix}$.

3.60. Trouver l'inverse de chacune des matrices: (i) $\begin{pmatrix} -1 & 2 & -3 \\ 2 & 1 & 0 \\ 4 & -2 & 5 \end{pmatrix}$, (ii) $\begin{pmatrix} 2 & 1 & -1 \\ 0 & 2 & 1 \\ 5 & 2 & -3 \end{pmatrix}$.

3.61. Trouver l'inverse de $\begin{pmatrix} 1 & 3 & 4 \\ 3 & -1 & 6 \\ -1 & 5 & 1 \end{pmatrix}$.

3.62. Montrer que les opérations inverse et transposition commutent, c'est-à-dire $(A^t)^{-1} = (A^{-1})^t$. Ainsi, en particulier, A est inversible si et seulement si A^t est inversible.

3.63. Quand une matrice diagonale $A = \begin{pmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_n \end{pmatrix}$ est-elle inversible et quelle est son inverse ?

3.64. Montrer que A est équivalente ligne à B si, et seulement si, il existe une matrice inversible P telle que $B = PA$.

3.65. Montrer que A est inversible si, et seulement si, le système $AX = 0$ admet uniquement la solution nulle.

PROBLEMES DIVERS

3.66. Démontrer le théorème 3.2 : 3) $(B + C)A = BA + CA$; 4) $k(AB) = (kA)B = A(kB)$ où k est un scalaire. Les parties 1) et 2) sont démontrées dans le problème 3.11 et 3.12.

3.67. Démontrer le théorème 3.3 : (i) $(A + B)^t = A^t + B^t$; (ii) $(A^t)^t = A$; (iii) $(kA)^t = kA^t$ pour k scalaire. La partie (iv) est démontrée dans le problème 3.16.

3.68. Supposons que $A = (A_{ik})$ et $B = (B_{kj})$ sont des matrices blocs pour lesquelles AB est défini et le nombre de colonnes de chaque bloc A_{ik} est égal au nombre de lignes de chaque bloc B_{kj}. Montrer que $AB = (C_{ij})$ où $C_{ij} = \sum_k A_{ik} B_{kj}$.

3.69. Les opérations suivantes sont appelées opérations élémentaires sur les colonnes.

$[E_1]$: On peut échanger la $i$ème colonne et la $j$ème colonne.

$[E_2]$: On peut multiplier la $i$ème colonne par un scalaire non nul k.

$[E_3]$: On peut remplacer la $i$ème colonne par k fois la $j$ème colonne plus la $l$ème colonne.

Montrer que chacune de ces opérations a une opération inverse du même type.

3.70. Une matrice A est dite équivalente à une matrice B si B peut être obtenue à partir de A par une suite finie d'opérations ou de combinaisons, chacune étant une opération élémentaire sur les lignes ou les colonnes. Montrer que l'équivalence de matrices est une relation d'équivalence.

3.71. Montrer que deux systèmes possibles d'équations linéaires ont le même ensemble de solutions si, et seulement si, leurs matrices augmentées sont équivalentes lignes. (Nous supposons que des lignes nulles sont ajoutées de telle manière que les deux matrices augmentées aient le même nombre de lignes).

REPONSES AUX PROBLEMES SUPPLEMENTAIRES

3.39. (i) $\begin{pmatrix} 5 & -1 & -1 \\ -1 & 1 & 7 \end{pmatrix}$ (ii) non défini (iii) $\begin{pmatrix} -13 & -3 & 18 \\ 4 & 17 & 0 \end{pmatrix}$

3.40. (i) Non défini (iii) $\begin{pmatrix} 9 \\ 9 \end{pmatrix}$ (v) $\begin{pmatrix} -1 \\ 9 \end{pmatrix}$

(ii) $\begin{pmatrix} -5 & -2 & 4 & 5 \\ 11 & -3 & -12 & 18 \end{pmatrix}$ (iv) $\begin{pmatrix} 11 & -12 & 0 & -5 \\ -15 & 5 & 8 & 4 \end{pmatrix}$ (vi) Non défini

3.41. (i) $\begin{pmatrix} 1 & 0 \\ -1 & 3 \\ 2 & 4 \end{pmatrix}$ (ii) Non défini (iii) $(9, 9)$ (iv) $\begin{pmatrix} 4 & -7 & 4 \\ 0 & -6 & -8 \\ -3 & 12 & 6 \end{pmatrix}$ (v) 14 (vi) $\begin{pmatrix} 4 & -2 & 6 \\ -2 & 1 & -3 \\ 6 & -3 & 9 \end{pmatrix}$

3.42. (1) (a_1, a_2, a_3, a_4) (2) (b_1, b_2, b_3, b_4) (3) (c_1, c_2, c_3, c_4)

3.44. (i) $\begin{pmatrix} 1 & 2 & -1 & 2 & 1 \\ 0 & 0 & 3 & -6 & 1 \\ 0 & 0 & 0 & -6 & 1 \end{pmatrix}$ et $\begin{pmatrix} 1 & 2 & 0 & 0 & 4/3 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1/6 \end{pmatrix}$

(ii) $\begin{pmatrix} 2 & 3 & -2 & 5 & 1 \\ 0 & -11 & 10 & -15 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$ et $\begin{pmatrix} 1 & 0 & 4/11 & 5/11 & 13/11 \\ 0 & 1 & -10/11 & 15/11 & -5/11 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$

3.45. (i) $\begin{pmatrix} 1 & 3 & -1 & 2 \\ 0 & 11 & -5 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ et $\begin{pmatrix} 1 & 0 & 4/11 & 13/11 \\ 0 & 1 & -5/11 & 3/11 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$

(ii) $\begin{pmatrix} 0 & 1 & 3 & -2 \\ 0 & 0 & -13 & 11 \\ 0 & 0 & 0 & 35 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ et $\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$

3.46. $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ or $\begin{pmatrix} 1 & k \\ 0 & 0 \end{pmatrix}$ où k est un scalaire quelconque.

3.48. $\begin{pmatrix} 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ est une matrice supérieure triangulaire mais non une matrice échelonnée.

3.50. (1) $A^2 = \begin{pmatrix} 10 & 2 \\ 3 & 7 \end{pmatrix}$, $A^3 = \begin{pmatrix} 26 & 18 \\ 27 & -1 \end{pmatrix}$; (2) $f(A) = \begin{pmatrix} -4 & 8 \\ 12 & -16 \end{pmatrix}$; (3), $g(A) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

3.51. (1) $f(B) = \begin{pmatrix} 31 & 12 \\ 20 & 39 \end{pmatrix}$; (2) $g(B) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$; (3) $u = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$ ou $\begin{pmatrix} 3k \\ 5k \end{pmatrix}$, $k \neq 0$.

3.52. Seulement les matrices de la forme $\begin{pmatrix} a & b \\ 0 & a \end{pmatrix}$ commutent avec $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

3.53. $A^n = \begin{pmatrix} 1 & 2n \\ 0 & 1 \end{pmatrix}$

3.54. (i) $A + B = \begin{pmatrix} 9 & 0 \\ 0 & 14 \end{pmatrix}$ (iii) $A^2 = \begin{pmatrix} 4 & 0 \\ 0 & 9 \end{pmatrix}$, $A^3 = \begin{pmatrix} 8 & 0 \\ 0 & 27 \end{pmatrix}$ (v) $f(A) = \begin{pmatrix} f(2) & 0 \\ 0 & f(3) \end{pmatrix}$

(ii) $AB = \begin{pmatrix} 14 & 0 \\ 0 & 33 \end{pmatrix}$ (iv) $A^n = \begin{pmatrix} 2^n & 0 \\ 0 & 3^n \end{pmatrix}$

3.55. (1) $DA = \begin{pmatrix} 3a_1 & 3a_2 & \dots & 3a_n \\ 3b_1 & 3b_2 & \dots & 3b_n \end{pmatrix} = 3A$ (2) $BD = \begin{pmatrix} 3c_1 & 3d_1 \\ 3c_2 & 3d_2 \\ \dots & \dots \\ 3c_n & 3d_n \end{pmatrix} = 3B$

3.59. (1) $\begin{pmatrix} 5 & -2 \\ -7 & 3 \end{pmatrix}$ (2) $\begin{pmatrix} 1/3 & 1/3 \\ -1/9 & 2/9 \end{pmatrix}$

3.60. (i) $\begin{pmatrix} -5 & 4 & -3 \\ 10 & -7 & 6 \\ 8 & -6 & 5 \end{pmatrix}$ (ii) $\begin{pmatrix} 8 & -1 & -3 \\ -5 & 1 & 2 \\ 10 & -1 & -4 \end{pmatrix}$

3.61. $\begin{pmatrix} 31/2 & -17/2 & -11 \\ 9/2 & -5/2 & -3 \\ -7 & 4 & 5 \end{pmatrix}$

3.62. Etant donné $AA^{-1} = I$, alors $I = I^t = (AA^{-1})^t = (A^{-1})^t A^t$. D'où $(A^{-1})^t = (A^t)^{-1}$.

3.63. A est inversible si et seulement si chaque $a_i \neq 0$. D'où $A^{-1} = \begin{pmatrix} a_1^{-1} & 0 & \dots & 0 \\ 0 & a_2^{-1} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_n^{-1} \end{pmatrix}$.

CHAPITRE 4

Espaces vectoriels et sous-espaces vectoriels

INTRODUCTION

Dans le premier chapitre, nous avons étudié les structures de $\mathbf{R}^n$ et $\mathbf{C}^n$ et nous en avons déduit différentes propriétés. Maintenant certaines de ces propriétés joueront le rôle d'axiomes dans la définition abstraite des "espaces vectoriels". En particulier les conclusions de (i) à (vii) du théorème 1-1, page 3, deviendront les axiomes $[A_1] - [A_4]$, $[M_1] - [M_4]$ dans la définition suivante. On peut donc remarquer que, dans un certain sens, on n'apporte dans cette question rien de nouveau. En fait, nous démontrons dans le chapitre 5 que tout espace vectoriel sur $\mathbf{R}$ de "dimension finie" peut être identifié avec $\mathbf{R}^n$ pour n donné.

La définition d'espace vectoriel suppose un corps arbitraire (voir Appendice B) dont les éléments sont appelés *scalaires*. Nous adopterons la notation suivante (à moins qu'il n'en soit spécifié autrement) :

K	corps des scalaires
a, b, c ou k	les éléments de K
V	l'espace vectoriel donné
u, v, w	les éléments de V

Remarquons que rien d'essentiel n'est oublié si le lecteur suppose que K est le corps des réels $\mathbf{R}$ ou le corps des complexes $\mathbf{C}$.

Nous indiquerons plus tard que "le produit scalaire" et des notions relatives à ce produit scalaire comme l'orthogonalité, ne sont pas parties intégrantes de la structure des espaces vectoriels, mais sont considérés comme une structure supplémentaire qui peut ou non être introduite. De tels espaces seront considérés dans la dernière partie de ce texte.

Définition : Soit K un corps donné, et soit V un ensemble non vide ayant les deux lois, addition et multiplication par un scalaire, qui font correspondre à $u, v \in V$ une somme $u + v \in V$ et à un $u \in V$ quelconque et $k \in K$ un produit $ku \in V$. V est alors appelé espace vectoriel sur K (et les éléments de V sont appelés vecteurs) si les axiomes suivants sont vérifiés :

$[A_1]$: Quels que soient les vecteurs $u, v, w \in V$, $(u + v) + w = u + (v + w)$.

$[A_2]$: Il existe un vecteur de V, noté 0 et appelé vecteur nul tel que $u + 0 = 0 + u = u$ quel que soit $u \in V$.

$[A_3]$: Quel que soit le vecteur $u \in V$, il existe un vecteur de V, noté $-u$, pour lequel $u + (-u) = 0$.

$[A_4]$: Quels que soient les vecteurs $v \in V$, $u + v = v + u$.

$[M_1]$: Quel que soit le scalaire $k \in K$ et quels que soient les vecteurs $u, v \in V$, $k(u + v) = ku + kv$.

$[M_2]$: Quels que soient les scalaires $a, b \in K$ et quel que soit le vecteur $u \in V$, $(a + b)u = au + bu$.

$[M_3]$: Quels que soient les scalaires $a, b \in K$ et quel que soit le vecteur $u \in V$, $(ab)u = a(bu)$.

$[M_4]$: Pour le scalaire $1 \in K$, $1u = u$ quel que soit le vecteur $u \in V$.

Les axiomes précédents se scindent naturellement en deux parties. Les quatre premiers concernent la structure additive de V et peuvent être résumés en disant que V est un groupe commutatif (voir Appendice B) par rapport à l'addition. Il s'ensuit que pour une somme quelconque de vecteurs de la forme

$$v_1 + v_2 + \cdots + v_m$$

il est inutile de mettre des parenthèses et que cette somme ne dépend pas de l'ordre de ses termes, le vecteur nul 0 est unique, et l'opposé $-u$ de u est unique, d'où la loi de régularité :

$$u + w = v + w \quad \text{implique} \quad u = v$$

quels que soient les vecteurs $u, v, w \in V$. De même, la soustraction est définie par

$$u - v = u + (-v)$$

D'autre part, les quatre axiomes restants apparaissent, comme l'indique leurs textes, définissant "l'action" du corps K sur V. En utilisant les axiomes de l'addition, on démontre les propriétés simples suivantes de l'espace vectoriel.

Théorème 4.1 : Soit V un espace vectoriel défini sur un corps K.

(i) Quel que soit le scalaire $k \in K$ et $0 \in V$, $k0 = 0$.

(ii) Pour $0 \in K$ et un vecteur quelconque $u \in V$, $0u = 0$.

(iii) Si $ku = 0$ avec $k \in K$ et $u \in K$, alors $k = 0$ ou $u = 0$.

(iv) Quel que soit le scalaire $k \in K$ et quel que soit $u \in V$, $(-k)u = k(-u) = -ku$

EXEMPLES D'ESPACES VECTORIELS

Dressons maintenant une liste des plus importants espaces vectoriels. Le premier exemple est une généralisation de l'espace $\mathbf{R}^n$.

Exemple 4.1 : Soit K un corps arbitraire. L'ensemble de tous les n-tuples dont les éléments appartiennent à K avec l'addition vectorielle et la multiplication scalaire définies par

$$(a_1, a_2, \ldots, a_n) + (b_1, b_2, \ldots, b_n) = (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n)$$

et

$$k(a_1, a_2, \ldots, a_n) = (ka_1, ka_2, \ldots, ka_n)$$

où $a_i, b_i, k \in K$, est un espace vectoriel sur K ; on note cet espace vectoriel K^n. Le vecteur nul de K^n étant le n-tuple dont tous les éléments sont nuls, $0 = (0, 0, \ldots, 0)$. Pour démontrer que K^n est un espace vectoriel il suffit d'appliquer le théorème 1.1. Nous pouvons maintenant dire que $\mathbf{R}^n$ muni des lois précédentes est un espace vectoriel sur $\mathbf{R}$.

Exemple 4.2 : Soit V l'ensemble des matrices $m \times n$ dont les éléments appartiennent à un corps arbitraire K. V muni des deux lois addition des matrices et multiplication d'une matrice par un scalaire est un espace vectoriel, d'après le théorème 3.1.

Exemple 4.3 : Soit V l'ensemble des polynômes $a_0 + a_1 t + a_2 t^2 + \ldots + a_n t^n$ dont les coefficients appartiennent à un corps K. V muni des deux lois addition des polynômes et multiplication par un scalaire est un espace vectoriel sur K.

Exemple 4.4 : Soit K un corps arbitraire et soit X un ensemble non vide. Considérons l'ensemble de toutes les fonctions de X dans K. La somme de deux fonctions quelconques $f, g \in V$ est la fonction $f + g \in V$ définie par

$$(f + g)(x) = f(x) + g(x)$$

et le produit d'une fonction $f \in V$ par un scalaire $k \in K$ est la fonction $kf \in V$ définie par

$$(kf)(x) = kf(x)$$

V muni des lois précédentes est un espace vectoriel sur K (Problème 4-5). Le vecteur nul de V est la fonction nulle qui à chaque $x \in X$ fait correspondre $0 \in K$. $\mathbf{0}(x) = 0$ quel que soit $x \in X$. De plus, pour une fonction quelconque $f \in V$, $-f$ est la fonction telle que $(-f)(x) = -f(x)$ pour tout $x \in X$.

Exemple 4.5 : Soit E un corps contenant un sous-corps K. E peut être considéré comme un espace vectoriel sur K, pour cela l'addition habituelle dans E jouera le rôle de l'addition vectorielle et nous définirons le produit par un scalaire kv avec $k \in K$ et $v \in E$. Ainsi le corps complexe $\mathbf{C}$ est un espace vectoriel sur le corps réel $\mathbf{R}$ et le corps réel $\mathbf{R}$ est un espace vectoriel sur le corps des rationnels $\mathbf{Q}$.

SOUS-ESPACES VECTORIELS

Soit W un sous-ensemble d'un espace vectoriel sur un corps K. W est appelé sous-espace de V si W est lui-même un espace vectoriel sur K par rapport aux lois d'addition vectorielle et de multiplication par un scalaire définies sur V. Il en découle les critères d'identification des sous-espaces vectoriels suivants :

Théorème 4.2 : W est un sous-espace vectoriel de V si et seulement si

(i) W n'est pas vide,

(ii) W est fermé pour l'addition ; $v, w \in W$ implique $v + w \in W$,

(iii) W est fermé pour la multiplication par un scalaire : $v \in W$ implique $kv \in W$ quel que soit $k \in K$.

Corollaire 4.3 : W est un sous-espace vectoriel de V si et seulement si (i) $0 \in W$ (ou $W \neq \emptyset$) et (ii) $v, w \in W$ implique $av + bw \in W$ quel que soit $a, b \in K$.

Exemple 4.6 : Soit V un espace vectoriel quelconque. L'ensemble $\{0\}$ constitué du seul vecteur nul, ainsi que l'espace entier V sont des sous-espaces vectoriels de V.

Exemple 4.7 : (i) Soit V l'espace vectoriel $\mathbf{R}^3$. L'ensemble des vecteurs W dont la dernière composante est nulle, $W = \{(a, b, 0) ; a, b \in \mathbf{R}\}$, est un sous-espace de V.

(ii) Soit V l'espace vectoriel des matrices $m \times n$ (voir exemple 4.2). L'ensemble W des matrices $A = (a_{ij})$ pour lesquelles $a_{ij} = a_{ji}$ appelées matrices symétriques est un sous-espace de V.

(iii) Soit V l'espace vectoriel des polynômes (voir exemple 4.3). L'ensemble W des polynômes de degré $\leqslant n$, n étant fixé, est un sous-espace vectoriel de V.

(iv) Soit V l'espace vectoriel de toutes les fonctions d'un ensemble non vide X vers le corps des réels $\mathbf{R}$. L'ensemble W constitué de toutes les fonctions bornées dans V est un sous-espace vectoriel de V. (Une fonction $f \in W$ est bornée s'il existe $M \in \mathbf{R}$ tel que $|f(x)| \leqslant M$ pour tout $x \in X$).

Exemple 4.8 : Considérons un système homogène quelconque d'équations linéaires à n inconnues et à coefficients réels :

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0\\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0\\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots &\\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0
\end{aligned}$$

Rappelons qu'une solution particulière quelconque de ce système peut être identifiée à un point de $\mathbf{R}^n$. L'ensemble W de toutes les solutions de ce système homogène est un sous-espace vectoriel de $\mathbf{R}^n$ (Problème 4.16) appelé espace vectoriel des solutions. Cependant l'ensemble des solutions d'un système non homogène d'équations linéaires à n inconnues n'est pas un sous-espace vectoriel de $\mathbf{R}^n$.

Exemple 4.9 : Soient U et W deux sous-espaces vectoriels d'un espace vectoriel V. Montrons que leur intersection $U \cap W$ est aussi un sous-espace de V. On a $0 \in U$ et $0 \in W$ puisque U et W sont des sous-espaces vectoriels; d'où $0 \in U \cap W$. Supposons maintenant $u, v \in U \cap W$ d'où $u, v \in U$ et $u, v \in W$ et puisque U et W sont des sous-espaces vectoriels,

$$au + bv \in U \qquad \text{et} \qquad au + bv \in W$$

quels que soient les scalaires $a, b \in K$. Donc $au + bv \in U \cap W$ d'où $U \cap W$ est un sous-espace vectoriel de V.

Le résultat de l'exemple précédent peut être généralisé dans le théorème suivant :

Théorème 4.4 : L'intersection d'un nombre quelconque de sous-espaces vectoriels d'un espace vectoriel de V est un sous-espace vectoriel de V.

COMBINAISONS LINEAIRES. GENERATEURS

Soit V un espace vectoriel sur le corps R et soient $v_1, \ldots, v_m \in V$. Un vecteur quelconque de V, de la forme

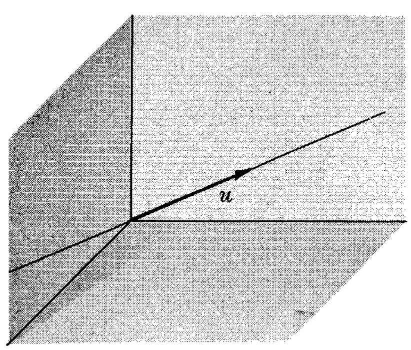
$$a_1v_1 + a_2v_2 + \cdots + a_mv_m$$

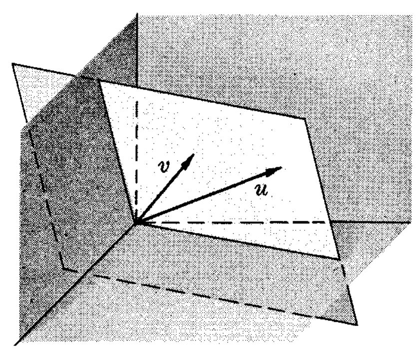
où les $a_i \in K$, est appelé une combinaison linéaire de $v_1, \ldots, v_m$. On applique le théorème suivant :

Théorème 4.5 : Soit S un sous-ensemble non vide de V. L'ensemble de toutes les combinaisons linéaires des vecteurs de S, noté par $L(S)$ est un sous-espace vectoriel de V contenant S. De plus, si W est un autre sous-espace vectoriel de V contenant S, alors $L(S) \subset W$.

En d'autres termes, $L(S)$ est le plus petit sous-espace de V contenant S ; il est appelé sous-espace engendré par S. Par convention on pose $L(\emptyset) = \{0\}$.

Exemple 4.10 : Soit V l'espace vectoriel $\mathbf{R}^3$. Un vecteur quelconque non nul u engendre l'ensemble des multiples scalaires de u ; géométriquement l'ensemble des multiples de u est la droite passant par l'origine et le point u. L'espace linéaire engendré par deux vecteurs u et v qui ne sont pas multiples l'un de l'autre est le plan passant par l'origine et contenant les points u et v.





Exemple 4.11 : Les vecteurs $e_1 = (1, 0, 0), e_2 = (0, 1, 0)$ et $e_3 = (0, 0, 1)$ engendrent l'espace vectoriel $\mathbf{R}^3$. Quel que soit un vecteur $(a, b, c) \in \mathbf{R}^3$, on peut l'écrire sous forme de combinaison linéaire des e_i ; en particulier

$$\begin{aligned}(a, b, c) &= a(1, 0, 0) + b(0, 1, 0) + c(0, 0, 1)\\ &= ae_1 + be_2 + ce_3\end{aligned}$$

Exemple 4.12 : Les polynômes $1, t, t^2, t^3, \ldots$ engendrent l'espace vectoriel V de tous les polynômes (en t) $V = L(1, t, t^2, \ldots)$. Chaque polynôme est une combinaison linéaire de 1 et de puissances de t.

Exemple 4.13 : Déterminer si le vecteur $v(3, 9, -4, -2)$ est une combinaison linéaire ou non des vecteurs $u_1 = (1, -2, 0, 3)$, $u_2 = (2, 3, 0, -1)$ et $u_3 = (2, -1, 2, 1)$, c'est-à-dire appartient à l'espace engendré par les u_i.

Soit v une combinaison linéaire des u_i, on utilisera les inconnues x, y et z d'où $v = xu_1 + yu_2 + zu_3$:

$$\begin{aligned}(3, 9, -4, -2) &= x(1, -2, 0, 3) + y(2, 3, 0, -1) + z(2, -1, 2, 1)\\ &= (x + 2y + 2z, -2x + 3y - z, 2z, 3x - y + z)\end{aligned}$$

Formons le système d'équations équivalent en égalant les composantes correspondantes, et réduisons-le à la forme échelonnée :

$$\begin{aligned} x + 2y + 2z &= 3\\ -2x + 3y - z &= 9\\ 2z &= -4\\ 3x - y + z &= -2 \end{aligned} \quad \text{ou} \quad \begin{aligned} x + 2y + 2z &= 3\\ 7y + 3z &= 15\\ 2z &= -4\\ -7y - 5z &= -11 \end{aligned} \quad \text{ou} \quad \begin{aligned} x + 2y + 2z &= 3\\ 7y + 3z &= 15\\ 2z &= -4\\ -2z &= 4 \end{aligned}$$

$$\text{ou} \quad \begin{aligned} x + 2y + 2z &= 3\\ 7y + 3z &= 15\\ 2z &= -4 \end{aligned}$$

Remarquons que le système précédent est possible et donc a une solution : d'où v est une combinaison linéaire des u_i. En résolvant ce système en x, y, z on obtient $x = 1$, $y = 3$, $z = -2$. Ainsi $v = u_1 + 3u_2 - 2u_3$.

Si le système d'équations linéaires n'est pas possible, c'est-à-dire s'il n'a pas de solution, le vecteur v n'est pas une combinaison linéaire des u_i.

ESPACE LIGNE D'UNE MATRICE

Soit A une matrice arbitraire $m \times n$ sur un corps K :

$$A = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{1n}\\ a_{21} & a_{22} & \dots & a_{2n}\\ \dots & \dots & \dots & \dots\\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

Les lignes de A sont

$$R_1 = (a_{11}, a_{21}, \dots, a_{1n}), \quad \dots, \quad R_m = (a_{m1}, a_{m2}, \dots, a_{mn})$$

et peuvent être considérées comme des vecteurs de K^n engendrant un sous-espace de K^n appelé espace ligne de A, c'est-à-dire que

$$\text{l'espace ligne de } A = L(R_1, R_2, \dots, R_m).$$

De manière analogue, les colonnes de A peuvent être considérées comme des vecteurs de K^m engendrant un sous-espace de K^m, appelé espace colonne de A.

Supposons maintenant que nous appliquions sur A les opérations élémentaires sur les lignes suivantes :

$$\text{(i) } R_i \leftrightarrow R_j, \quad \text{(ii) } R_i \to kR_i,\ k \neq 0, \quad \text{ou} \quad \text{(iii) } R_i \to kR_j + R_i$$

et que nous obtenions une matrice B. Chaque ligne de B est donc soit une ligne de A, soit une combinaison linéaire des lignes de A. En conséquence l'espace ligne de B est contenu dans l'espace ligne de A. D'autre part, nous pouvons appliquer les opérations élémentaires inverses des précédentes sur B et obtenir A ; d'où l'espace ligne de A est contenu dans l'espace ligne de B. En conclusion A et B ont le même espace ligne, ce qui nous conduit au théorème suivant.

Théorème 4.6 : Deux matrices équivalentes lignes ont le même espace ligne.

Nous démontrerons (problème 4.31), en particulier, le résultat fondamental concernant les matrices échelonnées réduite par les lignes (e.r.l.).

Théorème 4.7 : Des matrices e.r.l. ont le même espace ligne si et seulement si elles ont les mêmes lignes non nulles.

Ainsi chaque matrice est équivalente ligne à une unique matrice e.r.l, appelée forme canonique ligne.

Nous appliquons les résultats précédents dans l'exemple suivant.

Exemple 4.14 : Montrer que l'espace U engendré par les vecteurs

$$u_1 = (1, 2, -1, 3), \quad u_2 = (2, 4, 1, -2), \quad \text{et} \quad u_3 = (3, 6, 3, -7)$$

et l'espace V engendré par les vecteurs

$$v_1 = (1, 2, -4, 11) \quad \text{et} \quad v_2 = (2, 4, -5, 14)$$

sont égaux, c'est-à-dire $U = V$.

Méthode 1. Montrer que chaque u_i est une combinaison linéaire de v_1 et v_2 et montrer que chaque v_i est une combinaison linéaire des u_1, u_2, u_3. Remarquer que nous avons ainsi montré que les six systèmes d'équations linéaires sont compatibles.

Méthode 2. Formons la matrice A dont les lignes sont les u_i, et réduisons ligne A à sa forme ligne canonique :

$$A = \begin{pmatrix} 1 & 2 & -1 & 3 \\ 2 & 4 & 1 & -2 \\ 3 & 6 & 3 & -7 \end{pmatrix} \text{d'où} \begin{pmatrix} 1 & 2 & -1 & 3 \\ 0 & 0 & 3 & -8 \\ 0 & 0 & 6 & -16 \end{pmatrix} \text{d'où} \begin{pmatrix} 1 & 2 & -1 & 3 \\ 0 & 0 & 3 & -8 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\text{d'où} \begin{pmatrix} 1 & 2 & 0 & 1/3 \\ 0 & 0 & 1 & -8/3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Formons maintenant la matrice B dont les lignes sont v_1 et v_2 et réduisons les lignes de B à la forme canonique :

$$B = \begin{pmatrix} 1 & 2 & -4 & 11 \\ 2 & 4 & -5 & 14 \end{pmatrix} \text{d'où} \begin{pmatrix} 1 & 2 & -4 & 11 \\ 0 & 0 & 3 & -8 \end{pmatrix} \text{d'où} \begin{pmatrix} 1 & 2 & 0 & 1/3 \\ 0 & 0 & 1 & -8/3 \end{pmatrix}$$

Puisque les lignes non nulles des matrices réduites sont identiques, les espaces lignes de A et B sont égaux et donc $U = V$.

SOMMES ET SOMMES DIRECTES

Soient U et W deux sous-espaces d'un espace vectoriel V. La somme de U et W, écrite $U + W$ contient toutes les sommes $u + w$ où $u \in U$ et $w \in W$:

$$U + W = \{u + w : u \in U, w \in W\}$$

Remarquons que $0 = 0 + 0 \in U + W$ puisque $0 \in U$ et $0 \in W$. De plus supposons que $u + w$ et $u' + w'$ appartiennent à $U + W$, avec $u, u' \in U$ et $w, w' \in W$. Alors

$$(u+w) + (u'+w') = (u+u') + (w+w') \in U + W$$

et pour un scalaire quelconque $\quad k(u+w) = ku + kw \in U + W$

Ainsi le théorème précédent est démontré.

Théorème 4.8 : La somme $U + W$ de deux sous-espaces U et W est aussi un sous-espace de V.

Exemple 4.15 : Soit V l'espace vectoriel des matrices 2×2 sur $\mathbf{R}$. Soit U l'ensemble des matrices de V dont la seconde ligne est nulle, et soit W l'ensemble des matrices de V dont la seconde colonne est nulle :

$$U = \left\{\begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix} : a, b \in \mathbf{R}\right\}, \qquad W = \left\{\begin{pmatrix} a & 0 \\ c & 0 \end{pmatrix} : a, c \in \mathbf{R}\right\}$$

Donc U et W sont des sous-espaces de V; nous avons

$$\ldots b) : a, b, c \in \mathbf{R}\} \qquad \text{et} \qquad U \cap W = \left\{\begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} : a \in \mathbf{R}\right\}$$

$U + W$ ... W contient toutes les matrices dont le dernier élément à droite est nul ...ntient toutes les matrices dont la seconde ligne et la seconde colonne sont

...ctoriel V est dit être la somme directe des sous-espaces U et W, que l'on

$$V = U \oplus W$$

si chaque vecteur $v \in V$ peut être décomposé d'une manière unique en $v = u + w$ avec $u \in U$ et $w \in W$.

...démontre le théorème suivant :

Théorème 4.9 : L'espace vectoriel V est la somme directe des sous-espaces U et W si et seulement si si 1) $V = U + W$ et 2) $U \cap W = \{0\}$.

Exemple 4.16 : Dans l'espace vectoriel $\mathbf{R}^3$, soit U le plan xy et W le plan yz :

$$U = \{(a, b, 0) : a, b \in \mathbf{R}\} \qquad \text{et} \qquad W = \{(0, b, c) : b, c \in \mathbf{R}\}$$

Alors $\mathbf{R}^3 = U + W$ puisque chaque vecteur de $\mathbf{R}^3$ est la somme d'un vecteur de U et d'un vecteur de W. Cependant $\mathbf{R}^3$ n'est pas la somme directe de U et de W puisque de telles sommes ne sont pas uniques ; par exemple

$$(3, 5, 7) = (3, 1, 0) + (0, 4, 7) \quad \text{et aussi} \quad (3, 5, 7) = (3, -4, 0) + (0, 9, 7)$$

Exemple 4.17 : Dans $\mathbf{R}^3$, soit U le plan xy et W l'axe des z :

$$U = \{(a, b, 0) : a, b \in \mathbf{R}\} \qquad \text{et} \qquad W = \{(0, 0, c) : c \in \mathbf{R}\}$$

Un vecteur quelconque $(a, b, c) \in \mathbf{R}^3$ peut être écrit sous la forme d'une somme d'un vecteur de U et d'un vecteur de V d'une seule manière :

$$(a, b, c) = (a, b, 0) + (0, 0, c)$$

En conséquence $\mathbf{R}^3$ est la somme directe de U et W, c'est-à-dire $\mathbf{R}^3 = U \oplus W$.

PROBLEMES RESOLUS

ESPACES VECTORIELS

4.1. Démontrer le théorème 4.1 : Soit V un espace vectoriel sur le corps K.

1) Quel que soit le scalaire $k \in K$ et $0 \in V$, $k0 = 0$
2) Pour $0 \in K$ et un vecteur quelconque $u \in V$, $0u = 0$.
3) Si $ku = 0$, où $k \in K$ et $u \in V$, alors $k = 0$ ou $u = 0$.
4) Quel que soit $k \in K$ et quel que soit $u \in V$, $(k)u = k(-u) = -ku$.

1) D'après l'axiome $[A_2]$ avec $u = 0$, nous avons $0 + 0 = 0$. D'après l'axiome $[M_1]$, $k0 = k(0 + 0) = k0 + k0$. En ajoutant $-k0$ aux deux membres on obtient le résultat désiré.

2) D'après une propriété de K, $0 + 0 = 0$. D'où d'après l'axiome $[M_2]$, $0u = (0 + 0)u = 0u + 0u$. En ajoutant $-0u$ aux deux membres, on obtient le résultat demandé.

3) Supposons $ku = 0$ et $k \neq 0$. Il existe un scalaire k^{-1} tel que $k^{-1}k = 1$, d'où

$$u = 1u = (k^{-1}k)u = k^{-1}(ku) = k^{-1}0 = 0$$

4) En utilisant $u + (-u) = 0$, on obtient $0 = k0 = k[u + (-u)] = ku + k(-u)$. En ajoutant $-ku$ aux deux membres on a $-ku = k(-u)$.

En utilisant $k + (-k) = 0$, on obtient $0 = 0u = [k + (-k)]u = ku + (-k)u$. En ajoutant $-ku$ aux deux membres on obtient $-ku = (-k)u$. Ainsi $(-k)u = k(-u) = -ku$.

4.2. Montrer que quel que soit le scalaire k et quels que soient les vecteurs u et $v, k(u-v) = ku - kv$.

En utilisant la définition de la soustraction ($u - v = u + (-v)$) et le résultat du théorème 4.1 4) ($k(-v) = -kv$)

$$k(u-v) = k(u+(-v)) = ku + k(-v) = ku + (-kv) = ku - kv$$

4.3. Dans l'écriture de l'axiome $[M_2]$ $(a+b)u = au + bu$, quelle opération représente chaque signe plus ?

Le + dans $(a + b)u$ correspond à l'addition de deux scalaires a et b ; il représente donc l'addition dans le corps K. D'autre part, le + dans $au + bu$ représente l'addition de deux vecteurs au et bu ; ce plus représente donc l'addition vectorielle. En conclusion chaque + représente une opération différente.

4.4. Dans l'écriture de l'axiome $[M_3]$ $(ab)u = a(bu)$, quelle opération réprésente chaque produit ?

Dans $(ab)u$ le produit ab des scalaires a et b est une multiplication dans le corps K, tandis que le produit du scalaire ab et du vecteur u est une multiplication scalaire.

Dans $a(bu)$ le produit bu du scalaire b et du vecteur u est une multiplication scalaire, ainsi que le produit du scalaire a par le vecteur bu.

4.5. Soit V l'ensemble des fonctions d'un ensemble non vide X vers le corps K. Quelles que soient les fonctions $f, g \in V$ et quel que soit le scalaire $k \in K$, soient $f + g$ et kf les fonctions de V définies comme suit :

$$(f+g)(x) = f(x) + g(x) \quad \text{et} \quad (kf)(x) = kf(x), \quad \forall x \in X$$

(Le symbole $\forall$ veut dire "quel que soit" ou "pour tout"). Démontrer que V est un espace vectoriel sur K.

Puisque X n'est pas vide, V est aussi non vide. Montrons maintenant que tous les axiomes de l'espace vectoriel sont vérifiés.

$[A_1]$: Soit $f, g, h \in V$. Pour montrer que $(f + g) + h = f + (g + h)$ il est nécessaire de montrer que la fonction $f + (g + h)$ et la fonction $(f + g) + h$ donnent à chaque $x \in X$ la même valeur. D'où :

$$((f+g)+h)(x) = (f+g)(x) + h(x) = (f(x)+g(x)) + h(x), \quad \forall x \in X$$
$$(f+(g+h))(x) = f(x) + (g+h)(x) = f(x) + (g(x)+h(x)), \quad \forall x \in X$$

Mais $f(x)$, $g(x)$ et $h(x)$ sont des scalaires sur le corps K où l'addition des scalaires est associative, d'où

$$(f(x) + g(x)) + h(x) = f(x) + (g(x) + h(x))$$

En conséquence $(f + g) + h = f + (g + h)$.

$[A_2]$: Soit $\mathbf{0}$ la fonction nulle $\mathbf{0}(x) = 0, \forall x \in X$. Alors pour toute fonction $f \in V$,

$$(f+\mathbf{0})(x) = f(x) + \mathbf{0}(x) = f(x) + 0 = f(x), \quad \forall x \in X$$

Ainsi $f + \mathbf{0} = f$ et $\mathbf{0}$ est le vecteur nul de V.

$[A_3]$: Pour toute fonction $f \in V$, soit $-f$ la fonction définie par $(-f)(x) = -f(x)$. Donc

$$(f + (-f))(x) = f(x) + (-f)(x) = f(x) - f(x) = 0 = \mathbf{0}(x), \quad \forall x \in X$$

D'où $f + (-f) = \mathbf{0}$.

$[A_4]$: Soient f, $g \in V$, alors

$$(f+g)(x) = f(x) + g(x) = g(x) + f(x) = (g+f)(x), \quad \forall x \in X$$

Donc $f + g = g + f$. (Remarquons que $f(x) + g(x) = g(x) + f(x)$ découle du fait que $f(x)$ et $g(x)$ sont des scalaires sur le corps K où l'addition est commutative).

$[M_1]$: Soient f, $g \in V$ et $k \in K$, alors

$$(f+g))(x) = k((f+g)(x)) = k(f(x) + g(x)) = kf(x) + kg(x)$$
$$= (kf)(x) + (kg)(x) = (kf+kg)(x), \quad \forall x \in X$$

Donc $k(f + g) = kf + kg$. (Remarquons que $k(f(x) + g(x)) = kf(x) + kg(x)$ découle du fait que $k, f(x)$ et $g(x)$ sont des scalaires sur le corps K où la multiplication est distributive par rapport à l'addition).

$[M_2]$: Soient $f \in V$ et a, $b \in K$. D'où

$$((a+b)f)(x) = (a+b)f(x) = af(x) + bf(x) = (af)(x) + bf(x)$$
$$= (af+bf)(x), \quad \forall x \in X$$

Donc $(a + b)f = af + bf$.

$[M_3]$: Soient $f \in V$ et a, $b \in K$. D'où

$$((ab)f)(x) = (ab)f(x) = a(bf(x)) = a(bf)(x) = (a(bf))(x), \quad \forall x \in X$$

Donc $(ab)f = a(bf)$.

$[M_4]$: Soit $f \in V$. Pour l'élément neutre $1 \in K$, $(1f)(x) = 1f(x) = f(x)\ \forall x \in X$. Donc $1f = f$.

Tous les axiomes étant vérifiés, V est un espace vectoriel sur K.

4.6. Soit V l'ensemble de tous les couples de nombres réels $V = \{(a, b)\ ;\ a,\ b \in \mathbf{R}\}$. Montrer que V n'est pas un espace vectoriel sur R à l'aide des lois suivantes: addition dans V et multiplication scalaire sur V :

(i) $(a, b) + (c, d) = (a + c, b + d)$ et $k(a, b) = (ka, b)$;

(ii) $(a, b) + (c, d) = (a, b)$ et $k(a, b) = (ka, kb)$;

(iv) $(a, b) + (c, d) = (a + c, b + d)$ et $k(a, b) = (k^2a, k^2b)$.

Dans chaque cas nous montrerons qu'un des axiomes de l'espace vectoriel n'est pas vérifié.

(i) Soit $r = 1$, $s = 2$, $v = (3, 4)$. Alors

$$(r+s)v = 3(3, 4) = (9, 4)$$
$$rv + sv = 1(3, 4) + 2(3, 4) = (3, 4) + (6, 4) = (9, 8)$$

Puisque $(r + s)v \neq rv + sv$ l'axiome $[M_2]$ n'est pas vérifié.

(ii) Soit $v = (1, 2); w = (3, 4)$. Alors

$$v + w = (1, 2) + (3, 4) = (1, 2)$$
$$w + v = (3, 4) + (1, 2) = (3, 4)$$

Puisque $v + w \neq w + v$ l'axiome $[A_4]$ n'est pas vérifié.

(iii) Soit $r = 1$, $s = 2$, $v = (3, 4)$. Alors

$$(r+s)v = 3(3, 4) = (27, 36)$$
$$rv + sv = 1(3, 4) + 2(3, 4) = (3, 4) + (12, 16) = (15, 20)$$

Ainsi $(r + s)v \neq rv + sv$ et l'axiome $[M_2]$ n'est pas vérifié.

SOUS-ESPACES :

4.7 Démontrer le théorème 4.2 : W est un sous-espace de V si et seulement si (i) W n'est pas vide, (ii) v, $w \in W$ implique $v + w \in W$ et (iii) $v \in W$ implique $k\,v \in W$ pour tout scalaire $k \in K$.

Supposons que W satisfasse à (i), (ii), (iii). D'après (i) W n'est pas vide, et d'après (ii) et (iii) les opérations d'addition vectorielle et de multiplication par un scalaire sont bien définies dans W. D'ailleurs, les axiomes $[A_1]$, $[A_4]$, $[M_1]$, $[M_2]$, $[M_3]$ et $[M_4]$ sont vérifiés dans W puisque les vecteurs de W appartiennent à V. Ainsi il ne reste plus qu'à montrer que $[A_2]$ et $[A_3]$ sont aussi vérifiés dans W. D'après (i), W n'est pas vide, d'où $u \in W$. Alors d'après (iii) $0u = 0 \in W$ et $v + 0 = v$ pour tout $v \in W$. Ainsi W satisfait $[A_2]$. Finalement, si $v \in W$ alors $(-1)\,v = -v \in W$ et $v + (-v) = 0$; ainsi W satisfait $[A_3]$. D'où W est un sous-espace de V.

Réciproquement, si W est un sous-espace de V alors (i), (ii), (iii) sont clairement vérifiés.

4.8. Démontrer le corollaire 4.3 : W est un sous-espace de V si et seulement si (i) $0 \in W$ et (ii) v, $w \in W$ implique $av + bw \in W$ pour tous scalaires a, $b \in K$.

Supposons que W satisfasse (i) et (ii). Alors d'après (i) W n'est pas vide. De plus, si v, $w \in W$, alors d'après (ii) $v + w = 1v + 1w \in W$ et si $v \in W$ et $k \in K$ alors d'après (ii) $kv = kv + 0v \in W$. Ainsi d'après le théorème 4.2, W est sous-espace de V.

Réciproquement, si W est un sous-espace de V, alors (i) et (ii) sont vérifiés dans W.

4.9. Soit $V = \mathbf{R}^3$. Montrer que W est un sous-espace de V, où :

(i) $W = \{(a, b, 0)\ ;\ a, b \in \mathbf{R}\}$, c'est-à-dire W est le plan xy contenant ceux des vecteurs dont la troisième composante est 0.

(ii) $W = \{(a, b, c)\ ;\ a + b + c = 0\}$, c'est-à-dire W contient ceux des vecteurs qui ont pour propriété d'avoir la somme de leurs trois composantes nulle.

(i) $0 = (0, 0, 0) \in W$ puisque la troisième composante de 0 est nulle. Quels que soient les vecteurs $v = (a, b, 0)$, $w = (c, d, 0)$ dans W et quels que soient les scalaires (nombres réels) k et k',

$$\begin{aligned} kv + k'w &= k(a, b, 0) + k'(c, d, 0) \\ &= (ka, kb, 0) + (k'c, k'd, 0) = (ka + k'c, kb + k'd, 0) \end{aligned}$$

Ainsi $kv + k'w \in W$ et donc W est un sous-espace de V.

(ii) $0 = (0, 0, 0) \in W$ puisque $0 + 0 + 0 = 0$. Supposons que $v = (a, b, c)$, $w = (a', b', c')$ appartiennent à W c'est-à-dire que $a + b + c = 0$ et $a' + b' + c' = 0$. Alors pour n'importe quels scalaires k et k',

$$\begin{aligned} kv + k'w &= k(a, b, c) + k'(a', b', c') \\ &= (ka, kb, kc) + (k'a', k'b', k'c') \\ &= (ka + k'a', kb + k'b', kc + k'c') \end{aligned}$$

et de plus

$$\begin{aligned} (ka + k'a') + (kb + k'b') + (kc + k'c') &= k(a + b + c) + k'(a' + b' + c') \\ &= k0 + k'0 = 0 \end{aligned}$$

Ainsi $kv + k'w \in K$ et donc W est un sous-espace de V.

4.10. Soit $V = \mathbf{R}^3$. Montrer que W n'est pas un sous-espace de V, où :

(i) $W = \{(a, b, c)\ ;\ a \geqslant 0\}$; c'est-à-dire W contient ceux des vecteurs dont la première composante est positive ou nulle.

(ii) $W = \{(a, b, c) : a^2 + b^2 + c^2 \leqslant 1\}$; c'est-à-dire W contient ceux des vecteurs dont la longueur est inférieure ou égale à 1.

(iii) $W = \{(a, b, c) : a, b, c \in Q\}$; c'est-à-dire W contient ceux des vecteurs dont les composantes sont des nombres rationnels.

Dans chaque cas, montrons qu'une des propriétés du théorème 4.2 n'est pas vérifiée.

(i) $v = (1, 2, 3) \in W$ et $k = -5 \in \mathbf{R}$. Mais $kv = -5(1, 2, 3) = (-5, -10, -15)$ n'appartient pas à W puisque -5 est négatif. D'où W n'est pas un sous-espace de V.

(ii) $v = (1, 0, 0) \in W$ et $w = (0, 1, 0) \in W$. Mais $v + w = (1, 0, 0) + (0, 1, 0) = (1, 1, 0)$ n'appartient pas à W puisque $1^2 + 1^2 + 0^2 = 2 > 1$. D'où W n'est pas un sous-espace vectoriel de V.

(iii) $v = (1, 2, 3) \in W$ et $k = \sqrt{2} \in \mathbf{R}$. Mais $kv = \sqrt{2}(1, 2, 3) = (\sqrt{2}, 2\sqrt{2}, 3\sqrt{2})$ n'appartient pas à W puisque ses composantes ne sont pas des nombres rationnels. D'où W n'est pas un sous-espace vectoriel de V.

4.11. Soit V l'espace vectoriel des matrices carrés $n \times n$ sur le corps K. Montrer que W est un sous-espace de V où

(i) W contient les matrices symétriques ; c'est-à-dire toutes les matrices $A = (a_{ij})$ pour lesquelles $a_{ji} = a_{ij}$.

(ii) W est l'ensemble de toutes les matrices qui commutent avec une matrice donnée T ; c'est-à-dire $W = \{A \in V : AT = TA\}$.

(i) $0 \in W$ puisque tous les éléments de 0 sont nuls et donc égaux. Maintenant supposons que $A = (a_{ij})$ et $B = (b_{ij})$ appartiennent à W ; c'est-à-dire que $a_{ji} = a_{ij}$ et $b_{ji} = b_{ij}$. Donc quels que soient les scalaires $a, b \in K$, $aA + bB$ est une matrice dont le ij élément est $aa_{ij} + bb_{ij}$. Mais $aa_{ji} + bb_{ji} = aa_{ij} + bb_{ij}$. Ainsi $aA + bB$ est aussi symétrique et donc W est un sous-espace de V.

(ii) $0 \in W$ puisque $0T = 0 = T0$. Supposons maintenant que $A, B \in W$, c'est-à-dire que $AT = TA$ et $BT = TB$. Quels que soient les scalaires $a, b \in K$,

$$\begin{aligned}(aA+bB)T &= (aA)T + (bB)T = a(AT) + b(BT) = a(TA) + b(TB) \\ &= T(aA) + T(bB) = T(aA+bB)\end{aligned}$$

Ainsi $aA + bB$ commute avec T, c'est-à-dire appartient à W, donc W est un sous-espace de V.

4.12. Soit V l'espace vectoriel de toutes les matrices 2×2 sur un corps $\mathbf{R}$ réel. Montrer que W n'est pas un sous-espace vectoriel de V si

(i) W est l'ensemble de toutes les matrices ayant un déterminant nul.

(ii) W est l'ensemble de toutes les matrices A telles que $A^2 = A$.

(i) Rappelons que le déterminant de $\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$ Les matrices $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ et $B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ appartiennent à W puisque $\det(A) = 0$ et $\det(B) = 0$. Mais $A + B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ n'appartient pas à W puisque $\det(A + B) = 1$. Donc W n'est pas un sous-espace de V.

(ii) La matrice unitaire $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ appartient à W puisque

$$I^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$$

Mais $2I = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ n'appartient pas à W puisque

$$(2I)^2 = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix} \neq 2I$$

D'où W n'est pas un sous-espace de V.

4.13. Soit V l'espace vectoriel de toutes les fonctions du corps réel $\mathbf{R}$ sur $\mathbf{R}$. Montrer que W est un sous-espace de V où

(i) $W = \{f : f(3) = 0\}$; c'est-à-dire W contient les fonctions qui font correspondre à 3 le nombre 0.

(ii) $W = \{f : f(7) = f(1)\}$; c'est-à-dire W est l'ensemble des fonctions qui ont pour 7 et 1 la même image.

(iii) W est l'ensemble des fonctions impaires, c'est-à-dire les fonctions f pour lesquelles $f(-x) = -f(x)$.

Nous noterons ici $\mathbf{0}$ la fonction nulle $\mathbf{0}(x) = 0$ pour tout $x \in R$.

(i) $\mathbf{0} \in W$ puisque $\mathbf{0}(3) = 0$. Supposons $f, g \in W$ c'est-à-dire tels que $f(3) = 0$ et $g(3) = 0$. D'où quels que soient les réels a et b

$$(af + bg)(3) = af(3) + bg(3) = a0 + b0 = 0$$

Donc $af + bg \in W$ et W est un sous-espace de V.

(ii) $\mathbf{0} \in W$ puisque $\mathbf{0}(7) = 0 = \mathbf{0}(1)$. Supposons que $f, g \in W$ c'est-à-dire que $f(7) = f(1)$ et $g(7) = g(1)$. Alors quels que soient les nombres réels a et b,

$$(af + bg)(7) = af(7) + bg(7) = af(1) + bg(1) = (af + bg)(1)$$

Donc $af + bg \in W$ et donc W est un sous-espace de V.

(iii) $\mathbf{0} \in W$ puisque $\mathbf{0}(-x) = 0 = -0 = -\mathbf{0}(x)$. Supposons que $f, g \in W$, c'est-à-dire que $f(-x) = -f(x)$ et $g(-x) = -g(x)$. Alors quels que soient les réels a et b,

$$(af + bg)(-x) = af(-x) + bg(-x) = -af(x) - bg(x) = -(af(x) + bg(x)) = -(af + bg)(x)$$

Donc $af + bg \in W$ et donc W est un sous-espace de V.

4.14. Soit V l'espace vectoriel de toutes les fonctions numériques de $\mathbf{R}$ sur $\mathbf{R}$. Montrer que W n'est pas un sous-espace vectoriel de V si :

(i) $W = \{f : f(7) = 2 + f(1)\}$;

(ii) W est l'ensemble des fonctions non négatives ; c'est-à-dire toutes les fonctions f pour lesquelles $f(x) \geqslant 0, \forall x \in \mathbf{R}$.

(i) Supposons $f, g \in W$ c'est-à-dire $f(7) = 2 + f(1)$ et $g(7) = 2 + g(1)$. Alors

$$\begin{aligned}(f+g)(7) &= f(7) + g(7) = 2 + f(1) + 2 + g(1) \\ &= 4 + f(1) + g(1) = 4 + (f+g)(1) \neq 2 + (f+g)(1)\end{aligned}$$

Donc $f + g \notin W$ et donc W n'est pas un sous-espace de V.

(ii) Si $k = -2$, soit $f \in V$ définie par $f(x) = x^2$. Alors $f \in W$ puisque $f(x) = x^2 \geqslant 0, \forall x \in \mathbf{R}$. Mais $(kf)(5) = kf(5) = (-2)(5^2) = -50 < 0$. Donc $kf \notin W$ et donc W n'est pas un sous-espace de V.

4.15. Soit V l'espace vectoriel des polynômes $a_0 + a_1 t + a_2 t^2 + \ldots + a_n t^n$ à coefficients réels, c'est-à-dire $a_i \in \mathbf{R}$. Déterminer si oui ou non W est un sous-espace de V avec

(i) W ensemble de tous les polynômes à coefficients entiers.

(ii) W ensemble de tous les polynômes de degré $\leqslant 3$.

(iii) W ensemble de tous les polynômes $b_0 + b_1 t^2 + b_2 t^4 + \ldots + b_n t^{2n}$, c'est-à-dire des polynômes dont tous les termes sont élevés à des puissances paires.

(i) Non, puisque les multiples scalaires de vecteurs dans W n'appartiennent pas toujours à W. Par exemple $v = 3 + 5t + 7t^2 \in W$. Mais $\frac{1}{2}v = \frac{3}{2} + \frac{5}{2}t + \frac{7}{2}t^2 \notin W$. (Remarquons que W est "fermé" pour l'addition vectorielle, c'est-à-dire que les sommes d'éléments de W appartiennent à W).

(ii) et (iii) Oui. Pour chaque cas, W n'est pas vide, les sommes d'éléments de W appartiennent à W, et les multiples scalaires d'un élément quelconque de W appartiennent à W.

4.16. Considérons un système linéaire homogène à n inconnues $x_1, x_2 \ldots x_n$ sur un corps K :

$$\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0 \end{array}$$

Montrer que l'ensemble des solutions W est un sous-espace de l'espace vectoriel K^n.

$0 = (0, 0, 0 \ldots 0) \in W$ puisque l'on a très facilement

$$a_{i1}0 + a_{i2}0 + \cdots + a_{in}0 = 0, \quad \text{pour } i = 1, \ldots, m$$

Supposons que $u = (u_1, u_2, \ldots, u_n)$ et $v = (v_1, v_2, \ldots, v_n)$ appartiennent à W c'est-à-dire que pour $i = 1, \ldots, m$

$$a_{i1}u_1 + a_{i2}u_2 + \cdots + a_{in}u_n = 0$$
$$a_{i1}v_1 + a_{i2}v_2 + \cdots + a_{in}v_n = 0$$

Soient a et b des scalaires de K. Alors

$$au + bv = (au_1 + bv_1, au_2 + bv_2, \ldots, au_n + bv_n)$$

et pour $i = 1, \ldots, m$

$$\begin{aligned} & a_{i1}(au_1 + bv_1) + a_{i2}(au_2 + bv_2) + \cdots + a_{in}(au_n + bv_n) \\ &= a(a_{i1}u_1 + a_{i2}u_2 + \cdots + a_{in}u_n) + b(a_{i1}v_1 + a_{i2}v_2 + \cdots + a_{in}v_n) \\ &= a0 + b0 = 0 \end{aligned}$$

Ainsi $au + bv$ est une solution du système, c'est-à-dire appartient à W. Donc W est un sous-espace de K^n.

COMBINAISONS LINEAIRES

4.17. Ecrire le vecteur $v = (1, -2, 5)$ comme combinaison linéaire des vecteurs $e_1 = (1, 1, 1)$, $e_2 = (1, 2, 3)$ et $e_3 = (2, -1, 1)$.

Nous désirons exprimer v sous la forme $v = xe_1 + ye_2 + ze_3$ avec x, y, z qui sont des scalaires inconnus. Donc nous avons

$$\begin{aligned} (1, -2, 5) &= x(1, 1, 1) + y(1, 2, 3) + z(2, -1, 1) \\ &= (x, x, x) + (y, 2y, 3y) + (2z, -z, z) \\ &= (x + y + 2z, x + 2y - z, x + 3y + z) \end{aligned}$$

Formons le système équivalent d'équations en égalant les composantes correspondantes les unes aux autres, puis réduisons à une forme échelonnée :

$$\begin{aligned} x + y + 2z &= 1 \\ x + 2y - z &= -2 \\ x + 3y + z &= 5 \end{aligned} \quad \text{ou} \quad \begin{aligned} x + y + 2z &= 1 \\ y - 3z &= -3 \\ 2y - z &= 4 \end{aligned} \quad \text{ou} \quad \begin{aligned} x + y + 2z &= 1 \\ y - 3z &= -3 \\ 5z &= 10 \end{aligned}$$

Remarquons que le système précédent est possible et donc a une solution. En résolvant ce système on obtient pour les inconnues les valeurs $x = -6, y = 3, z = 2$. D'où $v = -6e_1 + 3e_2 + 2e_3$

4.18. Ecrire le vecteur $v = (2, -5, 3)$ de $\mathbf{R}^3$ comme combinaison linéaire des vecteurs $e_1 = (1, -3, 2)$, $e_2 = (2, -4, -1)$ et $e_3 = (1, -5, 7)$.

Représentons v comme une combinaison linéaire de e_i en utilisant les inconnues x, y, z sous la forme $v = xe_1 + ye_2 + ze_3$.

$$\begin{aligned} (2, -5, 3) &= x(1, -3, 2) + y(2, -4, -1) + z(1, -5, 7) \\ &= (x + 2y + z, -3x - 4y - 5z, 2x - y + 7z) \end{aligned}$$

Formons le système équivalent d'équations et réduisons-le à sa forme échelonnée :

$$\begin{aligned} x + 2y + z &= 2 \\ -3x - 4y - 5z &= -5 \\ 2x - y + 7z &= 3 \end{aligned} \quad \text{ou} \quad \begin{aligned} x + 2y + z &= 2 \\ 2y - 2z &= 1 \\ -5y + 5z &= -1 \end{aligned} \quad \text{ou} \quad \begin{aligned} x + 2y + z &= 2 \\ 2y - 2z &= 1 \\ 0 &= 3 \end{aligned}$$

Le système est impossible et n'a donc pas de solution. D'où v ne peut pas être écrit sous forme d'une combinaison linéaire des vecteurs e_1, e_2, e_3.

4.19. Pour quelle valeur de k le vecteur $u = (1, -2, k)$ de $\mathbf{R}^3$ est-il une combinaison linéaire des vecteurs $v = (3, 0, 2)$ et $w = (2, -1, -5)$?

Soit $u = xv + yw$.

$$(1, -2, k) = x(3, 0, -2) + y(2, -1, -5) = (3x + 2y, -y, -2x - 5y)$$

Formons le système d'équations équivalent :

$$3x + 2y = 1, \quad -y = -2, \quad -2x - 5y = k$$

D'après les deux premières équations on obtient $x = -1$, $y = 2$. En substituant dans la dernière équation on obtient $k = -8$.

4.20. Ecrire le polynôme $v = t^2 + 4t - 3$ de R comme combinaison linéaire des polynômes $e_1 = t^2 - 2t + 5$, $e_2 = 2t^2 - 3t$ et $e_3 = t + 3$

Ecrivons v comme une combinaison linéaire des e_i en utilisant les inconnues x, y et z : $v = xe_1 + ye_2 + ze_3$.

$$\begin{aligned} t^2 + 4t - 3 &= x(t^2 - 2t + 5) + y(2t^2 - 3t) + z(t + 3) \\ &= xt^2 - 2xt + 5x + 2yt^2 - 3yt + zt + 3z \\ &= (x + 2y)t^2 + (-2x - 3y + z)t + (5x + 3z) \end{aligned}$$

Egalons les uns aux autres les coefficients des mêmes puissances de t et réduisons le système obtenu à sa forme échelonnée :

$$\begin{array}{rcl} x + 2y &=& 1 \\ -2x - 3y + z &=& 4 \\ 5x + 3z &=& -3 \end{array} \quad \text{ou} \quad \begin{array}{rcl} x + 2y &=& 1 \\ y + z &=& 6 \\ -10y + 3z &=& -8 \end{array} \quad \text{ou} \quad \begin{array}{rcl} x + 2y &=& 1 \\ y + z &=& 6 \\ 13z &=& 52 \end{array}$$

Remarquons que le système étant possible admet une solution. En résolvant par rapport aux inconnues on obtient $x = -3$, $y = 2$, $z = 4$. D'où $v = -3e_1 + 2e_2 + 4e_3$.

4.21. Ecrire la matrice $E = \begin{pmatrix} 3 & 1 \\ 1 & -2 \end{pmatrix}$ sous forme de combinaison linéaire des matrices $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$ et $C = \begin{pmatrix} 0 & 2 \\ 0 & -1 \end{pmatrix}$.

Soit $E = xA + yB + zC$ une combinaison linéaire de A, B, C utilisant les inconnues x, y z.

$$\begin{aligned} \begin{pmatrix} 3 & 1 \\ 1 & -1 \end{pmatrix} &= x\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} + y\begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} + z\begin{pmatrix} 0 & 2 \\ 0 & -1 \end{pmatrix} \\ &= \begin{pmatrix} x & x \\ x & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ y & y \end{pmatrix} + \begin{pmatrix} 0 & 2z \\ 0 & -z \end{pmatrix} = \begin{pmatrix} x & x + 2z \\ x + y & y - z \end{pmatrix} \end{aligned}$$

Formons le système d'équations équivalent en égalant les éléments correspondants des deux membres :

$$x = 3, \quad x + y = 1, \quad x + 2z = 1, \quad y - z = -1$$

Remplaçons x par 3 dans la seconde et la troisième équations; on obtient $y = -2$ et $z = -1$. Puisque ces valeurs satisfont aussi la dernière équation, elles forment une solution du système. Donc $E = 3A - 2B - C$.

4.22. Supposons que u soit une combinaison linéaire des vecteurs $v_1, \ldots, v_m$ et supposons que chaque v_i est une combinaison linéaire des vecteurs $w_1, \ldots, w_n$:

$$u = a_1v_1 + a_2v_2 + \cdots + a_mv_m \quad \text{et} \quad v_i = b_{i1}w_1 + b_{i2}w_2 + \cdots + b_{in}w_n$$

Montrer que u est aussi une combinaison linéaire des w_i. Ainsi si $S \subset L(T)$ alors $L(S) \subset L(T)$.

$$\begin{aligned} u &= a_1v_1 + a_2v_2 + \cdots + a_mv_m \\ &= a_1(b_{11}w_1 + \cdots + b_{1n}w_n) + a_2(b_{21}w_1 + \cdots + b_{2n}w_n) + \cdots + a_m(b_{m1}w_1 + \cdots + b_{mn}w_n) \\ &= (a_1b_{11} + a_2b_{21} + \cdots + a_mb_{m1})w_1 + \cdots + (a_1b_{1n} + a_2b_{2n} + \cdots + a_mb_{mn})w_n \end{aligned}$$

ou plus simplement $u = \sum_{i=1}^{m} a_iv_i = \sum_{i=1}^{m} a_i\left(\sum_{j=1}^{n} b_{ij}w_j\right) = \sum_{j=1}^{n}\left(\sum_{i=1}^{m} a_ib_{ij}\right)w_j$

GENERATEURS

4.23. Montrer que les vecteurs $u = (1, 2, 3)$, $v = (0, 1, 2)$ et $w = (0, 0, 1)$ engendrent $\mathbf{R}^3$.

Il est nécessaire de montrer qu'un vecteur arbitraire $(a, b, c) \in \mathbf{R}^3$ est une combinaison linéaire de u, v et w.

Soit $(a, b, c) = xu + yv + zw$:

$$(a, b, c) = x(1, 2, 3) + y(0, 1, 2) + z(0, 0, 1) = (x, 2x + y, 3x + 2y + z)$$

Formons alors le système d'équations

$$\begin{array}{lcl} x & = & a \\ 2x + y & = & b \\ 3x + 2y + z & = & c \end{array} \quad \text{ou} \quad \begin{array}{rcl} z + 2y + 3x & = & c \\ y + 2x & = & b \\ x & = & a \end{array}$$

Le système précédent est sous forme échelonnée et est possible; en fait $x = a$, $y = b - 2a$, $z = c - 2b + a$ est une solution. Donc u, v et w engendrent $\mathbf{R}^3$.

4.24. Trouver à quelles conditions sur a, b et c le vecteur $(a, b, c) \in \mathbf{R}^3$ appartient à l'espace engendré par $u = (2, 1, 0)$, $v = (1, -1, 2)$ et $w = (0, 3, -4)$

Ecrivons (a, b, c) comme une combinaison linéaire des u, v et w en utilisant les inconnues x, y et z: $(a, b, c) = xu + yv + zw$.

$$(a, b, c) = x(2, 1, 0) + y(1, -1, 2) + z(0, 3, -4) = (2x + y, x - y + 3z, 2y - 4z)$$

Formons le système d'équations linéaires équivalent et réduisons-le à sa forme échelonnée:

$$\begin{array}{rcl} 2x + y & = & a \\ x - y + 3z & = & b \\ 2y - 4z & = & c \end{array} \quad \text{ou} \quad \begin{array}{rcl} 2x + y & = & a \\ 3y - 6z & = & a - 2b \\ 2y - 4z & = & c \end{array} \quad \text{ou} \quad \begin{array}{rcl} 2x + y & = & a \\ 3y - 6z & = & a - 2b \\ 0 & = & 2a - 4b - 3c \end{array}$$

Le vecteur (a, b, c) appartient à l'espace engendré par u, v et w si et seulement si le précédent système est possible et il est possible si et seulement si $2a - 4b - 3c = 0$. Remarquons, en particulier, que u, v et w n'engendrent pas l'espace entier $\mathbf{R}^3$.

4.25. Montrer que le plan xy, $W = \{(a, b, 0)\}$ dans $\mathbf{R}^3$ est engendré par u et v où 1) $u = (1, 2, 0)$ et $v = (0, 1, 0)$; 2) $u = (2, -1, 0)$ et $v = (1, 3, 0)$.

Dans chaque cas montrons qu'un vecteur arbitraire $(a, b, 0) \in W$ est une combinaison linéaire de u et v.

1) Soit $(a, b, 0) = xu + yv$:

$$(a, b, 0) = x(1, 2, 0) + y(0, 1, 0) = (x, 2x + y, 0)$$

Formons maintenant le système d'équations

$$\begin{array}{rcl} x & = & a \\ 2x + y & = & b \\ 0 & = & 0 \end{array} \quad \text{ou} \quad \begin{array}{rcl} y + 2x & = & b \\ x & = & a \end{array}$$

Le système est possible ; en fait $x = a$, $y = b - 2a$ est une solution. Ainsi u et v engendrent W.

2) Soit $(a, b, 0) = xu + yv$:

$$(a, b, 0) = x(2, -1, 0) + y(1, 3, 0) = (2x + y, -x + 3y, 0)$$

Formons le système équivalent et réduisons-le à sa forme échelonnée :

$$\begin{array}{rcl} 2x + y & = & a \\ -x + 3y & = & b \\ 0 & = & 0 \end{array} \quad \text{ou} \quad \begin{array}{rcl} 2x + y & = & a \\ 7y & = & a + 2b \end{array}$$

Le système est possible et a donc une solution. Donc W est engendré par u et v (Remarquons qu'il n'est pas nécessaire de résoudre le système en x et y, il est uniquement nécessaire de savoir si une solution existe).

4.26. Montrer que l'espace vectoriel V des polynômes sur un corps quelconque K ne peut être engendré par un nombre fini de vecteurs.

Un ensemble quelconque fini S de polynômes contient un polynôme de degré maximum, par exemple m. Alors l'espace vectoriel $L(S)$ engendré par S ne peut pas contenir des polynômes de degré supérieur à m. Donc $V \neq L(S)$ pour un ensemble quelconque fini S.

4.27. Démontrer le théorème 4.5 : Soit S un sous-ensemble non vide de V. Soit $L(S)$ l'ensemble de toutes les combinaisons linéaires des vecteurs de S; montrer que $L(S)$ est un sous-espace de V contenant S. De plus, si W est un autre sous-espace de V contenant S, alors $L(S) \subset W$.

Si $v \in S$ aors $1v = v \in L(S)$; donc S est un sous-ensemble de $L(S)$. De même $L(S)$ n'est pas vide puisque S n'est pas vide. Supposons maintenant $v,\ w \in L(S)$, c'est-à-dire

$$v = a_1v_1 + \cdots + a_mv_m \quad \text{et} \quad w = b_1w_1 + \cdots + b_nw_n$$

où $v_i, w_j \in S$ et a_i, b_j sont des scalaires. Alors

$$v + w = a_1v_1 + \cdots + a_mv_m + b_1w_1 + \cdots + b_nw_n$$

et pour un scalaire quelconque k

$$kv = k(a_1v_1 + \cdots + a_mv_m) = ka_1v_1 + \cdots + ka_mv_m$$

appartient à $L(S)$ puisque chaque élément est une combinaison linéaire des vecteurs de S. Donc $L(S)$ est un sous-espace de V.

Supposons maintenant que W soit un sous-espace de V contenant S et supposons que $v_1, \ldots, v_m \in S \subset W$. Alors tous les multiples $a_1\ v_1, \ldots, a_m\ v_m \in W$ où $a_i \in K$ et donc la somme $a_1\ v_1 + \cdots + a_m v_m \in W$. Donc W contient toutes les combinaisons linéaires des éléments de S. En conséquence $L(S) \subset W$ comme nous l'avions affirmé.

ESPACE LIGNE D'UNE MATRICE

4.28. Déterminer si oui ou non les matrices suivantes ont le même espace ligne :

$$A = \begin{pmatrix} 1 & 1 & 5 \\ 2 & 3 & 13 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & -1 & -2 \\ 3 & -2 & -3 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & -1 & -1 \\ 4 & -3 & -1 \\ 3 & -1 & 3 \end{pmatrix}$$

Réduisons chacune de ces matrices à leur forme canonique ligne.

$$A = \begin{pmatrix} 1 & 1 & 5 \\ 2 & 3 & 13 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & 1 & 5 \\ 0 & 1 & 3 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \end{pmatrix}$$

$$B = \begin{pmatrix} 1 & -1 & -2 \\ 3 & -2 & -3 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & -1 & -2 \\ 0 & 1 & 3 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 3 \end{pmatrix}$$

$$C = \begin{pmatrix} 1 & -1 & -1 \\ 4 & -3 & -1 \\ 3 & -1 & 3 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & -1 & -1 \\ 0 & 1 & 3 \\ 0 & 2 & 6 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & -1 & -1 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix}$$

Puisque les lignes non nulles de la forme réduite de A et de la forme réduite de C sont les mêmes, A et C ont le même espace ligne. D'autre part, les lignes non nulles de la forme réduite de B ne sont pas les mêmes que les lignes des autres, et donc B n'a pas le même espace ligne que A et C.

4.29. Considérons une matrice arbitraire $A = (a_{ij})$. Supposons que $u = (b_1, \ldots, b_n)$ soit une combinaison linéaire des lignes $R_1, \ldots, R_m$ de A ; c'est-à-dire $u = k_1R_1 + \cdots + k_m R_m$. Montrer que, pour chaque i, $b_i = k_1\ a_{1i} + k_2\ a_{2i} + \ldots\ k_m\ a_{mi}$ où les $a_{1i}, \ldots, a_{mi}$ sont les éléments de la $i^{\text{ème}}$ colonne de A.

On a donné $u = k_1 R_1 + \ldots + k_m R_m$; d'où

$$\begin{aligned} (b_1, \ldots, b_n) &= k_1(a_{11}, \ldots, a_{1n}) + \cdots + k_m(a_{m1}, \ldots, a_{mn}) \\ &= (k_1a_{11} + \cdots + k_ma_{m1}, \ldots, k_1a_{m1} + \cdots + k_ma_{mn}) \end{aligned}$$

En égalant les composantes correspondantes les unes aux autres, nous obtenons le résultat désiré.

4.30. Démontrer : Soit $A = (a_{ij})$ une matrice échelonnée dont les éléments distingués sont $a_{1j_1}, a_{2j_2}, \ldots, a_{rj_r}$, et soit $B = (b_{ij})$ une matrice échelonnée dont les éléments distingués sont $b_{1k_1}, b_{2k_2}, \ldots, b_{sk_s}$:

$$A = \begin{pmatrix} a_{1j_1} & * & * & * & * & * & * \\ & & a_{2j_2} & * & * & * & * \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ & & & & a_{rj_r} & * & * \\ & & & & & & \end{pmatrix}, \qquad B = \begin{pmatrix} b_{1k_1} & * & * & * & * & * & * \\ & & b_{2k_2} & * & * & * & * \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ & & & & b_{sk_s} & * & * \\ & & & & & & \end{pmatrix}$$

Supposons que A et B aient le même espace ligne. Montrer alors que les éléments distingués de A et les éléments distingués de B sont dans la même position : $j_1 = k_1, j_2 = k_2, \ldots, j_r = k_r$ et $r = s$.

$A = 0$ si et seulement si $B = 0$, et il nous reste donc à démontrer le théorème précédent lorsque $r \geqslant 1$ et $s \geqslant 1$. Montrons d'abord que $j_1 = k_1$. Supposons $j_1 < k_1$. Alors la $j^{\text{ième}}$ colonne de B est nulle. Puisque la première ligne de A fait partie de l'espace de B, d'après le précédent problème nous avons : $a_{1j_1} = c_1 0 + c_2 0 + \ldots + c_m 0 = 0$ pour des scalaires c_i, mais ceci est en contradiction avec le fait que l'élément distingué $a_{1j_1} \neq 0$. Donc $j_1 \geqslant k_1$ et de manière analogue $k_1 \geqslant j_1$. D'où $j_1 = k_1$.

Soit maintenant A' la sous-matrice de A obtenue en supprimant la première ligne de A et soit B' la sous-matrice de B obtenu en supprimant la première ligne de B. Nous allons démontrer que A' et B' ont le même espace ligne. Le théorème pourra s'appliquer par récurrence puisque A' et B' sont aussi des matrices échelonnées.

Soit $R = (a_1, a_2, \ldots, a_n)$ une ligne quelconque de A', et soit $R_1, \ldots, R_m$ les lignes de B. Puisque R fait partie de l'espace ligne de B, il existe des scalaires $d_1, \ldots, d_m$, tels que $R = d_1R_1 + d_2R_2 + \ldots + d_mR_m$. Or A est sous sa forme échelonnée et R n'est pas la première ligne de A, le $j^{\text{ème}}$ de R est zéro : $a_i = 0$ pour $i = j_1 = k_1$. De plus B étant sous sa forme échelonnée, tous les éléments dans la $k^{\text{ième}}$ colonne de B sont nuls sauf le premier : $b_{1k_1} \neq 0$, mais $b_{2k_1} = 0, \ldots, b_{mk_1} = 0$. Donc

$$0 = a_{k_1} = d_1 b_{1k_1} + d_2 0 + \cdots + d_m 0 = d_1 b_{1k_1}$$

D'où $b_{1k_1} \neq 0$ et donc $d_1 = 0$. Ainsi R est une combinaison linéaire des $R_2, \ldots, R_m$ et donc appartient à l'espace ligne de B'. Puisque R est une ligne quelconque de A', l'espace ligne de A' est contenu dans l'espace ligne de B'. De façon analogue, l'espace ligne de B' est contenu dans l'espace de A'. Donc A' et B' ont le même espace ligne et le théorème est démontré.

4.31. Démontrer le théorème 4.7 : Soit $A = (a_{ij})$ et $B = (b_{ij})$ deux matrices échelonnées réduites par les lignes (e.r.l.). A et B ont le même espace ligne si et seulement si elles ont les mêmes lignes non nulles.

Evidemment, si A et B ont les mêmes lignes non nulles, alors elles ont le même espace ligne. Nous n'avons ainsi plus qu'à démontrer la réciproque.

Supposons que A et B aient le même espace ligne et supposons que $R \neq 0$ soit la $i^{\text{ème}}$ ligne de A. Alors il existe des scalaires $c_1, c_2, \ldots, c_s$ de telle sorte que

$$R = c_1R_1 + c_2R_2 + \cdots + c_sR_s \tag{1}$$

où les R_i sont les lignes non nulles de B. Le théorème est démontré si nous montrons que $R = R_i$, ou $c_i = 1$ mais $c_k = 0$ pour $k \neq i$.

Soit a_{ij_i} l'élément distingué de R, c'est-à-dire le premier élément non nul de R. D'après (*1*) et le problème 4.29,

$$a_{ij_i} = c_1 b_{1j_i} + c_2 b_{2j_i} + \cdots + c_s b_{sj_i} \tag{2}$$

Mais d'après le problème précédent b_{ij_i} est un élément distingué de B et puisque B est réduite ligne, c'est le seul élément non nul dans la $j^{\text{ème}}$ colonne de B. Ainsi d'après (*2*) on obtient $a_{ij_i} = c_i b_{ij_i}$. Cependant $a_{ij_i} = 1$ et $b_{ij_i} = 1$ puisque A et B sont réduites lignes ; d'où $c_i = 1$.

Supposons maintenant $k \neq i$ et b_{kj_k} l'élément distingué de R_k. D'après (*1*) et le problème 4.29,

$$a_{ij_k} = c_1 b_{1j_k} + c_2 b_{2j_k} + \cdots + c_s b_{sj_k} \tag{3}$$

Puisque B est réduite ligne, b_{kj_k} est le seul élément non nul dans la $j_k$ème colonne de B ; ainsi d'après (3) $a_{ij_k} = c_k b_{kj_k}$. De plus, d'après le précédent problème a_{kj_k} est un élément distingué de A et puisque A est réduite ligne, $a_{ij_k} = 0$. Donc $c_k b_{kj_k} = 0$ et puisque $b_{kj_k} = 1$, $c_k = 0$, on a donc $R = R_i$ et le théorème est démontré.

4.32. Déterminer si les matrices suivantes ont le même espace colonne:

$$A = \begin{pmatrix} 1 & 3 & 5 \\ 1 & 4 & 3 \\ 1 & 1 & 9 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 2 & 3 \\ -2 & -3 & -4 \\ 7 & 12 & 17 \end{pmatrix}$$

Observons que A et B ont le même espace colonne si et seulement si les transposées A^t et B^t ont le même espace ligne. Réduisons A^t et B^t à leurs formes e.r.l. :

$$A^t = \begin{pmatrix} 1 & 1 & 1 \\ 3 & 4 & 1 \\ 5 & 3 & 9 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & -2 \\ 0 & -2 & 4 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix}$$

$$B^t = \begin{pmatrix} 1 & -2 & 7 \\ 2 & -3 & 12 \\ 3 & -4 & 17 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & -2 & 7 \\ 0 & 1 & -2 \\ 0 & 2 & -4 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & -2 & 7 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix}$$

Puisque A^t et B^t ont le même espace ligne, A et B ont le même espace colonne.

4.33. Soit R un vecteur ligne et B une matrice pour laquelle RB est définie. Montrer que RB est une combinaison linéaire des lignes de B. De plus, si A est une matrice pour laquelle AB est définie, montrer que l'espace ligne de AB est contenu dans l'espace ligne de B.

Supposons $R = (a_1, a_2, \ldots, a_m)$ et $B = (b_{ij})$. Soient $B_1, \ldots, B_m$ les lignes de B et $B^1, B^2, \ldots, B^m$ ses colonnes. Alors

$$\begin{aligned} RB &= (R \cdot B^1, R \cdot B^2, \ldots, R \cdot B^n) \\ &= (a_1b_{11} + a_2b_{21} + \cdots + a_mb_{m1}, a_1b_{12} + a_2b_{22} + \cdots + a_mb_{m2}, \ldots, a_1b_{1n} + a_2b_{2n} + \cdots + a_mb_{mn}) \\ &= a_1(b_{11}, b_{12}, \ldots, b_{1n}) + a_2(b_{21}, b_{22}, \ldots, b_{2n}) + \cdots + a_m(b_{m1}, b_{m2}, \ldots, b_{mn}) \\ &= a_1B_1 + a_2B_2 + \cdots + a_mB_m \end{aligned}$$

Ainsi RB est une combinaison linéaire des lignes de B, comme nous l'avions affirmé.

D'après le problème 3.27 les lignes de AB sont R_iB où R_i est la $i$ème ligne de A. Donc d'après le résultat précédent chaque ligne de AB est dans l'espace ligne de B. Ainsi l'espace ligne de AB est contenu dans l'espace ligne de B.

SOMMES ET SOMMES DIRECTES

4.34. Soit U et W des sous-espaces de l'espace vectoriel V. Montrer que :

(i) U et W sont contenus dans $U + W$.

(ii) $U + W$ est le plus petit sous-espace de V contenant U et W, c'est-à-dire $U + W$ est engendré par U et W : $U + W = L(U, W)$.

(i) Soit $u \in U$. Par hypothèse W est un sous-espace de V et ainsi $0 \in W$. Donc $u = u + 0 \in U + W$. Donc U est contenu dans $U + W$. De façon analogue W est contenu dans $U + W$.

(ii) Puisque $U + W$ est un sous-espace de V (Théorème 4.8) contenant à la fois U et W, il doit aussi contenir l'espace engendré par U et W: $L(U, W) \subset U + W$.

D'autre part, si $v \in U + W$ alors $v = u + w = 1u + 1w$ où $u \in U$ et $w \in W$; d'où v est une combinaison linéaire des éléments de $u \cup w$ et ainsi appartient à $L(U, W)$. Donc $U + W \subset L(U, W)$.

Les deux relations d'inclusion nous donnent le résultat demandé.

4.35. Supposons que U et W soient des sous-espaces de l'espace vectoriel V et supposons que $\{u_i\}$ engendrent U et les w_j engendrent W. Montrer que $\{u_i, w_j\}$ c'est-à-dire $\{u_i\} \cup \{w_j\}$ engendrent $U + W$.

Soit $v \in U + W$. Alors $v = u + w$, où $u \in U$ et $w \in W$. Puisque $\{u_i\}$ engendrent U, u est une combinaison linéaire des u_i et puisque $\{w_j\}$ engendrent W, w est une combinaison linéaire des w_j.

$$u = a_1 u_{i_1} + a_2 u_{i_2} + \cdots + a_n u_{i_n}, \qquad a_j \in K$$

$$w = b_1 w_{j_1} + b_2 w_{j_2} + \cdots + b_m w_{j_m}, \qquad b_j \in K$$

Ainsi $$v = u + w = a_1 u_{i_1} + a_2 u_{i_2} + \cdots + a_n u_{i_n} + b_1 w_{j_1} + b_2 w_{j_2} + \cdots + b_m w_{j_m}$$

et donc $\{u_i, w_j\}$ engendrent $U + W$.

4.36. Démontrer le théorème 4.9 : L'espace vectoriel V est la somme directe de ses sous-espaces U et W si et seulement si (i) $V = U + W$ et (ii) $U \cap W = \{0\}$.

Supposons $V = U \oplus W$. Alors un vecteur quelconque $v \in V$ peut être écrit uniquement sous la forme $v = u + w$ où $u \in U$ et $w \in W$. Ainsi en particulier $V = U + W$. Supposons maintenant que $v \in U \cap W$. Alors

$$(1)\ v = v + 0 \text{ où } v \in U,\ 0 \in W \text{ et } (2)\ v = 0 + v \text{ où } 0 \in U,\ v \in W.$$

Puisqu'une telle somme pour v doit être unique, $v = 0$. D'où $U \cap W = \{0\}$.

D'autre part, supposons $V = U + W$ et $U \cap W = \{0\}$. Soit $v \in V$. Puisque $V = U + W$ il existe $u \in U$ et $w \in W$ de telle sorte que $v = u + w$. Il nous reste à démontrer qu'une telle somme est unique. Supposons donc que $v = u' + w'$ où $u' \in U$ et $w' \in W$. Alors

$$u + w = u' + w' \qquad \text{d'où} \qquad u - u' = w' - w$$

Mais $u - u' \in U$ et $w' - w \in W$; donc puisque $U \cap W = \{0\}$,

$$u - u' = 0, \quad w' - w = 0 \qquad \text{d'où} \qquad u = u', \quad w = w'$$

Ainsi une telle somme pour $v \in V$ est unique et $V = U \oplus W$.

4.37. Soit U et W deux sous-espaces de $\mathbf{R}^3$ définis par

$$U = \{(a, b, c) : a = b = c\} \qquad \text{et} \qquad W = \{(0, b, c)\}$$

(Remarquons que W est le plan yz). Montrer que $\mathbf{R}^3 = U \oplus W$.

Remarquons d'abord que $U \cap W = \{0\}$, pour $v = (a, b, c) \in U \cap W$ implique que

$$a = b = c \quad \text{et} \quad a = 0 \qquad \text{qui implique} \qquad a = 0,\ b = 0,\ c = 0.$$

donc $v = (0, 0, 0)$.

Nous pouvons donc affirmer que $\mathbf{R}^3 = U + W$. Si $v = (a, b, c) \in \mathbf{R}^3$, alors $v = (a, a, a) + (0, b - a, c - a)$ où $(a, a, a) \in U$ et $(0, b - a, c - a) \in W$. Les deux conditions $U \cap W = \{0\}$ et $\mathbf{R}^3 = U + W$ impliquent $\mathbf{R}^3 = U \oplus W$.

4.38. Soit V l'espace vectoriel des n matrices carrées sur un corps $\mathbf{R}$. Soient U et W les sous-espaces des matrices symétriques et antisymétriques respectivement. Montrer que $V = U \oplus W$. (La matrice M est symétrique si et seulement si $M = M^t$ et antisymétrique si et seulement si $M^t = -M$).

Montrons d'abord que $V = U + W$. Soit A une matrice carrée n arbitraire. Remarquons que

$$A = \tfrac{1}{2}(A + A^t) + \tfrac{1}{2}(A - A^t)$$

Nous affirmons que $\frac{1}{2}(A + A^t) \in U$ et que $\frac{1}{2}(A - A^t) \in W$. Car

$$(\tfrac{1}{2}(A + A^t))^t = \tfrac{1}{2}(A + A^t)^t = \tfrac{1}{2}(A^t + A^{tt}) = \tfrac{1}{2}(A + A^t)$$

c'est-à-dire que $\frac{1}{2}(A + A^t)$ est symétrique. De plus

$$(\tfrac{1}{2}(A - A^t))^t = \tfrac{1}{2}(A - A^t)^t = \tfrac{1}{2}(A^t - A) = -\tfrac{1}{2}(A - A^t)$$

c'est-à-dire que $\frac{1}{2}(A - A^t)$ est antisymétrique.

Montrons ensuite que $U \cap W = \{0\}$. Supposons $M \in U \cap W$. Alors $M = M^t$ et $M^t = -M$ qui implique $M = -M$ ou $M = 0$. Donc $U \cap W = \{0\}$, d'où $V = U \oplus W$.

PROBLEMES SUPPLEMENTAIRES

ESPACES VECTORIELS

4.39. Soit V l'ensemble des suites infinies $(a_1, a_2 \ldots)$ sur un corps K avec une addition dans V et une multiplication scalaire sur V définies par

$$(a_1, a_2, \ldots) + (b_1, b_2, \ldots) = (a_1 + b_1, a_2 + b_2, \ldots)$$
$$k(a_1, a_2, \ldots) = (ka_1, ka_2, \ldots)$$

où a_i, b_j, $k \in K$. Montrer que V est un espace vectoriel sur K.

4.40. Soit V l'ensemble des couples $(a\ ,\ b)$ de nombres réels avec une addition dans V et une multiplication scalaire sur V définies par

$$(a, b) + (c, d) = (a + c, b + d) \qquad \text{et} \qquad k(a, b) = (ka, 0)$$

Montrer que V satisfait tous les axiomes de l'espace vectoriel sauf $[M_4]$: $1u = u$. Donc $[M_4]$ n'est pas une conséquence des autres axiomes.

4.41. Soit V l'ensemble des couples $(a\ ,\ b)$ de nombres réels. Montrer que V n'est pas un espace vectoriel sur $\mathbf{R}$ avec une addition dans V et une multiplication sur V définies par :

(i) $(a, b) + (c, d) = (a + d, b + c)$ et $k(a, b) = (ka, kb)$;

(ii) $(a, b) + (c, d) = (a + c, b + d)$ et $k(a, b) = (a, b)$;

(iii) $(a, b) + (c, d) = (0, 0)$ et $k(a, b) = (ka, kb)$;

(iv) $(a, b) + (c, d) = (ac, bd)$ et $k(a, b) = (ka, kb)$.

4.42. Soit V l'ensemble des couples (z_1, z_2) de nombres complexes. Montrer que V est un espace vectoriel sur le corps des réels $\mathbf{R}$ avec une addition dans V et une multiplication scalaire sur V définies par

$$(z_1, z_2) + (w_1, w_2) = (z_1 + w_1, z_2 + w_2) \qquad \text{and} \qquad k(z_1, z_2) = (kz_1, kz_2)$$

où z_1, z_2, w_1, $w_2 \in \mathbf{C}$ et $k \in \mathbf{R}$.

4.43. Soit V un espace vectoriel sur K, et soit F un sous-corps de K. Montrer que V est aussi un espace vectoriel sur F, où l'addition vectorielle par rapport à F est la même que celle par rapport à K, et où la multiplication scalaire par un élément $k \in F$ est la même que la multiplication par k élément de K.

4.44. Montrer que $[A_4]$, page 63, peut être obtenu à partir des autres axiomes de l'espace vectoriel.

4.45. Soient U et W des espaces vectoriels sur le corps K. Soit V l'ensemble des couples (u, w) où u appartient à U et w à W ; $V = \{(u, w) : u \in U, w \in W\}$. Montrer que V est un espace vectoriel sur K avec l'addition dans V et la multiplication scalaire sur V définies par

$$(u, w) + (u', w') = (u + u', w + w') \qquad \text{et} \qquad k(u, w) = (ku, kw)$$

où u, $u' \in U$, w, $w' \in W$ et $k \in K$. (Cet espace V est appelé la somme directe extérieure de U et W).

SOUS-ESPACES

4.46. Considérons l'espace vectoriel V du problème 4.39, des suites infinies $(a_1, a_2, \ldots)$ sur un corps K. Montrer que W est un sous-espace de V si :

(i) W contient toutes les suites dont la première composante est nulle.

(ii) W contient toutes les suites avec seulement un nombre fini de composantes non nulles.

4.47. Déterminer si oui ou non W est sous-espace de $\mathbf{R}^3$ si W contient les vecteurs $(a, b, c) \in \mathbf{R}^3$ pour lesquels (i) $a = 2b$; (ii) $a \leqslant b \leqslant c$; (iii) $ab = 0$; (iv) $a = b = c$; (v) $a = b^2$; (vi) $k_1 a + k_2 b + k_3 c = 0$ où $k_i \in \mathbf{R}$.

4.48. Soit V un espace vectoriel des matrices carrées n sur le corps K. Montrer que W est un sous-espace de V si W est l'ensemble des matrices (i) antisymétriques $A^t = -A$; (ii) triangulaires supérieures ; (iii) diagonales ; (iv) scalaires.

4.49. Soit $AX = B$ un système non homogène d'équations linéaires à n inconnues sur un corps K. Montrer que l'ensemble des solutions n'est pas un sous-espace de K^n.

4.50. Soit V l'espace vectoriel de toutes les fonctions du corps réel $\mathbf{R}$ vers $\mathbf{R}$. Montrer que W est un sous-espace de V dans chacun des cas suivants :

(i) W est l'ensemble de toutes les fonctions bornées. (Ici $f : \mathbf{R} \to \mathbf{R}$ est bornée s'il existe un $M \in \mathbf{R}$ tel que $|f(x)| \leqslant M, \forall x \in \mathbf{R}$).

(ii) W est l'ensemble de toutes les fonctions paires. (Ici : $f : \mathbf{R} \to \mathbf{R}$ est paire si $f(-x) = f(x), \forall x \in \mathbf{R}$).

(iii) W est l'ensemble de toutes les fonctions continues.

(iv) W est l'ensemble de toutes les fonctions différentiables.

(v) W est l'ensemble de toutes les fonctions intégrables, c'est-à-dire ici intégrables sur $0 \leqslant x \leqslant 1$.

(Les trois derniers cas demandent la connaissance d'un peu d'analyse)

4.51. Discuter si oui ou non $\mathbf{R}^2$ est un sous-espace de $\mathbf{R}^3$.

4.52. Démontrer le théorème 4.4 : L'intersection d'un nombre quelconque de sous-espaces d'un espace vectoriel V est un sous-espace de V.

4.53. Supposons U et W deux sous-espaces de l'espace vectoriel V, pour lesquels $U \cup W$ est aussi un sous-espace. Montrer que l'on a soit $U \subset W$ soit $W \subset U$.

COMBINAISONS LINEAIRES

4.54. Considérons les vecteurs $u = (1, -3, 2)$ et $v = (2, -1, 1)$ de $\mathbf{R}^3$.

(i) Ecrire $(1, 7, -4)$ comme une combinaison linéaire de u et v.

(ii) Ecrire $(2, -5, 4)$ comme combinaison linéaire de u et de v.

(iii) Pour quelle valeur de k le vecteur $(1, k, 5)$ est-il une combinaison linéaire de u et de v ?

(iv) Trouver une condition liant a, b, c de telle sorte que (a, b, c) soit une combinaison linéaire de u et de v.

4.55. Ecrire u comme une combinaison linéaire des polynômes $v = 2t^2 + 3t - 4$ et $w = t^2 - 2t - 3$ où (i) $u = 3t^2 + 8t - 5$; (ii) $u = 4t^2 - 6t - 1$.

4.56. Ecrire E comme une combinaison linéaire de $A = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}$ et $C = \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}$

où : (i) $E = \begin{pmatrix} 3 & -1 \\ 1 & -2 \end{pmatrix}$; (ii) $E = \begin{pmatrix} 2 & 1 \\ -1 & -2 \end{pmatrix}$.

GENERATEURS

4.57. Montrer que $(1, 1, 1)$, $(0, 1, 1)$ et $(0, 1, -1)$ engendrent $\mathbf{R}^3$, c'est-à-dire qu'un vecteur quelconque (a, b, c) est une combinaison linéaire des vecteurs donnés.

4.58. Montrer que le plan yz $W = \{(0, b, c)\}$ de $\mathbf{R}^3$ est engendré par (i) $(0, 1, 1)$ et $(0, 2, -1)$; (ii) $(0, 1, 2)$, $(0, 2, 3)$ et $(0, 3, 1)$.

4.59. Montrer que les nombres complexes $w = 2 + 3i$ et $z = 1 - 2i$ engendrent le corps complexe $\mathbf{C}$ considéré comme un espace vectoriel sur le corps des réels $\mathbf{R}$.

4.60. Montrer que les polynômes $(1 - t)^3$, $(1 - t)^2$, $1 - t$ et 1 engendrent l'espace des polynômes de degré $\leqslant 3$.

4.61. Trouver un vecteur de $\mathbf{R}^3$ qui engendre l'intersection de U et W où U est le plan xy : $U = \{(a, b, 0)\}$ et W est l'espace engendré par les vecteurs $(1, 2, 3)$ et $(1, -1, 1)$.

4.62. Démontrer que $L(S)$ est l'intersection de tous les sous-espaces de V contenant S.

4.63. Montrer que $L(S) = L(S \cup \{0\}$. Ce qui est équivalent à dire en adjoignant ou en supprimant d'un ensemble le vecteur nul, nous ne changeons pas l'espace engendré par l'ensemble.

4.64. Montrer que si $S \subset T$, alors $L(S) \subset L(T)$.

4.65. Montrer que $L(L(S)) = L(S)$.

ESPACE LIGNE D'UNE MATRICE

4.66. Déterminer lesquelles parmi les matrices suivantes ont le même espace ligne :

$$A = \begin{pmatrix} 1 & -2 & -1 \\ 3 & -4 & 5 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & -1 & 2 \\ 2 & 3 & -1 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & -1 & 3 \\ 2 & -1 & 10 \\ 3 & -5 & 1 \end{pmatrix}$$

4.67. Soient

$$u_1 = (1, 1, -1), \quad u_2 = (2, 3, -1), \quad u_3 = (3, 1, -5)$$
$$v_1 = (1, -1, -3), \quad v_2 = (3, -2, -8), \quad v_3 = (2, 1, -3)$$

Montrer que le sous-espace de $\mathbf{R}^3$ engendré par les u_i est le même que le sous-espace engendré par les v_i.

4.68. Montrer que si une ligne quelconque d'une matrice échelonnée (resp. e.r.l.) est supprimée, alors la matrice obtenue est toujours du même type.

4.69. Démontrer la réciproque du théorème 4.6: Des matrices ayant le même espace ligne (et les mêmes dimensions) sont équivalentes lignes.

4.70. Montrer que A et B ont le même espace colonne si et seulement si A^t et B^t ont le même espace ligne.

4.71. Soient A et B des matrices pour lesquelles le produit AB est défini. Montrer que l'espace colonne de AB est contenu dans l'espace colonne de A.

SOMMES ET SOMMES DIRECTES

4.72. Etendons la notion de somme à des sous-ensembles arbitraires non vides (non nécessairement des sous-espaces) S et T d'un espace vectoriel V en définissant $S + T = \{s + t \; ; s \in S, \; t \in T\}$. Montrer que cette opération vérifie les propriétés suivantes :

(i) la loi est commutative: $S + T = T + S$;

(ii) la loi est associative: $(S_1 + S_2) + S_3 = S_1 + (S_2 + S_3)$;

(iii) $S + \{0\} = \{0\} + S = S$;

(iv) $S + V = V + S = V$.

4.73. Montrer que pour un sous-espace quelconque W d'un espace vectoriel V, $W + W = W$

4.74. Donner un exemple d'un sous-ensemble S d'un espace vectoriel V, qui n'est pas un sous-espace de V, mais pour lequel (i) $S + S = S$, (ii) $S + S \subset S$ (inclus strictement).

4.75. Etendons la notion de somme de deux sous-espaces à plus de deux sous-espaces comme il suit. Si $W_1, W_2, \ldots, W_n$ sont des sous-espaces de V alors

$$W_1 + W_2 + \cdots + W_n = \{w_1 + w_2 + \cdots + w_n : \; w_i \in W_i\}$$

Montrer que :

(i) $L(W_1, W_2, \ldots, W_n) = W_1 + W_2 + \cdots + W_n$;

(ii) Si S_i engendre W_i, $i = 1, \ldots, n$, alors $S_1 \cup S_2 \cup \ldots \cup S_n$ engendre $W_1 + W_2 + \ldots + W_n$.

4.76. Supposons que U, V et W soient des sous-espaces d'un espace vectoriel. Démontrer que

$$(U \cap V) + (U \cap W) \subset U \cap (V + W)$$

Trouver des sous-espaces de $\mathbf{R}^2$, pour lesquels l'égalité n'est pas vérifiée.

4.77. Soient U, V et W les sous-espaces suivants de $\mathbf{R}^3$:

$$U = \{(a,b,c): a+b+c=0\}, \quad V = \{(a,b,c): a=c\}, \quad W = \{(0,0,c): c \in \mathbf{R}\}$$

Montrer que (i) $\mathbf{R}^3 = U + V$, (ii) $\mathbf{R}^3 = U + W$, (iii) $\mathbf{R}^3 = V + W$. Quand la somme est-elle directe ?

4.78. Soit V l'espace vectoriel de toutes les fonctions du corps réel $\mathbf{R}$ dans $\mathbf{R}$. Soit U le sous-espace des fonctions paires et W le sous-espace des fonctions impaires. Montrer que $V = U \oplus W$. (Rappelons que f est paire si et seulement si $f(-x) = f(x)$ et f est impaire si et seulement si $f(-x) = -f(x)$).

4.79. Soient $W_1, W_2, \ldots$ des sous-espaces d'un espace vectoriel V pour lesquels $W_1 \subset W_2 \subset \ldots$. Soit $W = W_1 \cup W_2 \cup \ldots$. Montrer que W est un sous-espace de V.

4.80. Dans le précédent problème supposons que S_i engendre W_i, $i = 1, 2 \ldots$ Montrer que $S = S_1 \cup S_2 \cup \ldots$ engendre W.

4.81. Soit V l'espace vectoriel des n-matrices carrées sur un corps K. Soit U un sous-espace de V formé des matrices triangulaires supérieures et W le sous-espace des matrices triangulaires inférieures. Trouver (i) $U + W$, (ii) $U \cap W$

4.82. Soit V la somme directe extérieure des espaces vectoriels U et W sur un corps K (voir problème 4.45). Soit

$$\hat{U} = \{(u,0): u \in U\}, \quad \hat{W} = \{(0,w): w \in W\}$$

Montrer que (i) $\hat{U}$ et $\hat{W}$ sont des sous-espaces de V, (ii) $V = \hat{U} \oplus \hat{W}$.

REPONSES DES PROBLEMES SUPPLEMENTAIRES

4.47. (i) Oui.

(ii) Non : exemple $(1, 2, 3) \in W$ mais $-2(1, 2, 3) \notin W$.

(iii) Non : exemple $(1, 0, 0), (0, 1, 0) \in W$ mais non leur somme.

(iv) Oui.

(v) Non : exemple $(9, 3, 0) \in W$ mais $2(9, 3, 0) \notin W$.

(vi) Oui.

4.50. (i) Soient $f, g \in W$ avec M_f et M_g bornées par f et g respectivement. Alors quels que soient les scalaires $a, b \in \mathbf{R}$,

$$|(af+bg)(x)| = |af(x)+bg(x)| \leq |af(x)| + |bg(x)| = |a|\,|f(x)| + |b|\,|g(x)| \leq |a|M_f + |b|M_g$$

C'est-à-dire, $|a|\,M_f + |b|\,M_g$ est borné par la fonction $af + bg$.

(ii) $$(af+bg)(-x) = af(-x) + bg(-x) = af(x) + bg(x) = (af+bg)(x)$$

4.51. Non. Bien qu'un seul permette d'identifier le vecteur $(a, b) \in \mathbf{R}^2$, par exemple $(a, b, 0)$ dans le plan xy de $\mathbf{R}^3$, ce sont des éléments distincts appartenant à des ensembles disjoints.

4.54. (i) $-3u+2v$. (ii) Impossible. (iii) $k = -8$. (iv) $a - 3b - 5c = 0$.

4.55. (i) $u = 2v - w$. (ii) Impossible.

4.56. (i) $E = 2A - B + 2C$. (ii) Impossible.

4.61. $(2, -5, 0)$.

4.66. A et C.

4.67. Former la matrice A dont les lignes sont les u_i et la matrice B dont les lignes sont les v_i, et montrer alors que A et B ont les mêmes formes canoniques lignes.

4.74. (i) Dans $\mathbf{R}^2$, soit $S = \{(0, 0), (0, 1), (0, 2), (0, 3), \ldots\}$.

(ii) Dans $\mathbf{R}^2$, soit $S = \{(0, 5)\ (0, 6), (0, 7) \ldots\}$.

4.77. La somme est directe dans ii) et iii).

4.78. $f(x) = \frac{1}{2} f(x) + f(-x)) + \frac{1}{2}(f(x) - f(-x))$, où $\frac{1}{2}(f(x) + f(-x))$ est paire et $\frac{1}{2}(f(x) - f(-x))$ est impaire.

4.81. (i) $V = U + W$. (ii) $U \cap W$ est l'espace des matrices diagonales.

CHAPITRE 5

Base et dimension

INTRODUCTION

Quelques-uns des résultats fondamentaux établis dans ce chapitre sont :

(1) On peut définir une notion de "dimension" d'un espace vectoriel.

(2) Si V a pour dimension n sur K, alors V est "isomorphe" à K^n (Théorème 5.12).

(3) Un système d'équations linéaires a une solution si et seulement si la matrice du système et la matrice augmentée ont le même "rang" (Théorème 5.10).

Ces concepts et ces résultats ne sont pas triviaux et répondent à certaines questions posées et étudiées par les mathématiciens de la fin du siècle dernier et du début de ce siècle.

Nous commencerons ce chapitre par la définition de la dépendance et de l'indépendance linéaires. Cette définition joue un rôle essentiel dans la théorie de l'algèbre linéaire et en mathématiques en général.

DEPENDANCE LINEAIRE

Définition : Soit V un espace vectoriel sur le corps K. Les vecteurs $v_1, \ldots, v_m \in V$ sont dits linéairement dépendants sur K, ou simplement dépendants s'il existe des scalaires $a_1, \ldots, a_m \in K$ non tous nuls, tels que :

$$a_1v_1 + a_2v_2 + \cdots + a_mv_m = 0 \qquad (*)$$

Sinon les vecteurs sont dits linéairement indépendants sur K, ou simplement indépendants.

Remarquons que la relation (*) est toujours vérifiée si les a_s sont tous nuls. Si cette relation est vérifiée seulement dans ce cas, c'est-à-dire

$$a_1v_1 + a_2v_2 + \cdots + a_mv_m = 0 \text{ seulement si } a_1 = 0, \ \ldots, \ a_m = 0$$

les vecteurs sont linéairement indépendants. Par contre si la relation (*) est vérifiée lorsque l'un des a_s n'est pas nul, alors les vecteurs sont linéairement dépendants.

Remarquons que si 0 est l'un des vecteurs $v_1, v_2, \ldots, v_m$, par exemple $v_1 = 0$, alors les vecteurs doivent être dépendants car :

$$1v_1 + 0v_2 + \cdots + 0v_m = 1 \cdot 0 + 0 + \cdots + 0 = 0$$

le coefficient de v_1 n'étant pas nul. D'autre part un vecteur quelconque v non nul est par lui-même indépendant :

$$kv = 0, \ v \neq 0, \text{ implique } k = 0.$$

D'autres exemples de vecteurs dépendants et indépendants suivent :

Exemple 5.1 : Les vecteurs $u = (1, -1, 0)$, $v = (1, 3, -1)$ et $w = (5, 3, -2)$ sont dépendants puisque $3u + 2v - w = 0$,

$$3(1, -1, 0) + 2(1, 3, -1) - (5, 3, -2) = (0, 0, 0)$$

Exemple 5.2 : Montrons que les vecteurs $u = (6, 2, 3, 4)$, $v = (0, 5, -3, 1)$ et $w = (0, 0, 7, -2)$ sont indépendants. Supposons que $xu + yv + zw = 0$ où x, y et z sont des scalaires inconnus. Alors

$$\begin{aligned}(0, 0, 0, 0) &= x(6, 2, 3, 4) + y(0, 5, -3, 1) + z(0, 0, 7, -2)\\ &= (6x, 2x+5y, 3x-3y+7z, 4x+y-2z)\end{aligned}$$

et ainsi en égalant les composantes correspondantes

$$\begin{aligned}6x \qquad\qquad &= 0\\ 2x + 5y \qquad &= 0\\ 3x - 3y + 7z &= 0\\ 4x + \ \ y - 2z &= 0\end{aligned}$$

La première équation donne $x = 0$; la seconde équation avec $x = 0$ donne $y = 0$; et la troisième équation avec $x = 0$ et $y = 0$ donne $z = 0$. Ainsi

$$xu + yv + zw = 0 \quad \text{implique} \quad x = 0,\ y = 0,\ z = 0$$

Les vecteurs u, v, w sont alors indépendants.

Remarquons que les vecteurs dans l'exemple précédent forme une matrice mise sous la forme échelonnée.

$$\begin{pmatrix} 6 & 2 & 3 & 4\\ 0 & 5 & -3 & 1\\ 0 & 0 & 7 & -2\end{pmatrix}$$

Ainsi, nous avons montré que les lignes (non nulles) de la matrice précédente sous la forme échelonnée sont indépendantes. Ce résultat reste vrai en général ; nous l'énonçons donc sous la forme d'un théorème à cause de son usage fréquent.

Théorème 5.1 : Les lignes non nulles d'une matrice sous sa forme échelonnée sont linéairement indépendantes.

On peut définir de la manière suivante pour plus d'un vecteur la notion de dépendance linéaire.

Les vecteurs $v_1, v_2, \ldots, v_m$ sont linéairement dépendants si et seulement si l'un d'entre eux est une combinaison linéaire des autres.

Supposons que v_i soit une combinaison linéaire des autres vecteurs :

$$v_i = a_1v_1 + \cdots + a_{i-1}v_{i-1} + a_{i+1}v_{i+1} + \cdots + a_mv_m$$

Ajoutons $-v_i$ aux deux membres, on obtient

$$a_1v_1 + \cdots + a_{i-1}v_{i-1} - v_i + a_{i+1}v_{i+1} + \cdots + a_mv_m = 0$$

où le coefficient de v_i n'est pas nul ; ainsi les vecteurs sont linéairement dépendants. Réciproquement supposons que les vecteurs soient linéairement dépendants, c'est-à-dire

$$b_1v_1 + \cdots + b_jv_j + \cdots + b_mv_m = 0 \qquad \text{où} \qquad b_j \neq 0$$

Alors $$v_j = -b_j^{-1}b_1v_1 - \cdots - b_j^{-1}b_{j-1}v_{j-1} - b_j^{-1}b_{j+1}v_{j+1} - \cdots - b_j^{-1}b_mv_m$$

et ainsi v_j est une combinaison linéaire des autres vecteurs.

Indiquons maintenant un lemme plus important que celui qui précède et qui a des conséquences multiples et importantes.

Lemme 5.2 : Les vecteurs non nuls $v_1, \ldots, v_m$ sont linéairement dépendants si et seulement si l'un d'entre eux, v_i, est une combinaison linéaire des vecteurs précédents :

$$v_i = k_1v_1 + k_2v_2 + \cdots + k_{i-1}v_{i-1}$$

Remarque 1 : L'ensemble $\{v_1, v_2, \ldots, v_m\}$ est dit un ensemble dépendant ou indépendant suivant que les vecteurs $v_1, v_2, \ldots, v_m$ sont dépendants ou indépendants. Nous définissons aussi l'ensemble vide $\emptyset$ comme ensemble indépendant.

Remarque 2 : Si deux des vecteurs $v_1, \ldots, v_m$ sont égaux, par exemple $v_1 = v_2$, alors les vecteurs sont dépendants. En effet

$$v_1 - v_2 + 0v_3 + \cdots + 0v_m = 0$$

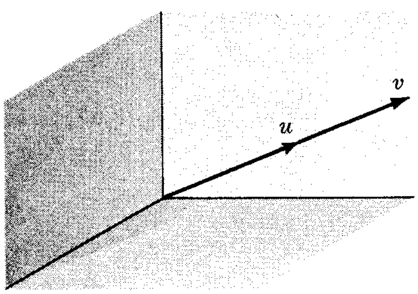
et le coefficient de v_1 n'est pas nul.

Remarque 3 : Deux vecteurs v_1 et v_2 sont dépendants si et seulement si l'un d'entre eux est un multiple de l'autre.

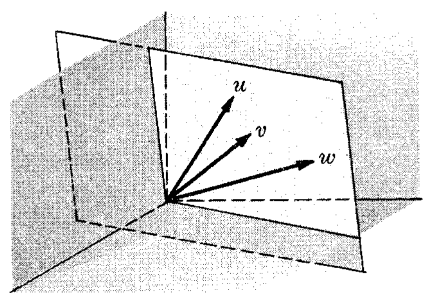
Remarque 4 : Un ensemble contenant un sous-ensemble dépendant est lui-même dépendant. Par suite un sous-ensemble quelconque d'un ensemble indépendant est indépendant.

Remarque 5 : Si l'ensemble $\{v_1, v_2, \ldots, v_m\}$ est indépendant, une permutation quelconque des vecteurs $\{v_{i_1}, v_{i_2}, \ldots, v_{i_m}\}$ est aussi indépendante.

Remarque 6 : Dans l'espace réel $\mathbf{R}^3$, la dépendance des vecteurs peut être représentée géométriquement de la manière suivante : deux vecteurs quelconques u et v sont dépendants si et seulement s'ils appartiennent à la même droite issue de l'origine ; trois vecteurs quelconques u, v et w sont dépendants si et seulement s'ils appartiennent au même plan passant par l'origine.



u et *v* sont dépendants



u, *v*, *w* sont dépendants.

BASE ET DIMENSION

Commençons par la définition.

Définition : Un espace vectoriel V est dit de dimension finie n, ou a pour dimension n, on écrit $\dim V = n$, s'il existe n vecteurs linéairement indépendants $e_1, e_2, e_3, \ldots, e_n$ qui engendrent V. La suite $\{e_1, e_2, \ldots, e_n\}$ est appelée une base de V.

La définition précédente de la dimension d'un espace vectoriel a un sens par suite du théorème suivant :

Théorème 5.3 : Soit V un espace vectoriel de dimension finie. Chaque base de V à alors le même nombre d'éléments.

L'espace vectoriel $\{0\}$ a par définition la dimension 0. (Dans un certain sens ceci est en accord avec la définition précédente, puisque par définition $\emptyset$ est indépendant et engendre $\{0\}$). Lorsqu'un espace vectoriel n'est pas de dimension finie, on dit qu'il s'agit d'un espace vectoriel de dimension infinie.

Exemple 5.3 : Soit K un corps quelconque. Considérons l'espace vectoriel K^n qui contient tous les n-tuples formés d'éléments de K. Les vecteurs

$$\begin{aligned} e_1 &= (1, 0, 0, \ldots, 0, 0) \\ e_2 &= (0, 1, 0, \ldots, 0, 0) \\ &\cdots\cdots\cdots\cdots\cdots \\ e_n &= (0, 0, 0, \ldots, 0, 1) \end{aligned}$$

forment une base, appelée la base usuelle ou base canonique de K^n. Ainsi K^n a pour dimension n.

Exemple 5.4 : Soit U l'espace vectoriel de toutes les matrices 2×3 sur un corps K. Les matrices

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

forment une base de U. Ainsi $\dim U = 6$. Plus généralement, soit V l'espace vectoriel de toutes les matrices $m \times n$ sur K et soit $E_{ij} \in V$ la matrice dont le ij élément est 1 et 0 partout ailleurs. Alors l'ensemble $\{E_{ij}\}$ est une base, appelée base usuelle de V (Problème 5.32) ; en conséquence $\dim V = mn$

Exemple 5.5 : Soit W l'espace vectoriel des polynômes en t de degré $\leqslant n$. L'ensemble $\{1, t, t^2, \ldots, t^n\}$ est linéairement indépendant et engendre W. Ainsi il s'agit d'une base de W et donc $\dim W = n + 1$.

L'espace vectoriel V de tous les polynômes n'est pas de dimension finie puisque (problème 4.26) aucun ensemble fini de polynômes n'engendre V.

Le précédent théorème fondamental sur la dimension est une conséquence de l'important lemme suivant :

Lemme 5.4 : Supposons que l'ensemble $\{v_1, v_2, \ldots, v_n\}$ engendre un espace vectoriel V. Si $\{w_1, \ldots, w_m\}$ est linéairement indépendant, alors $m \leqslant n$ et V est engendré par un ensemble de la forme

$$\{w_1, \ldots, w_m, v_{i_1}, \ldots, v_{i_{n-m}}\}$$

Ainsi, en particulier, $n + 1$ ou un nombre plus grand de vecteurs de V sont linéairement dépendants.

Remarquons que dans le lemme précédent nous avons remplacé m vecteurs dans l'ensemble engendrant l'espace vectoriel par m vecteurs indépendants et que nous avons obtenu toujours un ensemble engendrant l'espace vectoriel.

Supposons maintenant que S soit un sous-ensemble d'un espace vectoriel V. On dit que $\{v_1, \ldots, v_n\}$ est un sous-ensemble indépendant maximal de S si :

(i) c'est un sous-ensemble indépendant de S et

(ii) $\{v_1, \ldots, v_m, w\}$ est dépendant pour tout $w \in S$.

On a alors le théorème suivant :

Théorème 5.5 : Supposons que S engendre V et que $\{v_1, \ldots, v_m\}$ soit un sous-ensemble indépendant maximal de S. Alors $\{v_1, \ldots, v_m\}$ est une base de V.

La relation principale entre la dimension d'un espace vectoriel et ses sous-ensembles indépendants, est contenue dans le théorème suivant.

Théorème 5.6 : Soit V un espace vectoriel de dimension finie n. Alors

(i) Un ensemble quelconque de $n + 1$ ou d'un nombre supérieur de vecteurs est dépendant.

(ii) Un ensemble quelconque linéairement indépendant est une partie d'une base, c'est-à-dire peut être prolongé à une base.

(iii) Un ensemble linéairement indépendant de n éléments est une base.

Exemple 5.6 : Les quatre vecteurs de K^4

$$(1, 1, 1, 1), \quad (0, 1, 1, 1), \quad (0, 0, 1, 1), \quad (0, 0, 0, 1)$$

sont linéairement indépendants puisqu'ils forment une matrice sous forme échelonnée. De plus puisque $\dim K^4 = 4$, ils forment une base de K^4.

Exemple 5.7 : Les quatre vecteurs de $\mathbf{R}^3$

$$(257, -132, 58), \quad (43, 0, -17), \quad (521, -317, 94), \quad (328, -512, -731)$$

doivent être linéairement dépendants puisqu'ils sont extraits d'un espace vectoriel de dimension 3.

DIMENSION ET SOUS-ESPACES

Les théorèmes suivants établissent des relations de base importantes entre la dimension d'un espace vectoriel et la dimension d'un sous-espace.

Théorème 5.7 : Soit W un sous-espace d'un espace vectoriel V de dimension n. Alors $\dim W \leqslant n$. En particulier si $\dim W = n$, alors $W = V$.

Exemple 5.8 : Soit W un sous-espace de l'espace réel $\mathbf{R}^3$. Donc $\dim \mathbf{R}^3 = 3$; d'où d'après le théorème précédent la dimension de W peut être seulement 0, 1, 2 ou 3. Il s'ensuit les cas :

(i) $\dim W = 0$, alors $W = \{0\}$ est réduit au seul vecteur nul c'est-à-dire à l'origine.

(ii) $\dim W = 1$, alors W est une droite issue de l'origine.

(iii) $\dim W = 2$, W est alors un plan passant par l'origine.

(iv) $\dim W = 3$, W est alors l'espace entier $\mathbf{R}^3$.

Théorème 5.8 : Soient U et W des sous-espaces vectoriels de dimension finie d'un espace vectoriel V. Alors $U + W$ est de dimension finie et

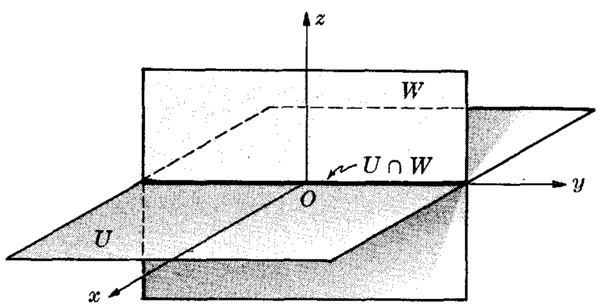
$$\dim (U + V) = \dim U + \dim W - \dim (U \cap W)$$

Remarquons que si V est la somme directe de U et W, c'est-à-dire $V = U \oplus W$, on a $\dim V = \dim U + \dim W$ (Problème 5.48).

Exemple 5.9 : Supposons que U et W soient respectivement le plan xy et le plan yz de $\mathbf{R}^3$: $U = \{(a, b, 0)\}$, $W = \{(0, b, c)\}$. Puisque $\mathbf{R}^3 = U + W$, $\dim (U + W) = 3$. On a aussi $\dim U = 2$ et $\dim W = 2$. D'après le théorème précédent,

$3 = 2 + 2 - \dim (U \cap W)$ c'est-à-dire $(U \cap W) = 1$.

Remarquons que ceci est en accord avec le fait que $U \cap W$ est l'axe des y. $U \cap W = \{(0, b, 0)\}$ et donc a pour dimension 1.



RANG D'UNE MATRICE

Soit A une matrice $m \times n$ quelconque sur un corps K. Rappelons que l'espace ligne de A est le sous-espace de K^n engendré par les lignes de A, et l'espace colonne de A est le sous-espace de K^m engendré par ses colonnes. Les dimensions de l'espace ligne et de l'espace colonne de A sont appelées, respectivement, le rang ligne et le rang colonne de la matrice A.

Théorème 5.9 : Le rang ligne et le rang colonne d'une matrice A sont égaux.

Définition : Le rang de la matrice A, que l'on écrit rang (A), est la valeur commune du rang ligne et du rang colonne de A.

Ainsi le rang d'une matrice donne le nombre maximum de lignes indépendantes ainsi que le nombre maximum de colonnes indépendantes. Nous pouvons obtenir le rang d'une matrice comme suit :

Supposons $A = \begin{pmatrix} 1 & 2 & 0 & -1 \\ 2 & 6 & -3 & -3 \\ 3 & 10 & -6 & -5 \end{pmatrix}$. Réduisons A à sa forme échelonnée, en utilisant des opérations élémentaires sur les lignes de la matrice.

$$A \text{ d'où } \begin{pmatrix} 1 & 2 & 0 & -1 \\ 0 & 2 & -3 & -1 \\ 0 & 4 & -6 & -2 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & 2 & 0 & -1 \\ 0 & 2 & -3 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Rappelons que les matrices équivalentes lignes ont le même espace ligne. Ainsi les lignes non nulles d'une matrice échelonnée, qui sont indépendantes d'après le théorème 5.1, forment une base de l'espace ligne de A. Donc le rang de A est 2.

APPLICATIONS AUX EQUATIONS LINEAIRES

Considérons un système de m équations linéaires à n inconnues $x_1, x_2, \ldots, x_n$ sur un corps K :

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

ou l'équation matricielle équivalente

$$AX = B$$

où $A = (a_{ij})$ est la matrice des coefficients et $X = (x_i)$ et $B = (b_i)$ sont les vecteurs colonnes représentant respectivement les inconnues et les constantes. Rappelons que la matrice complète ou augmentée de ce système est définie par la matrice

$$(A, B) = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{pmatrix}$$

Remarque 1 : Les équations linéaires précédentes sont dites dépendantes ou indépendantes suivant que les vecteurs correspondants, c'est-à-dire les lignes de la matrice complète, sont dépendantes ou indépendantes.

Remarque 2 : Deux systèmes d'équations linéaires sont équivalents si et seulement si leurs matrices complètes correspondantes sont équivalentes lignes, c'est-à-dire ont le même espace ligne.

Remarque 3 : Nous pouvons toujours remplacer un système d'équations par un système d'équations indépendantes, de telle manière que l'on ait un système sous forme échelonnée. Le nombre d'équations indépendantes sera toujours égal au rang de la matrice complète.

Remarquons que le système précédent est aussi équivalent à l'équation vectorielle

$$x_1\begin{pmatrix} a_{11} \\ a_{21} \\ \cdot\cdot \\ a_{m1} \end{pmatrix} + x_2\begin{pmatrix} a_{12} \\ a_{22} \\ \cdot\cdot \\ a_{m2} \end{pmatrix} + \cdots + x_n\begin{pmatrix} a_{1n} \\ a_{2n} \\ \cdot\cdot \\ a_{mn} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \cdot\cdot \\ b_m \end{pmatrix}$$

Ainsi le système $AX = B$ a une solution si et seulement si le vecteur colonne B est une combinaison linéaire des colonnes de la matrice A ; c'est-à-dire appartient à l'espace colonne de A. Ceci nous donne le théorème d'existence fondamental suivant :

Théorème 5.10 : Le système d'équations linéaires $AX = B$ a une solution si et seulement si la matrice des coefficients A et la matrice complète (A , B) ont le même rang.

Rappelons (théorème 2.1) que si le système $AX = B$ a une solution, par exemple v, alors sa solution générale est de la forme $v + W = \{v + w : w \in W\}$ où W est la solution générale du système homogène associé $AX = 0$. Maintenant W est un sous-espace de K^n et donc a une dimension. Le théorème suivant dont nous donnerons la démonstration dans le chapitre suivant (page 127) s'applique.

Théorème 5.11 : La dimension de l'espace solution W d'un système homogène d'équations linéaires $AX = 0$ est $n - r$ où n est le nombre d'inconnues et r le rang de la matrice des coefficients A.

Dans le cas où le système $AX = 0$ est sous sa forme échelonnée, alors il a précisément $n - r$ variables libres (voir page 21), par exemple $x_{i_1}, x_{i_2}, \ldots, x_{i_{n-r}}$. Les solutions $v_1, \ldots, v_{n-r}$ sont linéairement indépendantes (Problème 5-43) et donc forment une base de l'espace solution.

Exemple 5.10 : Trouver la dimension et une base de l'espace solution W du système d'équations linéaires.

$$\begin{aligned} x + 2y - 4z + 3r - s &= 0 \\ x + 2y - 2z + 2r + s &= 0 \\ 2x + 4y - 2z + 3r + 4s &= 0 \end{aligned}$$

Réduisons le système à sa forme échelonnée :

$$\begin{aligned} x + 2y - 4z + 3r - s &= 0 \\ 2z - r + 2s &= 0 \\ 6z - 3r + 6s &= 0 \end{aligned} \qquad \text{et alors} \qquad \begin{aligned} x + 2y - 4z + 3r - s &= 0 \\ 2z - r + 2s &= 0 \end{aligned}$$

Il y a 5 inconnues et 2 équations (non nulles) dans la forme échelonnée ; donc $\dim W = 5 - 2 = 3$. Remarquons que les inconnues libres sont y, r et s. Donc

$$\text{(i)}\ y = 1,\ r = 0,\ s = 0, \quad \text{(ii)}\ y = 0,\ r = 1,\ s = 0, \quad \text{(iii)}\ y = 0,\ r = 0,\ s = 1$$

qui permet d'obtenir les solutions suivantes :

$$v_1 = (-2, 1, 0, 0, 0), \quad v_2 = (-1, 0, \tfrac{1}{2}, 1, 0), \quad v_3 = (-3, 0, -1, 0, 1)$$

L'ensemble $\{v_1, v_2, v_3\}$ est une base de l'espace solution W.

COORDONNEES

Soit $\{e_1, e_2, \ldots, e_n\}$ une base d'un espace vectoriel V de dimension n sur un corps K, et soit v un vecteur quelconque de V. Puisque $\{e_i\}$ engendre V, v est une combinaison linéaire des e_i ;

$$v = a_1e_1 + a_2e_2 + \cdots + a_ne_n, \qquad a_i \in K$$

Les e_i étant indépendants, une telle représentation est unique (Problème 5-7) c'est-à-dire que les n scalaires $a_1, \ldots, a_n$ sont complètement déterminés par le vecteur v et la base $\{e_i\}$. Nous appelons ces scalaires les coordonnées de v sur la base $\{e_i\}$, et nous appelons le n-tuple $(a_1, a_2, \ldots, a_n)$ le vecteur-coordonnée de v relatif à la base $\{e_i\}$; on note $[v]_e$ ou plus simplement $[v]$ lorsqu'il n'y a pas d'ambiguïté :

$$[v]_e = (a_1, a_2, \ldots, a_n)$$

Exemple 5.11 : Soit V l'espace vectoriel des polynômes de degré $\leqslant 2$:

$$V = \{at^2 + bt + c : a, b, c \in \mathbf{R}\}$$

Les polynômes

$$e_1 = 1, \quad e_2 = t - 1 \quad \text{et} \quad e_3 = (t-1)^2 = t^2 - 2t + 1$$

forment une base de V. Soit $v = 2t^2 - 5t + 6$. Trouver $[v]_e$, les coordonnées du vecteur v relatives à la base $\{e_1, e_2, e_3\}$.

Ecrivons v comme une combinaison linéaire des e_i en utilisant les inconnues x, y et z : $v = xe_1 + ye_2 + ze_3$.

$$\begin{aligned} 2t^2 - 5t + 6 &= x(1) + y(t-1) + z(t^2 - 2t + 1) \\ &= x + yt - y + zt^2 - 2zt + z \\ &= zt^2 + (y - 2z)t + (x - y + z) \end{aligned}$$

Egalons les coefficients des mêmes puissances de t les uns aux autres :

$$\begin{aligned} x - y + z &= 6 \\ y - 2z &= -5 \\ z &= 2 \end{aligned}$$

La solution du système précédent est $x = 3$, $y = -1$, $z = 2$. Ainsi

$$v = 3e_1 - e_2 + 2e_3, \quad \text{et donc} \quad [v]_e = (3, -1, 2)$$

Exemple 5.12 : Considérons l'espace réel $\mathbf{R}^3$. Trouver les coordonnées du vecteur $v = (3, 1, -4)$ relativement à la base $f_1 = (1, 1, 1)$, $f_2 = (0, 1, 1)$, $f_3 = (0, 0, 1)$.

Ecrivons v comme combinaison linéaire des f_i en utilisant les inconnues x, y et z : $v = xf_1 + yf_2 + zf_3$.

$$\begin{aligned} (3, 1, -4) &= x(1, 1, 1) + y(0, 1, 1) + z(0, 0, 1) \\ &= (x, x, x) + (0, y, y) + (0, 0, z) \\ &= (x, x+y, x+y+z) \end{aligned}$$

En égalant les composantes correspondantes les unes aux autres on obtient le système équivalent d'équations

$$\begin{aligned} x \qquad &= 3 \\ x + y \quad &= 1 \\ x + y + z &= -4 \end{aligned}$$

ayant pour solution $x = 3$, $y = -2$, $z = -5$. Ainsi $[v]_f = (3, -2, -5)$.

Remarquons que relativement à la base usuelle $e_1 = (1, 0, 0), e_2 = (0, 1, 0)$, $e_3 = (0, 0, 1)$, les coordonnées du vecteur v sont identiques à v lui-même, $[v]_e = (3, 1, -4) = v$.

Nous avons montré qu'à chaque vecteur $v \in V$, il correspond relativement à une base donnée $\{e_1, \ldots, e_n\}$ un n-tuple $[v]_e$ de K^n. D'autre part, si $(a_1, a_2, \ldots, a_n) \in K^n$, il existe un vecteur de V de la forme $a_1 e_1 + \ldots + a_n e_n$. Ainsi la base $\{e_i\}$ détermine une correspondance biunivoque entre les vecteurs de V et les n-tuples de K^n. Remarquons aussi que si

$$v = a_1e_1 + \cdots + a_ne_n \quad \text{correspond à} \quad (a_1, \ldots, a_n)$$

et

$$w = b_1e_1 + \cdots + b_ne_n \quad \text{correspond à} \quad (b_1, \ldots, b_n)$$

alors

$v + w = (a_1 + b_1) e_1 + \cdots + (a_n + b_n)e_n$ correspond à $(a_1, \ldots, a_n) + (b_1, \ldots, b_n)$ et pour un scalaire quelconque $k \in K$

$$kv = (ka_1)e_1 + \cdots + (ka_n)e_n \quad \text{correspond à} \quad k(a_1, \ldots, a_n)$$

c'est-à-dire

$$[v + w]_e = [v]_e + [w]_e \quad \text{et} \quad [kv]_e = k[v]_e$$

Ainsi la correspondance biunivoque précédente entre V et K^n conserve les opérations de l'espace vectoriel, addition vectorielle et multiplication par un scalaire ; on dit alors que V et K^n sont isomorphes, on écrit $V \cong K^n$. On énonce formellement ce résultat par :

Théorème 5.12 : Soit V un espace vectoriel de dimension n sur un corps K. Alors V et K^n sont isomorphes.

L'exemple suivant donne une application pratique du résultat.

Exemple 5.13 : Déterminer si les matrices suivantes sont dépendantes ou indépendantes

$$A = \begin{pmatrix} 1 & 2 & -3 \\ 4 & 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 3 & -4 \\ 6 & 5 & 4 \end{pmatrix}, \quad C = \begin{pmatrix} 3 & 8 & -11 \\ 16 & 10 & 9 \end{pmatrix}$$

Les coordonnées des vecteurs des matrices précédentes sont relativement à la base de l'exemple 5.4 page 89 :

$$[A] = (1, 2, -3, 4, 0, 1), \quad [B] = (1, 3, -4, 6, 5, 4), \quad [C] = (3, 8, -11, 16, 10, 9)$$

Formons la matrice M, dont les lignes sont les coordonnées des vecteurs précédents :

$$M = \begin{pmatrix} 1 & 2 & -3 & 4 & 0 & 1 \\ 1 & 3 & -4 & 6 & 5 & 4 \\ 3 & 8 & -11 & 16 & 10 & 9 \end{pmatrix}$$

Réduisons M à sa forme échelonnée :

$$M \quad \text{d'où} \begin{pmatrix} 1 & 2 & -3 & 4 & 0 & 1 \\ 0 & 1 & -1 & 2 & 5 & 3 \\ 0 & 2 & -2 & 4 & 10 & 6 \end{pmatrix} \quad \text{d'où} \begin{pmatrix} 1 & 2 & -3 & 4 & 0 & 1 \\ 0 & 1 & -1 & 2 & 5 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Puisque la matrice échelonnée a seulement deux lignes non nulles, les vecteurs coordonnées $[A]$, $[B]$ et $[C]$ engendrent un espace de dimension 2 et donc sont dépendants. Donc les matrices initiales A, B et C sont dépendantes.

PROBLEMES RESOLUS

DEPENDANCE LINEAIRE

5.1. Déterminer si les vecteurs u et v sont ou non linéairement dépendants avec :

(i) $u = (3,4),\ v = (1,-3)$

(ii) $u = (2,-3),\ v = (6,-9)$

(iii) $u = (4,3,-2),\ v = (2,-6,7)$

(iv) $u = (-4,6,-2),\ v = (2,-3,1)$

(v) $u = \begin{pmatrix} 1 & -2 & 4 \\ 3 & 0 & -1 \end{pmatrix},\ v = \begin{pmatrix} 2 & -4 & 8 \\ 6 & 0 & -2 \end{pmatrix}$

(vi) $u = \begin{pmatrix} 1 & 2 & -3 \\ 6 & -5 & 4 \end{pmatrix},\ v = \begin{pmatrix} 6 & -5 & 4 \\ 1 & 2 & -3 \end{pmatrix}$

(vii) $u = 2 - 5t + 6t^2 - t^3,\ v = 3 + 2t - 4t^2 + 5t^3$

(viii) $u = 1 - 3t + 2t^2 - 3t^3,\ v = -3 + 9t - 6t^2 + 9t^3$

Deux vecteurs u et v sont dépendants si et seulement si l'un des vecteurs est un mulitple de l'autre.

1) non 2) oui : car $v = 3u$ 3) non 4) oui : car $u = -2v$ 5) oui : car $v = 2u$ 6) non 7) non 8) oui : car $v = -3u$.

5.2. Déterminer si oui ou non les vecteurs suivants de $\mathbf{R}^3$ sont linéairement dépendants :

(i) $(1,-2,1), (2,1,-1), (7,-4,1)$

(ii) $(1,-3,7), (2,0,-6), (3,-1,-1), (2,4,-5)$

(iii) $(1,2,-3), (1,-3,2), (2,-1,5)$

(iv) $(2,-3,7), (0,0,0), (3,-1,-4)$

1) **Méthode 1.** Ecrivons qu'une combinaison linéaire quelconque des vecteurs u, v, w est égale au vecteur nul en utilisant les scalaires inconnus x, y, z :

$$x(1, -2, 1) + y(2, 1, -1) + z(7, -4, 1) = (0, 0, 0)$$

Alors $$(x, -2x, x) + (2y, y, -y) + (7z, -4z, z) = (0, 0, 0)$$

ou $$(x + 2y + 7z, -2x + y - 4z, x - y + z) = (0, 0, 0)$$

En égalant les composantes correspondantes les unes aux autres on obtient le système homogène équivalent que l'on réduit à sa forme échelonnée :

$$\begin{aligned} x + 2y + 7z &= 0 \\ -2x + y - 4z &= 0 \\ x - y + z &= 0 \end{aligned} \quad \text{ou} \quad \begin{aligned} x + 2y + 7z &= 0 \\ 5y + 10z &= 0 \\ -3y - 6z &= 0 \end{aligned} \quad \text{ou} \quad \begin{aligned} x + 2y + 7z &= 0 \\ y + 2z &= 0 \end{aligned}$$

Le système réduit à sa forme échelonnée n'a que deux équations non nulles à trois inconnues ; d'où le système admet une solution non nulle. Ainsi les vecteurs unitiaux sont linéairement dépendants.

Méthode 2 : Formons la matrice dont les lignes sont les vecteurs donnés et réduisons-la à sa forme échelonnée en utilisant les opérations élémentaires sur les lignes :

$$\begin{pmatrix} 1 & -2 & 1 \\ 2 & 1 & -1 \\ 7 & -4 & 1 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & -2 & 1 \\ 0 & 5 & -3 \\ 0 & 10 & -6 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & -2 & 1 \\ 0 & 5 & -3 \\ 0 & 0 & 0 \end{pmatrix}$$

Puisque la matrice échelonnée a une ligne nulle, les vecteurs sont dépendants. (Les trois vecteurs donnés engendrent un espace de dimension 2).

2) Oui, puisque quatre vecteurs quelconques (ou plus) de $\mathbf{R}^3$ sont dépendants.

3) Formons la matrice dont les lignes sont les vecteurs donnés et réduisons cette matrice à sa forme échelonnée, grâce à des opérations élémentaires sur les lignes :

$$\begin{pmatrix} 1 & 2 & -3 \\ 1 & -3 & 2 \\ 2 & -1 & 5 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & 2 & -3 \\ 0 & -5 & 5 \\ 0 & -5 & 11 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & 2 & -3 \\ 0 & -5 & 5 \\ 0 & 0 & 6 \end{pmatrix}$$

Etant donné que la matrice échelonnée n'a pas de ligne nulle, les vecteurs sont indépendants. (Les trois vecteurs donnés engendrent un espace de dimension 3).

4) Puisque $0 = (0, 0, 0)$ est l'un des vecteurs, les vecteurs sont dépendants.

5.3. Soit V l'espace vectoriel, des matrices 2×2 sur $\mathbf{R}$. Déterminer si les matrices $A, B, C \in W$ sont dépendantes où :

(i) $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $C = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$

(ii) $A = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 3 & -1 \\ 2 & 2 \end{pmatrix}$, $C = \begin{pmatrix} 1 & -5 \\ -4 & 0 \end{pmatrix}$

1) Ecrivons qu'une combinaison linéaire des matrices A, B et C est égale à la matrice nulle en utilisant les scalaires inconnus x, y et z, c'est-à-dire $xA + yB + zC = 0$. Ainsi

$$x\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + y\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + z\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

ou $$\begin{pmatrix} x & x \\ x & x \end{pmatrix} + \begin{pmatrix} y & 0 \\ 0 & y \end{pmatrix} + \begin{pmatrix} z & z \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

ou $$\begin{pmatrix} x + y + z & x + z \\ x & x + y \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Egalons les composantes correspondantes les unes aux autres, on obtient ainsi le système homogène équivalent

$$\begin{aligned} x + y + z &= 0 \\ x + \phantom{y+{}} z &= 0 \\ x &= 0 \\ x + y &= 0 \end{aligned}$$

En résolvant le système précédent, on obtient seulement la solution nulle, $x = 0$, $y = 0$, $z = 0$. On a ainsi montré que $xA + yB + zC$ implique $x = 0$, $y = 0$, $z = 0$; les matrices A, B, C sont linéairement indépendantes.

2) Soit une combinaison linéaire des matrices A, B, C que nous égalons à zéro; on a $xA + yB + zC = 0$, D'où

$$x\begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} + y\begin{pmatrix} 3 & -1 \\ 2 & 2 \end{pmatrix} + z\begin{pmatrix} 1 & -5 \\ -4 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

ou
$$\begin{pmatrix} x & 2x \\ 3x & x \end{pmatrix} + \begin{pmatrix} 3y & -y \\ 2y & 2y \end{pmatrix} + \begin{pmatrix} z & -5z \\ -4z & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

ou
$$\begin{pmatrix} x + 3y + z & 2x - y - 5z \\ 3x + 2y - 4z & x + 2y \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Egalons les composantes correspondantes les unes aux autres. On obtient ainsi le système équivalent homogène que l'on réduit à sa forme échelonnée :

$$\begin{aligned} x + 3y + z &= 0 \\ 2x - y - 5z &= 0 \\ 3x + 2y - 4z &= 0 \\ x + 2y &= 0 \end{aligned} \quad \text{ou} \quad \begin{aligned} x + 3y + z &= 0 \\ -7y - 7z &= 0 \\ -7y - 7z &= 0 \\ -y - z &= 0 \end{aligned}$$

ou finalement
$$\begin{aligned} x + 3y + z &= 0 \\ y + z &= 0 \end{aligned}$$

Le système dans sa forme échelonnée a une inconnue libre, et donc une solution non nulle, par exemple $x = 2$, $y = -1$, $z = 1$. Nous avons ainsi montré que $xA + yB + zC = 0$ n'implique pas que $x = 0$, $y = 0$, $z = 0$, donc les matrices sont linéairement dépendantes.

5.4. Soit V l'espace vectoriel des polynômes de degré $\leqslant 3$ sur $\mathbf{R}$. Déterminer si $u, v, w \in V$ sont indépendants ou dépendants où :

(i) $u = t^3 - 3t^2 + 5t + 1$, $v = t^3 - t^2 + 8t + 2$, $w = 2t^3 - 4t^2 + 9t + 5$

(ii) $u = t^3 + 4t^2 - 2t + 3$, $v = t^3 + 6t^2 - t + 4$, $w = 3t^3 + 8t^2 - 8t + 7$

1) Ecrivons une combinaison linéaire des polynômes u, v et w que nous égalons au polynôme identiquement nul, en utilisant les scalaires inconnus x, y et z ; c'est-à-dire $xu + yv + zw = 0$. Ainsi on a :

$$x(t^3 - 3t^2 + 5t + 1) + y(t^3 - t^2 + 8t + 2) + z(2t^3 - 4t^2 + 9t + 5) = 0$$

ou
$$xt^3 - 3xt^2 + 5xt + x + yt^3 - yt^2 + 8yt + 2y + 2zt^3 - 4zt^2 + 9zt + 5z = 0$$

ou
$$(x + y + 2z)t^3 + (-3x - y - 4z)t^2 + (5x + 8y + 9z)t + (x + 2y + 5z) = 0$$

Les coefficients des puissances de t doivent être chacun nuls.

$$\begin{aligned} x + y + 2z &= 0 \\ -3x - y - 4z &= 0 \\ 5x + 8y + 9z &= 0 \\ x + 2y + 5z &= 0 \end{aligned}$$

En résolvant le système homogène précédent, nous obtenons seulement la solution nulle $x = 0$, $y = 0$, $z = 0$, ainsi u, v, w sont indépendants.

2) Ecrivons qu'une combinaison linéaire des polynômes u, v et w est égale au polynôme identiquement nul, en utilisant les scalaires inconnus x, y et z ; c'est-à-dire $xu + yv + zw = 0$. Donc

$$x(t^3+4t^2-2t+3) + y(t^3+6t^2-t+4) + z(3t^3+8t^2-8t+7) = 0$$

ou
$$xt^3 + 4xt^2 - 2xt + 3x + yt^3 + 6yt^2 - yt + 4y + 3zt^3 + 8zt^2 - 8zt + 7z = 0$$

ou
$$(x+y+3z)t^3 + (4x+6y+8z)t^2 + (-2x-y-8z)t + (3x+4y+7z) = 0$$

En égalant à zéro chacun des coefficients des puissances de t et en réduisant le système à la forme échelonnée :

$$\begin{aligned} x + y + 3z &= 0 \\ 4x + 6y + 8z &= 0 \\ -2x - y - 8z &= 0 \\ 3x + 4y + 7z &= 0 \end{aligned} \quad \text{ou} \quad \begin{aligned} x + y + 3z &= 0 \\ 2y - 4z &= 0 \\ y - 2z &= 0 \\ y - 2z &= 0 \end{aligned}$$

ou finalement
$$\begin{aligned} x + y + 3z &= 0 \\ y - 2z &= 0 \end{aligned}$$

Le système dans sa forme échelonnée a une inconnue libre et donc a une solution non nulle. Nous avons montré que $xu + yv + zw = 0$ n'implique pas que $x = 0$, $y = 0$, $z = 0$; ainsi les polynômes sont linéairement dépendants.

5.5. Soit V l'espace vectoriel des fonctions de $\mathbf{R}$ dans $\mathbf{R}$. Montrer que $f, g, h \in V$ sont indépendantes où 1) $f(t) = e^{2t}$, $g(t) = t^2$, $h(t) = t$, 2) $f(t) = \sin t$, $g(t) = \cos t$, $h(t) = t$.

Dans chaque cas, écrivons qu'une combinaison linéaire des fonctions est égale à la fonction nulle $\mathbf{0}$, en utilisant les scalaires inconnus $x, y, z : xf + yg + zh = \mathbf{0}$; et montrons donc que $x = 0$, $y = 0$, $z = 0$. On peut affirmer que $xf + yg + zh = \mathbf{0}$, veut dire que, pour chaque valeur de t, $xf(t) + yg(t) + zh(t) = 0$.

1) Dans l'équation $xe^{2t} + yt^2 + zt = 0$, remplaçons

$t = 0$ on obtient $xe^0 + y0 + z0 = 0$ ou $x = 0$
$t = 1$ on obtient $xe^2 + y + z = 0$
$t = 2$ on obtient $xe^4 + 4y + 2z = 0$

Résolvons le système $\begin{cases} x = 0 \\ xe^2 + y + z = 0 \\ xe^4 + 4y + 2z = 0 \end{cases}$ on obtient seulement la solution $x = 0$, $y = 0$, $z = 0$

Ainsi f, g, h sont indépendants.

2) **Méthode 1 :** Dans l'application $x \sin t + y \cos t + zt = 0$, remplaçons

$t = 0$ pour obtenir $x . 0 + y . 1 + z . 0 = 0$ ou $y = 0$
$t = \pi/2$ pour obtenir $x . 1 + y . 0 + z\pi/2 = 0$ ou $x + \pi z/2 = 0$
$t = \pi$ pour obtenir $x . 0 + y(-1) + z . \pi = 0$ ou $-y + \pi z = 0$

Résolvons le système $\begin{cases} y = 0 \\ x + \pi z/2 = 0 \\ -y + \pi z = 0 \end{cases}$ on obtient seulement la solution $x = 0$, $y = 0$, $z = 0$. Donc f, g et h sont indépendants.

Méthode 2 : Prenons les dérivées première, seconde et troisième de $x \sin t + y \cos t + zt = 0$ par rapport à t.

$$x \cos t - y \sin t + z = 0 \qquad (1)$$
$$-x \sin t - y \cos t = 0 \qquad (2)$$
$$-x \cos t + y \sin t = 0 \qquad (3)$$

Additionnons (*1*) et (*3*) ; on obtient $t = 0$. Multiplions (*2*) par $\sin t$ et (*3*) par $\cos t$, et ajoutons :

$$\begin{array}{lrcl} \sin t \times (2): & -x\sin^2 t - y\sin t\cos t & = & 0 \\ \cos t \times (3): & -x\cos^2 t + y\sin t\cos t & = & 0 \\ \hline & -x(\sin^2 t + \cos^2 t) & = & 0 \quad \text{ou} \quad x = 0 \end{array}$$

Enfin, multiplions (*2*) par $-\cos t$ et (*3*) par $\sin t$, et additionnons ; on obtient :

$$y(\cos^2 t + \sin^2 t) = 0 \quad \text{ou} \quad y = 0$$

Puisque $\quad x\sin t + y\cos t + zt = 0 \quad$ implique $\quad x = 0,\ y = 0,\ z = 0$

on en déduit que f, g, h sont indépendants.

5.6. Soient u, v et w des vecteurs indépendants. Montrer que $u + v$, $u - v$ et $u - 2v + w$ sont aussi indépendants.

Supposons $x(u + v) + y(u - v) + z(u - 2v + w) = 0$ où x, y et z sont des scalaires. Alors $xu + xv + yu - yv + zu - 2zv + zw = 0$ ou

$$(x + y + z)u + (x - y - 2z)v + zw = 0$$

Mais u, v et w sont linéairement indépendants ; donc les coefficients dans la relation précédente sont égaux chacun à 0 :

$$\begin{aligned} x + y + z &= 0 \\ x - y - 2z &= 0 \\ z &= 0 \end{aligned}$$

La seule solution du système précédent est $x = 0$, $y = 0$, $z = 0$. Donc $u + v$, $u - v$, et $u - 2v + w$ sont indépendants.

5.7. Soient $v_1, v_2, \ldots, v_m$ des vecteurs indépendants, et supposons que u soit une combinaison linéaire des v_i, c'est-à-dire $u = a_1 v_1 + a_2 v_2 + \cdots + a_m v_m$ où les a_i sont des scalaires. Montrer que la décomposition précédente de u est unique.

Supposons $u = b_1 v_1 + b_2 v_2 + \cdots + b_m v_m$ où les b_i sont des scalaires. En soustrayant,

$$0 = u - u = (a_1 - b_1)v_1 + (a_2 - b_2)v_2 + \cdots + (a_m - b_m)v_m$$

Mais les v_i sont linéairement indépendants ; donc les coefficients dans la relation précédente sont égaux chacun à 0 :

$$a_1 - b_1 = 0, \quad a_2 - b_2 = 0, \quad \ldots, \quad a_m - b_m = 0$$

Donc $a_1 = b_1, a_2 = b_2, \ldots, a_m = b_m$ et la décomposition précédente de u suivant une combinaison linéaire des v_i est unique.

5.8. Montrer que les vecteurs $v = (1 + i, 2i)$ et $w = (1, 1 + i)$ de $\mathbf{C}^2$ sont linéairement dépendants sur le corps complexe **C**, mais sont linéairement indépendants sur le corps réel **R**.

Rappelons que deux vecteurs sont dépendants si et seulement si l'un d'entre eux est un multiple de l'autre. La première coordonnée de w étant 1, v peut être un multiple de w ssi $v = (1 + i)w$. Mais $1 + i \notin R$; donc v et w sont indépendants sur **R**. Puisque

$$(1 + i)w = (1 + i)(1, 1 + i) = (1 + i, 2i) = v$$

et $1 + i \in \mathbf{C}$, ils sont dépendants sur **C**.

5.9. Supposons que $S = \{v_1, \ldots, v_m\}$ contient un sous-ensemble dépendant, par exemple $\{v_1, v_2, \ldots, v_r\}$. Montrer que S est aussi dépendant. Donc chaque sous-ensemble d'un ensemble indépendant est aussi indépendant.

Puisque $\{v_1, \ldots, v_r\}$ est dépendant, il existe des scalaires $a_1, \ldots, a_r$ non tous nuls, tels que

$$a_1 v_1 + a_2 v_2 + \cdots + a_r v_r = 0$$

Donc il existe des scalaires $a_1, a_2, \ldots, a_r, 0, 0, \ldots, 0$, non tous nuls, tels que

$$a_1v_1 + \cdots + a_rv_r + 0v_{r+1} + \cdots + 0v_m = 0$$

Donc S est dépendant.

5.10. Supposons que $\{v_1, v_2, \ldots, v_m\}$ soit indépendant, mais que $\{v_1, v_2, \ldots, v_m, w\}$ soit dépendant. Montrer que w est une combinaison linéaire des v_i.

Méthode 1 : Puisque le système $\{v_1, v_2, \ldots, v_m, w\}$ est dépendant, il existe alors des scalaires $a_1, a_2, \ldots a_m$, b non tous nuls, de telle sorte que $a_1v_1 + a_2v_2 + \cdots + a_mw_m + bw = 0$. Si $b = 0$, alors l'un des a_i n'est pas nul et $a_1v_1 + a_2v_2 + \cdots + a_mv_m = 0$. Mais ceci est en contradiction avec l'hypothèse que $\{v_1, v_2, v_3, \ldots, v_m\}$ est indépendant. Donc $b \neq 0$ et donc

$$w = b^{-1}(-a_1v_1 - \cdots - a_mv_m) = -b^{-1}a_1v_1 - \cdots - b^{-1}a_mv_m$$

C'est-à-dire que w est une combinaison linéaire des v_i.

Méthode 2 : Si $w = 0$, alors $w = 0v_1 + \cdots + 0v_m$. D'autre part, si $w \neq 0$, alors d'après le lemme 5.2, l'un des vecteurs de $\{v_1, v_2, \ldots, v_m, w\}$ est une combinaison linéaire des vecteurs précédents. Ce vecteur ne peut pas être l'un des v puisque $\{v_1, v_2, \ldots, v_m\}$ est indépendant. Donc w est une combinaison linéaire des v_i.

DEMONSTRATIONS DE THEOREMES

5.11. Démontrer le lemme 5.2. Les vecteurs non nuls $v_1, v_2, \ldots, v_m$ sont linéairement dépendants si et seulement si l'un d'entre eux, par exemple v_i, est une combinaison linéaire des vecteurs précédents : $v_i = a_1v_1 + a_2v_2 + \cdots + a_{i-1}v_{i-1}$.

Supposons $v_i = a_1v_1 + \cdots + a_{i-1}v_{i-1}$. Alors

$$a_1v_1 + \cdots + a_{i-1}v_{i-1} - v_i + 0v_{i+1} + \cdots + 0v_m = 0$$

et le coefficient de v_i n'est pas 0. Donc les v_i sont linéairement dépendants.

Réciproquement, supposons que les v_i soient linéairement dépendants. Il existe alors des scalaires $a_1, a_2, \ldots a_m$, non tous nuls, de telle sorte que $a_1v_1 + \cdots + a_mv_m = 0$. Soit k le plus grand entier tel que $a_k \neq 0$. Alors

$$a_1v_1 + \cdots + a_kv_k + 0v_{k+1} + \cdots + 0v_m = 0 \quad \text{ou} \quad a_1v_1 + \cdots + a_kv_k = 0$$

Supposons $k = 1$; alors $a_1v_1 = 0$, $a_1 \neq 0$, donc $v_1 = 0$. Mais les v_i ne sont pas des vecteurs nuls ; donc $k > 1$ et

$$v_k = -a_k^{-1}a_1v_1 - \cdots - a_k^{-1}a_{k-1}v_{k-1}$$

Donc v_k est une combinaison linéaire des vecteurs précédents.

5.12. Démontrer le théorème 5.1 : Les lignes non nulles $R_1, R_2, \ldots, R_n$ d'une matrice sous la forme échelonnée, sont linéairement indépendantes.

Supposons $\{R_n, R_{n-1}, \ldots, R_1\}$ dépendant. Alors l'une des lignes, par exemples R_m, est une combinaison linéaire des lignes précédentes :

$$R_m = a_{m+1}R_{m+1} + a_{m+2}R_{m+2} + \cdots + a_nR_n \qquad (*)$$

Supposons maintenant que la $k^{\text{ièmes}}$ composante de R_m soit le premier élément non nul. Alors puisque la matrice est sous sa forme échelonnée, les $k^{\text{ièmes}}$ composantes de $R_{m+1}, \ldots, R_n$ sont toutes nulles, et donc la $k^{\text{ième}}$ composante de (*) est $a_{m+1} \cdot 0 + a_{m+2} \cdot 0 + \cdots + a_n \cdot 0 = 0$. Mais ceci est en contradiction avec le fait que la $k^{\text{ième}}$ composante de R_m n'est pas nulle. Ainsi $R_1, R_2, \ldots, R_n$ sont indépendantes.

5.13. Supposons que $\{v_1, v_2, \ldots, v_n\}$ engendre l'espace vectoriel V. Démontrer :

1) Si $w \in V$, alors $\{w, v_1, \ldots, v_m\}$ est linéairement dépendant et engendre W.

2) Si v_i est une combinaison linéaire des vecteurs précédents, alors $\{v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_m\}$ engendre V.

1) Si $w \in V$, alors w est une combinaison linéaire des v_i puisque $\{v_i\}$ engendre V. Donc $\{w, v_1, \ldots, v_m\}$ est linéairement dépendant. Il est évident que w avec les v_i engendrent V puisque les v_i eux-mêmes engendrent V. Donc $\{w, v_1, \ldots, v_m\}$ engendre V.

(2) Supposons que $v_i = k_1 v_1 + \ldots + k_{i-1} v_{i-1}$. Soit $u \in V$. Puisque $\{v_i\}$ engendre V, u est une combinaison linéaire des v_i, c'est-à-dire $u = a_1 v_1 + \ldots + a_m v_m$. En remplaçant v_i par sa valeur, on obtient :

$$\begin{aligned} u &= a_1v_1 + \cdots + a_{i-1}v_{i-1} + a_i(k_1v_1 + \cdots + k_{i-1}v_{i-1}) + a_{i+1}v_{i+1} + \cdots + a_mv_m \\ &= (a_1 + a_ik_1)v_1 + \cdots + (a_{i-1} + a_ik_{i-1})v_{i-1} + a_{i+1}v_{i+1} + \cdots + a_mv_m \end{aligned}$$

Ainsi $\{v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_m\}$ engendre V. En d'autres termes, nous pouvons supprimer v_i de l'ensemble générateur et obtenir un ensemble générateur

5.14. Démontrer le lemme 5.4 : Supposons que $\{v_1, v_2, \ldots, v_n\}$ engendre l'espace vectoriel V. Si $\{w_1, w_2, \ldots, w_n\}$ est linéairement indépendant, alors $m \leqslant n$ et V est engendré par un ensemble de la forme $\{w_1, \ldots, w_m, v_{i_1}, \ldots, v_{i_{n-m}}\}$. Ainsi en particulier $n+1$ ou un nombre supérieur de vecteurs de V sont linéairement dépendants.

Il suffit de démontrer le théorème dans le cas où les v_i ne sont pas tous nuls. Puisque les $\{v_i\}$engendrent V, nous savons d'après le problème précédent que

$$\{w_1, v_1, \ldots, v_n\} \tag{1}$$

est linéairement dépendant et donc engendre V. D'après le lemme 5-2, l'un des vecteurs de (*1*) est une combinaison linéaire des vecteurs précédents. Ce vecteur ne peut pas être w_1, il doit donc être l'un des v, par exemple v_j. Ainsi d'après le précédent problème, nous pouvons supprimer v_j de l'ensemble générateur (*1*) et obtenir l'ensemble générateur

$$\{w_1, v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n\} \tag{2}$$

Répétons le même processus avec le vecteur w_2. Donc puisque (*2*) engendre V, l'ensemble

$$\{w_1, w_2, v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n\} \tag{3}$$

est linéairement indépendant et engendre aussi V. De nouveau à l'aide du lemme 5-2, l'un des vecteurs de (*3*) est une combinaison linéaire des vecteurs précédents. Ce vecteur ne peut être w_1 ou w_2, puisque $\{w_1, \ldots, w_m\}$ est indépendant ; donc ce vecteur doit être l'un des vecteurs v, par exemple v_k. Ainsi d'après le problème précédent, nous pouvons supprimer v_k de l'ensemble générateur (*3*) et on obtient le nouvel ensemble générateur

$$\{w_1, w_2, v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_{k-1}, v_{k+1}, \ldots, v_n\}$$

Répétons ce processus avec w_3 et ainsi de suite. A chaque étape nous pouvons ajouter l'un des vecteurs w et supprimer l'un des vecteurs v dans l'ensemble générateur. Si $m \leqslant n$, on obtient finalement un ensemble générateur de la forme demandée :

$$\{w_1, \ldots, w_m, v_{i_1}, \ldots, v_{i_{n-m}}\}$$

Finalement, nous montrons que $m > n$ n'est pas possible. D'autre part, après n des démarches ci-dessus, nous obtenons l'ensemble générateur $\{w_1, \ldots, w_n\}$. Cela implique que W_{n+1} est une combinaison linéaire de $w_1, \ldots, w_n$, ce qui est contraire à l'hypothèse que $\{w_1\}$ est linéairement indépendant.

5.15. Démontrer le théorème 5.3 : Soit V un espace vectoriel de dimension finie. Alors chaque base de V a le même nombre de vecteurs.

Supposons que $\{e_1, e_2, \ldots, e_n\}$ soit une base de V et supposons que $\{f_1, f_2, \ldots,\}$ soit une autre base de V. Puisque $\{e_i\}$ engendre V, la base $\{f_1, f_2, \ldots\}$ doit contenir n ou un nombre inférieur de vecteurs, ou sinon le système précédent est dépendant d'après le précédent problème. D'autre part, si la base $\{f_1, f_2, \ldots\}$ contient moins de n vecteurs, alors $\{e_1, \ldots, e_n\}$ est dépendant d'après le problème précédent. Donc la base $\{f_1, f_2, \ldots\}$ contient exactement n vecteurs et alors le théorème est vrai.

5.16. Démontrer le théorème 5.5 : Supposons que $\{v_1, \ldots, v_m\}$ soit un sous-ensemble indépendant maximal d'un ensemble S qui engendre un espace vectoriel V. Alors $\{v_1, v_2, \ldots, v_m\}$ est une base de V.

Supposons $w \in S$. Puisque $\{v_i\}$ est le sous-ensemble maximum indépendant de S, $\{v_1, v_2, \ldots, v_m, w\}$ est linéairement dépendant. D'après le problème 5.10, w est une combinaison linéaire des v_i, c'est-à-dire $w \in L(v_i)$. Donc $S \subset L(v_i)$. Ceci conduit à $V = L(S) \subset L(v_i) \subset V$. Donc, $\{v_i\}$ engendre V et, $\{v_i\}$ étant indépendant, c'est une base de V.

5.17. Supposons que V soit engendré par un ensemble fini S. Montrer que V est de dimension finie et en particulier qu'un sous-ensemble de S est une base de V.

Méthode 1. Parmi tous les sous-ensembles indépendants de S et il y a un nombre fini d'entre eux puisque S est fini, il y en a un au moins qui est maximal. D'après le précédent problème ce sous-ensemble de S est une base de V.

Méthode 2. Si S est indépendant, c'est une base de V. Si S est dépendant, l'un des vecteurs est une combinaison linéaire des vecteurs précédents. Nous pouvons supprimer ce vecteur et obtenir encore un ensemble générateur. Nous continuons ce processus jusqu'à obtenir un sous-ensemble qui est indépendant et engendre V c'est-à-dire une base de V.

5.18. Démontrer le théorème 5.6 : Soit V de dimension finie n. Alors

(i) Un ensemble quelconque de $n+1$ ou d'un nombre supérieur de vecteurs est linéairement dépendant.

(ii) Un ensemble quelconque linéairement indépendant est une partie de la base.

(iii) Un ensemble linéairement indépendant de n éléments est une base.

Supposons que $\{e_1, \ldots, e_n\}$ soit une base de V.

(i) Puisque $\{e_1, \ldots, e_n\}$ engendre V, $n+1$ vecteurs ou un nombre supérieur de vecteurs sont dépendants d'après le lemme 5.4.

(ii) Supposons que $\{v_1, \ldots, v_r\}$ soit indépendant. D'après le lemme 5.4, V est engendré par un ensemble de la forme

$$S = \{v_1, \ldots, v_r, e_{i_1}, \ldots, e_{i_{n-r}}\}$$

D'après le précédent problème, un sous-ensemble de S est une base. Mais S contient n éléments et chaque base de V et contient n éléments. Donc S est une base de V et contient comme sous-ensemble $\{v_1, v_2, \ldots, v_r\}$.

(iii) D'après (ii) un ensemble indépendant T de n éléments est une partie d'une base, Mais chaque base de V contient n éléments. Donc T est une base.

5.19. Démontrer le théorème 5.7 : Soit W un sous-espace d'un espace vectoriel V à n dimensions. Alors $\dim W \leqslant n$. En particulier si $\dim W = n$ alors $W = V$.

Puisque V est de dimension n, $n+1$ vecteurs ou un nombre supérieur de vecteurs sont linéairement dépendants. De plus, puisqu'une base de W contient uniquement des vecteurs indépendants, elle ne peut pas contenir plus de n éléments. Donc $\dim W \leqslant n$.

En particulier, si $\{w_1, \ldots, w_n\}$ est une base de W, puisque c'est un ensemble indépendant de n éléments, c'est aussi une base de V. Donc $W = V$ lorsque $\dim W = n$.

5.20. Démontrer le théorème 5.8 : $\dim(U+W) = \dim U + \dim W - \dim(U \cap W)$.

Remarquons que $U \cap W$ est un sous-espace à la fois de U et W. Supposons que $\dim U = m$, $\dim W = n$ et que $\dim (U \cap W) = r$. Supposons de plus que $\{v_1, v_2, \ldots, v_r\}$ soit une base de $U \cap W$. D'après le théorème 5.6 : (ii) nous pouvons étendre $\{v_i\}$ à une base de U et à une base de W ; par exemple :

$$\{v_1, \ldots, v_r, u_1, \ldots, u_{m-r}\} \qquad \text{et} \qquad \{v_1, \ldots, v_r, w_1, \ldots, w_{n-r}\}$$

sont des bases de U et W respectivement. Soit

$$B = \{v_1, \ldots, v_r, u_1, \ldots, u_{m-r}, w_1, \ldots, w_{n-r}\}$$

Remarquons que B a exactement $m+n-r$ éléments. Ainsi le théorème est démontré si nous pouvons montrer que B est une base de $U+W$. Puisque $\{v_i, u_j\}$ engendre U et $\{v_i, w_k\}$ engendre W, l'union $B = \{v_i, u_j, w_k\}$ engendre $U+W$. Il suffit alors de montrer que B est indépendant.

Supposons

$$a_1v_1 + \cdots + a_rv_r + b_1u_1 + \cdots + b_{m-r}u_{m-r} + c_1w_1 + \cdots + c_{n-r}w_{n-r} = 0 \tag{1}$$

où a_i, b_j, c_k sont des scalaires. Soit

$$v = a_1v_1 + \cdots + a_rv_r + b_1u_1 + \cdots + b_{m-r}u_{m-r} \tag{2}$$

D'après (*1*), nous avons aussi

$$v = -c_1w_1 - \cdots - c_{n-r}w_{n-r} \qquad (3)$$

Puisque $\{v_i, u_j\} \subset U$, $v \in U$ d'après (*2*) ; et puisque $\{w_k\} \subset W$, $v \in W$ d'après (*3*). Donc $v \in U \cap W$. Maintenant $\{v_i\}$ est une base de $U \cap W$ et donc il existe des scalaires $d_1, d_2, \ldots, d_r$ pour lesquels $v = d_1v_1 + d_2v_2 + \ldots d_rv_r$. Nous avons alors d'après (*3*)

$$d_1v_1 + \cdots + d_rv_r + c_1w_1 + \cdots + c_{n-r}w_{n-r} = 0$$

Mais $\{v_i, w_k\}$ est une base de W et donc est indépendante. Donc il s'ensuit les égalités principales $c_1 = 0, \ldots, c_{n-r} = 0$. Remplaçons dans (*1*) on obtient :

$$a_1v_1 + \cdots + a_rv_r + b_1u_1 + \cdots + b_{m-r}u_{m-r} = 0$$

Mais $\{v_i, u_j\}$ est une base de U et donc est indépendante. Il s'ensuit donc les égalités principales $a_1 = 0, \ldots, a_r = 0$, $b_1 = 0, \ldots, b_{m-r} = 0$.

Puisque l'équation (*1*) implique que les a_i, b_j et c_k sont tous nuls. $B = \{v_i, u_j, w_k\}$ est indépendante et le théorème est démontré.

5.21. Démontrer le théorème 5.9 : Le rang ligne et le rang colonne d'une quelconque matrice sont égaux.

Soit A une matrice quelconque $m \times n$

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$$

Soient $R_1, R_2, \ldots, R_m$ les différentes lignes de cette matrice :

$$R_1 = (a_{11}, a_{12}, \ldots, a_{1n}), \quad \ldots, \quad R_m = (a_{m1}, a_{m2}, \ldots, a_{mn})$$

Supposons que le rang ligne soit r et que les r vecteurs suivants forment une base de l'espace ligne:

$$S_1 = (b_{11}, b_{12}, \ldots, b_{1n}), \quad S_2 = (b_{21}, b_{22}, \ldots, b_{2n}), \quad \ldots, \quad S_r = (b_{r1}, b_{r2}, \ldots, b_{rn})$$

Alors chacun des vecteurs lignes est une combinaison linéaire des S_i :

$$\begin{aligned} R_1 &= k_{11}S_1 + k_{12}S_2 + \cdots + k_{1r}S_r \\ R_2 &= k_{21}S_1 + k_{22}S_2 + \cdots + k_{2r}S_r \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ R_m &= k_{m1}S_1 + k_{m2}S_2 + \cdots + k_{mr}S_r \end{aligned}$$

où les k_{ij} sont des scalaires. En égalant les $i$èmes composantes de chacune des équations vectorielles précédentes on obtient le système d'équations suivant, chacune étant valable pour $i = 1, \ldots, n$.

$$\begin{aligned} a_{1i} &= k_{11}b_{1i} + k_{12}b_{2i} + \cdots + k_{1r}b_{ri} \\ a_{2i} &= k_{21}b_{1i} + k_{22}b_{2i} + \cdots + k_{2r}b_{ri} \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{mi} &= k_{m1}b_{1i} + k_{m2}b_{2i} + \cdots + k_{mr}b_{ri} \end{aligned}$$

Ainsi pour $i = 1, \ldots, n$:

$$\begin{pmatrix} a_{1i} \\ a_{2i} \\ \cdot\cdot \\ a_{mi} \end{pmatrix} = b_{1i}\begin{pmatrix} k_{11} \\ k_{21} \\ \cdot\cdot \\ k_{m1} \end{pmatrix} + b_{2i}\begin{pmatrix} k_{12} \\ k_{22} \\ \cdot\cdot \\ k_{m2} \end{pmatrix} + \cdots + b_{ri}\begin{pmatrix} k_{1r} \\ k_{2r} \\ \cdot\cdot \\ k_{mr} \end{pmatrix}$$

En d'autres termes, chacune des colonnes de A est une combinaison linéaire des r vecteurs

$$\begin{pmatrix} k_{11} \\ k_{21} \\ \cdot\cdot \\ k_{m1} \end{pmatrix}, \quad \begin{pmatrix} k_{12} \\ k_{22} \\ \cdot\cdot \\ k_{m2} \end{pmatrix}, \quad \ldots, \quad \begin{pmatrix} k_{1r} \\ k_{2r} \\ \cdot\cdot \\ k_{mr} \end{pmatrix}$$

Ainsi l'espace colonne de la matrice A a une dimension au plus égale à r, c'est-à-dire que le rang colonne $\leqslant r$. Donc le rang colonne $\leqslant$ rang ligne.

De manière semblable (ou en considérant la matrice transposée A^t) on obtient rang ligne $\leqslant$ rang colonne. Donc le rang ligne et le rang colonne sont égaux.

BASE ET DIMENSION

5.22. Déterminer si oui ou non les vecteurs suivants forment une base de l'espace vectoriel $\mathbf{R}^3$:

(i) $(1, 1, 1)$ et $(1, -1, 5)$

(ii) $(1, 2, 3)$, $(1, 0, -1)$, $(3, -1, 0)$ et $(2, 1, -2)$

(iii) $(1, 1, 1)$, $(1, 2, 3)$ et $(2, -1, 1)$

(iv) $(1, 1, 2)$, $(1, 2, 5)$ et $(5, 3, 4)$

(i) et (ii) non ; car une base de $\mathbf{R}^3$ doit contenir exactement 3 éléments, puisque $\mathbf{R}^3$ est de dimension 3.

(iii) Les vecteurs forment une base si et seulement s'ils sont indépendants. Formons donc la matrice dont les lignes sont les vecteurs donnés, et réduisons les lignes à une forme échelonnée :

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 2 & -1 & 1 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & -3 & -1 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 5 \end{pmatrix}$$

La matrice échelonnée n'a pas de ligne nulle ; ainsi les trois vecteurs sont indépendants et donc forment une base de $\mathbf{R}^3$.

(iv) Formons la matrice dont les lignes sont les vecteurs donnés et réduisons la matrice à sa forme échelonnée, par des opérations élémentaires sur les lignes :

$$\begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & 5 \\ 5 & 3 & 4 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 3 \\ 0 & -2 & -6 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix}$$

La matrice échelonnée a une ligne nulle ; c'est-à-dire seulement deux lignes non nulles ; donc les trois vecteurs sont dépendants et donc ne forment pas une base de $\mathbf{R}^3$.

5.23. Soit W le sous-espace de $\mathbf{R}^4$ engendré par les vecteur $(1, -2, 5, -3)$, $(2, 3, 1, -4)$ et $(3, 8, -3, -5)$. (i) Trouver une base et la dimension de W. (ii) Etendre la base de W à une base de l'espace complet $\mathbf{R}^4$.

(i) Formons la matrice dont les lignes sont les vecteurs donnés et réduisons-la à sa forme échelonnée par des opérations élémentaires sur les lignes.

$$\begin{pmatrix} 1 & -2 & 5 & -3 \\ 2 & 3 & 1 & -4 \\ 3 & 8 & -3 & -5 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & -2 & 5 & 3 \\ 0 & 7 & -9 & 2 \\ 0 & 14 & -18 & 4 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & -2 & 5 & -3 \\ 0 & 7 & -9 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Les lignes non nulles $(1, -2, 5, -3)$ et $(0, 7, -9, 2)$ de la matrice échelonnée, forment une base de l'espace ligne, c'est-à-dire de W. Donc en particulier $\dim W = 2$.

(ii) Cherchons quatre vecteurs indépendants qui incluront les deux vecteurs précédents. Les vecteurs $(1, -2, 5, -3)$, $(0, 7, -9, 2)$, $(0, 0, 1, 0)$ et $(0, 0, 0, 1)$ sont indépendants (puisqu'ils forment une matrice échelonnée), et forment donc une base de $\mathbf{R}^4$ qui est une extension de la base de W.

5.24. Soit W un espace engendré par les polynômes

$$v_1 = t^3 - 2t^2 + 4t + 1 \qquad v_3 = t^3 + 6t - 5$$
$$v_2 = 2t^3 - 3t^2 + 9t - 1 \qquad v_4 = 2t^3 - 5t^2 + 7t + 5$$

Trouver une base et la dimension de W.

Les coordonnées des vecteurs polynômes relativement à la base $(t^3, t^2, t, 1)$ sont respectivement

$$[v_1] = (1, -2, 4, 1) \qquad [v_3] = (1, 0, 6, -5)$$
$$[v_2] = (2, -3, 9, -1) \qquad [v_4] = (2, -5, 7, 5)$$

Formons la matrice dont les lignes sont les coordonnées des vecteurs précédents et réduisons-la à sa forme échelonnée, par des opérations élémentaires sur les lignes :

$$\begin{pmatrix} 1 & -2 & 4 & 1 \\ 2 & -3 & 9 & -1 \\ 1 & 0 & 6 & -5 \\ 2 & -5 & 7 & 5 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & -2 & 4 & 1 \\ 0 & 1 & 1 & -3 \\ 0 & 2 & 2 & -6 \\ 0 & -1 & -1 & 3 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & -2 & 4 & 1 \\ 0 & 1 & 1 & -3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Les lignes non nulles $(1, -2, 4, 1)$ et $(0, 1, 1, -3)$ de la matrice sous forme échelonnée forment une base de l'espace engendré par les vecteurs correspondants et donc forment les polynômes

$$t^3 - 2t^2 + 4t + 1 \quad \text{et} \quad t^2 + t - 3$$

qui sont une base de W. Donc $\dim W = 2$.

5.25. Trouver la dimension et une base de l'espace solution W du système

$$\begin{aligned} x + 2y + 2z - s + 3t &= 0 \\ x + 2y + 3z + s + t &= 0 \\ 3x + 6y + 8z + s + 5t &= 0 \end{aligned}$$

Réduisons le système à sa forme échelonnée.

$$\begin{aligned} x + 2y + 2z - s + 3t &= 0 \\ z + 2s - 2t &= 0 \\ 2z + 4s - 4t &= 0 \end{aligned} \qquad \text{ou} \qquad \begin{aligned} x + 2y + 2z - s + 3t &= 0 \\ z + 2s - 2t &= 0 \end{aligned}$$

y s t

Le système dans sa forme échelonnée à 2 équations non nulles à 5 inconnues ; donc la dimension de l'espace solution W est $5 - 2 = 3$. Les inconnues libres sont ~~y, s, t~~. Donc y, s, t

(i) $y = 1,\ s = 0,\ t = 0$, (ii) $y = 0,\ s = 1,\ t = 0$, (iii) $y = 0,\ s = 0,\ t = 1$

ce qui permet d'obtenir les solutions respectives

$$v_1 = (-2, 1, 0, 0, 0), \quad v_2 = (5, 0, -2, 1, 0), \quad v_3 = (-7, 0, 2, 0, 1)$$

L'ensemble $\{v_1, v_2, v_3\}$ est une base de l'espace solution W.

5.26. Trouver un système homogène dont l'ensemble solution W est engendré par

$$\{(1, -2, 0, 3), (1, -1, -1, 4), (1, 0, -2, 5)\}$$

Méthode 1. Soit $v = (x, y, z, w)$. Formons la matrice M dont les premières lignes sont les vecteurs donnés et dont la dernière ligne est v, et réduisons-la à sa forme échelonnée :

$$M = \begin{pmatrix} 1 & -2 & 0 & 3 \\ 1 & -1 & -1 & 4 \\ 1 & 0 & -2 & 5 \\ x & y & z & w \end{pmatrix} \text{d'où} \begin{pmatrix} 1 & -2 & 0 & 3 \\ 0 & 1 & -1 & 1 \\ 0 & 2 & -2 & 2 \\ 0 & 2x+y & z & -3x+w \end{pmatrix} \text{d'où} \begin{pmatrix} 1 & -2 & 0 & 3 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 2x+y+z & -5x-y+w \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Les trois premières lignes montrent que W a pour dimension 2. Ainsi $v \in W$ si et seulement si la ligne supplémentaire n'augmente pas la dimension de l'espace ligne. Donc nous allons écrire que les deux derniers éléments de la troisième ligne sont nuls et nous obtenons ainsi le système demandé qui est homogène :

$$\begin{aligned} 2x + y + z \quad &= 0 \\ 5x + y \quad - w &= 0 \end{aligned}$$

Méthode 2. Nous savons que $v = (x, y, z, w) \in W$ si et seulement si v est une combinaison linéaire des générateurs de W :

$$(x, y, z, w) = r(1, -2, 0, 3) + s(1, -1, -1, 4) + t(1, 0, -2, 5)$$

L'équation vectorielle précédente dont les inconnues sont r, s et t est équivalente au système suivant :

$$\begin{aligned} r + s + t &= x \\ -2r - s \quad &= y \\ -s - 2t &= z \\ 3r + 4s + 5t &= w \end{aligned} \quad \text{ou} \quad \begin{aligned} r + s + t &= x \\ s + 2t &= 2x + y \\ -s - 2t &= z \\ s + 2t &= w - 3x \end{aligned} \quad \text{ou} \quad \begin{aligned} r + s + t &= x \\ s + 2t &= 2x + y \\ 0 &= 2x + y + z \\ 0 &= 5x + y - w \end{aligned} \qquad (1)$$

Ainsi $v \in W$ si et seulement si le système précédent a une solution, c'est-à-dire si

$$\begin{aligned} 2x + y + z &= 0 \\ 5x + y - w &= 0 \end{aligned}$$

Le système précédent est le système demandé.

Remarque : Observons que la matrice complète du système (*1*) est la transposée de la matrice M utilisée dans la première méthode.

5.27. Soient U et W les sous-espaces suivants de $\mathbf{R}^4$:

$$U = \{(a,b,c,d): b+c+d = 0\}, \qquad W = \{(a,b,c,d): a+b = 0, c = 2d\}$$

Trouver la dimension et une base de (i) U, (ii) W, (iii) $U \cap W$.

(i) Nous cherchons une base de l'ensemble des solutions (a, b, c, d) de l'équation

$$b + c + d = 0 \quad \text{ou} \quad 0 \cdot a + b + c + d = 0$$

Les inconnues libres sont a, c et d. Posons

$$(1)\ a = 1,\ c = 0,\ d = 0, \qquad (2)\ a = 0,\ c = 1,\ d = 0, \qquad (3)\ a = 0,\ c = 0,\ d = 1$$

pour obtenir les solutions respectives

$$v_1 = (1, 0, 0, 0), \qquad v_2 = (0, -1, 1, 0), \qquad v_3 = (0, -1, 0, 1)$$

L'ensemble $\{v_1, v_2, v_3\}$ est une base de U, et $\dim U = 3$.

(ii) Nous cherchons une base de l'ensemble des solutions (a, b, c, d) du système

$$\begin{aligned} a + b &= 0 \\ c &= 2d \end{aligned} \quad \text{ou} \quad \begin{aligned} a + b &= 0 \\ c - 2d &= 0 \end{aligned}$$

Les inconnues libres sont b et d. Posons

$$(1)\ b = 1,\ d = 0, \qquad (2)\ b = 0,\ d = 1$$

pour obtenir les solutions respectives

$$v_1 = (-1, 1, 0, 0), \qquad v_2 = (0, 0, 2, 1)$$

L'ensemble $\{v_1, v_2\}$ est une base de W et $\dim W = 2$.

(iii) $U \cap W$ contient les vecteurs (a, b, c, d) qui satisfont les conditions définies par U et celles définies par W, c'est-à-dire les trois équations

$$\begin{aligned} b + c + d &= 0 \\ a + b \quad &= 0 \\ c &= 2d \end{aligned} \quad \text{ou} \quad \begin{aligned} a + b \quad &= 0 \\ b + c + d &= 0 \\ c - 2d &= 0 \end{aligned}$$

L'inconnue libre est d. Posons $d = 1$; on obtient la solution $v = (3, -3, 2, 1)$. Ainsi $\{v\}$ est une base de $U \cap W$ est $\dim (U \cap W) = 1$.

5.28. Trouver la dimension de l'espace vectoriel engendré par

(i) $(1, -2, 3, -1)$ et $(1, 1, -2, 3)$

(ii) $(3, -6, 3, -9)$ et $(-2, 4, -2, 6)$

(iii) $t^3 + 2t^2 + 3t + 1$ et $2t^3 + 4t^2 + 6t + 2$

(iv) $t^3 - 2t^2 + 5$ et $t^2 + 3t - 4$

(v) $\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$ et $\begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$

(vi) $\begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$ et $\begin{pmatrix} -3 & -3 \\ 3 & 3 \end{pmatrix}$

(vii) 3 et -3

Deux vecteurs non nuls engendrent un espace W de dimension 2, s'ils sont indépendants, et de dimension 1 s'ils sont dépendants. Rappelons que deux vecteurs sont dépendants si et seulement si l'un d'entre eux est un multiple de l'autre. D'où (i) 2 ; (ii) 1 ; (iii) 1; (iv) 2 ; (v) 2 ; (vi) 1 ; (vii) 1.

5.29. Soit V l'espace vectoriel des 2×2 matrices symétriques sur K. Montrer que $\dim V = 3$. Rappelons que $A = (a_{ij})$ est symétrique si et seulement si $A = A^t$ ou $a_{ij} = a_{ji}$).

Une matrice quelconque 2×2 symétrique est de la forme $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ où $a, b, c \in K$. Remarquons qu'il y a 3 variables). Posons

$$\text{(i) } a = 1,\ b = 0,\ c = 0, \qquad \text{(ii) } a = 0,\ b = 1,\ c = 0, \qquad \text{(iii) } a = 0,\ b = 0,\ c = 1$$

Nous obtenons les matrices respectives

$$E_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad E_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad E_3 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Nous avons montré que $\{E_1, E_2, E_3\}$ est une base de V, c'est-à-dire qu'il engendre V (1) et qu'il est indépendant (2).

(1) Pour la matrice A précédente dans V, nous avons

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix} = aE_1 + bE_2 + cE_3$$

Donc $\{E_1, E_2, E_3\}$ engendre V.

(2) Supposons $xE_1 + yE_2 + zE_3 = 0$, où x, y, z sont des scalaires inconnus, c'est-à-dire supposons que

$$x\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + y\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + z\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{ou} \quad \begin{pmatrix} x & y \\ y & z \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

En égalant les éléments correspondants les uns aux autres, on obtient $x = 0$, $y = 0$, $z = 0$. En d'autres termes,

$$xE_1 + yE_2 + zE_3 = 0 \quad \text{implique} \quad x = 0,\ y = 0,\ z = 0$$

En conséquence $\{E_1, E_2, E_3\}$ est indépendant.

Donc $\{E_1, E_2, E_3\}$ est une base de V et donc la dimension de V est 3.

5.30. Soit V l'espace vectoriel des polynômes en t de degré $\leqslant n$. Montrer que chacun des ensembles suivants constitue une base de V :
(i) $\{1, t, t^2, \ldots, t^{n-1}, t^n\}$, (ii) $\{1, 1-t, (1-t)^2, \ldots, (1-t)^{n-1}, (1-t)^n\}$.
Donc $\dim V = n + 1$.

(i) Il est clair que chacun des polynômes de V est une combinaison linéaire de $1, t, \ldots, t^{n-1}$ et t^n. De plus $1, t, \ldots, t^{n-1}$ et t^n sont indépendants puisqu'aucun de ces polynômes n'est une combinaison linéaire des polynômes précédents. Ainsi $\{1, t, t^2, \ldots, t^n\}$ est une base de V.

(ii) (Remarquons que d'après (i) $\dim V = n + 1$; et donc un ensemble quelconque de $n + 1$ polynômes indépendants forme une base de V). Chaque polynôme de la suite $1, 1 - t, \ldots, (1 - t)^n$ est d'un degré supérieur au précédent et donc ne peut être une combinaison linéaire des précédents. Donc les $n + 1$ polynômes $1, 1 - t, \ldots, (1 - t)^n$ sont indépendants et forment une base de V.

5.31. Soit V l'espace vectoriel des couples de nombres complexes sur le corps réel $\mathbf{R}$ (voir problème 4.42). Montrer que V est de dimension 4.

Nous affirmons que l'ensemble suivant est une base de V:

$$B = \{(1, 0), (i, 0), (0, 1), (0, i)\}$$

Supposons $v \in V$. Alors $v = (z, w)$ où z, w sont des nombres complexes et donc de la forme $v = (a + bi, c + di)$ où a, b, c, d sont des nombres réels. Donc

$$v = a(1, 0) + b(i, 0) + c(0, 1) + d(0, i)$$

Ainsi B engendre V.

La démonstration est complète si nous montrons que B est indépendant. Supposons

$$x_1(1, 0) + x_2(i, 0) + x_3(0, 1) + x_4(0, i) = 0$$

où $x_1, x_2, x_3, x_4 \in R$. Alors

$$(x_1 + x_2 i, x_3 + x_4 i) = (0, 0) \quad \text{et donc} \quad \begin{cases} x_1 + x_2 i = 0 \\ x_3 + x_4 i = 0 \end{cases}$$

Donc $x_1 = 0$, $x_2 = 0$, $x_3 = 0$, $x_4 = 0$ et donc B est indépendant

5.32. Soit V l'espace vectoriel des matrices $m \times n$ sur un corps K. Soit $E_{ij} \in V$ la matrice qui contient 1 comme ij-élément et 0 partout ailleurs. Montrer que $\{E_{ij}\}$ est une base de V. Donc $\dim V = m \times n$.

Il nous faut donc démontrer que $\{E_{ij}\}$ engendre V et est indépendant.

Soit $A = (a_{ij})$ une matrice quelconque de V. D'où $A = \sum_{i,j} a_{ij} E_{ij}$. Donc $\{E_{ij}\}$engendre V.

Supposons maintenant que $\sum_{i,j} x_{ij} E_{ij} = 0$, où les x_{ij} sont les scalaires. Le ij-élément de $\sum_{i,j} x_{ij} E_{ij}$ est x_{ij} et le $ij^{\text{ème}}$élément de 0 est 0. Donc $x_{ij} = 0$, $i = 1, \ldots, m$, $j = 1, \ldots, n$. Les matrices sont en conséquence indépendantes.

D'où $\{E_{ij}\}$ est une base de V.

Remarque : En considérant un vecteur de K^n comme une $1 \times n$ matrice, nous avons montré par le précédent résultat que la base habituelle définie dans l'exemple 5.3 page 88 est une base de K^n et donc $\dim K^n = n$.

SOMMES ET INTERSECTIONS

5.33. Supposons que U et W soient deux sous-espaces distincts de dimension 4 d'un espace vectoriel V de dimension 6. Trouver les dimensions possibles de $U \cap W$.

Puisque U et W sont distincts, $U + W$ contient donc U et W ; donc $\dim(U + W) > 4$. Mais $\dim(U+W)$ ne peut pas être supérieure à 6, puisque $\dim V = 6$. Nous avons donc deux possibilités : 1) $\dim(U + W) = 5$, ou 2) $\dim(U + W) = 6$. En utilisant le théorème 5.8 : $\dim(U + W) = \dim U + \dim W - \dim(U \cap W)$, on obtient

$$1)\ 5 = 4 + 4 - \dim(U \cap W) \quad \text{ou} \quad \dim(U \cap W) = 3$$

$$2)\ 6 = 4 + 4 - \dim(U \cap W) \quad \text{ou} \quad \dim(U \cap W) = 2.$$

Donc la dimension de $U \cap W$ doit être soit 2 soit 3.

5.34. Soient U et W deux sous-espaces de $\mathbf{R}^4$ engendrés par

$$\{(1, 1, 0, -1), (1, 2, 3, 0), (2, 3, 3, -1)\} \quad \text{et} \quad \{(1, 2, 2, -2), (2, 3, 2, -3), (1, 3, 4, -3)\}$$

respectivement, Trouver 1) $\dim(U + W)$, 2) $\dim(U \cap W)$.

1) $U + W$ est l'espace engendré par les six vecteurs. Donc formons la matrice dont les lignes sont les six vecteurs donnés et réduisons-la à sa forme échelonnée.

$$\begin{pmatrix} 1 & 1 & 0 & -1 \\ 1 & 2 & 3 & 0 \\ 2 & 3 & 3 & -1 \\ 1 & 2 & 2 & -2 \\ 2 & 3 & 2 & -3 \\ 1 & 3 & 4 & -3 \end{pmatrix} \text{d'où} \begin{pmatrix} 1 & 1 & 0 & -1 \\ 0 & 1 & 3 & 1 \\ 0 & 1 & 3 & 1 \\ 0 & 1 & 2 & -1 \\ 0 & 1 & 2 & -1 \\ 0 & 2 & 4 & -2 \end{pmatrix} \text{d'où} \begin{pmatrix} 1 & 1 & 0 & -1 \\ 0 & 1 & 3 & 1 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \text{d'où} \begin{pmatrix} 1 & 1 & 0 & -1 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & -1 & -2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Puisque la matrice échelonnée a trois lignes non nulles, $\dim(U + W) = 3$.

2) Trouvons d'abord la dimension de U et la dimension de W. Formons les deux matrices dont les lignes sont les générateurs de U et de W respectivement et réduisons-là à sa forme échelonnée par des opérations élémentaires sur les lignes:

$$\begin{pmatrix} 1 & 1 & 0 & -1 \\ 1 & 2 & 3 & 0 \\ 2 & 3 & 3 & -1 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & 1 & 0 & -1 \\ 0 & 1 & 3 & 1 \\ 0 & 1 & 3 & 1 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & 1 & 0 & -1 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

et

$$\begin{pmatrix} 1 & 2 & 2 & -2 \\ 2 & 3 & 2 & -3 \\ 1 & 3 & 4 & -3 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & 2 & 2 & -2 \\ 0 & -1 & -2 & 1 \\ 0 & 1 & 2 & -1 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & 2 & 2 & -2 \\ 0 & -1 & -2 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Puisque chacune des matrices échelonnées a deux lignes non nulles, $\dim U = 2$ et $\dim W = 2$. En utilisant le théorème 5.8, on a $\dim(U + W) = \dim U + \dim W - \dim(U \cap W)$, ce qui donne

$$3 = 2 + 2 - \dim(U \cap W) \quad \text{ou} \quad \dim(U \cap W) = 1$$

5.35. Soit U un sous-espace de $\mathbf{R}^5$ engendré par

$$\{(1, 3, -2, 2, 3), (1, 4, -3, 4, 2), (2, 3, -1, -2, 9)\}$$

et soit W le sous-espace de $\mathbf{R}^5$ engendré par

$$\{(1, 3, 0, 2, 1), (1, 5, -6, 6, 3), (2, 5, 3, 2, 1)\}$$

Trouver une base et la dimension de 1) $U + W$, 2) $U \cap W$.

1) $U + W$ est l'espace engendré par l'ensemble des six vecteurs. Formons la matrice dont les lignes sont les six vecteurs et réduisons-la à sa forme échelonnée par des opérations élémentaires sur les lignes :

$$\begin{pmatrix} 1 & 3 & -2 & 2 & 3 \\ 1 & 4 & -3 & 4 & 2 \\ 2 & 3 & -1 & -2 & 9 \\ 1 & 3 & 0 & 2 & 1 \\ 1 & 5 & -6 & 6 & 3 \\ 2 & 5 & 3 & 2 & 1 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & 3 & -2 & 2 & 3 \\ 0 & 1 & -1 & 2 & -1 \\ 0 & -3 & 3 & -6 & 3 \\ 0 & 0 & 2 & 0 & -2 \\ 0 & 2 & -4 & 4 & 0 \\ 0 & -1 & 7 & -2 & -5 \end{pmatrix}$$

$$\text{d'où } \begin{pmatrix} 1 & 3 & -2 & 2 & 3 \\ 0 & 1 & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & -2 \\ 0 & 0 & -2 & 0 & 2 \\ 0 & 0 & 6 & 0 & -6 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & 3 & -2 & 2 & 3 \\ 0 & 1 & -1 & 2 & -1 \\ 0 & 0 & 2 & 0 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

L'ensemble des lignes non nulles de la matrice échelonnée

$$\{(1, 3, -2, 2, 3), (0, 1, -1, 2, -1), (0, 0, 2, 0, -2)\}$$

constitue une base de $U + W$; donc $\dim(U + W) = 3$.

2) Trouvons d'abord les systèmes homogènes dont les ensembles solutions sont U et W respectivement. Formons la matrice dont les premières lignes sont les générateurs de U et dont la dernière ligne est (x, y, z, s, t) et reduisons-la à sa forme échelonnée.

$$\begin{pmatrix} 1 & 3 & -2 & 2 & 3 \\ 1 & 4 & -3 & 4 & 2 \\ 2 & 3 & -1 & -2 & 9 \\ x & y & z & s & t \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & 3 & -2 & 2 & 3 \\ 0 & 1 & -1 & 2 & -1 \\ 0 & -3 & 3 & -6 & 3 \\ 0 & -3x + y & 2x + z & -2x + s & -3x + t \end{pmatrix}$$

$$\text{d'où } \begin{pmatrix} 1 & 3 & -2 & 2 & 3 \\ 0 & 1 & -1 & 2 & -1 \\ 0 & 0 & -x + y + z & 4x - 2y + s & -6x + y + t \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Ecrivons que les éléments de la troisième ligne sont égaux à 0 de manière à obtenir le système homogène dont l'ensemble solution est U :

$$-x+y+z = 0, \quad 4x-2y+s = 0, \quad -6x+y+t = 0$$

Formons maintenant la matrice dont les premières lignes sont les générateurs de W et dont la dernière ligne est (x, y, z, s, t) et réduisons-la à sa forme échelonnée:

$$\begin{pmatrix} 1 & 3 & 0 & 2 & 1 \\ 1 & 5 & -6 & 6 & 3 \\ 2 & 5 & 3 & 2 & 1 \\ x & y & z & s & t \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & 3 & 0 & 2 & 1 \\ 0 & 2 & -6 & 4 & 2 \\ 0 & -1 & 3 & -2 & -1 \\ 0 & -3x+y & z & -2x+s & -x+t \end{pmatrix}$$

$$\begin{pmatrix} 1 & 3 & 0 & 2 & 1 \\ 0 & 1 & -3 & 2 & 1 \\ 0 & 0 & -9x+3y+z & 4x-2y+s & 2x-y+t \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Ecrivons que la troisième ligne est composée d'éléments que nous égalerons à 0 pour obtenir le système homogène dont l'ensemble solution est W:

$$-9x+3y+z = 0, \quad 4x-2y+s = 0, \quad 2x-y+t = 0$$

En combinant les deux systèmes, nous obtenons le système homogène dont l'ensemble solution est $U \cap W$:

$$\begin{cases} -x+y+z = 0 \\ 4x-2y+s = 0 \\ -6x+y+t = 0 \\ -9x+3y+z = 0 \\ 4x-2y+s = 0 \\ 2x-y+t = 0 \end{cases} \quad \text{ou} \quad \begin{cases} -x+y+z = 0 \\ 2y+4z+s = 0 \\ -5y-6z+t = 0 \\ -6y-8z = 0 \\ 2y+4z+s = 0 \\ y+2z+t = 0 \end{cases}$$

$$\begin{cases} -x+y+z = 0 \\ 2y+4z+s = 0 \\ 8z+5s+2t = 0 \\ 4z+3s = 0 \\ s-2t = 0 \end{cases} \quad \text{ou} \quad \begin{cases} -x+y+z = 0 \\ 2y+4z+s = 0 \\ 8z+5s+2t = 0 \\ s-2t = 0 \end{cases}$$

Il y a une inconnue libre, qui est t ; donc $\dim(U \cap W) = 1$. Posons $t = 2$, on obtient la solution $x = 1$, $y = 4$, $z = -3$, $s = 4$, $t = 2$. Donc $\{(1, 4, -3, 4, 2)\}$ est une base de $U \cap W$.

VECTEURS COORDONNES

5.36. Trouver les coordonnées de v relativement à la base $\{(1, 1, 1) ; (1, 1, 0) ; (1, 0, 0)\}$ de $\mathbf{R}^3$ où 1) $v = (4, -3, 2)$, 2) $v = (a, b, c)$

Dans chaque cas écrivons que v est une combinaison linéaire des vecteurs de la base en utilisant les scalaires inconnus x, y et z:

$$v = x(1, 1, 1) + y(1, 1, 0) + z(1, 0, 0)$$

et résolvons en x, y, z pour obtenir les vecteurs (la solution est unique puisque les vecteurs de la base sont linéairement indépendants).

1) $$\begin{aligned}(4, -3, 2) &= x(1, 1, 1) + y(1, 1, 0) + z(1, 0, 0) \\ &= (x, x, x) + (y, y, 0) + (z, 0, 0) \\ &= (x+y+z, x+y, x)\end{aligned}$$

En égalant les composantes correspondantes les unes aux autres, on obtient le système

$$x+y+z = 4, \quad x+y = -3, \quad x = 2$$

Remplaçons x par 2 dans la seconde équation ; on obtient $y = -5$; remplaçons ensuite x par 2 et y par -5 dans la première équation ; on obtient $z = 7$. Ainsi $x = 2$, $y = -5$, $z = 7$ est la solution unique du système et les coordonnées du vecteur v relativement à la base donnée sont $[v] = (2, -5, 7)$.

2) $(a, b, c) = x(1, 1, 1) + y(1, 1, 0) + z(1, 0, 0) = (x+y+z, x+y, x)$

Alors $$x+y+z = a, \quad x+y = b, \quad x = c$$

d'où $x = c$, $y = b - c$, $z = a - b$. Ainsi $[v] = (c, b - c, a - c)$ c'est-à-dire $[(a, b, c)] = (c, b - c, a - b)$.

5.37. Soit V l'espace vectoriel des matrices 2×2 sur **R**. Trouver le vecteur coordonné de la matrice $A \in V$ relativement à la base

$$\left\{\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\right\} \quad \text{où} \quad A = \begin{pmatrix} 2 & 3 \\ 4 & -7 \end{pmatrix}$$

Posons A comme une combinaison linéaire des matrices de la base en utilisant les scalaires inconnus x, y, z, w:

$$\begin{aligned} A = \begin{pmatrix} 2 & 3 \\ 4 & -7 \end{pmatrix} &= x\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + y\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} + z\begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix} + w\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} x & x \\ x & x \end{pmatrix} + \begin{pmatrix} 0 & -y \\ y & 0 \end{pmatrix} + \begin{pmatrix} z & -z \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} w & 0 \\ 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} x+z+w & x-y-z \\ x+y & x \end{pmatrix} \end{aligned}$$

Ecrivons que les éléments correspondants sont égaux les uns aux autres. On obtient le système

$$x+z+w = 2, \quad x-y-z = 3, \quad x+y = 4, \quad x = -7$$

duquel on tire $x = -7$, $y = 11$, $z = -21$, $w = 30$. Ainsi $[A] = (-7, 11, -21, 30)$. (Remarquons que le vecteur coordonné de A doit être un vecteur de $\mathbf{R}^4$ puisque $\dim V = 4$).

5.38. Soit W l'espace vectoriel des matrices symétriques 2×2 sur **R** (voir problème 5.29). Trouver le vecteur coordonné de la matrice $A = \begin{pmatrix} 4 & -11 \\ -11 & -7 \end{pmatrix}$ relativement à la base $\left\{\begin{pmatrix} 1 & -2 \\ -2 & 1 \end{pmatrix}, \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}, \begin{pmatrix} 4 & -1 \\ -1 & -5 \end{pmatrix}\right\}$.

Ecrivons A sous forme de combinaison linéaire des matrices de la base en utilisant les scalaires inconnus x, y et z:

$$A = \begin{pmatrix} 4 & -11 \\ -11 & -7 \end{pmatrix} = x\begin{pmatrix} 1 & -2 \\ -2 & 1 \end{pmatrix} + y\begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix} + z\begin{pmatrix} 4 & -1 \\ -1 & -5 \end{pmatrix} = \begin{pmatrix} x+2y+4z & -2x+y-z \\ -2x+y-z & x+3y-5z \end{pmatrix}$$

Egalons les éléments correspondants entre eux; on obtient le système équivalent d'équations linéaires, et réduisons-le à sa forme échelonnée:

$$\begin{aligned} x+2y+4z &= 4 \\ -2x+y-z &= -11 \\ -2x+y-z &= -11 \\ x+3y-5z &= -7 \end{aligned} \quad \text{ou} \quad \begin{aligned} x+2y+4z &= 4 \\ 5y+7z &= -3 \\ y-9z &= -11 \end{aligned} \quad \text{ou} \quad \begin{aligned} x+2y+4z &= 4 \\ 5y+7z &= -3 \\ 52z &= 52 \end{aligned}$$

On obtient $z = 1$ d'après la troisième équation, puis $y = -2$ de la seconde équation et enfin $x = 4$ de la première équation. Donc la solution du système est $x = 4$, $y = -2$, $z = 1$. D'où $[A] = (4, -2, 1)$. (Puisque $\dim W = 3$ d'après le problème 5.29 le vecteur coordonné de A doit être un vecteur de $\mathbf{R}^3$).

5.39. Soit $\{e_1, e_2, e_3\}$ et $\{f_1, f_2, f_3\}$ des bases d'un espace vectoriel (de dimension 3). Supposons

$$\begin{aligned} e_1 &= a_1f_1 + a_2f_2 + a_3f_3 \\ e_2 &= b_1f_1 + b_2f_2 + b_3f_3 \\ e_3 &= c_1f_1 + c_2f_2 + c_3f_3 \end{aligned} \tag{1}$$

Soit P la matrice dont les lignes sont les coordonnées des vecteurs e_1, e_2, e_3 respectivement relativement à la base $\{f_i\}$.

$$P = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}$$

Montrer que, pour un vecteur quelconque $v \in V$, $[v]_e P = [v]_f$. C'est-à-dire qu'en multipliant le vecteur coordonné de v relativement à la base $\{e_i\}$ par la matrice P, on obtient le vecteur coordonné de v relativement à la base $\{f_i\}$. (La matrice P appelée la matrice de changement de base).

Supposons que $v = re_1 + se_2 + te_3$; alors $[v]_e = (r, s, t)$. En utilisant (*1*), nous avons

$$\begin{aligned} v &= r(a_1f_1 + a_2f_2 + a_3f_3) + s(b_1f_1 + b_2f_2 + b_3f_3) + t(c_1f_1 + c_2f_2 + c_3f_3) \\ &= (ra_1 + sb_1 + tc_1)f_1 + (ra_2 + sb_2 + tc_2)f_2 + (ra_3 + sb_3 + tc_3)f_3 \end{aligned}$$

D'où
$$[v]_f = (ra_1 + sb_1 + tc_1, ra_2 + sb_2 + tc_2, ra_3 + sb_3 + tc_3)$$

D'autre part
$$\begin{aligned} [v]_e P &= (r, s, t)\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} \\ &= (ra_1 + sb_1 + tc_1, ra_2 + sb_2 + tc_2, ra_3 + sb_3 + tc_3) \end{aligned}$$

En conséquence $[v]_e P = [v]_f$.

Remarque : Dans le chapitre 8, nous écrirons les vecteurs coordonnés comme des vecteurs colonnes, plutôt que de les écrire sous forme de vecteurs lignes. Alors d'après ce qui précède

$$Q[v]_e = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}\begin{pmatrix} r \\ s \\ t \end{pmatrix} = \begin{pmatrix} ra_1 + sb_1 + tc_1 \\ ra_2 + sb_2 + tc_2 \\ ra_3 + sb_3 + tc_3 \end{pmatrix} = [v]_f$$

où Q est la matrice dont les colonnes sont les vecteurs coordonnés de e_1, e_2 et e_3 respectivement, relativement à la base $\{f_i\}$. Remarquons que Q est la transposée de P et que Q est multipliée à droite par le vecteur $[v]_e$, tandis que P est multipliée à gauche par le vecteur $[v]_e$.

RANG D'UNE MATRICE

5.40. Trouver le rang de la matrice A où

(i) $A = \begin{pmatrix} 1 & 3 & 1 & -2 & -3 \\ 1 & 4 & 3 & -1 & -4 \\ 2 & 3 & -4 & -7 & -3 \\ 3 & 8 & 1 & -7 & -8 \end{pmatrix}$ (ii) $A = \begin{pmatrix} 1 & 2 & -3 \\ 2 & 1 & 0 \\ -2 & -1 & 3 \\ -1 & 4 & -2 \end{pmatrix}$ (iii) $A = \begin{pmatrix} 1 & 3 \\ 0 & -2 \\ 5 & -1 \\ -2 & 3 \end{pmatrix}$

1) Réduisons par des opérations élémentaires sur les lignes à la forme échelonnée :

$$\begin{pmatrix} 1 & 3 & 1 & -2 & -3 \\ 1 & 4 & 3 & -1 & -4 \\ 2 & 3 & -4 & -7 & -3 \\ 3 & 8 & 1 & -7 & -8 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & 3 & 1 & -2 & -3 \\ 0 & 1 & 2 & 1 & -1 \\ 0 & -3 & -6 & -3 & 3 \\ 0 & -1 & -2 & -1 & 1 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & 3 & 1 & -2 & -3 \\ 0 & 1 & 2 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Puisque la matrice échelonnée a deux lignes non nulles, rang $(A) = 2$.

2) Puisque le rang ligne est égal au rang colonne, il est facile de former la transposée de A et de réduire ligne à une forme échelonnée :

$$\begin{pmatrix} 1 & 2 & -2 & -1 \\ 2 & 1 & -1 & 4 \\ -3 & 0 & 3 & -2 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & 2 & -2 & -1 \\ 0 & -3 & 3 & 6 \\ 0 & 6 & -3 & -5 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & 2 & -2 & -1 \\ 0 & -3 & 3 & 6 \\ 0 & 0 & 3 & 7 \end{pmatrix}$$

Donc rang $(A) = 3$.

3) Les deux colonnes sont linéairement indépendantes, car l'une n'est pas un multiple de l'autre. Donc rang $(A) = 2$.

5.41. Soient A et B des matrices quelconques pour lesquelles le produit AB est défini. Montrer que rang $(AB) \leqslant$ rang (B) et rang $(AB) \leqslant$ rang (A).

D'après le problème 4.33 page 80, l'espace ligne de AB est contenu dans l'espace ligne de B ; d'où rang$(AB) \leqslant$ rang (B). De plus, d'après le problème 4.71 page 84, l'espace colonne de AB est contenu dans l'espace colonne de A ; d'où rang $(AB) \leqslant$ rang(A).

5.42. Soit A une n-matrice carrée. Montrer que A est inversible si et seulement si rang $(A) = n$.

Remarquons que les lignes de la n-matrice identité I_n sont linéairement indépendantes puisque I_n est sous sa forme échelonnée ; donc rang $(I_n) = n$. Maintenant si A est inversible, alors d'après le problème 3.36 page 57, A est équivalente ligne à I_n ; donc rang $(A) = n$. Mais si A n'est pas inversible, alors A est équivalente ligne à une matrice avec une ligne nulle ; donc rang $(A) < n$. Donc A est inversible si et seulement si rang $(A) = n$.

5.43. Soient $x_{i_1}, x_{i_2}, \ldots, x_{i_k}$ les inconnues libres d'un système homogène d'équations linéaires à n inconnues. Soit v_j la solution pour laquelle $x_{i_j} = 1$ et toutes les autres inconnues libres $= 0$. Montrer que les solutions $v_1, v_2, \ldots, v_k$ sont linéairement indépendantes.

Soit A la matrice dont les lignes sont les v_i respectivement. Echangeons la colonne 1 et la colonne i_1, puis la colonne 2 et la colonne $i_2, \ldots,$ et enfin la colonne k et la colonne i_k ; on obtient la matrice $k \times n$

$$B = (I, C) = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 & 0 & c_{1,k+1} & \ldots & c_{1n} \\ 0 & 1 & 0 & \ldots & 0 & 0 & c_{2,k+1} & \ldots & c_{2n} \\ \multicolumn{9}{c}{\dotfill} \\ 0 & 0 & 0 & \ldots & 0 & 1 & c_{k,k+1} & \ldots & c_{kn} \end{pmatrix}$$

La matrice précédente B est sous forme échelonnée, et donc ses lignes sont indépendantes, donc rang$(B) = k$. Puisque A et B sont équivalentes colonne, elles ont le même rang, c'est-à-dire rang $(A) = k$. Mais A a k lignes; d'où ces lignes, c'est-à-dire les v_i, sont linéairement indépendantes, comme nous l'affirmions.

PROBLEMES DIVERS

5.44. Le concept de dépendance linéaire peut être étendu à n'importe quel ensemble de vecteurs, fini ou infini, comme suit : l'ensemble des vecteurs $A = \{v_i\}$, est linéairement dépendant ssi, il existe des vecteurs $v_{i_1}, v_{i_2}, \ldots, v_{i_n} \in A$ et des scalaires $a_1, a_2, \ldots, a_n \in K$ non tous nuls, tels que :

$$a_1 v_{i_1} + a_2 v_{i_2} + \cdots + a_n v_{i_n} = 0$$

Sinon A est dit linéairement indépendant. Supposons que $A_1, A_2, \ldots,$ soient des ensembles de vecteurs linéairement indépendants et que $A_1 \subset A_2 \subset \ldots$ Montrons que l'union $A = A_1 \cup A_2 \cup \ldots$ est aussi linéairement indépendante.

Supposons A linéairement dépendant. Il existe alors des vecteurs $v_1, v_2, \ldots, v_n \in A$ et des scalaires $a_1, a_2, \ldots, a_n \in K$, non tous nuls, tels que

$$a_1 v_1 + a_2 v_2 + \cdots + a_n v_n = 0 \qquad (1)$$

Puisque $A = \cup A_i$ et $v_i \in A$, il existe des ensembles $A_{i_1}, \ldots, A_{i_n}$ tels que

$$v_1 \in A_{i_1}, \ v_2 \in A_{i_2}, \ \ldots, \ v_n \in A_{i_n}$$

Soit k l'indice maximum des ensembles A_{i_j} ; $k = \max(i_1, i_2, \ldots, i_n)$. Il s'ensuit, puisque $A_1 \subset A_2 \subset \ldots$, que chaque A_{i_j} est contenu dans A_k. Donc $v_1, v_2, \ldots, v_n \in A_k$ et donc d'après (1), A_k est linéairement dépendant, ce qui est en contradiction avec l'hypothèse. Ainsi A est linéairement indépendant.

5.45. Considérons une suite finie de vecteurs $S = \{v_1, v_2, \ldots, v_n\}$. Soit T la suite de vecteurs obtenue à partir de S par une "des opérations élémentaires" suivantes : (i) Echange de deux vecteurs, (ii) Multiplication d'un vecteur par un scalaire non nul, (iii) Addition d'un multiple d'un vecteur à un autre vecteur. Montrer que S et T engendrent le même espace W. Montrer aussi que T est indépendant si et seulement si S est indépendant.

Remarquons que, pour chacune des opérations, les vecteurs de T sont des combinaisons linéaires des vecteurs de S. D'autre part chacune des opérations a une inverse du même type (à démontrer) ; donc les vecteurs de S sont des combinaisons linéaires de vecteurs de T. D'où S et T engendrent le même espace W. Ainsi T est indépendant si et seulement si $\dim W = n$, ce qui est vérifié ssi S est aussi indépendant.

5.46. Soient $A = (a_{ij})$ et $B = (b_{ij})$ deux matrices $m \times n$ équivalentes ligne sur un corps K, et soient $v_1, v_2, \ldots, v_n$ des vecteurs quelconques d'un espace vectoriel V sur K. Soit

$$\begin{array}{llll} u_1 = a_{11}v_1 + a_{12}v_2 + \cdots + a_{1n}v_n & & w_1 = b_{11}v_1 + b_{12}v_2 + \cdots + b_{1n}v_n \\ u_2 = a_{21}v_1 + a_{22}v_2 + \cdots + a_{2n}v_n & & w_2 = b_{21}v_1 + b_{22}v_2 + \cdots + b_{2n}v_n \\ \cdots\cdots\cdots\cdots\cdots\cdots & & \cdots\cdots\cdots\cdots\cdots\cdots \\ u_m = a_{m1}v_1 + a_{m2}v_2 + \cdots + a_{mn}v_n & & w_m = b_{m1}v_1 + b_{m2}v_2 + \cdots + b_{mn}v_n \end{array}$$

Montrer que $\{u_i\}$ et $\{w_i\}$ engendrent le même espace.

Appliquer une "opération élémentaire" du problème précédent à $\{u_i\}$ est équivalent à appliquer une opération élémentaire sur les lignes de la matrice A. Puisque A et B sont équivalentes ligne, B peut être obtenue à partir de A par une suite d'opérations élémentaires sur les lignes ; donc $\{w_i\}$ peut être obtenu à partir de $\{u_i\}$ par la suite correspondante d'opérations. Donc $\{u_i\}$ et $\{w_i\}$ engendrent le même espace.

5.47. Soient $v_1, v_2, \ldots, v_n$ appartenant à un espace vectoriel V sur un corps K. Soit

$$\begin{aligned} w_1 &= a_{11}v_1 + a_{12}v_2 + \cdots + a_{1n}v_n \\ w_2 &= a_{21}v_1 + a_{22}v_2 + \cdots + a_{2n}v_n \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ w_n &= a_{n1}v_1 + a_{n2}v_2 + \cdots + a_{nn}v_n \end{aligned}$$

où $a_{ij} \in K$. Soit P la matrice carrée d'ordre n des coefficients, c'est-à-dire $P = (a_{ij})$.

(i) Supposons P inversible. Montrer que $\{w_i\}$ et $\{v_i\}$ engendrent le même espace ; donc $\{w_i\}$ est indépendant si et seulement si $\{v_i\}$ est indépendant.

(ii) Supposons P non inversible. Montrer que $\{w_i\}$ est dépendant.

(iii) Supposons $\{w_i\}$ indépendant. Montrer que P est inversible.

(i) Puisque P est inversible, P est équivalente ligne à la matrice identité I. Donc d'après le problème précédent $\{w_i\}$ et $\{v_i\}$ engendrent le même espace. Ainsi l'un de ces ensembles est indépendant si et seulement si l'autre l'est.

(ii) Puisque P n'est pas inversible, elle est équivalente ligne à une matrice ayant une ligne nulle. Ce qui veut dire que $\{w_i\}$ engendre un espace dont un ensemble générateur a moins de n éléments. Ainsi $\{w_i\}$ est dépendant.

(iii) Il s'agit de la réciproque du théorème (ii), et donc elle découle de (ii).

5.48. Supposons que V soit la somme directe de ses sous-espaces U et W, c'est-à-dire $V = U \oplus W$. Montrer que : (i) si $\{u_1, \ldots, u_m\} \subset U$ et $\{w_1, \ldots, w_n\} \subset W$ sont indépendants, alors $\{u_i, w_j\}$ est aussi indépendant. (ii) $\dim V = \dim U + \dim W$.

(i) Supposons $a_1u_1 + \cdots + a_mu_m + b_1w_1 + \cdots + b_nw_n = 0$, où a_i, b_j sont des scalaires. Alors

$$0 = (a_1u_1 + \cdots + a_mu_m) + (b_1w_1 + \cdots + b_nw_n) = 0 + 0$$

où 0, $a_1u_1 + \cdots + a_mu_m \in U$ et 0, $b_1w_1 + \cdots + b_nw_n \in W$. Puisqu'une telle somme pour 0 est unique, ceci conduit à :

$$a_1u_1 + \cdots + a_mu_m = 0, \quad b_1w_1 + \cdots + b_nw_n = 0$$

L'indépendance des u_i implique que les a_i sont tous nuls, et l'indépendance des w_j implique que les b_j sont tous nuls. En conséquence $\{u_i, w_j\}$ est indépendant.

(ii) **Méthode 1.** Puisque $V = U \oplus W$, nous avons $V = U + W$ et $U \cap W = \{0\}$. Ainsi, d'après le théorème 5.8 page 90,

$$\dim V = \dim U + \dim W - \dim(U \cap W) = \dim U + \dim W - 0 = \dim U + \dim W$$

Méthode 2. Supposons que $\{u_1, \ldots, u_r\}$ et $\{w_1, \ldots, w_s\}$ soient les bases de U et W respectivement. Puisqu'ils engendrent U et W respectivement, $\{u_i, w_j\}$ engendre $V = U + W$. D'autre part d'après (i) $\{u_i, w_j\}$ est indépendant. Donc $\{u_i, w_j\}$ est une base de V ; d'où $\dim V = \dim U + \dim W$.

5.49. Soit U un sous-espace de l'espace vectoriel V de dimension finie. Montrer qu'il existe un sous-espace W de V tel que $V = U \oplus W$.

Soit $\{u_1, u_2, \ldots, u_r\}$ une base de U. Puisque $\{u_i\}$ est linéairement indépendant, on peut l'étendre à une base de V, par exemple, $\{u_1, \ldots, u_r, w_1, \ldots, w_s\}$. Soit W l'espace engendré par $\{w_1, \ldots, w_s\}$. Puisque $\{u_i, w_j\}$ engendre V, $V = U + W$. D'autre part, $U \cap W = \{0\}$. (Problème 5-62). Donc $V = U \oplus W$.

5.50. Rappelons (page 65) que si K est un sous-corps d'un corps E (où E est une extension de K), alors E peut être considéré comme un espace vectoriel sur K. (i) Montrer que le corps des complexes **C** est un espace vectoriel de dimension 2 sur le corps des réels **R**. (ii) Montrer que le corps des réels **R** est un espace vectoriel de dimension infinie sur le corps des rationnels **Q**.

(i) Nous affirmons que $\{1, i\}$ est une base de **C** sur **R**. Car si $v \in \mathbf{C}$ alors $v = a + bi = a \cdot 1 + b \cdot i$ où $a, b \in \mathbf{R}$; c'est-à-dire que $\{1, i\}$ engendre **C** sur **R**. De plus, si $x \cdot 1 + y \cdot i = 0$ ou $x + yi = 0$, où $x, y \in \mathbf{R}$, alors $x = 0$ et $y = 0$; c'est-à-dire, $\{1, i\}$ est linéairement indépendant sur **R**. Ainsi $\{1, i\}$ est une base de **C** sur **R**, et donc **C** est de dimension 2 sur **R**.

(ii) Nous affirmons que, quel que soit n, $\{1, \pi, \pi^2, \ldots, \pi^n\}$ est linéairement indépendant sur **Q**. En effet supposons que $a_0 1 + a_1\pi + a_2\pi^2 + \cdots + a_n\pi^n = 0$, où les $a_i \in \mathbf{Q}$ et ne sont pas tous nuls. π est alors une racine du polynôme non nul suivant sur **Q** : $a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$. Mais l'on peut montrer que π est un nombre transcendant ; c'est-à-dire que π ne peut être la racine d'un quelconque polynôme non nul sur **Q**. En conséquence les $n + 1$ nombres réels $1, \pi, \pi^2, \ldots, \pi^n$ sont linéairement indépendants sur **Q**. Ainsi quel que soit n fini, **R** ne peut être de dimension n sur **Q**, c'est-à-dire **R** est de dimension infinie sur **Q**.

5.51. Soit K un sous-corps d'un corps L et L étant lui-même un sous-corps d'un corps E : $K \subset L \subset E$: Supposons que E soit de dimension n sur L et L soit de dimension m sur K. Montrer que E est de dimension mn sur K.

Supposons que $\{v_1, \ldots, v_n\}$ soit une base de E sur L et que $\{a_1, \ldots, a_m\}$ soit une base de L sur K. Nous affirmons que $\{a_iv_j : i = 1, \ldots, m,\ j = 1, \ldots, n\}$ est une base de E sur K. Remarquons que $\{a_iv_j\}$ contient mn éléments.

Soit w un élément quelconque de E. Puisque $\{v_1, \ldots, v_n\}$ engendre E sur L, W est une combinaison linéaire des v_i dont les coefficients appartiennent à L :

$$w = b_1v_1 + b_2v_2 + \cdots + b_nv_n, \qquad b_i \in L \tag{1}$$

Puisque $\{a_1, \ldots, a_m\}$ engendre L sur K, chaque $b_i \in L$ est une combinaison linéaire des a_j dont les coefficients appartiennent à K :

$$\begin{aligned} b_1 &= k_{11}a_1 + k_{12}a_2 + \cdots + k_{1m}a_m \\ b_2 &= k_{21}a_1 + k_{22}a_2 + \cdots + k_{2m}a_m \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ b_n &= k_{n1}a_1 + k_{n2}a_2 + \cdots + k_{nm}a_m \end{aligned}$$

où $k_{ij} \in K$. En remplaçant dans (*1*), on obtient

$$\begin{aligned} w &= (k_{11}a_1 + \cdots + k_{1m}a_m)v_1 + (k_{21}a_1 + \cdots + k_{2m}a_m)v_2 + \cdots + (k_{n1}a_1 + \cdots + k_{nm}a_m)v_n \\ &= k_{11}a_1v_1 + \cdots + k_{1m}a_mv_1 + k_{21}a_1v_2 + \cdots + k_{2m}a_mv_2 + \cdots + k_{n1}a_1v_n + \cdots + k_{nm}a_mv_n \\ &= \sum_{i,j} k_{ji}(a_iv_j) \end{aligned}$$

où $k_{ji} \in K$. w est donc une combinaison linéaire des a_iv_j dont les coefficients appartiennent à K ; donc $\{a_i, v_j\}$ engendre E sur K.

La démonstration sera complète si nous montrons que $\{a_iv_j\}$ est linéairement indépendant sur K. Supposons que, pour les scalaires $x_{ji} \in K$, $\sum_{i,j} x_{ji}(a_iv_j) = 0$; c'est-à-dire

$$(x_{11}a_1v_1 + x_{12}a_2v_1 + \cdots + x_{1m}a_mv_1) + \cdots + (x_{n1}a_1v_n + x_{n2}a_2v_n + \cdots + x_{nm}a_mv_n) = 0$$

ou
$$(x_{11}a_1 + x_{12}a_2 + \cdots + x_{1m}a_m)v_1 + \cdots + (x_{n1}a_1 + x_{n2}a_2 + \cdots + x_{nm}a_m)v_n = 0$$

Puisque $\{v_1, \ldots, v_n\}$ est linéairement indépendant sur L, et puisque les coefficients des v_i appartiennent à L, chaque coefficient doit être nul:

$$x_{11}a_1 + x_{12}a_2 + \cdots + x_{1m}a_m = 0, \quad \ldots, \quad x_{n1}a_1 + x_{n2}a_2 + \cdots + x_{nm}a_m = 0$$

Mais $\{a_1, \ldots, a_m\}$ est linéairement indépendant sur K ; ainsi puisque $x_{ji} \in K$,

$$x_{11} = 0, \; x_{12} = 0, \; \ldots, \; x_{1m} = 0, \; \ldots, \; x_{n1} = 0, \; x_{n2} = 0, \; \ldots, \; x_{nm} = 0$$

Donc $\{a_i, v_j\}$ est linéairement indépendant sur K, et le théorème est démontré.

PROBLEMES SUPPLEMENTAIRES

DEPENDANCE LINEAIRE

5.52. Déterminer si u et v sont linéairement dépendants où

(i) $u = (1, 2, 3, 4)$, $v = (4, 3, 2, 1)$

(ii) $u = (-1, 6, -12)$, $v = (\frac{1}{2}, -3, 6)$

(iii) $u = (0, 1)$, $v = (0, -3)$

(iv) $u = (1, 0, 0)$, $v = (0, 0, -3)$

(v) $u = \begin{pmatrix} 4 & -2 \\ 0 & -1 \end{pmatrix}$, $v = \begin{pmatrix} -2 & 1 \\ 0 & \frac{1}{2} \end{pmatrix}$

(vi) $u = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $v = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$

(vii) $u = -t^3 + \frac{1}{2}t^2 - 16$, $v = \frac{1}{2}t^3 - \frac{1}{4}t^2 + 8$

(viii) $u = t^3 + 3t + 4$, $v = t^3 + 4t + 3$

5.53. Déterminer si les vecteurs suivants de $\mathbf{R}^4$ sont linéairement dépendants ou indépendants : (i) $(1, 3, -1, 4)$, $(3, 8, -5, 7)$, $(2, 9, 4, 23)$; (ii) $(1, -2, 4, 1)$, $(2, 1, 0, -3)$, $(3, -6, 1, 4)$.

5.54. Soit V l'espace vectoriel des matrices 2×3 sur $\mathbf{R}$. Déterminer si les matrices $A, B, C \in V$ sont linéairement dépendantes ou indépendantes où :

(i) $A = \begin{pmatrix} 1 & -2 & 3 \\ 2 & 4 & -1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & -1 & 4 \\ 4 & 5 & -2 \end{pmatrix}$, $C = \begin{pmatrix} 3 & -8 & 7 \\ 2 & 10 & -1 \end{pmatrix}$

(ii) $A = \begin{pmatrix} 2 & 1 & -1 \\ 3 & -2 & 4 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 1 & -3 \\ -2 & 0 & 5 \end{pmatrix}$, $C = \begin{pmatrix} 4 & -1 & 2 \\ 1 & -2 & -3 \end{pmatrix}$

5.55. Soit V l'espace vectoriel des polynômes de degré $\leqslant 3$ sur $\mathbf{R}$. Déterminer si $u, v, w \in W$ sont linéairement dépendants ou indépendants avec :

(i) $u = t^3 - 4t^2 + 2t + 3$, $v = t^3 + 2t^2 + 4t - 1$, $w = 2t^3 - t^2 - 3t + 5$

(ii) $u = t^3 - 5t^2 - 2t + 3$, $v = t^3 - 4t^2 - 3t + 4$, $w = 2t^3 - 7t^2 - 7t + 9$

5.56. Soit V l'espace vectoriel des fonctions de **R** dans **R**. Montrer que $f, g, h \in V$ sont linéairement indépendantes avec (i) $f(t) = e^t$, $g(t) = \sin t$, $h(t) = t^2$; (ii) $f(t) = e^t$, $g(t) = e^{2t}$, $h(t) = t$; (iii) $f(t) = e^t$, $g(t) = \sin t$, $h(t) = \cos t$.

5.57. Montrer que (i) les vecteurs $(1 - i, i)$ et $(2, -1 + i)$ dans $\mathbf{C}^2$ sont linéairement dépendants sur le corps complexe **C**, mais sont linéairement indépendants sur le corps des réels **R** ; (ii) les vecteurs $(3 + \sqrt{2}, 1 + \sqrt{2})$ et $(7, 1 + 2\sqrt{2})$ dans $\mathbf{R}^2$ sont linéairement dépendants sur le corps des réels **R** mais sont linéairement indépendants sur le corps des rationnels **Q**.

5.58. Supposons u, v et w vecteurs linéairement indépendants. Montrer que :

(i) $u + v - 2w$, $u - v - w$ et $u + w$ sont linéairement indépendants.

(ii) $u + v - 3w$, $u + 3v - w$ et $v + w$ sont linéairement dépendants.

5.59. Démontrer ou trouver un contre-exemple : Si les vecteurs non nuls u, v et w sont linéairement dépendants, alors w est une combinaison linéaire de u et v.

5.60. Supposons $v_1, v_2, \ldots, v_n$ n vecteurs linéairement indépendants. Démontrer que

(i) $\{a_1 v_1, a_2 v_2, \ldots, a_n v_n\}$ est linéairement indépendant où chaque $a_i \neq 0$.

(ii) $\{v_1, \ldots, v_{i-1}, w, v_{i+1}, \ldots, v_n\}$ est linéairement indépendant où $w = b_1 v_1 + \cdots + b_i v_i + \cdots + b_n v_n$ et $b_i \neq 0$.

5.61. Soient $v = (a, b)$ et $w = (c, d)$ appartenant à K^2. Montrer que $\{v, w\}$ est linéairement dépendant si et seulement si $ad - bc = 0$.

5.62. Supposons que $\{u_1, \ldots, u_r, w_1, \ldots, w_s\}$ soit un sous-ensemble linéairement indépendant d'un espace vectoriel V. Montrer que $L(u_i) \cap L(w_j) = \{0\}$. (Rappelons que $L(u_i)$ est l'espace linéaire engendré par les u_i).

5.63. Supposons que $(a_{11}, \ldots, a_{1n}), \ldots, (a_{m1}, \ldots, a_{mn})$ soient des vecteurs linéairement indépendants de K^n, et supposons que $v_1, \ldots, v_n$ soient des vecteurs linéairement indépendants d'un espace vectoriel V sur K. Montrer que les vecteurs

$$w_1 = a_{11}v_1 + \cdots + a_{1n}v_n, \quad \ldots, \quad w_m = a_{m1}v_1 + \cdots + a_{mn}v_n$$

sont aussi linéairement indépendants.

BASE ET DIMENSION

5.64. Déterminer si oui ou non chacune des formes suivantes est une base de $\mathbf{R}^2$:

(i) $(1, 1)$ et $(3, 1)$

(ii) $(2, 1), (1, -1)$ et $(0, 2)$

(iii) $(0, 1)$ et $(0, -3)$

(iv) $(2, 1)$ et $(-3, 87)$

5.65. Déterminer si oui ou non les formes suivantes constituent des bases de $\mathbf{R}^3$:

(i) $(1, 2, -1)$ et $(0, 3, 1)$

(ii) $(2, 4, -3), (0, 1, 1)$ et $(0, 1, -1)$

(iii) $(1, 5, -6), (2, 1, 8), (3, -1, 4)$ et $(2, 1, 1)$

(iv) $(1, 3, -4), (1, 4, -3)$ et $(2, 3, -11)$

5.66. Trouver une base et la dimension du sous-espace W de $\mathbf{R}^4$ engendré par

(i) $(1, 4, -1, 3), (2, 1, -3, -1)$ et $(0, 2, 1, -5)$

(ii) $(1, -4, -2, 1), (1, -3, -1, 2)$ et $(3, -8, -2, 7)$

5.67. Soit V l'espace des matrices 2×2 sur **R** et soit W le sous-espace engendré par :

$$\begin{pmatrix} 1 & -5 \\ -4 & 2 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 \\ -1 & 5 \end{pmatrix}, \quad \begin{pmatrix} 2 & -4 \\ -5 & 7 \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} 1 & -7 \\ -5 & 1 \end{pmatrix}$$

Trouver une base et la dimension de W.

5.68. Soit W l'espace engendré par les polynômes

$$u = t^3 + 2t^2 - 2t + 1, \quad v = t^3 + 3t^2 - t + 4 \quad \text{et} \quad w = 2t^3 + t^2 - 7t - 7$$

Trouver une base et la dimension de W.

5.69. Trouver une base et la dimension de l'espace solution W de chacun des systèmes homogènes suivants :

$$\begin{array}{ccc} \begin{aligned} x+3y+2z&=0\\ x+5y+z&=0\\ 3x+5y+8z&=0 \end{aligned} & \begin{aligned} x-2y+7z&=0\\ 2x+3y-2z&=0\\ 2x-y+z&=0 \end{aligned} & \begin{aligned} x+4y+2z&=0\\ 2x+y+5z&=0 \end{aligned} \\ \text{(i)} & \text{(ii)} & \text{(iii)} \end{array}$$

5.70. Trouver une base et la dimension de l'espace solution W de chacun des systèmes homogènes :

$$\begin{array}{cc} \begin{aligned} x+2y-2z+2s-t&=0\\ x+2y-z+3s-2t&=0\\ 2x+4y-7z+s+t&=0 \end{aligned} & \begin{aligned} x+2y-z+3s-4t&=0\\ 2x+4y-2z-s+5t&=0\\ 2x+4y-2z+4s-2t&=0 \end{aligned} \\ \text{(i)} & \text{(ii)} \end{array}$$

5.71. Trouver un système homogène dont l'ensemble solution W est engendré par

$$\{(1,-2,0,3,-1),\ (2,-3,2,5,-3),\ (1,-2,1,2,-2)\}$$

5.72. Soient V et W les sous-espaces suivants de $\mathbf{R}^4$:

$$V=\{(a,b,c,d):\ b-2c+d=0\},\qquad W=\{(a,b,c,d):\ a=d,\ b=2c\}$$

Trouver une base et la dimension de 1) V, 2) W, 3) $V\cap W$.

5.73. Soit V l'espace vectoriel des polynômes en t de degré $\leqslant n$. Déterminer si oui ou non chacun des systèmes suivants est une base de V:

(i) $\{1,\ 1+t,\ 1+t+t^2,\ 1+t+t^2+t^3,\ \ldots,\ 1+t+t^2+\cdots+t^{n-1}+t^n\}$

(ii) $\{1+t,\ t+t^2,\ t^2+t^3,\ \ldots,\ t^{n-2}+t^{n-1},\ t^{n-1}+t^n\}$.

SOMMES ET INTERSECTIONS

5.74. Supposons que U et W soient des sous-espaces de dimension 2 de $\mathbf{R}^3$. Montrer que $U\cap W\neq\{0\}$.

5.75. Supposons que U et W soient des sous-espaces de V et que $\dim U=4$, $\dim W=5$ et $\dim V=7$. Trouver les dimensions possibles de $U\cap W$.

5.76. Soient U et W des sous-espaces de $\mathbf{R}^3$ pour lesquels $\dim U=1$, $\dim W=2$ et $U\not\subset W$. Montrer que $\mathbf{R}^3=U\oplus W$.

5.77. Soit U le sous-espace de $\mathbf{R}^5$ engendré par

$$\{(1,3,-3,-1,-4),\ (1,4,-1,-2,-2),\ (2,9,0,-5,-2)\}$$

et soit W le sous-espace engendré par

$$\{(1,6,2,-2,3),\ (2,8,-1,-6,-5),\ (1,3,-1,-5,-6)\}$$

Trouver 1) $\dim(U+W)$, 2) $\dim(U\cap W)$.

5.78. Soit V l'espace vectoriel des polynômes sur $\mathbf{R}$. Soient U et W les sous-espaces engendrés par

$\{t^3+4t^2-t+3,\ t^3+5t^2+5,\ 3t^3+10t^2-5t+5\}$ et $\{t^3+4t^2+6,\ t^3+2t^2-t+5,\ 2t^3+2t^2-3t+9\}$

respectivement. Trouver 1) $\dim(U+W)$, 2) $\dim(U\cap W)$.

5.79. Soit U le sous-espace de $\mathbf{R}^5$ engendré par

$$\{(1,-1,-1,-2,0),\ (1,-2,-2,0,-3),\ (1,-1,-2,-2,1)\}$$

et soit W le sous-espace engendré par

$$\{(1,-2,-3,0,-2),\ (1,-1,-3,2,-4),\ (1,-1,-2,2,-5)\}$$

1) Trouver deux systèmes homogènes dont les espaces solutions sont U et W respectivement.

2) Trouver une base et la dimension de $U\cap W$.

VECTEURS COORDONNES

5.80. Considérons la base suivante de $\mathbf{R}^3$: $\{(2, 1), (1, -1)\}$. Trouver le vecteur coordonné de $v \in \mathbf{R}^2$ relativement à la base précédente où 1) $v = (2, 3)$; 2) $v = (4, -1)$; 3) $(3, -3)$; 4) $v = (a, b)$.

5.81. Dans l'espace vectoriel V des polynômes en t de degré $\leqslant 3$, considérons la base suivante: $\{1, 1-t, (1-t)^2, (1-t)^3\}$. Trouver le vecteur coordonné de $v \in V$ relativement à la base précédente si 1) $v = 2 - 3t + t^2 + 2t^3$; 2) $v = 3 - 2t - t^2$; 3) $v = a + bt + ct^2 + dt^3$.

5.82. Dans l'espace vectoriel W des matrices symétriques 2×2 sur $\mathbf{R}$, considérons la base suivante :

$$\left\{\begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}, \begin{pmatrix} 4 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 3 & -2 \\ -2 & 1 \end{pmatrix}\right\}$$

Trouver le vecteur coordonné de la matrice $A \in W$ relativement à la base précédente si

(i) $A = \begin{pmatrix} 1 & -5 \\ -5 & 5 \end{pmatrix}$ (ii) $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$

5.83. Considérons les deux bases suivantes de $\mathbf{R}^3$:

$$\{e_1 = (1,1,1),\ e_2 = (0,2,3),\ e_3 = (0,2,-1)\} \quad \text{et} \quad \{f_1 = (1,1,0),\ f_2 = (1,-1,0),\ f_3 = (0,0,1)\}$$

1) Trouver le vecteur coordonné de $v = (3, 5, -2)$ relativement à chaque base : $[v]_e$ et $[v]_f$.
2) Trouver la matrice P dont les lignes sont respectivement les vecteurs coordonnés des e_i relativement à la base $\{f_1, f_2, f_3\}$.
3) Vérifier que $[v]_e\, P = [v]_f$.

5.84. Supposons que $\{e_1, \ldots, e_n\}$ et $\{f_1, \ldots, f_n\}$ soient des bases de l'espace vectoriel V (de dimension n). Soit P la matrice dont les lignes sont respectivement les vecteurs coordonnés des e relativement à la base $\{f_i\}$. Démontrer que pour un vecteur quelconque $v \in V$, $[v]_e\, P = [v]_f$. (Ce résultat est démontré dans le problème 5.39 dans le cas $n = 3$).

5.85. Montrer que le vecteur coordonné de $0 \in V$ relativement à une base quelconque de V est toujours le n-tuple nul $(0, 0, 0, \ldots, 0)$.

RANG D'UNE MATRICE

5.86. Trouver le rang de chacune des matrices :

$$\underset{\text{(i)}}{\begin{pmatrix} 1 & 3 & -2 & 5 & 4 \\ 1 & 4 & 1 & 3 & 5 \\ 1 & 4 & 2 & 4 & 3 \\ 2 & 7 & -3 & 6 & 13 \end{pmatrix}} \quad \underset{\text{(ii)}}{\begin{pmatrix} 1 & 2 & -3 & -2 & -3 \\ 1 & 3 & -2 & 0 & -4 \\ 3 & 8 & -7 & -2 & -11 \\ 2 & 1 & -9 & -10 & -3 \end{pmatrix}} \quad \underset{\text{(iii)}}{\begin{pmatrix} 1 & 1 & 2 \\ 4 & 5 & 5 \\ 5 & 8 & 1 \\ -1 & -2 & 2 \end{pmatrix}} \quad \underset{\text{(iv)}}{\begin{pmatrix} 2 & 1 \\ 3 & -7 \\ -6 & 1 \\ 5 & -8 \end{pmatrix}}$$

5.87. Soient A et B deux matrices arbitraires $m \times n$. Montrer que $\operatorname{rang}(A + B) \leqslant \operatorname{rang}(A) + \operatorname{rang}(B)$.

5.88. Donner des exemples de matrices 2×2 A et B telles que :
(i) $\operatorname{rang}(A+B) < \operatorname{rang}(A), \operatorname{rang}(B)$ (ii) $\operatorname{rang}(A + B) = \operatorname{rang}(A) = \operatorname{rang}(B)$
(iii) $\operatorname{rang}(A + B) > \operatorname{rang}(A), \operatorname{rang}(B)$.

PROBLEMES DIVERS

5.89. Soit W l'espace vectoriel des matrices symétriques 3×3 sur K. Montrer que $\dim W = 6$ en donnant une base de W (Rappelons que $A = (a_{ij})$ est symétrique ssi $a_{ij} = a_{ji}$).

5.90. Soit W l'espace vectoriel des matrices antisymétriques 3×3 sur K. Montrer que $\dim W = 3$ en donnant une base de W. (Rappelons que $A = (a_{ij})$ est antisymétrique ssi $a_{ij} = -a_{ji}$).

5.91. Supposons $\dim V = n$. Montrer qu'un ensemble générateur de n éléments est une base. (Comparer avec le théorème 5.6 (3) page 89).

5.92. Soient $t_1, t_2, \ldots, t_n$ des symboles et soit K un corps quelconque. Soit V l'ensemble des expressions $a_1 t_1 + a_2 t_2 + \cdots + a_n t_n$ où $a_i \in K$. Définissons une addition dans V par

$$(a_1t_1 + a_2t_2 + \cdots + a_nt_n) + (b_1t_1 + b_2t_2 + \cdots + b_nt_n)$$
$$= (a_1 + b_1)t_1 + (a_2 + b_2)t_2 + \cdots + (a_n + b_n)t_n$$

Définissons une multiplication scalaire dans V par

$$k(a_1t_1 + a_2t_2 + \cdots + a_nt_n) = ka_1t_1 + ka_2t_2 + \cdots + ka_nt_n$$

Montrer que V est un espace vectoriel sur K avec les opérations précédentes. Montrer aussi que $\{t_1, \ldots, t_n\}$ est une base de V où, pour $i = 1, \ldots, n$,

$$t_i = 0t_1 + \cdots + 0t_{i-1} + 1t_i + 0t_{i+1} + \cdots + 0t_n$$

5.93. Soit V un espace vectoriel de dimension n sur un corps K, et supposons que K soit un espace vectoriel de dimension m sur un sous-corps F. (D'où V peut être aussi regardé comme un espace vectoriel sur le sous-corps F). Démontrer que la dimension de V sur F est mn.

5.94. Soient U et W deux espaces vectoriels sur le même corps K, et soit V la somme directe extérieure de U et W (voir problème 4.45). Soient $\hat{U}$ et $\hat{W}$ les sous-espaces de V définis par $\hat{U} = \{(u, 0) : u \in U\}$ et $\hat{W} = \{(0, w) : w \in W\}$.

1) Montrer que U est isomorphe à $\hat{U}$ par la correspondance biunivoque $u \leftrightarrow (u, 0)$ et que W est isomorphe à $\hat{W}$ par la correspondance biunivoque $w \leftrightarrow (0, w)$.

2) Montrer que $\dim V = \dim U + \dim W$.

5.95. Supposons $V = U \oplus W$. Soit $\hat{V}$ le produit direct extérieur de U et W. Montrer que V est isomorphe à $\hat{V}$ par la correspondance biunivoque $v = u + w \leftrightarrow (u, w)$.

REPONSES AUX PROBLEMES SUPPLEMENTAIRES

5.52. 1) non 2) oui 3) oui 4) non 5) oui 6) non 7) oui 8) non.

5.53. 1) dépendant 2) indépendant.

5.54. 1) dépendant 2) indépendant.

5.55. 1) indépendant 2) dépendant.

5.57. (i) $(2, -1 + i) = (1 + i)(1 - i, i)$; (ii) $(7, 1 + 2\sqrt{2}) = (3 - \sqrt{2})(3 + \sqrt{2}, 1 + \sqrt{2})$.

5.59. L'énoncé est faux. Contre-exemple : $u = (1, 0), v = (2, 0)$ et $w = (1, 1)$ dans $\mathbf{R}^2$. Le lemme 5.2 suppose qu'un des vecteurs non nuls u, v, w est une combinaison linéaire des vecteurs précédents. Dans ce cas $v = 2u$.

5.64. 1) oui 2) non 3) non 4) oui.

5.65. 1) non 2) oui 3) non 4) non.

5.66. 1) $\dim W = 3$. 2) $\dim W = 2$.

5.67. $\dim W = 2$.

5.68. $\dim W = 2$.

5.69. 1) base $\{(7, -1, -2)\}$; $\dim W = 1$; 2) $\dim W = 0$; 3) base $\{(18, -1, -7)\}$; $\dim W = 1$.

5.70. (i) base $\{(2, -1, 0, 0, 0), (4, 0, 1, -1, 0), (3, 0, 1, 0, 1)\}$; $\dim W = 3$

(ii) base $\{(2, -1, 0, 0, 0), (1, 0, 1, 0, 0)\}$; $\dim W = 2$.

5.71. $\begin{cases} 5x + y - z - s = 0 \\ x + y - z - t = 0 \end{cases}$

5.72. (i) base $\{(1,0,0,0), (0,2,1,0), (0,-1,0,1)\}$; $\dim V = 3$.

(ii) base $\{(1,0,0,1), (0,2,1,0)\}$; $\dim W = 2$.

(iii) base $\{(0,2,1,0)\}$; $\dim(V \cap W) = 1$. *Rép.* $V \cap W$ doit satisfaire les trois conditions sur a, b, c et d.

5.73. 1) oui 2) non. Car $\dim V = n + 1$, mais l'ensemble contient seulement n éléments.

5.75. $\dim(U \cap W) = 2$, 3 ou 4.

5.77. $\dim(U + W) = 3$, $\dim(U \cap W) = 2$.

5.78. $\dim(U + W) = 3$, $\dim(U \cap W) = 1$.

5.79. (i) $\begin{cases} 3x + 4y - z \quad - t = 0 \\ 4x + 2y \quad + s \quad = 0 \end{cases}$, $\begin{cases} 4x + 2y \quad - s \quad = 0 \\ 9x + 2y + z \quad + t = 0 \end{cases}$

(ii) $\{(1, -2, -5, 0, 0), (0, 0, 1, 0, -1)\}$. $\dim(U \cap W) = 2$.

5.80. (i) $[v] = (5/3, -4/3)$, (ii) $[v] = (1, 2)$, (iii) $[v] = (0, 3)$, (iv) $[v] = ((a + b)/3, (a - 2b)/3)$.

5.81. (i) $[v] = (2, -5, 7, -2)$, (ii) $[v] = (0, 4, -1, 0)$, (iii) $[v] = (a + b + c + d, -b - 2c - 3d, c + 3d, -d)$.

5.82. (i) $[A] = (2, -1, 1)$, (ii) $[A] = (3, 1, -2)$.

5.83. (i) $[v]_e = (3, -1, 2)$, $[v]_f = (4, -1, -2)$; (ii) $P = \begin{pmatrix} 1 & 0 & 1 \\ 1 & -1 & 3 \\ 1 & -1 & -1 \end{pmatrix}$.

5.86. (i) 3, (ii) 2, (iii) 3, (iv) 2.

5.88. (i) $A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} -1 & -1 \\ 0 & 0 \end{pmatrix}$ (iii) $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$

(ii) $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}$

5.89. $\left\{ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right\}$

5.90. $\left\{ \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \right\}$

5.93. La démonstration est identique à celle donnée dans le problème 5.48 page 113 pour un cas particulier (lorsque v est une extension du corps K).

CHAPITRE 6

Applications linéaires

FONCTIONS

Soient A et B deux ensembles quelconques. Supposons qu'à chaque $a \in A$ on associe un élément unique de B ; l'ensemble de ces correspondances est appelé une fonction ou application de A dans B et on l'écrit

$$f: A \to B \qquad \text{ou} \qquad A \xrightarrow{f} B$$

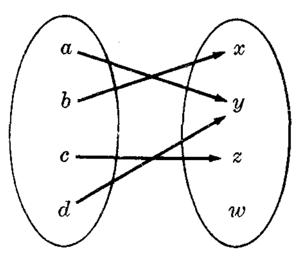
On écrit $f(a)$, lu "f de a", pour l'élément de B associé à $a \in A$ par f ; c'est la valeur donnée par f à a, où l'image de a par f. Si A' est un sous-ensemble quelconque de A, alors $f(A')$ est l'ensemble des images des éléments de A' ; et si B' est un sous-ensemble de B, alors $f^{-1}(B')$ est l'ensemble des éléments de A dont les images appartiennent à B':

$$f(A') = \{f(a) : a \in A'\} \qquad \text{et} \qquad f^{-1}(B') = \{a \in A : f(a) \in B'\}$$

Nous appelons $f(A')$ l'image de A' et $f^{-1}(B')$ l'image réciproque ou la préimage de B'. En particulier l'ensemble de toutes les images, c'est-à-dire $f(A)$, est appelé l'image de f. De plus A est appelé le domaine de l'application $f : A \to B$ et B est appelé son co-domaine.

A chaque fonction $f : A \to B$, il correspond le sous-ensemble de $A \times B$ défini par $\{a, f(a) : a \in A\}$. Nous appelons cet ensemble le graphe de f. Deux fonctions $f : A \to B$ et $g : A \to B$ sont dites égales par définition si $f(a) = g(a)$ pour chaque $a \in A$ et on écrit $f = g$, c'est-à-dire qu'elles ont le même graphe. Ainsi nous ne distinguerons pas une fonction de son graphe. La négation de la propriété $f = g$ est écrite $f \neq g$ et signifie qu'il existe un a de A tel que $f(a) \neq g(a)$.

Exemple 6.1 : Soient $A = \{a, b, c, d\}$ et $B = \{x, y, z, w\}$. Le diagramme suivant définit une application f de A dans B:



Ici $f(a) = y$, $f(b) = x$, $f(c) = z$ et $f(d) = y$. Aussi

$$f(\{a, b, d\}) = \{f(a), f(b), f(d)\} = \{y, x, y\} = \{x, y\}$$

L'image de f est l'ensemble $\{x, y, z\}$: $f(A) = \{x, y, z\}$.

Exemple 6.2 : Soit $f : \mathbf{R} \to \mathbf{R}$ l'application qui associe à chaque nombre réel x son carré x^2:

$$x \mapsto x^2 \qquad \text{ou} \qquad f(x) = x^2$$

Ici l'image de -3 est 9 et nous pouvons écrire $f(-3) = 9$.

Nous utilisons la flèche $\mapsto$ pour indiquer l'image d'un élément arbitraire $x \in A$ par la fonction $f : A \to B$ et on écrit

$$x \mapsto f(x)$$

comme dans l'exemple précédent.

Exemple 6.3 : Considérons la 2×3 matrice $A = \begin{pmatrix} 1 & -3 & 5 \\ 2 & 4 & -1 \end{pmatrix}$. Si nous écrivons les vecteurs dans $\mathbf{R}^3$ et $\mathbf{R}^2$ comme des vecteurs colonnes, alors A détermine l'application $T : \mathbf{R}^3 \to \mathbf{R}^2$ définie par

$$v \mapsto Av \quad \text{c'est-à-dire } T(v) = Av, \quad v \in \mathbf{R}^3$$

Donc si $v = \begin{pmatrix} 3 \\ 1 \\ -2 \end{pmatrix}$, alors $T(v) = Av = \begin{pmatrix} 1 & -3 & 5 \\ 2 & 4 & -1 \end{pmatrix}\begin{pmatrix} 3 \\ 1 \\ -2 \end{pmatrix} = \begin{pmatrix} -10 \\ 12 \end{pmatrix}$.

Remarque : Chaque matrice $m \times n$ A sur un corps K détermine l'application $T : K^n \to K^m$ définie par

$$v \mapsto Av$$

où les vecteurs dans K^n et K^m sont écrits comme des vecteurs colonnes. Pour la commodité nous noterons habituellement l'application précédente par A, qui est le même symbole utilisé pour la matrice.

Exemple 6.4 : Soit V l'espace vectoriel des polynômes de la variable t sur le corps des réels $\mathbf{R}$. Alors la dérivation définit une application $D : V \to V$, où à chaque polynôme $f \in V$ on associe $D(f) = df/dt$. Par exemple $D(3t^2 - 5t + 2) = 6t - 5$.

Exemple 6.5 : Soit V l'espace vectoriel des polynômes en t sur $\mathbf{R}$ (comme dans l'exemple précédent). Alors l'intégrale de 0 à 1 définit une application $\mathcal{J} : V \to \mathbf{R}$, où à chaque polynôme $f \in V$ on associe $\mathcal{J}(f) = \int_0^1 f(t)\,dt$. Par exemple

$$\mathcal{J}(3t^2 - 5t + 2) = \int_0^1 (3t^2 - 5t + 2)\,dt = \tfrac{1}{2}$$

Remarquons que cette application est une application de l'espace vectoriel V dans le corps des scalaires $\mathbf{R}$, tandis que l'application dans l'exemple précédent est une application de V dans lui-même.

Exemple 6.6 : Considérons les deux applications $f : A \to B$ et $g : B \to C$ représentées ci-dessous :

$$A \xrightarrow{f} B \xrightarrow{g} C$$

Soit $a \in A$: donc $f(a) \in B$, le domaine de g. Donc, nous pouvons obtenir l'image de $f(a)$ par l'application g, qui est $g(f(a))$. Cette application

$$a \mapsto g(f(a))$$

de A dans C est appelée l'application composée ou produit de f et de g et est notée $g \circ f$. En d'autres termes $(g \circ f) : A \to C$ est l'application définie par

$$(g \circ f)(a) = g(f(a))$$

Notre premier théorème montrera que la composition des applications est associative.

Théorème 6.1 : Soit $f : A \to B$, $g : B \to C$ et $h : C \to D$. Alors $h \circ (g \circ f) = (h \circ g) \circ f$.

Démontrons maintenant ce théorème. Si $a \in A$, alors

$$(h \circ (g \circ f))(a) = h((g \circ f)(a)) = h(g(f(a)))$$

et

$$((h \circ g) \circ f)(a) = (h \circ g)(f(a)) = h(g(f(a)))$$

D'où $(h \circ (g \circ f))(a) = ((h \circ g) \circ f)(a)$ quel que soit $a \in A$ et donc $h \circ (g \circ f) = (h \circ g) \circ f$.

Remarque : Soit $F : A \to B$. Certains auteurs écrivent aF à la place de $F(a)$ pour l'image de a par F. Avec cette notation la composition de fonctions $F : A \to B$ et $G : B \to C$ est notée par $F \circ G$ et non par $G \circ F$ comme on l'utilisera par la suite dans le texte.

Nous introduirons maintenant des types spéciaux d'applications.

Définition : Une application $f : A \to B$ est dite injective si des éléments distincts de A ont des images distinctes ; c'est-à-dire

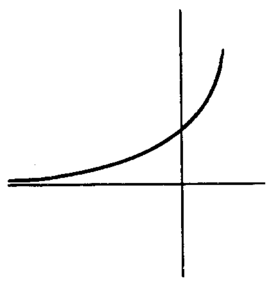
$$\text{si } a \neq a' \text{ implique } f(a) \neq f(a')$$

ou ce qui est équivalent $\quad$ si $f(a) \neq f(a')$ implique $a = a'$

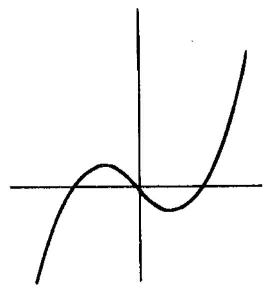
Définition : Une application $f : A \to B$ est dite surjective si chaque $b \in B$ est l'image d'au moins un $a \in A$.

Une application qui est à la fois injective et surjective est dite bijective.

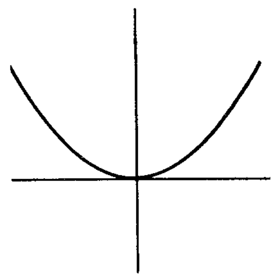
Exemple 6.7 : Soient $f : \mathbf{R} \to \mathbf{R}$, $g : \mathbf{R} \to \mathbf{R}$ et $k : \mathbf{R} \to \mathbf{R}$ définies par $f(x) = 2^x$, $g(x) = x^3 - x$ et $h(x) = x^2$, dont les graphes sont les suivants :



$f(x) = 2^x$



$g(x) = x^3 - x$



$h(x) = x^2$

L'application f est injective ; géométriquement ceci veut dire que toute droite horizontale ne contient pas plus d'un point de f. L'application g est surjective : géométriquement ceci veut dire que toute droite horizontale contient au moins un point de g. L'application h est ni injective, ni surjective, par exemple 2 et -2 ont la même image 4, et -16 n'est pas l'image d'un élément quelconque de $\mathbf{R}$.

Exemple 6.8 : Soit A un ensemble quelconque. L'application $f : A \to A$ définie par $f(a) = a$, c'est-à-dire qui associe à chaque élément de A lui-même, est appelée l'application identique de A et est notée par 1_A ou 1 ou I.

Exemple 6.9 : Soit $f : A \to B$. Nous appelons $g : B \to A$ l'inverse de f, écrite f^{-1}, si

$$f \circ g = 1_B \quad \text{et} \quad g \circ f = 1_A$$

Nous affirmons que f admet un inverse si et seulement si f est à la fois injective et surjective (Problème 6.9). Ainsi si $b \in B$ alors $f^{-1}(b) = a$, où a est l'élément unique de A pour lequel $f(a) = b$.

APPLICATIONS LINEAIRES

Soient U et V deux espaces vectoriels sur le même corps K. Une application $F : V \to U$ est appelée une application linéaire (ou transformation linéaire ou homomorphisme d'espace vectoriel) si elle satisfait aux deux conditions suivantes :

(1) Quels que soient $v, w \in V$, $F(v + w) = F(v) + F(w)$.

(2) Quel que soit $k \in K$ et quel que soit $v \in V$, $F(kv) = kF(v)$.

En d'autres termes $F : V \to U$ est linéaire si elle conserve les deux opérations de base d'un espace vectoriel, c'est-à-dire l'addition vectorielle et la multiplication par un scalaire.

En remplaçant k par 0 dans (2) on obtient $F(0) = 0$. En conséquence, chaque application linéaire donne du vecteur nul une image qui est le vecteur nul.

Maintenant quels que soient les scalaires a, $b \in K$ et quels que soient les vecteurs v, $w \in W$, nous obtenons en appliquant les deux conditions de linéarité à la fois.

$$F(av + bw) = F(av) + F(bw) = aF(v) + bF(w)$$

Plus généralement, quels que soient les scalaires $a_i \in K$ et quels que soient les vecteurs $v_i \in V$, nous obtenons la propriété de base des applications linéaires:

$$F(a_1v_1 + a_2v_2 + \cdots + a_nv_n) = a_1F(v_1) + a_2F(v_2) + \cdots + a_nF(v_n)$$

Remarquons que la condition $F(av + bw) = aF(v) + bF(w)$ caractérise complètement les applications linéaires et peut être utilisée comme la définition des applications linéaires.

Exemple 6.10 : Soit A une matrice quelconque $m \times n$ sur le corps K. Comme on l'a dit précédemment, A détermine une application $T : K^n \to K^m$ par la correspondance $v \mapsto Av$. (Ici les vecteurs de K^n et K^m sont écrits sous forme de vecteurs colonnes). Nous affirmons que T est une application linéaire. Car, d'après les propriétés des matrices,

$$T(v+w) = A(v+w) = Av + Aw = T(v) + T(w)$$

et

$$T(kv) = A(kv) = kAv = kT(v)$$

où v, $w \in K^n$ et $k \in K$.

Dans le chapitre suivant, nous montrerons que toute application linéaire d'un espace vectoriel fini dans un autre, peut être représentée par une application linéaire du type précédent.

Exemple 6.11 : Soit $F : \mathbf{R}^3 \to \mathbf{R}^3$ l'application "projection" dans le plan xy : $F(x, y, z) = (x, y, 0)$. Montrons que F est linéaire. Soient $v = (a, b, c)$ et $w = (a', b', c')$. Alors

$$\begin{aligned} F(v+w) &= F(a+a', b+b', c+c') = (a+a', b+b', 0) \\ &= (a, b, 0) + (a', b', 0) = F(v) + F(w) \end{aligned}$$

et quel que soit $k \in \mathbf{R}$,

$$F(kv) = F(ka, kb, kc) = (ka, kb, 0) = k(a, b, 0) = kF(v)$$

Donc F est linéaire

Exemple 6.12 : Soit $F : \mathbf{R}^2 \to \mathbf{R}^2$ l'application "translation" définie par $F(x, y) = (x + 1, y + 2)$. Remarquons que $F(0) = F(0, 0) = (1, 2) \neq 0$. En d'autres termes le vecteur nul n'est pas transformé en vecteur nul. Donc F n'est pas linéaire.

Exemple 6.13 : Soit $F : V \to U$ l'application qui associe $0 \in U$ à chaque $v \in V$. Alors quels que soient v, $w \in V$ et quel que soit $k \in K$ nous avons

$$F(v+w) = 0 = 0 + 0 = F(v) + F(w) \quad \text{et} \quad F(kv) = 0 = k0 = kF(v)$$

Ainsi F est linéaire. Nous appelons F l'application nulle et on la notera habituellement par 0.

Exemple 6.14 : Considérons l'application identité $I : V \to V$ qui transforme $v \in V$ en lui-même. Quels que soient v, $w \in V$ et quels que soient a, $b \in K$ nous avons

$$I(av + bw) = av + bw = aI(v) + bI(w)$$

Donc I est linéaire.

Exemple 6.15 : Soit V l'espace vectoriel des polynômes de la variable t sur le corps des réels $\mathbf{R}$. L'application différentielle $D : V \to V$ et l'application intégrale $\mathcal{J} : V \to R$ définies dans les exemples 6.4 et 6.5 sont linéaires. On démontre en effet en analyse que quels que soient u, $v \in V$ et $k \in \mathbf{R}$,

$$\frac{d(u+v)}{dt} = \frac{du}{dt} + \frac{dv}{dt} \quad \text{et} \quad \frac{d(ku)}{dt} = k\frac{du}{dt}$$

c'est-à-dire $D(u + v) = D(u) + D(v)$ et $D(ku) = k\,D(u)$; et aussi

$$\int_0^1 (u(t) + v(t))\,dt = \int_0^1 u(t)\,dt + \int_0^1 v(t)\,dt$$

$$\int_0^1 k\,u(t)\,dt = k\int_0^1 u(t)\,dt$$

c'est-à-dire $\mathcal{J}(u+v) = \mathcal{J}(u) + \mathcal{J}(v)$ et $\mathcal{J}(ku) = k\,\mathcal{J}(u)$.

Exemple 6.16 : Soit $F : V \to U$ une application linéaire qui est à la fois injective et surjective. Alors l'application inverse $F^{-1} : U \to V$ existe. Nous montrerons (Problème 6.17) que l'application inverse est aussi linéaire.

Lorsque nous avons discuté des coordonnées d'un vecteur relativement à une base donnée nous avons aussi introduit la notion de deux espaces isomorphes. Donnons maintenant une définition formelle.

Définition : Une application linéaire $F : V \to U$ est appelée un isomorphisme si elle est injective. Les espaces vectoriels V, U sont dits isomorphes s'il existe un isomorphisme surjectif de V **sur** U.

Exemple 6.17 : Soit V un espace vectoriel sur K de dimension n et soit $\{e_1, e_2, \ldots, e_n\}$ une base de V. Alors comme on l'a dit précédemment l'application $v \mapsto [v]_e$, c'est-à-dire l'application qui à tout $v \in V$ associe son vecteur coordonné relativement à la base $\{e_i\}$, est un isomorphe surjectif de V sur K^n.

Le théorème suivant nous donne une foule d'exemples d'applications linéaires ; en particulier il nous apprend qu'une application linéaire est complètement déterminée par les images des éléments d'une base.

Théorème 6.2 : Soient V et U deux espaces vectoriels sur un corps K. Soit $\{v_1, v_2, \ldots, v_n\}$ une base de V et soit $u_1, u_2, \ldots, u_n$ des vecteurs quelconques de U. Il existe alors une application linéaire unique $F : V \to U$ telle que
$F(v_1) = u_1, F(v_2) = u_2, \ldots, F(v_n) = u_n$.

Il faut bien remarquer que les vecteurs $u_1, \ldots, u_n$ du précédent théorème sont arbitraires ; ils peuvent être linéairement dépendants, ou égaux les uns aux autres.

NOYAU ET IMAGE D'UNE APPLICATION LINEAIRE

Nous commençons par les deux propriétés suivantes :

Définition : Soit $F : V \to U$ une application linéaire. L'image de F, que l'on note Im F, est l'ensemble des éléments-images de U:

$$\text{Im } F = \{u \in U : F(v) = u \text{ pour un certain } v \in V\}$$

Le noyau de F, noté Ker F, est l'ensemble des éléments de V dont l'image est $0 \in U$:

$$\text{Ker } F = \{v \in V ; F(v) = 0\}$$

Le théorème suivant se démontre facilement (Problème 6.22).

Théorème 6.3 : Soit $F : V \to U$ une application linéaire. L'image de F est un sous-espace de U et le noyau de F est un sous-espace de V.

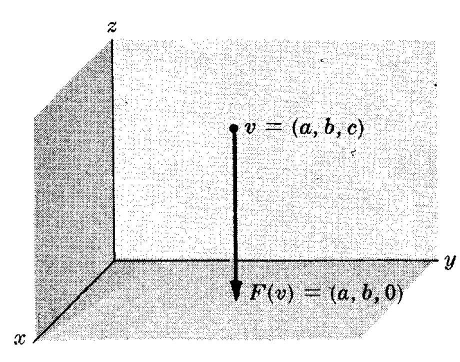
Exemple 6.18 : Soit $F : \mathbf{R}^3 \to \mathbf{R}^3$ l'application projection dans le plan xy : $F(x, y, z) = (x, y, 0)$. Il est clair que l'image de F est le plan xy entier:

$$\text{Im } F = \{(a, b, 0) : a, b \in \mathbf{R}\}$$

Remarquons que le noyau de F est l'axe des z:

$$\text{Ker } F = \{(0, 0, c) : c \in \mathbf{R}\}$$

puisque ces points et seulement ces points ont pour image le vecteur nul $0 = (0, 0, 0)$.



Supposons maintenant que les vecteurs $v_1, \ldots, v_n$ engendrent V et que $F : V \to U$ soit linéaire. Nous allons montrer que les vecteurs $F(v_1), \ldots, F(v_n) \in U$ engendrent ImF. Supposons en effet $u \in \mathrm{Im}F$; alors $F(v) = u$ pour un certain vecteur $v \in V$. Puisque les v_i engendrent V et puisque $v \in V$, il existe des scalaires $a_1, \ldots, a_n$ pour lesquels $v = a_1 v_1 + a_2 v_2 + \cdots + a_n v_n$. En conséquence

$$u = F(v) = F(a_1v_1 + a_2v_2 + \cdots + a_nv_n) = a_1 F(v_1) + a_2 F(v_2) + \cdots + a_n F(v_n)$$

et donc les vecteurs $F(v_1), \ldots, F(v_n)$ engendrent ImF.

Exemple 6.19 : Considérons une matrice arbitraire 4×3 A sur un corps K:

$$A = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \\ d_1 & d_2 & d_3 \end{pmatrix}$$

que nous considérons comme une application linéaire $A : K^3 \to K^4$. La base usuelle $\{e_1, e_2, e_3\}$ de K^3 engendre K^3 et donc les images des éléments de cette base Ae_1, Ae_2, Ae_3 par A engendrent l'image de A. Mais les vecteurs Ae_1, Ae_2, Ae_3 sont les colonnes de A:

$$Ae_1 = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \\ d_1 & d_2 & d_3 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} a_1 \\ b_1 \\ c_1 \\ d_1 \end{pmatrix}, \qquad Ae_2 = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \\ d_1 & d_2 & d_3 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a_2 \\ b_2 \\ c_2 \\ d_2 \end{pmatrix}$$

$$Ae_3 = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \\ d_1 & d_2 & d_3 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} a_3 \\ b_3 \\ c_3 \\ d_3 \end{pmatrix}$$

donc l'image de A est précisément l'espace colonne de A.

Il est à remarquer que si A est une matrice quelconque $m \times n$ sur K, qui peut être considérée comme une application linéaire $A : K^n \to K^m$, l'image de A est précisément l'espace colonne de A.

Nous n'avons cependant pas parlé de la notion de dimension attachée à une application linéaire $F : V \to U$. Dans le cas où V est de dimension finie, nous avons le théorème fondamental suivant :

Théorème 6.4 : Soit V de dimension finie et soit $F : V \to U$ une application linéaire. Alors :

$$\dim V = \dim(\mathrm{Ker}\, F) + \dim(\mathrm{Im}\, F)$$

C'est-à-dire, la somme des dimensions de l'image et du noyau d'une application linéaire est égale à la dimension de son domaine de définition. Cette formule est visiblement vérifiée pour l'application projection F dans l'exemple 6.18. Alors l'image (plan xy) et le noyau (axe des z) de F ont pour dimensions respectives 2 et 1, tandis que le domaine de F qui est $\mathbf{R}^3$ a pour dimension 3.

Remarque : Soit $F : V \to U$ une application linéaire. Le rang de F est défini comme étant la dimension de son image, et la nullité de F est définie comme étant la dimension de son noyau:

$$\mathrm{rang}\,(F) = \dim(\mathrm{Im}\, F) \quad \text{et} \quad \text{nullité}\,(F) = \dim(\mathrm{Ker}\, F).$$

Ainsi le théorème précédent vérifie la formule suivante pour F lorsque V est de dimension finie:

$$\mathrm{rang}\,(F) + \text{nullité}\,(F) = \dim V.$$

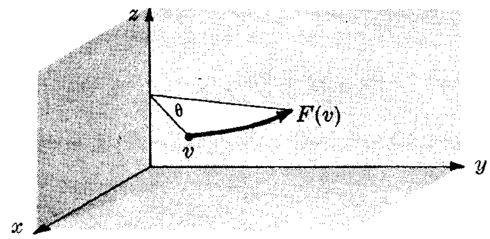
Rappelons qu'habituellement le rang d'une matrice A est défini comme la dimension de son espace colonne et de son espace ligne. Remarquons que si nous considérons maintenant A comme une application linéaire, les deux définitions se correspondent puisque l'image de A est précisément son espace colonne.

APPLICATIONS SINGULIERES ET NON SINGULIERES

Une application linéaire $F : V \to U$ est dite singulière s'il existe un vecteur non nul dont l'image par F est 0 ; c'est-à-dire s'il existe $v \in V$ avec $v \neq 0$ et pour lequel $F(v) = 0$. De même $F : V \to U$ est non singulière si seul le vecteur $0 \in V$ se transforme en $0 \in U$, ou ce qui est équivalent, si le noyau de F se réduit au seul vecteur nul : $\operatorname{Ker} F = \{0\}$.

Exemple 6.20 : Soit $F : \mathbf{R}^3 \to \mathbf{R}^3$ une application linéaire qui fait tourner un vecteur autour de l'axe des z de l'angle θ:

$$F(x, y, z) = (x \cos\theta - y \sin\theta,\ x \sin\theta + y \cos\theta,\ z)$$



Remarquons que seul le vecteur nul est transformé en le vecteur nul: donc F est non singulière.

Si l'application linéaire $F : V \to U$ est injective, alors seul $0 \in V$ peut être transformé en $0 \in U$ et donc F est non singulière. La réciproque est aussi vraie. En effet supposons que F soit non singulière et que $F(v) = F(w)$; alors $F(v - w) = F(v) - F(w) = 0$ d'où $v - w = 0$ ou $v = w$. Ainsi $F(v) = F(w)$ implique $v = w$; c'est-à-dire que F est injective. D'après la définition (page 125) une application linéaire injective est appelée un isomorphisme. Nous avons ainsi démontré

Théorème 6.5 : Une application linéaire $F : V \to U$ est un isomorphisme si et seulement si elle est non singulière

Nous remarquons que les applications non singulières peuvent aussi être caractérisées comme étant des applications qui transforment des ensembles indépendants en des ensembles indépendants (Problème 6.26).

APPLICATIONS LINEAIRES ET SYSTEMES D'EQUATIONS LINEAIRES

Considérons un système de m équations linéaires à n inconnues sur un corps K.

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

qui est équivalent à l'équation matricielle

$$Ax = b$$

où $A = (a_{ij})$ est la matrice des coefficients et $x = (x_i)$ et $b = (b_i)$ sont les vecteurs colonnes des inconnues et des constantes respectivement. La matrice A peut être aussi considérée comme une application linéaire

$$A : K^n \to K^m$$

En conséquence la solution de l'équation $Ax = b$ peut être considérée comme l'image réciproque ou préimage de $b \in K^m$ par l'application linéaire $A : K^n \to K^m$. De plus la solution du système homogène associé $Ax = 0$ peut être considérée comme le noyau de l'application linéaire $A : K^n \to K^m$.

D'après le théorème 6.4,

$$\dim(\operatorname{Ker} A) = \dim K^n - \dim(\operatorname{Im} A) = n - \operatorname{rang} A$$

Mais n est exactement le nombre d'inconnues dans le système homogène $Ax = 0$. Nous avons ainsi le théorème suivant sur les équations linéaires que l'on a rencontré dans le chapitre 5.

Théorème 5.11 : La dimension de l'espace solution W d'un système homogène d'équations linéaires $AX = 0$ est $n - r$, où n est le nombre d'inconnues et r le rang de la matrice coefficient A.

OPERATIONS SUR LES APPLICATIONS LINEAIRES

Nous pouvons combiner entre elles des applications linéaires de différentes manières, de façon à obtenir de nouvelles applications linéaires. Ces opérations sont très importantes et seront utilisées dans toute la suite.

Supposons $F : V \to U$ et $G : V \to U$ deux applications linéaires d'espaces vectoriels sur un corps K. Nous définissons la somme $F + G$ qui est l'application linéaire de V dans U qui associe $F(v) + G(v)$ à $v \in V$:

$$(F+G)(v) = F(v) + G(v)$$

De plus, pour tout scalaire $k \in K$, nous définissons le produit kF comme étant l'application de V dans U qui associe $kF(v)$ à $v \in V$:

$$(kF)(v) = kF(v)$$

Montrons que si F et G sont linéaires, alors $F + G$ et kF sont aussi linéaires. Nous avons, quels que soient $v, w \in V$ et quels que soient les scalaires $a, b \in K$,

$$\begin{aligned}(F+G)(av+bw) &= F(av+bw) + G(av+bw)\\ &= aF(v) + bF(w) + aG(v) + bG(w)\\ &= a(F(v)+G(v)) + b(F(w)+G(w))\\ &= a(F+G)(v) + b(F+G)(w)\end{aligned}$$

et

$$\begin{aligned}(kF)(av+bw) &= kF(av+bw) = k(aF(v)+bF(w))\\ &= akF(v) + bkF(w) = a(kF)(v) + b(kF)(w)\end{aligned}$$

Donc $F + G$ et kF sont linéaires

On a alors le théorème suivant :

Théorème 6.6 : Soient V et U deux espaces vectoriels sur le corps K. L'ensemble de toutes les applications linéaires de V dans U avec les opérations précédentes d'addition et de multiplication par un scalaire forme un espace vectoriel sur K.

L'espace dans le théorème précédent est habituellement noté par

$$\mathrm{Hom}\,(V, U)$$

Ici Hom provient du mot homomorphisme. Dans le cas où V et U sont de dimension finie, nous avons le théorème suivant

Théorème 6.7 : Supposons $\dim V = m$ et $\dim U = n$. Alors $\dim \mathrm{Hom}\,(V, U) = mn$.

Supposons maintenant que V, U et W soient des espaces vectoriels sur un même corps K, et que $F : V \to U$ et $G : U \to W$ soient des applications linéaires:

$$V \xrightarrow{F} U \xrightarrow{G} W$$

Rappelons que la fonction composé $G \circ F$ est l'application de V dans W définie par $(G \circ F)(v) = G(F(v))$. Montrons que $G \circ F$ est linéaire si F et G sont linéaires. Nous avons, quels que soient les vecteurs $v, w \in V$ et quels que soient les scalaires $a, b \in K$,

$$\begin{aligned}(G\circ F)(av+bw) &= G(F(av+bw)) = G(aF(v)+bF(w))\\ &= aG(F(v)) + bG(F(w)) = a(G\circ F)(v) + b(G\circ F)(w)\end{aligned}$$

Donc $G \circ F$ est linéaire.

Les propriétés de la composition d'applications linéaires, de l'addition et de la multiplication par un scalaire sont résumées ci-dessous :

Théorème 6.8 : Soient V, U et W des espaces vectoriels sur K. Soient F et F' des applications linéaires de V dans U et G, G' des applications linéaires de U dans W et soit $k \in K$. Alors

$$\text{(i)} \quad G \circ (F + F') = G \circ F + G \circ F'$$
$$\text{(ii)} \quad (G + G') \circ F = G \circ F + G' \circ F$$
$$\text{(iii)} \quad k(G \circ F) = (kG) \circ F = G \circ (kF).$$

ALGEBRE DES OPERATEURS LINEAIRES

Soit V un espace vectoriel sur un corps K. Considérons le cas particulier d'applications linéaires $T : V \to V$ c'est-à-dire de V dans lui-même. Elles sont aussi appelées opérateurs linéaires ou transformations linéaires sur V. Nous écrirons $A(V)$, au lieu de $\mathrm{Hom}(V, V)$, pour l'espace déterminé par toutes ces applications particulières.

D'après le théorème 6.6, $A(V)$ est un espace vectoriel sur K ; sa dimension est n^2, si V a pour dimension n. Si T, $S \in A(V)$ alors la composée $S \circ T$ existe et est aussi une application linéaire de V dans lui-même, c'est-à-dire $S \circ T \in A(V)$. Nous avons ainsi une "multiplication" définie dans $A(V)$. (Nous écrirons ST au lieu de $S \circ T$ dans l'espace $(A(V))$.

Remarquons qu'une algèbre A sur un corps K est un espace vectoriel sur K dans lequel l'opération de multiplication est définie et satisfait, pour chaque F, G, $H \in A$ et pour chaque $k \in K$,

1) $F(G + H) = FG + FH.$
2) $(G + H)F = GF + HF$
3) $k(GF) = (kG)F = G(kF).$

Si l'associativité reste aussi vraie pour la multiplication, c'est-à-dire si pour tout F, G, $H \in A$,

4) $(FG)H = F(GH)$

alors l'algèbre A est dite associative. Ainsi d'après les théorèmes 6.8 et 6.1 $A(V)$ est une algèbre associative sur K par rapport à la composition des applications ; on l'appelle souvent l'algèbre des opérateurs linéaires sur V.

Remarquons que l'application identique $I : V \to V$ appartient à $A(V)$. Quel que soit $T \in A(V)$ nous avons $TI = IT = T$. Nous pouvons aussi former des puissances de T en utilisant la notation $T^2 = T \circ T$; $T^3 = T \circ T \circ T$. De plus, quel que soit le polynôme

$$p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n, \qquad a_i \in K$$

on peut former l'opérateur $p(T)$ défini par

$$p(T) = a_0 I + a_1 T + a_2 T^2 + \cdots + a_n T^n$$

(Pour un scalaire $k \in K$, l'opérateur kI est fréquemment noté simplement k). En particulier, si $p(T) = 0$, l'application nulle, on dit que T est un zéro du polynôme $p(x)$.

Exemple 6.21 : Soit $T : \mathbf{R}^3 \to \mathbf{R}^3$ définie par $T(x, y, z) = (0, x, y)$. Si (a, b, c) est un élément quelconque de $\mathbf{R}^3$ alors

$$(T + I)(a, b, c) = (0, a, b) + (a, b, c) = (a, a + b, b + c)$$

et $$T^3(a, b, c) = T^2(0, a, b) = T(0, 0, a) = (0, 0, 0)$$

Donc $T^3 = 0$ est l'application nulle de V dans lui-même. En d'autres termes, T est un zéro du polynôme $p(x) = x^3$.

OPERATEURS INVERSIBLES

Un opérateur linéaire $T : V \to V$ est dit inversible s'il a un inverse, c'est-à-dire s'il existe $T^{-1} \in A(V)$ tel que $TT^{-1} = T^{-1}T = I$.

Or T est inversible si et seulement si T est bijectif. Ainsi, en particulier, si T est inversible, alors seulement $0 \in V$ peut être transformé en lui-même ; c'est-à-dire T est non singulier. D'autre part, supposons T non singulier c'est-à-dire Ker $T = \{0\}$. Rappelons (page 127) que T est aussi injectif. De plus en supposant V de dimension finie, nous avons, d'après le théorème 6.4,

$$\begin{aligned}\dim V &= \dim(\operatorname{Im} T) + \dim(\operatorname{Ker} T) = \dim(\operatorname{Im} T) + \dim(\{0\}) \\ &= \dim(\operatorname{Im} T) + 0 = \dim(\operatorname{Im} T)\end{aligned}$$

Donc Im $T = V$, c'est-à-dire que l'image de T est V ; donc T est surjectif. T est donc à la fois injectif et surjectif et donc est inversible. Nous avons démontré le

Théorème 6.9 : Un opérateur linéaire $T : V \to V$ sur un espace vectoriel de dimension finie est inversible si et seulement s'il est non singulier.

Exemple 6.22 : Soit T l'opérateur de $\mathbf{R}^2$ défini par $T(x, y) = (y, 2x - y)$. Le noyau de T est $\{(0,0)\}$; donc T est non singulier et, d'après le théorème précédent, il est inversible. Il s'agit de trouver l'expression de T^{-1}. Supposons que (s, t) soit l'image de (x, y) par T ; donc (x, y) est l'image de (s, t) par T^{-1} : $T(x, y) = (s, t)$ et $T^{-1}(s, t) = (x, y)$. Nous avons

$$T(x, y) = (y, 2x - y) = (s, t) \quad \text{et donc} \quad y = s, \; 2x - y = t$$

En résolvant en x et y, on obtient $x = \frac{1}{2}s + \frac{1}{2}t$, $y = s$. Donc T^{-1} est donné par la formule $T^{-1}(s, t) = \left(\frac{1}{2}s + \frac{1}{2}t, s\right)$.

La dimension de V doit être finie dans le théorème précédent, comme on peut le voir dans l'exemple suivant.

Exemple 6.23 : Soit V l'espace vectoriel des polynômes sur K et soit T l'opérateur de V défini par

$$T(a_0 + a_1t + \cdots + a_nt^n) = a_0t + a_1t^2 + \cdots + a_nt^{n+1}$$

c'est-à-dire que T augmente la puissance de chaque terme de 1. T est une application linéaire non singulière. Toutefois, T n'est pas surjectif et donc non inversible

Donnons maintenant une application importante du théorème précédent aux systèmes d'équations linéaires sur le corps K. Considérons un système qui a le même nombre d'équations que d'inconnues, c'est-à-dire n. Nous pouvons représenter ce système par l'équation matricielle

$$Ax = b \tag{*}$$

où A est une matrice carrée d'ordre n sur K que nous pouvons considérer comme un opérateur linéaire de K^n. Supposons que la matrice A soit non singulière ; c'est-à-dire que l'équation matricielle $Ax = 0$ admette seulement la solution nulle. Alors. d'après le théorème 6.9, l'application linéaire A est injective et surjective. Ceci veut dire que le système (*) a une solution unique quel que soit $b \in K^n$. D'autre part, supposons que la matrice A soit singulière, c'est-à-dire que l'équation matricielle $Ax = 0$ admette une solution non nulle ; l'application linéaire A n'est alors pas surjective. Ce qui veut dire qu'il existe $b \in K^n$ pour lequel (*) n'a pas de solution. De plus, si une solution existe, elle n'est pas unique. Nous avons ainsi démontré le résultat fondamental suivant :

Théorème 6.10 : Considérons le système suivant d'équations linéaires :

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots & \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n \end{aligned}$$

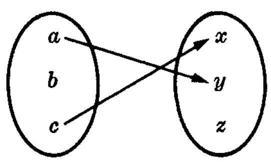
(i) Si le système homogène associé admet seulement la solution nulle, le système précédent a une solution unique quelles que soient les valeurs de b_i.

(ii) Si le système homogène associé admet une solution non nulle alors (i) il y a des valeurs de b_i pour lesquelles le système précédent n'a pas de solution ; (ii) cependant s'il existe une solution du système précédent, cette solution n'est pas unique.

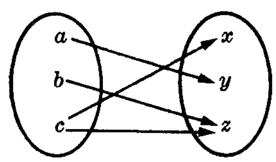
PROBLEMES RESOLUS

APPLICATIONS

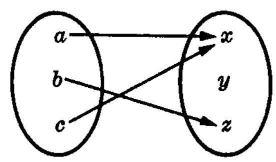
6.1. Rechercher si oui ou non chacun des diagrammes suivants définit une application de $A = \{a, b, c\}$ dans $B = \{x, y, z\}$.



(i)



(ii)



(iii)

(i) Non : L'élément $b \in A$ n'a pas d'image.

(ii) Non : Deux éléments x et z ayant le même antécédent $c \in A$.

(iii) Oui

6.2. Utiliser une formule pour définir chacune des fonctions suivantes de R dans R.

(i) A chaque nombre f associe son cube.

(ii) A chaque nombre g associe le nombre 5.

(iii) A chaque nombre positif h associe son carré, et à chaque nombre non positif h associe le nombre 6.

Trouver aussi les valeurs prises par chaque fonction pour 4, -2 et 0.

(i) Puisque f associe à un nombre quelconque x, son cube x^3, nous pouvons définir f par $f(x) = x^3$. Ainsi

$$f(4) = 4^3 = 64, \quad f(-2) = (-2)^3 = -8, \quad f(0) = 0^3 = 0$$

(ii) Puisque g associe 5 à tout nombre x, nous pouvons définir g par $g(x) = 5$. Ainsi les valeurs prises par g pour chacun des nombres 4, -2 et 0 sont égales à 5 :

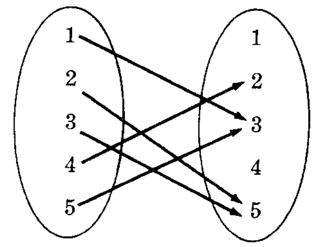
$$g(4) = 5, \quad g(-2) = 5, \quad g(0) = 5$$

(iii) On utilise deux formules distinctes pour définir h comme suit :

$$h(x) = \begin{cases} x^2 & \text{si } x > 0 \\ 6 & \text{si } x \leqslant 0 \end{cases}$$

Puisque $4 > 0$, $h(4) = 4^2 = 16$. D'autre part, -2, $0 \leqslant 0$ et donc $h(-2) = 6$, $h(0) = 6$.

6.3. Soit $A = \{1, 2, 3, 4, 5\}$ et soit $f : A \to A$ l'application définie par le diagramme de droite. (i) Trouver l'image de f. (ii) Trouver le graphe de f.



(i) L'image $f(A)$ de l'application f contient toutes les images des éléments de A. Or seulement 2, 3 et 5 sont les images d'éléments de A, d'où $f(A) = \{2, 3, 5\}$.

(ii) Le graphe de f contient tous les couples $(a, f(a))$ où $a \in A$. Or $f(1) = 3$, $f(2) = 5$, $f(3) = 5$, $f(4) = 2$, $f(5) = 3$, d'où le graphe de

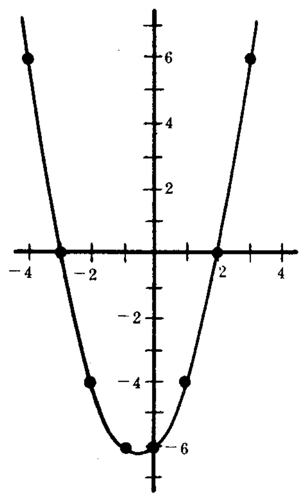
$$f = \{(1,3), (2,5), (3,5), (4,2), (5,3)\}$$

6.4. Tracer le graphe de (i) $f(x) = x^2 + x - 6$, (ii) $g(x) = x^3 - 3x^2 - x + 3$.

Remarquons qu'il s'agit de "fonctions polynômes". Dans chaque cas dressons une table de valeurs qui à x fait correspondre $f(x)$. Plaçons les points dans un repère cartésien, et joignons ces différents points obtenus par une courbe continue régulière.

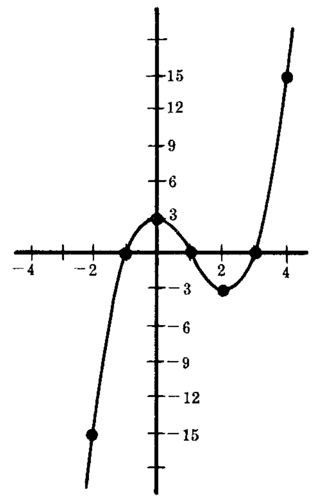
(i)

x	$f(x)$
-4	6
-3	0
-2	-4
-1	-6
0	-6
1	-4
2	0
3	6

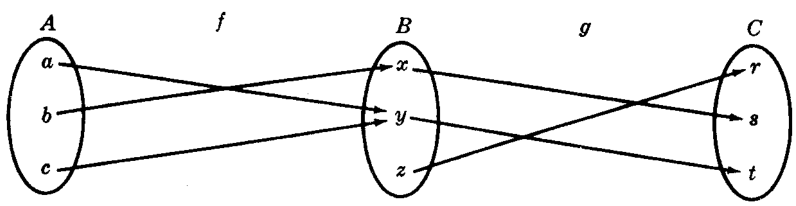


(ii)

x	$g(x)$
-2	-15
-1	0
0	3
1	0
2	-3
3	0
4	15



6.5. Soient les applications $f : A \to B$ et $g : B \to C$ définies par le diagramme



(i) Trouver l'application composée $(g \circ f) : A \to C$. (ii) Trouver l'image de chacune des applications f, g et $g \circ f$.

(i) Utilisons pour calculer la définition de la composition des applications :

$$(g \circ f)(a) = g(f(a)) = g(y) = t$$
$$(g \circ f)(b) = g(f(b)) = g(x) = s$$
$$(g \circ f)(c) = g(f(c)) = g(y) = t$$

Remarquons que nous arrivons à la même réponse que celle obtenue à l'aide des flèches dans le diagramme:

$$a \to y \to t, \qquad b \to x \to s, \qquad c \to y \to t$$

(ii) D'après le diagramme les images par l'application f sont x et y et les images par g sont r, s et t, donc

$$\text{image de } f = \{x, y\} \text{ et image de } g = \{r, s, t\}.$$

D'après (i) les images par l'application $g \circ f$ sont t et s ; donc l'image de $g \circ f = \{s, t\}$. Remarquons que les images de g et $g \circ f$ sont différentes.

6.6. Soient les applications f et g définies par $f(x) = 2x + 1$ et $g(x) = x^2 - 2$. (i) Trouver $(g \circ f)$ (4) et $(f \circ g)(4)$. (ii) Trouver les formules définissant la composition des applications $g \circ f$ et $f \circ g$.

(i) $f(4) = 2 \cdot 4 + 1 = 9$. Donc $(g \circ f)(4) = g(f(4)) = g(9) = g(9) = 9^2 - 2 = 79$

$g(4) = 4^2 - 2 = 14$. Donc $(f \circ g)(4) = f(g(4)) = f(14) = 2 \cdot 14 + 1 = 29$.

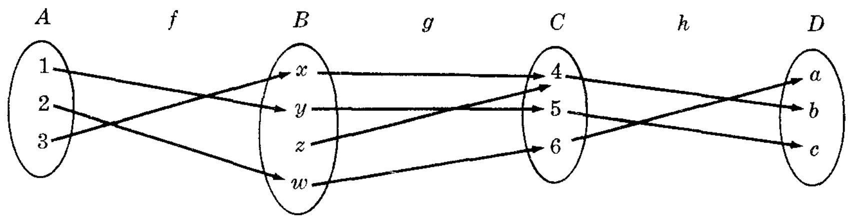
(ii) Calculons la formule donnant $g \circ f$ comme suit :

$$(g \circ f)(x) = g(f(x)) = g(2x + 1) = (2x + 1)^2 - 2 = 4x^2 + 4x - 1.$$

Remarquons que la même réponse peut être trouvée en écrivant $y = f(x) = 2x + 1$ et $z = g(y) = y^2 - 2$, et en éliminant y : $z = y^2 - 2 = (2x + 1)^2 - 2 = 4x^2 + 4x - 1$.

$(f \circ g)(x) = f(g(x)) = f(x^2 - 2) = 2(x^2 - 2) + 1 = 2x^2 - 3$. Remarquons que $f \circ g \neq g \circ f$.

6.7. Soient les applications $f : A \to B$, $g : B \to C$ et $h : C \to D$ définies par le diagramme



Déterminer si chacune de ces applications (i) est injective, (ii) surjective, (iii) a un inverse.

(i) L'application $f : A \to B$ est injective puisque chaque élément de A a une image distincte. L'application $g : B \to C$ n'est pas injective puisque x et z ont tous deux pour images le même élément 4. L'application $h : C \to D$ est injective.

(ii) L'application $f : A \to B$ n'est pas surjective puisque $z \in B$ n'est pas l'image d'un élément quelconque de A. L'application $g : B \to C$ est surjective puisque chaque élément de C est l'image d'un élément de B. L'application $h : C \to D$ est aussi surjective.

(iii) Une application admet une application inverse si et seulement si elle est à la fois injective et surjective. Ainsi seulement h a un inverse.

6.8. Supposons $f : A \to B$ et $g : B \to C$; donc l'application composée $(g \circ f) : A \to C$ existe. Démontrer que : (i) Si f et g sont injectives, alors $g \circ f$ est injective. (ii) si f et g sont surjectives alors $g \circ f$ est surjective. (iii) Si $g \circ f$ est injective alors f est injective. (iv) Si $g \circ f$ est surjective alors g est surjective.

(i) Supposons $(g \circ f)(x) = (g \circ f)(y)$. Alors $g(f(x)) = g(f(y))$. Puisque g est injective $f(x) = f(y)$. Puisque f est injective $x = y$. Nous avons démontré que $(g \circ f)(x) = (g \circ f)(y)$ implique $x = y$; donc $g \circ f$ est injective.

(ii) Supposons $c \in C$. Puisque g est surjective, il existe $b \in B$ pour lequel $g(b) = c$. Puisque f est surjective, il existe $a \in A$ pour lequel $f(a) = b$. Donc $(g \circ f)(a) = g(f(a)) = g(b) = c$, donc $g \circ f$ est surjective.

(iii) Supposons f non injective. Il existe alors des éléments distincts $x, y \in A$ pour lesquels $f(x) = f(y)$. Ainsi $(g \circ f)(x) = g(f(x)) = g(f(y)) = (g \circ f)(y)$ donc $g \circ f$ n'est pas injective. En conséquence si $g \circ f$ est injective alors f doit être aussi injective.

(iv) Si $a \in A$, alors $(g \circ f)(a) = g(f(a)) \in g(B)$; donc $(g \circ f)(A) \subset g(B)$. Supposons g non surjective. Alors $g(B)$ est contenu dans C et ainsi $(g \circ f)(A)$ est contenu dans C ; donc $g \circ f$ n'est pas surjective. En conséquence si $g \circ f$ est surjective, alors g doit être surjective.

6.9. Démontrer qu'une application $f : A \to B$ admet une inverse si et seulement si elle est injective et surjective.

Supposons que f admette une inverse ; c'est-à-dire qu'il existe une fonction $f^{-1} : B \to A$ pour laquelle $f^{-1} \circ f = 1_A$ et $f \circ f^{-1} = 1_B$. Puisque 1_A est injective, f est injective d'après le problème 6.8 (iii) et puisque 1_B est surjective, f est surjective d'après le problème 6.8 (iv). Donc f est à la fois injective et surjective.

Supposons maintenant que f soit à la fois injective et surjective. Alors chaque $b \in B$ est l'image d'un élément unique de A, appelé $\hat{b}$. Ainsi si $f(a) = b$ alors $a = \hat{b}$; donc $f(\hat{b}) = b$. Soit g l'application de B dans A définie par $b \mapsto \hat{b}$. Nous avons

(1) $(g \circ f)(a) = g(f(a)) = g(b) = \hat{b} = a$, pour chaque $a \in A$; donc $g \circ f = 1_A$

(2) $(f \circ g)(b) = f(g(b)) = f(\hat{b}) = b$, pour chaque $b \in B$; donc $f \circ g = 1_B$.

En conséquence, f a une inverse. Son inverse est l'application g.

6.10. Soit $f : \mathbf{R} \to \mathbf{R}$ définie par $f(x) = 2x - 3$. f est injective et surjective ; donc f a une application inverse f^{-1}. Trouver la formule définissant f^{-1}.

Soit y l'image de x par l'application f : $y = f(x) = 2x - 3$. En conséquence x sera l'image de y par l'application inverse f^{-1}. Résolvons l'équation précédente en exprimant x en fonction de y : $x = (y + 3)/2$. Cette formule définit la fonction inverse qui est $f^{-1}(y) = (y + 3)/2$.

APPLICATIONS LINEAIRES

6.11. Montrer que les applications F suivantes sont linéaires :

(i) $F : \mathbf{R}^2 \to \mathbf{R}^2$ définie par $F(x, y) = (x + y, x)$.

(ii) $F : \mathbf{R}^3 \to \mathbf{R}$ définie par $F(x, y, z) = 2x - 3y + 4z$.

(i) Soit $v = (a, b)$ et $w = (a', b')$; donc

$$v + w = (a + a', b + b') \quad \text{et} \quad kv = (ka, kb), \quad k \in \mathbf{R}$$

Nous avons $F(v) = (a + b, a)$ et $F(w) = (a' + b', a')$. Ainsi

$$\begin{aligned} F(v + w) &= F(a + a', b + b') = (a + a' + b + b', a + a') \\ &= (a + b, a) + (a' + b', a') = F(v) + F(w) \end{aligned}$$

et
$$F(kv) = F(ka, kb) = (ka + kb, ka) = k(a + b, a) = kF(v)$$

Puisque v, w et k étaient arbitraires, F est linéaire.

(ii) Soient $v = (a, b, c)$ et $w = (a', b', c')$; donc

$$v + w = (a + a', b + b', c + c') \quad \text{et} \quad kv = (ka, kb, kc), \quad k \in \mathbf{R}$$

Nous avons $F(v) = 2a - 3b + 4c$ et $F(w) = 2a' - 3b' + 4c'$. Ainsi

$$\begin{aligned} F(v + w) &= F(a + a', b + b', c + c') = 2(a + a') - 3(b + b') + 4(c + c') \\ &= (2a - 3b + 4c) + (2a' - 3b' + 4c') = F(v) + F(w) \end{aligned}$$

et
$$F(kv) = F(ka, kb, kc) = 2ka - 3kb + 4kc = k(2a - 3b + 4c) = kF(v)$$

En conséquence, F est linéaire.

6.12. Montrer que les applications suivantes F ne sont pas linéaires :

(i) $F : \mathbf{R}^2 \to \mathbf{R}$ définie par $F(x, y) = xy$.

(ii) $F : \mathbf{R}^2 \to \mathbf{R}^3$ définie par $F(x, y) = (x + 1, 2y, x + y)$.

(iii) $F : \mathbf{R}^3 \to \mathbf{R}^2$ définie par $F(x, y, z) = (|x|, 0)$.

(i) Soient $v = (1, 2)$ et $w = (3, 4)$; alors $v + w = (4, 6)$.

Nous avons $F(v) = 1 \cdot 2 = 2$ et $F(w) = 3 \cdot 4 = 12$. Donc

$$F(v+w) = F(4,6) = 4 \cdot 6 = 24 \neq F(v) + F(w)$$

En conséquence, F n'est pas linéaire.

(ii) Puisque $F(0, 0) = (1, 0, 0) \neq (0, 0, 0)$, F ne peut être linéaire.

(iii) Soit $v = (1, 2, 3)$ et $k = -3$, donc $kv = (-3, -6, -9)$.

Nous avons $F(v) = (1, 0)$ et donc aussi $kF(v) = -3(1, 0) = (-3, 0)$. Alors

$$F(kv) = F(-3, -6, -9) = (3, 0) \neq kF(v)$$

et donc F n'est pas linéaire.

6.13. Soit V l'espace vectoriel des matrices carrées n sur K. Soit M une matrice arbitraire de V. Soit $T : V \to V$ définie par $T(A) = AM + MA$ où $A \in V$. Montrer que T est linéaire.

Quels que soient $A, B \in V$ et quel que soit $k \in K$, nous avons

$$\begin{aligned} T(A+B) &= (A+B)M + M(A+B) = AM + BM + MA + MB \\ &= (AM+MA) + (BM+MB) = T(A) + T(B) \end{aligned}$$

et $\quad T(kA) = (kA)M + M(kA) = k(AM) + k(MA) = k(AM+MA) = kT(A)$

Donc, T est linéaire.

6.14. Démontrer le théorème 6.2 : Soient V et U deux espaces vectoriels sur le corps K. Soit $\{v_1, v_2, \ldots, v_n\}$ une base de V et soit $u_1, \ldots, u_n$ des vecteurs quelconques de U. Il existe alors une application linéaire unique $F : V \to U$, telle que $F(v_1) = u_1, F(v_2) = u_2, \ldots, F(v_n) = u_n$.

Il y a trois étapes dans la démonstration de ce théorème : (1) Définir une application $F : V \to U$ telle que $F(v_i) = u_i$, $i = 1, \ldots, n$. (2) Montrer que F est linéaire. (3) Montrer que F est unique.

Première étape. Soit $v \in V$. Puisque $\{v_1, v_2, \ldots, v_n\}$ est une base de V, il existe des scalaires uniques. $a_1, \ldots, a_n \in K$ pour lesquels $v = a_1v_1 + a_2v_2 + \cdots + a_nv_n$. Nous définissons $F : V \to U$ par $F(v) = a_1u_1 + a_2u_2 + \cdots + a_nu_n$.(Puisque les a_i sont uniques, l'application F est bien définie). Maintenant, pour $i = 1, \ldots, n$,

$$v_i = 0v_1 + \cdots + 1v_i + \cdots + 0v_n$$

Donc $\quad F(v_i) = 0u_1 + \cdots + 1u_i + \cdots + 0u_n = u_i$

Ainsi la première étape de la démonstration est complète.

Deuxième étape : Supposons $v = a_1v_1 + a_2v_2 + \cdots + a_nv_n$ et $w = b_1v_1 + b_2v_2 + \cdots + b_nv_n$. Alors

$$v + w = (a_1+b_1)v_1 + (a_2+b_2)v_2 + \cdots + (a_n+b_n)v_n$$

et pour k quelconque $k \in K$, $kv = ka_1v_1 + ka_2v_2 + \cdots + ka_nv_n$. D'après la définition de l'application F,

$$F(v) = a_1u_1 + a_2u_2 + \cdots + a_nu_n \quad \text{et} \quad F(w) = b_1u_1 + b_2v_2 + \cdots + b_nu_n$$

Donc
$$\begin{aligned} F(v+w) &= (a_1+b_1)u_1 + (a_2+b_2)u_2 + \cdots + (a_n+b_n)u_n \\ &= (a_1u_1 + a_2u_2 + \cdots + a_nu_n) + (b_1u_1 + b_2u_2 + \cdots + b_nu_n) \\ &= F(v) + F(w) \end{aligned}$$

et $\quad F(kv) = k(a_1u_1 + a_2u_2 + \cdots + a_nu_n) = kF(v)$

D'où F est bien linéaire.

Troisième étape Supposons maintenant $G : V \to U$ linéaire et $G(v_i) = u_i$, $i = 1, \ldots, n$. Si $v = a_1v_1 + a_2v_2 + \cdots + a_nv_n$, alors

$$\begin{aligned} G(v) &= G(a_1v_1 + a_2v_2 + \cdots + a_nv_n) = a_1G(v_1) + a_2G(v_2) + \cdots + a_nG(v_n) \\ &= a_1u_1 + a_2u_2 + \cdots + a_nu_n = F(v) \end{aligned}$$

Puisque $G(v) = F(v)$ quel que soit $v \in V$, $G = F$. Ainsi F est unique et le théorème est démontré.

6.15. Soit $T : \mathbf{R}^2 \to \mathbf{R}$ l'application linéaire pour laquelle

$$T(1,1) = 3 \quad \text{et} \quad T(0,1) = -2 \tag{1}$$

(Puisque $\{(1,1), (0,1)\}$ est une base de $\mathbf{R}^2$, une telle application linéaire existe et est unique d'après le théorème 6.2). Trouver $T(a, b)$.

Ecrivons d'abord (a, b) comme une combinaison linéaire de $(1, 1)$ et $(0, 1)$ en utilisant les inconnues scalaires x et y:

$$(a, b) = x(1, 1) + y(0, 1) \tag{2}$$

Alors $\quad (a, b) = (x, x) + (0, y) = (x, x + y) \quad$ et $\quad$ donc $x = a,\ x + y = b$

Résolvons en x et y en fonction de a et b, nous obtenons

$$x = a \quad \text{et} \quad y = b - a \tag{3}$$

En utilisant maintenant (1) et (2)

$$T(a, b) = T(x(1, 1) + y(0, 1)) = xT(1, 1) + yT(0, 1) = 3x - 2y$$

Finalement en utilisant (3) nous avons

$$T(a, b) = 3x - 2y = 3(a) - 2(b - a) = 5a - 2b.$$

6.16. Soit $T : V \to U$ une application linéaire, et supposons que $v_1, v_2, \ldots, v_n \in V$ aient la propriété que leurs images $T(v_1), \ldots, T(v_n)$ soient linéairement indépendantes. Montrer que les vecteurs $v_1, \ldots, v_n$ sont aussi linéairement indépendants.

Supposons que, pour les scalaires $a_1, \ldots, a_n$, $a_1 v_1 + a_2 v_2 + \cdots + a_n v_n = 0$. Alors

$$0 = T(0) = T(a_1v_1 + a_2v_2 + \cdots + a_nv_n) = a_1T(v_1) + a_2T(v_2) + \cdots + a_nT(v_n)$$

Puisque les $T(v_i)$ sont linéairement indépendants, tous les $a_i = 0$. Ainsi les vecteurs $v_1, \ldots, v_n$ sont linéairement indépendants.

6.17. Supposons que l'application linéaire $F : V \to U$ soit injective et surjective. Montrer que l'application inverse $F^{-1} : U \to V$ est aussi linéaire.

Supposons $u, u' \in U$. Puisque F est injective et surjective, il existe des vecteurs uniques $v, v' \in V$ pour lesquels $F(v) = u$ et $F(v') = u'$. Puisque F est linéaire, nous avons aussi

$$F(v + v') = F(v) + F(v') = u + u' \quad \text{et} \quad F(kv) = kF(v) = ku$$

Par définition de l'application inverse $F^{-1}(u) = v$, $F^{-1}(u') = v'$, $F^{-1}(u + u') = v + v'$ et $F^{-1}(ku) = kv$. Alors

$$F^{-1}(u + u') = v + v' = F^{-1}(u) + F^{-1}(u') \quad \text{et} \quad F^{-1}(ku) = kv = kF^{-1}(u)$$

et donc F^{-1} est linéaire.

IMAGE ET NOYAU D'APPLICATIONS LINEAIRES

6.18. Soit $F : \mathbf{R}^4 \to \mathbf{R}^3$ une application linéaire définie par

$$F(x, y, s, t) = (x - y + s + t,\ x + 2s - t,\ x + y + 3s - 3t)$$

Trouver une base et la dimension de 1) image U de F, 2) noyau W de F.

1) Les images des générateurs suivants de $\mathbf{R}^4$ engendrent l'image U de F:

$$F(1, 0, 0, 0) = (1, 1, 1) \qquad F(0, 0, 1, 0) = (1, 2, 3)$$
$$F(0, 1, 0, 0) = (-1, 0, 1) \qquad F(0, 0, 0, 1) = (1, -1, -3)$$

Formons la matrice dont les lignes sont les générateurs de U et réduisons-la par des opérations élémentaires sur les lignes à sa forme échelonnée:

$$\begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ 1 & 2 & 3 \\ 1 & -1 & -3 \end{pmatrix} \text{ donne } \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \\ 0 & -2 & -4 \end{pmatrix} \text{ donne } \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Ainsi $\{(1, 1, 1), (0, 1, 2)\}$ est une base de U ; donc dim $U = 2$.

2) Cherchons l'ensemble des (x, y, s, t) tel que $F(x, y, z, t) = (0, 0, 0)$ c'est-à-dire

$$F(x, y, s, t) = (x - y + s + t,\ x + 2s - t,\ x + y + 3s - 3t) = (0, 0, 0)$$

En égalant les unes aux autres les composantes correspondantes, on obtient le système homogène suivant dont l'espace solution est le noyau W de F:

$$\begin{aligned} x - y + s + t &= 0 \\ x + 2s - t &= 0 \\ x + y + 3s - 3t &= 0 \end{aligned} \quad \text{ou} \quad \begin{aligned} x - y + s + t &= 0 \\ y + s - 2t &= 0 \\ 2y + 2s - 4t &= 0 \end{aligned} \quad \text{ou} \quad \begin{aligned} x - y + s + t &= 0 \\ y + s - 2t &= 0 \end{aligned}$$

Les inconnues libres sont s et t ; donc $\dim W = 2$. Posons

a) $s = -1,\ t = 0$ on obtient la solution $(2, 1, -1, 0)$

b) $s = 0,\ t = 1$ on obtient la solution $(1, 2, 0, 1)$

Ainsi $\{(2, 1, -1, 0)$ et $(1, 2, 0, 1)\}$ est une base de W. (Remarquons que $\dim U + \dim W = 2 + 2 = 4$ qui est la dimension du domaine $\mathbf{R}^4$ de F).

6.19. Soit $T : \mathbf{R}^3 \to \mathbf{R}^3$ l'application linéaire définie par

$$T(x, y, z) = (x + 2y - z,\ y + z,\ x + y - 2z)$$

Trouver une base et la dimension de 1) image U de T, 2) noyau W de T.

1) Les images des générateurs de $\mathbf{R}^3$ engendrent l'image U de T:

$$T(1, 0, 0) = (1, 0, 1), \quad T(0, 1, 0) = (2, 1, 1), \quad T(0, 0, 1) = (-1, 1, -2)$$

Formons la matrice dont les lignes sont les générateurs de U et réduisons-la par des opérations élémentaires sur les lignes à sa forme échelonnée :

$$\begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 1 \\ -1 & 1 & -2 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 1 & -1 \end{pmatrix} \text{ d'où } \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

Ainsi $\{(1, 0, 1), (0, 1, -1)\}$ est une base de U et donc $\dim U = 2$.

2) Cherchons l'ensemble des (x, y, z) tels que $T(x, y, z) = (0, 0, 0)$ c'est-à-dire

$$T(x, y, z) = (x + 2y - z,\ y + z,\ x)$$

En égalant les composantes correspondantes les unes aux autres, on forme le système homogène dont l'espace solution est le noyau W de T:

$$\begin{aligned} x + 2y - z &= 0 \\ y + z &= 0 \\ x + y - 2z &= 0 \end{aligned} \quad \text{ou} \quad \begin{aligned} x + 2y - z &= 0 \\ y + z &= 0 \\ -y - z &= 0 \end{aligned} \quad \text{ou} \quad \begin{aligned} x + 2y - z &= 0 \\ y + z &= 0 \end{aligned}$$

La seule inconnue libre est z ; donc $\dim W = 1$. Soit $z = 1$; alors $y = -1$ et $x = 3$. Donc $\{(3, -1, 1)\}$ est une base de W. (Remarquons que $\dim U + \dim W = 2 + 1 = 3$, qui est la dimension du domaine $\mathbf{R}^3$ de T).

6.20. Trouver une application linéaire $F : \mathbf{R}^3 \to \mathbf{R}^4$ dont l'image est engendrée par $(1, 2, 0, -4)$ et $(2, 0, -1, -3)$.

Méthode 1 .

Considérons la base habituelle de $\mathbf{R}^3$: $e_1 = (1, 0, 0), e_2 = (0, 1, 0), e_3 = (0, 0, 1)$. Soient $F(e_1) = (1, 2, 0, -4)$, $F(e_2) = (2, 0, -1, -3)$ et $F(e_3) = (0, 0, 0, 0)$. D'après le théorème 6.2, une telle application linéaire F existe et est unique. De plus, l'image de F est engendrée par les $F(e_i)$; donc F a la propriété requise. Nous trouvons la formule générale pour $F(x, y, z)$:

$$\begin{aligned} F(x, y, z) &= F(xe_1 + ye_2 + ze_3) = xF(e_1) + yF(e_2) + zF(e_3) \\ &= x(1, 2, 0, -4) + y(2, 0, -1, -3) + z(0, 0, 0, 0) \\ &= (x + 2y,\ 2x,\ -y,\ -4x - 3y) \end{aligned}$$

Méthode 2.

Formons une matrice 4 x 3 dont les colonnes sont uniquement les vecteurs donnés ; c'est-à-dire

$$A = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 0 & 0 \\ 0 & -1 & -1 \\ -4 & -3 & -3 \end{pmatrix}$$

Rappelons que A détermine une application linéaire $A : \mathbf{R}^3 \to \mathbf{R}^4$ dont l'image est engendrée par les colonnes A. Ainsi A satisfait à la condition requise.

6.21. Soit V l'espace vectoriel des matrices 2 x 2 sur R et soit $M = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$. Soit $F : V \to V$ l'application linéaire définie par $F(A) = AM - MA$. Trouver une base et la dimension du noyau W de F.

Nous cherchons l'ensemble des $\begin{pmatrix} x & y \\ s & t \end{pmatrix}$ tels que $F\begin{pmatrix} x & y \\ s & t \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.

$$\begin{aligned} F\begin{pmatrix} x & y \\ s & t \end{pmatrix} &= \begin{pmatrix} x & y \\ s & t \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix} - \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} x & y \\ s & t \end{pmatrix} \\ &= \begin{pmatrix} x & 2x+3y \\ s & 2s+3t \end{pmatrix} - \begin{pmatrix} x+2s & y+2t \\ 3s & 3t \end{pmatrix} \\ &= \begin{pmatrix} -2s & 2x+2y-2t \\ -2s & 2s \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \end{aligned}$$

Ainsi $$\begin{aligned} 2x + 2y - 2t &= 0 \\ 2s &= 0 \end{aligned} \quad \text{ou} \quad \begin{aligned} x + y - t &= 0 \\ s &= 0 \end{aligned}$$

Les inconnues libres sont y et t ; donc dim $W = 2$. Pour obtenir une base de W posons

a) $y = -1$, $t = 0$, on obtient la solution $x = 1$, $y = -1$, $s = 0$, $t = 0$.

b) $y = 0$, $t = 1$, on obtient la solution $x = 1$, $y = 0$, $s = 0$, $t = 1$.

Donc : $\left\{ \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix} ; \begin{pmatrix} 1, & 0 \\ 0 & 1 \end{pmatrix} \right\}$ est une base de W.

6.22. Démontrer le théorème 6.3 : Soit $F : V \to U$ une application linéaire. Alors 1) l'image de F est un sous-espace de U et 2) le noyau de F est un sous-espace de V.

1) Puisque $F(0) = 0$, $0 \in \operatorname{Im} F$. Supposons maintenant $u, u' \in \operatorname{Im} F$ et $a, b \in K$. Puisque u et u' appartiennent à l'image de F, il existe des vecteurs $v, v' \in V$ tels que $F(v) = u$ et $F(v') = u'$. D'où

$$F(av + bv') = aF(v) + bF(v') = au + bu' \in \operatorname{Im} F$$

Ainsi l'image de F est un sous-espace de U.

2) Puisque $F(0) = 0$, $0 \in \operatorname{Ker} f$. Supposons maintenant $v, w \in \operatorname{Ker} F$ et $a, b \in K$. Puisque v et w appartiennent au noyau de F, $F(v) = 0$ et $F(w) = 0$. Donc

$$F(av + bw) = aF(v) + bF(w) = a0 + b0 = 0 \quad \text{et donc} \quad av + bw \in \operatorname{Ker} F$$

Donc le noyau de F est un sous-espace de V.

6.23. Démontrer le théorème 6.4 : Soit V un espace vectoriel de dimension finie et soit $F: V \to U$ une application linéaire d'image U' et de noyau W. Alors dim U' + dim W = dim V.

Supposons dim $V = n$. Puisque W est un sous-espace de V, sa dimension est finie, c'est-à-dire dim $W = r \leqslant n$. Nous avons maintenant besoin de démontrer que dim $U' = n - r$.

Soit $\{w_1, \ldots, w_r\}$ une base de W. Etendons-la à une base de V.

$$\{w_1, \ldots, w_r, v_1, \ldots, v_{n-r}\}$$

Soit
$$B = \{F(v_1), F(v_2), \ldots, F(v_{n-r})\}$$

Le théorème est démontré si nous montrons que B est une base de l'image U' de F.

Démontrons que B engendre U'. Soit $u \in U'$. Alors il existe $v \in V$ tel que $F(v) = u$. Puisque $\{w_i, v_i\}$ engendre V et puisque $v \in V$,

$$v = a_1w_1 + \cdots + a_rw_r + b_1v_1 + \cdots + b_{n-r}v_{n-r}$$

où les a_i, b_i sont des scalaires. Remarquons que $F(w_i) = 0$, puisque w_i appartient au noyau de F. D'où

$$\begin{aligned} u = F(v) &= F(a_1w_1 + \cdots + a_rw_r + b_1v_1 + \cdots + b_{n-r}v_{n-r}) \\ &= a_1F(w_1) + \cdots + a_rF(w_r) + b_1F(v_1) + \cdots + b_{n-r}F(v_{n-r}) \\ &= a_10 + \cdots + a_r0 + b_1F(v_1) + \cdots + b_{n-r}F(v_{n-r}) \\ &= b_1F(v_1) + \cdots + b_{n-r}F(v_{n-r}) \end{aligned}$$

En conséquence les $F(v_i)$ engendrent l'image de F.

Démontrons que B est linéairement indépendant. Supposons

$$a_1F(v_1) + a_2F(v_2) + \cdots + a_{n-r}F(v_{n-r}) = 0$$

Alors $F(a_1 v_1 + a_2 v_2 + \cdots + a_{n-r} v_{n-r}) = 0$ et donc $a_1 v_1 + \cdots + a_{n-r} v_{n-r}$ appartient au noyau W de F. Puisque $\{w_i\}$ engendre W, il existe des scalaires $b_1, \ldots, b_r$ tels que

$$a_1v_1 + a_2v_2 + \cdots + a_{n-r}v_{n-r} = b_1w_1 + b_2w_2 + \cdots + b_rw_r$$

ou
$$a_1v_1 + \cdots + a_{n-r}v_{n-r} - b_1w_1 - \cdots - b_rw_r = 0 \qquad (*)$$

Puisque $\{w_i, v_i\}$ est une base de V, il s'agit d'un système linéairement indépendant ; donc les coefficients des w_i et v_i dans (*) sont tous nuls. En particulier $a_1 = 0, \ldots, a_{n-r} = 0$. En conséquence les $F(v_i)$ sont linéairement indépendants.

Ainsi B est une base de U' et donc $\dim U' = n - r$ et le théorème est démontré

6.24. Supposons que $f : V \to U$ soit une application linéaire de noyau W et que $f(v) = u$. Montrer que la "classe" $v + W = \{v + w \; ; \; w \in W\}$ est la préimage ou l'image réciproque de u, c'est-à-dire que $f^{-1}(u) = v + W$.

Nous devons démontrer que 1) $f^{-1}(u) \subset v + W$ et 2) $v + W \subset f^{-1}(u)$. Prouvons premièrement 1). Supposons $v' \in f^{-1}(u)$. Alors $f(v') = u$ et donc $f(v' - v) = f(v') - f(v) = u - u = 0$, c'est-à-dire $v' - v \in W$. Ainsi $v' = v + (v' - v) \in v + W$ et donc $f^{-1}(u) \subset v + W$.

Démontrons maintenant 2). Supposons $v' \in v + W$. Alors $v' = v + w$ où $w \in W$. Puisque W est le noyau de f, $f(w) = 0$. En conséquence $f(v') = f(v + w) = f(v) + f(w) = f(v) + 0 = f(v) = u$. Ainsi $v' = f^{-1}(u)$ et donc $v + W \subset f^{-1}(u)$.

APPLICATIONS SINGULIERES ET NON SINGULIERES

6.25. Supposons $F : V \to U$ une application linéaire, et V de dimension finie. Montrer que V et l'image de F ont la même dimension si et seulement si F est non singulière. Déterminer toutes les applications non singulières $T : \mathbf{R}^4 \to \mathbf{R}^3$.

D'après le théorème 6.4 dim V = dim (Im F) + dim (Ker f). Donc V et Im F ont la même dimension si et seulement si dim (Ker F) = 0 ou Ker $F = \{0\}$ c'est-à-dire si et seulement si F est non singulière.

Puisque la dimension de $\mathbf{R}^3$ est inférieure à la dimension de $\mathbf{R}^4$, elle est donc la dimension de l'image de T. En conséquence, aucune application linéaire $T : \mathbf{R}^4 \to \mathbf{R}^3$ ne peut être non singulière.

6.26. Démontrer qu'une application linéaire $F : V \to U$ est non singulière si et seulement si l'image d'un ensemble indépendant est un ensemble indépendant.

Supposons F non singulière et supposons que $\{v_1, \ldots, v_n\}$ est un sous-ensemble indépendant de V. Nous affirmons que les vecteurs $F(v_1), \ldots, F(v_n)$ sont indépendants. Supposons $a_1 F(v_1) + a_2 F(v_2) + \cdots a_n F(v_n) = 0$ où $a_i \in K$. Puisque F est linéaire $F(a_1 v_1 + a_2 v_2 + \cdots + a_n v_n) = 0$, donc

$$a_1v_1 + a_2v_2 + \cdots + a_nv_n \in \text{Ker } F$$

Mais F est non singulière c'est-à-dire $\operatorname{Ker} F = \{0\}$; donc $a_1 v_1 + a_2 v_2 + \cdots + a_n v_n = 0$. Puisque les v_i sont linéairement indépendants, tous les a_i sont nuls. En conséquence, les $F(v_i)$ sont linéairement indépendants. En d'autres termes, l'image d'un ensemble indépendant $\{v_1, v_2, \ldots, v_n\}$ est indépendant.

D'autre part, supposons que l'image d'un ensemble quelconque indépendant soit indépendant. Alors $\{F(v)\}$ est indépendant et donc $F(v) \neq 0$. En conséquence, F est non singulière.

OPERATIONS SUR DES APPLICATIONS LINEAIRES

6.27. Soit $F : \mathbf{R}^3 \to \mathbf{R}^2$ et $G : \mathbf{R}^3 \to \mathbf{R}^2$ définies par $F(x, y, z) = (2x, y + z)$ et $G(x, y, z) = (x - z, y)$. Trouver les formules définissant les applications $F + G$, $3F$ et $2F - 5G$.

$$\begin{aligned}(F+G)(x, y, z) &= F(x, y, z) + G(x, y, z) \\ &= (2x, y+z) + (x-z, y) = (3x-z, 2y+z)\end{aligned}$$

$$(3F)(x, y, z) = 3F(x, y, z) = 3(2x, y+z) = (6x, 3y+3z)$$

$$\begin{aligned}(2F-5G)(x, y, z) &= 2F(x, y, z) - 5G(x, y, z) = 2(2x, y+z) - 5(x-z, y) \\ &= (4x, 2y+2z) + (-5x+5z, -5y) = (-x+5z, -3y+2z)\end{aligned}$$

6.28. Soit $F : \mathbf{R}^3 \to \mathbf{R}^2$ et $G : \mathbf{R}^2 \to \mathbf{R}^2$ définies par $F(x, y, z) = (2x, y + z)$ et $G(x, y) = (y, x)$. En déduire les formules définissant les applications $G \circ F$ et $F \circ G$.

$$(G \circ F)(x, y, z) = G(F(x, y, z)) = G(2x, y+z) = (y+z, 2x)$$

L'application $F \circ G$ n'est pas définie puisque l'image de G n'est pas contenue dans le domaine de F.

6.29. Montrer 1) l'application nulle 0, définie par $0(v) = 0$ pour tout $v \in V$ est l'élément nul de $\operatorname{Hom}(V, U)$. 2) La symétrique ou opposée de $F \in \operatorname{Hom}(V, U)$ est l'application $(-1)F$, c'est-à-dire $-F = (-1)F$.

1) Soit $F \in \operatorname{Hom}(V, U)$. Alors pour chaque $v \in V$

$$(F+0)(v) = F(v) + 0(v) = F(v) + 0 = F(v)$$

Puisque $(F + 0)(v) = F(v)$ pour tout $v \in V$, $F + 0 = F$.

2) Pour tout $v \in V$,

$$(F+(-1)F)(v) = F(v) + (-1)F(v) = F(v) - F(v) = 0 = 0(v)$$

Puisque $(F + (-1)F)(v) = 0(v)$ pour tout $v \in V$, $F + (-1)F = 0$. Ainsi $(-1)F$ est l'opposée de F.

6.30. Montrer que pour $F_1, \ldots, F_n \in \operatorname{Hom}(V, U)$ et $a_1, \ldots, a_n \in K$ et pour $v \in V$ quelconque,

$$(a_1F_1 + a_2F_2 + \cdots + a_nF_n)(v) = a_1F_1(v) + a_2F_2(v) + \cdots + a_nF_n(v)$$

Par définition de l'application a_1F_1, $(a_1F_1)(v) = a_1F_1(v) = a_1F_1(v)$; donc le théorème reste vrai pour $n = 1$. Ainsi par récurrence,

$$\begin{aligned}(a_1F_1 + a_2F_2 + \cdots + a_nF_n)(v) &= (a_1F_1)(v) + (a_2F_2 + \cdots + a_nF_n)(v) \\ &= a_1F_1(v) + a_2F_2(v) + \cdots + a_nF_n(v)\end{aligned}$$

6.31. Soit $F : \mathbf{R}^3 \to \mathbf{R}^2$, $G : \mathbf{R}^3 \to \mathbf{R}^2$ et $H : \mathbf{R}^3 \to \mathbf{R}^2$ définies par $F(x, y, z) = (x + y + z, x + y)$, $G(x, y, z) = (2x + z, x + y)$ et $H(x, y, z) = (2y, x)$. Montrer que $F, G, H \in \operatorname{Hom}(\mathbf{R}^3, \mathbf{R}^2)$ sont linéairement indépendantes.

Supposons pour les scalaires $a, b, c, \in K$

$$aF + bG + cH = 0 \tag{1}$$

(Ici 0 est l'application nulle). Pour $e_1 = (1, 0, 0) \in \mathbf{R}^3$, nous avons

$$\begin{aligned}(aF + bG + cH)(e_1) &= aF(1, 0, 0) + bG(1, 0, 0) + cH(1, 0, 0) \\ &= a(1, 1) + b(2, 1) + c(0, 1) = (a+2b, a+b+c)\end{aligned}$$

et $0(e_1) = (0, 0)$. Donc d'après (*1*), $(a + 2b, a + b + c) = (0, 0)$ et ainsi

$$a + 2b = 0 \quad \text{et} \quad a + b + c = 0 \tag{2}$$

De façon analogue pour $e_2 = (0, 1, 0) \in \mathbf{R}^3$, nous avons

$$\begin{aligned}(aF + bG + cH)(e_2) &= aF(0, 1, 0) + bG(0, 1, 0) + cH(0, 1, 0)\\ &= a(1, 1) + b(0, 1) + c(2, 0) = (a + 2c, a + b) = 0(e_2) = (0, 0)\end{aligned}$$

Ainsi
$$a + 2c = 0 \quad \text{et} \quad a + b = 0 \tag{3}$$

En utilisant (*2*) et (*3*) nous obtenons $\quad a = 0, \quad b = 0, \quad c = 0 \qquad$ (*4*)

Puisque (*1*) implique (*4*), les applications F, G et H sont linéairement indépendantes.

6.32. Démontrer le théorème 6.7 : Supposons $\dim V = m$ et $\dim U = n$. Alors $\dim \operatorname{Hom}(V, U) = mn$.

Supposons que $\{v_1, \ldots, v_m\}$ soit une base de V et $\{u_1, \ldots, u_n\}$ une base de U. D'après le théorème 6.2 une application linéaire dans $\operatorname{Hom}(V, U)$ est uniquement déterminée en désignant arbitrairement des éléments de V qui sont les images d'éléments de la base de V. On définit

$$F_{ij} \in \operatorname{Hom}(V, U), \quad i = 1, \ldots, m, \quad j = 1, \ldots, n$$

comme l'application linéaire pour laquelle $F_{ij}(v_i) = u_j$ et $F_{ij}(v_k) = 0$ pour $k \neq i$, c'est-à-dire F_{ij} transforme v_i en u_j et les autres v en 0. Remarquons que $\{F_{ij}\}$ contient exactement mn éléments, donc le théorème est démontré si nous montrons qu'il s'agit d'une base de $\operatorname{Hom}(V, U)$.

Démontrons que $\{F_{ij}\}$ *engendre* $\operatorname{Hom}(V, U)$: Soit $F \in \operatorname{Hom}(V, U)$. Supposons $F(v_1) = w_1$, $F(v_2) = w_2, \ldots, F(v_m) = w_m$. Puisque $w_k \in U$ c'est une combinaison linéaire des u, c'est-à-dire

$$w_k = a_{k1}u_1 + a_{k2}u_2 + \cdots + a_{kn}u_n, \qquad k = 1, \ldots, m, \quad a_{ij} \in K \tag{1}$$

Considérons l'application linéaire $G = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}F_{ij}$. Puisque G est une combinaison linéaire des F_{ij}, la preuve que $\{F_{ij}\}$ engendre $\operatorname{Hom}(V, U)$ est complète si nous démontrons que $F = G$.

Calculons maintenant $G(v_k)$, $k = 1, \ldots, m$. Puisque $F_{ij}(v_k) = 0$ pour $k \neq i$ et $F_{ki}(v_k) = u_i$,

$$\begin{aligned}G(v_k) &= \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}F_{ij}(v_k) = \sum_{j=1}^{n} a_{kj}F_{kj}(v_k) = \sum_{j=1}^{n} a_{kj}u_j\\ &= a_{k1}u_1 + a_{k2}u_2 + \cdots + a_{kn}u_n\end{aligned}$$

Ainsi d'après (*1*) $G(v_k) = w_k$ pour tout k. Mais $F(v_k) = w_k$ pour tout k. En conséquence, d'après le théorème 6.2, $F = G$; donc $\{F_{ij}\}$ engendre $\operatorname{Hom}(V, U)$.

Démontrons que $\{F_{ij}\}$ *est linéairement indépendant* : Supposons que pour les scalaires $a_{ij} \in K$,

$$\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}F_{ij} = 0$$

Pour v_k, $k = 1, \ldots, m$,

$$\begin{aligned}0 = 0(v_k) &= \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}F_{ij}(v_k) = \sum_{j=1}^{n} a_{kj}F_{kj}(v_k) = \sum_{j=1}^{n} a_{kj}u_j\\ &= a_{k1}u_1 + a_{k2}u_2 + \cdots + a_{kn}u_n\end{aligned}$$

Mais les u_i sont linéairement indépendants ; donc pour $k = 1, \ldots, m$ nous avons $a_{k1} = 0$ $a_{k2} = 0, \ldots, a_{kn} = 0$. En d'autres termes, tous les $a_{ij} = 0$ et donc $\{F_{ij}\}$ est linéairement indépendant.

Ainsi $\{F_{ij}\}$ est une base de $\operatorname{Hom}(V, U)$; donc $\dim(V, U) = mn$.

6.33. Démontrer le théorème 6.8 : Soient V, U et W des espaces vectoriels sur K. Soient F, F' des applications linéaires de V dans U et soient G, G' des applications linéaires de U dans W ; et soit $k \in K$. Alors 1) $G \circ (F + F') = G \circ F + G \circ F'$; 2) $(G + G') \circ F = G \circ F + G' \circ F$; 3) $k(G \circ F) = G \circ (kF)$.

1) Quel que soit $v \in V$,

$$(G\circ(F+F'))(v) = G((F+F')(v)) = G(F(v)+F'(v))$$
$$= G(F(v))+G(F'(v)) = (G\circ F)(v)+(G\circ F')(v) = (G\circ F+G\circ F')(v)$$

Puisque $(G\circ(F+F'))(v) = (G\circ F+G\circ F')(v)$ quel que soit $v\in V$, $G\circ(F+F') = G\circ F+G\circ F'$

2) Quel que soit $v\in V$,

$$((G+G')\circ F)(v) = (G+G')(F(v)) = G(F(v))+G'(F(v))$$
$$= (G\circ F)(v)+(G'\circ F)(v) = (G\circ F+G'\circ F)(v)$$

Puisque $((G+G')\circ F)(v) = (G\circ F+G\circ F')(v)$ quel que soit $v\in V$, $(G+G')\circ F = G\circ F+G'\circ F$.

3) Quel que soit $v\in V$,

$$(k(G\circ F))(v) = k(G\circ F)(v) = k(G(F(v))) = (kG)(F(v)) = (kG\circ F)(v)$$

et $$(k(G\circ F))(v) = k(G\circ F)(v) = k(G(F(v))) = G(kF(v)) = G((kF)(v)) = (G\circ kF)(v)$$

En conséquence $k(G\circ F) = (kG)\circ F = G\circ(kF)$.(Nous pouvons affirmer que deux applications sont égales en montrant qu'elles associent la même image à chaque point du domaine de définition).

6.34. Soit $F: V\to U$ et $G: U\to W$ deux applications linéaires. Donc $(G\circ F): V\to W$ est linéaire. Montrer que 1) rang $(G\circ F)\leqslant$ rang G, 2) rang $(G\circ F)\leqslant$ rang F.

1) Puisque $F(V)\subset U$, nous avons aussi $G(F(V))\subset G(U)$ et donc $\dim G(F(V))\leqslant \dim G(U)$. Alors

$$\text{rang }(G\circ F) = \dim((G\circ F)(V)) = \dim(G(F(V))) \leq \dim G(U) = \text{rang } G$$

2) D'après le théorème 6.4 : $\dim(G(F(V)))\leqslant \dim F(V)$. Donc

$$\text{rang }(G\circ F) = \dim((G\circ F)(V)) = \dim(G(F(V))) \leq \dim F(V) = \text{rang } F$$

ALGEBRE DES OPERATEURS LINEAIRES

6.35 Soient S et T des opérateurs linéaires sur $\mathbf{R}^2$ définis par $S(x,y) = (y,x)$ et $T(x,y) = (0,x)$. Trouver les formules définissant les opérateurs $S+T$, $2S-3T$, ST, TS, S^2 et T^2

$(S+T)(x,y) = S(x,y)+T(x,y) = (y,x)+(0,x) = (y,2x)$.

$(2S-3T)(x,y) = 2S(x,y)-3T(x,y) = 2(y,x)-3(0,x) = (2y,-x)$.

$(ST)(x,y) = S(T(x,y)) = S(0,x) = (x,0)$.

$(TS)(x,y) = T(S(x,y)) = T(y,x) = (0,y)$.

$S^2(x,y) = S(S(x,y)) = S(y,x) = (x,y)$. Remarquons que $S^2 = I$, application identique

$T^2(x,y) = T(T(x,y)) = T(0,x) = (0,0)$. Remarquons que $T^2 = 0$, application nulle.

6.36. Soit T un opérateur linéaire de $\mathbf{R}^2$ défini par

$$T(3,1) = (2,-4) \quad \text{et} \quad T(1,1) = (0,2) \tag{1}$$

(D'après le théorème 6.2 un tel opérateur linéaire existe et est unique). Trouver $T(a,b)$. En particulier trouver $T(7,4)$

Ecrivons d'abord (a,b) comme combinaison linéaire de $(3,1)$ et $(1,1)$ en utilisant les inconnues scalaires x et y:

$$(a,b) = x(3,1)+y(1,1) \tag{2}$$

Donc $(a,b) = (3x,x)+(y,y) = (3x+y, x+y)$ et donc $\begin{Bmatrix} 3x+y = a \\ x+y = b \end{Bmatrix}$

Résolvons en x et y en fonction de a et b,

$$x = \frac{1}{2}a-\frac{1}{2}b \quad \text{et} \quad y = -\frac{1}{2}a+\frac{3}{2}b \tag{3}$$

Utilisons maintenant (2), (1) et (3),

$$T(a,b) = xT(3,1)+yT(1,1) = x(2,-4)+y(0,2)$$
$$= (2x,-4x)+(0,2y) = (2x,-4x+2y) = (a-b, 5b-3a)$$

Ainsi $T(7,4) = (7-4, 20-21) = (3,-1)$.

6.37. Soit T un opérateur de $\mathbf{R}^3$ défini par $T(x, y, z) = (2x, 4x - y, 2x + 3y - z)$. 1) Montrer que T est inversible. 2) Trouver une formule donnant T^{-1}.

1) Le noyau W de T est l'ensemble de tous les (x, y, z) tels que $T(x, y, z) = (0, 0, 0)$ c'est-à-dire :

$$T(x, y, z) = (2x, 4x - y, 2x + 3y - z) = (0, 0, 0)$$

W est l'espace solution du système homogène

$$2x = 0, \quad 4x - y = 0, \quad 2x + 3y - z = 0$$

qui admet seulement la solution triviale $(0, 0, 0)$. Donc $W = \{0\}$ et T est non singulière et donc d'après le théorème 6.9 inversible.

2) Soit (r, s, t) l'image de (x, y, z) par T ; alors (x, y, z) est l'image de (r, s, t) par T^{-1} : $T(x, y, z) = (r, s, t)$ et $T^{-1}(r, s, t) = (x, y, z)$. Nous trouverons les valeurs de x, y, z en fonction de r, s, t, et en remplaçant dans la formule précédente T^{-1}. Formons

$$T(x, y, z) = (2x, 4x - y, 2x + 3y - z) = (r, s, t)$$

nous trouvons $x = \dfrac{1}{2}r$, $y = 2r - s$, $z = 7r - 3s - t$. Ainsi T^{-1} est donné par

$$T^{-1}(r, s, t) = (\tfrac{1}{2}r, 2r - s, 7r - 3s - t)$$

6.38. Soit V de dimension finie et soit T un opérateur linéaire sur V. Rappelons que T est inversible si et seulement si T est non singulière ou injective. Montrer que T est inversible si et seulement si T est surjective.

D'après le théorème 6.4 $\dim V = \dim(\operatorname{Im} T) + \dim(\operatorname{Ker} T)$. Donc les résultats suivants sont équivalents : 1) T est surjective ; 2) $\operatorname{Im} T = V$; 3) $\dim(\operatorname{Im} T) = \dim V$; 4) $\dim(\operatorname{Ker} T) = 0$; 5) $\operatorname{Ker} T = \{0\}$; 6) T est non singulière ; 7) T est inversible

6.39. Soit V de dimension finie et soit T un opérateur linéaire sur V pour lequel $TS = I$, quel que soit l'opérateur S de V. (Nous appelons S l'inverse à droite de T). 1) Montrer que T est inversible 2) Montrer que $S = T^{-1}$. 3) Donner un exemple montrant que le résultat précédent ne reste plus valable si V est de dimension infinie.

1) Soit $\dim V = n$. D'après le problème précédent, T est inversible si et seulement si T est surjective ; donc T est inversible si et seulement si rang $T = n$. Nous avons $n = \text{rang } I = \text{rang } TS \leqslant \text{rang } T \leqslant n$. Donc rang $T = n$ et T est inversible.

2) $TT^{-1} = T^{-1}T = I$. Alors $S = IS = (T^{-1}T)S = T^{-1}(TS) = T^{-1}I = T^{-1}$.

3) Soit V l'espace des polynômes en t sur K ; soit $p(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n$. Soient T et S deux opérateurs de V définis par

$$T(p(t)) = 0 + a_1 + a_2 t + \cdots + a_n t^{n-1} \quad \text{et} \quad S(p(t)) = a_0 t + a_1 t^2 + \cdots + a_n t^{n+1}$$

Nous avons

$$\begin{aligned}(TS)(p(t)) &= T(S(p(t))) = T(a_0 t + a_1 t^2 + \cdots + a_n t^{n+1}) \\ &= a_0 + a_1 t + \cdots + a_n t^n = p(t)\end{aligned}$$

et donc $TS = I$, l'application identique. D'autre part, si $k \in K$ et $k \neq 0$ alors $ST(k) = S(T(k)) = S(0) = 0 \neq k$. En conséquence $ST \neq I$.

6.40. Soient T et S des opérateurs linéaires sur $\mathbf{R}^2$ définis par $S(x, y) = (0, x)$ et $T(x, y) = (x, 0)$. Montrer que $TS = 0$ mais $ST \neq 0$. Montrer aussi que $T^2 = T$.

$(TS)(x, y) = T(S(x, y)) = T(0, x) = (0, 0)$. Puisque TS associe $0 = (0, 0)$ à chaque $(x, y) \in \mathbf{R}^2$, c'est l'application identique $TS = 0$.

$(ST)(x, y) = S(T(x, y)) = S(x, 0) = (0, x)$. Par exemple $ST(4, 2) = (0, 4)$. Ainsi $ST \neq 0$, puisqu'on n'associe pas $0 = (0, 0)$ à chaque élément de $\mathbf{R}^2$.

Pour tout $(x, y) \in \mathbf{R}^2$, $(T^2(x, y) = T(T(x, y)) = T(x, 0) = x, 0) = T(x, y)$. Donc $T^2 = T$.

PROBLEMES DIVERS

6.41. Soit $\{e_1, e_2, e_3\}$ une base de V et $\{f_1, f_2\}$ une base de U. Soit $T : V \to U$ un opérateur linéaire. De plus, supposons

$$\begin{aligned} T(e_1) &= a_1f_1 + a_2f_2 \\ T(e_2) &= b_1f_1 + b_2f_2 \\ T(e_3) &= c_1f_1 + c_2f_2 \end{aligned} \qquad \text{et} \qquad A = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix}$$

Montrer que, pour tout $v \in V$, $A[v]_e = [T(v)]_f$ où les vecteurs dans K^2 et K^3 sont écrits comme vecteurs colonnes.

Supposons $v = k_1e_1 + k_2e_2 + k_3e_3$; alors $[v]_e = \begin{pmatrix} k_1 \\ k_2 \\ k_3 \end{pmatrix}$. Aussi

$$\begin{aligned} T(v) &= k_1T(e_1) + k_2T(e_2) + k_3T(e_3) \\ &= k_1(a_1f_1 + a_2f_2) + k_2(b_1f_1 + b_2f_2) + k_3(c_1f_1 + c_2f_2) \\ &= (a_1k_1 + b_1k_2 + c_1k_3)f_1 + (a_2k_1 + b_2k_2 + c_2k_3)f_2 \end{aligned}$$

En conséquence,
$$[T(v)]_f = \begin{pmatrix} a_1k_1 + b_1k_2 + c_1k_3 \\ a_2k_1 + b_2k_2 + c_2k_3 \end{pmatrix}$$

Calculons, on obtient

$$A[v]_e = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix}\begin{pmatrix} k_1 \\ k_2 \\ k_3 \end{pmatrix} = \begin{pmatrix} a_1k_1 + b_1k_2 + c_1k_3 \\ a_2k_1 + b_2k_2 + c_2k_3 \end{pmatrix} = [T(v)]_f$$

6.42. Soit k un scalaire non nul. Montrer qu'une application linéaire T est singulière si et seulement si kT est singulière. Donc T est singulière si et seulement si $-T$ est singulière.

Supposons T singulière. Alors $T(v) = 0$ quel que soit le vecteur $v \neq 0$. Donc $(kT)(v) = kT(v) = k0 = 0$ et donc kT est singulière

Supposons maintenant kT singulière. Alors $(kT)(w) = 0$ quel que soit le vecteur $w \neq 0$; donc $T(kw) = kT(w) = (kT)(w) = 0$. Mais $k \neq 0$ et $w \neq 0$ implique $kw \neq 0$; ainsi T est aussi singulière.

6.43. Soit E un opérateur linéaire sur V pour lequel $E^2 = E$. (Un tel opérateur est appelé une projection). Soit U l'image de E et W le noyau. Montrer que : 1) si $u \in U$ alors $E(u) = u$ c'est-à-dire que E est l'application identique sur U ; 2) si $E \neq I$, alors E est singulière, c'est-à-dire $E(v) = 0$ pour au moins un $v \neq 0$. 3) $V = U \oplus W$.

1) Si $u \in U$, l'image de E, alors $E(v) = u$ quel que soit $v \in V$. En utilisant $E^2 = E$, nous avons

$$u = E(v) = E^2(v) = E(E(v)) = E(u)$$

2) Si $E \neq I$, alors quel que soit $v \in V$, $E(v) = u$ où $v \neq u$. D'après (*1*) $E(u) = u$. Ainsi

$$E(v-u) = E(v) - E(u) = u - u = 0 \qquad \text{où} \qquad v - u \neq 0$$

3) Montrons d'abord que $V = U + W$. Soit $v \in V$. Posons $u = E(v)$ et $w = v - E(v)$. Alors

$$v = E(v) + v - E(v) = u + w$$

Par définition, $u = E(v) \in U$, l'image de E. Montrons maintenant que $w \in W$, noyau de E :

$$E(w) = E(v - E(v)) = E(v) - E^2(v) = E(v) - E(v) = 0$$

et ainsi $w \in W$. Donc $V = U + W$.

Montrons maintenant que $U \cap W = \{0\}$. Soit $v \in U \cap W$. Puisque $v \in U, E(v) = v$ d'après (*1*). Puisque $v \in W$, $E(v) = 0$ et ainsi $U \cap W = \{0\}$.

Les deux propriétés précédentes impliquent que $V = U \oplus W$.

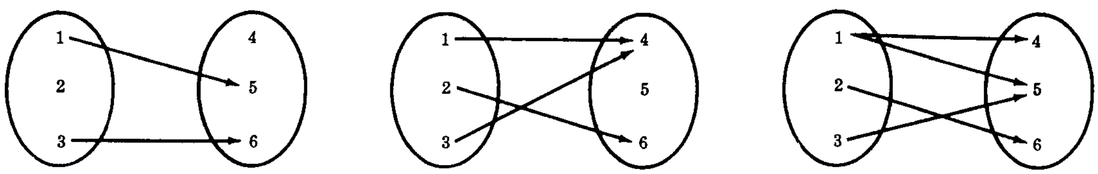
6.44. Montrer qu'une matrice carrée A est inversible si et seulement si elle est non singulière. (Comparer avec le théorème 6.9 page 130).

Rappelons que A est inversible si et seulement si A est équivalente ligne à une matrice identité I. Donc les résultats suivants sont équivalents: 1) A est inversible. 2) A et I sont équivalentes lignes. 3) Les équations $AX = 0$ et $IX = 0$ ont le même espace solution. 4) $AX = 0$ a seulement la solution nulle. 5) A est non singulière.

PROBLEMES SUPPLEMENTAIRES

APPLICATIONS

6.45. Dire si chacun des diagrammes suivants définit une application de $\{1, 2, 3\}$ dans $\{4, 5, 6\}$.



6.46. Définir chacune des applications suivantes $f : \mathbf{R} \to \mathbf{R}$ par une formule.

1) A chaque nombre f associe son carré plus 3.
2) A chaque nombre f associe son cube plus deux fois le nombre.
3) A chaque nombre $\geqslant 3$, f associe le carré de ce nombre et à chaque nombre < 3, f associe le nombre 2.

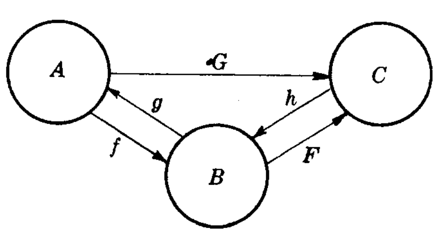
6.47. Soit $f : \mathbf{R} \to \mathbf{R}$ définie par $f(x) = x^2 - 4x + 3$. Trouver 1) $f(4)$, 2) $f(-3)$, 3) $f(y - 2x)$, 4) $f(x - 2)$.

6.48. Déterminer le nombre d'applications différentes de $\{a, b\}$ sur $\{1, 2, 3\}$.

6.49. Soit l'application g associant à chaque nom de l'ensemble {Berthe, Marie, David, Alain, Rebecca} le nombre de lettres distinctes nécessaire pour épeler le nom. Trouver 1) le graphe de g, 2) l'image de g.

6.50. Dresser le graphe de chacune des applications 1) $f(x) = \frac{1}{2}x - 1$, 2) $g(x) = 2x^2 - 4x - 3$.

6.51. Les applications $f : A \to B$, $g : B \to A$, $h : C \to B$, $F : B \to C$ et $G : A \to C$ sont illustrées par le diagramme ci-dessous.



Déterminer si chacune des formules suivantes définit une application composée et si oui trouver le domaine et le co-domaine: 1) $g \circ f$, 2) $h \circ f$, 3) $F \circ f$, 4) $G \circ f$, 5) $g \circ h$, 6) $h \circ G \circ g$.

6.52. Soit $f : \mathbf{R} \to \mathbf{R}$ et $g : \mathbf{R} \to \mathbf{R}$ définies par $f(x) = x^2 + 3x + 1$ et $g(x) = 2x - 3$. Trouver les formules définissant l'application composée 1) $f \circ g$, 2) $g \circ f$, 3) $g \circ g$, 4) $f \circ f$.

6.53. Quelle que soit l'application $f : A \to B$, montrer que $1_B \circ f = f = f \circ 1_A$.

6.54. Pour chacune des applications suivantes $f : \mathbf{R} \to \mathbf{R}$ trouver une formule définissant l'application inverse : 1) $f(x) = 3x - 7$, 2) $f(x) = x^3 + 2$.

APPLICATIONS LINEAIRES

6.55. Montrer que les applications suivantes F sont linéaires :

1) $F : \mathbf{R}^2 \to \mathbf{R}^2$ définie par $F(x, y) = (2x - y, x)$.
2) $F : \mathbf{R}^3 \to \mathbf{R}^2$ définie par $F(x, y, z) = (z, x + y)$.
3) $F : \mathbf{R} \to \mathbf{R}^2$ définie par $F(x) = (2x, 3x)$.
4) $F : \mathbf{R}^2 \to \mathbf{R}^2$ définie par $F(x, y) = (ax + by, cx + dy)$ où $a, b, c, d \in \mathbf{R}$.

6.56. Montrer que les applications suivantes F ne sont pas linéaires:

1) $F : \mathbf{R}^2 \to \mathbf{R}^2$ définie par $F(x, y) = (x^2, y^2)$.
2) $F : \mathbf{R}^3 \to \mathbf{R}^2$ définie par $F(x, y, z) = (x + 1, y + z)$.
3) $F : \mathbf{R} \to \mathbf{R}^2$ définie par $F(x) = (x, 1)$.
4) $F : \mathbf{R}^2 \to \mathbf{R}$ définie par $F(x, y) = |x - y|$.

6.57. Soit V l'espace vectoriel des polynômes en t sur K. Montrer que les applications $T : V \to V$ et $S : V \to V$ définies ci-dessous sont linéaires:

$$T(a_0 + a_1 t + \cdots + a_n t^n) = a_0 t + a_1 t^2 + \cdots + a_n t^{n+1}$$
$$S(a_0 + a_1 t + \cdots + a_n t^n) = 0 + a_1 + a_2 t + \cdots + a_n t^{n-1}$$

6.58. Soit V l'espace vectoriel des matrices $n \times n$ sur K ; et soit M une matrice arbitraire de V. Montrer que les deux premières applications sont linéaires. $T : V \to V$, mais que la troisième n'est pas linéaire (à moins que $M = 0$). 1) $T(A) = MA$, 2) $T(A) = MA - AM$, 3) $T(A) = M + A$.

6.59. Trouver $T(a, b)$ où $T : \mathbf{R}^2 \to \mathbf{R}^3$ est définie par $T(1, 2) = (3, -1, 5)$ et $T(0, 1) = (2, 1, -1)$.

6.60. Trouver $T(a, b, c)$ où $T : \mathbf{R}^3 \to \mathbf{R}$ est définie par

$$T(1, 1, 1) = 3, \quad T(0, 1, -2) = 1 \quad \text{et} \quad T(0, 0, 1) = -2$$

6.61. Supposons $F : V \to U$ linéaire. Montrer que, pour tout $v \in V$, $F(-v) = -F(v)$.

6.62. Soit W un sous-espace de V. Montrer que l'application inclusion de W dans V notée par $i : W \subset V$ et définie par $i(w) = w$ est linéaire

NOYAU ET IMAGE D'APPLICATIONS LINEAIRES

6.63. Pour chacune des applications linéaires F suivantes, trouver une base et la dimensions de (*a*) son image U et (*b*) son noyau W :

1) $F : \mathbf{R}^3 \to \mathbf{R}^3$ définie par $F(x, y, z) = (x + 2y, y - z, x + 2z)$.
2) $F : \mathbf{R}^2 \to \mathbf{R}^2$ définie par $F(x, y) = (x + y, x + y)$.
3) $F : \mathbf{R}^3 \to \mathbf{R}^2$ définie par $F(x, y, z) = (x + y, y + z)$

6.64. Soit V l'espace vectoriel des matrices 2×2 sur $\mathbf{R}$ et soit $M = \begin{pmatrix} 1 & -1 \\ -2 & 2 \end{pmatrix}$. Soit $F : V \to V$ l'application linéaire définie par $F(A) = MA$. Trouver une base et la dimension de (*1*) le noyau W de F et (*2*) l'image U de F.

6.65. Trouver une application linéaire $F : \mathbf{R}^3 \to \mathbf{R}^3$, dont l'image est engendrée par $(1, 2, 3)$ et $(4, 5, 6)$.

6.66. Trouver une application linéaire $F : \mathbf{R}^4 \to \mathbf{R}^3$ dont le noyau est engendré par $(1, 2, 3, 4)$ et $(0, 1, 1, 1)$.

6.67. Soit V l'espace vectoriel des polynômes en t sur $\mathbf{R}$. Soit $D : V \to V$ l'opérateur différentiel: $D(f) = df/dt$. Trouver le noyau et l'image de D.

6.68. Soit $F : V \to U$ une application linéaire. Montrer que 1) l'image sous-espace quelconque de V est un sous-espace de U et 2) la préimage d'un sous-espace quelconque de U est un sous-espace de V.

6.69. Chacune des matrices suivantes détermine une application linéaire de $\mathbf{R}^4$ dans $\mathbf{R}^3$:

$$\text{(i)}\quad A = \begin{pmatrix} 1 & 2 & 0 & 1 \\ 2 & -1 & 2 & -1 \\ 1 & -3 & 2 & -2 \end{pmatrix} \qquad \text{(ii)}\quad B = \begin{pmatrix} 1 & 0 & 2 & -1 \\ 2 & 3 & -1 & 1 \\ -2 & 0 & -5 & 3 \end{pmatrix}$$

Trouver une base et la dimension de l'image U et du noyau W de chaque application.

6.70. Soit $T : \mathbf{C} \to \mathbf{C}$ l'application conjuguée sur le corps des complexes $\mathbf{C}$, c'est-à-dire $T(z) = \bar{z}$ où $z \in C$, ou $T(a+ib) = a - ib$ où $a, b \in \mathbf{R}$. 1) Montrer que T est non linéaire si $\mathbf{C}$ est considéré comme un espace vectoriel sur lui-même; 2) montrer que T est linéaire si $\mathbf{C}$ est considéré comme un espace vectoriel sur le corps des réels $\mathbf{R}$.

OPERATIONS SUR LES APPLICATIONS LINEAIRES

6.71. Soit $F : \mathbf{R}^3 \to \mathbf{R}^2$ et $G : \mathbf{R}^3 \to \mathbf{R}^2$ définies par $F(x, y, z) = (y, z + x)$ et $G(x, y, z) = (2z, x - y)$. Trouver les formules définissant les applications $F + G$ et $3F - 2G$.

6.72. Soit $H : \mathbf{R}^2 \to \mathbf{R}^2$ définie par $H(x, y) = (y, 2x)$. En utilisant les applications F et G du problème précédent, trouver les formules définissant les applications 1) $H \circ F$ et $H \circ G$, 2) $F \circ H$ et $G \circ H$, 3) $H \circ (F + G)$ et $H \circ F + H \circ G$.

6.73. Montrer que les applications suivantes F, G et H sont linéairement indépendantes:

1) $F, G, H \in \mathrm{Hom}(\mathbf{R}^2, \mathbf{R}^2)$ définies par :
$$F(x, y) = (x, 2y),\ G(x, y) = (y, x + y),\ H(x, y) = (0, x)$$

2) $F, G, H \in \mathrm{Hom}(\mathbf{R}^3, \mathbf{R})$ définies par
$$F(x, y, z) = x + y + z,\ G(x, y, z) = y + z,\ H(x, y, z) = x - z$$

6.74. Pour $F, G \in \mathrm{Hom}(V, U)$, montrer que rang $(F + G) \leqslant$ rang F + rang G. (Ici V est de dimension finie)

6.75. Soit $F : V \to U$ et $G : U \to V$ deux applications linéaires. Montrer que si F et G sont non singulières, alors $G \circ F$ est non singulière. Donner un exemple où $G \circ F$ est non singulière mais G est singulière.

6.76. Montrer que Hom (V, U) satisfait bien tous les axiomes d'un espace vectoriel, c'est-à-dire démontrer le théorème 6.6 page 128.

ALGEBRE D'OPERATEURS LINEAIRES

6.77. Soient S et T deux opérateurs linéaires de $\mathbf{R}^2$ définis par $S(x, y) = (x + y, 0)$ et $T(x, y) = (-y, x)$. Trouver les formules définissant les opérateurs $S + T$, $5S - 3T$, ST, TS, S^2 et T^2.

6.78. Soit T un opérateur linéaire sur $\mathbf{R}^2$ défini par $T(x, y) = (x + 2y, 3x + 4y)$. Trouver $p(T)$ où $p(t) = t^2 - 5t - 2$.

6.79. Montrer que chacun des opérateurs suivants T de $\mathbf{R}^3$ est inversible et trouver une formule pour T^{-1} : 1) $T(x, y, z) = (x - 3y - 2z, y - 4z, z)$; 2) $T(x, y, z) = (x + z, x - z, y)$.

6.80. Supposons S et T des opérateurs linéaires sur V, et supposons que S soit non singulier. Nous supposons que V est de dimension finie. Montrer que rang (ST) = rang (TS) = rang T.

6.81. Supposons $V = U \oplus W$. Soient E_1 et E_2 des opérateurs linéaires sur V définis par $E_1(v) = u$, $E_2(v) = w$, où $v = u + w$, $u \in U, w \in W$. Montrer que 1) $E_1^2 = E_1$ et $E_2^2 = E_2$ c'est-à-dire que E_1 et E_2 sont des "projections". 2) $E_1 + E_2 = I$, l'application identique. 3) $E_1 E_2 = 0$ et $E_2 E_1 = 0$.

6.82. Soient E_1 et E_2 des opérateurs linéaires sur V satisfaisant 1), 2) et 3) du problème 6.81. Montrer que V est la somme directe de l'image de E_1 et de l'image de E_2 : $V = \mathrm{Im}\, E_1 \oplus \mathrm{Im}\, E_2$.

6.83. Montrer que si S et T sont des opérateurs linéaires inversibles, alors ST est inversible et $(ST)^{-1} = T^{-1} S^{-1}$.

6.84. Soit V de dimension finie, et soit T un opérateur linéaire sur V, tel que rang (T^2) = rang T. Montrer que $\operatorname{Ker} T \cap \operatorname{Im} T = \{0\}$.

PROBLEMES DIVERS

6.85. Supposons $T : K^n \to K^m$ une application linéaire. Soit $\{e_1, \ldots, e_n\}$ la base usuelle de K^n et soit A la matrice $m \times n$ dont les colonnes sont les vecteurs $T(e_1), \ldots, T(e_n)$ respectivement. Montrer que, pour tout vecteur $v \in K^n$, $T(v) = Av$, où v est écrit sous la forme de vecteur colonne.

6.86. Supposons $F : V \to U$ une application linéaire et k un scalaire non nul. Montrer que les applications F et kF ont le même noyau et la même image.

6.87. Montrer que si $F : V \to U$ est surjective, alors $\dim U \leqslant \dim V$. Déterminer toutes les applications linéaires $T : \mathbf{R}^3 \to \mathbf{R}^4$ qui sont surjectives.

6.88. Trouver ceux des théorèmes du chapitre 3 qui montrent que l'espace des matrices carrées n sur K est une algèbre associative de K.

6.89. Soit $T : V \to U$ une application linéaire et soit W un sous-espace de V. La restriction de T à W est l'application $T_W : W \to U$ définie par $T_W(w) = T(w)$, pour tout $w \in W$. Démontrer que 1) T_W est linéaire; 2) $\operatorname{Ker} T_W = \operatorname{Ker} T \cap W$; 3) $\operatorname{Im} T_W = T(W)$.

6.90. Deux opérateurs $S, T \in A(V)$ sont dits semblables, s'il existe un opérateur inversible $P \in A(V)$ pour lequel $S = P^{-1}TP$. Déterminer que : 1) La similitude des opérateurs est une relation d'équivalence. 2) Des opérateurs semblables ont le même rang (quand V est de dimension finie).

REPONSES AUX PROBLEMES SUPPLEMENTAIRES

6.45. 1) non 2) oui 3) non

6.46. (i) $f(x) = x^2 + 3$, (ii) $f(x) = x^3 + 2x$, (iii) $f(x) = \begin{cases} x^2 & \text{if } x \geqq 3 \\ -2 & \text{if } x < 3 \end{cases}$

6.47. (i) 3, (ii) 24, (iii) $y^2 - 4xy + 4x^2 - 4y + 8x + 3$, (iv) $x^2 - 8x + 15$.

6.48.

6.49. 1) {(Berthe, 5), (Marie, 5), (David 4), (Alain, 4), (Rebecca, 5)}
2) Image de $g = \{3, 4, 5, 6\}$.

6.51. 1) $(g \circ f) : A \to A$ 2) non 3) $(F \circ f) : A \to C$ 4) non 5) $(g \circ h) : C \to A$ 6) $(h \circ G \circ g) : B \to B$.

6.52. (i) $(f \circ g)(x) = 4x^2 - 6x + 1$ (iii) $(g \circ g)(x) = 4x - 9$
(ii) $(g \circ f)(x) = 2x^2 + 6x - 1$ (iv) $(f \circ f)(x) = x^4 + 6x^3 + 14x^2 + 15x + 5$

6.54. (i) $f^{-1}(x) = (x + 7)/3$, (ii) $f^{-1}(x) = \sqrt[3]{x - 2}$.

6.59. $T(a, b) = (-a + 2b, -3a + b, 7a - b)$.

6.60. $T(a, b, c) = 8a - 3b - 2c$.

6.61. $F(v) + F(-v) = F(v + (-v)) = F(0) = 0$; d'où $F(-v) = -F(v)$.

6.63. (i) (a) $\{(1,0,1), (0,1,-2)\}$, $\dim U = 2$; (b) $\{(2,-1,-1)\}$, $\dim W = 1$.
(ii) (a) $\{(1,1)\}$, $\dim U = 1$; (b) $\{(1,-1)\}$, $\dim W = 1$.
(iii) (a) $\{(1,0), (0,1)\}$, $\dim U = 2$; (b) $\{(1,-1,1)\}$, $\dim W = 1$.

6.64. (i) $\left\{\begin{pmatrix}1 & 0\\ 1 & 0\end{pmatrix}, \begin{pmatrix}0 & 1\\ 0 & 1\end{pmatrix}\right\}$ base de Ker F dim (Ker F) = 2

(ii) $\left\{\begin{pmatrix}1 & 0\\ -2 & 0\end{pmatrix}, \begin{pmatrix}0 & 1\\ 0 & -2\end{pmatrix}\right\}$ base de Im F ; dim (Im F) = 2.

6.65. $F(x, y, z) = (x+4y, 2x+5y, 3x+6y)$.

6.66. $F(x, y, z, w) = (x+y-z, 2x+y-w, 0)$.

6.67. Le noyau de D est l'ensemble des polynômes constants. L'image de D est l'espace entier V.

6.69. 1) $a)\ \{(1, 2, 1), (0, 1, 1)\}$ base de ImA ; dim (ImA) = 2.
$b)\ \{(4, -2, -5, 0), (1, -3, 0, 5)\}$ base de Ker A ; dim (Ker A) = 2.
2) $a)$ Im $B = \mathbf{R}^3$; $b)\ \{(-1, 2/3, 1, 1)\}$ base de Ker B ; dim (Ker B) = 1.

6.71. $(F+G)(x, y, z) = (y+2z, 2x-y+z)$, $(3F-2G)(x, y, z) = (3y-4z, x+2y+3z)$.

6.72. (i) $(H\circ F)(x, y, z) = (x+z, 2y)$, $(H\circ G)(x, y, z) = (x-y, 4z)$. (ii) Non définie
(iii) $(H\circ(F+G))(x, y, z) = (H\circ F + H\circ G)(x, y, z) = (2x-y+z, 2y+4z)$.

6.77. $(S+T)(x, y) = (x, x)$ $(ST)(x, y) = (x-y, 0)$
$(5S-3T)(x, y) = (5x+8y, -3x)$ $(TS)(x, y) = (0, x+y)$

$S^2(x, y) = (x+y, 0)$; remarquons que $S^2 = S$.
$T^2(x, y) = (-x, -y)$; remarquons que $T^2 + I = 0$ donc I est un zéro de x^2+1

6.78. $p(T) = 0$.

6.79. (i) $T^{-1}(r, s, t) = (14t+3s+r, 4t+s, t)$, (ii) $T^{-1}(r, s, t) = (\tfrac{1}{2}r+\tfrac{1}{2}s, t, \tfrac{1}{2}r-\tfrac{1}{2}s)$.

6.87. Il n'y a pas d'applications linéaires de $\mathbf{R}^3$ dans $\mathbf{R}^4$ qui soient surjectives.

CHAPITRE 7

Matrices et opérateurs linéaires

INTRODUCTION

Supposons $\{e_1, \ldots, e_n\}$ une base de l'espace vectoriel V sur le corps K, et pour $v \in V$, supposons que $v = a_1 e_1 + a_2 e_2 + \cdots + a_n e_n$. Alors, le vecteur coordonné de v relativement à $\{e_i\}$ que l'on écrit sous forme de vecteur colonne, à moins qu'il en soit spécifié autrement, est

$$[v]_e = \begin{pmatrix} a_1 \\ a_2 \\ \cdot\cdot \\ a_n \end{pmatrix}$$

Rappelons que l'application $v \mapsto [v]_e$ déterminée par la base $\{e_i\}$ est un isomorphisme de V sur l'espace K^n.

Dans ce chapitre, nous montrons qu'il y a aussi un isomorphisme déterminé par la base $\{e_i\}$ de l'algèbre $A(V)$ des opérateurs linéaires sur V sur l'algèbre $\mathcal{A}$ des matrices carrées sur K.

Un résultat analogue est aussi valable pour les applications linéaires $F : V \to U$ d'un espace dans un autre.

REPRESENTATION MATRICIELLE D'UN OPERATEUR LINEAIRE

Soit T un opérateur linéaire d'un espace V sur un corps K et supposons que $\{e_1, e_2, \ldots, e_n\}$ soit une base de V.

Donc $T(e_1), \ldots, T(e_n)$ sont des vecteurs de V et donc chacun est une combinaison linéaire des éléments de la base $\{e_i\}$:

$$\begin{aligned} T(e_1) &= a_{11}e_1 + a_{12}e_2 + \cdots + a_{1n}e_n \\ T(e_2) &= a_{21}e_1 + a_{22}e_2 + \cdots + a_{2n}e_n \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ T(e_n) &= a_{n1}e_1 + a_{n2}e_2 + \cdots + a_{nn}e_n \end{aligned}$$

On a la définition suivante :

Définition : La transposée de la matrice précédente des coefficients, notée $[T]_e$ ou $[T]$, est appelée la représentation matricielle de T relativement à la base $\{e_i\}$ ou simplement la matrice de T sur la base $\{e_i\}$:

$$[T]_e = \begin{pmatrix} a_{11} & a_{21} & \ldots & a_{n1} \\ a_{12} & a_{22} & \ldots & a_{n2} \\ \ldots & \ldots & \ldots & \ldots \\ a_{1n} & a_{2n} & \ldots & a_{nn} \end{pmatrix}$$

Exemple 7.1 : Soit V l'espace vectoriel des polynômes en t sur $\mathbf{R}$ de degré $\leqslant 3$ et soit $D : V \to V$ l'opérateur différentiel défini par $D(p(t)) = d(p(t))/dt$. Calculons la matrice de D dans la base $\{1, t, t^2, t^3\}$. Nous avons :

$$\begin{aligned} D(1) &= 0 = 0 + 0t + 0t^2 + 0t^3 \\ D(t) &= 1 = 1 + 0t + 0t^2 + 0t^3 \\ D(t^2) &= 2t = 0 + 2t + 0t^2 + 0t^3 \\ D(t^3) &= 3t^2 = 0 + 0t + 3t^2 + 0t^3 \end{aligned}$$

En conséquence,

$$[D] = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Exemple 7.2 : Soit T l'opérateur linéaire sur $\mathbf{R}^2$ défini par $T(x, y) = (4x - 2y, 2x + y)$. Calculons la matrice de T dans la base $\{f_1 = (1, 1), f_2 = (-1, 0)\}$; nous avons

$$T(f_1) = T(1, 1) = (2, 3) = 3(1, 1) + (-1, 0) = 3f_1 + f_2$$

$$T(f_2) = T(-1, 0) = (-4, -2) = -2(1, 1) + 2(-1, 0) = -2f_1 + 2f_2$$

En conséquence, $[T]_f = \begin{pmatrix} 3 & -2 \\ 1 & 2 \end{pmatrix}$.

Remarque : Rappelons qu'une matrice A quelconque carrée d'ordre n sur K définit un opérateur de K^n par l'application $v \mapsto Av$ (où v est écrit sous forme de vecteur colonne). Nous montrerons (problème 7.7) que la représentation matricielle de cet opérateur est précisément la matrice A si on utilise la base usuelle de K^n.

Notre premier théorème nous indique que "l'action" d'un opérateur T sur un vecteur v est conservée par sa représentation matricielle.

Théorème 7.1 : Soit $\{e_1, e_2, \ldots, e_n\}$ une base de V et soit T un opérateur linéaire quelconque sur V. Alors quel que soit le vecteur $v \in V$, $[T]_e\ [v]_e = [T(v)]_e$.

C'est-à-dire que si nous multiplions le vecteur coordonné de v par la représentation matricielle de T, nous obtenons le vecteur coordonné de $T(v)$.

Exemple 7.3 : Considérons l'opérateur différentiel $D : V \to V$ dans l'exemple 7.1. Soit :

$$p(t) = a + bt + ct^2 + dt^3 \quad \text{et donc} \quad D(p(t)) = b + 2ct + 3dt^2$$

Donc, relativement à la base $\{1, t, t^2, t^3\}$,

$$[p(t)] = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} \quad \text{et} \quad [D(p(t))] = \begin{pmatrix} b \\ 2c \\ 3d \\ 0 \end{pmatrix}$$

Montrons que le théorème 7.1 s'applique bien ici :

$$[D][p(t)] = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} b \\ 2c \\ 3d \\ 0 \end{pmatrix} = [D(p(t))]$$

Exemple 7.4 : Considérons l'opérateur linéaire $T : \mathbf{R}^2 \to \mathbf{R}^2$ dans l'exemple 7.2 : $T(x, y) = (4x - 2y, 2x + y)$. Soit $v = (5, 7)$. Alors,

$$v = (5, 7) = 7(1, 1) + 2(-1, 0) = 7f_1 + 2f_2$$

$$T(v) = (6, 17) = 17(1, 1) + 11(-1, 0) = 17f_1 + 11f_2$$

où $f_1 = (1, 1)$ et $f_2 = (-1, 0)$. Donc, relativement à la base $\{f_1, f_2\}$,

$$[v]_f = \begin{pmatrix} 7 \\ 2 \end{pmatrix} \quad \text{et} \quad [T(v)]_f = \begin{pmatrix} 17 \\ 11 \end{pmatrix}$$

En utilisant la matrice $[T]_f$ dans l'exemple 7.2, nous vérifions que le théorème 7.1 reste applicable ici :

$$[T]_f\,[v]_f = \begin{pmatrix} 3 & -2 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 7 \\ 2 \end{pmatrix} = \begin{pmatrix} 17 \\ 11 \end{pmatrix} = [T(v)]_f$$

Nous avons maintenant associé une matrice $[T]_e$ à chaque T dans $A(V)$, algèbre des opérateurs linéaires sur V. D'après le premier théorème, l'action d'un opérateur T est préservée par sa représentation matricielle. Les deux théorèmes suivants montrent que les trois opérations de base avec ces opérateurs :

1) addition, 2) multiplication par un scalaire, 3) composition,

sont aussi conservées.

Théorème 7.2 : Soit $\{e_1, \ldots, e_n\}$ une base de V sur K et soit $\mathcal{A}$ l'algèbre des matrices carrées n sur K. L'application $T \to [T]_e$ est un isomorphisme d'espaces vectoriels de $A(V)$ sur $\mathcal{A}$ C'est-à-dire que l'application est injective et surjective quels que soient S, $T \in A(V)$ et $k \in K$.

$$[T+S]_e = [T]_e + [S]_e \qquad \text{et} \qquad [kT]_e = k[T]_e$$

Théorème 7.3 : Quels que soient les opérateurs S, $T \in A(V)$, $[ST]_e = [S]_e\,[T]_e$.

Illustrons les théorèmes précédents dans le cas où dim $V = 2$. Supposons $\{e_1, e_2\}$ une base de V et T et S des opérateurs sur V pour lesquels :

$$\begin{aligned} T(e_1) &= a_1e_1 + a_2e_2 \\ T(e_2) &= b_1e_1 + b_2e_2 \end{aligned}, \qquad \begin{aligned} S(e_1) &= c_1e_1 + c_2e_2 \\ S(e_2) &= d_1e_1 + d_2e_2 \end{aligned}$$

Alors
$$[T]_e = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} \qquad \text{et} \qquad [S]_e = \begin{pmatrix} c_1 & d_1 \\ c_2 & d_2 \end{pmatrix}$$

Nous avons maintenant
$$\begin{aligned} (T+S)(e_1) &= T(e_1) + S(e_1) = a_1e_1 + a_2e_2 + c_1e_1 + c_2e_2 \\ &= (a_1+c_1)e_1 + (a_2+c_2)e_2 \\ (T+S)(e_2) &= T(e_2) + S(e_2) = b_1e_1 + b_2e_2 + d_1e_1 + d_2e_2 \\ &= (b_1+d_1)e_1 + (b_2+d_2)e_2 \end{aligned}$$

Ainsi
$$[T+S]_e = \begin{pmatrix} a_1+c_1 & b_1+d_1 \\ a_2+c_2 & b_2+d_2 \end{pmatrix} = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} + \begin{pmatrix} c_1 & d_1 \\ c_2 & d_2 \end{pmatrix} = [T]_e + [S]_e$$

Pour $k \in K$, nous avons aussi :

$$\begin{aligned} (kT)(e_1) &= kT(e_1) = k(a_1e_1 + a_2e_2) = ka_1e_1 + ka_2e_2 \\ (kT)(e_2) &= kT(e_2) = k(b_1e_1 + b_2e_2) = kb_1e_1 + kb_2e_2 \end{aligned}$$

D'où
$$[kT]_e = \begin{pmatrix} ka_1 & kb_1 \\ ka_2 & kb_2 \end{pmatrix} = k\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} = k[T]_e$$

Finalement, nous avons
$$\begin{aligned} (ST)(e_1) &= S(T(e_1)) = S(a_1e_1 + a_2e_2) = a_1S(e_1) + a_2S(e_2) \\ &= a_1(c_1e_1 + c_2e_2) + a_2(d_1e_1 + d_2e_2) \\ &= (a_1c_1 + a_2d_1)e_1 + (a_1c_2 + a_2d_2)e_2 \\ (ST)(e_2) &= S(T(e_2)) = S(b_1e_1 + b_2e_2) = b_1S(e_1) + b_2S(e_2) \\ &= b_1(c_1e_1 + c_2e_2) + b_2(d_1e_1 + d_2e_2) \\ &= (b_1c_1 + b_2d_1)e_1 + (b_1c_2 + b_2d_2)e_2 \end{aligned}$$

En conséquence,
$$[ST]_e = \begin{pmatrix} a_1c_1 + a_2d_1 & b_1c_1 + b_2d_1 \\ a_1c_2 + a_2d_2 & b_1c_2 + b_2d_2 \end{pmatrix} = \begin{pmatrix} c_1 & d_1 \\ c_2 & d_2 \end{pmatrix}\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} = [S]_e\,[T]_e$$

CHANGEMENT DE BASE

Nous avons montré que l'on peut représenter des vecteurs par des n-tuples (vecteurs colonnes) et des opérateurs linéaires par des matrices, lorsqu'on a choisi une base. Posons-nous la question suivante, qui est naturelle : comment la représentation matricielle change-t-elle si nous choisissons une autre base ? Pour répondre à cette question nous avons d'abord besoin de la définition suivante :

Définition : Soit $\{e_1, \ldots, e_n\}$ une base de V et soit $\{f_1, f_2, \ldots, f_n\}$ une autre base. Supposons

$$\begin{aligned} f_1 &= a_{11}e_1 + a_{12}e_2 + \cdots + a_{1n}e_n \\ f_2 &= a_{21}e_1 + a_{22}e_2 + \cdots + a_{2n}e_n \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ f_n &= a_{n1}e_1 + a_{n2}e_2 + \cdots + a_{nn}e_n \end{aligned}$$

La transposée P de la matrice des coefficients ci-dessus est appelée matrice de passage de l'ancienne base $\{e_i\}$ à la nouvelle base $\{f_i\}$.

$$P = \begin{pmatrix} a_{11} & a_{21} & \ldots & a_{n1} \\ a_{12} & a_{22} & \ldots & a_{n2} \\ \ldots & \ldots & \ldots & \ldots \\ a_{1n} & a_{2n} & \ldots & a_{nn} \end{pmatrix}$$

Nous pouvons dire, puisque les vecteurs $f_1, \ldots, f_n$ sont linéairement indépendants, que la matrice P est inversible (Problème 5.47). En fait, son inverse P^{-1} est la matrice de passage de la base $\{f_I\}$ à la base $\{e_i\}$.

Exemple 7.5 : Considérons les deux bases suivantes de $\mathbf{R}^2$:

$$\{e_1 = (1,0),\ e_2 = (0,1)\} \quad \text{et} \quad \{f_1 = (1,1),\ f_2 = (-1,0)\}$$

Alors
$$\begin{aligned} f_1 &= (1,1) = (1,0) + (0,1) = e_1 + e_2 \\ f_2 &= (-1,0) = -(1,0) + 0(0,1) = -e_1 + 0e_2 \end{aligned}$$

Donc la matrice de passage P de la base $\{e_i\}$ à la base $\{f_i\}$ est :

$$P = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}$$

Nous avons aussi
$$\begin{aligned} e_1 &= (1,0) = 0(1,1) - (-1,0) = 0f_1 - f_2 \\ e_2 &= (0,1) = (1,1) + (-1,0) = f_1 + f_2 \end{aligned}$$

Donc la matrice de passage Q de la base $\{f_i\}$ à la base $\{e_i\}$ est :

$$Q = \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}$$

Observons que P et Q sont inverses:

$$PQ = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$$

Montrons maintenant comment les vecteurs coordonnées sont affectés par un changement de base.

Théorème 7.4 : Soit P la matrice de passage de la base $\{e_i\}$ à la base $\{f_i\}$ dans un espace vectoriel V. Pour un vecteur quelconque $v \in V$, $P[v]_f = [v]_e$. Donc $[v]_f = P^{-1}[v]_e$.

Insistons sur le fait que, quoique la matrice P soit la matrice de passage de l'ancienne base $\{e_i\}$ à la nouvelle base $\{f_i\}$, elle a pour effet de transformer les coordonnées d'un vecteur de la nouvelle base $\{f_i\}$ en les coordonnées de ce vecteur dans la base $\{e_i\}$.

Illustrons le précédent théorème dans le cas de dim $V = 3$. Supposons que P soit la matrice de passage de la base $\{e_1, e_2, e_3\}$ de V à la base $\{f_1, f_2, f_3\}$ de V ; c'est-à-dire :

$$\begin{aligned} f_1 &= a_1e_1 + a_2e_2 + a_3e_3 \\ f_2 &= b_1e_1 + b_2e_2 + b_3e_3 . \\ f_3 &= c_1e_1 + c_2e_2 + c_3e_3 \end{aligned} \qquad \text{Donc} \quad P = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$$

Supposons maintenant $v \in V$, c'est-à-dire $v = k_1 f_1 + k_2 f_2 + k_3 f_3$ et remplaçons dans cette expression les f_i par leurs valeurs, on obtient :

$$\begin{aligned} v &= k_1(a_1e_1 + a_2e_2 + a_3e_3) + k_2(b_1e_1 + b_2e_2 + b_3e_3) + k_3(c_1e_1 + c_2e_2 + c_3e_3) \\ &= (a_1k_1 + b_1k_2 + c_1k_3)e_1 + (a_2k_1 + b_2k_2 + c_2k_3)e_2 + (a_3k_1 + b_3k_2 + c_3k_3)e_3 \end{aligned}$$

Ainsi,

$$[v]_f = \begin{pmatrix} k_1 \\ k_2 \\ k_3 \end{pmatrix} \qquad \text{et} \qquad [v]_e = \begin{pmatrix} a_1k_1 + b_1k_2 + c_1k_3 \\ a_2k_1 + b_2k_2 + c_2k_3 \\ a_3k_1 + b_3k_2 + c_3k_3 \end{pmatrix}$$

En conséquence,

$$P[v]_f = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}\begin{pmatrix} k_1 \\ k_2 \\ k_3 \end{pmatrix} = \begin{pmatrix} a_1k_1 + b_1k_2 + c_1k_3 \\ a_2k_1 + b_2k_2 + c_2k_3 \\ a_3k_1 + b_3k_2 + c_3k_3 \end{pmatrix} = [v]_e$$

Multiplions aussi l'équation précédente par P^{-1}, nous avons :

$$P^{-1}[v]_e = P^{-1}P[v]_f = I[v]_f = [v]_f$$

Exemple 7.6 : Soit $v = (a, b) \in \mathbf{R}^2$. Alors, pour les bases de $\mathbf{R}^2$ de l'exemple précédent,

$$\begin{aligned} v &= (a, b) = a(1, 0) + b(0, 1) = ae_1 + be_2 \\ v &= (a, b) = b(1, 1) + (b - a)(-1, 0) = bf_1 + (b - a)f_2 \end{aligned}$$

Donc $$[v]_e = \begin{pmatrix} a \\ b \end{pmatrix} \qquad \text{et} \qquad [v]_f = \begin{pmatrix} b \\ b - a \end{pmatrix}$$

D'après l'exemple précédent, la matrice de passage P de $\{e_i\}$ à $\{f_i\}$ et son inverse P^{-1} sont données par

$$P = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix} \qquad \text{et} \qquad P^{-1} = \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}$$

Nous vérifions ainsi le résultat du théorème 7.4 :

$$P[v]_f = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} b \\ b - a \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix} = [v]_e$$

$$P^{-1}[v]_e = \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} b \\ b - a \end{pmatrix} = [v]_f$$

Le théorème suivant montre comment les représentations matricielles d'opérateurs linéaires sont affectées par un changement de base.

Théorème 7.5 : Soit P la matrice de passage d'une base $\{e_i\}$ à une base $\{f_i\}$ dans un espace vectoriel V. Alors, quel que soit l'opérateur linéaire T de V, $[T]_f = P^{-1}[T]_e P$.

Exemple 7.7 : Soit T un opérateur linéaire de $\mathbf{R}^2$ défini par $T(x, y) = (4x - 2y, 2x + y)$. Alors, pour les bases de $\mathbf{R}^2$ dans l'exemple 7.5 nous avons :

$$\begin{aligned} T(e_1) &= T(1, 0) = (4, 2) = 4(1, 0) + 2(0, 1) = 4e_1 + 2e_2 \\ T(e_2) &= T(0, 1) = (-2, 1) = -2(1, 0) + (0, 1) = -2e_1 + e_2 \end{aligned}$$

En conséquence $$[T]_e = \begin{pmatrix} 4 & -2 \\ 2 & 1 \end{pmatrix}$$

Calculons $[T]_f$, en utilisant le théorème 7.5 :

$$[T]_f = P^{-1}[T]_e P = \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 4 & -2 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 3 & -2 \\ 1 & 2 \end{pmatrix}$$

Remarquons que ce résultat concorde avec la manière dont on obtient $[T]_f$ dans l'exemple 7.2.

Remarque : Supposons $P = (a_{ij})$ une matrice quelconque carrée d'ordre n inversible sur un corps K. Si $\{e_1, e_2, \ldots, e_n\}$ est une base d'un espace vectoriel V sur K, alors les n vecteurs

$$f_i = a_{1i}e_1 + a_{2i}e_2 + \cdots + a_{ni}e_n, \qquad i = 1, \ldots, n$$

sont linéairement indépendants (Problème 5.47) et donc forment une autre base de V. De plus, P est la matrice de passage de la base $\{e_i\}$ à la base $\{f_i\}$. En conséquence, si A est la représentation matricielle quelconque d'un opérateur linéaire T sur V, alors la matrice $B = P^{-1}AP$ est aussi une représentation matricielle de T.

MATRICES SEMBLABLES

Supposons que A et B soient des matrices carrées pour lesquelles il existe une matrice inversible P telle que $B = P^{-1}AP$. On dit alors que B est semblable à A. On dit aussi que B est obtenue à partir de A par une similitude. Nous montrerons (Problème 7.22) que la similitude des matrices est une relation d'équivalence. Ainsi, d'après le théorème 7.5 et la remarque précédente, nous avons le résultat fondamental suivant :

Théorème 7.6 : Deux matrices A et B représentent le même opérateur linéaire T si, et seulement si, elles sont semblables.

C'est-à-dire que toutes les représentations matricielles de l'opérateur linéaire T forment une classe d'équivalence de matrices semblables.

Un opérateur linéaire T est dit diagonalisable, s'il existe une base $\{e_i\}$ sur laquelle il est représenté par une matrice diagonale ; la base $\{e_i\}$ diagonalise T. Le théorème précédent nous donne le résultat suivant :

Théorème 7.7. : Soit A une représentation matricielle d'un opérateur linéaire T. T est alors diagonalisable si, et seulement si, il existe une matrice inversible P telle que $P^{-1}AP$ soit une matrice diagonale.

C'est-à-dire que T est diagonalisable si, et seulement si, sa représentation matricielle peut être diagonalisée par une similitude.

Remarquons que tous les opérateurs ne sont pas diagonalisables. Cependant, nous montrerons (Chapitre 10) que tout opérateur T peut être représenté par certaines matrices "standard" appelées ses formes normale ou canonique. Il faut dire que cette discussion demandera quelques connaissances de la théorie des corps, polynômes et déterminants.

Supposons que f soit une fonction de matrices carrées qui associe la même valeur à des matrices semblables ; c'est-à-dire que $f(A) = f(B)$ si A est semblable à B. Donc f introduit une fonction notée aussi f, sur les opérateurs linéaires T de façon naturelle : $f(T) = f([T]_e)$ où $\{e_i\}$ est une base quelconque. La fonction est bien définie par le précédent théorème.

Le *déterminant* est peut-être le plus important exemple du type précédent de fonctions. Voici un autre exemple important d'une telle fonction.

Exemple 7.8 : La *trace* d'une matrice carrée $A = (a_{ij})$ écrite $\mathrm{tr}(A)$ est définie comme étant la somme de ses éléments situés sur la diagonale principale

$$\mathrm{tr}(A) = a_{11} + a_{22} + \cdots + a_{nn}$$

Nous montrerons (problème 7.22) que des matrices semblables ont la même trace. Nous pouvons ainsi parler de la trace d'un opérateur linéaire T : c'est la trace d'une quelconque de ses représentations matricielles : $\mathrm{tr}(T) = \mathrm{tr}([T]_e)$.

MATRICES ET APPLICATIONS LINEAIRES

Considérons maintenant le cas général d'applications linéaires d'un espace dans un autre espace. Soient V et U des espaces vectoriels sur le même corps K et soit $\dim V = m$ et $\dim U = n$. De plus, soit $\{e_1, \ldots, e_m\}$ et $\{f_1, \ldots, f_n\}$ des bases arbitraires mais fixées de V et U respectivement.

Supposons $F : V \to U$ une application linéaire. Alors, les vecteurs $F(e_1), \ldots, F(e_m)$ appartiennent à U et donc chacun est une combinaison linéaire des f_i:

$$\begin{aligned} F(e_1) &= a_{11}f_1 + a_{12}f_2 + \cdots + a_{1n}f_n \\ F(e_2) &= a_{21}f_1 + a_{22}f_2 + \cdots + a_{2n}f_n \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ F(e_m) &= a_{m1}f_1 + a_{m2}f_2 + \cdots + a_{mn}f_n \end{aligned}$$

La transposée de la matrice des coefficients ci-dessous, notée par $[F]_e^f$, est appelée la représentation matricielle de F relativement aux bases $\{e_i\}$ et $\{f_i\}$, ou la matrice de F dans les bases $\{e_i\}$ et $\{f_i\}$:

$$[F]_e^f = \begin{pmatrix} a_{11} & a_{21} & \ldots & a_{m1} \\ a_{12} & a_{22} & \ldots & a_{m2} \\ \ldots & \ldots & \ldots & \ldots \\ a_{1n} & a_{2n} & \ldots & a_{mn} \end{pmatrix}$$

On a alors les théorèmes suivants.

Théorème 7.8 : Quel que soit le vecteur $v \in V$, $[F]_e^f\,[v]_e = [F(v)]_f$.

C'est-à-dire, en multipliant le vecteur coordonné de v dans la base $\{e_i\}$ par la matrice $[F]_e^f$, nous obtenons le vecteur coordonné de $F(v)$ dans la base $\{f_i\}$.

Théorème 7.9 : L'application $F \mapsto [F]_e^f$ est un isomorphisme de Hom (V, U) sur l'espace vectoriel des matrices $n \times m$ sur le corps K. C'est-à-dire que l'application est injective et surjective et quels que soient $F, G \in$ Hom (V, U) et quel que soit $k \in K$,

$$[F+G]_e^f = [F]_e^f + [G]_e^f \qquad \text{et} \qquad [kF]_e^f = k[F]_e^f$$

Remarque : Rappelons qu'une matrice quelconque $n \times m$ A sur K a été identifiée à un opérateur linéaire de K^m dans K^n donné par $v \mapsto Av$. Supposons maintenant que V et U soient des espaces vectoriels sur K de dimensions respectives m et n et supposons que $\{e_i\}$ soit une base de V et $\{f_i\}$ une base de U. Alors, eu égard au théorème précédent, nous identifierons aussi A avec l'application linéaire $F : V \to U$ donnée par $[F(v)]_f = A\,[v]_e$. Nous pouvons dire que si d'autres bases de V et U sont données, alors A est identifiée avec une autre application linéaire de V dans U.

Théorème 7.10 : Soient $\{e_i\}$, $\{f_i\}$ et $\{g_i\}$ des bases de V, U et W respectivement. Soit $F : V \to U$ et $G : U \to W$ des applications linéaires. Alors,

$$[G \circ F]_e^g = [G]_f^g\,[F]_e^f$$

C'est-à-dire, relativement à des bases appropriées, que la représentation matricielle de la composition de deux applications linéaires est égale au produit des représentations matricielles de chacune des applications.

Nous montrons enfin comment la représentation matricielle d'une application linéaire $F : V \to U$ est affectée lorsqu'on prend de nouvelles bases.

Théorème 7.11 : Soit P la matrice de passage d'une base $\{e_i\}$ à une base $\{e'_i\}$ dans V, et soit Q la matrice de passage d'une base $\{f_i\}$ à une base $\{f'_i\}$ dans U. Alors, pour une application linéaire quelconque $F : V \to U$:

$$[F]_{e'}^{f'} = Q^{-1}\,[F]_e^f\,P$$

Ainsi, en particulier,

$$[F]_e^{f'} = Q^{-1}[F]_e^f$$

c'est-à-dire lorsque le changement de base opère uniquement dans U, et

$$[F]_{e'}^f = [F]_e^f P$$

lorsque le changement de base opère uniquement dans V.

Remarquons que les théorèmes 7.1, 7.2, 7.3 et 7.5 sont des cas particuliers des théorèmes 7.8, 7,9, 7.10 et 7.11 respectivement.

Le théorème suivant montre que chaque application linéaire d'un espace dans un autre peut être représentée par une matrice très simple.

Théorème 7.12 : Soit $F : V \to U$ une application linéaire, et rang $F = r$. Il existe alors des bases de V et de U telles que la représentation matricielle de F a la forme

$$A = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}$$

où I est la matrice carrée identité d'ordre r. Nous appelons A la forme normale ou canonique de F.

REMARQUE

Comme il était dit précédemment, on rencontre dans certains textes le symbole opérateur T à droite du vecteur v, on écrit alors :

$$vT \quad \text{au lieu de} \quad T(v)$$

Dans de tels textes, les vecteurs et opérateurs sont représentés par des n-tuples et des matrices qui sont les transposées des matrices que nous avons utilisées. C'est-à-dire, si :

$$v = k_1e_1 + k_2e_2 + \cdots + k_ne_n$$

alors on écrit

$$[v]_e = (k_1, k_2, \ldots, k_n) \quad \text{au lieu de} \quad [v]_e = \begin{pmatrix} k_1 \\ k_2 \\ \cdots \\ k_n \end{pmatrix}$$

Et si,

$$\begin{aligned} T(e_1) &= a_1e_1 + a_2e_2 + \cdots + a_ne_n \\ T(e_2) &= b_1e_1 + b_2e_2 + \cdots + b_ne_n \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ T(e_n) &= c_1e_1 + c_2e_2 + \cdots + c_ne_n \end{aligned}$$

alors on écrit

$$[T]_e = \begin{pmatrix} a_1 & a_2 & \ldots & a_n \\ b_1 & b_2 & \ldots & b_n \\ \ldots & \ldots & \ldots & \ldots \\ c_1 & c_2 & \ldots & c_n \end{pmatrix} \quad \text{au lieu de} \quad [T]_e = \begin{pmatrix} a_1 & b_1 & \ldots & c_1 \\ a_2 & b_2 & \ldots & c_2 \\ \ldots & \ldots & \ldots & \ldots \\ a_n & b_n & \ldots & c_n \end{pmatrix}$$

Ceci est aussi vrai pour la matrice de passage d'une base à une autre et pour les représentations matricielles d'applications linéaires $F : V \to U$. De tels textes ont évidemment des théorèmes analogues à ceux que nous avons donnés ici.

PROBLEMES RESOLUS

REPRESENTATIONS MATRICIELLES D'OPERATEURS LINEAIRES :

7.1. Trouver la représentation matricielle de chacun des opérateurs T de $\mathbf{R}^2$ relativement à la base usuelle $\{e_1 = (1, 0), e_2 = (0, 1)\}$.

(i) $T(x, y) = (2y, 3x - y)$, (ii) $T(x, y) = (3x - 4y, x + 5y)$.

Remarquons d'abord que si $(a, b) \in \mathbf{R}^2$, alors $(a, b) = ae_1 + be_2$.

(i) $$\begin{aligned} T(e_1) &= T(1,0) = (0,3) = 0e_1 + 3e_2 \\ T(e_2) &= T(0,1) = (2,-1) = 2e_1 - e_2 \end{aligned} \quad \text{et} \quad [T]_e = \begin{pmatrix} 0 & 2 \\ 3 & -1 \end{pmatrix}.$$

(ii) $$\begin{aligned} T(e_1) &= T(1,0) = (3,1) = 3e_1 + e_2 \\ T(e_2) &= T(0,1) = (-4,5) = -4e_1 + 5e_2 \end{aligned} \quad \text{et} \quad [T]_e = \begin{pmatrix} 3 & -4 \\ 1 & 5 \end{pmatrix}.$$

7.2. Trouver la représentation matricielle de chacun des opérateurs T dans le précédent problème relativement à la base $\{f_1 = (1, 3), f_2 = (2, 5)\}$.

Nous devons d'abord trouver les coordonnées d'un vecteur arbitraire $(a, b) \in \mathbf{R}^2$ relativement à la base $\{f_i\}$. Nous avons

$$(a, b) = x(1, 3) + y(2, 5) = (x + 2y, 3x + 5y)$$

ou $$x + 2y = a \quad \text{et} \quad 3x + 5y = b$$

ou $$x = 2b - 5a \quad \text{et} \quad y = 3a - b$$

Ainsi $$(a, b) = (2b - 5a)f_1 + (3a - b)f_2$$

(i) Nous avons $T(x, y) = (2y, 3x - y)$. Donc

$$\begin{aligned} T(f_1) &= T(1,3) = (6,0) = -30f_1 + 18f_2 \\ T(f_2) &= T(2,5) = (10,1) = -48f_1 + 29f_2 \end{aligned} \quad \text{et} \quad [T]_f = \begin{pmatrix} -30 & -48 \\ 18 & 29 \end{pmatrix}$$

(ii) Nous avons $T(x, y) = (3x - 4y, x + 5y)$. Donc

$$\begin{aligned} T(f_1) &= T(1,3) = (-9,16) = 77f_1 - 43f_2 \\ T(f_2) &= T(2,5) = (-14,27) = 124f_1 - 69f_2 \end{aligned} \quad \text{et} \quad [T]_f = \begin{pmatrix} 77 & 124 \\ -43 & -69 \end{pmatrix}$$

7.3. Supposons que T soit un opérateur linéaire de $\mathbf{R}^3$ défini par

$$T(x, y, z) = (a_1x + a_2y + a_3z,\ b_1x + b_2y + b_3z,\ c_1x + c_2y + c_3z)$$

Montrons que la matrice de T dans la base usuelle $\{e_i\}$ est donnée par

$$[T]_e = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}$$

C'est-à-dire, les lignes de $[T]_e$ sont obtenues à partir des coefficients de x, y, z dans les composantes de $T(x, y, z)$.

$$\begin{aligned} T(e_1) &= T(1, 0, 0) = (a_1, b_1, c_1) = a_1e_1 + b_1e_2 + c_1e_3 \\ T(e_2) &= T(0, 1, 0) = (a_2, b_2, c_2) = a_2e_1 + b_2e_2 + c_2e_3 \\ T(e_3) &= T(0, 0, 1) = (a_3, b_3, c_3) = a_3e_1 + b_3e_2 + c_3e_3 \end{aligned}$$

En conséquence $$[T]_e = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}$$

Remarque : Cette propriété reste vraie pour un espace quelconque K^n, mais uniquement relativement à la base usuelle

$$\{e_1 = (1, 0, \ldots, 0),\ e_2 = (0, 1, 0, \ldots, 0),\ \ldots,\ e_n = (0, \ldots, 0, 1)\}$$

7.4. Trouver la représentation matricielle de chacun des opérateurs linéaires suivants T sur $\mathbf{R}^3$ relativement à la base usuelle $\{e_1 = (1, 0, 0), e_2 = (0, 1, 0), e_3 = (0, 0, 1)\}$:

(i) $T(x, y, z) = (2x - 3y + 4z, 5x - y + 2z, 4x + 7y)$,

(ii) $T(x, y, z) = (2y + z, x - 4y, 3x)$.

D'après le problème 7.3 : (i) $[T]_e = \begin{pmatrix} 2 & -3 & 4 \\ 5 & -1 & 2 \\ 4 & 7 & 0 \end{pmatrix}$, (ii) $[T]_e = \begin{pmatrix} 0 & 2 & 1 \\ 1 & -4 & 0 \\ 3 & 0 & 0 \end{pmatrix}$.

7.5. Soit T un opérateur linéaire sur $\mathbf{R}^3$ défini par $T(x, y, z) = (2y + z, x - 4y, 3x)$.

(i) Trouver la matrice de T dans la base $\{f_1 = (1, 1, 1), f_2 = (1, 1, 0), f_3 = (1, 0, 0)\}$.

(ii) Vérifier que $[T]_f\,[v]_f = [T(v)]_f$ pour un vecteur quelconque $v \in \mathbf{R}^3$.

Nous devons d'abord trouver les coordonnées d'un vecteur arbitraire $(a, b, c) \in \mathbf{R}^3$, par rapport à la base $\{f_1, f_2, f_3\}$. Ecrivons (a, b, c) sous forme de combinaison linéaire des f_i à l'aide des scalaires inconnus x, y, z:

$$\begin{aligned}(a, b, c) &= x(1, 1, 1) + y(1, 1, 0) + z(1, 0, 0) \\ &= (x + y + z, x + y, x)\end{aligned}$$

En égalant les composantes correspondantes entre elles, on obtient le système d'équations

$$x + y + z = a, \qquad x + y = b, \qquad x = c$$

En résolvant ce système en x, y, z en fonction de a, b, c on trouve $x = c$, $y = b - c$, $z = a - b$. Ainsi

$$(a, b, c) = cf_1 + (b - c)f_2 + (a - b)f_3$$

(i) Puisque $T(x, y, z) = (2y + z, x - 4y, 3x)$

$$\begin{aligned} T(f_1) &= T(1,1,1) = (3,-3,3) = 3f_1 - 6f_2 + 6f_3 \\ T(f_2) &= T(1,1,0) = (2,-3,3) = 3f_1 - 6f_2 + 5f_3 \\ T(f_3) &= T(1,0,0) = (0,1,3) = 3f_1 - 2f_2 - f_3 \end{aligned} \quad \text{et} \quad [T]_f = \begin{pmatrix} 3 & 3 & 3 \\ -6 & -6 & -2 \\ 6 & 5 & -1 \end{pmatrix}$$

(ii) Supposons $v = (a, b, c)$; alors

$$v = (a, b, c) = cf_1 + (b - c)f_2 + (a - b)f_3 \quad \text{et donc} \quad [v]_f = \begin{pmatrix} c \\ b - c \\ a - b \end{pmatrix}$$

On a aussi,

$$\begin{aligned} T(v) &= T(a, b, c) = (2b + c, a - 4b, 3a) \\ &= 3af_1 + (-2a - 4b)f_2 + (-a + 6b + c)f_3 \end{aligned} \quad \text{et donc} \quad [T(v)]_f = \begin{pmatrix} 3a \\ -2a - 4b \\ -a + 6b + c \end{pmatrix}$$

Ainsi

$$[T]_f\,[v]_f = \begin{pmatrix} 3 & 3 & 3 \\ -6 & -6 & -2 \\ 6 & 5 & -1 \end{pmatrix}\begin{pmatrix} c \\ b - c \\ a - b \end{pmatrix} = \begin{pmatrix} 3a \\ -2a - 4b \\ -a + 6b + c \end{pmatrix} = [T(v)]_f$$

7.6. Soit $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ et T l'opérateur linéaire de $\mathbf{R}^2$ défini par $T(v) = Av$ (où v est écrit sous forme de vecteur colonne). Trouver la matrice de T dans chacune des bases suivantes :

(i) $\{e_1 = (1, 0), e_2 = (0, 1)\}$ c'est-à-dire la base usuelle ;

(ii) $\{f_1 = (1, 3), f_2 = (2, 5)\}$.

(i)
$$\begin{aligned} T(e_1) &= \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix} = 1e_1 + 3e_2 \\ T(e_2) &= \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix} = 2e_1 + 4e_2 \end{aligned} \quad \text{et ainsi} \quad [T]_e = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}.$$

Observons que la matrice de T dans la base usuelle est précisément la matrice initiale A qui définissait T. Ceci n'est pas inhabituel. En fait, nous montrerons dans le problème suivant que ceci est vrai pour une matrice quelconque A lorsqu'on utilise la base usuelle.

(ii) D'après le problème 7.2, $(a, b) = (2b - 5a) f_1 + (3a - b) f_2$. Donc

$$T(f_1) = \begin{pmatrix}1 & 2\\3 & 4\end{pmatrix}\begin{pmatrix}1\\3\end{pmatrix} = \begin{pmatrix}7\\15\end{pmatrix} = -5f_1 + 6f_2$$
$$T(f_2) = \begin{pmatrix}1 & 2\\3 & 4\end{pmatrix}\begin{pmatrix}2\\5\end{pmatrix} = \begin{pmatrix}12\\26\end{pmatrix} = -8f_1 + 10f_2$$
et donc
$$[T]_f = \begin{pmatrix}-5 & -8\\6 & 10\end{pmatrix}$$

7.7. Rappelons qu'une matrice carrée quelconque n, $A = (a_{ij})$ peut être considérée comme un opérateur linéaire T sur K^n défini par $T(v) = Av$, où v est écrit sous la forme d'un vecteur colonne. Montrer que la représentation matricielle de T relativement à la base usuelle $\{e_i\}$ de K^n est la matrice A, c'est-à-dire $[T]_e = A$.

$$T(e_1) = Ae_1 = \begin{pmatrix}a_{11} & a_{12} & \dots & a_{1n}\\a_{21} & a_{22} & \dots & a_{2n}\\ \dots & \dots & \dots & \dots \\a_{n1} & a_{n2} & \dots & a_{nn}\end{pmatrix}\begin{pmatrix}1\\0\\ \vdots\\0\end{pmatrix} = \begin{pmatrix}a_{11}\\a_{21}\\ \vdots\\a_{n1}\end{pmatrix} = a_{11}e_1 + a_{21}e_2 + \cdots + a_{n1}e_n$$

$$T(e_2) = Ae_2 = \begin{pmatrix}a_{11} & a_{12} & \dots & a_{1n}\\a_{21} & a_{22} & \dots & a_{2n}\\ \dots & \dots & \dots & \dots \\a_{n1} & a_{n2} & \dots & a_{nn}\end{pmatrix}\begin{pmatrix}0\\1\\ \vdots\\0\end{pmatrix} = \begin{pmatrix}a_{12}\\a_{22}\\ \vdots\\a_{n2}\end{pmatrix} = a_{12}e_1 + a_{22}e_2 + \cdots + a_{n2}e_n$$

....

$$T(e_n) = Ae_n = \begin{pmatrix}a_{11} & a_{12} & \dots & a_{1n}\\a_{21} & a_{22} & \dots & a_{2n}\\ \dots & \dots & \dots & \dots \\a_{n1} & a_{n2} & \dots & a_{nn}\end{pmatrix}\begin{pmatrix}0\\0\\ \vdots\\1\end{pmatrix} = \begin{pmatrix}a_{1n}\\a_{2n}\\ \vdots\\a_{nn}\end{pmatrix} = a_{1n}e_1 + a_{2n}e_2 + \cdots + a_{nn}e_n$$

(C'est-à-dire $T(e_i) = Ae_i$ est la $i$ème colonne de A). Donc

$$[T]_e = \begin{pmatrix}a_{11} & a_{12} & \dots & a_{1n}\\a_{21} & a_{22} & \dots & a_{2n}\\ \dots & \dots & \dots & \dots \\a_{n1} & a_{n2} & \dots & a_{nn}\end{pmatrix} = A$$

7.8. Chacun des ensembles (i) $\{1, t, e^t, te^t\}$ et (ii) $\{e^{3t}, te^{3t}, t^2e^{3t}\}$ est une base d'un espace vectoriel V de fonctions $f : \mathbf{R} \to \mathbf{R}$. Soit D l'opérateur différentiel de V, tel que $D(f) = df/dt$. Trouver la matrice de D dans la base donnée.

(i)
$$\begin{aligned} D(1) &= 0 &&= 0(1) + 0(t) + 0(e^t) + 0(te^t)\\ D(t) &= 1 &&= 1(1) + 0(t) + 0(e^t) + 0(te^t)\\ D(e^t) &= e^t &&= 0(1) + 0(t) + 1(e^t) + 0(te^t)\\ D(te^t) &= e^t + te^t &&= 0(1) + 0(t) + 1(e^t) + 1(te^t)\end{aligned}$$
et
$$[D] = \begin{pmatrix}0 & 1 & 0 & 0\\0 & 0 & 0 & 0\\0 & 0 & 1 & 1\\0 & 0 & 0 & 1\end{pmatrix}$$

(ii)
$$\begin{aligned} D(e^{3t}) &= 3e^{3t} &&= 3(e^{3t}) + 0(te^{3t}) + 0(t^2e^{3t})\\ D(te^{3t}) &= e^{3t} + 3te^{3t} &&= 1(e^{3t}) + 3(te^{3t}) + 0(t^2e^{3t})\\ D(t^2e^{3t}) &= 2te^{3t} + 3t^2e^{3t} &&= 0(e^{3t}) + 2(te^{3t}) + 3(t^2e^{3t})\end{aligned}$$
et
$$[D] = \begin{pmatrix}3 & 1 & 0\\0 & 3 & 2\\0 & 0 & 3\end{pmatrix}$$

7.9. Démontrer le théorème 7.1 : Supposons que $\{e_1, \ldots, e_n\}$ soit une base de V et T un opérateur linéaire sur V. Alors pour tout $v \in V$, $[T]_e\,[v]_e = [T(v)]_e$.

Supposons que, pour $i = 1, \ldots, n$.

$$T(e_i) = a_{i1}e_1 + a_{i2}e_2 + \cdots + a_{in}e_n = \sum_{j=1}^{n} a_{ij}e_j$$

Alors $[T]_e$ est la matrice carrée d'ordre n dont la $j$ème ligne est :

$$(a_{1j}, a_{2j}, \ldots, a_{nj}) \tag{1}$$

Supposons maintenant que $v = k_1e_1 + k_2e_2 + \cdots + k_ne_n = \sum_{i=1}^{n} k_ie_i$

En écrivant un vecteur colonne comme transposé d'un vecteur ligne,

$$[v]_e = (k_1, k_2, \ldots, k_n)^t \tag{2}$$

De plus en utilisant la linéarité de T,

$$\begin{aligned} T(v) &= T\left(\sum_{i=1}^{n} k_ie_i\right) = \sum_{i=1}^{n} k_iT(e_i) = \sum_{i=1}^{n} k_i\left(\sum_{j=1}^{n} a_{ij}e_j\right) \\ &= \sum_{j=1}^{n}\left(\sum_{i=1}^{n} a_{ij}k_i\right)e_j = \sum_{j=1}^{n}(a_{1j}k_1 + a_{2j}k_2 + \cdots + a_{nj}k_n)e_j \end{aligned}$$

Ainsi $[T(v)]_e$ est un vecteur colonne, dont le $j$ème élément est

$$a_{1j}k_1 + a_{2j}k_2 + \cdots + a_{nj}k_n \tag{3}$$

D'autre part, le $j$ème élément de $[T]_e\,[v]_e$ est obtenu en multipliant la $j$ième ligne de $[T]_e$ par $[v]_e$ c'est-à-dire (*1*) par (*2*). Mais le produit de (*1*) et (*2*) est (*3*) ; donc $[T]_e\,[v]_e$ et $[T(v)]_e$ ont les mêmes éléments. Ainsi $[T]_e\,[v]_e = [T(v)]_e$.

7.10. Démontrer le théorème 7.2 : Soit $\{e_1, e_2, \ldots, e_n\}$ une base de V sur K, et soit $\mathcal{A}$ l'algèbre des matrices carrées $n \times n$ sur K. Alors l'application $T \mapsto [T]_e$ est un isomorphisme d'espace vectoriel de $A(V)$ sur $\mathcal{A}$. C'est-à-dire que l'application est à la fois injective et surjective et quels que soient $S, T \in A(V)$ et quel que soit $k \in K$, $[T+S]_e = [T]_e + [S]_e$ et $[kT]_e = k[T]_e$.

L'application est injective puisque, d'après le théorème 8.1, une application linéaire est complètement déterminée par les images de la base. L'application est surjective puisque chaque matrice $M \in \mathcal{A}$ est l'image de l'opérateur linéaire

$$F(e_i) = \sum_{j=1}^{n} m_{ij}\,e_j \qquad i = 1, \ldots, n$$

où (m_{ij}) est la transposée de la matrice M.

Supposons maintenant que pour $i = 1, \ldots, n$,

$$T(e_i) = \sum_{j=1}^{n} a_{ij}e_j \quad \text{et} \quad S(e_i) = \sum_{j=1}^{n} b_{ij}e_j$$

Soient A et B les matrices $A = (a_{ij})$ et $B = (b_{ij})$. Donc $[T]_e = A^t$ et $[S]_e = B^t$. Nous avons pour $i = 1, \ldots, n$,

$$(T+S)(e_i) = T(e_i) + S(e_i) = \sum_{j=1}^{n}(a_{ij} + b_{ij})e_j$$

Remarquons que $A + B$ est la matrice $(a_{ij} + b_{ij})$. En conséquence

$$[T+S]_e = (A+B)^t = A^t + B^t = [T]_e + [S]_e$$

Nous avons aussi, pour $i = 1, \ldots, n$,

$$(kT)(e_i) = k\,T(e_i) = k\sum_{j=1}^{n} a_{ij}e_j = \sum_{j=1}^{n}(ka_{ij})e_j$$

Remarquons que kA est la matrice (ka_{ij}). En conséquence

$$[kT]_e = (kA)^t = kA^t = k[T]_e$$

Le théorème est ainsi démontré.

7.11. Démontrer le théorème 7.3 : Soit $\{e_1, \ldots, e_n\}$ une base de V. Alors quels que soient les opérateurs linéaires S, $T \in A(V)$, $[ST]_e = [S]_e\ [T]_e$.

Supposons $T(e_i) = \sum_{j=1}^{n} a_{ij}\, e_j$ et $S(e_j) = \sum_{k=1}^{n} b_{jk}\, e_k$. Soient A et B les matrices $A = (a_{ij})$ et $B = (b_{ij})$. Alors $[T]_e = A^t$ et $[S]_e = B^t$. Nous avons

$$\begin{aligned} (ST)(e_i) &= S(T(e_i)) = S\left(\sum_{j=1}^{n} a_{ij}e_j\right) = \sum_{j=1}^{n} a_{ij}S(e_j) \\ &= \sum_{j=1}^{n} a_{ij}\left(\sum_{k=1}^{n} b_{jk}e_k\right) = \sum_{k=1}^{n}\left(\sum_{j=1}^{n} a_{ij}b_{jk}\right)e_k \end{aligned}$$

Rappelons que AB est la matrice $AB = (c_{ik})$ où $\sum_{j=1}^{n} a_{ij}\, b_{jk}$. En conséquence,

$$[ST]_e = (AB)^t = B^tA^t = [S]_e[T]_e$$

CHANGEMENT DE BASE, MATRICES SEMBLABLES

7.12. Considérons les bases suivantes de $\mathbf{R}^2$: $\{e_1 = (1, 0), e_2 = (0, 1)\}$ et $\{f_1 = (1, 3)\ f_2 = (2, 5)\}$. 1) Trouver la matrice de passage P de la base $\{e_i\}$ à la base $\{f_i\}$. 2) Trouver la matrice de passage Q de $\{f_i\}$ à $\{e_i\}$. 3) Vérifier que $Q = P^{-1}$. 4) Montrer que $[v]_f = P^{-1}[v]_e$ pour tout vecteur $v \in \mathbf{R}^2$. 5) Montrer que $[T]_f = P^{-1}[T]_eP$. pour l'opérateur T de $\mathbf{R}^2$ défini par $T(x, y) = (2y, 3x - y)$ (Voir problèmes 7.1 et 7.2).

1)
$$\begin{aligned} f_1 &= (1,3) = 1e_1 + 3e_2 \\ f_2 &= (2,5) = 2e_1 + 5e_2 \end{aligned} \qquad \text{et} \qquad P = \begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix}$$

2) D'après le problème 7.2, $(a, b) = (2b - 5a)\ f_1 + (3a - b)\ f_2$. Ainsi

$$\begin{aligned} e_1 &= (1,0) = -5f_1 + 3f_2 \\ e_2 &= (0,1) = \ \ 2f_1 - f_2 \end{aligned} \qquad \text{et} \qquad Q = \begin{pmatrix} -5 & 2 \\ 3 & -1 \end{pmatrix}$$

3)
$$PQ = \begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix}\begin{pmatrix} -5 & 2 \\ 3 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$$

4) Si $v = (a, b)$, alors $[v]_e = \begin{pmatrix} a \\ b \end{pmatrix}$ et $[v]_f = \begin{pmatrix} 2b - 5a \\ 3a - b \end{pmatrix}$. Ainsi

$$P^{-1}[v]_e = \begin{pmatrix} -5 & 2 \\ 3 & -1 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} -5a + 2b \\ 3a - b \end{pmatrix} = [v]_f$$

5) D'après les problèmes 7.1 et 7.2, $[T]_e = \begin{pmatrix} 0 & 2 \\ 3 & -1 \end{pmatrix}$ et $[T]_f = \begin{pmatrix} -30 & -48 \\ 18 & 29 \end{pmatrix}$. Donc

$$P^{-1}[T]_eP = \begin{pmatrix} -5 & 2 \\ 3 & -1 \end{pmatrix}\begin{pmatrix} 0 & 2 \\ 3 & -1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix} = \begin{pmatrix} -30 & -48 \\ 18 & 29 \end{pmatrix} = [T]_f$$

7.13. Considérons les bases suivantes de $\mathbf{R}^3$: $\{e_1 = (1, 0, 0), e_2 = (0, 1, 0), e_3 = (0, 0, 1)\}$ et $\{f_1 = (1, 1, 1), f_2 = (1, 1, 0), f_3 = (1, 0, 0)\}$. 1) Trouver la matrice de passage P de $\{e_i\}$ à $\{f_i\}$. 2) Trouver la matrice de passage Q de $\{f_i\}$ à $\{e_i\}$. 3) Vérifier que $Q = P^{-1}$. 4) Montrer que $[v]_f = P^{-1}[v]_e$ pour tout vecteur $v \in \mathbf{R}^3$. 5) Montrer que $[T]_f = P^{-1}[T]_eP$ pour T défini par $T(x, y, z) = (2y + z, x - 4y, 3x)$ (voir Problèmes 7.4 et 7.5).

1)
$$\begin{aligned} f_1 &= (1,1,1) = 1e_1 + 1e_2 + 1e_3 \\ f_2 &= (1,1,0) = 1e_1 + 1e_2 + 0e_3 \\ f_3 &= (1,0,0) = 1e_1 + 0e_2 + 0e_3 \end{aligned} \qquad \text{et} \qquad P = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

2) D'après le problème 7.5, $(a, b, c) = cf_1 + (b - c) f_2 + (a - b) f_3$. Ainsi

$$\begin{aligned} e_1 &= (1,0,0) = 0f_1 + 0f_2 + 1f_3 \\ e_2 &= (0,1,0) = 0f_1 + 1f_2 - 1f_3 \\ e_3 &= (0,0,1) = 1f_1 - 1f_2 + 0f_3 \end{aligned} \quad \text{et} \quad Q = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{pmatrix}$$

3) $$PQ = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I$$

4) Si $v = (a, b, c)$, alors $[v]_e = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$ et $[v]_f = \begin{pmatrix} c \\ b - c \\ a - b \end{pmatrix}$. Ainsi

$$P^{-1}[v]_e = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} c \\ b-c \\ a-b \end{pmatrix} = [v]_f$$

5) D'après les problèmes 7.4 (2) et 7.5, $[T]_e = \begin{pmatrix} 0 & 2 & 1 \\ 1 & -4 & 0 \\ 3 & 0 & 0 \end{pmatrix}$ et $[T]_f = \begin{pmatrix} 3 & 3 & 3 \\ -6 & -6 & -2 \\ 6 & 5 & -1 \end{pmatrix}$. Ainsi

$$P^{-1}[T]_e P = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 2 & 1 \\ 1 & -4 & 0 \\ 3 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 3 & 3 \\ -6 & -6 & -2 \\ 6 & 5 & -1 \end{pmatrix} = [T]_f$$

7.14. Démontrer le théorème 7.4 : Soit P la matrice de passage de la base $\{e_i\}$ à la base $\{f_i\}$ dans un espace vectoriel V. Pour tout vecteur $v \in V$, $P[v]_f = [v]_e$. Aussi $[v]_f = P^{-1}[v]_e$.

Supposons, pour $i = 1, \ldots, n$, $f_i = a_{i1} e_1 + a_{i2} e_2 + \cdots + a_{in} e_n = \sum_{j=1}^{n} a_{ij} e_j$. Alors P est la matrice carrée $n \times n$ dont la $j^{\text{ème}}$ ligne est

$$(a_{1j}, a_{2j}, \ldots, a_{nj}) \tag{1}$$

Supposons aussi $v = k_1 f_1 + k_2 f_2 + \ldots + k_n f_n = \sum_{i=1}^{n} k_i f_i$. Alors en écrivant un vecteur colonne comme le transposé d'un vecteur ligne,

$$[v]_f = (k_1, k_2, \ldots, k_n)^t \tag{2}$$

En remplaçant les f_i par leurs valeurs dans l'équation donnant v,

$$\begin{aligned} v &= \sum_{i=1}^{n} k_i f_i = \sum_{i=1}^{n} k_i \left(\sum_{j=1}^{n} a_{ij} e_j \right) = \sum_{j=1}^{n} \left(\sum_{i=1}^{n} a_{ij} k_i \right) e_j \\ &= \sum_{j=1}^{n} (a_{1j}k_1 + a_{2j}k_2 + \cdots + a_{nj}k_n) e_j \end{aligned}$$

En conséquence $[v]_e$ est le vecteur colonne dont le $j^{\text{ème}}$ élément est :

$$a_{1j} k_1 + a_{2j} k_2 + \cdots + a_{nj} k_n \tag{3}$$

D'autre part, le $j^{\text{ème}}$ élément de $P[v]_f$ est obtenu en multipliant la $j^{\text{ème}}$ ligne de P par $[v]_f$ c'est-à-dire (1) et (2). Mais le produit de (1) et (2) est (3) ; donc $P[v]_f$ et $[v]_e$ ont les mêmes éléments et donc $P[v]_f = [v]_e$.

De plus, en multipliant l'égalité précédente par P^{-1} on obtient $P^{-1}[v]_e = P^{-1}P[v]_f = [v]_f$.

7.15. Démontrer le théorème 7.5 : Soit P la matrice de passage d'une base $\{e_i\}$ à la base $\{f_i\}$ dans l'espace vectoriel V. Alors, quel que soit l'opérateur linéaire T sur V, $[T]_f = P^{-1}[T]_e P$.

Pour un vecteur quelconque $v \in V$, $P^{-1}[T]_e P[v]_f = P^{-1}[T]_e[v]_e = [T(v)]_e = [T(v)]_f$.

Mais $[T]_f[v]_f = [T(v)]_f$; donc $P^{-1}[T]_e P[v]_f = [T]_f[v]_f$.

Puisque l'application $v \mapsto [v]_f$ est surjective sur K^n , $P^{-1}[T]_e PX = [T]_f X$ quel que soit $X \in K^n$.

En conséquence $P^{-1}[T]_e P = [T]_f$.

7.16. Montrer que la similitude des matrices est une relation d'équivalence, c'est-à-dire que 1) A est semblable à A, 2) si A est semblable à B, alors B est semblable à A, 3) si A est semblable à B et B semblable à C, alors A est semblable à C.

1) La matrice identique I est inversible et $I = I^{-1}$. Puisque $A = I^{-1}AI$, A est semblable à A.

2) Puisque A est semblable à B, il existe une matrice inversible P telle que $A = P^{-1}BP$. Donc $B = PAP^{-1} = (P^{-1})^{-1}AP^{-1}$ et P^{-1} est inversible. Ainsi B est semblable à A.

3) Puisque A est semblable à B, il existe une matrice inversible P telle que $A = P^{-1}BP$ et puisque B est semblable à C, il existe une matrice inversible Q telle que $B = Q^{-1}CQ$. Donc $A = P^{-1}BP = P^{-1}(Q^{-1}CQ)P$ $(QP)^{-1}C(QP)$ et QP est inversible. Donc A est semblable à C.

TRACE

7.17. La trace d'une matrice carrée $A = (a_{ij})$, que l'on écrit $\mathrm{tr}(A)$, est la somme de ses éléments diagonaux ; $\mathrm{tr}(A) = a_{11} + a_{22} + \cdots + a_{nn}$. Montrer que 1) $\mathrm{tr}(AB) = \mathrm{tr}(BA)$, 2) si A est semblable à B alors $\mathrm{tr}(A) = \mathrm{tr}(B)$.

1) Supposons $A = (a_{ij})$ et $B = (b_{ij})$. Alors $AB = (c_{ik})$ où $c_{ik} = \sum_{j=1}^{n} a_{ij}b_{jk}$. Ainsi

$$\mathrm{tr}(AB) = \sum_{i=1}^{n} c_{ii} = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}b_{ji}$$

D'autre part, $BA = (d_{jk})$ où $d_{jk} = \sum_{i=1}^{n} b_{ji}a_{ik}$. Ainsi

$$\mathrm{tr}(BA) = \sum_{j=1}^{n} d_{jj} = \sum_{j=1}^{n}\sum_{i=1}^{n} b_{ji}a_{ij} = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}b_{ji} = \mathrm{tr}(AB)$$

2) Si A est semblable à B, il existe une matrice inversible P telle que $A = P^{-1}BP$. En utilisant 1),

$$\mathrm{tr}(A) = \mathrm{tr}(P^{-1}BP) = \mathrm{tr}(BPP^{-1}) = \mathrm{tr}(B)$$

7.18 Trouver la trace de l'opérateur suivant de $\mathbf{R}^3$:

$$T(x, y, z) = (a_1x + a_2y + a_3z,\ b_1x + b_2y + b_3z,\ c_1x + c_2y + c_3z)$$

Nous devons d'abord trouver une représentation matricielle de T. En choisissant la base usuelle $\{e_i\}$,

$$[T]_e = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}$$

et $\mathrm{tr}(T) = \mathrm{tr}([T]_e) = a_1 + b_2 + c_3$.

7.19. Soit V l'espace des matrices 2×2 sur $\mathbf{R}$ et soit $M = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. Soit T l'opérateur linéaire de V défini par $T(A) = MA$. Trouver la trace de T.

Nous devons d'abord trouver une représentation matricielle de T. En choisissant la base usuelle de V:

$$\left\{E_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},\ E_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},\ E_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},\ E_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right\}$$

Alors

$$T(E_1) = ME_1 = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 3 & 0 \end{pmatrix} = 1E_1 + 0E_2 + 3E_3 + 0E_4$$

$$T(E_2) = ME_2 = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 3 \end{pmatrix} = 0E_1 + 1E_2 + 0E_3 + 3E_4$$

$$T(E_3) = ME_3 = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 4 & 0 \end{pmatrix} = 2E_1 + 0E_2 + 4E_3 + 0E_4$$

$$T(E_4) = ME_4 = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ 0 & 4 \end{pmatrix} = 0E_1 + 2E_2 + 0E_3 + 4E_4$$

Donc

$$[T]_E = \begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 2 \\ 3 & 0 & 4 & 0 \\ 0 & 3 & 0 & 4 \end{pmatrix}$$

et $\operatorname{tr}(T) = 1 + 1 + 4 + 4 = 10$.

REPRESENTATIONS MATRICIELLES D'APPLICATIONS LINEAIRES

7.20. Soit $F: \mathbf{R}^3 \to \mathbf{R}^2$ l'application linéaire définie par $F(x, y, z) = (3x + 2y - 4z, x - 5y + 3z)$.

1) Trouver la matrice de F dans les bases suivantes de $\mathbf{R}^3$ et $\mathbf{R}^2$

$$\{f_1 = (1,1,1),\ f_2 = (1,1,0),\ f_3 = (1,0,0)\},\quad \{g_1 = (1,3),\ g_2 = (2,5)\}$$

2) Vérifier que l'action de F est préservée par sa représentation matricielle, c'est-à-dire, que pour un vecteur quelconque, $v \in \mathbf{R}^3$, $[F]_f^g\,[v]_f = [F(v)]_g$.

1) D'après le problème 7.2, $(a, b) = (2b - 5a)g_1 + (3a - b)g_2$. Donc

$$\begin{aligned} F(f_1) &= F(1,1,1) = (1,-1) = -7g_1 + 4g_2 \\ F(f_2) &= F(1,1,0) = (5,-4) = -33g_1 + 19g_2 \\ F(f_3) &= F(1,0,0) = (3,1) = -13g_1 + 8g_2 \end{aligned} \quad \text{et} \quad [F]_f^g = \begin{pmatrix} -7 & -33 & -13 \\ 4 & 19 & 8 \end{pmatrix}$$

2) Si $v = (x, y, z)$, alors, d'après le problème 7.5, $v = zf_1 + (y - z)f_2 + (x - y)f_3$. Aussi

$$F(v) = (3x + 2y - 4z, x - 5y + 3z) = (-13x - 20y + 26z)g_1 + (8x + 11y - 15z)g_2$$

Donc $[v]_f = \begin{pmatrix} z \\ y - z \\ x - y \end{pmatrix}$ et $[F(v)]_g = \begin{pmatrix} -13x - 20y + 26z \\ 8x + 11y - 15z \end{pmatrix}$. Ainsi

$$[F]_f^g\,[v]_f = \begin{pmatrix} -7 & -33 & -13 \\ 4 & 19 & 8 \end{pmatrix}\begin{pmatrix} z \\ y - z \\ x - y \end{pmatrix} = \begin{pmatrix} -13x - 20y + 26z \\ 8x + 11y - 15z \end{pmatrix} = [F(v)]_g$$

7.21. Soit $F : K^n \to K^m$ l'application linéaire définie par :

$$F(x_1, x_2, \ldots, x_n) = (a_{11}x_1 + \cdots + a_{1n}x_n,\ a_{21}x_1 + \cdots + a_{2n}x_n,\ \ldots,\ a_{m1}x_1 + \cdots + a_{mn}x_n)$$

Montrer que la représentation matricielle de F relativement aux bases habituelles de K^n et de K^m est donnée par

$$[F] = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$$

C'est-à-dire que les lignes de $[F]$ sont obtenues à partir des coefficients des x_i dans les composantes de $F(x_1, x_2, \ldots, x_n)$ respectivement,

$$\begin{array}{lcl} F(1,0,\ldots,0) & = & (a_{11}, a_{21}, \ldots, a_{m1}) \\ F(0,1,\ldots,0) & = & (a_{12}, a_{22}, \ldots, a_{m2}) \\ \multicolumn{3}{c}{\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots} \\ F(0,0,\ldots,1) & = & (a_{1n}, a_{2n}, \ldots, a_{mn}) \end{array} \quad \text{et} \quad [F] = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$$

7.22. Trouver la représentation matricielle de chacune des applications linéaires suivantes relativement aux bases habituelles de $\mathbf{R}^n$:

1) $F : \mathbf{R}^2 \to \mathbf{R}^3$ définie par $F(x, y) = (3x - y, 2x + 4y, 5x - 6y)$.

2) $F : \mathbf{R}^4 \to \mathbf{R}^2$ définie par $F(x, y, s, t) = (3x - 4y + 2s - 5t, 5x + 7y - s - 2t)$.

3) $F : \mathbf{R}^3 \to \mathbf{R}^4$ définie par $F(x, y, z) = (2x + 3y - 8z, x + y + z, 4x - 5z, 6y)$.

D'après le problème 7.21, nous avons seulement besoin des coefficients des inconnues dans $F(x, y, \ldots)$. Ainsi :

$$\text{(i)}\quad [F] = \begin{pmatrix} 3 & -1 \\ 2 & 4 \\ 5 & -6 \end{pmatrix} \qquad \text{(ii)}\quad [F] = \begin{pmatrix} 3 & -4 & 2 & -5 \\ 5 & 7 & -1 & -2 \end{pmatrix} \qquad \text{(iii)}\quad [F] = \begin{pmatrix} 2 & 3 & -8 \\ 1 & 1 & 1 \\ 4 & 0 & -5 \\ 0 & 6 & 0 \end{pmatrix}$$

7.23. Soit $T : \mathbf{R}^2 \to \mathbf{R}^2$ définie par $T(x, y) = (2x - 3y, x + 4y)$. Trouver la matrice de T dans les bases $\{e_1 = (1, 0), e_2 = (0, 1)\}$ et $\{f_1 = (1, 3), f_2 = (2, 5)\}$ de $\mathbf{R}^2$ respectivement. (Nous pouvons considérer T comme une application linéaire d'un espace dans un autre, chacun ayant sa propre base).

D'après le Problème 7.2, $(a, b) = (2b - 5a) f_1 + (3a - b) f_2$. Alors

$$\begin{array}{lcl} T(e_1) = T(1,0) = (2,1) & = & -8f_1 + 5f_2 \\ T(e_2) = T(0,1) = (-3,4) & = & 23f_1 - 13f_2 \end{array} \quad \text{et} \quad [T]_e^f = \begin{pmatrix} -8 & 23 \\ 5 & -13 \end{pmatrix}$$

7.24. Soit $A = \begin{pmatrix} 2 & 5 & -3 \\ 1 & -4 & 7 \end{pmatrix}$. Rappelons que A détermine une application linéaire $F : \mathbf{R}^3 \to \mathbf{R}^2$ définie par $F(v) = Av$ où v est écrit comme un vecteur colonne.

1) Montrer que la représentation matricielle de F relativement à la base usuelle de $\mathbf{R}^3$ et de $\mathbf{R}^2$ est la matrice A elle-même: $[F] = A$.

2) Trouver la représentation matricielle de F relativement aux bases suivantes de $\mathbf{R}^3$ et $\mathbf{R}^2$.

$$\{f_1 = (1,1,1),\ f_2 = (1,1,0),\ f_3 = (1,0,0)\}, \quad \{g_1 = (1,3),\ g_2 = (2,5)\}$$

1)

$$F(1,0,0) = \begin{pmatrix} 2 & 5 & -3 \\ 1 & -4 & 7 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} = 2e_1 + 1e_2$$

$$F(0,1,0) = \begin{pmatrix} 2 & 5 & -3 \\ 1 & -4 & 7 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 5 \\ -4 \end{pmatrix} = 5e_1 - 4e_2$$

$$F(0,0,1) = \begin{pmatrix} 2 & 5 & -3 \\ 1 & -4 & 7 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -3 \\ 7 \end{pmatrix} = -3e_1 + 7e_2$$

pour lequel $[F] = \begin{pmatrix} 2 & 5 & -3 \\ 1 & -4 & 7 \end{pmatrix} = A$. (Comparer avec le problème 7.7).

2) D'après le problème 7.2, $(a, b) = (2b - 5a) g_1 + (3a - b) g_2$. Alors

$$F(f_1) = \begin{pmatrix} 2 & 5 & -3 \\ 1 & -4 & 7 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \end{pmatrix} = -12g_1 + 8g_2$$

$$F(f_2) = \begin{pmatrix} 2 & 5 & -3 \\ 1 & -4 & 7 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 7 \\ -3 \end{pmatrix} = -41g_1 + 24g_2$$

$$F(f_3) = \begin{pmatrix} 2 & 5 & -3 \\ 1 & -4 & 7 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} = -8g_1 + 5g_2$$

et $\quad [F]^g_f = \begin{pmatrix} -12 & -41 & -8 \\ 8 & 24 & 5 \end{pmatrix}$.

7.25. Démontrer le théorème 7.12 : Soit $F : V \to U$ une application linéaire. Alors il existe une base de V et une base de U telle que la représentation matricielle A de F ait la forme $A = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}$ où I est la matrice identité carrée $r \times r$ et r est le rang de F.

Supposons dim $V = m$ et dim $U = n$. Soit W le noyau de F et U' l'image de F. Nous supposons que rang $F = r$; donc la dimension du noyau de F est $m - r$. Soit $\{w_1, \ldots, w_{m-r}\}$ une base du noyau de F et prolongeons-la pour obtenir une base de V :

$$\{v_1, \ldots, v_r, w_1, \ldots, w_{m-r}\}$$

Posons $\qquad u_1 = F(v_1), u_2 = F(v_2), \ldots, u_r = F(v_r)$

Remarquons que $\{u_1, \ldots, u_n\}$ est une base de U', l'image de F. Prolongeons-la pour obtenir une base.

$$\{u_1, \ldots, u_r, u_{r+1}, \ldots u_n\}$$

de U. Remarquons que

$$\begin{aligned}
F(v_1) &= u_1 = 1u_1 + 0u_2 + \cdots + 0u_r + 0u_{r+1} + \cdots + 0u_n \\
F(v_2) &= u_2 = 0u_1 + 1u_2 + \cdots + 0u_r + 0u_{r+1} + \cdots + 0u_n \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
F(v_r) &= u_r = 0u_1 + 0u_2 + \cdots + 1u_r + 0u_{r+1} + \cdots + 0u_n \\
F(w_1) &= 0 = 0u_1 + 0u_2 + \cdots + 0u_r + 0u_{r+1} + \cdots + 0u_n \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
F(w_{m-r}) &= 0 = 0u_1 + 0u_2 + \cdots + 0u_r + 0u_{r+1} + \cdots + 0u_n
\end{aligned}$$

Ainsi la matrice de F dans les bases précédentes a la forme demandée.

PROBLEMES SUPPLEMENTAIRES

REPRESENTATIONS MATRICIELLES D'OPERATEURS LINEAIRES

7.26. Trouver la matrice de chacun des opérateurs linéaires suivants T de $\mathbf{R}^2$ par rapport à la base ordinaire $\{e_1 = (1, 0), e_2 = (0, 1)\}$: (i) $T(x, y) = (2x - 3y, x + y)$, (ii) $T(x, y) = (5x + y, 3x - 2y)$.

7.27. Trouver la matrice de chacun des opérateurs T dans le précédent problème par rapport à la base $\{f_1 = (1, 2), f_2 = (2, 3)\}$. Dans chaque cas, vérifier que $[T]_f [v]_f = [T(v)]_f$ quel que soit $v \in \mathbf{R}^2$.

7.28. Trouver la matrice de chacun des opérateurs T du problème 7.26 dans la base $\{g_1 = (1, 3), g_2 = (1, 4)\}$.

7.29. Trouver la représentation matricielle de chacun des opérateurs linéaires suivants T de $\mathbf{R}^3$ relativement à la base usuelle.

(i) $T(x, y, z) = (x, y, 0)$

(ii) $T(x, y, z) = (2x - 7y - 4z, 3x + y + 4z, 6x - 8y + z)$

(iii) $T(x, y, z) = (z, y + z, x + y + z)$

7.30. Soit D l'opérateur différentiel, c'est-à-dire $D(f) = df/dt$. Chacun des ensembles suivants constitue une base d'un espace vectoriel V des fonctions $f : \mathbf{R} \to \mathbf{R}$. Trouver la matrice de D dans chacune des bases (i) $\{e^t, e^{2t}, te^{2t}\}$, (ii) $\{\sin t, \cos t\}$, (iii) $\{e^{5t}, te^{5t}, t^2 e^{5t}\}$, (iv) $\{1, t \sin 3t, \cos 3t\}$.

7.31. Considérons le corps des complexes $\mathbf{C}$ comme un espace vectoriel sur le corps des réels $\mathbf{R}$. Soit T l'opérateur de conjugaison sur $\mathbf{C}$ tel que $T(z) = \bar{z}$. Trouver la matrice T dans chaque base (i) $\{1, i\}$, (ii) $\{1 + i, 1 + 2i\}$.

7.32. Soit V l'espace vectoriel des matrices 2×2 sur $\mathbf{R}$ et soit $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Trouver la matrice de chacun des opérateurs linéaires suivants de T sur V dans la base habituelle (voir problème 7.19) de V : (i) $T(A) = MA$, (ii) $T(A) = AM$. (iii) $T(A) = MA - AM$.

7.33. Soit 1_V et 0_V respectivement l'opérateur identique et l'opérateur nul de l'espace vectoriel V. Montrer que, quelle que soit la base $\{e_i\}$ de V, (i) $[1_V]_e = I$ la matrice identité, (ii) $[0_V]_e = 0$ la matrice nulle.

CHANGEMENT DE BASE, MATRICES SEMBLABLES

7.34. Considérons les bases suivantes de $\mathbf{R}^2$: $\{e_1 = (1, 0), e_2 = (0, 1)\}$ et $\{f_1 = (1, 2), f_2 = (2, 3)\}$.

(i) Trouver les matrices de passage P et Q de $\{e_i\}$ à $\{f_i\}$ et de $\{f_i\}$ à $\{e_i\}$ respectivement. Vérifier $Q = P^{-1}$.

(ii) Montrer que $[v]_e = P[v]_f$, quel que soit $v \in \mathbf{R}^2$.

(iii) Montrer que $[T]_f = P^{-1} [T]_e P$ quel que soit l'opérateur T du problème 7.26.

7.35. Reprendre le problème 7.34 pour les bases $\{f_1 = (1, 2), f_2 = (2, 3)\}$ et $\{g_1 = (1, 3), g_2 = (1, 4)\}$.

7.36. Supposons que $\{e_1, e_2\}$ soit une base de V et $T : V \to V$ un opérateur linéaire pour lequel $T(e_1) = 3e_1 - 2e_2$ et $T(e_2) = e_1 + 4e_2$. Supposons que $\{f_1, f_2\}$ soit une base de V pour laquelle $f_1 = e_1 + e_2$ et $f_2 = 2e_1 + 3e_2$. Trouver la matrice de T dans la base $\{f_1, f_2\}$.

7.37. Considérons les bases $B = \{1, i\}$ et $B' = \{1 + i, 1 + 2i\}$ du corps des complexes $\mathbf{C}$ sur le corps des réels $\mathbf{R}$. (i) Trouver les matrices de passage P et Q de B à B' et de B' à B, respectivement. Vérifier que $Q = P^{-1}$. (ii) Montrer que $[T]_{B'} = P^{-1} [T]_B P$ pour l'opérateur conjugué T du problème 7.31.

7.38. Supposons que $\{e_i\}$, $\{f_i\}$ et $\{g_i\}$ soient des bases de V et soient P et Q les matrices de passage de $\{e_i\}$ à $\{f_i\}$ et de $\{f_i\}$ à $\{g_i\}$ respectivement. Montrer que PQ est la matrice de passage de $\{e_i\}$ à $\{g_i\}$.

7.39. Soit A une matrice 2×2 telle que seulement A soit semblable à elle-même. Montrer que A est de la forme

$$A = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$$

Généraliser aux matrices $n \times n$.

7.40. Montrer que toutes les matrices semblables à une matrice inversible sont inversibles. Plus généralement, montrer que des matrices semblables ont le même rang.

REPRESENTATIONS MATRICIELLES D'APPLICATIONS LINEAIRES

7.41. Trouver la représentation matricielle des applications linéaires suivantes relativement aux bases habituelles de $\mathbf{R}^n$:

(i) $F : \mathbf{R}^3 \to \mathbf{R}^2$ définie par $F(x, y, z) = (2x - 4y + 9z, 5x + 3y - 2z)$

(ii) $F : \mathbf{R}^2 \to \mathbf{R}^4$ définie par $F(x, y) = (3x + 4y, 5x - 2y, x + 7y, 4x)$

(iii) $F : \mathbf{R}^4 \to \mathbf{R}$ définie par $F(x, y, s, t) = 2x + 3y - 7s - t$

(iv) $F : \mathbf{R} \to \mathbf{R}^2$ définie par $F(x) = (3x, 5x)$

7.42. Soit $F : \mathbf{R}^3 \to \mathbf{R}^2$ l'application linéaire définie par $F(x, y, z) = (2x + y - z, 3x - 2y + 4z)$.

(i) Trouver la matrice de F dans les bases suivantes de $\mathbf{R}^3$ et $\mathbf{R}^2$:

$$\{f_1 = (1, 1, 1), f_2 = (1, 1, 0), f_3 = (1, 0, 0)\} \text{ et } \{g_1 = (1, 3), g_2 = (1, 4)\}$$

(ii) Vérifier que, pour un vecteur quelconque $v \in \mathbf{R}^3$, $[F]_f^g\,[v]_f = [F(v)]_g$.

7.43. Soient $\{e_i\}$ et $\{f_i\}$ des bases de V, et soit 1_V l'application identique sur V. Montrer que la matrice de 1_V dans les bases $\{e_i\}$ et $\{f_i\}$ est l'inverse de la matrice de passage P de $\{e_i\}$ à $\{f_i\}$; c'est-à-dire $[1_V]_e^f = P^{-1}$.

7.44. Démontrer le théorème 7.7 page 155. (voir Problème 7.9 page 161).

7.45. Démontrer le théorème 7.8. (voir Problème 7.10).

7.46. Démontrer le théorème 7.9 (voir Problème 7.11).

7.47. Démontrer le théorème 7.10 (voir Problème 7.15).

PROBLEMES DIVERS

7.48. Soit T un opérateur linéaire sur V et soit W un sous-espace de V invariant par T, c'est-à-dire tel que $T(W) \subset W$. Supposons dim $W = m$. Montrer que T a une représentation matricielle de la forme $\begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$ où A est une sous-matrice $m \times m$.

7.49. Soit $V = U \oplus W$ et soit U et W chacun étant invariant par l'opérateur linéaire $T : V \to V$. Supposons dim $U = m$ et dim $V = n$. Montrer que T a une représentation matricielle de la forme $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$ où A et B sont des sous-matrices $m \times m$ et $n \times n$ respectivement.

7.50. Rappelons que deux opérateurs linéaires F et G sur V sont dits semblables s'il existe un opérateur inversible T sur V tel que $G = T^{-1}FT$.

(i) Montrer que des opérateurs linéaires F et G sont semblables si et seulement si, pour une base $\{e_i\}$ quelconque de V, les représentations matricielles $[F]_e$ et $[G]_e$ sont des matrices semblables.

(ii) Montrer que si un opérateur F est diagonalisable, alors un opérateur quelconque semblable G est aussi diagonalisable.

7.51. Deux matrices $m \times n$ A et B sur K sont dites équivalentes s'il existe une matrice carrée m inversible Q et une matrice carrée n inversible P telle que $B = QAP$.

(i) Montrer que l'équivalence des matrices est une relation d'équivalence.

(ii) Montrer que A et B peuvent être les représentations matricielles du même opérateur linéaire $F : V \to U$ si et seulement si A et B sont équivalents.

(iii) Montrer que toute matrice A est équivalente à une matrice de la forme $\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}$ où I est la matrice carré r identité et $r = \text{rang } A$.

7.52. Deux algèbres A et B sur un corps K sont dites algèbres isomorphiques s'il existe une application bijective $f : \mathbf{A} \to \mathbf{B}$ telle que pour $u, v \in A$ et $k \in K$, (i) $f(u + v) = f(u) + f(v)$, (ii) $f(ku) = kf(u)$, (iii) $f(uv) = f(u)\,f(v)$. (C'est-à-dire f conserve les trois opérations d'une algèbre : addition vectorielle, multiplication scalaire, et multiplication vectorielle). L'application f est alors appelée un isomorphisme de A sur B. Montrer que la relation d'isomorphisme entre algèbres est une relation d'équivalence.

7.53. Soit $\mathcal{A}$ l'algèbre des matrices carrés n sur K et soit P une matrice inversible de $\mathcal{A}$. Montrer que l'application $A \mapsto P^{-1}AP$, où $A \in \mathcal{A}$ est un isomorphisme d'algèbre de $\mathcal{A}$ dans lui-même.

REPONSES AUX PROBLEMES SUPPLEMENTAIRES

7.26. (i) $\begin{pmatrix} 2 & -3 \\ 1 & 1 \end{pmatrix}$ (ii) $\begin{pmatrix} 5 & 1 \\ 3 & -2 \end{pmatrix}$

7.27. Ici $(a, b) = (2b-3a)f_1 + (2a-b)f_2$. (i) $\begin{pmatrix} 18 & 25 \\ -11 & -15 \end{pmatrix}$ (ii) $\begin{pmatrix} -23 & -39 \\ 15 & 26 \end{pmatrix}$

7.28. Ici $(a, b) = (4a-b)g_1 + (b-3a)g_2$. (i) $\begin{pmatrix} -32 & -45 \\ 25 & 35 \end{pmatrix}$ (ii) $\begin{pmatrix} 35 & 41 \\ -27 & -32 \end{pmatrix}$

7.29. (i) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ (ii) $\begin{pmatrix} 2 & -7 & -4 \\ 3 & 1 & 4 \\ 6 & -8 & 1 \end{pmatrix}$ (iii) $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$

7.30. (i) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$ (ii) $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ (iii) $\begin{pmatrix} 5 & 1 & 0 \\ 0 & 5 & 2 \\ 0 & 0 & 5 \end{pmatrix}$ (iv) $\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -3 \\ 0 & 0 & 3 & 0 \end{pmatrix}$

7.31. (i) $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ (ii) $\begin{pmatrix} -3 & 4 \\ -2 & -3 \end{pmatrix}$

7.32. (i) $\begin{pmatrix} a & 0 & b & 0 \\ 0 & a & 0 & b \\ c & 0 & d & 0 \\ 0 & c & 0 & d \end{pmatrix}$ (ii) $\begin{pmatrix} a & c & 0 & 0 \\ b & d & 0 & 0 \\ 0 & 0 & a & c \\ 0 & 0 & b & d \end{pmatrix}$ (iii) $\begin{pmatrix} 0 & -c & b & 0 \\ -b & a-d & 0 & b \\ c & 0 & d-a & -c \\ 0 & c & -b & 0 \end{pmatrix}$

7.34. $P = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$, $Q = \begin{pmatrix} -3 & 2 \\ 2 & -1 \end{pmatrix}$

7.35. $P = \begin{pmatrix} 3 & 5 \\ -1 & -2 \end{pmatrix}$, $Q = \begin{pmatrix} 2 & 5 \\ -1 & -3 \end{pmatrix}$

7.36. $\begin{pmatrix} 8 & 11 \\ -2 & -1 \end{pmatrix}$

7.37. $P = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$, $Q = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$

7.41. (i) $\begin{pmatrix} 2 & -4 & 9 \\ 5 & 3 & -2 \end{pmatrix}$ (ii) $\begin{pmatrix} 3 & 4 \\ 5 & -2 \\ 1 & 7 \\ 4 & 0 \end{pmatrix}$ (iii) $(2, 3, -7, -1)$ (iv) $\begin{pmatrix} 3 \\ 5 \end{pmatrix}$

7.42. (i) $\begin{pmatrix} 3 & 11 & 5 \\ -1 & -8 & -3 \end{pmatrix}$

CHAPITRE 8

Déterminants

INTRODUCTION

A chaque matrice carrée sur un corps K on peut associer un scalaire appelé le déterminant de A, noté habituellement

$$\det(A) \quad \text{ou} \quad |A|$$

Cette application déterminant a été découverte pour la première fois lors de l'étude de systèmes d'équations linéaires. Nous aurons l'occasion de voir dans les chapitres suivants que le déterminant est indispensable pour l'examen, la recherche et l'obtention des propriétés d'un opérateur linéaire.

Notons que la définition du déterminant et la plupart de ses propriétés sont également valables dans le cas où les éléments d'une matrice appartiennent à un anneau (voir l'Appendice B).

Nous commencerons le chapitre par l'étude des permutations, ce qui est absolument nécessaire dans la définition du déterminant.

PERMUTATIONS

Une application bijective σ de l'ensemble $\{1, 2, \ldots, n\}$ dans lui-même est appelée une permutation. Nous noterons la permutation σ par

$$\sigma = \begin{pmatrix} 1 & 2 & \ldots & n \\ j_1 & j_2 & \ldots & j_n \end{pmatrix} \quad \text{ou bien} \quad \sigma = j_1 j_2 \ldots j_n, \qquad \text{où} \qquad j_i = \sigma(i)$$

Observons que, puisque σ est injective et surjective, la suite $j_1, j_2, \ldots, j_n$ est simplement un réarrangement des nombres $1, 2, \ldots, n$. Remarquons que le nombre de ces permutations est $n!$ et que leur ensemble est noté habituellement S_n. Remarquons aussi que si $\sigma \in S_n$, alors l'application réciproque $\sigma^{-1} \in S_n$; et si $\sigma, \tau \in S_n$ alors l'application composée $\sigma \circ \tau \in S_n$. En particulier, l'application identique

$$\epsilon = \sigma \circ \sigma^{-1} = \sigma^{-1} \circ \sigma$$

appartient aussi à S_n. (En fait $\epsilon = 1\ 2\ \ldots.\ n$).

Exemple 8.1 : Il y a $2! = 2 \times 1 = 2$ permutations dans S_2 : 12 et 21.

Exemple 8.2 : Il y a $3! = 3 \times 2 \times 1 = 6$ permutations dans S_3 : 123, 132, 213, 231, 312, 321.

Considérons une permutation quelconque σ de S_n : $\sigma = j_1\ j_2 \ldots . j_n$. Nous dirons que σ est paire ou impaire selon que nous avons un nombre pair ou impair de couples (i, k) pour lesquels

$$i > k \quad \text{mais} \quad i \text{ précède } k \text{ dans } \sigma \qquad (*)$$

Nous définissons alors la signature ou la parité de σ, et on écrit sgn σ, par

$$\operatorname{sgn} \sigma = \begin{cases} 1 & \text{si } \sigma \text{ est paire} \\ -1 & \text{si } \sigma \text{ est impaire} \end{cases}$$

Exemple 8.3 : Considérons la permutation $\sigma = 35142$ de S_5.

3 et 5 précèdent 1 et sont supérieurs à 1 ; donc (3, 1) et (5, 1) satisfont (*).

3, 5 et 4 précèdent 2 et sont supérieurs à 2 ; donc (3, 2) , (5, 2) et (4, 2) satisfont (*).

5 précède 4 et est supérieur à lui ; donc (5, 4) satisfait (*).

Il existe donc exactement six couples satisfaisant (*), σ est paire et sgn $\sigma = 1$.

Exemple 8.4 : La permutation identique $\epsilon = 12\ldots n$ est paire puisqu'aucune paire ne satisfait (*).

Exemple 8.5 : Dans S_2, 12 est paire, 21 est impaire

Dans S_3, 123, 231 et 312 sont paires, et 132, 213 et 321 sont impaires.

Exemple 8.6 : Soit τ la permutation qui échange les deux nombres i et j en laissant les autres nombres fixes :

$$\tau(i) = j, \quad \tau(j) = i, \quad \tau(k) = k, \quad k \neq i, j$$

On appelle τ une transposition. Si $i < j$, on a alors

$$\tau = 12\ldots(i-1)\,j(i+1)\ldots(j-1)\,i(j+1)\ldots n$$

Il y a $2(j-i-1)+1$ couples satisfaisant (*) :

$$(j, i), \ (j, x), (x, i), \quad \text{où} \quad x = i+1, \ldots, j-1$$

Donc la transposition τ est impaire.

DETERMINANT

Soit $A = (a_{ij})$ une matrice carrée n sur un corps K:

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}$$

Considérons un produit de n éléments de A tel qu'un seul et unique élément appartienne à chaque ligne et un seul et unique élément appartienne à chaque colonne. Un tel produit peut être écrit sous la forme

$$a_{1j_1}\, a_{2j_2} \ldots a_{nj_n}$$

où les facteurs proviennent des lignes successives et donc les premiers indices sont dans l'ordre naturel $1, 2, \ldots, n$. Mais à partir du moment où les facteurs proviennent des colonnes différentes, la suite des deuxièmes indices forme une permutation $\sigma = j_1 j_2 \ldots j_n$ de S_n. Réciproquement chaque permutation de S_n détermine un produit de la forme précédente. Ainsi la matrice A contient $n!$ produits.

Définition : Le déterminant d'une matrice carrée $n \times n$, $A = (a_{ij})$, noté par $\det(A)$ ou $|A|$, est la somme suivante, qui est obtenue par la somme sur toutes les permutations possibles $\sigma = j_1 j_2 \ldots j_n$ de S_n :

$$|A| = \sum_{\sigma} (\operatorname{sgn} \sigma)\, a_{1j_1} a_{2j_2} \ldots a_{nj_n}$$

C'est-à-dire

$$|A| = \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma)\, a_{1\sigma(1)}\, a_{2\sigma(2)} \ldots a_{n\sigma(n)}$$

Le déterminant de la matrice carrée A $n \times n$ est dit d'ordre n, et est noté fréquemment par

$$\begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix}$$

Soulignons qu'un tableau carré de scalaires entouré de segments de droite n'est pas une matrice mais plutôt le scalaire que le déterminant associe à la matrice formée par ce même tableau de scalaires.

Exemple 8.7 : Le déterminant d'une matrice 1×1 $A = (a_{11})$ est le scalaire a_{11} lui-même : $|A| = a_{11}$. (Remarquons que la seule permutation dans S_1 est paire).

Exemple 8.8. : Dans S_2 la permutation 12 est paire et la permutation 21 est impaire. D'où

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

Ainsi $\begin{vmatrix} 4 & -5 \\ -1 & -2 \end{vmatrix} = 4(-2) - (-5)(-1) = -13$ et $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$

Exemple 8.9 : Dans S_3, les permutations 123, 231 et 312 sont paires et les permutations 321, 213 et 132 sont impaires. D'où

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{array}{l} a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ \quad - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32} \end{array}$$

ce qui peut également s'écrire

$$a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31})$$

ou

$$a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

qui est une combinaison linéaire des trois déterminants du second ordre dont les coefficients (avec des signes alternés) forment la première ligne de la matrice donnée. Remarquons que chaque matrice 2×2 peut être obtenue en supprimant, dans la matrice initiale, la ligne et la colonne contenant le coefficient :

$$a_{11}\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

Exemple 8.10 :

(i) $$\begin{vmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \\ 8 & 9 & 1 \end{vmatrix} = 2\begin{vmatrix} 6 & 7 \\ 9 & 1 \end{vmatrix} - 3\begin{vmatrix} 5 & 7 \\ 8 & 1 \end{vmatrix} + 4\begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix}$$
$$= 2(6-63) - 3(5-56) + 4(45-48) = 27$$

(ii) $$\begin{vmatrix} 2 & 3 & -4 \\ 0 & -4 & 2 \\ 1 & -1 & 5 \end{vmatrix} = 2\begin{vmatrix} -4 & 2 \\ -1 & 5 \end{vmatrix} - 3\begin{vmatrix} 0 & 2 \\ 1 & 5 \end{vmatrix} + (-4)\begin{vmatrix} 0 & -4 \\ 1 & -1 \end{vmatrix}$$
$$= 2(-20+2) - 3(0-2) - 4(0+4) = -46$$

Quand n augmente, le nombre de termes dans le déterminant devient très grand. En conséquence, on utilise des méthodes indirectes pour calculer les déterminants, plutôt que la définition. En fait nous allons démontrer un certain nombre de propriétés sur les déterminants qui vont nous permettre d'écourter considérablement le calcul. En particulier, nous démontrerons qu'un déterminant d'ordre n est égal à une combinaison linéaire de déterminants d'ordre $n - 1$, comme dans le cas où $n = 3$ ci-dessus.

PROPRIETES DES DETERMINANTS

Voici maintenant la liste des propriétés fondamentales du déterminant.

Théorème 8.1 : Les déterminants d'une matrice A et de sa transposée A^t sont égaux : $|A| = |A^t|$.

D'après ce théorème, tout autre théorème concernant le déterminant d'une matrice A qui s'applique aux lignes de A aura un théorème analogue s'appliquant aux colonnes de A.

Le théorème suivant donne certains cas dans lesquels le déterminant peut être obtenu immédiatement.

Théorème 8.2 : Soit A une matrice carrée.

1) Si A a une ligne (colonne) de zéros, alors $|A| = 0$.
2) Si A a deux lignes (colonnes) identiques, alors $|A| = 0$.
3) Si A est triangulaire, c'est-à-dire A est constituée de zéros au-dessus ou au-dessous de la diagonale, alors $|A|$ = produit des éléments diagonaux. Ainsi en particulier $|I| = 1$ où I est la matrice identique.

Le théorème suivant montre comment le déterminant d'une matrice est affecté par les opérations "élémentaires".

Théorème 8.3 : Soit B la matrice obtenue à partir de la matrice A :

1) par multiplication d'une ligne (colonne) de A par un scalaire k ; alors $|B| = k\,|A|$.
2) en échangeant deux lignes (colonnes) de A : alors $|B| = -\,|A|$.
3) en additionnant un multiple d'une ligne (colonne) de A à un autre ; alors $|B| = |A|$.

Enonçons maintenant deux des théorèmes les plus importants et les plus utiles concernant les déterminants.

Théorème 8.4 : Soit A une matrice carrée $n \times n$ quelconque. Les énoncés suivants sont équivalents.

1) A est inversible, c'est-à-dire A a une matrice A^{-1} inverse.
2) A est non singulière, c'est-à-dire $AX = 0$ admet seulement la solution nulle, ou le rang $A = n$, ou les lignes (colonnes) de A sont linéairement indépendantes.
3) Le déterminant de A n'est pas nul : $|A| \neq 0$.

Théorème 8.5 : Le déterminant est une fonction multiplicative, c'est-à-dire que le déterminant d'un produit de deux matrices A et B est égal au produit de leurs déterminant: $|AB| = |A|\,.\,|B|$.

Nous démontrerons les deux théorèmes précédents en utilisant la théorie des matrices élémentaires (voir page 56) et le lemme suivant.

Lemme 8.6 : Soit E une matrice élémentaire. Alors, quelle que soit la matrice A, $|EA| = |E|\,.\,|A|$.

Remarquons que l'on peut démontrer les deux théorèmes précédents directement, sans utiliser la théorie des matrices élémentaires.

MINEURS ET COFACTEURS

Considérons une matrice carrée $n \times n$ $A = (a_{ij})$. Soit M_{ij} la sous-matrice carrée $(n-1) \times (n-1)$ de A, obtenue en supprimant la $i^{\text{ème}}$ ligne et la $j^{\text{ème}}$ colonne de A. Le déterminant $|M_{ij}|$ est appelé le mineur de l'élément a_{ij} de A, et nous définissons le cofacteur de a_{ij} noté A_{ij} comme étant le mineur affecté de sa signature :

$$A_{ij} = (-1)^{i+j}\,|M_{ij}|$$

Notons que les "signatures" $(-1)^{i+j}$ des mineurs forment un arrangement en échiquier, avec les + sur la diagonale principale :

$$\begin{pmatrix} + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ + & - & + & - & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}$$

Il faut souligner que M_{ij} note une matrice alors que A_{ij} note un scalaire.

Exemple 8.11 : Soit $A = \begin{pmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \\ 8 & 9 & 1 \end{pmatrix}$. Alors $M_{23} = \begin{pmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \\ 8 & 9 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 8 & 9 \end{pmatrix}$ et

$$A_{23} = (-1)^{2+3}\begin{vmatrix} 2 & 3 \\ 8 & 9 \end{vmatrix} = -(18-24) = 6$$

Le théorème suivant s'applique alors.

Théorème 8.7 : Le déterminant de la matrice $A = (a_{ij})$ est égal à la somme des produits obtenus en multipliant les éléments d'une ligne (colonne) quelconque par leurs cofacteurs respectifs :

$$|A| = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in} = \sum_{j=1}^{n} a_{ij}A_{ij}$$

et

$$|A| = a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj} = \sum_{i=1}^{n} a_{ij}A_{ij}$$

Les formules précédentes, appelées développements de Laplace du déterminant de A suivant la $i^{\text{ème}}$ ligne ou la $j^{\text{ème}}$ colonne respectivement, offrent une méthode simplifiant le calcul de $|A|$, c'est-à-dire qu'en additionnant un multiple d'une ligne (ou colonne) à une autre ligne (colonne), nous pouvons réduire A à une matrice contenant une ligne ou une colonne avec un seul élément égal à 1 et les autres éléments étant nuls. En développant par rapport à cette ligne ou colonne, nous réduisons le calcul de $|A|$ au calcul d'un déterminant d'un ordre immédiatement inférieur.

Exemple 8.12 : Calculons le déterminant de $A = \begin{pmatrix} 5 & 4 & 2 & 1 \\ 2 & 3 & 1 & -2 \\ -5 & -7 & -3 & 9 \\ 1 & -2 & -1 & 4 \end{pmatrix}$.

Remarquons que 1 apparaît dans la seconde ligne et la troisième colonne. Effectuons les opérations suivantes sur A, où R_i note la $i^{\text{ème}}$ ligne :

(1) addition $-2R_2$ à R_1, (2) addition de $3R_2$ à R_3, (3) addition de $1R_2$ à R_4.

D'après le théorème 8.3 (3) la valeur du déterminant ne change pas lorsqu'on fait ces opérations ; c'est-à-dire

$$|A| = \begin{vmatrix} 5 & 4 & 2 & 1 \\ 2 & 3 & 1 & -2 \\ -5 & -7 & -3 & 9 \\ 1 & -2 & -1 & 4 \end{vmatrix} = \begin{vmatrix} 1 & -2 & 0 & 5 \\ 2 & 3 & 1 & -2 \\ 1 & 2 & 0 & 3 \\ 3 & 1 & 0 & 2 \end{vmatrix}$$

Si nous développons par rapport à la troisième colonne, nous devons négliger tous les termes contenant 0. Ainsi

$$|A| = (-1)^{2+3}\begin{vmatrix} 1 & -2 & 0 & 5 \\ 2 & 3 & 1 & -2 \\ 1 & 2 & 0 & 3 \\ 3 & 1 & 0 & 2 \end{vmatrix} = -\begin{vmatrix} 1 & -2 & 5 \\ 1 & 2 & 3 \\ 3 & 1 & 2 \end{vmatrix}$$

$$= -\left\{\begin{vmatrix} 2 & 3 \\ 1 & 2 \end{vmatrix} - (-2)\begin{vmatrix} 1 & 3 \\ 3 & 2 \end{vmatrix} + 5\begin{vmatrix} 1 & 2 \\ 3 & 1 \end{vmatrix}\right\} = 38$$

MATRICE ADJOINTE CLASSIQUE (OU COMMATRICE)

Considérons une matrice carrée $n \times n$ $A = (a_{ij})$ sur un corps K :

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

La transposée de la matrice des cofacteurs des éléments a_{ij} de A, notée par adj A, est appelée adjointe classique de A:

$$\operatorname{adj} A = \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n1} \\ \dots & \dots & \dots & \dots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix}$$

Nous disons "adjointe classique" au lieu de simplement "adjointe" parce que le terme adjointe sera utilisé dans le chapitre 13 pour un concept entièrement différent.

Exemple 8.13 : Soit $A = \begin{pmatrix} 2 & 3 & -4 \\ 0 & -4 & 2 \\ 1 & -1 & 5 \end{pmatrix}$. Les cofacteurs des neuf éléments de A sont

$$A_{11} = +\begin{vmatrix} -4 & 2 \\ -1 & 5 \end{vmatrix} = -18, \quad A_{12} = -\begin{vmatrix} 0 & 2 \\ 1 & 5 \end{vmatrix} = 2, \quad A_{13} = +\begin{vmatrix} 0 & -4 \\ 1 & -1 \end{vmatrix} = 4$$

$$A_{21} = -\begin{vmatrix} 3 & -4 \\ -1 & 5 \end{vmatrix} = -11, \quad A_{22} = +\begin{vmatrix} 2 & -4 \\ 1 & 5 \end{vmatrix} = 14, \quad A_{23} = -\begin{vmatrix} 2 & 3 \\ 1 & -1 \end{vmatrix} = 5$$

$$A_{31} = +\begin{vmatrix} 3 & -4 \\ -4 & 2 \end{vmatrix} = -10, \quad A_{32} = -\begin{vmatrix} 2 & -4 \\ 0 & 2 \end{vmatrix} = -4, \quad A_{33} = +\begin{vmatrix} 2 & 3 \\ 0 & -4 \end{vmatrix} = -8$$

Pour obtenir la matrice adjointe classique de A, nous prenons la transposée de la matrice précédente:

$$\operatorname{adj} A = \begin{pmatrix} -18 & -11 & -10 \\ 2 & 14 & -4 \\ 4 & 5 & -8 \end{pmatrix}$$

Théorème 8.8 : Pour une matrice carrée quelconque A,

$$A \cdot (\operatorname{adj} A) = (\operatorname{adj} A) \cdot A = |A| I$$

où I est la matrice identique. Ainsi, si $|A| \neq 0$,

$$A^{-1} = \frac{1}{|A|} (\operatorname{adj} A)$$

Remarquons que le théorème précédent nous donne une méthode importante pour obtenir l'inverse d'une matrice donnée.

Exemple 8.14 : Considérons la matrice A de l'exemple précédent pour laquelle $|A| = -46$. Nous avons

$$A (\operatorname{adj} A) = \begin{pmatrix} 2 & 3 & -4 \\ 0 & -4 & 2 \\ 1 & -1 & 5 \end{pmatrix} \begin{pmatrix} -18 & -11 & -10 \\ 2 & 14 & -4 \\ 4 & 5 & -8 \end{pmatrix} = \begin{pmatrix} -46 & 0 & 0 \\ 0 & -46 & 0 \\ 0 & 0 & -46 \end{pmatrix} = -46 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$
$$= -46I = |A| I$$

Nous avons aussi, d'après le Théorème 8.8,

$$A^{-1} = \frac{1}{|A|} (\operatorname{adj} A) = \begin{pmatrix} -18/-46 & -11/-46 & -10/-46 \\ 2/-46 & 14/-46 & -4/-46 \\ 4/-46 & 5/-46 & -8/-46 \end{pmatrix} = \begin{pmatrix} 9/23 & 11/46 & 5/23 \\ -1/23 & -7/23 & 2/23 \\ -2/23 & -5/46 & 4/23 \end{pmatrix}$$

APPLICATIONS AUX EQUATIONS LINEAIRES

Considérons un système de n équations linéaires à n inconnues:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n \end{aligned}$$

Appelons Δ le déterminant de la matrice $A = (a_{ij})$ des coefficients : $\Delta = |A|$. Et appelons Δ_i le déterminant de la matrice obtenue en remplaçant la $i$ème colonne de A par la colonne des termes constants. Il s'ensuit la relation fondamentale entre les déterminants et la solution du système précédent.

Théorème 8-9 : Le système précédent a une solution unique si et seulement si $\Delta \neq 0$. Dans ce cas la solution unique est donnée par

$$x_1 = \frac{\Delta_1}{\Delta}, \quad x_2 = \frac{\Delta_2}{\Delta}, \quad \ldots, \quad x_n = \frac{\Delta_n}{\Delta}$$

Le théorème précédent est connu sous le nom de "Règle de Cramer" pour la résolution des systèmes d'équations linéaires. Soulignons que ce théorème s'applique uniquement à un système qui a le même nombre d'équations que d'inconnues, et qu'il donne une solution unique dans le cas où seulement $\Delta \neq 0$. En fait, si $\Delta = 0$, le théorème n'indique pas si le système admet ou non une solution. Cependant dans le cas d'un système homogène, nous avons le résultat suivant qui s'avère d'un grand intérêt.

Théorème 8.10 : Le système homogène $Ax = 0$ admet une solution non nulle si et seulement si $\Delta = |A| = 0$.

Exemple 8.15 : Résoudre en utilisant les déterminants : $\begin{cases} 2x - 3y = 7 \\ 3x + 5y = 1 \end{cases}$.

Calculons d'abord le déterminant Δ de la matrice des coefficients :

$$\Delta = \begin{vmatrix} 2 & -3 \\ 3 & 5 \end{vmatrix} = 10 + 9 = 19$$

Puisque $\Delta \neq 0$ le système a une solution unique. Nous avons aussi

$$\Delta_x = \begin{vmatrix} 7 & -3 \\ 1 & 5 \end{vmatrix} = 38, \qquad \Delta_y = \begin{vmatrix} 2 & 7 \\ 3 & 1 \end{vmatrix} = -19$$

En conséquence, la solution unique du système est

$$x = \frac{\Delta_x}{\Delta} = \frac{38}{19} = 2, \qquad y = \frac{\Delta_y}{\Delta} = \frac{-19}{19} = -1$$

Nous pouvons remarquer que le théorème précédent est intéressant plus pour des raisons théoriques et historiques que pour des raisons pratiques. La méthode déjà indiquée, de résolution des systèmes d'équations linéaires par réduction du système à sa forme échelonnée, est plus efficace que celle qui utilise les déterminants.

DETERMINANT D'UN OPERATEUR LINEAIRE

En utilisant la propriété multiplicative d'un déterminant (Théorème 8.5), nous obtenons

Théorème 8.11 : Supposons que A et B soient des matrices semblables. Alors $|A| = |B|$.

Supposons maintenant que T soit un opérateur linéaire quelconque sur un espace vectoriel V. Nous définissons le déterminant de T, écrit det (T), par

$$\det(T) = |[T]_e|$$

où $[T]_e$ est la matrice de T dans une base $\{e_i\}$. D'après le théorème précédent cette définition est indépendante de la base particulière qui est choisie.

Le théorème suivant découle des théorèmes analogues sur les matrices.

Théorème 8.12 : Soit T et S des opérateurs linéaires sur un espace vectoriel V. Alors

(i) $\det(S \circ T) = \det(S) \cdot \det(T)$,

(ii) T est inversible si et seulement si $\det(T) \neq 0$.

Remarquons aussi que $\det(1_V) = 1$ où 1_V est l'application identique, et que $\det(T^{-1}) = \det(T)^{-1}$ si T est inversible.

Exemple 8.16 : Soit T l'opérateur linéaire de $\mathbf{R}^3$ défini par

$$T(x, y, z) = (2x - 4y + z,\ x - 2y + 3z,\ 5x + y - z)$$

La matrice de T dans la base usuelle de $\mathbf{R}^3$ est $[T] = \begin{pmatrix} 2 & -4 & 1 \\ 1 & -2 & 3 \\ 5 & 1 & -1 \end{pmatrix}$. Alors

$$\det(T) = \begin{vmatrix} 2 & -4 & 1 \\ 1 & -2 & 3 \\ 5 & 1 & -1 \end{vmatrix} = 2(2-3) + 4(-1-15) + 1(1+10) = -55$$

MULTILINEARITE ET DETERMINANTS

Soit $\mathcal{A}$ l'ensemble de toutes les matrices carrées $n \times n$ A sur un corps K. Nous pouvons considérer A comme un n-tuple constitué de ses vecteurs lignes $A_1, A_2, \ldots, A_n$:

$$A = (A_1, A_2, \ldots, A_n)$$

Donc $\mathcal{A}$ peut être considérée comme l'ensemble des n-tuples de K:

$$\mathcal{A} = (K^n)^n$$

Il s'ensuit les définitions suivantes.

Définition : Une fonction $D : \mathcal{A} \to K$ est dite multilinéaire si elle est linéaire suivant chacune de ses composantes, c'est-à-dire

(i) si ligne $A_i = B + C$, alors

$$D(A) = D(\ldots, B + C, \ldots) = D(\ldots, B, \ldots) + D(\ldots, C, \ldots) ;$$

(ii) si ligne $A_i = kB$, où $k \in K$, alors

$$D(A) = D(\ldots, kB, \ldots) = kD(\ldots, B, \ldots).$$

Nous pouvons dire aussi n-linéaire au lieu de multilinéaire si elle est linéaire pour chacune de ses n composantes.

Définition : Une fonction $D : \mathcal{A} \to K$ est dite alternée si $D(A) = 0$ toutes les fois que A a deux lignes identiques :

$$D(A_1, A_2, \ldots, A_n) = 0 \quad \text{toutes les fois que } A_i = A_j,\ i \neq j$$

Nous obtenons le résultat fondamental suivant : ici I note la matrice identique.

Théorème 8.13 : Il existe une fonction unique $D : \mathcal{A} \to K$ telle que

(i) D est multilinéaire, (ii) D est alternée, (iii) $D(I) = 1$.

Cette fonction D n'est pas autre chose que la fonction déterminant ; c'est-à-dire la fonction qui à une matrice quelconque $A \in \mathcal{A}$ associe $D(A) = |A|$.

PROBLEMES RESOLUS

CALCUL DE DETERMINANTS

8.1. Calculer le déterminant de chacune des matrices (i) $\begin{pmatrix} 3 & -2 \\ 4 & 5 \end{pmatrix}$, (ii) $\begin{pmatrix} a-b & a \\ a & a+b \end{pmatrix}$.

(i) $\begin{vmatrix} 3 & -2 \\ 4 & 5 \end{vmatrix} = 3 \cdot 5 - (-2) \cdot 4 = 23.$ (ii) $\begin{vmatrix} a-b & a \\ a & a+b \end{vmatrix} = (a-b)(a+b) - a \cdot a = -b^2.$

8.2. Déterminer les valeurs de k pour lesquelles $\begin{vmatrix} k & k \\ 4 & 2k \end{vmatrix} = 0$.

$\begin{vmatrix} k & k \\ 4 & 2k \end{vmatrix} = 2k^2 - 4k = 0$, ou $2k(k-2) = 0$. Donc $k = 0$; et $k = 2$. c'est-à-dire si $k = 0$ ou $k = 2$, le déterminant est nul.

8.3. Calculer le déterminant de chacune des matrices

(i) $\begin{pmatrix} 1 & 2 & 3 \\ 4 & -2 & 3 \\ 2 & 5 & -1 \end{pmatrix}$, (ii) $\begin{pmatrix} 2 & 0 & 1 \\ 4 & 2 & -3 \\ 5 & 3 & 1 \end{pmatrix}$, (iii) $\begin{pmatrix} 2 & 0 & 1 \\ 3 & 2 & -3 \\ -1 & -3 & 5 \end{pmatrix}$, (iv) $\begin{pmatrix} 1 & 0 & 0 \\ 3 & 2 & -4 \\ 4 & 1 & 3 \end{pmatrix}$.

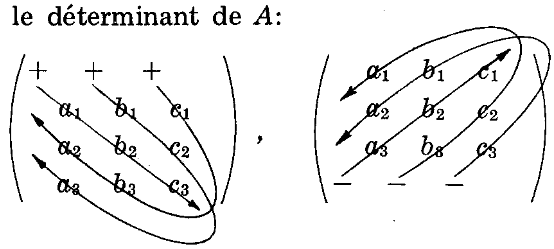
(i) $$\begin{vmatrix} 1 & 2 & 3 \\ 4 & -2 & 3 \\ 2 & 5 & -1 \end{vmatrix} = 1\begin{vmatrix} -2 & 3 \\ 5 & -1 \end{vmatrix} - 2\begin{vmatrix} 4 & 3 \\ 2 & -1 \end{vmatrix} + 3\begin{vmatrix} 4 & -2 \\ 2 & 5 \end{vmatrix}$$
$$= 1(2-15) - 2(-4-6) + 3(20+4) = 79$$

(ii) $$\begin{vmatrix} 2 & 0 & 1 \\ 4 & 2 & -3 \\ 5 & 3 & 1 \end{vmatrix} = 2\begin{vmatrix} 2 & -3 \\ 3 & 1 \end{vmatrix} - 0\begin{vmatrix} 4 & -3 \\ 5 & 1 \end{vmatrix} + 1\begin{vmatrix} 4 & 2 \\ 5 & 3 \end{vmatrix} = 24$$

(iii) $$\begin{vmatrix} 2 & 0 & 1 \\ 3 & 2 & -3 \\ -1 & -3 & 5 \end{vmatrix} = 2(10-9) + 1(-9+2) = -5$$

(iv) $$\begin{vmatrix} 1 & 0 & 0 \\ 3 & 2 & -4 \\ 4 & 1 & 3 \end{vmatrix} = 1(6+4) = 10$$

8.4. Considérons la matrice carrée 3×3 $A = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$. Montrer que le diagramme ci-dessous peut être utilisé pour obtenir le déterminant de A:



Formons le produit de chacun des trois nombres liés par une flèche dans le diagramme de gauche et faisons précéder chaque produit par un signe plus, comme il suit :

$$+ \; a_1b_2c_3 \; + \; b_1c_2a_3 \; + \; c_1a_2b_3$$

Formons maintenant le produit de chacun des trois nombres liés par une flèche dans le diagramme de droite, et faisons précéder chaque produit par un signe moins comme il suit:

$$-\ a_3b_2c_1\ -\ b_3c_2a_1\ -\ c_3a_2b_1$$

Alors le déterminant de A est précisément la somme des deux expressions précédentes :

$$|A| = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1b_2c_3 + b_1c_2a_3 + c_1a_2b_3 - a_3b_2c_1 - b_3c_2a_1 - c_3a_2b_1$$

La méthode précédente de calcul de $|A|$ ne s'applique pas à des déterminants d'un ordre supérieure à 3.

8.5. Calculer le déterminant de chacune des matrices :

$$\text{(i)} \begin{pmatrix} 2 & 0 & -1 \\ 3 & 0 & 2 \\ 4 & -3 & 7 \end{pmatrix}, \quad \text{(ii)} \begin{pmatrix} a & b & c \\ c & a & b \\ b & c & a \end{pmatrix}, \quad \text{(iii)} \begin{pmatrix} 3 & 2 & -4 \\ 1 & 0 & -2 \\ -2 & 3 & 3 \end{pmatrix}$$

(i) Développons le déterminant suivant la deuxième colonne en négligeant les termes contenant un zéro.

$$\begin{vmatrix} 2 & 0 & -1 \\ 3 & 0 & 2 \\ 4 & -3 & 7 \end{vmatrix} = -(-3)\begin{vmatrix} 2 & -1 \\ 3 & 2 \end{vmatrix} = 3(4+3) = 21$$

(ii) En utilisant la méthode du précédent problème :

$$\begin{vmatrix} a & b & c \\ c & a & b \\ b & c & a \end{vmatrix} = a^3 + b^3 + c^3 - abc - abc - abc = a^3 + b^3 + c^3 - 3abc$$

(iii) En additionnant le double de la première colonne à la troisième colonne, et en développant par rapport à la seconde ligne :

$$\begin{vmatrix} 3 & 2 & -4 \\ 1 & 0 & -2 \\ -2 & 3 & 3 \end{vmatrix} = \begin{vmatrix} 3 & 2 & -4+2(3) \\ 1 & 0 & -2+2(1) \\ -2 & 3 & 3+2(-2) \end{vmatrix} = \begin{vmatrix} 3 & 2 & 2 \\ 1 & 0 & 0 \\ -2 & 3 & -1 \end{vmatrix} = -1\begin{vmatrix} 2 & 2 \\ 3 & -1 \end{vmatrix} = 8$$

8.6. Calculer le déterminant de $A = \begin{pmatrix} \frac{1}{2} & -1 & -\frac{1}{3} \\ \frac{3}{4} & \frac{1}{2} & -1 \\ 1 & -4 & 1 \end{pmatrix}$.

Multiplions d'abord la première ligne par 6 et la seconde ligne par 4. Alors

$$6 \cdot 4\,|A| = 24\,|A| = \begin{vmatrix} 3 & -6 & -2 \\ 3 & 2 & -4 \\ 1 & -4 & 1 \end{vmatrix} = \begin{vmatrix} 3 & -6+4(3) & -2-(3) \\ 1 & 2+4(3) & -4-(3) \\ 1 & -4+4(1) & 1-(1) \end{vmatrix} = \begin{vmatrix} 3 & 6 & -5 \\ 1 & 14 & -7 \\ 1 & 0 & 0 \end{vmatrix}$$

$$= +\begin{vmatrix} 6 & -5 \\ 14 & -7 \end{vmatrix} = 28, \qquad \text{et} \quad |A| = 28/24 = 7/6.$$

8.7. Calculer le déterminant de $A = \begin{pmatrix} 2 & 5 & -3 & -2 \\ -2 & -3 & 2 & -5 \\ 1 & 3 & -2 & 2 \\ -1 & -6 & 4 & 3 \end{pmatrix}$.

Remarquons que 1 est dans la troisième ligne, première colonne. Appliquons les opérations suivantes à A (où R_i est la $i$ème ligne) : (i) ajoutons $-2R_3$ à R_1, (ii) ajoutons $2R_3$ à R_2, (iii) ajoutons $1R_3$ à R_4. Ainsi on a

$$|A| = \begin{vmatrix} 2 & 5 & -3 & -2 \\ -2 & -3 & 2 & -5 \\ 1 & 3 & -2 & 2 \\ -1 & -6 & 4 & 3 \end{vmatrix} = \begin{vmatrix} 0 & -1 & 1 & -6 \\ 0 & 3 & -2 & -1 \\ 1 & 3 & -2 & 2 \\ 0 & -3 & 2 & 5 \end{vmatrix} = +\begin{vmatrix} -1 & 1 & -6 \\ 3 & -2 & -1 \\ -3 & 2 & 5 \end{vmatrix}$$

$$= \begin{vmatrix} -1+1 & 1 & -6+6(1) \\ 3-2 & -2 & -1+6(-2) \\ -3+2 & 2 & 5+6(2) \end{vmatrix} = \begin{vmatrix} 0 & 1 & 0 \\ 1 & -2 & -13 \\ -1 & 2 & 17 \end{vmatrix} = -\begin{vmatrix} 1 & -13 \\ -1 & 17 \end{vmatrix} = -4$$

8.8. Calculer le déterminant de $A = \begin{pmatrix} 3 & -2 & -5 & 4 \\ -5 & 2 & 8 & -5 \\ -2 & 4 & 7 & -3 \\ 2 & -3 & -5 & 8 \end{pmatrix}$.

Tout d'abord réduisons A à une matrice admettant 1 comme élément ; pour cela ajoutons le double de la première ligne à la seconde ligne et procédons alors comme dans le problème précédent.

$$|A| = \begin{vmatrix} 3 & -2 & -5 & 4 \\ -5 & 2 & 8 & -5 \\ -2 & 4 & 7 & -3 \\ 2 & -3 & -5 & 8 \end{vmatrix} = \begin{vmatrix} 3 & -2 & -5 & 4 \\ -5+2(3) & 2+2(-2) & 8+2(-5) & -5+2(4) \\ -2 & 4 & 7 & -3 \\ 2 & -3 & -5 & 8 \end{vmatrix}$$

$$= \begin{vmatrix} 3 & -2 & -5 & 4 \\ 1 & -2 & -2 & 3 \\ -2 & 4 & 7 & -3 \\ 2 & -3 & -5 & 8 \end{vmatrix} = \begin{vmatrix} 3 & -2+2(3) & -5+2(3) & 4-3(3) \\ 1 & -2+2(1) & -2+2(1) & 3-3(1) \\ -2 & 4+2(-2) & 7+2(-2) & -3-3(-2) \\ 2 & -3+2(2) & -5+2(2) & 8-3(2) \end{vmatrix}$$

$$= \begin{vmatrix} 3 & 4 & 1 & -5 \\ 1 & 0 & 0 & 0 \\ -2 & 0 & 3 & 3 \\ 2 & 1 & -1 & 2 \end{vmatrix} = -\begin{vmatrix} 4 & 1 & -5 \\ 0 & 3 & 3 \\ 1 & -1 & 2 \end{vmatrix} = -\begin{vmatrix} 4 & 1 & -5-(1) \\ 0 & 3 & 3-(3) \\ 1 & -1 & 2-(-1) \end{vmatrix}$$

$$= -\begin{vmatrix} 4 & 1 & -6 \\ 0 & 3 & 0 \\ 1 & -1 & 3 \end{vmatrix} = -3\begin{vmatrix} 4 & -6 \\ 1 & 3 \end{vmatrix} = -3(12+6) = -54$$

8.9. Calculer le déterminant de $A = \begin{pmatrix} t+3 & -1 & 1 \\ 5 & t-3 & 1 \\ 6 & -6 & t+4 \end{pmatrix}$.

Ajoutons la seconde colonne à la première ligne, et ensuite ajoutons la troisième colonne à la seconde colonne ; on obtient

$$|A| = \begin{vmatrix} t+2 & 0 & 1 \\ t+2 & t-2 & 1 \\ 0 & t-2 & t+4 \end{vmatrix}$$

Mettons en facteur $t+2$ qui se trouve dans la première colonne et $t-2$ qui se trouve dans la seconde colonne ; on obtient

$$|A| = (t+2)(t-2)\begin{vmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & t+4 \end{vmatrix}$$

Finalement soustrayons la première colonne de la troisième colonne ; on obtient

$$|A| = (t+2)(t-2)\begin{vmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & t+4 \end{vmatrix} = (t+2)(t-2)(t+4)$$

COFACTEURS

8.10. Trouver le cofacteur de 7 dans la matrice $\begin{pmatrix} 2 & 1 & -3 & 4 \\ 5 & -4 & 7 & -2 \\ 4 & 0 & 6 & -3 \\ 3 & -2 & 5 & 2 \end{pmatrix}$.

$$(-1)^{2+3}\begin{vmatrix} 2 & 1 & -3 & 4 \\ 5 & -4 & 7 & -2 \\ 4 & 0 & 6 & -3 \\ 3 & -2 & 5 & 2 \end{vmatrix} = -\begin{vmatrix} 2 & 1 & 4 \\ 4 & 0 & -3 \\ 3 & -2 & 2 \end{vmatrix} = -\begin{vmatrix} 2 & 1 & 4 \\ 4 & 0 & -3 \\ 7 & 0 & 10 \end{vmatrix} = -\left(-\begin{vmatrix} 4 & -3 \\ 7 & 10 \end{vmatrix}\right) = 61$$

L'exposant $2+3$ provient du fait que 7 apparaît dans la seconde ligne, troisième colonne.

8.11. Considérons la matrice $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 1 & 5 & 7 \end{pmatrix}$. (i) Calculer $|A|$. (ii) Trouver adj A. (iii) Vérifier que $A(\text{adj } A) = |A| \cdot I$. (iv) Trouver A^{-1}.

(i) $$|A| = 1\begin{vmatrix} 3 & 4 \\ 5 & 7 \end{vmatrix} - 2\begin{vmatrix} 2 & 4 \\ 1 & 7 \end{vmatrix} + 3\begin{vmatrix} 2 & 3 \\ 1 & 5 \end{vmatrix} = 1 - 20 + 21 = 2$$

(ii) $$\text{adj } A = \begin{pmatrix} +\begin{vmatrix} 3 & 4 \\ 5 & 7 \end{vmatrix} & -\begin{vmatrix} 2 & 4 \\ 1 & 7 \end{vmatrix} & +\begin{vmatrix} 2 & 3 \\ 1 & 5 \end{vmatrix} \\ -\begin{vmatrix} 2 & 3 \\ 5 & 7 \end{vmatrix} & +\begin{vmatrix} 1 & 3 \\ 1 & 7 \end{vmatrix} & -\begin{vmatrix} 1 & 2 \\ 1 & 5 \end{vmatrix} \\ +\begin{vmatrix} 2 & 3 \\ 3 & 4 \end{vmatrix} & -\begin{vmatrix} 1 & 3 \\ 2 & 4 \end{vmatrix} & +\begin{vmatrix} 1 & 2 \\ 2 & 3 \end{vmatrix} \end{pmatrix}^t = \begin{pmatrix} 1 & -10 & 7 \\ 1 & 4 & -3 \\ -1 & 2 & -1 \end{pmatrix}^t = \begin{pmatrix} 1 & 1 & -1 \\ -10 & 4 & 2 \\ 7 & -3 & -1 \end{pmatrix}$$

C'est-à-dire que adj A est la transposée de la matrice des cofacteurs. Remarquons que les signes dans la matrice des cofacteurs ont une forme en échiquier $\begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix}$.

(iii) $$A \cdot (\text{adj } A) = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 1 & 5 & 7 \end{pmatrix}\begin{pmatrix} 1 & 1 & -1 \\ -10 & 4 & 2 \\ 7 & -3 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} = 2\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = |A|\, I$$

(iv) $$A^{-1} = \frac{1}{|A|}(\text{adj } A) = \frac{1}{2}\begin{pmatrix} 1 & 1 & -1 \\ -10 & 4 & 2 \\ 7 & -3 & -1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ -5 & 2 & 1 \\ \frac{7}{2} & -\frac{3}{2} & -\frac{1}{2} \end{pmatrix}$$

8.12. Considérons une matrice quelconque 2×2 $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

(i) Trouver adj A. (ii) Montrer que adj (adj A) $= A$.

(i) $$\text{adj } A = \begin{pmatrix} +|d| & -|c| \\ -|b| & +|a| \end{pmatrix}^t = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}^t = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

(ii) $$\text{adj}(\text{adj } A) = \text{adj}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} +|a| & -|-c| \\ -|-b| & +|d| \end{pmatrix}^t = \begin{pmatrix} a & c \\ b & d \end{pmatrix}^t = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = A$$

DETERMINANTS ET SYSTEMES D'EQUATIONS LINEAIRES

8.13. Résoudre les systèmes suivants en x et y en utilisant les déterminants :

$$\text{(i)} \quad \begin{array}{rl} 2x + y &= 7 \\ 3x - 5y &= 4 \end{array} \qquad \text{(ii)} \quad \begin{array}{rl} ax - 2by &= c \\ 3ax - 5by &= 2c \end{array}, \quad \text{où} \quad ab \neq 0.$$

(i) $\Delta = \begin{vmatrix} 2 & 1 \\ 3 & -5 \end{vmatrix} = -13$, $\Delta_x = \begin{vmatrix} 7 & 1 \\ 4 & -5 \end{vmatrix} = -39$, $\Delta_y = \begin{vmatrix} 2 & 7 \\ 3 & -4 \end{vmatrix} = -13$. Alors $x = \Delta_x/\Delta = 3$, $y = \Delta_y/\Delta = 1$.

(ii) $\Delta = \begin{vmatrix} a & -2b \\ 3a & -5b \end{vmatrix} = ab$, $\Delta_x = \begin{vmatrix} c & -2b \\ 2c & -5b \end{vmatrix} = -bc$, $\Delta_y = \begin{vmatrix} a & c \\ 3a & 2c \end{vmatrix} = -ac$. Alors $x = \Delta_x/\Delta = -c/a$, $y = \Delta_y/\Delta = -c/b$.

8.14. Résoudre en utilisant les déterminants : $\begin{cases} 3y + 2x = z + 1 \\ 3x + 2z = 8 - 5y \\ 3z - 1 = x - 2y \end{cases}$.

Tout d'abord mettons le système sous une forme standard de manière que les inconnues apparaissent en colonnes :

$$\begin{aligned} 2x + 3y - z &= 1 \\ 3x + 5y + 2z &= 8 \\ x - 2y - 3z &= -1 \end{aligned}$$

Calculons le déterminant Δ de la matrice A des coefficients :

$$\Delta = \begin{vmatrix} 2 & 3 & -1 \\ 3 & 5 & 2 \\ 1 & -2 & -3 \end{vmatrix} = 2(-15+4) - 3(-9-2) - 1(-6-5) = 22$$

Puisque $\Delta \neq 0$ le système admet une solution unique. Pour obtenir Δ_x, Δy et Δ_z, remplaçons les coefficients de l'inconnue considérée dans la matrice A par la colonne des constantes. Ainsi

$$\Delta_x = \begin{vmatrix} 1 & 3 & -1 \\ 8 & 5 & 2 \\ -1 & -2 & -3 \end{vmatrix} = 66, \quad \Delta_y = \begin{vmatrix} 2 & 1 & -1 \\ 3 & 8 & 2 \\ 1 & -1 & -3 \end{vmatrix} = -22, \quad \Delta_z = \begin{vmatrix} 2 & 3 & 1 \\ 3 & 5 & 8 \\ 1 & -2 & -1 \end{vmatrix} = 44$$

et $x = \Delta_x/\Delta = 3$, $y = \Delta_y/\Delta = -1$, $z = \Delta_z/\Delta = 2$.

DEMONSTRATIONS DES THEOREMES

8.15. Démontrons le théorème 8.1 : $|A^t| = |A|$.

Supposons $A = (a_{ij})$. Alors $A^t = (b_{ij})$ où $b_{ij} = a_{ji}$. Donc

$$\begin{aligned} |A^t| &= \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma)\, b_{1\sigma(1)}\, b_{2\sigma(2)} \ldots b_{n\sigma(n)} \\ &= \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma)\, a_{\sigma(1),1}\, a_{\sigma(2),2} \ldots a_{\sigma(n),n} \end{aligned}$$

Soit $\tau = \sigma^{-1}$. D'après le problème 8.36 sgn τ = sgn σ et

$$a_{\sigma(1),1}\, a_{\sigma(2),2} \ldots a_{\sigma(n),n} = a_{1\tau(1)}\, a_{2\tau(2)} \ldots a_{n\tau(n)}$$

Donc $$|A^t| = \sum_{\sigma \in S_n} (\operatorname{sgn} \tau)\, a_{1\tau(1)}\, a_{2\tau(2)} \ldots a_{n\tau(n)}$$

Comme σ parcourt tous les éléments de S_n, $\tau = \sigma^{-1}$ parcourt aussi tous les éléments de S_n. Ainsi $|A^t| = A$.

8.16. Démontrer le théorème 8.3 (ii) : Soit B la matrice obtenue à partir de la matrice carrée A en échangeant deux lignes (colonnes) de A. Alors $|B| = -|A|$.

Nous démontrons ce théorème dans le cas où deux colonnes sont échangées. Soit τ la transposition qui échange les deux nombres correspondant aux deux colonnes de A qui sont échangées. Si $A = (a_{ij})$ et $B = (b_{ij})$ alors $b_{ij} = a_{i\tau(j)}$. Donc, pour une permutation quelconque σ,

$$b_{1\sigma(1)}\, b_{2\sigma(2)} \ldots b_{n\sigma(n)} = a_{1\tau\sigma(1)}\, a_{2\tau\sigma(2)} \ldots a_{n\tau\sigma(n)}$$

Ainsi
$$|B| = \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma)\, b_{1\sigma(1)}\, b_{2\sigma(2)} \ldots b_{n\sigma(n)} = \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma)\, a_{1\tau\sigma(1)}\, a_{2\tau\sigma(2)} \ldots a_{n\tau\sigma(n)}$$

Puisque la transposition τ est une permutation impaire, $\operatorname{sgn} \tau\sigma = \operatorname{sgn} \tau \cdot \operatorname{sgn} \sigma = -\operatorname{sgn} \sigma$. Donc $\operatorname{sgn} \sigma = -\operatorname{sgn} \tau\sigma$ et donc

$$|B| = -\sum_{\sigma \in S_n} (\operatorname{sgn} \tau\sigma)\, a_{1\tau\sigma(1)}\, a_{2\tau\sigma(2)} \ldots a_{n\tau\sigma(n)}$$

Mais comme σ parcourt tous les éléments de S_n, $\tau\sigma$ parcourt aussi tous les éléments de S_n ; donc $|B| = -|A|$.

8.17. Démontrer le théorème 8.2 : (i) Si A a une ligne (colonne) de zéros, alors $|A| = 0$. (ii) Si A a deux lignes identiques (colonnes) alors $|A| = 0$. (iii) Si A est triangulaire alors $|A|$ = produit des éléments diagonaux. Ainsi en particulier $|I| = 1$ où I est la matrice identité.

(i) Chaque terme dans $|A|$ contient un facteur provenant de chaque ligne et donc de la ligne de zéros. Donc chaque terme de $|A|$ est nul et donc $|A| = 0$.

(ii) Supposons $1 + 1 \neq 0$ dans K. Si nous échangeons les deux lignes identiques de A, nous obtenons toujours la matrice A. Donc d'après le problème précédent $|A| = -|A|$ et donc $|A| = 0$.

Supposons maintenant $1 + 1 = 0$ dans K. Alors $\operatorname{sgn} \sigma = 1$ pour tout $\sigma \in S_n$. Puisque A a deux lignes identiques, nous pouvons arranger les termes de A en paires de termes égaux. Puisque chaque paire est nulle le déterminant de A est zéro.

(iii) Supposons que $A = (a_{ij})$ soit une matrice triangulaire inférieure, c'est-à-dire que les éléments au-dessus de la diagonale sont tous nuls : $a_{ij} = 0$ toutes les fois que $i < j$. Considérons un terme t du déterminant de A :

$$t = (\operatorname{sgn} \sigma)\, a_{1i_1}\, a_{2i_2} \ldots a_{ni_n}, \qquad \text{où} \quad \sigma = i_1 i_2 \ldots i_n$$

Supposons $i_1 \neq 1$. Alors $1 < i_1$ et donc $a_{1i_1} = 0$; d'où $t = 0$. En somme, chaque terme pour lequel $i_1 \neq 1$ est nul.

Supposons maintenant $i_1 = 1$, mais $i_2 \neq 2$. Alors $2 < i_2$ et donc $a_{2i_2} = 0$, d'où $t = 0$. Ainsi chaque terme pour lequel $i_1 \neq 1$ ou $i_2 \neq 2$ est nul.

De façon analogue, nous obtenons que chaque terme, pour lequel $i_1 \neq 1$ ou $i_2 \neq 2$ ou ... ou $i_n \neq n$, est nul. En conséquence $|A| = a_{11}\, a_{22} \ldots a_{nn}$ = produit des éléments diagonaux.

8.18. Démontrer le théorème 8.3 : Soit B obtenue à partir de A en

(i) multipliant une ligne (colonne) de A par un scalaire k ; alors $|B| = k\,|A|$.

(ii) échangeant deux lignes (colonnes) de A ; alors $|B| = -|A|$.

(iii) additionnant un multiple d'une ligne (colonne) de A à une autre ; alors $|B| = |A|$.

(i) Si la $j$ème ligne de A est multipliée par k, alors chaque terme de $|A|$ est multiplié par k et donc $|B| = k\,|A|$. C'est-à-dire

$$|B| = \sum_{\sigma} (\operatorname{sgn} \sigma)\, a_{1i_1}\, a_{2i_2} \ldots (k a_{ji_j}) \ldots a_{ni_n} = k \sum_{\sigma} (\operatorname{sgn} \sigma)\, a_{1i_1}\, a_{2i_2} \ldots a_{ni_n} = k\,|A|$$

(ii) Démontrée dans le problème 8.16.

(iii) Supposons que l'on additionne la $k$ième ligne multipliée par c, à la $j$ème ligne de A. En utilisant le symbole $\wedge$ pour noter la $j$ème position dans le déterminant, nous avons

$$|B| = \sum_{\sigma} (\operatorname{sgn} \sigma)\, a_{1i_1}\, a_{2i_2} \ldots \widehat{(c a_{ki_k} + a_{ji_j})} \ldots a_{ni_n}$$
$$= c \sum_{\sigma} (\operatorname{sgn} \sigma)\, a_{1i_1}\, a_{2i_2} \ldots \widehat{a_{ki_k}} \ldots a_{ni_n} + \sum_{\sigma} (\operatorname{sgn} \sigma)\, a_{1i_1}\, a_{2i_2} \ldots \widehat{a_{ji_j}} \ldots a_{ni_n}$$

La première somme est le déterminant d'une matrice dont les $k$ième et $j$èmes ligne sont identiques ; donc d'après le théorème 8.2 (ii) la somme est nulle. La seconde somme est le déterminant de A. Donc $|B| = c \cdot 0 + |A| = |A|$.

8.19. Démontrer le lemme 8.6 : Pour une quelconque matrice élémentaire E, $|EA| = |E| . |A|$.

Considérons les opérations élémentaires suivantes : 1) multiplication d'une ligne par une constante $k \neq 0$; 2) échange de deux lignes; 3) addition d'un multiple d'une ligne par une autre. Soient E_1, E_2 et E_3 les matrices élémentaires correspondantes. C'est-à-dire que E_1, E_2 et E_3 sont obtenues en appliquant les opérations précédentes, respectivement à la matrice identité I. D'après le problème précédent,

$$|E_1| = k\,|I| = k, \qquad |E_2| = -|I| = -1, \qquad |E_3| = |I| = 1$$

Rappelons (page 56) que $E_i A$ est identique à la matrice obtenue en appliquant l'opération correspondante à A. Donc d'après le problème précédent,

$$|E_1 A| = k\,|A| = |E_1|\,|A|, \qquad |E_2 A| = -|A| = |E_2|\,|A|, \qquad |E_3 A| = |A| = 1\,|A| = |E_3|\,|A|$$

et le lemme est démontré.

8.20. Supposons que B soit équivalente ligne à A ; d'où $B = E_n . E_{n-1} \ldots E_2\, E_1\, A$, où les E_i sont les matrices élémentaires. Montrer que

1) $|B| = |E_n|\,|E_{n-1}| \ldots |E_2|\,|E_1|\,|A|$, 2) $|B| \neq 0$ si et seulement si $|A| \neq 0$.

1) D'après le problème précédent $|E_1 A| = |E_1| . |A|$. Ainsi par récurrence.

$$|B| = |E_n|\,|E_{n-1} \ldots E_2 E_1 A| = |E_n|\,|E_{n-1}| \ldots |E_2|\,|E_1|\,|A|$$

2) D'après le problème précédent $E_i \neq 0$ pour chaque i. Donc $|B| \neq 0$ si et seulement si $|A| \neq 0$.

8.21. Démontrer le théorème 8.4 : Soit A une matrice carrée d'ordre n. Les affirmations suivantes sont équivalentes : 1) A est inversible; 2) A est non singulière ; 3) $|A| \neq 0$.

D'après le problème 6.44 1) et 2) sont équivalentes. Donc il suffit de montrer que 1) et 3) sont équivalentes.

Supposons que A soit inversible. Alors A est équivalente ligne à la matrice identité I ; mais $|I| \neq 0$, donc d'après le problème précédent $|A| \neq 0$. D'autre part supposons A non inversible. Alors A est équivalente ligne à une matrice B qui a une ligne nulle. D'après le théorème 8.2 , 1) $|B| = 0$; alors d'après le problème précédent $|A| = 0$. Donc 1) et 3) sont équivalentes.

8.22. Démontrer le théorème 8.5 : $|AB| = |A| . |B|$.

Si A est singulière, alors AB est aussi singulière et donc $|AB| = 0 = |A| . |B|$. D'autre part si A est non singulière, alors $A = E_n \ldots E_2 E_1$ est un produit de matrices élémentaires. Ainsi d'après le problème 8.20,

$$|A| = |E_n \ldots E_2 E_1 I| = |E_n| \ldots |E_2|\,|E_1|\,|I| = |E_n| \ldots |E_2|\,|E_1|$$

et donc

$$|AB| = |E_n \ldots E_2 E_1 B| = |E_n| \ldots |E_2|\,|E_1|\,|B| = |A|\,|B|$$

8.23. Démontrer le théorème 8.7 : Soit $A = (a_{ij})$; alors $|A| = a_{i1} A_{i1} + a_{i2} A_{i2} + \cdots + a_{in} A_{in}$, où A_{ij} est le cofacteur de a_{ij}.

Chaque terme dans $|A|$ contient un et un seul élément de la $i^{\text{ème}}$ ligne $(a_{i1}, a_{i2}, \ldots, a_{in})$ de A. Donc nous pouvons écrire $|A|$ sous la forme

$$|A| = a_{i1} A_{i1}^* + a_{i2} A_{i2}^* + \cdots + a_{in} A_{in}^*$$

(Remarquons que A_{ij}^* est une somme de termes ne contenant aucun élément de la $i^{\text{ème}}$ ligne de A). Ainsi le théorème est démontré si nous pouvons montrer que

$$A_{ij}^* = A_{ij} = (-1)^{i+j}\,|M_{ij}|$$

où M_{ij} est la matrice obtenue en supprimant la ligne et la colonne contenant l'élément a_{ij}. (Historiquement, l'expression A_{ij}^* était définie comme le cofacteur de a_{ij} et le théorème se réduit à démontrer que les deux définitions du cofacteur sont équivalentes).

Considérons d'abord le cas où $i = n$, $j = n$. La somme des termes dans $|A|$ contenant a_{nn} est

$$a_{nn}A^*_{nn} = a_{nn} \sum_{\sigma} (\operatorname{sgn} \sigma)\, a_{1\sigma(1)}\, a_{2\sigma(2)} \cdots a_{n-1,\sigma(n-1)}$$

où nous sommons sur toutes les permutations $\sigma \in S_n$ pour lesquelles $\sigma(n) = n$, ce qui est équivalent (Problème 8.63) à sommer sur toutes les permutations de $\{1, \ldots, n-1\}$. Ainsi $A^*_{nn} = |M_{nn}| = (-1)^{n+n}|M_{nn}|$

Considérons maintenant i et j quelconques. Echangeons la $i^{\text{ème}}$ ligne avec la ligne qui lui succède et ainsi de suite jusqu'à ce qu'elle soit la dernière, et nous échangeons de même la $j^{\text{ème}}$ colonne avec celle lui succédant jusqu'à ce qu'elle soit la dernière. Remarquons que le déterminant $|M_{ij}|$ n'est pas affecté puisque les positions relatives des autres lignes et colonnes ne sont pas affectées par ces échanges. De toute façon, le "signe" de $|A|$ et de A^*_{ij} est changé $n - i$ et $n - j$ fois. En conséquence :

$$A^*_{ij} = (-1)^{n-i+n-j}\,|M_{ij}| = (-1)^{i+j}\,|M_{ij}|$$

8.24. Soit $A = (a_{ij})$ et soit B la matrice obtenue à partir de A en remplaçant la $i^{\text{ème}}$ ligne de A par le vecteur ligne $(b_{i1}, \ldots, b_{in})$. Montrer que

$$|B| = b_{i1}A_{i1} + b_{i2}A_{i2} + \cdots + b_{in}A_{in}$$

De plus, montrer que, pour $j \neq i$,

$$a_{j1}A_{i1} + a_{j2}A_{i2} + \cdots + a_{jn}A_{in} = 0$$

et

$$a_{1j}A_{1i} + a_{2j}A_{2i} + \cdots + a_{nj}A_{ni} = 0$$

Soit $B = (b_{ij})$. D'après le problème précédent,

$$|B| = b_{i1}B_{i1} + b_{i2}B_{i2} + \cdots + b_{in}B_{in}$$

Puisque B_{ij} ne dépend pas de la $i^{\text{ème}}$ ligne de B, $B_{ij} = A_{ij}$ pour $j = 1, \ldots, n$. Donc

$$|B| = b_{i1}A_{i1} + b_{i2}A_{i2} + \cdots + b_{in}A_{in}$$

Soit A' obtenue à partir de A en remplaçant la $i^{\text{ème}}$ ligne de A par la $j^{\text{ème}}$ ligne de A. Puisque A' a deux lignes identiques, $|A'| = 0$. Ainsi d'après le résultat précédent,

$$|A'| = a_{j1}A_{i1} + a_{j2}A_{i2} + \cdots + a_{jn}A_{in} = 0$$

En utilisant $|A^t| = |A|$, nous obtenons aussi que $a_{1j}A_{1i} + a_{2j}A_{2i} + \cdots + a_{nj}A_{ni} = 0$.

8.25. Démontrer le théorème 8.8 : $A \,.\, (\operatorname{adj} A) = (\operatorname{adj} A) \,.\, A = |A| \,.\, I$. Ainsi si $|A| \neq 0$, $A^{-1} = (1/|A|)\,(\operatorname{adj} A)$.

Soit $A = (a_{ij})$ et soit $A \,.\, (\operatorname{adj} A) = (b_{ij})$. La $i^{\text{ème}}$ ligne de A est

$$(a_{i1}, a_{i2}, \ldots, a_{in}) \tag{1}$$

Puisque adj A est la transposée de la matrice des cofacteurs, la $j^{\text{ème}}$ colonne de adj A est la transposée des cofacteurs de la $j^{\text{ème}}$ ligne de A:

$$(A_{j1}, A_{j2}, \ldots, A_{jn})^t \tag{2}$$

Maintenant b_{ij}, le ij élément de $A.(\operatorname{adj} A)$, est obtenu en multipliant (*1*) et (*2*):

$$b_{ij} = a_{i1}A_{j1} + a_{i2}A_{j2} + \cdots + a_{in}A_{jn}$$

Donc d'après le théorème 8.7 et le problème précédent,

$$b_{ij} = \begin{cases} |A| & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

En conséquence $A \,.\, (\operatorname{adj} A)$ est la matrice diagonale dont chaque élément diagonal est $|A|$. En d'autres termes, $A.(\operatorname{adj} A) = |A| \,.\, I$. De façon analogue $(\operatorname{adj} A) \,.\, A = |A|\, I$.

8.26. Démontrer le théorème 8.9 : Le système d'équations linéaires $Ax = b$ admet une solution unique si et seulement si $\Delta = |A| \neq 0$. Dans ce cas la solution unique est donnée par $x_1 = \Delta_1/\Delta,\ x_2 = \Delta_2/\Delta, \ldots, x_n = \Delta_n/\Delta$.

D'après les résultats précédents, $Ax = b$ a une solution unique si et seulement si A est inversible, et A est inversible si et seulement si $\Delta = |A| \neq 0$.

Supposons maintenant $\Delta \neq 0$. D'après le problème 8.25 $A^{-1} = (1/\Delta)$ (adj A). En multipliant $Ax = b$ par A^{-1}, on obtient

$$x = A^{-1}Ax = (1/\Delta)(\text{adj } A)b \tag{1}$$

Remarquons que la $i^{\text{ème}}$ ligne de $(1/\Delta)$ (adj A) est $(1/\Delta)$ $(A_{1i}, A_{2i}, \ldots, A_{ni})$. Si $b = (b_1, b_2, \ldots b_n)^t$ alors, d'après (1),

$$x_i = (1/\Delta)(b_1A_{1i} + b_2A_{2i} + \cdots + b_nA_{ni})$$

De toute façon, comme dans le problème 8.24,

$$b_1A_{1i} + b_2A_{2i} + \cdots + b_nA_{ni} = \Delta_i$$

le déterminant de la matrice obtenue en remplaçant la $i^{\text{ème}}$ colonne de A par le vecteur colonne b. Ainsi $x_i = (1/\Delta)\ \Delta_i$ comme il était exigé.

8.27. Supposons P inversible. Montrer que $|P^{-1}| = |P|^{-1}$.

$P^{-1}P = I$. Donc $1 = |I| = |P^{-1}P| = |P^{-1}|\,|P|$, et donc $|P^{-1}| = |P|^{-1}$.

8.28. Démontrer le théorème 8.11 : Supposons que A et B soient des matrices semblables. Alors $|A| = |B|$.

Comme A et B sont semblables, il existe une matrice inversible P telle que $B = P^{-1}AP$. Alors, d'après le précédent problème, $|B| = |P^{-1}AP| = |P^{-1}|\,|A|\,|P| = |A|\,|P^{-1}|\,|P| = |A|$.

Remarquons que, bien que les matrices P^{-1} et A ne commutent pas, leurs déterminants $|P^{-1}|$ et $|A|$ commutent puisqu'ils sont scalaires dans le champ K.

8.29. Démontrer le théorème 8.13: Il existe une fonction unique $D : \mathcal{A} \to K$ telle que 1) D est multilinéaire, 2) D est alternée, 3) $D(I) = 1$. Cette fonction D est la fonction déterminant, c'est-à-dire $D(A) = |A|$.

Soit D la fonction déterminant : $D(A) = |A|$. Nous devons montrer que D satisfait 1), 2) et 3) et que D est la seule fonction vérifiant 1), 2) et 3).

D'après les résultats précédents, D satisfait 2) et 3) ; donc il nous reste à montrer que D est multilinéaire. Supposons $A = (a_{ij}) = (A_1, A_2, \ldots, A_n)$ où A_k est la $k^{\text{ième}}$ ligne de A. De plus, supposons que pour i fixé,

$$A_i = B_i + C_i, \qquad \text{où} \qquad B_i = (b_1, \ldots, b_n) \quad \text{et} \quad C_i = (c_1, \ldots, c_n)$$

En conséquence $\quad a_{i1} = b_1 + c_1, \quad a_{i2} = b_2 + c_2, \quad \ldots, \quad a_{in} = b_n + c_n$

Développons $D(A) = |A|$ par rapport à la $i^{\text{ème}}$ ligne.

$$\begin{aligned} D(A) &= D(A_1, \ldots, B_i + C_i, \ldots, A_n) = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in} \\ &= (b_1 + c_1)A_{i1} + (b_2 + c_2)A_{i2} + \cdots + (b_n + c_n)A_{in} \\ &= (b_1A_{i1} + b_2A_{i2} + \cdots + b_nA_{in}) + (c_1A_{i1} + c_2A_{i2} + \cdots + c_nA_{in}) \end{aligned}$$

De toute façon d'après le problème 8.24, les deux sommes précédentes sont les déterminants des matrices obtenues à partir de A en remplaçant la $i^{\text{ème}}$ ligne par B_i et C_i respectivement. C'est-à-dire

$$\begin{aligned} D(A) &= D(A_1, \ldots, B_i + C_i, \ldots, A_n) \\ &= D(A_1, \ldots, B_i, \ldots, A_n) + D(A_1, \ldots, C_i, \ldots, A_n) \end{aligned}$$

De plus, d'après le théorème 8.3 1),

$$D(A_1, \ldots, kA_i, \ldots, A_n) = k\,D(A_1, \ldots, A_i, \ldots, A_n)$$

Ainsi D est multilinéaire, c'est-à-dire que D satisfait 3).

Nous devons maintenant démontrer l'unicité de D. Supposons que D satisfasse à 1), 2) et 3). Si $\{e_1, \ldots, e_n\}$ est la base usuelle de K^n, alors d'après 3) $D(e_1, e_2, \ldots, e_n) = D(I) = 1$. En utilisant 2) nous avons aussi (Problème 8.73)

$$D(e_{i_1}, e_{i_2}, \ldots, e_{i_n}) = \text{sgn}\,\sigma, \qquad \text{où} \quad \sigma = i_1 i_2 \ldots i_n \tag{1}$$

Supposons maintenant que $A = (a_{ij})$. Remarquons que la $k^{\text{ième}}$ ligne A_k de A est

$$A_k = (a_{k1}, a_{k2}, \ldots, a_{kn}) = a_{k1}e_1 + a_{k2}e_2 + \cdots + a_{kn}e_n$$

Ainsi $\qquad D(A) = D(a_{11}e_1 + \cdots + a_{1n}e_n,\, a_{21}e_1 + \cdots + a_{2n}e_n,\, \ldots,\, a_{n1}e_1 + \cdots + a_{nn}e_n)$

En utilisant la multilinéarité de D, nous pouvons écrire $D(A)$ sous forme d'une somme de termes :

$$\begin{aligned} D(A) &= \sum D(a_{1i_1}e_{i_1}, a_{2i_2}e_{i_2}, \ldots, a_{ni_n}e_{i_n}) \\ &= \sum (a_{1i_1} a_{2i_2} \ldots a_{ni_n})\, D(e_{i_1}, e_{i_2}, \ldots, e_{i_n}) \end{aligned} \tag{2}$$

où la somme est sommée sur toutes les suites $i_1, i_2, \ldots, i_n$ où $i_k \in \{1, \ldots, n\}$. Si deux des indices sont égaux, par exemple $i_j = i_k$, mais $j \neq k$, alors d'après (2)

$$D(e_{i_1}, e_{i_2}, \ldots, e_{i_n}) = 0$$

En utilisant 1) nous avons finalement

$$\begin{aligned} D(A) &= \sum_\sigma (a_{1i_1} a_{2i_2} \ldots a_{ni_n})\, D(e_{i_1}, e_{i_2}, \ldots, e_{i_n}) \\ &= \sum_\sigma (\text{sgn}\,\sigma)\, a_{1i_1} a_{2i_2} \ldots a_{ni_n}, \qquad \text{où} \quad \sigma = i_1 i_2 \ldots i_n \end{aligned}$$

Donc D est la fonction déterminant et donc le théorème est démontré.

PERMUTATIONS

8.30 Déterminer la parité de $\sigma = 542163$.

Méthode 1.

Il nous faut obtenir le nombre de paires (i, j), pour lesquelles $i > j$ et i précède j dans σ, c'est-à-dire :

3 nombres (5, 4 et 2) sont supérieurs à 1 et précèdent 1.

2 nombres (5 et 4) sont supérieurs à 2 et précèdent 2.

3 nombres (5, 4 et 6) sont supérieurs à 3 et précèdent 3.

1 nombre (5) est supérieur à 4 et précède 4.

0 nombres sont supérieurs à 5 et précèdent 5.

0 nombres sont supérieurs à 6 et précèdent 6.

Puisque $3 + 2 + 3 + 1 + 0 + 0 = 9$ est impair, σ est une permutation impaire et donc $\text{sgn}\,\sigma = -1$.

Méthode 2.

Transposons 1 à la première position comme suit :

$$5\ 4\ 2\ 1\ 6\ 3 \quad \text{d'où} \quad 1\ 5\ 4\ 2\ 6\ 3$$

Transposons 2 à la seconde position :

$$1\ 5\ 4\ 2\ 6\ 3 \quad \text{d'où} \quad 1\ 2\ 5\ 4\ 6\ 3$$

Transposons 3 à la troisième position :

$$1\ 2\ 5\ 4\ 6\ 3 \quad \text{d'où} \quad 1\ 2\ 3\ 5\ 4\ 6$$

Transposons 4 à la quatrième position :

$$1\ 2\ 3\ 5\ 4\ 6 \quad \text{d'où} \quad 1\ 2\ 3\ 4\ 5\ 6$$

Remarquons que 5 et 6 sont dans leurs positions normales. Dénombrons le nombre des nombres "sautés": $3 + 2 + 3 + 1 = 9$. Puisque 9 est impair, σ est une permutation impaire. (Remarque : cette méthode est essentiellement la même que la méthode précédente).

Méthode 3.

L'échange de deux nombres dans une permutation est équivalent au produit de la permutation par une transposition. Donc transformons σ en la permutation identique en utilisant des transpositions ; telles que

$$\begin{array}{cccccc}
5 & 4 & 2 & 1 & 6 & 3 \\
1 & 4 & 2 & 5 & 6 & 3 \\
1 & 2 & 4 & 5 & 6 & 3 \\
1 & 2 & 3 & 5 & 6 & 4 \\
1 & 2 & 3 & 4 & 6 & 5 \\
1 & 2 & 3 & 4 & 5 & 6
\end{array}$$

Puisqu'un nombre impair 5, de transpositions est utilisé, σ est une permutation impaire

8.31. Soient $\sigma = 24513$ et $\tau = 41352$ des permutations de S_5. Trouver 1) les permutations composées $\tau \circ \sigma$ et $\sigma \circ \tau$, 2) σ^{-1}.

Rappelons que $\sigma = 24513$ et $\tau = 41352$ sont des écritures condensées de

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 5 & 1 & 3 \end{pmatrix} \qquad \tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 1 & 3 & 5 & 2 \end{pmatrix}$$

ce qui veut dire

$$\sigma(1) = 2, \quad \sigma(2) = 4, \quad \sigma(3) = 5, \quad \sigma(4) = 1 \quad \text{et} \quad \sigma(5) = 3$$

et

$$\tau(1) = 4, \quad \tau(2) = 1, \quad \tau(3) = 3, \quad \tau(4) = 5 \quad \text{et} \quad \tau(5) = 2$$

(i)

$$\begin{array}{cccccc}
 & 1 & 2 & 3 & 4 & 5 \\
\sigma & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\
 & 2 & 4 & 5 & 1 & 3 \\
\tau & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\
 & 1 & 5 & 2 & 4 & 3
\end{array}
\qquad \text{et} \qquad
\begin{array}{cccccc}
 & 1 & 2 & 3 & 4 & 5 \\
\tau & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\
 & 4 & 1 & 3 & 5 & 2 \\
\sigma & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\
 & 1 & 2 & 5 & 3 & 4
\end{array}$$

Ainsi $\tau \circ \sigma = 15243$ et $\sigma \circ \tau = 12534$.

(ii)

$$\sigma^{-1} = \begin{pmatrix} 2 & 4 & 5 & 1 & 3 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 1 & 5 & 2 & 3 \end{pmatrix}$$

c'est-à-dire $\sigma^{-1} = 41523$.

8.32. Considérons une permutation quelconque $\sigma = j_1\, j_2 \ldots j_n$. Montrer que pour toute paire (i, k) telle que

$$i > k \quad \text{et} \quad i \text{ précède } k \text{ dans } \sigma$$

il existe une paire (i^*, k^*) telle que

$$i^* < k^* \quad \text{et} \quad \sigma(i^*) > \sigma(k^*) \tag{1}$$

et vice versa. Ainsi σ est paire ou impaire suivant qu'il y a un nombre pair ou impair de paires satisfaisant *(1)*.

Choisissons i^* et k^* de telle sorte que $\sigma(i^*) = i$ et $\sigma(k^*) = k$. Alors $i > k$ si et seulement si $\sigma(i^*) > \sigma(k^*)$ et i précède k dans σ si et seulement si $i^* < k^*$.

8.33. Considérons le polynôme $g = g(x_1, \ldots, x_n) = \prod_{i<j} (x_i - x_j)$. Ecrivons explicitement le polynôme $g = g(x_1, x_2, x_3, x_4)$.

Le symbole Π est utilisé pour un produit de termes, de la même façon que le symbole Σ est utilisé pour une somme de termes. C'est-à-dire $\prod_{i<j} (x_i - x_j)$ représente le produit de tous les facteurs $(x_i - x_j)$ pour lesquels $i < j$. Donc

$$g = g(x_1, \ldots, x_4) = (x_1 - x_2)(x_1 - x_3)(x_1 - x_4)(x_2 - x_3)(x_2 - x_4)(x_3 - x_4)$$

8.34. Soit σ une permutation arbitraire. En considérant le polynôme g du problème précédent défini par $\sigma(g) = \prod_{i<j} (x_{\sigma(i)} - x_{\sigma(j)})$. Montrer que

$$\sigma(g) = \begin{cases} g & \text{si } \sigma \text{ est paire} \\ -g & \text{si } \sigma \text{ est impaire} \end{cases}$$

En conséquence $\sigma(g) = (\operatorname{sgn} \sigma)\, g$.

Puisque σ est injective et surjective,

$$\sigma(g) = \prod_{i<j} (x_{\sigma(i)} - x_{\sigma(j)}) = \prod_{i<j \text{ ou } i>j} (x_i - x_j)$$

Ainsi $\sigma(g) = g$ ou $\sigma(g) = -g$ suivant qu'il y a un nombre pair ou impair de facteurs de la forme $(x_i - x_j)$ où $i > j$. Remarquons que pour chaque paire (i, j) pour laquelle

$$i < j \quad \text{et} \quad \sigma(i) > \sigma(j) \tag{1}$$

il y a un facteur $(x_{\sigma(i)} - x_{\sigma(j)})$ dans $\sigma(g)$ pour lequel $\sigma(i) > \sigma(j)$. Puisque σ est paire si et seulement si il y a un nombre pair de paires satisfaisant (*1*), nous avons $\sigma(g) = g$ si et seulement si σ est paire ; en conséquence $\sigma(g) = -g$ si et seulement si σ est impaire.

8.35. Soit $\sigma, \tau \in S_n$. Montrer que $\operatorname{sgn}(\tau \circ \sigma) = (\operatorname{sgn} \tau) . (\operatorname{sgn} \sigma)$. Ainsi le produit de deux permutations paires ou impaires est pair, et le produit d'une permutation impaire par une permutation paire est impair.

En utilisant le précédent problème, nous avons

$$\operatorname{sgn}(\tau \circ \sigma) g = (\tau \circ \sigma)(g) = \tau(\sigma(g)) = \tau((\operatorname{sgn} \sigma) g) = (\operatorname{sgn} \tau)(\operatorname{sgn} \sigma) g$$

En conséquence $\operatorname{sgn}(\tau \circ \sigma) = (\operatorname{sgn} \tau)(\operatorname{sgn} \sigma)$.

8.36. Considérons la permutation $\sigma = j_1 j_2 \ldots, j_n$. Montrer que $\operatorname{sgn} \sigma^{-1} = \operatorname{sgn} \sigma$, et pour les scalaires a_{ij},

$$a_{j_1 1} a_{j_2 2} \ldots a_{j_n n} = a_{1k_1} a_{2k_2} \ldots a_{nk_n} \qquad \text{où} \quad \sigma^{-1} = k_1 k_2 \ldots k_n$$

Nous avons $\sigma^{-1} \circ \sigma = \epsilon$, la permutation identique. Puisque ϵ est paire, σ^{-1} et σ sont toutes les deux paires ou toutes les deux impaires. Donc $\operatorname{sgn} \sigma^{-1} = \operatorname{sgn} \sigma$.

Puisque $\sigma = j_1 j_2 \ldots j_n$ est une permutation, $a_{j_1 1}, a_{j_2 2}, \ldots, a_{j_n n} = a_{1k_1} a_{2k_2} \ldots a_{nk_n}$. Alors $k_1, k_2, \ldots, k_n$ ont la propriété suivante :

$$\sigma(k_1) = 1, \quad \sigma(k_2) = 2, \quad \ldots, \quad \sigma(k_n) = n$$

Soit $\tau = k_1 k_2 \ldots k_n$. Alors pour $i = 1, \ldots, n$,

$$(\sigma \circ \tau)(i) = \sigma(\tau(i)) = \sigma(k_i) = i$$

Ainsi $\sigma \circ \tau = \epsilon$, la permutation identique ; donc $\tau = \sigma^{-1}$.

PROBLEMES DIVERS

8.37. Trouver $\det(T)$ pour chacun des opérateurs linéaires T suivants :

1) T est l'opérateur de $\mathbf{R}^3$ défini par

$$T(x, y, z) = (2x - z, x + 2y - 4z, 3x - 3y + z)$$

2) T est l'opérateur de l'espace vectoriel V des matrices carrées 2×2 sur K défini par $T(A) = MA$ où $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

1) Trouvons la représentation matricielle de T, relativement à la base usuelle qui est $[T] = \begin{pmatrix} 2 & 0 & -1 \\ 1 & 2 & -4 \\ 3 & -3 & 1 \end{pmatrix}$.

Alors

$$\det(T) = \begin{vmatrix} 2 & 0 & -1 \\ 1 & 2 & -4 \\ 3 & -3 & 1 \end{vmatrix} = 2(2-12) - 1(-3-6) = -11$$

2) Trouvons une représentation matricielle de T dans une base quelconque de V, par exemple

$$\left\{E_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \; E_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \; E_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \; E_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right\}$$

Alors
$$T(E_1) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} a & 0 \\ c & 0 \end{pmatrix} = aE_1 + 0E_2 + cE_3 + 0E_4$$

$$T(E_2) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & a \\ 0 & c \end{pmatrix} = 0E_1 + aE_2 + 0E_3 + cE_4$$

$$T(E_3) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} b & 0 \\ d & 0 \end{pmatrix} = bE_1 + 0E_2 + dE_3 + 0E_4$$

$$T(E_4) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & b \\ 0 & d \end{pmatrix} = 0E_1 + bE_2 + 0E_3 + dE_4$$

Ainsi $[T]_E = \begin{pmatrix} a & 0 & c & 0 \\ 0 & a & 0 & c \\ b & 0 & d & 0 \\ 0 & b & 0 & d \end{pmatrix}$ et

$$\det(T) = \begin{vmatrix} a & 0 & c & 0 \\ 0 & a & 0 & c \\ b & 0 & d & 0 \\ 0 & b & 0 & d \end{vmatrix} = a\begin{vmatrix} a & 0 & c \\ 0 & d & 0 \\ b & 0 & d \end{vmatrix} + c\begin{vmatrix} 0 & a & c \\ b & 0 & 0 \\ 0 & b & d \end{vmatrix} = a^2d^2 + b^2c^2 - 2abcd$$

8.38. Trouver l'inverse de $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$.

L'inverse de A est de la forme (Problème 8.53) $A^{-1} = \begin{pmatrix} 1 & x & y \\ 0 & 1 & z \\ 0 & 0 & 1 \end{pmatrix}$.

Posons $AA^{-1} = I$, matrice identité :

$$AA^{-1} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & x & y \\ 0 & 1 & z \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & x+1 & y+z+1 \\ 0 & 1 & z+1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I$$

Egalons les éléments correspondants entre eux; on obtient le système

$$x + 1 = 0, \quad y + z + 1 = 0, \quad z + 1 = 0$$

La solution du système est $x = -1$, $y = 0$, $z = -1$. Donc $A^{-1} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$

A^{-1} peut être aussi trouvé grâce à la formule $A^{-1} = (\text{adj } A)/|A|$.

8.39. Soit D une fonction bilinéaire alternée. Montrer que $D(A, B) = -D(B, A)$. Plus généralement, montrer que si D est multilinéaire et alternée alors

$$D(\ldots, A, \ldots, B, \ldots) = -D(\ldots, B, \ldots, A, \ldots).$$

C'est-à-dire que le signe est changé chaque fois que deux composantes, deux termes sont échangés.

Puisque D est alternée, $D(A + B, A + B) = 0$. De plus, puisque D est multilinéaire,

$$\begin{aligned} 0 = D(A+B, A+B) &= D(A, A+B) + D(B, A+B) \\ &= D(A, A) + D(A, B) + D(B, A) + D(B, B) \end{aligned}$$

Mais $D(A, A) = 0$ et $D(B, B) = 0$, donc

$$0 = D(A, B) + D(B, A) \quad \text{ou} \quad D(A, B) = -D(B, A)$$

De façon analogue

$$\begin{aligned} 0 &= D(\ldots, A+B, \ldots, A+B, \ldots) \\ &= D(\ldots, A, \ldots, A, \ldots) + D(\ldots, A, \ldots, B, \ldots) \\ &\quad + D(\ldots, B, \ldots, A, \ldots) + D(\ldots, B, \ldots, B, \ldots) \\ &= D(\ldots, A, \ldots, B, \ldots) + D(\ldots, B, \ldots, A, \ldots) \end{aligned}$$

et donc $D(\ldots, A \ldots, B, \ldots) = -D(\ldots, B, \ldots, A, \ldots)$

8.40. Soit V l'espace vectoriel des matrices 2×2 $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ sur $\mathbf{R}$. Déterminer si oui ou non $D : V \to \mathbf{R}$ est bilinéaire (par rapport aux lignes) si (i) $D(M) = a + d$, (ii) $D(M) = ad$.

(i) Non. Par exemple, supposons $A = (1, 1)$ et $B = (3, 3)$; alors

$$D(A, B) = D\begin{pmatrix} 1 & 1 \\ 3 & 3 \end{pmatrix} = 4 \quad \text{et} \quad D(2A, B) = D\begin{pmatrix} 2 & 2 \\ 3 & 3 \end{pmatrix} = 5 \neq 2D(A, B)$$

(ii) Oui : Soient $A = (a_1, a_2)$, $B = (b_1, b_2)$ et $C = (c_1, c_2)$; alors

$$D(A, C) = D\begin{pmatrix} a_1 & a_2 \\ c_1 & c_2 \end{pmatrix} = a_1c_2 \quad \text{et} \quad D(B, C) = D\begin{pmatrix} b_1 & b_2 \\ c_1 & c_2 \end{pmatrix} = b_1c_2$$

Donc quels que soient les scalaires s, $t \in \mathbf{R}$,

$$\begin{aligned} D(sA + tB, C) &= D\begin{pmatrix} sa_1 + tb_1 & sa_2 + tb_2 \\ c_1 & c_2 \end{pmatrix} = (sa_1 + tb_1)c_2 \\ &= s(a_1c_2) + t(b_1c_2) = s\,D(A, C) + t\,D(B, C) \end{aligned}$$

C'est-à-dire, D est linéaire par rapport à la première ligne.

De plus

$$D(C, A) = D\begin{pmatrix} c_1 & c_2 \\ a_1 & a_2 \end{pmatrix} = c_1a_2 \quad \text{et} \quad D(C, B) = D\begin{pmatrix} c_1 & c_2 \\ b_1 & b_2 \end{pmatrix} = c_1b_2$$

Donc quels que soient les scalaires s, $t \in \mathbf{R}$,

$$\begin{aligned} D(C, sA + tB) &= D\begin{pmatrix} c_1 & c_2 \\ sa_1 + tb_1 & sa_2 + tb_2 \end{pmatrix} = c_1(sa_2 + tb_2) \\ &= s(c_1a_2) + t(c_1b_2) = s\,D(C, A) + t\,D(C, B) \end{aligned}$$

Donc D est linéaire par rapport à la seconde ligne.

Les deux conditions de linéarité impliquent que D est bilinéaire.

PROBLEMES SUPPLEMENTAIRES

CALCUL DE DETERMINANTS

8.41. Calculer le déterminant de chacune des matrices : (i) $\begin{pmatrix} 2 & 5 \\ 4 & 1 \end{pmatrix}$, (ii) $\begin{pmatrix} 6 & 1 \\ 3 & -2 \end{pmatrix}$.

8.42. Calculer le déterminant de chacune des matrices: (i) $\begin{pmatrix} t-2 & -3 \\ -4 & t-1 \end{pmatrix}$, (ii) $\begin{pmatrix} t-5 & 7 \\ -1 & t+3 \end{pmatrix}$.

8.43. Pour chacune des matrices du problème précédent, trouver les valeurs de t pour lesquelles le déterminant est nul.

8.44. Calculer le déterminant de chacune des matrices:

$$(i) \begin{pmatrix} 2 & 1 & 1 \\ 0 & 5 & -2 \\ 1 & -3 & 4 \end{pmatrix}, \quad (ii) \begin{pmatrix} 3 & -2 & -4 \\ 2 & 5 & -1 \\ 0 & 6 & 1 \end{pmatrix}, \quad (iii) \begin{pmatrix} -2 & -1 & 4 \\ 6 & -3 & -2 \\ 4 & 1 & 2 \end{pmatrix}, \quad (iv) \begin{pmatrix} 7 & 6 & 5 \\ 1 & 2 & 1 \\ 3 & -2 & 1 \end{pmatrix}.$$

8.45. Calculer le déterminant de chacune des matrices:

$$(i) \begin{pmatrix} t-2 & 4 & 3 \\ 1 & t+1 & -2 \\ 0 & 0 & t-4 \end{pmatrix}, \quad (ii) \begin{pmatrix} t-1 & 3 & -3 \\ -3 & t+5 & -3 \\ -6 & 6 & t-4 \end{pmatrix}, \quad (iii) \begin{pmatrix} t+3 & -1 & 1 \\ 7 & t-5 & 1 \\ 6 & -6 & t+2 \end{pmatrix}.$$

8.46. Pour chacune des matrices du problème précédent, déterminer les valeurs de t pour lesquelles le déterminant est nul.

8.47. Calculer le déterminant de chacune des matrices: (i) $\begin{pmatrix} 1 & 2 & 2 & 3 \\ 1 & 0 & -2 & 0 \\ 3 & -1 & 1 & -2 \\ 4 & -3 & 0 & 2 \end{pmatrix}$, (ii) $\begin{pmatrix} 2 & 1 & 3 & 2 \\ 3 & 0 & 1 & -2 \\ 1 & -1 & 4 & 3 \\ 2 & 2 & -1 & 1 \end{pmatrix}$.

COFACTEURS, ADJOINTS CLASSIQUES, INVERSES

8.48. Pour la matrice $\begin{pmatrix} 1 & 2 & -2 & 3 \\ 3 & -1 & 5 & 0 \\ 4 & 0 & 2 & 1 \\ 1 & 7 & 2 & -3 \end{pmatrix}$, trouver le cofacteur de :

(i) l'élément 4, (ii) l'élément 5, (iii) l'élément 7.

8.49. Soit $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 2 & 1 \end{pmatrix}$. Trouver (i) adj A, (ii) A^{-1}.

8.50. Soit $A = \begin{pmatrix} 1 & 2 & 2 \\ 3 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$. Trouver (i) adj A, (ii) A^{-1}.

8.51. Trouver l'adjoint classique de chaque matrice dans le problème 8.47.

8.52. Déterminer la matrice générale 2×2 A pour laquelle $A = \text{adj } A$.

8.53. Supposons A diagonale et B triangulaire ; c'est-à-dire

$$A = \begin{pmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_n \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} b_1 & c_{12} & \dots & c_{1n} \\ 0 & b_2 & \dots & c_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & b_n \end{pmatrix}$$

(i) Montrer que adj A est diagonale et adj B est triangulaire.

(ii) Montrer que B est inversible si tous les $b_i \neq 0$; donc A est inversible ssi tous les $a_i \neq 0$.

(iii) Montrer que les inverses de A et B (s'ils existent) sont de la forme

$$A^{-1} = \begin{pmatrix} a_1^{-1} & 0 & \dots & 0 \\ 0 & a_2^{-1} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_n^{-1} \end{pmatrix}, \quad B^{-1} = \begin{pmatrix} b_1^{-1} & d_{12} & \dots & d_{1n} \\ 0 & b_2^{-1} & \dots & d_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & b_n^{-1} \end{pmatrix}$$

C'est-à-dire que, les éléments diagonaux de A^{-1} et B^{-1} sont les inverses des éléments correspondants diagonaux de A et B.

DETERMINANT D'UN OPERATEUR LINEAIRE

8.54. Soit T l'opérateur linéaire de $\mathbf{R}^3$ définie par

$$T(x, y, z) = (3x - 2z, 5y + 7z, x + y + z)$$

Trouver det (T).

8.55. Soit $D : V \to V$ l'opérateur différentiel ; c'est-à-dire $D(V) = dv/dt$. Trouver det (D) si V est l'espace engendré par (i) $\{1, t, \dots, t^n\}$. (ii) $\{e^t, e^{2t}, e^{3t}\}$. (iii) $\{\sin t, \cos t\}$.

8.56. Démontrer le théorème 8.12 : Soient T et S des opérateurs linéaires de V. Alors :
(i) det $(S \circ T)$ = det (S) · det (T) ; (ii) T est inversible si et seulement si det $(T) \neq 0$.

8.57. Montrer que (i) det $(1_V) = 1$ où 1_V est l'opérateur identique de V, (ii) det (T^{-1}) = det $(T)^{-1}$ si T est inversible.

DETERMINANTS ET EQUATIONS LINEAIRES

8.58. Résoudre à l'aide des déterminants : (i) $\begin{cases} 3x + 5y = 8 \\ 4x - 2y = 1 \end{cases}$ (ii) $\begin{cases} 2x - 3y = -1 \\ 4x + 7y = -1 \end{cases}$.

8.59. Résoudre à l'aide des déterminants : (i) $\begin{cases} 2x - 5y + 2z = 7 \\ x + 2y - 4z = 3 \\ 3x - 4y - 6z = 5 \end{cases}$ (ii) $\begin{cases} 2z + 3 = y + 3x \\ x - 3z = 2y + 1 \\ 3y + z = 2 - 2x \end{cases}$.

8.60. Démontrer le théorème 8.10 : Le système homogène $Ax = 0$ a une solution non nulle si et seulement si $\Delta = |A| = 0$.

PERMUTATIONS

8.61. Déterminer la parité des permutations suivantes de S_5 : (i) $\sigma = 3\ 2\ 1\ 5\ 4$, (ii) $\tau = 1\ 3\ 5\ 2\ 4$, (iii) $\pi = 4\ 2\ 5\ 3\ 1$.

8.62. Pour les permutations σ, τ et π du problème 8.61, trouver (i) $\tau \circ \sigma$, (ii) $\pi \circ \sigma$, (iii) σ^{-1}, (iv) τ^{-1}.

8.63. Soit $\tau \in S_n$. Montrer que $\tau \circ \sigma$ parcourt S_n lorsque σ parcourt S_n ; c'est-à-dire $S_n = \{\tau \circ \sigma : \sigma \in S_n\}$.

8.64. Soit $\sigma \in S_n$ qui a la propriété $\sigma(n) = n$. Soit $\sigma^* \in S_{n-1}$ définie par $\sigma^*(x) = \sigma(x)$. (i) Montrer que sgn σ^* = sgn σ. (ii) Montrer que lorsque σ parcourt S_n avec $\sigma(n) = n$, σ^* parcourt S_{n-1} ; c'est-à-dire $S_{n-1} = \{\sigma^* : \sigma \in S_n, \sigma(n) = n\}$.

MULTILINEARITE

8.65. Soit $V = (K^m)^m$ c'est-à-dire que V est l'espace des matrices carrées $m \times m$ considérées comme m-tuples de vecteurs lignes. Soit $D : V \to K$.

(i) Montrer que :

$$D(A_1, A_2, \dots, A_n) = 0 \quad \text{avec} \quad A_i = A_{i+1} \quad \text{pour } i \text{ quelconque}$$

est équivalent à dire que D est alterné.

(ii) Supposons D m-linéaire et alternée. Montrer que si $A_1, A_2, \dots, A_m$ sont linéairement dépendantes alors $D(A_1, A_2, \dots, A_m) = 0$.

8.66. Soit V l'espace vectoriel des matrices 2×2 $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ sur $\mathbf{R}$. Déterminer si oui ou non $D : V \to \mathbf{R}$ est bilinéaire (par rapport aux lignes) si (i) $D(M) = ac - bd$, (ii) $D(M) = ab - cd$, (iii) $D(M) = 0$, (iv) $D(M) = 1$.

8.67. Soit V l'espace des matrices carrées $n \times n$ sur K. Supposons $B \in V$ inversible et donc $\det(B) \neq 0$. Définissons $D : V \to K$ par $D(A) = \det(AB)/\det(B)$ où $A \in V$. Donc

$$D(A_1, A_2, \ldots, A_n) = \det(A_1B, A_2B, \ldots, A_nB)/\det(B)$$

où A_i est la $i$ème ligne de A et donc A_iB est la $i$ème ligne de AB. Montrer que D est multilinéaire et alternée et que $D(I) = 1$. (Ainsi d'après le 8.13, $D(A) = \det(A)$ et donc $\det(AB) = \det(A)\det(B)$. Cette méthode est utilisée dans quelques livres pour démontrer le Théorème 8.5, c'est-à-dire $|AB| = |A|\,|B|$).

PROBLEMES DIVERS

8.68. Soit A une matrice carrée $n \times n$. Démontrer que $|kA| = k^n\,|A|$.

8.69. Démontrer que :

$$\begin{vmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \ldots & x_2^{n-1} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & x_n & x_n^2 & \ldots & x_n^{n-1} \end{vmatrix} = (-1)^n \prod_{i<j} (x_i - x_j)$$

Le déterminant précédent est appelé déterminant de Vandermonde d'ordre n.

8.70. Considérons la matrice bloc $M = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$ où A et C sont des matrices carrées. Démontrer que $|M| = |A|\,|C|$. Plus généralement, démontrer que si M est une matrice bloc triangulaire avec des matrices carrées $A_1, \ldots, A_m$ sur la diagonale, alors $|M| = |A_1| \cdot |A_2| \ldots |A_m|$.

8.71. Soient A, B, C et D des matrices carrées $n \times n$ commutables. Considérons la matrice bloc carrée $2n \times 2n$ $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$. Démontrer que $|M| = |A| \cdot |D| - |B| \cdot |C|$.

8.72. Supposons que A soit orthogonale, c'est-à-dire que $A^tA = I$. Montrer que $|A| = \pm 1$.

8.73. Considérons une permutation $\sigma = j_1 j_2 \ldots j_n$. Soit $\{e_i\}$ la base usuelle de K^n, et soit A la matrice dont les $i$èmes lignes sont e_{ji}, c'est-à-dire $A = (e_{j1}, e_{j2}, \ldots, e_{jn})$. Montrer que $|A| = \operatorname{sgn} \sigma$.

8.74. Soit A une matrice carrée $n \times n$. Le rang déterminant de A est l'ordre de la plus grande sous-matrice de A (obtenue en supprimant des lignes et des colonnes de A) dont le déterminant n'est pas nul. Montrer que le rang déterminant de A est égal au rang de A, c'est-à-dire au nombre maximum de lignes linéairement indépendantes (ou de colonnes).

REPONSES AUX PROBLEMES SUPPLEMENTAIRES

8.41. (i) -18, (ii) -15.

8.42. (i) $t^2 - 3t - 10$, (ii) $t^2 - 2t - 8$.

8.43. (i) $t = 5$, $t = -2$; (ii) $t = 4$, $t = -2$.

8.44. (i) 21, (ii) -11, (iii) 100, (iv) 0.

8.45. (i) $(t+2)(t-3)(t-4)$, (ii) $(t+2)^2(t-4)$, (iii) $(t+2)^2(t-4)$.

8.46. (i) 3, 4, -2; (ii) 4, -2; (iii) 4, -2.

8.47. (i) -131, (ii) -55.

8.48. (i) -135, (ii) -103, (iii) -31.

8.49. $\text{adj}\, A = \begin{pmatrix} -1 & -1 & 1 \\ -1 & 1 & -1 \\ 2 & -2 & 0 \end{pmatrix}$, $A^{-1} = (\text{adj}\, A)/|A| = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ -1 & 1 & 0 \end{pmatrix}$.

8.50. $\text{adj}\, A = \begin{pmatrix} 1 & 0 & -2 \\ -3 & -1 & 6 \\ 2 & 1 & -5 \end{pmatrix}$, $A^{-1} = \begin{pmatrix} -1 & 0 & 2 \\ 3 & 1 & -6 \\ -2 & -1 & 5 \end{pmatrix}$.

8.51. (i) $\begin{pmatrix} -16 & -29 & -26 & -2 \\ -30 & -38 & -16 & 29 \\ -8 & 51 & -13 & -1 \\ -13 & 1 & 28 & -18 \end{pmatrix}$ (ii) $\begin{pmatrix} 21 & -14 & -17 & -19 \\ -44 & 11 & 33 & 11 \\ -29 & 1 & 13 & 21 \\ 17 & 7 & -19 & -18 \end{pmatrix}$

8.52. $A = \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}$

8.54. $\det(T) = 4$.

8.55. (i) 0, (ii) 6, (iii) 1.

8.58. (i) $x = 21/26$, $y = 29/26$; (ii) $x = -5/13$, $y = 1/13$.

8.59. (i) $x = 5$, $y = 1$, $z = 1$. (ii) Puisque $\Delta = 0$ le système ne peut être résolu à l'aide des déterminants.

8.61. $\text{sgn}\, \sigma = 1$, $\text{sgn}\, \tau = -1$, $\text{sgn}\, \pi = -1$.

8.62. (i) $\tau \circ \sigma = 53142$, (ii) $\pi \circ \sigma = 52413$, (iii) $\sigma^{-1} = 32154$, (iv) $\tau^{-1} = 14253$.

8.66. (i) Oui (ii) Non (iii) Oui (iv) Non.

CHAPITRE 9

Valeurs propres et vecteurs propres

INTRODUCTION

Dans ce chapitre nous approfondirons la théorie d'un opérateur linéaire T sur un espace vectoriel de dimension finie. En particulier, nous trouverons les conditions pour que T soit diagonalisable. Comme nous l'avons vu dans le chapitre 7, cette question est en relation étroite avec la théorie des transformations de similitude pour les matrices.

Nous associerons aussi certains polynômes avec un opérateur T : son polynôme caractéristique et son polynôme minimal. Ces polynômes et leurs racines tiennent une très grande place dans l'étude de T. Le corps K tient aussi une place importante dans la théorie puisque l'existence des racines d'un polynôme dépend de K.

POLYNOMES DE MATRICES ET OPERATEURS LINEAIRES

Considérons un polynôme $f(t)$ sur un corps K : $f(t) = a_n t^n + \cdots + a_1 t + a_0$. Si A est une matrice carrée sur K, nous définissons alors

$$f(A) = a_n A^n + \cdots + a_1 A + a_0 I$$

où I est la matrice identité. En particulier, on dit que A est une racine ou un zéro du polynôme $f(t)$ si $f(A) = 0$.

Exemple 9.1 Soit $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, et soit $f(t) = 2t^2 - 3t + 7$, $g(t) = t^2 - 5t - 2$. Alors

$$f(A) = 2\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}^2 - 3\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + 7\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 18 & 14 \\ 21 & 39 \end{pmatrix}$$

et

$$g(A) = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}^2 - 5\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} - 2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Donc A est un zéro de $g(t)$.

On a le théorème suivant :

Théorème 9.1 : Soient f et g deux polynômes sur K, et soit A une matrice carrée $n \times n$ sur K. Alors

(i) $(f + g)(A) = f(A) + g(A)$

(ii) $(fg)(A) = f(A)\,g(A)$

et pour tout scalaire $k \in K$

(iii) $(kf)(A) = kf(A)$

De plus, puisque $f(t)\,g(t) = g(t)\,f(t)$ quels que soient les polynômes $f(t)$ et $g(t)$,

$$f(A)\,g(A) = g(A)\,f(A)$$

C'est-à-dire, deux polynômes quelconques de la matrice A commutent.

Supposons maintenant que $T : V \to V$ est un opérateur linéaire sur un espace vectoriel V sur K. Si $f(t) = a_n t^n + \cdots + a_1 t + a_0$, nous définissons $f(T)$ de la même manière que nous avons fait pour les matrices :

$$f(T) = a_n T^n + \cdots + a_1 T + a_0 I$$

où I désigne l'application identité. Nous disons aussi que T est un zéro ou une racine de $f(t)$ si $f(T) = 0$. Remarquons que les relations du théorème 9.1 restent vraies pour des opérateurs comme d'ailleurs pour des matrices ; donc deux polynômes quelconques dans T commutent.

De plus, si A est la représentation matricielle de T, alors $f(A)$ est la représentation matricielle de $f(T)$. En particulier, $f(T) = 0$ si et seulement si $f(A) = 0$.

VALEURS PROPRES ET VECTEURS PROPRES

Soit $T : V \to V$ un opérateur linéaire sur un espace vectoriel V sur un corps K. Un scalaire $\lambda \in K$ est appelé valeur propre de T s'il existe un vecteur non nul $v \in V$ pour lequel

$$T(v) = \lambda v$$

Tout vecteur satisfaisant cette relation est alors appelé un vecteur propre de T correspondant à la valeur propre λ. Remarquons que tout multiple kv de v est aussi un vecteur propre :

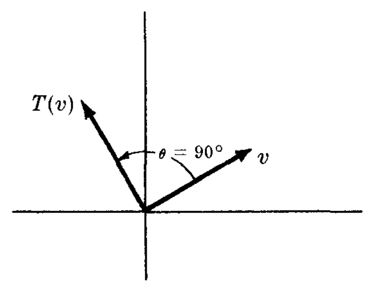
$$T(kv) = k\,T(v) = k(\lambda v) = \lambda(kv)$$

L'ensemble de tous les vecteurs propres est un sous-espace de V (Problème 9.6) appelé espace propre associé à λ.

On emploie aussi fréquemment la terminologie valeur caractéristique et vecteur caractéristique au lieu de valeur propre et vecteur propre.

Exemple 9.2 : Soit $I : V \to V$ l'application identique. Pour tout $v \in V$, $I(v) = v = 1v$. Donc 1 est une valeur propre de I et tout vecteur de V est un vecteur propre associé au scalaire 1.

Exemple 9.3 : Soit $T : \mathbf{R}^2 \to \mathbf{R}^2$ l'opérateur linéaire qui fait subir à chaque vecteur $v \in \mathbf{R}^2$ une rotation d'un angle $\theta = 90°$. Remarquons qu'aucun vecteur non nul n'est transformé en un multiple de lui-même. Donc T n'a aucune valeur propre et donc aussi aucun vecteur propre.



Exemple 9.4 : Soit D l'opérateur différentiel sur l'espace vectoriel V des fonctions différentiables. Nous avons $D(e^{5t}) = 5e^{5t}$. Donc 5 est une valeur propre de D avec un vecteur propre e^{5t}.

Si A est une matrice carrée $n \times n$ sur K, alors une valeur propre de A est considérée comme une valeur propre de A considérée comme un opérateur de K^n. C'est-à-dire, $\lambda \in K$ est une valeur propre de A si, pour un vecteur (colonne) non nul $v \in K^n$,

$$Av = \lambda v$$

Dans ce cas v est un vecteur propre de A associé à λ.

Exemple 9.5 : Trouver les valeurs propres et les vecteurs propres non nuls associés à la matrice $A = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$.

Cherchons un scalaire t et un vecteur non nul $X = \begin{pmatrix} x \\ y \end{pmatrix}$ tel que $AX = tX$:

$$\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = t\begin{pmatrix} x \\ y \end{pmatrix}$$

L'équation matricielle précédente est équivalente au système homogène

$$\begin{cases} x + 2y = tx \\ 3x + 2y = ty \end{cases} \quad \text{ou} \quad \begin{cases} (t-1)x - 2y = 0 \\ -3x + (t-2)y = 0 \end{cases} \qquad (1)$$

Rappelons que le système homogène a une solution non nulle si et seulement si le déterminant de la matrice des coefficients est 0 :

$$\begin{vmatrix} t-1 & -2 \\ -3 & t-2 \end{vmatrix} = t^2 - 3t - 4 = (t-4)(t+1) = 0$$

Ainsi t est une valeur propre de A si et seulement si $t = 4$ ou $t = -1$.

Posons $t = 4$ dans (*1*),

$$\begin{cases} 3x - 2y = 0 \\ -3x + 2y = 0 \end{cases} \quad \text{ou simplement} \quad 3x - 2y = 0$$

Ainsi $v = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$ est un vecteur propre non nul correspondant à la valeur propre $t = 4$, et tout vecteur propre correspondant à $t = 4$ est un multiple de v.

Posons $t = -1$ dans (*1*),

$$\begin{cases} -2x - 2y = 0 \\ -3x - 3y = 0 \end{cases} \quad \text{ou simplement} \quad x + y = 0$$

Ainsi $w = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ est un vecteur propre non nul correspondant à la valeur propre $t = -1$, et tout vecteur propre correspondant à $t = -1$ est un multiple de w.

Le théorème suivant donne une propriété caractéristique des valeurs propres qui est fréquemment utilisée comme définition.

Théorème 9.2 : Soit $T : V \to V$ un opérateur linéaire sur l'espace vectoriel sur K. $\lambda \in K$ est une valeur propre de T si et seulement si l'opérateur $\lambda I - T$ est singulier. L'espace propre de λ est le noyau de $\lambda I - T$.

Démonstration : λ est une valeur propre de T si et seulement si il existe un vecteur non nul v tel que

$$T(v) = \lambda v \quad \text{ou} \quad (\lambda I)(v) - T(v) = 0 \quad \text{ou} \quad (\lambda I - T)(v) = 0$$

c'est-à-dire $\lambda I - T$ est singulier. Nous avons aussi démontré que v est un espace propre de λ si et seulement si les relations précédentes sont vérifiées ; donc v appartient au noyau de $\lambda I - T$.

Nous allons maintenant démontrer un théorème très utile qui est démontré (Problème 9.14) par récurrence.

Théorème 9.3 : Des vecteurs propres distincts correspondant à des valeurs propres distinctes sont linéairement indépendants.

Exemple 9.6 : Considérons les fonctions $e^{a_1 t}, e^{a_2 t}, \ldots, e^{a_n t}$ où $a_1, a_2, \ldots, a_n$ sont des nombres réels distincts. Si D est l'opérateur différentiel, alors $D(e^{a_k t}) = a_k e^{a_k t}$. En conséquence $e^{a_1 t}, \ldots, e^{a_n t}$ sont les vecteurs propres de D correspondant aux valeurs propres distinctes $a_1, a_2, \ldots, a_n$ et donc d'après le théorème 9.3 sont linéairement indépendants.

Remarquons que des vecteurs propres indépendants peuvent être associés à la même valeur propre (voir Problème 9.7).

DIAGONALISATION ET VECTEURS PROPRES

Soit $T : V \to V$ un opérateur linéaire d'un espace vectoriel V de dimension finie n. Remarquons que T peut être représenté par une matrice diagonale

$$\begin{pmatrix} k_1 & 0 & \ldots & 0 \\ 0 & k_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & k_n \end{pmatrix}$$

si et seulement si il existe une base $\{v_1, \ldots, v_n\}$ de V pour laquelle

$$\begin{aligned} T(v_1) &= k_1 v_1 \\ T(v_2) &= \quad k_2 v_2 \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ T(v_n) &= \qquad\quad k_n v_n \end{aligned}$$

ce qui veut dire que les vecteurs $v_1, v_2, \ldots, v_n$ sont des vecteurs propres de T correspondant respectivement aux valeurs propres $k_1, k_2, \ldots, k_n$. En d'autres termes:

Théorème 9.4 : Un opérateur linéaire $T : V \to V$ peut être représenté par une matrice diagonale B si et seulement si V a une base formée des vecteurs propres de T. Dans ce cas les éléments diagonaux de B sont les valeurs propres correspondantes.

Nous obtenons le théorème équivalent suivant.

Autre forme du théorème 9.4 : Une matrice carrée $n \times n$ A est semblable à une matrice diagonale B si et seulement si A a n vecteurs propres linéairement indépendants. Dans ce cas les éléments diagonaux de B sont les valeurs propres correspondantes.

Dans le théorème précédent, si P est la matrice dont les colonnes sont les n vecteurs propres indépendants de A, alors $B = P^{-1}AP$.

Exemple 9.7 : Considérons la matrice $A = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$. D'après l'exemple 9.5. A a deux vecteurs propres indépendants $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$ et $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Soit $P = \begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix}$ et donc $P^{-1} = \begin{pmatrix} 1/5 & 1/5 \\ 3/5 & -2/5 \end{pmatrix}$.

Alors A est semblable à la matrice diagonale

$$B = P^{-1}AP = \begin{pmatrix} 1/5 & 1/5 \\ 3/5 & -2/5 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 0 & -1 \end{pmatrix}$$

Comme on devait le prévoir, les éléments diagonaux 4 et -1 de la matrice diagonale B sont les valeurs propres correspondant aux vecteurs propres donnés.

POLYNOME CARACTERISTIQUE. THEOREME DE CAYLEY-HAMILTON

Considérons une matrice $n \times n$ A sur un corps K:

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}$$

La matrice $tI_n - A$, où I_n est la matrice carrée $n \times n$ identité et t une inconnue, est appelée la matrice caractéristique de A:

$$tI_n - A = \begin{pmatrix} t - a_{11} & -a_{12} & \ldots & -a_{1n} \\ -a_{21} & t - a_{22} & \ldots & -a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ -a_{n1} & -a_{n2} & \ldots & t - a_{nn} \end{pmatrix}$$

Son déterminant $\quad \Delta_A(t) = \det(tI_n - A)$

qui est un polynôme en t, est appelé polynôme caractéristique de A, nous appelons aussi

$$\Delta_A(t) = \det(tI_n - A) = 0$$

l'équation caractéristique de A.

Chaque terme dans le déterminant contient un et seulement un élément de chaque ligne et de chaque colonne ; donc le polynôme caractéristique précédent est de la forme

$$\Delta_A(t) = (t - a_{11})(t - a_{22}) \ldots (t - a_{nn})$$

$$+ \text{ des termes avec plus de } (n-2) \text{ facteurs de la forme } t - a_{ii}$$

En conséquence

$$\Delta_A(t) = t^n - (a_{11} + a_{22} + \cdots + a_{nn})\, t^{n-1} + \text{termes de degré inférieur.}$$

Rappelons que la trace de A est la somme de ses éléments diagonaux. Ainsi le polynôme caractéristique $\Delta_A(t) = \det(tI_n - A)$ de A est un polynôme normé de degré n, et le coefficient de t^{n-1} est l'opposé de la trace de A. (Un polynôme est normé si son coefficient de terme de plus haut degré est 1).

De plus, si nous posons $t = 0$ dans $\Delta_A(t)$, nous obtenons

$$\Delta_A(0) = |-A| = (-1)^n |A|$$

Mais $\Delta_A(0)$ est le terme constant du polynôme $\Delta_A(t)$. Ainsi le terme constant du polynôme caractéristique de la matrice A est $(-1)^n\,|A|$ où n est l'ordre de A.

Exemple 9.8 : Le polynôme caractéristique de la matrice $A = \begin{pmatrix} 1 & 3 & 0 \\ -2 & 2 & -1 \\ 4 & 0 & -2 \end{pmatrix}$ est

$$\Delta(t) = |tI - A| = \begin{vmatrix} t-1 & -3 & 0 \\ 2 & t-2 & 1 \\ -4 & 0 & t+2 \end{vmatrix} = t^3 - t^2 + 2t + 28$$

Comme prévu, $\Delta(t)$ est un polynôme normé de degré 3.

Nous allons maintenant énoncer l'un des théorèmes les plus importants de l'algèbre linéaire.

Théorème de Cayley-Hamilton 9.5 : Chaque matrice est un zéro de son polynôme caractéristique.

Exemple 9.9 : Le polynôme caractéristique de la matrice $A = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$ est

$$\Delta(t) = |tI - A| = \begin{vmatrix} t-1 & -2 \\ -3 & t-2 \end{vmatrix} = t^2 - 3t - 4$$

Comme prévu, d'après le théorème de Cayley-Hamilton, A est un zéro de $\Delta(t)$:

$$\Delta(A) = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}^2 - 3\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix} - 4\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Le théorème suivant montre la relation principale entre les polynômes caractéristiques et les valeurs propres.

Théorème 9.6 : Soit A une matrice $n \times n$ sur un corps K. Un scalaire $\lambda \in K$ est une valeur propre de A si et seulement si λ est une racine du polynôme caractéristique $\Delta(t)$ de A.

Démonstration. D'après le théorème 9.2, λ est une valeur propre de A si et seulement si $\lambda I - A$ est singulière. De plus, d'après le théorème 8.4, $\lambda I - A$ est singulière si et seulement si $|\lambda I - A| = 0$, c'est-à-dire λ est une racine de $\Delta(t)$. Ainsi le théorème est démontré.

En utilisant les théorèmes 9.3, 9.4 et 9.6 on obtient

Corollaire 9.7 : Si le polynôme caractéristique $\Delta(t)$ d'une matrice carrée $n \times n$ A est un produit de facteurs linéaires distincts.

$$\Delta(t) = (t-a_1)(t-a_2)\cdots(t-a_n)$$

c'est-à-dire, si $a_1, a_2, \ldots, a_n$ sont des racines distinctes de $\Delta(t)$, alors A est semblable à la matrice diagonale dont les éléments diagonaux sont les a_i.

De plus, en utilisant le théorème fondamental de l'algèbre (chaque polynôme de **C** admet une racine) et le théorème précédent, on obtient :

Corollaire 9.8 : Soit A une matrice carrée $n \times n$ sur le corps des complexes **C**. Alors A a au moins une valeur propre

Exemple 9.10 : Soit $A = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & -5 \\ 0 & 1 & -2 \end{pmatrix}$. Son polynôme caractéristique est

$$\Delta(t) = \begin{vmatrix} t-3 & 0 & 0 \\ 0 & t-2 & 5 \\ 0 & -1 & t+2 \end{vmatrix} = (t-3)(t^2+1)$$

Considérons deux cas :

1) A est une matrice sur le corps des réels **R**. Alors A admet seulement une valeur propre 3. Puisque 3 a seulement un vecteur propre indépendant, A n'est pas diagonalisable.

2) A est une matrice sur le corps des complexes **C**. Alors A admet trois valeurs propres distinctes : 3, i et $-i$. Il existe alors une matrice inversible P sur le corps des complexes **C** pour laquelle

$$P^{-1}AP = \begin{pmatrix} 3 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & -i \end{pmatrix}$$

c'est-à-dire que A est diagonalisable.

Supposons maintenant que A et B soient des matrices semblables, c'est-à-dire $B = P^{-1}AP$ où P est inversible. Montrons que A et B ont le même polynôme caractéristique. Utilisons $tI = P^{-1}\,tIP$,

$$\begin{aligned} |tI - B| &= |tI - P^{-1}AP| = |P^{-1}tIP - P^{-1}AP| \\ &= |P^{-1}(tI-A)P| = |P^{-1}|\,|tI-A|\,|P| \end{aligned}$$

Puisque les déterminants sont des scalaires et commutent et puisque $|P^{-1}|\,|P| = 1$, nous obtenons finalement

$$|tI - B| = |tI - A|$$

Nous avons ainsi démontré

Théorème 9.9 : Des matrices semblables ont le même polynôme caractéristique.

POLYNOME MINIMAL

Soit A une matrice carrée $n \times n$ sur un corps K. Remarquons qu'il y a des polynômes non nuls $f(t)$ pour lesquels $f(A) = 0$; par exemple le polynôme caractéristique de A. Parmi ces polynômes, nous considérons ceux de plus bas degré et parmi ceux-ci nous choisissons celui qui admet pour premier coefficient 1 ; c'est-à-dire qui est normé. Un tel polynôme $m(t)$ existe et est unique (Problème 9.25) ; nous l'appelons le polynôme minimal de A.

Théorème 9.10 : Le polynôme minimal $m(t)$ de A divise tout polynôme qui admet A comme zéro. En particulier, $m(t)$ divise le polynôme caractéristique $\Delta(t)$ de A.

Il s'agit là d'une très importante relation entre $m(t)$ et $\Delta(t)$.

Théorème 9.11 : Les polynômes caractéristique et minimal d'une matrice A ont les mêmes facteurs irréductibles.

Ce théorème n'implique pas que $m(t) = \Delta(t)$; il implique seulement qu'un facteur irréductible quelconque de l'un doit diviser l'autre. En particulier puisqu'un facteur linéaire est irréductible, $m(t)$ et $\Delta(t)$ ont les mêmes facteurs linéaires ; donc ils ont les mêmes racines. Ainsi d'après le théorème 9.6 on obtient

Théorème 9.12 : Un scalaire λ est une valeur propre pour une matrice A si et seulement si λ est une racine du polynôme minimal de A.

Exemple 9.11 : Trouver le polynôme minimal $m(t)$ de la matrice $A = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix}$.

Le polynôme caractéristique de A est $\Delta(t) = |tI - A| = (t-2)^3(t-5)$. D'après le théorème 9.11 à la fois $t-2$ et $t-5$ doivent être des facteurs de $m(t)$. Mais d'après le théorème 9.10 $m(t)$ doit diviser $\Delta(t)$; donc $m(t)$ doit être l'un des trois polynômes suivants :

$$m_1(t) = (t-2)(t-5), \qquad m_2(t) = (t-2)^2(t-5), \qquad m_3(t) = (t-2)^3(t-5)$$

Nous savons d'après le théorème de Cayley-Hamilton que $m_3(A) = \Delta(A) = 0$. Le lecteur peut vérifier que $m_1(A) \neq 0$ mais $m_2(A) = 0$. En conséquence $m_2(t) = (t-2)^2(t-5)$ est le polynôme minimal de A.

Exemple 9.12 : Soit A une matrice 3×3 sur le corps des réels **R**. Montrons que A ne peut être un zéro du polynôme $f(t) = t^2 + 1$. D'après le théorème de Cayley-Hamilton, A est un zéro de son polynôme caractéristique $\Delta(t)$. Remarquons que $\Delta(t)$ est de degré 3 ; donc il a au moins une racine réelle.

Supposons maintenant que A soit un zéro de $f(t)$. Puisque $f(t)$ est irréductible sur **R**, $f(t)$ doit être le polynôme minimal de A. Mais $f(t)$ n'a pas de racine réelle. Ceci est en contradiction avec le fait que les polynômes caractéristique et minimal ont les mêmes racines. Donc A n'est pas un zéro de $f(t)$.

Le lecteur peut vérifier que la matrice 3×3 sur le corps des complexes **C** est un zéro de $f(t)$:

$$\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & i \end{pmatrix}$$

POLYNOMES CARACTERISTIQUE ET MINIMAL D'OPERATEURS LINEAIRES

Supposons maintenant $T : V \to V$ un opérateur linéaire sur un espace vectoriel V de dimension finie. Nous définissons le polynôme caractéristique $\Delta(t)$ de T comme étant le polynôme caractéristique d'une quelconque représentation matricielle de T. D'après le théorème 9.9, $\Delta(t)$ est indépendant de la base particulière dans laquelle la représentation matricielle est calculée. Remarquons que le degré de $\Delta(t)$ est égal à la dimension de V. Nous avons les théorèmes suivants pour T qui sont semblables à ceux appliqués sur les matrices.

Théorème 9.5' : T est un zéro de son polynôme caractéristique.

Théorème 9.6' : Le scalaire $\lambda \in K$ est une valeur propre de T si et seulement si λ est une racine du polynôme caractéristique de T.

L'ordre de multiplicité algébrique d'une valeur propre $\lambda \in K$ de T est défini comme étant la multiplicité de λ considéré comme une racine du polynôme caractéristique de T. L'ordre de multiplicité géométrique d'une valeur propre λ est défini comme étant la dimension de son espace propre.

Théorème 9.13 : La multiplicité géométrique d'une valeur propre λ n'est pas supérieure à sa multiplicité algébrique.

Exemple 9.13 : Soit V l'espace vectoriel des fonctions, qui a comme base $\{\sin\theta, \cos\theta\}$ et soit D l'opérateur différentiel sur V. Alors

$$D(\sin\theta) = \cos\theta = 0(\sin\theta) + 1(\cos\theta)$$
$$D(\cos\theta) = -\sin\theta = -1(\sin\theta) + 0(\cos\theta)$$

La matrice A de D dans la base précédente est donc $A = [D] = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Ainsi

$$\det(tI - A) = \begin{vmatrix} t & -1 \\ 1 & t \end{vmatrix} = t^2 + 1$$

et le polynôme caractéristique de D est $\Delta(t) = t^2 + 1$.

D'autre part, le polynôme minimal $m(t)$ de l'opérateur T est défini indépendamment de la théorie des matrices, comme étant le polynôme de plus bas degré et admettant comme coefficient 1 qui est un zéro de T. Cependant pour un polynôme quelconque $f(t)$,

$$f(T) = 0 \quad \text{si et seulement si} \quad f(A) = 0$$

où A est une représentation matricielle quelconque de T. En conséquence T et A ont le même polynôme minimal. Remarquons que tous les théorèmes dans ce chapitre sur le polynôme minimal d'une matrice restent aussi valables pour le polynôme minimal de l'opérateur T.

PROBLEMES RESOLUS

POLYNOMES DE MATRICES ET OPERATEURS LINEAIRES

9.1. Trouver $f(A)$ où $A = \begin{pmatrix} 1 & -2 \\ 4 & 5 \end{pmatrix}$ et $f(t) = t^2 - 3t + 7$.

$$f(A) = A^2 - 3A + 7I = \begin{pmatrix} 1 & -2 \\ 4 & 5 \end{pmatrix}^2 - 3\begin{pmatrix} 1 & -2 \\ 4 & 5 \end{pmatrix} + 7\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -3 & -6 \\ 12 & 9 \end{pmatrix}$$

9.2. Montrer que $A = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}$ est un zéro de $f(t) = t^2 - 4t - 5$.

$$f(A) = A^2 - 4A - 5I = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}^2 - 4\begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix} - 5\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

9.3. Soit V l'espace vectoriel des fonctions qui admettent $\{\sin\theta, \cos\theta\}$ comme une base et soit D l'opérateur différentiel de V. Montrer que D est un zéro de $f(t) = t^2 + 1$.

Appliquons $f(D)$ à chaque vecteur de la base:

$$f(D)(\sin\theta) = (D^2 + I)(\sin\theta) = D^2(\sin\theta) + I(\sin\theta) = -\sin\theta + \sin\theta = 0$$
$$f(D)(\cos\theta) = (D^2 + I)(\cos\theta) = D^2(\cos\theta) + I(\cos\theta) = -\cos\theta + \cos\theta = 0$$

Puisque chaque vecteur de la base est transformé en 0, chaque vecteur $v \in V$ est aussi transformé en 0 par $f(D)$. Ainsi $f(D) = 0$.

Ce résultat est normal puisque, d'après l'exemple 9.13, $f(t)$ est le polynôme caractéristique de D.

9.4. Soit A une représentation matricielle d'un opérateur T. Montrer que $f(A)$ est la représentation matricielle de $f(T)$, pour un polynôme quelconque $f(t)$.

Soit ϕ l'application $T \mapsto A$, c'est-à-dire qui transforme l'opérateur T en sa représentation matricielle A. Il nous reste à démontrer que $\phi(f(T)) = f(A)$. Supposons $f(t) = a_n t^n + \cdots + a_1 t + a_0$. La démonstration se fait par récurrence, sur n, degré de $f(t)$.

Supposons $n = 0$. Rappelons que $\phi(I') = I$ où I' est l'application identique et I la matrice identique. Ainsi,

$$\phi(f(T)) = \phi(a_0 I') = a_0\phi(I') = a_0 I = f(A)$$

et donc le théorème est vrai pour $n = 0$.

Supposons maintenant que le théorème soit vrai pour des polynômes de degré inférieur à n. Alors puisque ϕ est un isomorphisme,

$$\begin{aligned}\phi(f(T)) &= \phi(a_nT^n + a_{n-1}T^{n-1} + \cdots + a_1T + a_0I')\\ &= a_n\,\phi(T)\,\phi(T^{n-1}) + \phi(a_{n-1}T^{n-1} + \cdots + a_1T + a_0I')\\ &= a_nAA^{n-1} + (a_{n-1}A^{n-1} + \cdots + a_1A + a_0I) = f(A)\end{aligned}$$

et le théorème est démontré

9.5. Démontrer le théorème 9.1: Soient f et g des polynômes de K. Soit A une matrice carrée sur le corps K. Alors : 1) $(f+g)(A) = f(A) + g(A)$; 2) $(fg)(A) = f(A)\,g(A)$ et 3) $(kf)(A) = kf(A)$ où $k \in K$.

Supposons $f = a_n t^n + \cdots + a_1 t + a_0$ et $g = b_m t^m + \cdots + b_1 t + b_0$. Alors d'après la définition

$$f(A) = a_nA^n + \cdots + a_1A + a_0I \qquad \text{et} \qquad g(A) = b_mA^m + \cdots + b_1A + b_0I$$

1) Supposons $m \leqslant n$ et soit $b_i = 0$ si $i > m$. Alors

$$f + g = (a_n + b_n)t^n + \cdots + (a_1 + b_1)t + (a_0 + b_0)$$

Donc

$$\begin{aligned}(f+g)(A) &= (a_n + b_n)A^n + \cdots + (a_1 + b_1)A + (a_0 + b_0)I\\ &= a_nA^n + b_nA^n + \cdots + a_1A + b_1A + a_0I + b_0I = f(A) + g(A)\end{aligned}$$

2) Par définition, $fg = c_{n+m}t^{n+m} + \cdots + c_1t + c_0 = \sum_{k=0}^{n+m} c_kt^k$ où $c_k = a_0b_k + a_1b_{k-1} + \cdots + a_kb_0 = \sum_{i=0}^{k} a_ib_{k-i}$. Donc $(fg)(A) = \sum_{k=0}^{n+m} c_kA^k$ et

$$f(A)\,g(A) = \left(\sum_{i=0}^{n} a_iA^i\right)\left(\sum_{j=0}^{m} b_jA^j\right) = \sum_{i=0}^{n}\sum_{j=0}^{m} a_ib_jA^{i+j} = \sum_{k=0}^{n+m} c_kA^k = (fg)(A)$$

3) Par définition, $kf = ka_nt^n + \cdots + ka_1t + ka_0$, et donc

$$(kf)(A) = ka_nA^n + \cdots + ka_1A + ka_0I = k(a_nA^n + \cdots + a_1A + a_0I) = k\,f(A)$$

VALEURS PROPRES ET VECTEURS PROPRES

9.6. Soit λ une valeur propre d'un opérateur $T : V \to V$. Soit V_λ l'ensemble de tous les vecteurs propres de T correspondant à la valeur propre λ (appelé espace propre de λ). Montrer que V_λ est un sous-espace de V.

Supposons $v, w \in V_\lambda$; c'est-à-dire $T(v) = \lambda v$ et $T(w) = \lambda w$. Alors pour des scalaires quelconques $a, b \in K$,

$$T(av + bw) = a\,T(v) + b\,T(w) = a(\lambda v) + b(\lambda w) = \lambda(av + bw)$$

Ainsi $av + bw$ est un vecteur propre correspondant à λ ; c'est-à-dire $av + bw \in V_\lambda$. Donc V_λ est un sous-espace de V.

9.7. Soit $A = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}$. 1) Trouver toutes les valeurs propres de A et leurs vecteurs propres correspondants. 2) Trouver une matrice inversible P telle que $P^{-1}AP$ soit diagonale.

(i) Formons la matrice caractéristique $tI - A$ de A :

$$tI - A = \begin{pmatrix} t & 0 \\ 0 & t \end{pmatrix} - \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} t-1 & -4 \\ -2 & t-3 \end{pmatrix} \tag{1}$$

Le polynôme caractéristique $\Delta(t)$ de A est son déterminant :

$$\Delta(t) = |tI - A| = \begin{vmatrix} t-1 & -4 \\ -2 & t-3 \end{vmatrix} = t^2 - 4t - 5 = (t-5)(t+1)$$

Les racines de $\Delta(t)$ sont 5 et -1 et donc ces nombres sont les valeurs propres de A.

Nous obtenons les vecteurs propres correspondant à la valeur propre 5. Remplaçons d'abord t par 5 dans la matrice caractéristique (*1*) on obtient la matrice $\begin{pmatrix} 4 & -4 \\ -2 & 2 \end{pmatrix}$. Les vecteurs propres correspondant à la valeur 5 forment la solution du système homogène déterminé par la matrice précédente c'est-à-dire

$$\begin{pmatrix} 4 & -4 \\ -2 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{ou} \quad \begin{cases} 4x - 4y = 0 \\ -2x + 2y = 0 \end{cases} \quad \text{ou} \quad x - y = 0$$

(En d'autres termes, les vecteurs propres correspondant à 5 forment le noyau de l'opérateur $tI - A$ pour $t = 5$). Le système précédent admet une seule solution indépendante ; par exemple $x = 1, y = 1$. Donc $v = (1, 1)$ est un vecteur propre qui engendre l'espace propre de 5, c'est-à-dire que chaque vecteur correspondant à la valeur propre 5 est un multiple de v.

Nous obtenons les vecteurs propres correspondants à la valeur propre -1 en remplaçant t par -1 dans (*1*), ce qui donne le système homogène

$$\begin{pmatrix} -2 & -4 \\ -2 & -4 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{ou} \quad \begin{cases} -2x - 4y = 0 \\ -2x - 4y = 0 \end{cases} \quad \text{ou} \quad x + 2y = 0$$

Le système admet une seule solution indépendante ; par exemple $x = 2$, $y = -1$. Ainsi $w = (2, -1)$ est un vecteur propre qui engendre l'espace propre de -1.

(ii) Soit P la matrice dont les colonnes sont les précédents vecteurs propres : $P = \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}$. Alors $B = P^{-1}AP$ est la matrice diagonale dont les éléments diagonaux sont les valeurs propres respectives :

$$B = P^{-1}AP = \begin{pmatrix} 1/3 & 2/3 \\ 1/3 & -1/3 \end{pmatrix}\begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 5 & 0 \\ 0 & -1 \end{pmatrix}$$

(*Remarque* : Ici P est la matrice de passage de la base usuelle de $\mathbf{R}^2$ à la base formée des vecteurs propres $\{v, w\}$. Donc B est la représentation matricielle de l'opérateur A dans la nouvelle base).

9.8. Pour chacune des matrices suivantes, trouver toutes les valeurs propres et une base de chacun des espaces propres :

$$\text{(i)} \quad A = \begin{pmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{pmatrix}, \qquad \text{(ii)} \quad B = \begin{pmatrix} -3 & 1 & -1 \\ -7 & 5 & -1 \\ -6 & 6 & -2 \end{pmatrix}$$

Quelles matrices sont-elles diagonalisables, et pourquoi ?

(i) Formons la matrice caractéristique $tI - A$ et calculons son déterminant, pour obtenir le polynôme caractéristique $\Delta(t)$ de A :

$$\Delta(t) = |tI - A| = \begin{vmatrix} t-1 & 3 & -3 \\ -3 & t+5 & -3 \\ -6 & 6 & t-4 \end{vmatrix} = (t+2)^2(t-4)$$

Les racines de $\Delta(t)$ sont -2 et 4 ; donc ces nombres sont les valeurs propres de A.

Trouvons une base de l'espace propre de la valeur propre -2. Remplaçons t par -2 dans la matrice caractéristique $tI - A$ de manière à obtenir le système homogène

$$\begin{pmatrix} -3 & 3 & -3 \\ -3 & 3 & -3 \\ -6 & 6 & -6 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{ou} \quad \begin{cases} -3x + 3y - 3z = 0 \\ -3x + 3y - 3z = 0 \\ -6x + 6y - 6z = 0 \end{cases} \quad \text{ou} \quad x - y + z = 0$$

Le système admet deux solutions indépendantes par exemple $x = 1$, $y = 1$, $z = 0$ et $x = 1$, $y = 0$, $z = -1$. Ainsi $u = (1, 1, 0)$ et $v = (1, 0, -1)$ sont des vecteurs propres indépendants qui engendrent l'espace propre de -2. C'est-à-dire que u et v forment une base de l'espace propre de -2. Ceci veut dire que chaque vecteur propre correspondant à -2 est une combinaison linéaire de u et v.

Trouvons une base de l'espace propre de la valeur propre 4. Remplaçons t par 4 dans la matrice caractéristique $tI - A$; on obtient alors le système homogène

$$\begin{pmatrix} 3 & 3 & -3 \\ -3 & 9 & -3 \\ -6 & 6 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{ou} \quad \begin{cases} 3x + 3y - 3z = 0 \\ -3x + 9y - 3z = 0 \\ -6x + 6y = 0 \end{cases} \quad \text{ou} \quad \begin{cases} x + y - z = 0 \\ 2y - z = 0 \end{cases}$$

Le système a une seule inconnue libre ; donc n'importe quelle solution non nulle, par exemple $x = 1$, $y = 1$, $z = 2$ engendre son espace solution. Donc $w = (1, 1, 2)$ est un vecteur propre qui engendre et donc forme une base de l'espace propre de 4.

Puisque A a trois vecteurs propres indépendants. A est diagonalisable. En fait, soit P la matrice dont les colonnes sont constituées des trois vecteurs propres indépendants :

$$P = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & -1 & 2 \end{pmatrix}. \quad \text{Alors} \quad P^{-1}AP = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

Comme prévu, les éléments diagonaux de $P^{-1}AP$ sont les valeurs propres de A correspondant aux colonnes de P.

(ii) $$\Delta(t) = |tI - B| = \begin{vmatrix} t+3 & -1 & 1 \\ 7 & t-5 & 1 \\ 6 & -6 & t+2 \end{vmatrix} = (t+2)^2(t-4)$$

Les valeurs propres de B sont donc -2 et 4.

Trouvons une base de l'espace propre de la valeur propre -2. Remplaçons t par -2 dans $tI - B$; on obtient le système homogène

$$\begin{pmatrix} 1 & -1 & 1 \\ 7 & -7 & 1 \\ 6 & -6 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{ou} \quad \begin{cases} x - y + z = 0 \\ 7x - 7y + z = 0 \\ 6x - 6y = 0 \end{cases} \quad \text{ou} \quad \begin{cases} x - y + z = 0 \\ x - y = 0 \end{cases}$$

Le système admet seulement une solution indépendante, par exemple $x = 1$, $y = 1$, $z = 0$. Ainsi $u = (1, 1, 0)$ forme une base de l'espace propre de -2.

Trouvons une base de l'espace propre de la valeur propre 4. Remplaçons t par 4 dans $tI - B$ on obtient le système homogène.

$$\begin{pmatrix} 7 & -1 & 1 \\ 7 & -1 & 1 \\ 6 & -6 & 6 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{ou} \quad \begin{cases} 7x - y + z = 0 \\ 7x - y + z = 0 \\ 6x - 6y + 6z = 0 \end{cases} \quad \text{ou} \quad \begin{cases} 7x - y + z = 0 \\ x = 0 \end{cases}$$

Le système admet une seule solution indépendante, par exemple $x = 0$, $y = 1$, $z = 1$. Ainsi $v = (0, 1, 1)$ forme une base de l'espace propre de 4.

Remarquons que B n'est pas semblable à une matrice diagonale puisque B admet seulement deux vecteurs propres indépendants. De plus, puisque A peut être diagonalisée, alors que B ne peut pas être diagonalisée, A et B ne sont pas des matrices semblables, quoiqu'elles aient le même polynôme caractéristique.

9.9. Soient $A = \begin{pmatrix} 3 & -1 \\ 1 & 1 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & -1 \\ 2 & -1 \end{pmatrix}$. Trouver toutes les valeurs propres et les vecteurs propres correspondants de A et de B considérés comme des matrices (i) sur le corps des réels **R**, (ii) sur le corps des complexes **C**.

(i) $$\Delta_A(t) = |tI - A| = \begin{vmatrix} t-3 & 1 \\ -1 & t-1 \end{vmatrix} = t^2 - 4t + 4 = (t-2)^2$$

Donc 2 est la seule valeur propre. Remplaçons t par 2 dans $tI - A$ on obtient le système homogène

$$\begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{ou} \quad \begin{cases} -x + y = 0 \\ -x + y = 0 \end{cases} \quad \text{ou} \quad x - y = 0$$

Le système admet seulement une solution indépendante, par exemple $x = 1$, $y = 1$. Ainsi $v = (1, 1)$ est un vecteur propre qui engendre l'espace propre de 2, c'est-à-dire chaque vecteur propre de 2 est un multiple (complexe) de v.

Nous avons ainsi $$\Delta_B(t) = |tI - B| = \begin{vmatrix} t-1 & 1 \\ -2 & t+1 \end{vmatrix} = t^2 + 1$$

(ii) Puisque $t^2 + 1$ n'admet pas de racine dans **R**, B n'a pas de valeur propre sur **R**.

Puisque $\Delta_A(t) = (t-2)^2$ a seulement une racine réelle 2, les résultats sont les mêmes que dans (*i*). C'est-à-dire, 2 est une valeur propre de A, et $v = (1, 1)$ est un vecteur propre engendrant l'espace propre de 2, c'est-à-dire chaque vecteur propre de 2 est un multiple (complexe) de v.

La matrice caractéristique de B est $\Delta_B(t) = |tI - B| = t^2 + 1$. Donc i et $-i$ sont les valeurs propres de B.

Nous trouvons les vecteurs propres associés à $t = i$ en substituant $t = i$ dans $tI - B$ on obtient le système homogène

$$\begin{pmatrix} i-1 & 1 \\ -2 & i+1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{ou} \quad \begin{cases} (i-1)x + y = 0 \\ -2x + (i+1)y = 0 \end{cases} \quad \text{ou} \quad (i-1)x + y = 0$$

Le système admet seulement une solution indépendante ; par exemple $x = 1$, $y = 1 - i$. Ainsi $w' = (1, 1 - i)$ est un vecteur propre qui engendre l'espace propre de i.

Remplaçons maintenant t par $-i$ dans $tI - B$, on obtient le système homogène

$$\begin{pmatrix} -i-1 & 1 \\ -2 & -i-1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{ou} \quad \begin{cases} (-i-1)x + y = 0 \\ -2x + (-i-1)y = 0 \end{cases} \quad \text{ou} \quad (-i-1)x + y = 0$$

Le système admet seulement une solution indépendante, par exemple $x = 1$, $y = 1 + i$. Ainsi $w' = (1, 1 + i)$ est un vecteur propre qui engendre l'espace propre de $-i$.

9.10. Trouver toutes les valeurs propres et une base de l'espace propre de l'opérateur $T : \mathbf{R}^3 \to \mathbf{R}^3$ défini par $T(x, y, z) = (2x + y, y - z, 2y + 4z)$.

Trouvons d'abord une représentation matricielle de T, par exemple relativement à la base usuelle de $\mathbf{R}^3$:

$$A = [T] = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 2 & 4 \end{pmatrix}$$

Le polynôme caractéristique $\Delta(t)$ de T est alors

$$\Delta(t) = |tI - A| = \begin{vmatrix} t-2 & -1 & 0 \\ 0 & t-1 & 1 \\ 0 & -2 & t-4 \end{vmatrix} = (t-2)^2(t-3)$$

Ainsi 2 et 3 sont les valeurs propres de T.

Nous trouvons alors une base de l'espace propre de valeur propre 2. En remplaçant t par 2 dans $tI - A$ on obtient le système homogène

$$\begin{pmatrix} 0 & -1 & 0 \\ 0 & 1 & 1 \\ 0 & -2 & -2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{ou} \quad \begin{cases} -y = 0 \\ y + z = 0 \\ -2y - 2z = 0 \end{cases} \quad \text{ou} \quad \begin{cases} y = 0 \\ y + z = 0 \end{cases}$$

Le système admet une seule solution indépendante, par exemple $x = 1$, $y = 0$, $z = 0$. Ainsi $u = (1, 0, 0)$ forme une base de l'espace propre de 2.

Trouvons une base de l'espace propre correspondant à la valeur propre 3. Remplaçons t par 3 dans $tI - A$; on obtient le système homogène.

$$\begin{pmatrix} 1 & -1 & 0 \\ 0 & 2 & 1 \\ 0 & -2 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{ou} \quad \begin{cases} x - y = 0 \\ 2y + z = 0 \\ -2y - z = 0 \end{cases} \quad \text{ou} \quad \begin{cases} x - y = 0 \\ 2y + z = 0 \end{cases}$$

Le système admet une seule solution indépendante, par exemple $x = 1$, $y = 1$, $z = -2$. Donc $v = (1, 1, -2)$ forme une base de l'espace propre de 3.

Remarquons que T n'est pas diagonalisable, puisque T a seulement deux vecteurs propres linéairement indépendants.

9.11. Montrer que 0 est une valeur propre de T si et seulement si T est singulier.

0 est une valeur propre de T si et seulement si il existe un vecteur non nul v tel que $T(v) = 0v = 0$, c'est-à-dire que T est alors singulier.

9.12. Soient A et B deux matrices carrées $n \times n$. Montrer que AB et BA ont les mêmes valeurs propres.

D'après le problème 9.11 et le fait que le produit de matrices non singulières est une matrice non singulière, les proprositions suivantes sont équivalentes : (i) 0 est une valeur propre de AB, (ii) AB est singulière, (iii) A ou B est singulière, (iv) BA est singulière, (v) 0 est une valeur propre de BA.

Supposons maintenant que λ soit une valeur propre non nulle de AB. Alors il existe un vecteur non nul v tel que $ABv = \lambda v$. Posons $w = Bv$. Puisque $\lambda \neq 0$ et $v \neq 0$,

$$Aw = ABv = \lambda v \neq 0 \quad \text{et donc} \quad w \neq 0$$

Mais w est un vecteur propre de BA correspondant à la valeur propre λ puisque

$$BAw = BABv = B\lambda v = \lambda Bv = \lambda w$$

Donc λ est une valeur propre de BA. De manière analogue, une valeur propre de BA est aussi une valeur propre de AB.

Ainsi AB et BA ont les mêmes valeurs propres.

9.13. Supposons que λ soit une valeur propre d'un opérateur inversible T. Montrer que λ^{-1} est une valeur propre de T^{-1}.

Puisque T est inversible, il est aussi non singulier ; donc d'après le Problème 9.11, $\lambda \neq 0$.

D'après la définition d'une valeur propre, il existe un vecteur non nul v pour lequel $T(v) = \lambda v$. Composons les deux membres avec T^{-1}, on obtient $v = T^{-1}(\lambda v) = \lambda T^{-1}(v)$. Donc $T^{-1}(v) = \lambda^{-1} v$; c'est-à-dire λ^{-1} est une valeur propre de T^{-1}.

9.14. Démontrer le théorème 9.3 : Soient $v_1, \ldots, v_n$ les vecteurs propres non nuls d'un opérateur $T : V \to V$ correspondant aux valeurs propres distinctes $\lambda_1, \ldots, \lambda_n$. Alors $v_1, \ldots, v_n$ sont linéairement indépendants.

La démonstration se fait par récurrence sur n. Si $n = 1$, alors v_1 est linéairement indépendant puisque $v_1 \neq 0$. Posons $n > 1$. Supposons

$$a_1 v_1 + a_2 v_2 + \cdots + a_n v_n = 0 \qquad (1)$$

où les a_i sont des scalaires. En appliquant T à la relation précédente, nous obtenons par linéarité

$$a_1 T(v_1) + a_2 T(v_2) + \cdots + a_n T(v_n) = T(0) = 0$$

Mais par hypothèse $T(v_i) = \lambda_i v_i$; donc

$$a_1\lambda_1 v_1 + a_2\lambda_2 v_2 + \cdots + a_n\lambda_n v_n = 0 \qquad (2)$$

D'autre part, multiplions (*1*) par λ_n,

$$a_1\lambda_n v_1 + a_2\lambda_n v_2 + \cdots + a_n\lambda_n v_n = 0 \qquad (3)$$

Soustrayons maintenant (*3*) de (*2*),

$$a_1(\lambda_1 - \lambda_n)v_1 + a_2(\lambda_2 - \lambda_n)v_2 + \cdots + a_{n-1}(\lambda_{n-1} - \lambda_n)v_{n-1} = 0$$

Par récurrence, chacun des coefficients précédents est nul. Puisque les λ_i sont distincts, $\lambda_i - \lambda_n \neq 0$ pour $i \neq n$. Donc $a_1 = \cdots = a_{n-1} = 0$. Remplaçons les $a_1 \cdots a_{n-1}$ par leurs valeurs dans (*1*), on obtient $a_n v_n = 0$ et donc $a_n = 0$. En résumé les v_i sont linéairement indépendants.

POLYNOME CARACTERISTIQUE. THEOREME DE CAYLEY-HAMILTON

9.15. Considérons une matrice triangulaire

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}$$

Trouver son polynôme caractéristique $\Delta(t)$ et ses valeurs propres.

Puisque A est triangulaire et tI diagonale, $tI - A$ est aussi triangulaire avec pour éléments diagonaux $t - a_{ii}$:

$$tI - A = \begin{pmatrix} t-a_{11} & -a_{12} & \dots & -a_{1n} \\ 0 & t-a_{22} & \dots & -a_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & t-a_{nn} \end{pmatrix}$$

Alors $\Delta(t) = |tI - A|$ est le produit des éléments diagonaux $t - a_{ii}$.

$$\Delta(t) = (t-a_{11})(t-a_{22})\cdots(t-a_{nn})$$

Donc les valeurs propres de A sont $a_{11}, a_{22}, \dots, a_{nn}$, c'est-à-dire les éléments diagonaux de A.

9.16. Soit $A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 3 \\ 0 & 0 & 3 \end{pmatrix}$. Est-ce que A est semblable à une matrice diagonale ? S'il en est ainsi trouver cette matrice diagonale.

Puisque A est triangulaire, les valeurs propres de A sont les éléments diagonaux 1, 2 et 3. Puisqu'elles sont distinctes, A est semblable à la matrice diagonale dont les éléments diagonaux sont 1, 2, 3 ; par exemple

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

9.17. Pour chacune des matrices suivantes, trouver un polynôme ayant la matrice donnée comme racine :

(i) $A = \begin{pmatrix} 2 & 5 \\ 1 & -3 \end{pmatrix}$, (ii) $B = \begin{pmatrix} 2 & -3 \\ 7 & -4 \end{pmatrix}$, (iii) $C = \begin{pmatrix} 1 & 4 & -3 \\ 0 & 3 & 1 \\ 0 & 2 & -1 \end{pmatrix}$.

D'après le théorème de Cayley-Hamilton, chaque matrice est une racine de son polynôme caractéristique. Trouvons donc le polynôme caractéristique $\Delta(t)$ dans chacun des cas.

(i) $\Delta(t) = |tI - A| = \begin{vmatrix} t-2 & -5 \\ -1 & t+3 \end{vmatrix} = t^2 + t - 11$

(ii) $\Delta(t) = |tI - B| = \begin{vmatrix} t-2 & 3 \\ -7 & t+4 \end{vmatrix} = t^2 + 2t + 13$

(iii) $\Delta(t) = |tI - C| = \begin{vmatrix} t-1 & -4 & 3 \\ 0 & t-3 & -1 \\ 0 & -2 & t+1 \end{vmatrix} = (t-1)(t^2 - 2t - 5)$

9.18. Démontrer le théorème de Cayley-Hamilton Théorème 9.5 : Toute matrice est un zéro de son polynôme caractéristique.

Soit A une matrice carrée arbitraire $n \times n$ et soit $\Delta(t)$ son polynôme caractéristique; par exemple,

$$\Delta(t) = |tI - A| = t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0$$

Soit $B(t)$ l'adjoint classique de la matrice $tI - A$. Les éléments de $B(t)$ sont les cofacteurs de la matrice $tI - A$ et donc sont des polynômes en t de degré inférieur ou égal à $n - 1$. Ainsi

$$B(t) = B_{n-1}t^{n-1} + \cdots + B_1 t + B_0$$

où les B_i sont des matrices carrées $n \times n$ sur K qui sont indépendantes de t. D'après la propriété fondamentale de l'adjoint classique (Théorème 8.8),

$$(tI - A)B(t) = |tI - A|I$$

$$(tI - A)(B_{n-1}t^{n-1} + \cdots + B_1 t + B_0) = (t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0)I$$

En supprimant les parenthèses et en égalant les coefficients des puissances correspondantes de t,

$$\begin{aligned}
B_{n-1} &= I \\
B_{n-2} - AB_{n-1} &= a_{n-1}I \\
B_{n-3} - AB_{n-2} &= a_{n-2}I \\
\cdots\cdots\cdots&\cdots\cdots \\
B_0 - AB_1 &= a_1 I \\
-AB_0 &= a_0 I
\end{aligned}$$

En multipliant les équations matricielles par $A^n, A^{n-1}, \ldots, A, I$ respectivement on a

$$\begin{aligned}
A^n B_{n-1} &= A^n \\
A^{n-1}B_{n-2} - A^n B_{n-1} &= a_{n-1}A^{n-1} \\
A^{n-2}B_{n-3} - A^{n-1}B_{n-2} &= a_{n-2}A^{n-2} \\
\cdots\cdots\cdots\cdots&\cdots\cdots \\
AB_0 - A^2 B_1 &= a_1 A \\
-AB_0 &= a_0 I
\end{aligned}$$

En additionnant les équations matricielles précédentes,

$$0 = A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I$$

En d'autres termes, $\Delta(A) = 0$. C'est-à-dire que A est un zéro de son polynôme caractéristique.

9.19. Montrer qu'une matrice A et sa transposée A^t ont le même polynôme caractéristique.

D'après l'opération de transposition $(tI - A)^t = tI^t - A^t = tI - A^t$. Puisqu'une matrice et sa transposée ont le même déterminant $|tI - A| = |(tI - A^t)| = |tI - A^t|$. Donc A et A^t ont le même polynôme caractéristique.

9.20. Supposons $M = \begin{pmatrix} A_1 & B \\ 0 & A_2 \end{pmatrix}$, où A_1 et A_2 sont des matrices carrées. Montrer que le polynôme caractéristique de M est le produit des polynômes caractéristiques de A_1 et de A_2. Généraliser

$tI - M = \begin{pmatrix} tI - A_1 & -B \\ 0 & tI - A_2 \end{pmatrix}$. Donc d'après le problème 8.70 $|tI - M| = \begin{vmatrix} tI - A_1 & -B \\ 0 & tI - A_2 \end{vmatrix} =$ $|tI - A|\ |tI - B|$, comme il était demandé.

Par récurrence, le polynôme caractéristique de la matrice triangulaire bloc

$$M = \begin{pmatrix} A_1 & B & \dots & C \\ 0 & A_2 & \dots & D \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & A_n \end{pmatrix}$$

où les A_i sont des matrices carrées, est le produit des polynômes caractéristiques des divers A_i.

POLYNOME MINIMAL

9.21. Trouver le polynôme minimal $m(t)$ de $A = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -2 & 4 \end{pmatrix}$.

Le polynôme caractéristique de A est :

$$\Delta(t) = \begin{vmatrix} t-2 & -1 & 0 & 0 \\ 0 & t-2 & 0 & 0 \\ 0 & 0 & t-1 & -1 \\ 0 & 0 & 2 & t-4 \end{vmatrix} = \begin{vmatrix} t-2 & -1 \\ 0 & t-2 \end{vmatrix} \begin{vmatrix} t-1 & -1 \\ 2 & t-4 \end{vmatrix} = (t-3)(t-2)^3$$

Le polynôme minimal $m(t)$ doit diviser $\Delta(t)$. Donc, chaque facteur irréductible de $\Delta(t)$ par exemple $t-2$ et $t-3$, doit être un facteur de $m(t)$. Ainsi $m(t)$ doit être l'un des polynômes suivants :

$$f(t) = (t-3)(t-2), \quad g(t) = (t-3)(t-2)^2, \quad h(t) = (t-3)(t-2)^3$$

Nous avons

$$f(A) = (A-3I)(A-2I) = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & -2 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & -2 & 2 \end{pmatrix} \neq 0$$

$$g(A) = (A-3I)(A-2I)^2 = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & -2 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & -2 & 2 \end{pmatrix}^2 = 0$$

Ainsi $g(t) = (t-3)(t-2)^2$ est le polynôme minimal de A.

Remarque : Nous savons que $h(A) = \Delta(A) = 0$ d'après le théorème de Cayley-Hamilton. Cependant, le degré de $g(t)$ est inférieur au degré de $h(t)$; donc $g(t)$ et non $h(t)$ est le polynôme minimal de A.

9.22. Trouver le polynôme minimal $m(t)$ de chaque matrice (où $a \neq 0$):

(i) $A = \begin{pmatrix} \lambda & a \\ 0 & \lambda \end{pmatrix}$, (ii) $B = \begin{pmatrix} \lambda & a & 0 \\ 0 & \lambda & a \\ 0 & 0 & \lambda \end{pmatrix}$, (iii) $C = \begin{pmatrix} \lambda & a & 0 & 0 \\ 0 & \lambda & a & 0 \\ 0 & 0 & \lambda & a \\ 0 & 0 & 0 & \lambda \end{pmatrix}$.

(i) Le polynôme caractéristique de A est $\Delta(t) = (t-\lambda)^2$. Nous trouvons alors $A - \lambda I \neq 0$; donc $m(t) = \Delta(t) = (t-\lambda)^2$.

(ii) Le polynôme caractéristique de B est $\Delta(t) = (t-\lambda)^3$. (Remarquons que $m(t)$ doit être l'un des polynômes $t-\lambda$, $(t-\lambda)^2$ ou $(t-\lambda)^3$). On a $(B-\lambda I)^2 \neq 0$; ainsi $m(t) = \Delta(t) = (t-\lambda)^3$.

(iii) Le polynôme caractéristique de C est $\Delta(t) = (t-\lambda)^4$. Or on trouve $(C-\lambda I)^3 \neq 0$ donc $m(t) = \Delta(t) = (t-\lambda)^4$.

9.23. Soit $M = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$ où A et B sont des matrices carrées. Montrer que le polynôme minimal $m(t)$ de M est le plus petit commun multiple des polynômes minimaux $g(t)$ et $h(t)$ de A et B respectivement. Généraliser.

Puisque $m(t)$ est le polynôme minimal de M, $m(M) = \begin{pmatrix} m(A) & 0 \\ 0 & m(B) \end{pmatrix} = 0$ et donc $m(A) = 0$ et $m(B) = 0$. Puisque $g(t)$ est le polynôme minimal de A, $g(t)$ divise $m(t)$. De façon analogue $h(t)$ divise $m(t)$. Ainsi $m(t)$ est un multiple de $g(t)$ et $h(t)$.

Soit maintenant $f(t)$ un autre multiple de $g(t)$ et $h(t)$; alors $f(M) = \begin{pmatrix} f(A) & 0 \\ 0 & f(B) \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0.$ Mais $m(t)$ est le polynôme minimal de M ; donc $m(t)$ divise $f(t)$. Ainsi $m(t)$ est le plus petit commun multiple de $g(t)$ et $h(t)$.

Nous pouvons alors dire par récurrence, que le polynôme minimal de

$$M = \begin{pmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & A_n \end{pmatrix}$$

où les A_i sont des matrices carrées, est le plus petit commun multiple des polynômes minimaux des A_i.

9.24. Trouver le polynôme minimal $m(t)$ de

$$M = \begin{pmatrix} 2 & 8 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 5 \end{pmatrix}$$

Posons $A = \begin{pmatrix} 2 & 8 \\ 0 & 2 \end{pmatrix}$, $B = \begin{pmatrix} 4 & 2 \\ 1 & 3 \end{pmatrix}$, $C = \begin{pmatrix} 0 & 3 \\ 0 & 0 \end{pmatrix}$, $D = (5)$. Les polynômes minimaux de A, C et D sont $(t-2)^2$, t^2 et $t-5$ respectivement. Le polynôme caractéristique de B est

$$|tI - B| = \begin{vmatrix} t-4 & -2 \\ -1 & t-3 \end{vmatrix} = t^2 - 7t + 10 = (t-2)(t-5)$$

et donc, il est aussi le polynôme minimal de B.

Remarquons que $M = \begin{pmatrix} A & 0 & 0 & 0 \\ 0 & B & 0 & 0 \\ 0 & 0 & C & 0 \\ 0 & 0 & 0 & D \end{pmatrix}$. Ainsi $m(t)$ est le plus petit commun multiple des polynômes minimaux de A, B, C et D. En conséquence $m(t) = t^2(t-2)^2(t-5)$.

9.25. Montrer que le polynôme minimal d'une matrice (opérateur) A existe et est unique.

D'après le théorème de Cayley-Hamilton, A est un zéro d'un polynôme non nul (problème 9.31). Soit n le plus bas degré pour lequel un polynôme $f(t)$ existe et tel que $f(A) = 0$. Divisons $f(t)$ par le coefficient de son terme de plus haut degré, nous obtenons un polynôme normalisé $m(t)$ de degré n qui admet A comme zéro. Supposons que $m'(t)$ soit un autre polynôme normalisé de degré n pour lequel $m'(A) = 0$. Alors la différence $m(t) - m'(t)$ est un polynôme non nul de degré inférieur à n qui admet A comme zéro. Ceci est en contradiction avec le degré n du polynôme minimal ; donc $m(t)$ est un polynôme minimal unique.

9.26. Démontrer le théorème 9.10 : Le polynôme minimal $m(t)$ d'une matrice (opérateur) A divise tout polynôme admettant A comme zéro. En particulier $m(t)$ divise le polynôme caractéristique de A.

Supposons que $f(t)$ soit un polynôme pour lequel $f(A) = 0$. D'après l'algorithme de la division il existe des polynômes $q(t)$ et $r(t)$ pour lesquels $f(t) = m(t)\,q(t) + r(t)$ et $r(t) = 0$ ou $\deg r(t) < \deg m(t)$. Remplaçons t par A dans cette équation et utilisons le fait que $f(A) = 0$ et $m(A) = 0$, on obtient $r(A) = 0$. Si $r(t) \neq 0$, alors $r(t)$ est un polynôme de degré inférieur à $m(t)$ qui admet A comme zéro ; ce qui est en contradiction avec la définition du polynôme minimal. Donc $r(t) = 0$ et aussi $f(t) = m(t)\,q(t)$, c'est-à-dire $m(t)$ divise $f(t)$.

9.27. Soit $m(t)$ le polynôme minimal d'une matrice carrée A ; $n \times n$. Montrer que le polynôme caractéristique de A divise $(m(t))^n$.

Supposons $m(t) = t^r + c_1 t^{r-1} + \cdots + c_{r-1}t + c_r$. Considérons les matrices suivantes :

$$\begin{aligned} B_0 &= I \\ B_1 &= A + c_1 I \\ B_2 &= A^2 + c_1 A + c_2 I \\ &\cdots\cdots\cdots\cdots \\ B_{r-1} &= A^{r-1} + c_1 A^{r-2} + \cdots + c_{r-1} I \end{aligned}$$

Alors

$$\begin{aligned} B_0 &= I \\ B_1 - AB_0 &= c_1 I \\ B_2 - AB_1 &= c_2 I \\ &\cdots\cdots\cdots\cdots \\ B_{r-1} - AB_{r-2} &= c_{r-1} I \end{aligned}$$

D'où

$$\begin{aligned} -AB_{r-1} &= c_r I - (A^r + c_1 A^{r-1} + \cdots + c_{r-1}A + c_r I) \\ &= c_r I - m(A) \\ &= c_r I \end{aligned}$$

Posons

$$B(t) = t^{r-1}B_0 + t^{r-2}B_1 + \cdots + tB_{r-2} + B_{r-1}$$

Alors

$$\begin{aligned} (tI - A)\cdot B(t) &= (t^r B_0 + t^{r-1}B_1 + \cdots + tB_{r-1}) - (t^{r-1}AB_0 + t^{r-2}AB_1 + \cdots + AB_{r-1}) \\ &= t^r B_0 + t^{r-1}(B_1 - AB_0) + t^{r-2}(B_2 - AB_1) + \cdots + t(B_{r-1} - AB_{r-2}) - AB_{r-1} \\ &= t^r I + c_1 t^{r-1} I + c_2 t^{r-2} I + \cdots + c_{r-1} t I + c_r I \\ &= m(t) I \end{aligned}$$

Le déterminant des deux membres donne $|tI - A|\ |B(t)| = |m(t)I| = (m(t))^n$. Puisque $|B(t)|$ est un polynôme, $|tI - A|$ divise $(m(t))^n$; c'est-à-dire, le polynôme caractéristique de A divise $(m(t))^n$.

9.28. Démontrer le théorème 9.11: Le polynôme caractéristique $\Delta(t)$ et le polynôme minimal $m(t)$ d'une matrice A ont les mêmes facteurs irréductibles.

Supposons que $f(t)$ soit un polynôme irréductible. Si $f(t)$ divise $m(t)$ alors, puisque $m(t)$ divise $\Delta(t)$, $f(t)$ divise $\Delta(t)$. D'autre part, si $f(t)$ divise $\Delta(t)$ alors, d'après le problème précédent, $f(t)$ divise $(m(t))^n$. Mais $f(t)$ est irréductible ; donc $f(t)$ divise aussi $m(t)$. Ainsi $m(t)$ et $\Delta(t)$ ont les mêmes facteurs irréductibles.

9.29. Soit T un opérateur linéaire sur un espace vectoriel V de dimension finie. Montrer que T est inversible si et seulement si le terme constant du polynôme minimal (caractéristique) de T est non nul.

Supposons que le polynôme minimal (caractéristique) de T soit $f(t) = t^r + a_{n-1}t^{r-1} + \cdots + a_1 t + a_0$. Chacun des résultats suivants est équivalent au suivant d'après les résultats précédents. (i) T est inversible. (ii) T est non singulière. (iii) 0 n'est pas une valeur propre de T. (iv) 0 n'est pas une racine de $m(t)$. (v) le terme constant a_0 n'est pas nul. Le théorème est donc démontré.

9.30. Supposons dim $V = n$. Soit $T : V \to V$ un opérateur inversible. Montrer que T^{-1} est égal à un polynôme en T de degré n'excédant pas n.

Soit $m(t)$ le polynôme minimal de T. Alors $m(t) = t^r + a_{r-1}\,t^{r-1} + \dots + a_1 t + a_0$, où $r \leqslant n$. Puisque T est inversible, $a_0 \neq 0$, nous avons :

$$m(T) = T^r + a_{r-1}T^{r-1} + \cdots + a_1T + a_0I = 0$$

Donc

$$-\frac{1}{a_0}(T^{r-1} + a_{r-1}T^{r-2} + \cdots + a_1I)T = I \qquad \text{et} \qquad T^{-1} = -\frac{1}{a_0}(T^{r-1} + a_{r-1}T^{r-2} + \cdots + a_1I)$$

PROBLEMES DIVERS

9.31. Soit T un opérateur linéaire sur un espace vectoriel V de dimension n. Sans utiliser le théorème de Cayley-Hamilton, montrer que T est un zéro d'un polynôme non nul.

Soit $N = n^2$. Considérons les $N + 1$ opérateurs suivants de V ; $I, T, T^2, \dots, T^N$. Rappelons que l'espace vectoriel $A(V)$ des opérateurs de V a pour dimension $N = n^2$. Ainsi les $N + 1$ opérateurs sont linéairement dépendants. Donc il existe des scalaires $a_0, a_1, \dots, a_N$ pour lesquels $a_N T^N + \dots + a_1 T + a_0 I = 0$. En conséquence, T est un zéro du polynôme $f(t) = a_N t^N + \dots + a_1 t + a_0$.

9.32. Démontrer le théorème 9.13 : Soit λ une valeur propre d'un opérateur $T : V \to V$. L'ordre de multiplicité géométrique de λ n'excède pas son ordre de multiplicité algébrique.

Supposons que l'ordre de multiplicité géométrique de λ soit r. Alors λ contient r vecteurs propres linéairement indépendants $v_1, \dots, v_r$. Etendons l'ensemble $\{v_i\}$ à une base de V : $\{v_1, \dots, v_r\ ;\ w_1, \dots, w_s\}$. Nous avons

$$\begin{aligned}
T(v_1) &= \lambda v_1 \\
T(v_2) &= \qquad \lambda v_2 \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
T(v_r) &= \qquad\qquad \lambda v_r \\
T(w_1) &= a_{11}v_1 + \cdots + a_{1r}v_r + b_{11}w_1 + \cdots + b_{1s}w_s \\
T(w_2) &= a_{21}v_1 + \cdots + a_{2r}v_r + b_{21}w_1 + \cdots + b_{2s}w_s \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
T(w_s) &= a_{s1}v_1 + \cdots + a_{sr}v_r + b_{s1}w_1 + \cdots + b_{ss}w_s
\end{aligned}$$

La matrice de T dans la base précédente est

$$M = \left(\begin{array}{cccc|cccc}
\lambda & 0 & \dots & 0 & a_{11} & a_{21} & \dots & a_{s1} \\
0 & \lambda & \dots & 0 & a_{12} & a_{22} & \dots & a_{s2} \\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\
0 & 0 & \dots & \lambda & a_{1r} & a_{2r} & \dots & a_{sr} \\
\hline
0 & 0 & \dots & 0 & b_{11} & b_{21} & \dots & b_{r1} \\
0 & 0 & \dots & 0 & b_{12} & b_{22} & \dots & b_{r2} \\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\
0 & 0 & \dots & 0 & b_{1s} & b_{2s} & \dots & b_{ss}
\end{array}\right) = \left(\begin{array}{c|c} \lambda I_r & A \\ \hline 0 & B \end{array}\right)$$

où $A = (a_{ij})^t$ et $B = (b_{ij})^t$.

D'après le problème 9.20, le polynôme caractéristique de λI_r, qui est $(t - \lambda)^r$, doit diviser le polynôme caractéristique de M et donc T. Donc l'ordre de multiplicité algébrique de λ pour l'opérateur T est au moins r, comme demandé.

9.33. Montrer que $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ n'est pas diagonalisable.

Le polynôme caractéristique de A est $\Delta(t) = (t - 1)^2$; donc 1 est la seule valeur propre de A. Nous trouvons donc une base de l'espace propre de valeur propre 1. Remplaçons t par 1 dans la matrice $tI - A$; on obtient le système homogène.

$$\begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{ou} \quad \begin{cases} -y = 0 \\ \ \ 0 = 0 \end{cases} \quad \text{ou} \quad y = 0$$

Le système a seulement une solution indépendante, par exemple $x = 1$, $y = 0$, Donc $u = (1, 0)$ forme une base de l'espace propre de 1.

Puisque A a au plus un vecteur propre indépendant, A ne peut être diagonalisée.

9.34. Soit F une extension du corps K. Soit A une matrice carrée $n \times n$ sur K. Remarquons que A peut être aussi considérée comme une matrice $\hat{A}$ sur F, C'est-à-dire $|tI - A| = |tI - \hat{A}|$, d'où A et $\hat{A}$ ont le même polynôme caractéristique. Montrer que A et $\hat{A}$ ont le même polynôme minimal.

Soient $m(t)$ et $m'(t)$ les polynômes minimaux de A et $\hat{A}$ respectivement. $m'(t)$ divise tout polynôme de F qui admet A comme zéro. Puisque $m(t)$ admet A comme zéro et puisque $m(t)$ peut être considéré comme un polynôme sur F, $m'(t)$ divise $m(t)$. Montrons maintenant que $m(t)$ divise $m'(t)$.

Puisque $m'(t)$ est un polynôme sur F, qui est une extension de K, nous pouvons écrire

$$m'(t) = f_1(t)\, b_1 + f_2(t)\, b_2 + \cdots + f_n(t)\, b_n$$

où $f_i(t)$ sont des polynômes sur K, et $b_1, \ldots, b_n$ appartiennent à F et linéairement indépendants sur K. Nous avons

$$m'(A) = f_1(A)\, b_1 + f_2(A)\, b_2 + \cdots + f_n(A)\, b_n = 0 \tag{1}$$

Soit $a_{ij}^{(k)}$ le ij élément de $f_k(A)$. L'équation matricielle précédente implique que, pour chaque paire (i, j),

$$a_{ij}^{(1)} b_1 + a_{ij}^{(2)} b_2 + \cdots + a_{ij}^{(n)} b_n = 0$$

Puisque les b_i sont linéairement indépendants sur K et puisque les $a_{ij}^{(k)} \in K$, chaque $a_{ij}^{(k)} = 0$. D'où

$$f_1(A) = 0,\ f_2(A) = 0,\ \ldots,\ f_n(A) = 0$$

Puisque les $f_i(t)$ sont des polynômes sur K qui admettent A comme zéro, et puisque $m(t)$ est le polynôme minimal de A considérée comme une matrice sur K, $m(t)$ divise chacun des $f_i(t)$. En conséquence, d'après (1), $m(t)$ doit aussi diviser $m'(t)$. Mais les polynômes normalisés qui se divisent les uns les autres sont nécessairement égaux. Donc $m(t) = m'(t)$.

9.35. Soit $\{v_1, \ldots, v_n\}$ une base de V. Soit $T : V \to V$ un opérateur pour lequel $T(v_1) = 0$, $T(v_2) = a_{21} v_1$, $T(v_3) = a_{31} v_1 + a_{32} v_2, \ldots, T(v_n) = a_{n1} v_1 + \cdots + a_{n,n-1} v_{n-1}$. Montrer que $T^n = 0$.

Il suffit de montrer que

$$T^j(v_j) = 0 \tag{*}$$

pour $j = 1, \ldots, n$. Il s'ensuit alors que

$$T^n(v_j) = T^{n-j}(T^j(v_j)) = T^{n-j}(0) = 0, \quad \text{pour} \quad j = 1, \ldots, n$$

et puisque $\{v_1, \ldots, v_n\}$ est une base $T^n = 0$.

Nous démontrons (*) par récurrence sur j. Le cas $j = 1$ est vraie par hypothèse. Il s'ensuit par récurrence (pour $j = 2, \ldots, n$)

$$\begin{aligned} T^j(v_j) &= T^{j-1}(T(v_j)) = T^{j-1}(a_{j1} v_1 + \cdots + a_{j,j-1} v_{j-1}) \\ &= a_{j1} T^{j-1}(v_1) + \cdots + a_{j,j-1} T^{j-1}(v_{j-1}) \\ &= a_{j1} 0 + \cdots + a_{j,j-1} 0 = 0 \end{aligned}$$

Remarque : Observons que la représentation matricielle de T dans la base précédente est triangulaire avec des éléments diagonaux nuls:

$$\begin{pmatrix} 0 & a_{21} & a_{31} & \ldots & a_{n1} \\ 0 & 0 & a_{32} & \ldots & a_{n2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \ldots & a_{n,n-1} \\ 0 & 0 & 0 & \ldots & 0 \end{pmatrix}$$

PROBLEMES SUPPLEMENTAIRES

POLYNOMES DE MATRICES ET OPERATEURS LINEAIRES

9.36. Soit $f(t) = 2t^2 - 5t + 6$ et $g(t) = t^3 - 2t^2 + t + 3$. Trouver $f(A), g(A)$, $f(B)$ et $g(B)$ où $A = \begin{pmatrix} 2 & -3 \\ 5 & 1 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$.

9.37. Soit $T : \mathbf{R}^2 \to \mathbf{R}^2$ défini par $T(x, y) = (x + y, 2x)$. Soit $f(t) = t^2 - 2t + 3$. Trouver $f(T)(x, y)$.

9.38. Soit V l'espace vectoriel des polynômes $v(x) = ax^2 + bx + c$. Soit $D : V \to V$ l'opérateur différentiel. Soit $f(t) = t^2 + 2t - 5$. Trouver $f(D)(v(x))$.

9.39. Soit $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Trouver A^2, A^3, A^n.

9.40. Soit $B = \begin{pmatrix} 8 & 12 & 0 \\ 0 & 8 & 12 \\ 0 & 0 & 8 \end{pmatrix}$. Trouver une matrice réelle A, telle que $B = A^3$.

9.41. Considérons une matrice diagonale M et une matrice triangulaire N:

$$M = \begin{pmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_n \end{pmatrix} \quad \text{et} \quad N = \begin{pmatrix} a_1 & b & \dots & c \\ 0 & a_2 & \dots & d \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_n \end{pmatrix}$$

Montrer que, pour un polynôme quelconque $f(t)$, $f(M)$ et $f(N)$ sont de la forme

$$f(M) = \begin{pmatrix} f(a_1) & 0 & \dots & 0 \\ 0 & f(a_2) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & f(a_n) \end{pmatrix} \quad \text{et} \quad f(N) = \begin{pmatrix} f(a_1) & x & \dots & y \\ 0 & f(a_2) & \dots & z \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & f(a_n) \end{pmatrix}$$

9.42. Considérons une matrice diagonale bloc M et une matrice triangulaire bloc N:

$$M = \begin{pmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & A_n \end{pmatrix} \quad \text{et} \quad N = \begin{pmatrix} A_1 & B & \dots & C \\ 0 & A_2 & \dots & D \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & A_n \end{pmatrix}$$

où les A_i sont des matrices carrées. Montrer que, pour tout polynôme $f(t)$, $f(M)$ et $f(N)$ sont de la forme

$$f(M) = \begin{pmatrix} f(A_1) & 0 & \dots & 0 \\ 0 & f(A_2) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & f(A_n) \end{pmatrix} \quad \text{et} \quad f(N) = \begin{pmatrix} f(A_1) & X & \dots & Y \\ 0 & f(A_2) & \dots & Z \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & f(A_n) \end{pmatrix}$$

9.43. Montrer que pour une matrice carrée quelconque (ou opérateur) A, $(P^{-1}AP)^n = P^{-1}A^nP$ où P est inversible. Plus généralement, montrer que $f(P^{-1}AP) = P^{-1}f(A)P$ pour tout polynôme $f(t)$.

9.44. Soit $f(t)$ un polynôme quelconque. Montrer que (i) $f(A^t) = (f(A))^t$, (ii) si A est symétrique, c'est-à-dire $A^t = A$, montrer que $f(A)$ est symétrique.

VALEURS PROPRES ET VECTEURS PROPRES

9.45. Pour chaque matrice, trouver toutes les valeurs propres et les vecteurs propres linéairement indépendants:

$$\text{(i)}\ A = \begin{pmatrix} 2 & 2 \\ 1 & 3 \end{pmatrix}, \quad \text{(ii)}\ B = \begin{pmatrix} 4 & 2 \\ 3 & 3 \end{pmatrix}, \quad \text{(iii)}\ C = \begin{pmatrix} 5 & -1 \\ 1 & 3 \end{pmatrix}$$

Trouver les matrices inversibles P_1, P_2 et P_3 telles que $P_1^{-1}AP_1$, $P_2^{-1}BP_2$ et $P_3^{-1}CP_3$ soient diagonales.

9.46. Pour chacune des matrices suivantes, trouver toutes les valeurs propres et une base pour chacun des espaces propres :

$$\text{(i)}\ A = \begin{pmatrix} 3 & 1 & 1 \\ 2 & 4 & 2 \\ 1 & 1 & 3 \end{pmatrix}, \quad \text{(ii)}\ B = \begin{pmatrix} 1 & 2 & 2 \\ 1 & 2 & -1 \\ -1 & 1 & 4 \end{pmatrix}, \quad \text{(iii)}\ C = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Lorsque c'est possible, trouver les matrices inversibles P_1, P_2 et P_3 telles que $P_1^{-1} AP_1$, $P_2^{-1} BP_2$, et $P_3^{-1} CP_3$ soient des matrices diagonales.

9.47. Considérons $A = \begin{pmatrix} 2 & -1 \\ 1 & 4 \end{pmatrix}$ et $B = \begin{pmatrix} 3 & -1 \\ 13 & -3 \end{pmatrix}$ comme des matrices sur le corps des réels $\mathbf{R}$. Trouver toutes les valeurs propres et les vecteurs propres linéairement indépendants.

9.48. Considérons A et B du précédent problème comme des matrices sur le corps des complexes $\mathbf{C}$. Trouver toutes les valeurs propres et les vecteurs propres linéairement indépendants.

9.49. Pour chacun des opérateurs suivants $T : \mathbf{R}^2 \to \mathbf{R}^2$, trouver toutes les valeurs propres et une base pour chacun des espaces propres: (i) $T(x, y) = (3x + 3y, x + 5y)$; (ii) $T(x, y) = (y, x)$; (iii) $T(x, y) = (y, -x)$.

9.50. Pour chacun des opérateurs suivants $T : \mathbf{R}^3 \to \mathbf{R}^3$, trouver toutes les valeurs propres et une base pour chacun des espaces propres: (i) $T(x, y, z) = (x + y + z,\ 2y + z,\ 2y + 3z)$; (ii) $T(x, y, z) = T(x + y, y + z, -2y - y)$; (iii) $T(x, y, z) = (x - y,\ 2x + 3y + 2z,\ x + y + 2z)$.

9.51. Pour chacune des matrices suivantes sur le corps des complexes $\mathbf{C}$, trouver toutes les valeurs propres et les vecteurs propres linéairement indépendants :

$$\text{(i)}\ \begin{pmatrix} 1 & i \\ 0 & i \end{pmatrix}, \quad \text{(ii)}\ \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}, \quad \text{(iii)}\ \begin{pmatrix} 1 & -3i \\ i & -1 \end{pmatrix}, \quad \text{(iv)}\ \begin{pmatrix} 1 & -2 \\ 1 & -1 \end{pmatrix}.$$

9.52. Supposons que v soit un vecteur propre des opérateurs S et T. Montrer que v est aussi un vecteur propre de l'opérateur $aS + bT$ où a et b sont des scalaires quelconques.

9.53. Supposons que v soit un vecteur propre d'un opérateur T correspondant à la valeur propre λ. Montrer que pour $n > 0$, v est aussi un vecteur propre de T^n correspondant à λ^n.

9.54. Supposons que λ soit une valeur propre d'un opérateur T. Montrer que $f(\lambda)$ est une valeur propre de $f(T)$.

9.55. Montrer que des matrices semblables ont les mêmes valeurs propres.

9.56. Montrer que les matrices A et A^t ont les mêmes valeurs propres. Donner un exemple où A et A^t ont différents vecteurs propres.

9.57. Soient S et T des opérateurs linéaires tels que $ST = TS$. Soit λ une valeur propre de T et soit W son espace propre. Montrer que W est invariant par S, c'est-à-dire que $S(W) \subset W$.

9.58. Soit V un espace vectoriel de dimension finie sur le corps des complexes $\mathbf{C}$. Soit $W \neq \{0\}$ un sous-espace vectoriel de V invariant par l'opérateur linéaire $T : V \to V$. Montrer que W contient un vecteur propre non nul de T.

9.59. Soit A une matrice carrée $n \times n$ sur K. Soient $v_1, \ldots, v_n \in K^n$ des vecteurs propres linéairement indépendants de A correspondant aux valeurs propres $\lambda_1, \lambda_2, \ldots, \lambda_n$ respectivement. Soit P la matrice dont les colonnes sont les vecteurs $v_1, \ldots, v_n$. Montrer que $P^{-1} PA$ est la matrice diagonale dont les éléments diagonaux sont les valeurs propres $\lambda_1, \ldots, \lambda_n$.

POLYNOME MINIMAL ET CARACTERISTIQUE

9.60. Pour chacune des matrices suivantes, trouver un polynôme pour lequel la matrice donnée est une racine:

$$\text{(i)}\ A = \begin{pmatrix} 3 & -7 \\ 4 & 5 \end{pmatrix}, \quad \text{(ii)}\ B = \begin{pmatrix} 5 & -1 \\ 8 & 3 \end{pmatrix}, \quad \text{(iii)}\ C = \begin{pmatrix} 2 & 3 & -2 \\ 0 & 5 & 4 \\ 1 & 0 & -1 \end{pmatrix}.$$

9.61. Considérons la matrice carrée $n \times n$

$$A = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix}$$

Montrer que $f(t) = (t - \lambda)^n$ est à la fois le polynôme caractéristique et le polynôme minimal de A.

9.62. Trouver les polynômes minimal et caractéristique de chacune des matrices

$$A = \begin{pmatrix} 2 & 5 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 4 & 2 & 0 \\ 0 & 0 & 3 & 5 & 0 \\ 0 & 0 & 0 & 0 & 7 \end{pmatrix}, \quad B = \begin{pmatrix} 3 & 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 3 & 1 & 0 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix}, \quad C = \begin{pmatrix} \lambda & 0 & 0 & 0 & 0 \\ 0 & \lambda & 0 & 0 & 0 \\ 0 & 0 & \lambda & 0 & 0 \\ 0 & 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & 0 & \lambda \end{pmatrix}$$

9.63. Soient $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ et $B = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 1 \end{pmatrix}$. Montrer que A et B ont des polynômes caractéristiques différents (et donc ne sont pas semblables), mais ont le même polynôme minimal. Ainsi des matrices non semblables peuvent avoir le même polynôme minimal.

9.64. L'application $T : V \to V$ définie par $T(v) = kv$ est appelée l'application scalaire correspondant à $k \in K$. Montrer que T est l'application scalaire correspondant à $k \in K$ si et seulement si le polynôme minimal de T est $m(t) = t - k$.

9.65. Soit A une matrice carrée $n \times n$ pour laquelle $A^k = 0$ pour $k > n$. Montrer que $A^n = 0$.

9.66. Montrer qu'une matrice A et sa transposée A^t ont le même polynôme minimal.

9.67. Supposons que $f(t)$ soit un polynôme irréductible normalisé pour lequel $f(T) = 0$, où T est un opérateur linéaire $T : V \to V$. Montrer que $f(t)$ est le polynôme minimal de T.

9.68. Considérons la matrice bloc $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$. Montrer que $tI - M = \begin{pmatrix} tI - A & -B \\ -C & tI - D \end{pmatrix}$ est la matrice caractéristique de M.

9.69. Soit T un opérateur linéaire sur un espace vectoriel V de dimension finie. Soit W un sous-espace de V invariant par T c'est-à-dire $T(W) \subset W$. Soit $T_W : W \to W$ la restriction de T à W. (i) Montrer que le polynôme caractéristique de T_W divise le polynôme caractéristique de T. (ii) Montrer que le polynôme minimal de T_W divise le polynôme minimal de T.

9.70. Soit $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$. Montrer que le polynôme caractéristique de A est

$$\Delta(t) = t^3 - (a_{11} + a_{22} + a_{33})t^2 + \left(\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} \right) t - \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

9.71. Soit A une matrice carrée $n \times n$. Le déterminant de la matrice d'ordre $n - m$ obtenue par suppression des lignes et des colonnes passant par m éléments diagonaux de A est appelé un mineur principal de degré $n - m$. Montrer que le coefficient de t^m dans le polynôme caractéristique $\Delta(t) = |tI - A|$ est la somme de tous les principaux mineurs de A de degré $n - m$ multipliés par $(-1)^{n-m}$. (Remarquons que le problème précédent est un cas particulier de ce résultat.)

9.72. Considérons un polynôme arbitraire normalisé $f(t) = t^n + a_{n-1} t^{n-1} + \cdots + a_1 t + a_0$. La matrice carrée $n \times n$ A est appelée la matrice compagnon de $f(t)$:

$$A = \begin{pmatrix} 0 & 0 & \ldots & 0 & -a_0 \\ 1 & 0 & \ldots & 0 & -a_1 \\ 0 & 1 & \ldots & 0 & -a_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 & -a_{n-1} \end{pmatrix}$$

Montrer que $f(t)$ est le polynôme minimal de A.

9.73. Trouver une matrice A dont le polynôme minimal est (i) $t^3 - 5t^2 + 6t + 8$. (ii) $t^4 - 5t^3 - 2t^2 + 7t + 4$

DIAGONALISATION

9.74. Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ une matrice sur le corps des réels **R**. Trouver les conditions nécessaires et suffisantes en fonction de a, b, c et d de telle manière que A soit diagonalisable, c'est-à-dire admette deux vecteurs propres linéairement indépendants.

9.75. Revoir le problème précédent dans le cas où A est une matrice sur le corps des complexes **C**.

9.76. Montrer qu'une matrice (opérateur) est diagonalisable si et seulement si son polynôme minimal est un produit de facteurs linéaires distincts.

9.77. Soient A et B des matrices carrées $n \times n$ sur K telles que (i) $AB = BA$ et (ii) A et B sont toutes deux diagonalisables. Montrer que A et B peuvent être simultanément diagonalisées c'est-à-dire qu'il existe une base de K^n dans laquelle A et B sont représentées par des matrices diagonales. (Voir Problème 9.57).

9.78. Soit $E : V \to V$ un opérateur de projection c'est-à-dire $E^2 = E$. Montrer que E est diagonalisable et peut être représenté par la matrice diagonale $A = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ où r est le rang de E.

REPONSES AUX PROBLEMES SUPPLEMENTAIRES

9.36. $f(A) = \begin{pmatrix} -26 & -3 \\ 5 & -27 \end{pmatrix}$, $g(A) = \begin{pmatrix} -40 & 39 \\ -65 & -27 \end{pmatrix}$, $f(B) = \begin{pmatrix} 3 & 6 \\ 0 & 9 \end{pmatrix}$, $g(B) = \begin{pmatrix} 3 & 12 \\ 0 & 15 \end{pmatrix}$.

9.37. $f(T)(x, y) = (4x - y, -2x + 5y)$.

9.38. $f(D)(v(x)) = -5ax^2 + (4a - 5b)x + (2a + 2b - 5c)$.

9.39. $A^2 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$, $A^3 = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$, $A^n = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$.

9.40 Soit $A = \begin{pmatrix} 2 & a & b \\ 0 & 2 & c \\ 0 & 0 & 2 \end{pmatrix}$. Posons $B = A^3$ et on obtient alors les conditions sur a, b et c.

9.41. (ii) Utilisons (i), nous avons $(f(A))^t = f(A^t) = f(A)$.

9.45. (i) $\lambda_1 = 1$, $u = (2, -1)$; $\lambda_2 = 4$, $v = (1, 1)$.
(ii) $\lambda_1 = 1$, $u = (2, -3)$; $\lambda_2 = 6$, $v = (1, 1)$.
(iii) $\lambda = 4$, $u = (1, 1)$.

Soit $P_1 = \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix}$ et $P_2 = \begin{pmatrix} 2 & 1 \\ -3 & 1 \end{pmatrix}$. P_3 n'existe pas puisque C a seulement un vecteur propre indépendant et donc ne peut être diagonalisé.

9.46. (i) $\lambda_1 = 2$, $u = (1, -1, 0)$, $v = (1, 0, -1)$; $\lambda_2 = 6$, $w = (1, 2, 1)$.
(ii) $\lambda_1 = 3$, $u = (1, 1, 0)$, $v = (1, 0, 1)$; $\lambda_2 = 1$, $w = (2, -1, 1)$.
(iii) $\lambda = 1$, $u = (1, 0, 0)$, $v = (0, 0, 1)$.

Soit $P_1 = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & 2 \\ 0 & -1 & 1 \end{pmatrix}$ et $P_2 = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix}$. P_3 n'existe pas puisque C a au plus deux vecteurs propres linéairement indépendants, et donc ne peut être diagonalisée.

9.47. (i) $\lambda = 3$, $u = (1, -1)$. (ii) B n'a pas de valeurs propres (dans **R**).

9.48. (i) $\lambda = 3$, $u = (1, -1)$. (ii) $\lambda_1 = 2i$, $u = (1, 3-2i)$; $\lambda_2 = -2i$, $v = (1, 3+2i)$.

9.49. (i) $\lambda_1 = 2$, $u = (3, -1)$; $\lambda_2 = 6$, $v = (1, 1)$. (ii) $\lambda_1 = 1$, $u = (1, 1)$; $\lambda_2 = -1$, $v = (1, -1)$. (iii) Il n'y a pas de valeurs propres (dans **R**).

9.50. (i) $\lambda_1 = 1$, $u = (1, 0, 0)$; $\lambda_2 = 4$, $v = (1, 1, 2)$.
(ii) $\lambda = 1$, $u = (1, 0, 0)$. Il n'y a pas d'autres valeurs propres (dans **R**).
(iii) $\lambda_1 = 1$, $u = (1, 0, -1)$; $\lambda_2 = 2$, $v = (2, -2, -1)$; $\lambda_3 = 3$, $w = (1, -2, -1)$.

9.51. (i) $\lambda_1 = 1$, $u = (1, 0)$; $\lambda_2 = i$, $v = (1, 1+i)$. (ii) $\lambda = 1$, $u = (1, 0)$. (iii) $\lambda_1 = 2$, $u = (3, i)$; $\lambda_2 = -2$, $v = (1, -i)$. (iv) $\lambda_1 = i$, $u = (2, 1-i)$; $\lambda_2 = -i$, $v = (2, 1+i)$.

9.56. Soit $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Alors $\lambda = 1$ est la seule valeur propre et $v = (1, 0)$ engendre l'espace propre de $\lambda = 1$.

D'autre part, pour $A^t = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$, $\lambda = 1$ est encore la seule valeur propre, mais $w = (0, 1)$ engendre l'espace propre de $\lambda = -1$.

9.57. Soit $v \in W$, et donc $T(v) = \lambda v$. Alors $T(Sv) = S(Tv) = S(\lambda v) = \lambda(Sv)$, c'est-à-dire, Sv est un vecteur propre de T correspondant a la valeur propre λ. En d'autres termes $Sv \in W$ et donc $S(W) \subset W$.

9.58. Soit $\hat{T} : W \to W$ la restriction de T à W. Le polynôme caractéristique de $\hat{T}$ est un polynôme sur le corps des complexes **C**, qui d'après le théorème fondamental de l'algèbre, a une racine λ. Alors λ est une valeur propre de $\hat{T}$ et donc $\hat{T}$ admet un vecteur propre non nul dans W qui est aussi un vecteur propre de T.

9.59. Supposons $T(v) = \lambda v$. Alors $(kT)(v) = kT(v) = k(\lambda v) = (k\lambda)v$.

9.60. (i) $f(t) = t^2 - 8t + 43$, (ii) $g(t) = t^2 - 8t + 23$, (iii) $h(t) = t^3 - 6t^2 + 5t - 12$.

9.62. (i) $\Delta(t) = (t-2)^3(t-7)^2$; $m(t) = (t-2)^2(t-7)$. (ii) $\Delta(t) = (t-3)^5$; $m(t) = (t-3)^3$. (iii) $\Delta(t) = (t-\lambda)^5$; $m(t) = t - \lambda$.

9.73. En utilisant le résultat du problème 9.72 : (i) $A = \begin{pmatrix} 0 & 0 & -8 \\ 1 & 0 & -6 \\ 0 & 1 & 5 \end{pmatrix}$, (ii) $A = \begin{pmatrix} 0 & 0 & 0 & -4 \\ 1 & 0 & 0 & -7 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 5 \end{pmatrix}$.

9.77. Utiliser le résultat du problème 9.57.

CHAPITRE 10

Formes canoniques

INTRODUCTION

Soit T un opérateur linéaire sur un espace vectoriel de dimension finie. Comme on l'a vu dans le précédent chapitre, T peut ne pas avoir de représentation matricielle sous forme diagonale. Cependant, il est toujours possible de simplifier la représentation matricielle de T de différentes manières. Ceci est le principal sujet de ce chapitre. En particulier, nous obtiendrons le théorème de la décomposition primaire, triangulaire, de Jordan et les formes rationnelles canoniques.

Nous montrerons que les formes canoniques triangulaires et de Jordan existent pour T si et seulement si le polynôme caractéristique $\Delta(t)$ de T admet toutes ses racines dans le corps de référence K ce qui est toujours possible si K est le corps des complexes $\mathbf{C}$, mais n'est pas toujours possible si K est le corps des réels $\mathbf{R}$.

Nous introduirons aussi la notion d'espace quotient. Il s'agit d'une notion très puissante qui sera utilisée dans la démonstration de l'existence de formes canoniques triangulaires et rationnelles.

FORMES TRIANGULAIRES

Soit T un opérateur linéaire sur un espace vectoriel V de dimension n. Supposons que T puisse être représenté par une matrice triangulaire

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ & a_{22} & \dots & a_{2n} \\ & & \dots & \dots \\ & & & a_{nn} \end{pmatrix}$$

Alors le polynôme caractéristique de T

$$\Delta(t) = |tI - A| = (t - a_{11})(t - a_{22}) \dots (t - a_{nn})$$

est un produit de facteurs linéaires. La réciproque est également vraie et est un théorème important :

Théorème 10.1 : Soit $T : V \to V$ un opérateur linéaire dont les facteurs du polynôme caractéristique sont des polynômes linéaires. Il existe alors une base de V dans laquelle T est représenté par une matrice triangulaire.

Autre forme du théorème 10-1 : Soit A une matrice carrée, dont les facteurs du polynôme caractéristique peuvent se mettre sous la forme de polynômes linéaires. Alors A est semblable à une matrice triangulaire, c'est-à-dire, il existe une matrice inversible P telle que $P^{-1}AP$ soit triangulaire.

Nous disons qu'un opérateur T peut être mis sous forme triangulaire, s'il peut être représenté par une matrice triangulaire. Remarquons que dans ce cas les valeurs propres de T sont précisément les éléments de la diagonale principale. Donnons une application de cette remarque.

Exemple 10.1 : Soit A une matrice carrée sur le corps des complexes **C**. Supposons que λ soit une valeur propre de A^2. Montrons que $\sqrt{\lambda}$ ou $-\sqrt{\lambda}$ est une valeur propre de A. Nous savons d'après le théorème précédent que A est semblable à une matrice triangulaire

$$B = \begin{pmatrix} \mu_1 & * & \dots & * \\ & \mu_2 & \dots & * \\ & & \dots\dots\dots & \\ & & & \mu_n \end{pmatrix}$$

Donc A^2 est semblable à la matrice

$$B^2 = \begin{pmatrix} \mu_1^2 & * & \dots & * \\ & \mu_2^2 & \dots & * \\ & & \dots\dots\dots & \\ & & & \mu_n^2 \end{pmatrix}$$

Puisque des matrices semblables ont les mêmes valeurs propres, $\lambda = \mu_i^2$ pour un certain i. On a donc $\mu_i = \sqrt{\lambda}$ ou $\mu_i = -\sqrt{\lambda}$; c'est-à-dire que $\sqrt{\lambda}$ ou $-\sqrt{\lambda}$ est une valeur propre de A.

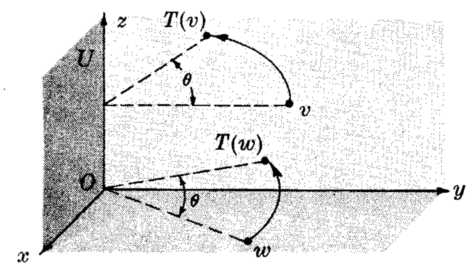
INVARIANCE

Soit $T : V \to V$ linéaire. Un sous-espace W de V est invariant par T ou T-invariant si T transforme W en lui-même c'est-à-dire si $v \in W$ implique $T(v) \in W$. Dans ce cas T restreint à W définit un opérateur linéaire sur W, c'est-à-dire que T induit un opérateur linéaire $\hat{T} : W \to W$ défini par $\hat{T}(w) = T(w)$ pour tout $w \in W$.

Exemple 10.2 : Soit $T : \mathbf{R}^3 \to \mathbf{R}^3$ un opérateur linéaire qui fait subir à un vecteur quelconque une rotation autour de l'axe des z d'un angle θ:

$$T(x, y, z) = (x \cos\theta - y \sin\theta,\ x \sin\theta + y \cos\theta,\ z)$$

Remarquons que chaque vecteur $w = (a, b, 0)$ dans le plan xy, W reste dans W lorsqu'on lui applique la transformation T, c'est-à-dire que W est invariant par T. Remarquons aussi que U, l'axe des z, est invariant par T. De plus la restriction de T à W donne de chaque vecteur son transformé par une rotation autour de l'origine O, et la restriction de T à U est l'application identique sur U.



Exemple 10.3 : Les vecteurs propres non nuls d'un opérateur linéaire $T : V \to V$ peuvent être caractérisés comme les générateurs des sous-espaces vectoriels à une dimension invariants par T. Supposons $T(v) = \lambda v$, $v \neq 0$. Alors $W = \{kv,\ k \in K\}$, le sous-espace à une dimension engendré par v est invariant par T, car

$$T(kv) = k\,T(v) = k(\lambda v) = k\lambda v \in W$$

Réciproquement, supposons dim $U = 1$ et $u \neq 0$ engendre U, U étant invariant par T. Alors $T(u) \in U$ et donc $T(u)$ est un multiple de u, c'est-à-dire $T(u) = \mu u$. Donc u est un vecteur propre de T.

Le théorème suivant donne une classe importante de sous-espaces invariants.

Théorème 10.2 : Soit $T : V \to V$ une application linéaire, et soit $f(t)$ un polynôme quelconque. Alors le noyau de $f(T)$ est invariant par T.

La notion d'invariance peut s'exprimer sous forme de matrices comme il suit.

Théorème 10.3 : Supposons que W soit un sous-espace invariant de $T : V \to V$. Alors T admet une représentation matricielle bloc $\begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$ où A est la représentation matricielle de la restriction de T à W.

DECOMPOSITIONS EN SOMMES DIRECTES INVARIANTES

Un espace vectoriel V est la somme directe de ses sous-espaces $W_1, \ldots, W_r$, et on écrit

$$V = W_1 \oplus W_2 \oplus \ldots \oplus W_r$$

si tout vecteur $v \in V$ peut être écrit de manière unique sous la forme

$$v = w_1 + w_2 + \cdots + w_r \quad \text{avec} \quad w_i \in W_i$$

Le théorème suivant s'applique alors.

Théorème 10.4 : Supposons que $W_1, \ldots, W_r$ soient des sous-espaces de V, et supposons que

$$\{w_{11}, \ldots, w_{1n_1}\}, \ \ldots, \ \{w_{r1}, \ldots, w_{rn_r}\}$$

soient des bases de $W_1, \ldots, W_r$ respectivement. V est alors la somme directe des W_i si et seulement si l'ensemble $\{w_{11}, \ldots, w_{1n_1}, \ldots, w_{r1}, \ldots, w_{rn_r}\}$ est une base de V.

Supposons maintenant $T : V \to V$ linéaire et V la somme directe de sous-espaces (non nuls) invariants par T, $W_1, \ldots, W_r$:

$$V = W_1 \oplus \cdots \oplus W_r \quad \text{et} \quad T(W_i) \subset W_i, \quad i = 1, \ldots, r$$

Notons T_i la restriction de T à W_i. Alors T est dit décomposable suivant les opérateurs T_i, ou T est dit être la somme directe des T_i et on l'écrit $T = T_1 \oplus T_2 \oplus \ldots \oplus T_r$. On dit aussi que les sous-espaces $W_1, \ldots, W_r$ réduisent T ou forment une décomposition de V en somme directe de sous-espaces invariants par T.

Considérons le cas particulier où deux sous-espaces U et W réduisent un opérateur $T : V \to V$; par exemple, $\dim U = 2$ et $\dim W = 3$ et supposons que $\{u_1, u_2\}$ et $\{w_1, w_2, w_3\}$ soient des bases de U et de W respectivement. Si T_1 et T_2 sont les restrictions de T à U et W respectivement, alors

$$\begin{aligned} T_1(u_1) &= a_{11}u_1 + a_{12}u_2 \\ T_1(u_2) &= a_{21}u_1 + a_{22}u_2 \end{aligned} \qquad \begin{aligned} T_2(w_1) &= b_{11}w_1 + b_{12}w_2 + b_{13}w_3 \\ T_2(w_2) &= b_{21}w_1 + b_{22}w_2 + b_{23}w_3 \\ T_2(w_3) &= b_{31}w_1 + b_{32}w_2 + b_{33}w_3 \end{aligned}$$

Donc

$$A = \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} b_{11} & b_{21} & b_{31} \\ b_{12} & b_{22} & b_{32} \\ b_{13} & b_{23} & b_{33} \end{pmatrix}$$

sont les représentations matricielles de T_1 et T_2 respectivement. D'après le théorème précédent $\{u_1, u_2, w_1, w_2, w_3\}$ est une base de V. Puisque $T(u_i) = T_1(u_i)$ et $T(w_j) = T_2(w_j)$, la matrice de T dans cette base est la matrice bloc diagonale $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$.

La généralisation de la démonstration précédente nous donne le théorème suivant.

Théorème 10.5 : Supposons $T : V \to V$ linéaire et V une somme directe de sous-espaces invariants par T, $W_1, \ldots, W_r$. Si A_i est une représentation matricielle de la restriction de T à W_i, alors T peut être représenté par la matrice bloc diagonale

$$M = \begin{pmatrix} A_1 & 0 & \ldots & 0 \\ 0 & A_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & A_r \end{pmatrix}$$

La matrice diagonale bloc M dont les éléments diagonaux sont $A_1, \ldots, A_r$ est quelquefois appelée la somme directe des matrices $A_1, \ldots, A_r$ et est notée $M = A_1 \oplus \ldots \oplus A_r$.

DECOMPOSITION PRIMAIRE

Le théorème précédent montre que tout opérateur $T : V \to V$ est décomposable en opérateurs dont les polynômes minimaux sont des puissances de polynômes irréductibles. Il s'agit ici de la première étape pour l'obtention de la forme canonique de T.

Théorème de décomposition primaire 10.6 : Soit $T : V \to V$ un opérateur linéaire ayant pour polynôme minimal

$$m(t) = f_1(t)^{n_1} f_2(t)^{n_2} \ldots f_r(t)^{n_r}$$

où les $f_i(t)$ sont des polynômes normalisés irréductibles distincts. Alors V est la somme directe des sous-espaces T-invariants $W_1, \ldots, W_r$ où W_i est le noyau de $f_i(T)^{n_i}$. De plus $f_i(t)^{n_i}$ est le polynôme minimal de la restriction de T à W_i.

Puisque les polynômes $f_i(t)^{n_i}$ sont premiers entre eux, le résultat fondamental précédent découle (Problème 10.11) des deux théorèmes suivants.

Théorème 10.7 : Supposons $T : V \to V$ linéaire et supposons que $f(t) = g(t)h(t)$ sont des polynômes tels que $f(T) = 0$ avec $g(t)$ et $h(t)$ premiers entre eux. Alors V est la somme directe des sous-espaces T-invariants U et W, où $U = \operatorname{Ker} g(T)$ et $W = \operatorname{Ker} h(T)$.

Théorème 10.8 : Dans le théorème 10.7, si $f(t)$ est le polynôme minimal de T [avec $g(t)$ et $h(t)$ normalisés], alors $g(t)$ et $h(t)$ sont les polynômes minimaux de la restriction de T à U et W respectivement.

Nous utiliserons aussi le théorème de la décomposition primaire pour démontrer la propriété caractéristique des opérateurs diagonalisables.

Théorème 10.9 : Un opérateur linéaire $T : V \to V$ a une représentation matricielle diagonale si et seulement si son polynôme minimal $m(t)$ est un produit de polynômes linéaires distincts.

Autre forme du théorème 10.9 : Une matrice A est semblable à une matrice diagonale si et seulement si son polynôme minimal est un produit de polynômes linéaires distincts.

Exemple 10.4 : Supposons $A \neq I$, A étant une matrice carrée pour laquelle $A^3 = I$. Déterminer si oui ou non A est semblable à une matrice diagonale, lorsque A est une matrice sur (*1*) le corps des réels $\mathbf{R}$ (*2*) le corps des complexes $\mathbf{C}$.

Puisque $A^3 = I$, A est un zéro du polynôme $f(t) = t^3 - 1 = (t-1)(t^2+t+1)$. Le polynôme minimal $m(t)$ de A ne peut être $t-1$ puisque $A \neq I$. Donc

$$m(t) = t^2 + t + 1 \quad \text{ou} \quad m(t) = t^3 - 1.$$

Puisque aucun de ces polynômes n'est le produit de polynômes linéaires de $\mathbf{R}$, A n'est pas diagonalisable sur $\mathbf{R}$. D'autre part, chacun de ces polynômes est un produit de polynômes distincts linéaires sur $\mathbf{C}$. Donc A est diagonalisable sur $\mathbf{C}$.

OPERATEURS NILPOTENTS

Un opérateur linéaire $T : V \to V$ est dit nilpotent si $T^n = 0$ pour une puissance n donnée positive ; on appelle k indice de nilpotence de T si $T^k = 0$ mais $T^{k-1} \neq 0$. De façon analogue, une matrice carrée A est dite nilpotente si $A^n = 0$ pour une puissance n donnée positive, et d'indice k si $A^k = 0$ mais $A^{k-1} \neq 0$. Plus clairement le polynôme minimal d'un opérateur, ou d'une matrice, nilpotent d'indice k est $m(t) = t^k$, donc 0 est sa seule valeur propre.

Il en résulte le résultat fondamental suivant sur les opérateurs nilpotents.

Théorème 10.10 : Soit $T : V \to V$ un opérateur nilpotent d'indice k. T a alors une représentation matricielle sous forme de matrice diagonale bloc dont les éléments diagonaux sont de la forme suivante :

$$N = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix}$$

(c'est-à-dire que tous les éléments de N sont nuls sauf ceux qui sont directement au-dessus de la diagonale principale et qui sont égaux à 1). Il y a au moins une matrice N d'ordre k et toutes les autres matrices N sont d'ordre $\leqslant k$. Le nombre de matrices N de tous les ordres possibles est uniquement fonction de T. De plus, le nombre total de N de tous les ordres est égal à la nullité de T.

Dans la démonstration du théorème précédent, nous montrerons que le nombre des N d'ordre i est $2m_i - m_{i+1} - m_{i-1}$, où m_i est la nullité de T^i.

Nous remarquons que la matrice précédente N est elle-même nilpotente et que son indice de nilpotence est égal à son ordre (Problème 10-13). Notons que la matrice N d'ordre 1 est justement la matrice nulle 1×1, c'est-à-dire la matrice (0).

FORME CANONIQUE DE JORDAN

Un opérateur T peut être mis sous la forme canonique de Jordan si ses polynômes caractéristique et minimal contiennent des facteurs qui sont uniquement des polynômes linéaires. Ceci est toujours vrai si K est le corps des complexes C. Dans tous les cas, nous pouvons toujours étendre le corps de référence K à un corps dans lequel les polynômes caractéristique et minimal se transforment en produit de facteurs linéaires ; ainsi au sens large chaque opérateur a une forme canonique de Jordan. De façon analogue, chaque matrice est semblable à une matrice dans la forme canonique de Jordan.

Théorème 10.11 : Soit $T : V \to V$ un opérateur linéaire dont les polynômes caractéristique et minimal sont respectivement

$$\Delta(t) = (t-\lambda_1)^{n_1} \dots (t-\lambda_r)^{n_r} \quad \text{et} \quad m(t) = (t-\lambda_1)^{m_1} \dots (t-\lambda_r)^{m_r}$$

où les λ_i sont des scalaires distincts. Alors T admet une représentation matricielle diagonale bloc J dont les éléments diagonaux sont de la forme

$$J_{ij} = \begin{pmatrix} \lambda_i & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda_i & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda_i & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda_i \end{pmatrix}$$

Pour chaque λ_i les blocs correspondants J_{ij} ont les propriétés suivantes.

1) Il y a au moins une J_{ij} d'ordre m_i, toutes les autres J_{ij} sont d'ordre $\leqslant m_i$.

2) La somme des ordres des J_{ij} est n_i.

3) Le nombre des J_{ij} est égal à l'ordre de multiplicité géométrique des λ_i .

4) Le nombre des J_{ij} de chaque ordre possible est uniquement déterminé par T.

La matrice J qui apparaît dans le théorème précédent est appelée la forme canonique de Jordan de l'opérateur T. Une matrice diagonale bloc J_{ij} est appelée une matrice bloc de Jordan correspondant à la valeur propre λ_i. Remarquons que

$$\begin{pmatrix} \lambda_i & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda_i & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda_i & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda_i \end{pmatrix} = \begin{pmatrix} \lambda_i & 0 & \dots & 0 & 0 \\ 0 & \lambda_i & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_i & 0 \\ 0 & 0 & \dots & 0 & \lambda_i \end{pmatrix} + \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix}$$

C'est-à-dire

$$J_{ij} = \lambda_i I + N$$

où N est la matrice bloc nilpotente apparaissant dans le théorème 10.10. En fait, nous démontrons le théorème précédent (Problème 10.18) en montrant que T peut être décomposé en opérateurs, chacun étant la somme d'un scalaire et d'un opérateur nilpotent.

Exemple 10.5 : Supposons que les polynômes caractéristique et minimal d'un opérateur T soient respectivement

$$\Delta(t) = (t-2)^4(t-3)^3 \quad \text{et} \quad m(t) = (t-2)^2(t-3)^2$$

Alors la forme canonique de Jordan de T est l'une des matrices suivantes :

$$\begin{pmatrix} 2 & 1 & & & & & \\ 0 & 2 & & & & & \\ & & 2 & 1 & & & \\ & & 0 & 2 & & & \\ & & & & 3 & 1 & \\ & & & & 0 & 3 & \\ & & & & & & 3 \end{pmatrix} \quad \text{ou} \quad \begin{pmatrix} 2 & 1 & & & & & \\ 0 & 2 & & & & & \\ & & 2 & & & & \\ & & & 2 & & & \\ & & & & 3 & 1 & \\ & & & & 0 & 3 & \\ & & & & & & 3 \end{pmatrix}$$

La première matrice se présente si T a deux vecteurs propres indépendants correspondant à la valeur propre 2 ; et la seconde matrice se présente si T a trois vecteurs propres indépendants correspondant à 2.

SOUS-ESPACES CYCLIQUES

Soit T un opérateur linéaire sur un espace vectoriel V de dimension finie sur K. Supposons $v \in V$ et $v \neq 0$. L'ensemble de tous les vecteurs de la forme $f(T)(v)$ où $f(t)$ décrit l'ensemble des polynômes de K, est un sous-espace T-invariant de V appelé le sous-espace T cyclique de V engendré par v; nous le noterons par $Z(v, T)$ et noterons par T_v la restriction de T à $Z(v, T)$. Nous pouvons définir de manière équivalente $Z(v, T)$ comme l'intersection de tous les sous-espaces T-invariants de V contenant v.

Considérons maintenant la suite

$$v,\ T(v),\ T^2(v),\ T^3(v),\ \ldots$$

des puissances de $T(v)$. Soit k le plus petit entier tel que $T^k(v)$ soit une combinaison linéaire des vecteurs le précédant dans la suite ; c'est-à-dire

$$T^k(v) = -a_{k-1}T^{k-1}(v) - \cdots - a_1 T(v) - a_0 v$$

Alors

$$m_v(t) = t^k + a_{k-1}t^{k-1} + \cdots + a_1 t + a_0$$

est le seul polynôme normalisé de plus bas degré pour lequel $m_v(T)(v) = 0$. Nous appelons $m_v(t)$ l'annihilateur T de v et $Z(v, T)$.

Le théorème suivant s'applique alors.

Théorème 10.12 : Soient $Z(v, T)$, T_v et $m_v(t)$ définies ci-dessus. Alors

1) L'ensemble $\{v, T(v), \ldots, T^{k-1}(v)\}$ est une base de $Z(v, T)$; donc $\dim Z(v, T) = k$.

2) Le polynôme minimal de T_v est $m_v(t)$.

3) La représentation matricielle de T_v dans la base précédente est

$$C = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & 0 & \dots & 0 & -a_1 \\ 0 & 1 & 0 & \dots & 0 & -a_2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & -a_{k-2} \\ 0 & 0 & 0 & \dots & 1 & -a_{k-1} \end{pmatrix}$$

La matrice précédente est appelée la matrice, compagnon du polynôme $m_v(t)$, ou matrice associée au polynôme $m_v(t)$.

FORMES RATIONNELLES CANONIQUES

Dans cette partie, nous présentons la forme rationnelle canonique d'un opérateur linéaire $T : V \to V$. Soulignons que cette forme existe même si le polynôme minimal ne peut être factorisé en polynômes linéaires. (Rappelons que ce n'est pas le cas pour la forme canonique de Jordan).

Lemme 10.13 : Soit $T : V \to V$ un opérateur linéaire, dont le polynôme minimal est $f(t)^n$ où $f(t)$ est un polynôme normalisé irréductible. V est alors la somme directe

$$V = Z(v_1, T) \oplus \cdots \oplus Z(v_r, T)$$

de sous-espaces T-cycliques $Z(v_i, T)$ avec pour T-annihilateurs correspondants

$$f(t)^{n_1}, f(t)^{n_2}, \dots, f(t)^{n_r}, \qquad n = n_1 \geq n_2 \geq \cdots \geq n_r$$

Toute autre décomposition de V en sous-espaces T-cycliques a le même nombre de composantes et le même ensemble de T-annihilateurs.

Soulignons que le lemme précédent n'indique pas que les vecteurs v_i ou les sous-espaces T-cycliques $Z(v_i, T)$ sont uniquement déterminés par T ; mais il indique par contre que l'ensemble des T-annihilateurs est uniquement déterminé par T . Ainsi T a une représentation matricielle unique

$$\begin{pmatrix} C_1 & & & \\ & C_2 & & \\ & & \ddots & \\ & & & C_r \end{pmatrix}$$

où C_i sont les matrices compagnon. En fait, les C_i sont les matrices compagnon des polynômes $f(t)^{n_i}$.

En utilisant le théorème de la décomposition primaire et le lemme précédent, nous obtenons le résultat fondamental suivant.

Théorème 10.14 : Soit $T : V \to V$ un opérateur linéaire dont le polynôme minimal

$$m(t) = f_1(t)^{m_1} f_2(t)^{m_2} \dots f_s(t)^{m_s}$$

où les $f_i(t)$ sont des polynômes distincts normalisés irréductibles. T admet alors une représentation matricielle unique sous forme de matrice bloc diagonale

$$\begin{pmatrix} C_{11} & & & & & \\ & \ddots & & & & \\ & & C_{1r_1} & & & \\ & & & \ddots & & \\ & & & & C_{s1} & & \\ & & & & & \ddots & \\ & & & & & & C_{sr_s} \end{pmatrix}$$

où les C_{ij} sont les matrices compagnon. En particulier, les C_{ij} sont les matrices compagnon des polynômes $f_i(t)^{n_{ij}}$ où

$$m_1 = n_{11} \geq n_{12} \geq \cdots \geq n_{1r_1}, \quad \dots, \quad m_s = n_{s1} \geq n_{s2} \geq \cdots \geq n_{sr_s}$$

La représentation matricielle précédente de T est dite sa forme rationnelle canonique. Les polynômes $f_i(t)^{n_{ij}}$ sont appelés les diviseurs élémentaires de T.

Exemple 10.6 : Soit V un espace vectoriel de dimension 6 sur $\mathbf{R}$, et soit T un opérateur linéaire dont le polynôme minimal est $m(t) = (t^2 - t + 3)(t-2)^2$. Alors la forme rationnelle canonique de T est l'une des sommes directes suivantes de matrices compagnon :

(i) $C(t^2 - t + 3) \oplus C(t^2 - t + 3) \oplus C((t-2)^2)$

(ii) $C(t^2 - t + 3) \oplus C((t-2)^2) \oplus C((t-2)^2)$

(iii) $C(t^2 - t + 3) \oplus C((t-2)^2) \oplus C(t-2) \oplus C(t-2)$

où $C(f(t))$ est la matrice compagnon de $f(t)$; c'est-à-dire

$$\begin{pmatrix} 0 & -3 & & & & \\ 1 & 1 & & & & \\ & & 0 & -3 & & \\ & & 1 & 1 & & \\ & & & & 0 & -4 \\ & & & & 1 & 4 \end{pmatrix} \qquad \begin{pmatrix} 0 & -3 & & & & \\ 1 & 1 & & & & \\ & & 0 & -4 & & \\ & & 1 & 4 & & \\ & & & & 0 & -4 \\ & & & & 1 & 4 \end{pmatrix} \qquad \begin{pmatrix} 0 & -3 & & & & \\ 1 & 1 & & & & \\ & & 0 & -4 & & \\ & & 1 & 4 & & \\ & & & & 2 & \\ & & & & & 2 \end{pmatrix}$$

(i) (ii) (iii)

ESPACES QUOTIENTS

Soit V un espace vectoriel sur un corps K et soit W un sous-espace de V. Si v est un vecteur quelconque de V, nous pouvons écrire $v + W$ comme l'ensemble des sommes $v + w$ avec $w \in W$:

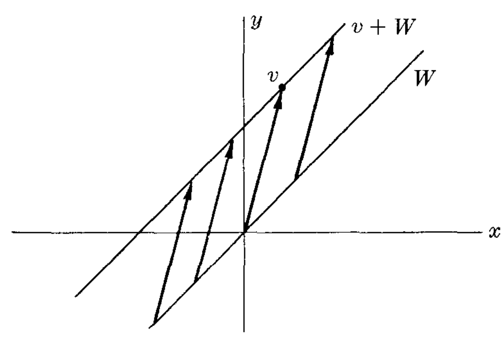
$$v + W = \{v + w : w \in W\}$$

Ces ensembles sont appelés les classes de W dans V. Nous montrerons (Problème 10.22) que ces classes réalisent une partition de V en sous-ensembles deux à deux disjoints.

Exemple 10.7 : Soit W le sous-espace de $\mathbf{R}^2$ défini par

$$W = \{(a, b) \; ; \; a = b\}$$

W est donc la droite dont l'équation est $x - y = 0$. Nous pouvons considérer $v + W$ comme la translatée de cette droite, obtenue en faisant subir à chaque point de W une translation de vecteur v. Comme on le remarque sur le diagramme de droite, $v + W$ est aussi une droite qui est parallèle à W. Ainsi toutes les classes de W dans $\mathbf{R}^2$ sont constituées de toutes les droites parallèles à W.



Dans le théorème suivant, nous utiliserons les classes d'un sous-espace W d'un espace vectoriel V pour définir un nouvel espace vectoriel ; qui sera appelé l'espace quotient de V par W et est noté V/W.

Théorème 10.15 : Soit W un sous-espace d'un espace vectoriel sur un corps K. Les classes de W dans V forment un espace vectoriel sur K, en prenant pour addition et multiplication par un scalaire :

(i) $(u + W) + (v + W) = (u + v) + W$

(ii) $k(u + W) = ku + W$, où $k \in K$.

Remarquons que, dans la démonstration du théorème précédent, il est d'abord nécessaire de montrer que les opérations sont bien définies ; c'est-à-dire que lorsque $u + W = u' + W$ et $v + W = v' + W$, alors

(i) $(u + v) + W = (u' + v') + W$ et (ii) $ku + W = ku' + W$, pour $k \in K$ quelconque

Dans le cas d'un sous-espace invariant, nous avons le résultat suivant.

Théorème 10.16 : Supposons que W soit un sous-espace invariant par rapport à l'opérateur linéaire $T : V \to V$. T induit alors un opérateur linéaire $\overline{T}$ sur V/W défini par $\overline{T}(v + W) = T(v) + W$. De plus, si T est un zéro d'un polynôme quelconque, alors il en est de même pour $\overline{T}$. Donc le polynôme minimal de $\overline{T}$ divise le polynôme minimal de T.

PROBLEMES RESOLUS

SOUS-ESPACES INVARIANTS

10.1. Supposons $T : V \to V$ linéaire. Montrer que chacun des sous-espaces suivants est invariant par T.

1) $\{0\}$; 2) V; 3) le noyau de T; 4) l'image de T.

1) Nous avons $T(0) = 0 \in \{0\}$; donc $\{0\}$ est invariant par T.

2) Pour tout $v \in V$, $T(v) \in V$; donc V est invariant par T.

3) Soit $u \in \operatorname{Ker} T$. Alors $T(u) = 0 \in \operatorname{Ker} T$ puisque le noyau de T est un sous-espace de V. Ainsi $\operatorname{Ker} T$ est invariant par T.

4) Puisque $T(v) \in \operatorname{Im} T$ pour tout $v \in V$, ceci est vérifié si $v \in \operatorname{Im} T$; donc l'image de T est invariante par T.

10.2. Supposons $\{W_i\}$ un système de sous-espaces invariants par T d'un espace vectoriel V. Montrer alors que l'intersection $W = \cap_i W_i$ est aussi invariante, par T.

Supposons $v \in W$; alors $v \in W_i$ pour tout i. Puisque W_i est invariant par T, $T(v) \in W_i$ pour tout i. Ainsi $T(v) \in W = \cap_i W_i$ et donc W est invariant par T.

10.3. Démontrer le théorème 10.2 : Soit $T : V \to V$ un opérateur linéaire quelconque et soit $f(t)$ un polynôme quelconque. Le noyau de $f(T)$ est alors invariant par T.

Supposons $v \in \operatorname{Ker} f(T)$ c'est-à-dire $f(T)(v) = 0$. Cherchons à montrer que $T(v)$ appartient aussi au noyau de $f(T)$ c'est-à-dire que $f(T)\,(T(v)) = 0$. Puisque $f(t)t = tf(t)$ nous avons $f(T)T = Tf(T)$. Ainsi

$$f(T)T(v) = Tf(T)(v) = T(0) = 0$$

comme il est demandé.

10.4. Trouver tous les sous-espaces invariants de $A = \begin{pmatrix} 2 & -5 \\ 1 & -2 \end{pmatrix}$ considérée comme un opérateur de $\mathbf{R}^2$.

Tout d'abord, $\mathbf{R}^2$ et $\{0\}$ sont invariants par A. Si A admet maintenant d'autres sous-espaces invariants, ils doivent être de dimension 1. Le polynôme caractéristique de A est

$$\Delta(t) = |tI - A| = \begin{vmatrix} t-2 & 5 \\ -1 & t+2 \end{vmatrix} = t^2 + 1$$

Donc A n'a pas de valeurs propres (dans $\mathbf{R}$) et donc A n'a pas de vecteurs propres. Mais les sous-espaces invariants à une dimension correspondent aux vecteurs propres ; donc $\mathbf{R}^2$ et $\{0\}$ sont les seuls sous-espaces invariants de A.

10.5. Démontrer le théorème 10.3 : Supposons que W soit un sous-espace invariant de $T : V \to V$. Alors T a une représentation matricielle sous forme de matrice diagonale bloc $\begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$ où A est une représentation matricielle de la restriction $\hat{T}$ de T à W.

Choisissons une base $\{w_1, \ldots, w_r\}$ de W et étendons-la à une base $\{w_1, \ldots, w_r, v_1, \ldots, v_s\}$ de V. Nous avons

$$\begin{array}{rcrcl}
\hat{T}(w_1) & = & T(w_1) & = & a_{11}w_1 + \cdots + a_{1r}w_r \\
\hat{T}(w_2) & = & T(w_2) & = & a_{21}w_1 + \cdots + a_{2r}w_r \\
\multicolumn{5}{c}{\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots} \\
\hat{T}(w_r) & = & T(w_r) & = & a_{r1}w_1 + \cdots + a_{rr}w_r \\
 & & T(v_1) & = & b_{11}w_1 + \cdots + b_{1r}w_r + c_{11}v_1 + \cdots + c_{1s}v_s \\
 & & T(v_2) & = & b_{21}w_1 + \cdots + b_{2r}w_r + c_{21}v_1 + \cdots + c_{2s}v_s \\
 & & \multicolumn{3}{c}{\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots} \\
 & & T(v_s) & = & b_{s1}w_1 + \cdots + b_{sr}w_r + c_{s1}v_1 + \cdots + c_{ss}v_s
\end{array}$$

Mais la matrice de T dans cette base est la transposée de la matrice des coefficients dans le système précédent d'équations (voir page 150). Donc elle a la forme $\begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$ où A est la transposée de la matrice des coefficients du sous-système évident. Par le même raisonnement, A est la matrice de $\hat{T}$ relativement à la base $\{w_i\}$ de W.

10.6. Soit $\hat{T}$ la restriction d'un opérateur T a un sous-espace invariant W, tel que $\hat{T}(w) = T(w)$ pour tout $w \in W$. Démontrer que

(i) pour tout polynôme $f(t)$, $f(\hat{T})\,(w) = f(T)\,(w)$.

(ii) le polynôme minimal de $\hat{T}$ divise le polynôme minimal de T.

(i) Si $f(t) = 0$ ou si $f(t)$ est une constante, donc de degré 1, le résultat est évident. Supposons que $\deg f = n > 1$, et que le résultat soit vrai pour les polynômes de degré inférieur à n. On a

$$f(t) = a_n t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0$$

Alors
$$\begin{aligned}
f(\hat{T})(w) &= (a_n\hat{T}^n + a_{n-1}\hat{T}^{n-1} + \cdots + a_0 I)(w) \\
&= (a_n\hat{T}^{n-1})(\hat{T}(w)) + (a_{n-1}\hat{T}^{n-1} + \cdots + a_0 I)(w) \\
&= (a_n T^{n-1})(T(w)) + (a_{n-1}T^{n-1} + \cdots + a_0 I)(w) \\
&= f(T)(w)
\end{aligned}$$

(ii) Soit $m(t)$ le polynôme minimal de T. Alors d'après (i), $m(\hat{T})(w) = m(T)(w) = 0(w) = 0$ pour tout $w \in W$; c'est-à-dire, que $\hat{T}$ est un zéro du polynôme $m(t)$. Donc le polynôme minimal de $\hat{T}$ divise $m(t)$.

DECOMPOSITIONS EN SOMMES DIRECTES INVARIANTES

10.7. Démontrer le théorème 10.4 : Supposons que $W_1, \ldots, W_r$ soient des sous-espaces de V et supposons, que pour $i = 1, \ldots, r, \{w_{i1}, \ldots, w_{in_i}\}$ soit une base de W_i. Alors V est la somme directe des W_i, si et seulement si l'union

$$B = \{w_{11}, \ldots, w_{1n_1}, \ldots, w_{r1}, \ldots, w_{rn_r}\}$$

est une base de V.

Supposons que B soit une base de V. Alors pour tout $v \in V$,

$$v = a_{11}w_{11} + \cdots + a_{1n_1}w_{1n_1} + \cdots + a_{r1}w_{r1} + \cdots + a_{rn_r}w_{rn_r} = w_1 + w_2 + \cdots + w_r$$

où $w_i = a_{i1}w_{i1} + \cdots + a_{in_i}w_{in_i} \in W_i$. Montrons ensuite qu'une telle somme est unique. Supposons

$$v = w_1' + w_2' + \cdots + w_r' \qquad \text{où} \qquad w_i' \in W_i$$

Puisque $\{w_{i1}, \ldots, w_{in_i}\}$ est une base de W_i, $w_i' = b_{i1}w_{i1} + \cdots + b_{in_i}w_{in_i}$ et donc

$$v = b_{11}w_{11} + \cdots + b_{1n_1}w_{1n_1} + \cdots + b_{r1}w_{r1} + \cdots + b_{rn_r}w_{rn_r}$$

Puisque B est une base de V, $a_{ij} = b_{ij}$ pour tout i et j. Donc $w_i = w_i'$ et la somme donnant v est unique. En conséquence, V est la somme directe des W_i.

Réciproquement, supposons que V soit la somme directe des W_i. Alors pour tout $v \in V$, $v = w_1 + \ldots + w_r$ où $w_i \in W_i$. Or $\{w_{ij}\}$ est une base de W_i, chaque w_i est une combinaison linéaire des w_{ij} et donc v est une combinaison linéaire des éléments de B. Donc B engendre V. Montrons maintenant que B est linéairement indépendant. Supposons

$$a_{11}w_{11} + \cdots + a_{1n_1}w_{1n_1} + \cdots + a_{r1}w_{r1} + \cdots + a_{rn_r}w_{rn_r} = 0$$

Remarquons que $a_{i1} w_{i1} + \cdots + a_{in_i} w_{in_i} \in W_i$. Nous avons aussi $0 = 0 + 0 + \cdots + 0$ où $0 \in W_i$. Puisqu'une telle somme pour 0 est unique,

$$a_{i1}w_{i1} + \cdots + a_{in_i}w_{in_i} = 0 \qquad \text{pour } i = 1, \ldots, r$$

L'indépendance des bases $\{w_{ij_i}\}$ implique que tous les a sont nuls. Ainsi B est linéairement indépendant et donc est une base de V.

10.8. Supposons $T : V \to V$ linéaire et supposons que $T = T_1 \oplus T_2$ par rapport à une décomposition en sommes directes T-invariante. $V = U \oplus W$. Montrer que

(i) $m(t)$ est le plus petit commun multiple de $m_1(t)$ et $m_2(t)$ où $m(t)$, $m_1(t)$, $m_2(t)$ sont les polynômes minimaux de T, T_1 et T_2 respectivement.

(ii) $\Delta(t) = \Delta_1(t)\, \Delta_2(t)$ où $\Delta(t)$, $\Delta_1(t)$ et $\Delta_2(t)$ sont les polynômes caractéristiques de T, T_1 et T_2 respectivement.

(i) D'après le problème 10.6 chacun des $m_1(t)$ et $m_2(t)$ divise $m(t)$. Supposons maintenant que $f(t)$ soit un multiple à la fois de $m_1(t)$ et de $m_2(t)$; alors $f(T_1)(U) = 0$ et $f(T_2)(W) = 0$. Soit $v \in V$, alors $v = u + w$ avec $u \in U$ et $w \in W$. On a donc

$$f(T)\,v = f(T)\,u + f(T)\,w = f(T_1)\,u + f(T_2)\,w = 0 + 0 = 0$$

C'est-à-dire que T est un zéro de $f(t)$. Donc $m(t)$ divise $f(t)$ et donc $m(t)$ est le plus petit commun multiple de $m_1(t)$ et $m_2(t)$.

(ii) D'après le théorème 10.5, T a une représentation matricielle $M = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$ où A et B sont les représentations matricielles de T_1 et T_2 respectivement. Alors d'après le problème 9.66

$$\Delta(t) = |tI - M| = \begin{vmatrix} tI - A & 0 \\ 0 & tI - B \end{vmatrix} = |tI - A|\,|tI - B| = \Delta_1(t)\,\Delta_2(t)$$

comme il est demandé.

10.9. Démontrer le théorème 10.7. Supposons $T : V \to V$ linéaire, et supposons que $f(t) = g(t)\,h(t)$ soient des polynômes tels que $f(T) = 0$, $g(t)$ et $h(t)$ soient premiers entre eux. Alors V est la somme directe des sous-espaces invariants par T, U et W où $U = \operatorname{Ker} g(T)$ et $W = \operatorname{Ker} h(T)$.

Remarquons d'abord que U et W sont invariants par T d'après le théorème 10.2. Puisque $g(t)$ et $h(t)$ sont premiers entre eux, il existe des polynômes $r(t)$ et $s(t)$ tels que

$$r(t)\,g(t) + s(t)\,h(t) = 1$$

Donc, pour l'opérateur T,

$$r(T)\,g(T) + s(T)\,h(T) = I \qquad (*)$$

Soit $v \in V$; alors d'après (*)

$$v = r(T)\,g(T)\,v + s(T)\,h(T)\,v$$

Mais le premier terme dans cette somme appartient à $W = \operatorname{Ker} h(T)$ puisque

$$h(T)\,r(T)\,g(T)\,v = r(T)\,g(T)\,h(T)\,v = r(T)\,f(T)\,v = r(T)\,0\,v = 0$$

De façon analogue, le second terme appartient à U. Donc V est la somme de U et W.

Pour démontrer que $V = U \oplus W$, nous devons montrer qu'une somme $v = u + w$ avec $u \in U$ et $w \in W$ uniquement déterminée par v. Appliquons l'opérateur $r(T)\,g(T)$ à $v = u + w$ en utilisant le fait que $g(T)\,u = 0$; on obtient

$$r(T)\,g(T)\,v = r(T)\,g(T)\,u + r(T)\,g(T)\,w = r(T)\,g(T)\,w$$

Appliquons aussi (*) à w seul en utilisant le fait que $h(T)\,w = 0$; on obtient

$$w = r(T)\,g(T)\,w + s(T)\,h(T)\,w = r(T)\,g(T)\,w$$

Les deux formules précédentes nous donnent $w = r(T)\,g(T)\,v$ et donc w est uniquement déterminé par v. De façon analogue u est uniquement déterminé par v. Donc $V = U \oplus W$ comme demandé.

10.10. Démontrer le théorème 10.8 : Dans le théorème 10.7 (Problème 10.9), si $f(t)$ est le polynôme minimal de T (avec $g(t)$ et $h(t)$ normalisés), alors $g(t)$ est le polynôme minimal de la restriction T_1 de T à U et $h(t)$ est le polynôme minimal de la restriction T_2 de T à W.

Soient $m_1(t)$ et $m_2(t)$ les polynômes minimaux de T_1 et T_2 respectivement. Remarquons que $g(T_1) = 0$ et $h(T_2) = 0$ car $U = \text{Ker } g(T)$ et $W = \text{Ker } h(T)$. Ainsi

$$m_1(t) \text{ divise } g(t) \quad \text{et} \quad m_2(t) \text{ divise } h(t) \qquad (1)$$

D'après le Problème 10.9, $f(t)$ est le plus petit commun multiple de $m_1(t)$ et $m_2(t)$. Mais $m_1(t)$ et $m_2(t)$ sont premiers entre eux, puisque $g(t)$ et $h(t)$ sont premiers entre eux. En conséquence, $f(t) = m_1(t)m_2(t)$. Nous avons aussi $f(t) = g(t)\, h(t)$. Ces deux équations avec (*1*) et le fait que tous les polynômes sont normalisés impliquent que $g(t) = m_1(t)$ et $h(t) = m_2(t)$ comme il est demandé.

10.11. Démontrer le théorème 10.6 sur la décomposition primaire : Soit $T : V \to V$ un opérateur linéaire avec pour polynôme minimal

$$m(t) = f_1(t)^{n_1} f_2(t)^{n_2} \ldots f_r(t)^{n_r}$$

où les $f_i(t)$ sont des polynômes distincts irréductibles et normalisés. Alors V est la somme directe des sous-espaces invariants par T, $W_1, W_2, \ldots, W_r$ où W_i est le noyau de $f_i(T)^{n_i}$. De plus $f_i(t)^{n_i}$ est le polynôme minimal de la restriction de T à W_i.

La démonstration se fait par récurrence sur r. Le cas $r = 1$ est trivial. Supposons que le théorème ait été démontré pour $r - 1$. D'après le théorème 10.7 nous pouvons écrire V comme la somme directe des sous-espaces T-invariants W_1 et V_1 où W_1 est le noyau de $f_1(T)^{n_1}$ et où V_1 est le noyau de $f_2(T)^{n_2} \ldots f_r(T)^{n_r}$. D'après le théorème 10.8 les polynômes minimaux des restrictions de T à W_1 et V_1 sont respectivement $f_1(t)^{n_1}$ et $f_2(t)^{n_2} \ldots f_r(t)^{n_r}$.

Notons la restriction de T à V_1 par T_1. D'après l'hypothèse de récurrence, V_1 est la somme directe des sous-espaces $W_2, \ldots, W_r$ telle que W_i est le noyau de $f_i(T_1)^{n_i}$ et telle que $f_i(t)^{n_i}$ est le polynôme minimal de la restriction de T_1 à W_i. Mais le noyau de $f_i(T)^{n_i}$ pour $i = 2, \ldots, r$ est nécessairement contenu dans V_1, puisque $f_i(t)^{n_i}$ divise $f_2(t)^{n_2} \ldots f_r(t)^{n_r}$. Ainsi le noyau de $f_i(T)^{n_i}$ est le même que le noyau de $f_i(T_1)^{n_i}$ qui est W_i. Ainsi, la restriction de T à W_i est la même que la restriction de T_1 à W_i pour $i = 2, \ldots, r)$; donc $f_i(t)^{n_i}$ est aussi le polynôme minimal pour la restriction de T à W_i. Donc $V = W_1 \oplus W_2 \oplus \cdots \oplus W_r$ est la décomposition demandée de T.

10.12. Démontrer le théorème 10.9 : Un opérateur linéaire $T : V \to V$ a une représentation matricielle sous forme diagonale si et seulement si son polynôme minimal $m(t)$ est un produit de polynômes linéaires distincts.

Supposons que $m(t)$ soit un produit de polynômes linéaires distincts ; par exemple

$$m(t) = (t - \lambda_1)(t - \lambda_2) \ldots (t - \lambda_r)$$

où les λ_i sont des scalaires distincts. D'après le théorème de la décomposition primaire, V est la somme directe des sous-espaces $W_1, \ldots, W_r$ où $W_i = \text{Ker } (T - \lambda_i I)$. Ainsi si $v \in W_i$ alors $(T - \lambda_i I)(v) = 0$ où $T(v) = \lambda_i v$. En d'autres termes, chaque vecteur de W_i est un vecteur propre appartenant à la valeur propre λ_i. D'après le théorème 10.4, l'union des bases de $W_1, \ldots, W_r$ est une base de V. Cette base est une base de vecteurs propres et donc T est diagonalisable.

Réciproquement, supposons que T soit diagonalisable, c'est-à-dire que V soit une base contenant les vecteurs propres de T. Soit $\lambda_1, \ldots, \lambda_s$ les valeurs propres distinctes de T. Alors l'opérateur

$$f(T) = (T - \lambda_1 I)(T - \lambda_2 I) \ldots (T - \lambda_s I)$$

transforme chaque vecteur de base en 0. Ainsi $f(T) = 0$ et donc le polynôme minimal $m(t)$ de T divise le polynôme

$$f(t) = (t - \lambda_1)(t - \lambda_2) \ldots (t - \lambda_s I)$$

En conséquence, $m(t)$ est un produit de polynômes linéaires distincts.

OPERATEURS NILPOTENTS. FORME CANONIQUE DE JORDAN

10.13. Soit $T : V \to V$ linéaire. Supposons que pour $v \in V$, $T^k(v) = 0$, mais $T^{k-1}(v) \neq 0$. Démontrer que

(1) L'ensemble $S = \{v, T(v) \ldots T^{k-1}(v)\}$ est linéairement indépendant.

(2) Le sous-espace W engendré par S est invariant par T.

(3) La restriction $\hat{T}$ de T à W est nilpotente et d'indice k.

(4) Relativement à la base $\{T^{k-1}(v), \ldots, T(v), v\}$ de W, la matrice de T est de la forme

$$\begin{pmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 0 & 1 \\ 0 & 0 & 0 & \ldots & 0 & 0 \end{pmatrix}$$

Donc la matrice $k \times k$ précédente est nilpotente et d'indice k.

(1) Supposons

$$av + a_1 T(v) + a_2 T^2(v) + \cdots + a_{k-1} T^{k-1}(v) = 0 \qquad (*)$$

En appliquant T^{k-1} à (*) et en utilisant le fait que $T^k(v) = 0$, on obtient a $T^{k-1}(v)$; puisque $T^{k-1}(v) \neq 0$, $a = 0$. Appliquons maintenant T^{k-2} à (*) et utilisons le fait que $T^k(v) = 0$ et $a = 0$, nous trouvons $a_1 T^{k-1}(v) = 0$ donc $a_1 = 0$. En appliquant ensuite T^{k-3} à (*) et en utilisant le fait que $T^k(v) = 0$ et $a = a_1 = 0$, nous obtenons $a_2 T^{k-1}(v) = 0$, donc $a_2 = 0$. En continuant la même démonstration nous trouvons que tous les a sont nuls donc S est indépendant.

(2) Soit $v \in W$. Alors

$$v = bv + b_1 T(v) + b_2 T^2(v) + \cdots + b_{k-1} T^{k-1}(v)$$

En utilisant $T^k(v) = 0$, nous avons

$$T(v) = b\,T(v) + b_1 T^2(v) + \cdots + b_{k-2} T^{k-1}(v) \in W$$

Donc W est invariant par T.

(3) D'après l'hypothèse que $T^k(v) = 0$, donc pour $i = 0, \ldots, k-1$,

$$\hat{T}^k(T^i(v)) = T^{k+i}(v) = 0$$

C'est-à-dire qu'en appliquant $\hat{T}^k$ à chaque générateur de W, nous obtenons 0 ; donc $\hat{T}^k = 0$ et donc $\hat{T}$ est nilpotent d'indice au plus égal à k. D'autre part $\hat{T}^{k-1}(v) = T^{k-1}(v) \neq 0$; donc T est nilpotent d'indice égal exactement à k.

(4) Pour la base $\{T^{k-1}(v), T^{k-2}(v), \ldots, T(v), v\}$ de W,

$$\begin{array}{lcl} \hat{T}(T^{k-1}(v)) & = & T^k(v) = 0 \\ \hat{T}(T^{k-2}(v)) & = & T^{k-1}(v) \\ \hat{T}(T^{k-3}(v)) & = & T^{k-2}(v) \\ \ldots & \ldots & \ldots \\ \hat{T}(T(v)) & = & T^2(v) \\ \hat{T}(v) & = & T(v) \end{array}$$

Donc la matrice de T dans cette base est

$$\begin{pmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 0 & 1 \\ 0 & 0 & 0 & \ldots & 0 & 0 \end{pmatrix}$$

10.14. Soit $T : V \to V$ linéaire. Soit $U = \text{Ker } T^i$ et $W = \text{Ker } T^{i+1}$. Montrer que (1) $U \subset W$ (2) $T(W) \subset U$.

(1) Supposons $u \in U = \text{Ker } T^i$. Alors $T^i(u) = 0$ et donc $T^{i+1}(u) = T(T^i(u)) = T(0) = 0$. Donc $u \in \text{Ker } T^{i+1} = W$. Mais ceci est vrai pour tout $u \in U$; donc $U \subset W$.

(2) De façon analogue, si $w \in W = \text{Ker } T^{i+1}$, alors $T^{i+1}(w) = 0$. Donc $T^{i+1}(w) = T^i(T(w)) = 0$ et donc $T(W) \subset U$.

10.15. Soit $T : V \to V$ linéaire. Soit $X = \text{Ker } T^{i-2}$, $Y = \text{Ker } T^{i-1}$ et $Z = \text{Ker } T^i$. D'après le problème précédent $X \subset Y \subset Z$. Supposons que

$$\{u_1, \ldots, u_r\}, \ \{u_1, \ldots, u_r, v_1, \ldots, v_s\}, \ \{u_1, \ldots, u_r, v_1, \ldots, v_s, w_1, \ldots, w_t\}$$

soient des bases de X, Y et Z respectivement. Montrer que

$$S = \{u_1, \ldots, u_r, T(w_1), \ldots, T(w_t)\}$$

est contenu dans Y et est linéairement indépendant.

D'après le problème précédent $T(Z) \subset Y$ et donc $S \subset Y$. Supposons maintenant que S soit linéairement dépendant. Alors il existe une relation

$$a_1u_1 + \cdots + a_ru_r + b_1\,T(w_1) + \cdots + b_t\,T(w_t) = 0$$

où au moins un coefficient n'est pas nul. De plus, puisque $\{u_i\}$ est indépendant au moins un des b_k ne doit pas être nul. En transposant, nous trouvons

$$b_1\,T(w_1) + \cdots + b_t\,T(w_t) = -a_1u_1 - \cdots - a_ru_r \in X = \text{Ker } T^{i-2}$$

Donc $$T^{i-2}(b_1\,T(w_1) + \cdots + b_t\,T(w_t)) = 0$$

Ainsi $$T^{i-1}(b_1w_1 + \cdots + b_tw_t) = 0 \quad \text{et donc} \quad b_1w_1 + \cdots + b_tw_t \in Y = \text{Ker } T^{i-1}$$

Puisque $\{u_i, v_j\}$ engendre Y, nous obtenons une relation entre les u_i, v_j et w_k où l'un des coefficients, c'est-à-dire l'un des b_k n'est pas nul. Ceci contredit le fait que $\{u_i, v_j, w_k\}$ est indépendant. Donc S doit être aussi indépendant.

10.16. Démontrer le théorème 10.10 : Soit $T : V \to V$ un opérateur nilpotent d'indice k. Alors T a une représentation matricielle sous forme de matrice diagonale bloc dont les éléments diagonaux sont de la forme

$$N = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ \multicolumn{6}{c}{\dotfill} \\ 0 & 0 & 0 & \ldots & 0 & 1 \\ 0 & 0 & 0 & \ldots & 0 & 0 \end{pmatrix}$$

Il existe au moins un N d'ordre k et tous les autres N sont d'ordre $\leqslant k$. Le nombre des N de chaque ordre possible est uniquement déterminé par T. De plus le nombre total des N de tous les ordres est la nullité de T.

Supposons $\dim V = n$. Soit $W_1 = \text{Ker } T$, $W_2 = \text{Ker } T^2, \ldots, W_k = \text{Ker } T^k$. Posons $m_i = \dim W_i$ pour $i = 1, \ldots, k$. Puisque T est d'indice k, $W_k = V$ et $W_{k-1} \neq V$ et donc $m_{k-1} < m_k = n$. D'après le problème 10.17,

$$W_1 \subset W_2 \subset \cdots \subset W_k = V$$

Ainsi, par récurrence, nous pouvons choisir une base $\{u_1, \ldots, u_n\}$ de V telle que $\{u_1, \ldots, u_{m_i}\}$ soit une base de W_i.

Choisissons maintenant une nouvelle base pour V dans laquelle T a la forme désirée. Il sera pratique d'indiquer les membres de cette nouvelle base par des couples d'indices. Commençons en posant

$$v(1, k) = u_{m_{k-1}+1}, \ v(2, k) = u_{m_{k-1}+2}, \ \ldots, \ v(m_k - m_{k-1}, k) = u_{m_k}$$

et posons

$$v(1, k-1) = Tv(1, k),\ \ v(2, k-1) = Tv(2, k),\ \ \ldots,\ \ v(m_k - m_{k-1}, k-1) = Tv(m_k - m_{k-1}, k)$$

D'après le problème précédent,

$$S_1 \;=\; \{u_1 \;\ldots,\; u_{m_{k-2}},\; v(1, k-1),\; \ldots,\; v(m_k - m_{k-1}, k-1)\}$$

est un sous-espace de W_{k-1} linéairement indépendant. Etendons S_1 à une base de W_{k-1} en adjoignant de nouveaux éléments (si nécessaire) que nous noterons par

$$v(m_k - m_{k-1} + 1, k-1),\ \ v(m_k - m_{k-1} + 2, k-1),\ \ \ldots,\ \ v(m_{k-1} - m_{k-2}, k-1)$$

Posons maintenant

$$v(1, k-2) = Tv(1, k-1),\ \ v(2, k-2) = Tv(2, k-1),\ \ \ldots,$$
$$v(m_{k-1} - m_{k-2}, k-2) = Tv(m_{k-1} - m_{k-2}, k-1)$$

D'après le problème précédent

$$S_2 \;=\; \{u_1,\; \ldots,\; u_{m_{k-3}},\; v(1, k-2),\; \ldots,\; v(m_{k-1} - m_{k-2}, k-2)\}$$

est un sous-espace de W_{k-2} linéairement indépendant, que nous pouvons étendre à une base de W_{k-2} en adjoignant les éléments

$$v(m_{k-1} - m_{k-2} + 1, k-2),\ \ v(m_{k-1} - m_{k-2} + 2, k-2),\ \ \ldots,\ \ v(m_{k-2} - m_{k-3}, k-2)$$

En continuant de cette manière, nous obtenons une nouvelle base pour V, que nous écrivons comme suit pour des raisons de commodité :

$$\begin{array}{lllll}
v(1, k), & \ldots,\ v(m_k - m_{k-1}, k) & & & \\
v(1, k-1), & \ldots,\ v(m_k - m_{k-1}, k-1), & \ldots,\ v(m_{k-1} - m_{k-2}, k-1) & & \\
\cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\
v(1, 2), & \ldots,\ v(m_k - m_{k-1}, 2), & \ldots,\ v(m_{k-1} - m_{k-2}, 2), & \ldots,\ v(m_2 - m_1, 2) & \\
v(1, 1), & \ldots,\ v(m_k - m_{k-1}, 1), & \ldots,\ v(m_{k-1} - m_{k-2}, 1), & \ldots,\ v(m_2 - m_1, 1), & \ldots,\ v(m_1, 1)
\end{array}$$

La dernière ligne forme une base de W_1, les deux dernières lignes forment une base de W_2, etc. mais ce qui est important pour nous c'est que T transforme chaque vecteur en le vecteur immédiatement sous lui dans le tableau précédent ou en 0 si le vecteur est dans la dernière ligne, c'est-à-dire

$$Tv(i, j) \;=\; \begin{cases} v(i, j-1) & \text{pour } j > 1 \\ 0 & \text{pour } j = 1 \end{cases}$$

Il est clair maintenant (voir problème 10.13 (4)) que T aura la forme désirée si les $v(i, j)$ sont ordonnés par l'ordre lexicographique ; en commençant par $v(1, 1)$ et décrivant la première colonne jusqu'à $v(1, k)$; puis en passant à $v(2, 1)$ et en décrivant la seconde colonne, etc.

De plus, nous aurons exactement

	$m_k - m_{k-1}$	éléments diagonaux d'ordre k
$(m_{k-1} - m_{k-2}) - (m_k - m_{k-1})$	$2m_{k-1} - m_k - m_{k-2}$	éléments diagonaux d'ordre $k-1$
	....	....
	$2m_2 - m_1 - m_3$	éléments diagonaux d'ordre 2
	$2m_1 - m_2$	éléments diagonaux d'ordre 1

comme nous pouvons le lire directement sur le tableau précédent. En particulier, puisque les nombres $m_1, \ldots, m_k$ sont uniquement déterminés par T, le nombre des éléments diagonaux de chaque ordre est uniquement déterminé par T. Finalement l'identité

$$m_1 \;=\; (m_k - m_{k-1}) + (2m_{k-1} - m_k - m_{k-2}) + \cdots + (2m_2 - m_1 - m_3) + (2m_1 - m_2)$$

montre que la nullité m_1 de T est le nombre total des éléments diagonaux de T.

10.17. Soit $A = \begin{pmatrix} 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$. Alors $A^2 = \begin{pmatrix} 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$ et $A^3 = 0$;

donc A est nilpotente d'indice 2. Trouver la matrice nilpotente M dans la forme canonique qui est semblable à A

Puisque A est nilpotente et d'indice 2, M contient un bloc diagonal d'ordre 2 et aucun d'ordre supérieur à 2. Remarquons que rang $A = 2$; donc la nullité de $A = 5 - 2 = 3$. Donc M contient 3 blocs diagonaux. En conséquence M doit contenir 2 blocs diagonaux d'ordre 2, et 1 bloc d'ordre 1, c'est-à-dire

$$M = \left(\begin{array}{cc|cc|c} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

10.18. Démontrer le théorème 10.11, page 226, sur la forme canonique de Jordan d'un opérateur T.

D'après le théorème de la décomposition primaire, T est décomposable en opérateurs $T_1, T_2 \dots T_r$, c'est-à-dire $T = T_1 \oplus \dots \oplus T_r$ où $(t - \lambda_i)^{m_i}$ est le polynôme minimal de T_i. Ainsi en particulier

$$(T_1 - \lambda_1 I)^{m_1} = 0, \quad \dots, \quad (T_r - \lambda_r I)^{m_r} = 0$$

Posons $N_i = T_i - \lambda_i I$. Alors pour $i = 1, \dots, r$

$$T_i = N_i + \lambda_i I, \qquad \text{où} \qquad N_i^{m_i} = 0$$

c'est-à-dire que T_i est la somme des opérateurs scalaires $\lambda_i I$ et d'un opérateur nilpotent N_i, qui est d'indice m_i, puisque $(t - \lambda_i)^{m_i}$ est le polynôme minimal de T_i.

D'après le théorème 10.10 sur les opérateurs nilpotents, nous pouvons choisir une base de sorte que N_i soit sous sa forme canonique. Dans cette base $T_i = N_i + \lambda_i I$ est représentée par une matrice diagonale bloc M_i dont les éléments diagonaux sont les matrices J_{ij}. La somme directe J des matrices M_i est sous la forme canonique de Jordan et, d'après le théorème 10.5, est une représentation matricielle de T.

Nous montrons maintenant que les blocs J_{ij} satisfont les propriétés demandées. La propriété (1) découle du fait que N_i est d'indice m_i. La propriété (2) est vraie car T et J ont le même polynôme caractéristique. La propriété (3) est vraie puisque la nullité de $N_i = T_i - \lambda_i I$ est égale à la multiplicité géométrique de la valeur propre λ_i. La propriété (4) découle du fait que T_i et donc N_i sont uniquement déterminés par T.

10.19. Déterminer toutes les formes canoniques de Jordan possibles pour un opérateur linéaire $T : V \to V$ dont le polynôme caractéristique est $\Delta(t) = (t - 2)^3 (t - 5)^2$.

Puisque $t - 2$ a pour exposant 3 dans $\Delta(t)$, 2 doit donc apparaître trois fois dans la diagonale principale. De façon analogue 5 doit apparaître deux fois. Ainsi les formes canoniques de Jordan possibles sont

$$\left(\begin{array}{ccc|cc} 2 & 1 & & & \\ & 2 & 1 & & \\ & & 2 & & \\ \hline & & & 5 & 1 \\ & & & & 5 \end{array}\right) \qquad \left(\begin{array}{cc|c|cc} 2 & 1 & & & \\ & 2 & & & \\ \hline & & 2 & & \\ \hline & & & 5 & 1 \\ & & & & 5 \end{array}\right) \qquad \left(\begin{array}{c|c|c|cc} 2 & & & & \\ \hline & 2 & & & \\ \hline & & 2 & & \\ \hline & & & 5 & 1 \\ & & & & 5 \end{array}\right)$$

(i) (ii) (iii)

$$\left(\begin{array}{ccc|c|c} 2 & 1 & & & \\ & 2 & 1 & & \\ & & 2 & & \\ \hline & & & 5 & \\ \hline & & & & 5 \end{array}\right) \qquad \left(\begin{array}{cc|c|c|c} 2 & 1 & & & \\ & 2 & & & \\ \hline & & 2 & & \\ \hline & & & 5 & \\ \hline & & & & 5 \end{array}\right) \qquad \left(\begin{array}{c|c|c|c|c} 2 & & & & \\ \hline & 2 & & & \\ \hline & & 2 & & \\ \hline & & & 5 & \\ \hline & & & & 5 \end{array}\right)$$

(iv) (v) (vi)

10.20. Déterminer toutes les formes canoniques de Jordan possibles J pour une matrice d'ordre 5 dont le polynôme minimal est $m(t) = (t-2)^2$.

J doit avoir une matrice bloc de Jordan d'ordre 2 et les autres doivent être d'ordre 2 ou 1. Ainsi il y a donc seulement deux possiblités :

$$J = \begin{pmatrix} 2 & 1 & & & \\ & 2 & & & \\ & & 2 & 1 & \\ & & & 2 & \\ & & & & 2 \end{pmatrix} \quad \text{ou} \quad J = \begin{pmatrix} 2 & 1 & & & \\ & 2 & & & \\ & & 2 & & \\ & & & 2 & \\ & & & & 2 \end{pmatrix}$$

Remarquons que tous les éléments diagonaux doivent être 2, puisque 2 est la seule valeur propre.

ESPACE QUOTIENT ET FORME TRIANGULAIRE

10.21. Soit W un sous-espace d'un espace vectoriel V. Montrer que les résultats suivants sont équivalents :

(1) $u \in v + W$ (2) $u - v \in W$ (3) $v \in u + W$

Supposons $u \in v + W$. Alors il existe $w_0 \in W$ tel que $u = v + w_0$. Donc $u - v = w_0 \in W$. Réciproquement, supposons $u - v \in W$. Alors $u - v = w_0$ où $w_0 \in W$. Donc $u = v + w_0 \in v + W$. Donc (1) et (2) sont équivalents.

Nous avons aussi $u - v \in W$ ssi $-(u - v) = v - u \in W$ ssi $v \in u + W$. Donc (2) et (3) sont aussi équivalents.

10.22. Démontrer : Les classes de W dans V réalisent une partition de V en ensembles mutuellement disjoints, c'est-à-dire

(1) deux classes quelconques $u + W$ et $v + W$ sont soit identiques, soit disjointes, et

(2) chaque $v \in V$ appartient à une classe ; en fait $v \in v + W$.

De plus $u + W = v + W$ si et seulement si $u - v \in W$ et donc $(v + w) + W = v + W$ pour un $w \in W$ quelconque.

Soit $v \in V$. Puisque $0 \in W$, nous avons $v = v + 0 \in v + W$ qui prouve (2).

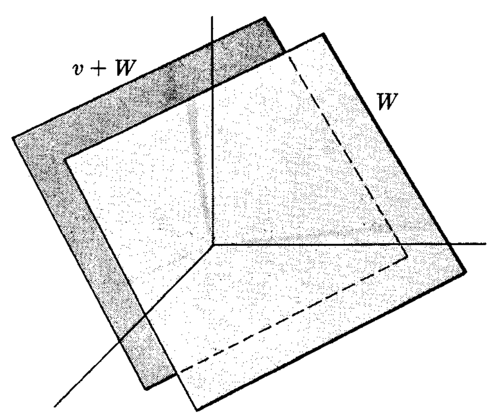
Supposons maintenant que les classes $u + W$ et $v + W$ ne soient pas disjointes c'est-à-dire que le vecteur x appartient à la fois aux deux classes $u + W$ et $v + W$. Alors $u - x \in W$ et $x - v \in W$. La démonstration de (1) sera complète si nous montrons que $u + W = v + W$. Soit $u + w_0$ un élément quelconque de la classe $u + W$. Puisque $u - x$, $x - v$ et w_0 appartiennent à W,

$$(u + w_0) - v = (u - x) + (x - v) + w_0 \in W$$

Ainsi $u + w_0 \in w + W$ et dont la classe $u + W$ est contenue dans la classe $v + W$. De façon analogue $v + W$ est contenu dans $u + W$ et donc $u + W = v + W$.

Le dernier résultat découle du fait que $u + W = v + W$ si et seulement si $u \in v + W$ et d'après le problème précédent ceci est équivalent à $u - v \in W$.

10.23. Soit W l'espace solution de l'équation homogène $2x + 3y + 4z = 0$. Donner les classes de W dans $\mathbf{R}^3$.



W est un plan passant par l'origine $O = (0, 0, 0)$ et les classes de W sont les plans parallèles à W. De façon équivalente, les classes de W sont les ensembles solutions de la famille d'équations

$$2x + 3y + 4z = k \qquad k \in \mathbf{R}$$

En particulier la classe $v + W$ où $v = (a, b, c)$ est l'ensemble solution de l'équation linéaire

$$2x + 3y + 4z = 2a + 3b + 4c$$

ou $\quad 2(x - a) + 3(y - b) + 4(z - c) = 0$

10.24. Supposons que W soit un sous-espace d'un espace vectoriel V. Montrer que les opérations dans le théorème 10.15 page 229 sont bien définies ; montrer donc que si $u + W = u' + W$ et $v + W = v' + W$ alors

(1) $(u + v) + W = (u' + v') + W$ et (2) $ku + W = ku' + W$ pour un $k \in K$ quelconque.

(1) Puisque $u + W = u' + W$ et $v + W = v' + W$, à la fois $u - u'$ et $v - v'$ appartiennent à W. Mais alors $(u + v) - (u' + v') = (u - u') + (v - v') \in W$. Donc $(u + v) + W = (u' + v') + W$.

(2) Puisque aussi $u - u' \in W$ implique $k(u - u') \in W$ alors $ku - ku' = k(u - u') \in W$ donc $ku + W = ku' + W$.

10.25. Soit V un espace vectoriel et W un sous-espace de V. Montrer que l'application $\eta : V \to V/W$ définie par $\eta(v) = v + W$ est linéaire.

Pour tout $u, v \in V$ et pour tout $k \in K$, nous avons

$$\eta(u + v) = u + v + W = u + W + v + W = \eta(u) + \eta(v)$$

et

$$\eta(kv) = kv + W = k(v + W) = k\,\eta(v)$$

En conséquence, η est linéaire.

10.26. Soit W un sous-espace d'un espace vectoriel V. Supposons que $\{w_1, \ldots, w_r\}$ soit une base de W, et l'ensemble des classes $\{\bar{v}_1, \ldots, \bar{v}_s\}$ où $\bar{v}_j = v_j + W$ soit une base de l'espace quotient. Montrer que $B = \{v_1, \ldots, v_s, w_1, \ldots, w_r\}$, est une base de V. Ainsi $\dim V = \dim W + \dim (V/W)$.

Supposons que $u \in V$. Puisque $\{\bar{v}_j\}$ est une base de V/W.

$$\bar{u} = u + W = a_1\bar{v}_1 + a_2\bar{v}_2 + \cdots + a_s\bar{v}_s$$

Donc $u = a_1 v_1 + \cdots + a_s v_s + w$ où $w \in W$. Puisque $\{w_i\}$ est une base de W,

$$u = a_1v_1 + \cdots + a_sv_s + b_1w_1 + \cdots + b_rw_r$$

En conséquence B engendre V.

Montrons maintenant que B est linéairement indépendant. Supposons

$$c_1v_1 + \cdots + c_sv_s + d_1w_1 + \cdots + d_rw_r = 0 \qquad (1)$$

Alors

$$c_1\bar{v}_1 + \cdots + c_s\bar{v}_s = \bar{0} = W$$

Puisque $\{\bar{v}_j\}$ est indépendant les c sont tous nuls. En remplaçant dans (1) on trouve $d_1w_1 + \cdots + d_rw_r = 0$. Puisque $\{w_i\}$ est indépendant, les d sont tous nuls. Ainsi B est linéairement indépendant et donc constitue une base de V.

10.27. Démontrer le théorème 10.16 : Supposons que W soit un sous-espace invariant par rapport à l'opérateur linéaire $T : V \to V$. Alors T induit un opérateur linéaire $\bar{T}$ sur V/W défini par $\bar{T}(v + W) = T(v) + W$. De plus, si T est un zéro d'un polynôme quelconque, il en est de même pour $\bar{T}$. Ainsi le polynôme minimal de $\bar{T}$ divise le polynôme minimal de T.

Montrons d'abord que $\bar{T}$ est bien défini, c'est-à-dire que si $u + W = v + W$ alors $\bar{T}(u + W) = \bar{T}(v + W)$. Si $u + W = v + W$ alors $u - v \in W$ et puisque W est invariant par T, $T(u - v) = T(u) - T(v) \in W$. En conséquence

$$\bar{T}(u + W) = T(u) + W = T(v) + W = \bar{T}(v + W)$$

comme demandé.

Montrons ensuite que $\bar{T}$ est linéaire. Nous avons

$$\begin{aligned}\bar{T}((u + W) + (v + W)) &= \bar{T}(u + v + W) = T(u + v) + W = T(u) + T(v) + W \\ &= T(u) + W + T(v) + W = \bar{T}(u + W) + \bar{T}(v + W)\end{aligned}$$

et

$$\bar{T}(k(u + W)) = \bar{T}(ku + W) = T(ku) + W = kT(u) + W = k(T(u) + W) = k\,\bar{T}(u + W)$$

Ainsi $\bar{T}$ est linéaire.

Pour une classe quelconque $u + W$ dans V/W,

$$\overline{T^2}(u+W) = T^2(u) + W = T(T(u)) + W = \bar{T}(T(u)+W) = \bar{T}(\bar{T}(u+W)) = \bar{T}^2(u+W)$$

Donc $\overline{T^2} = \bar{T}^2$. De manière analogue $\overline{T^n} = \bar{T}^n$ quel que soit n. Donc pour un polynôme quelconque,

$$f(t) = a_n t^n + \cdots + a_0 = \sum a_i t^i,$$

$$\begin{aligned}\overline{f(T)}(u+W) &= f(T)(u) + W = \sum a_i T^i(u) + W = \sum a_i (T^i(u) + W) \\ &= \sum a_i \overline{T^i}(u+W) = \sum a_i \bar{T}^i(u+W) = \left(\sum a_i \bar{T}^i\right)(u+W) = f(\bar{T})(u+W)\end{aligned}$$

et donc $\overline{f(T)} = f(\bar{T})$. En conséquence, si T est une racine de $f(t)$ alors $\overline{f(T)} = \bar{0} = W = f(\bar{T})$ c'est-à-dire que $\bar{T}$ est aussi une racine de $f(t)$. Ainsi le théorème est démontré.

10.28. Démontrer le théorème 10.1 : Soit $T : V \to V$ un opérateur linéaire dont les facteurs du polynôme caractéristique sont des polynômes linéaires. Alors V admet une base dans laquelle T est représentée par une matrice triangulaire.

La démonstration se fait par récurrence sur la dimension de V. Si $\dim V = 1$ alors toute représentation matricielle de T est une matrice 1×1 qui est triangulaire.

Supposons maintenant que $\dim V = n > 1$ et que le théorème soit vérifié pour des espaces de dimension inférieure à n. Puisque le polynôme caractéristique de T se décompose en produit de polynômes linéaires, T a au moins une valeur propre et donc au moins un vecteur propre non nul v, par exemple $T(v) = a_{11} v$. Soit W le sous-espace à une dimension engendré par v. Posons $\bar{V} = V/W$. Donc (Problème 10.26) $\dim \bar{V} = \dim V - \dim W = n - 1$. Remarquons aussi que W est invariant par T. D'après le théorème 10.16, T induit un opérateur linéaire $\bar{T}$ sur $\bar{V}$ dont le polynôme minimal divise le polynôme minimal de T. Puisque le polynôme caractéristique de T est un produit de polynômes linéaires, donc est son polynôme minimal ; et donc sont aussi les polynômes minimal et caractéristique de $\bar{T}$. Ainsi $\bar{V}$ et $\bar{T}$ satisfont les hypothèses du théorème. Donc par récurrence, il existe une base $\{\bar{v}_2, \bar{v}_3, \ldots, \bar{v}_n\}$ de $\bar{V}$ telle que

$$\begin{aligned}\bar{T}(\bar{v}_2) &= a_{22}\bar{v}_2 \\ \bar{T}(\bar{v}_3) &= a_{32}\bar{v}_2 + a_{33}\bar{v}_3 \\ &\cdots\cdots\cdots\cdots \\ \bar{T}(\bar{v}_n) &= a_{n2}\bar{v}_2 + a_{n3}\bar{v}_3 + \cdots + a_{nn}\bar{v}_n\end{aligned}$$

Soient maintenant $v_2, \ldots, v_n$ des éléments de V qui appartiennent respectivement aux classes $\bar{v}_2, \ldots, \bar{v}_n$. Alors $\{v, v_2, \ldots, v_n\}$ est une base de V (Problème 10.26). Puisque $\bar{T}(\bar{v}_2) = a_{22}\bar{v}_2$, nous avons

$$\bar{T}(\bar{v}_2) - a_{22}\bar{v}_2 = 0 \quad \text{et aussi} \quad T(v_2) - a_{22}v_2 \in W$$

Mais W est engendré par v ; donc $T(v_2) - a_{22}v_2$ est un multiple de v, c'est-à-dire

$$T(v_2) - a_{22}v_2 = a_{21}v \quad \text{et aussi} \quad T(v_2) = a_{21}v + a_{22}v_2$$

De façon analogue, pour $i = 3, \ldots, n$.

$$T(v_i) - a_{i2}v_2 - a_{i3}v_3 - \cdots - a_{ii}v_i \in W \quad \text{et aussi} \quad T(v_i) = a_{i1}v + a_{i2}v_2 + \cdots + a_{ii}v_i$$

Ainsi

$$\begin{aligned}T(v) &= a_{11}v \\ T(v_2) &= a_{21}v + a_{22}v_2 \\ &\cdots\cdots\cdots\cdots \\ T(v_n) &= a_{n1}v + a_{n2}v_2 + \cdots + a_{nn}v_n\end{aligned}$$

et donc la matrice de T est triangulaire dans cette base.

SOUS-ESPACES CYCLIQUES, FORMES RATIONNELLES CANONIQUES

10.29. Démontrer le théorème 10.12 : Soit $Z(v, T)$, un sous-espace T-cyclique, T_v la restriction de T à $Z(v, T)$ et $m_v(t) = t^k + a_{k-1}t^{k-1} + \ldots + a_0$ le T-annihilateur de v. Alors

(i) L'ensemble $\{v, T(v), \ldots, T^{k-1}(v)\}$ est une base de $Z(v, T)$; donc $\dim Z(v, T) = k$.

(ii) Le polynôme minimal de T_v est $m_v(t)$.

(iii) La matrice de T_v dans la base précédente est

$$C = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & 0 & \dots & 0 & -a_1 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & -a_{k-2} \\ 0 & 0 & 0 & \dots & 1 & -a_{k-1} \end{pmatrix}$$

(i) D'après la définition de $m_v(t)$, $T^k(v)$ est le premier vecteur de la suite v, $T(v)$, $T^2(v), \dots$ qui est une combinaison linéaire des vecteurs qui le précèdent dans la suite ; donc l'ensemble $B = \{v, T(v), \dots, T^{k-1}(v)\}$ est linéairement indépendant. Il nous reste uniquement à montrer que $Z(v, T) = L(B)$ est le sous-espace engendré par B. D'après ce qui précède, $T^k(v) \in L(B)$. Démontrons par récurrence que $T^n(v) \in L(B)$ pour tout n. Supposons $n > k$ et $T^{n-1}(v) \in L(B)$ c'est-à-dire $T^{n-1}(v)$ est une combinaison linéaire de $v, \dots, T^{k-1}(v)$. Alors $T^n(v) = T(T^{n-1}(v))$ est une combinaison linéaire de $T(v), \dots, T^{k-1}(v)$. Mais $T^k(v) \in L(B)$; donc $T^n(v) \in L(B)$, pour tout n. En conséquence $f(T)(v) \in L(B)$ pour tout polynôme $f(t)$. Ainsi $Z(v, T) = L(B)$ et donc B est aussi la base demandée comme nous l'avions affirmé.

(ii) Supposons $m(t) = t^s + b_{s-1}t^{s-1} + \dots + b_0$ soit le polynôme minimal de T_v. Alors puisque $v \in Z(v, T)$

$$0 = m(T_v)(v) = m(T)(v) = T^s(v) + b_{s-1}T^{s-1}(v) + \dots + b_0 v$$

Ainsi $T^s(v)$ est une combinaison linéaire de v, $T(v), \dots, T^{s-1}(v)$ et donc $k \leqslant s$. Cependant $m_v(T)v = 0$ et aussi $m_v(T_v) = 0$. Alors $m(t)$ divise $m_v(t)$ et donc $s \leqslant k$. En conséquence $k = s$ et donc $m_v(t) = m(t)$.

(iii)

$$\begin{aligned}
T_v(v) &= T(v) \\
T_v(T(v)) &= T^2(v) \\
&\dots\dots\dots \\
T_v(T^{k-2}(v)) &= T^{k-1}(v) \\
T_v(T^{k-1}(v)) &= T^k(v) = -a_0 v - a_1 T(v) - a_2 T^2(v) - \dots - a_{k-1}T^{k-1}(v)
\end{aligned}$$

D'après la définition, la matrice de T_v dans cette base est la transposée de la matrice des coefficients du système précédent d'équations ; donc il s'agit de C, comme il était demandé.

10.30. Soit $T : V \to V$ linéaire. Soit W un sous-espace invariant par T de V et $\overline{T}$ l'opérateur induit dans V/W. Démontrer (i) Le T-annihilateur de $v \in V$ divise le polynôme minimal de T (ii) Le $\overline{T}$-annihilateur de $\overline{v} \in V/W$ divise le polynôme minimal de T.

(i) Le T-annihilateur de $v \in V$ est le polynôme minimal de la restriction de T à $Z(v, T)$ et donc, d'après le Problème 10.6, il divise le polynôme minimal de T.

(ii) Le $\overline{T}$-annihilateur de $\overline{v} \in V/W$ divise le polynôme minimal de $\overline{T}$, qui divise le polynôme minimal de T d'après le théorème 10.16.

Remarque. Dans le cas où le polynôme minimal de T est $f(t)^n$ où $f(t)$ est un polynôme normalisé irréductible, alors le T-annihilateur de $v \in V$ et le T-annihilateur de $\overline{v} \in V/W$ sont de la forme $f(t)^m$ où $m \leqslant n$.

10-31. Démontrer le lemme 10-13 : Soit $T : V \to V$ un opérateur linéaire dont le polynôme minimal est $f(t)^n$ où $f(t)$ est un polynôme normalisé irréductible. Alors V est la somme directe des sous-espaces T cycliques $Z_i = Z(v_i, T)$, $i = 1, \dots, r$ avec les T-annihilateurs correspondants

$$f(t)^{n_1},\ f(t)^{n_2},\ \dots,\ f(t)^{n_r}, \quad n = n_1 \geqq n_2 \geqq \dots \geqq n_r$$

Toute autre décomposition de V en somme directe de sous-espaces T cycliques a le même nombre de composantes et le même ensemble de T-annihilateurs.

La démonstration se fait par récurrence sur la dimension de V. Si $\dim V = 1$, alors V est lui-même T-cyclique et le lemme est vrai. Supposons maintenant $\dim V > 1$ et que le lemme soit vrai pour les espaces vectoriels de dimension inférieure à celle de V.

Puisque le polynôme minimal de T est $f(t)^n$, il existe $v_1 \in V$ tel que $f(T)^{n-1}(v_1) \neq 0$; donc le T-annihilateur de v_1 est $f(t)^n$. Soit $Z_1 = Z(v_1, T)$ et rappelons que Z_1 est invariant par T. Soit $\bar{V} = V/Z_1$ et soit $\hat{T}$ l'opérateur linéaire sur $\bar{V}$ induit par T. D'après le théorème 10.16, le polynôme minimal de $\bar{T}$ divise $f(t)^n$; donc l'hypothèse est vérifiée pour $\bar{V}$ et $\bar{T}$. En conséquence, par récurrence, $\bar{V}$ est la somme directe des $\bar{T}$ sous-espaces cycliques, c'est-à-dire

$$\bar{V} = Z(\bar{v}_2, \bar{T}) \oplus \cdots \oplus Z(\bar{v}_r, \bar{T})$$

où les $\bar{T}$ annihilateurs correspondants sont $f(t)^{n_2}, \ldots, f(t)^{n_r}$, $n \geq n_2 \geq \cdots \geq n_r$.

Nous affirmons qu'il y a un vecteur v_2 dans la classe $\bar{v}_2$ dont le T-annihilateur est $f(t)^{n_2}$, le $\bar{T}$-annihilateur de $\bar{v}_2$. Alors $f(T)^{n_2}(w) \in Z_1$. Donc il existe un polynôme $g(t)$ pour lequel

$$f(T)^{n_2}(w) = g(T)(v_1) \tag{1}$$

Puisque $f(t)^n$ est le polynôme minimal de T, nous avons d'après (1),

$$0 = f(T)^n(w) = f(T)^{n-n_2} g(T)(v_1)$$

Mais $f(t)^n$ est le T-annihilateur de v_1 ; donc $f(t)^n$ divise $f(t)^{n-n_2} g(t)$ et donc $g(t) = f(t)^{n_2} h(t)$ pour un certain polynôme $h(t)$. Nous posons

$$v_2 = w - h(T)(v_1)$$

Puisque $w - v_2 = h(T)(v_1) \in Z_1$, v_2 appartient aussi à la classe $\bar{v}_2$. Ainsi le T-annihilateur de v_2 est un multiple du $\bar{T}$-annihilateur de $\bar{v}_2$. D'autre part, d'après (1)

$$f(T)^{n_2}(v_2) = f(T)^{n_2}(w - h(T)(v_1)) = f(T)^{n_2}(w) - g(T)(v_1) = 0$$

En conséquence le T-annihilateur de v_2 est $f(t)^{n_2}$ comme nous l'affirmions.

De façon analogue, il existe des vecteurs $v_3, \ldots, v_r \in V$ tels que $v_i \in \bar{v}_i$ et tel que le T-annihilateur de v_i est $f(t)^{n_i}$, le $\bar{T}$-annihilateur de $\bar{v}_i$. Nous posons

$$Z_2 = Z(v_2, T), \ \ldots, \ Z_r = Z(v_r, T)$$

Soit d le degré de $f(t)$ et donc le degré de $f(t)^{n_i}$ sera dn_i. Alors puisque $f(t)^{n_i}$ est à la fois le T-annihilateur de v_i et le $\bar{T}$-annihilateur de $\bar{v}_i$, nous savons que

$$\{v_i, T(v_i), \ldots, T^{dn_i-1}(v_i)\} \quad \text{et} \quad \{\bar{v}_i, \bar{T}(\bar{v}_i), \ldots, \bar{T}^{dn_i-1}(\bar{v}_i)\}$$

sont des bases pour $Z(v_i, T)$ et $Z(\bar{v}_i, \bar{T})$ respectivement, pour $i = 2, \ldots, r$. Mais $\bar{V} = Z(\bar{v}_2, \bar{T}) \oplus \cdots \oplus Z(\bar{v}_r, \bar{T})$; donc

$$\{\bar{v}_2, \ldots, \bar{T}^{dn_2-1}(\bar{v}_2), \ldots, \bar{v}_r, \ldots, \bar{T}^{dn_r-1}(\bar{v}_r)\}$$

est une base pour $\bar{V}$. Donc d'après le problème 10.26 et la relation $\bar{T}^i(\bar{v}) = \overline{T^i(v)}$ (voir problème 10.27),

$$\{v_1, \ldots, T^{dn_1-1}(v_1), v_2, \ldots, T^{dn_2-1}(v_2), \ldots, v_r, \ldots, T^{dn_r-1}(v_r)\}$$

est une base pour V. Ainsi d'après le théorème 10.4 $V = Z(v_1, T) \oplus \cdots \oplus Z(v_r, T)$ comme prévu.

Il reste à montrer que les exposants $n_1, \ldots, n_r$ sont uniquement déterminés par T. Soit d le degré de $f(t)$,

$$\dim V = d(n_1 + \cdots + n_r) \quad \text{et} \quad \dim Z_i = dn_i, \quad i = 1, \ldots, r.$$

Ainsi, si s est un exposant quelconque entier positif (Problème 10.59) $f(T)^s(Z_i)$ est un sous-espace cyclique engendré par $f(T)^s(v_i)$ et il a pour dimension $d(n_i - s)$ si $n_i > s$ et pour dimension 0 si $n_i \leq s$.

Un vecteur quelconque $v \in V$ peut être écrit sous forme unique $v = w_1 + \cdots + w_r$ où $w_i \in Z_i$. Donc un vecteur quelconque de $f(T)^s(V)$ peut être écrit de manière unique sous la forme

$$f(T)^s(v) = f(T)^s(w_1) + \cdots + f(T)^s(w_r)$$

où $f(T)^s(w_i) \in f(T)^s(Z_i)$. Soit t l'exposant entier, dépendant de s, pour lequel

$$n_1 > s, \ \ldots, \ n_t > s, \ n_{t+1} \leq s$$

Alors
$$f(T)^s(V) = f(T)^s(Z_1) \oplus \cdots \oplus f(T)^s(Z_t)$$

et donc
$$\dim(f(T)^s(V)) = d[(n_1 - s) + \cdots + (n_t - s)] \tag{*}$$

Les nombres à gauche de (*) sont uniquement déterminés par T. Posons $s = n - 1$ et (*) le nombre des n égaux à n. Posons ensuite $s = n - 2$ (*) déterminant le nombre des n_i égaux à $n - 1$. Répétons ce raisonnement jusqu'à $s = 0$ et déterminons le nombre des n_i égaux à 1. Ainsi les n_i sont uniquement déterminés par T et V et le lemme est démontré.

10.32. Soit V un espace vectoriel de dimension 7 sur $\mathbf{R}$ et soit $T : V \to V$ un opérateur linéaire ayant pour polynôme minimal $m(t) = (t^2 + 2)\,(t + 3)^2$. Trouver toutes les formes canoniques rationnelles possibles pour T.

La somme des degrés des matrices compagnons doit donner 7. Ainsi une matrice compagnon doit être associée à $t^2 + 2$ et une doit l'être à $(t + 3)^3$. Ainsi la forme rationnelle canonique de T est exactement l'une des sommes directes suivantes des matrices compagnons suivantes :

(i) $C(t^2+2) \oplus C(t^2+2) \oplus C((t+3)^3)$

(ii) $C(t^2+2) \oplus C((t+3)^3) \oplus C((t+3)^2)$

(iii) $C(t^2+2) \oplus C((t+3)^3) \oplus C(t+3) \oplus C(t+3)$

c'est-à-dire

$$\begin{pmatrix} 0 & -2 & & & & & \\ 1 & 0 & & & & & \\ & & 0 & -2 & & & \\ & & 1 & 0 & & & \\ & & & & 0 & 0 & -27 \\ & & & & 1 & 0 & -27 \\ & & & & 0 & 1 & -9 \end{pmatrix} \qquad \begin{pmatrix} 0 & -2 & & & & & \\ 1 & 0 & & & & & \\ & & 0 & 0 & -27 & & \\ & & 1 & 0 & -27 & & \\ & & 0 & 1 & -9 & & \\ & & & & & 0 & -9 \\ & & & & & 1 & -6 \end{pmatrix} \qquad \begin{pmatrix} 0 & -2 & & & & & \\ 1 & 0 & & & & & \\ & & 0 & 0 & -27 & & \\ & & 1 & 0 & -27 & & \\ & & 0 & 1 & -9 & & \\ & & & & & -3 & \\ & & & & & & -3 \end{pmatrix}$$

(i) (ii) (iii)

PROJECTIONS

10.33. Supposons $V = W_1 \oplus \cdots \oplus W_r$. La projection de V dans son sous-espace W_k est l'application $E : V \to V$ définie par $E(v) = w_k$ où $v = w_1 + \cdots + w_r$, $w_i \in W_i$. Montrer que (i) E est linéaire (ii) $E^2 = E$.

(i) Puisque la somme $v = w_1 + \cdots + w_r$, $w_i \in W$ est uniquement déterminé par v, l'application E est bien définie. Supposons que pour $u \in V$, $u = w'_1 + \cdots + w'_r$, $w'_i \in W_i$. Alors

$$v + u = (w_1 + w'_1) + \cdots + (w_r + w'_r) \quad \text{et} \quad kv = kw_1 + \cdots + kw_r, \qquad kw_i,\ w_i + w'_i \in W_i$$

sont les sommes uniques correspondant à $v + u$ et kv. Donc

$$E(v+u) = w_k + w'_k = E(v) + E(u) \quad \text{et} \quad E(kv) = kw_k = kE(v)$$

et donc E est linéaire.

(ii) Nous savons que $\qquad w_k = 0 + \cdots + 0 + w_k + 0 + \cdots + 0$

est la somme unique correspondant à $w_k \in W_k$; donc $E(w_k) = w_k$. Alors pour tout $v \in V$

$$E^2(v) = E(E(v)) = E(w_k) = w_k = E(v)$$

Ainsi $E^2 = E$, comme demandé.

10.34. Supposons $E : V \to V$ linéaire et $E^2 = E$. Montrer que (i) $E(u) = u$ pour tout $u \in \operatorname{Im} E$, c'est-à-dire la restriction à E de son image est l'application identique, (ii) V est la somme directe de l'image et du noyau de E : $V = \operatorname{Im} E \oplus \operatorname{Ker} E$; (iii) E est la projection de V sur $\operatorname{Im} E$, son image. Ainsi, d'après le problème précédent, une application linéaire $T : V \to V$ est une projection si et seulement si $T^2 = T$; cette caractérisation d'une projection est fréquemment utilisée comme définition.

(i) Si $u \in \operatorname{Im} E$, alors il existe $v \in V$ pour lequel $E(v) = u$; donc

$$E(u) = E(E(v)) = E^2(v) = E(v) = u$$

comme demandé.

(ii) Si $v \in V$. Nous pouvons écrire v sous la forme $v = E(v) + v - E(v)$. Maintenant $E(v) \in \operatorname{Im} E$ et puisque

$$E(v - E(v)) = E(v) - E^2(v) = E(v) - E(v) = 0$$

$v - E(v) \in \operatorname{Ker} E$. En conséquence $V = \operatorname{Im} E + \operatorname{Ker} E$.

Supposons maintenant $w \in \text{Im } E \cap \text{Ker } E$. D'après (i) $E(w) = w$ car $w \in \text{Im } E$. D'autre part $E(w) = 0$, car $W \in \text{Ker } E$. Ainsi $w = 0$ et donc $\text{Im } E \cap \text{Ker } E = \{0\}$. Ces deux conditions impliquent que V est la somme directe de l'image et du noyau de E.

(iii) Soit $v \in V$ est supposons $v = u + w$ où $u \in \text{Im } E$ et $w \in \text{Ker } E$. Remarquons que $E(u) = u$ d'après (i) et $E(w) = 0$ car $w \in \text{Ker } E$. Donc

$$E(v) = E(u+w) = E(u) + E(w) = u + 0 = u$$

C'est-à-dire, E est la projection de V sur son image.

10.35. Supposons $V = U \oplus W$ et supposons que $T : V \to V$ soit linéaire. Montrer que U et W sont tous deux invariants par T si et seulement si $TE = ET$ où E est la projection de V sur U.

Remarquons que $E(v) \in U$ pour tout $v \in V$ et que (i) $E(v) = v$ ssi $v \in U$ (ii) $E(v) = 0$ ssi $v \in W$.

Supposons $ET = TE$. Soit $u \in U$. Puisque $E(u) = u$

$$T(u) = T(E(u)) = (TE)(u) = (ET)(u) = E(T(u)) \in U$$

Donc U est invariant par T. Soit $w \in W$. Puisque $E(w) = 0$.

$$E(T(w)) = (ET)(w) = (TE)(w) = T(E(w)) = T(0) = 0 \quad \text{et ainsi} \quad T(w) \in W$$

Donc W est invariant par T.

Réciproquement supposons à la fois U et W tous deux invariants par T. Soit $v \in V$ et supposons $v = u + w$ où $u \in T$ et $w \in W$. Alors $T(u) \in U$ et $T(w) \in W$; donc $E(T(u)) = T(u)$ et $E(T(w)) = 0$. Ainsi

$$(ET)(v) = (ET)(u+w) = (ET)(u) + (ET)(w) = E(T(u)) + E(T(w)) = T(u)$$

et

$$(TE)(v) = (TE)(u+w) = T(E(u+w)) = T(u)$$

C'est-à-dire, $(ET)(v) = (TE)(v)$ pour tout $v \in V$; donc $ET = TE$ comme demandé.

PROBLEMES SUPPLEMENTAIRES

SOUS-ESPACES INVARIANTS

10.36. Supposons W invariant par $T : V \to V$. Montrer que W est invariant par $f(T)$ pour tout polynôme $f(t)$.

10.37. Montrer que chaque sous-espace de V est invariant par I et 0 ; qui sont les opérateurs identique et nul.

10.38. Supposons W invariant par $S : V \to V$ et $T : V \to V$. Montrer aussi que W est invariant par $S + T$ et ST.

10.39. Soit $T : V \to V$ une application linéaire et soit W le sous-espace propre correspondant à la valeur propre λ de T. Montrer que W est invariant par T.

10.40. Soit V un espace vectoriel de dimension impaire (supérieure à 1) sur le corps des réels $\mathbf{R}$. Montrer que tout opérateur linéaire sur V a un sous-espace invariant autre que V ou $\{0\}$.

10.41. Déterminer les sous-espaces invariants de $A = \begin{pmatrix} 2 & -4 \\ 5 & 2 \end{pmatrix}$ considéré comme un opérateur linéaire sur (i) $\mathbf{R}^2$ (ii) $\mathbf{C}^2$.

10.42. Supposons que $\dim V = n$. Montrer que $T : V \to V$ a une représentation matricielle triangulaire si et seulement si il existe des sous-espaces invariants par T, $W_1 \subset W_2 \subset \cdots \subset W_n = V$ pour lesquels $\dim W_k = k$, $k = 1, \ldots, n$.

SOMMES DIRECTES INVARIANTES

10.43. Les sous-espaces $W_1, \ldots, W_r$ sont dits indépendants si $w_1 + \cdots + w_r = 0$, $w_i \in W_i$ ce qui implique que chaque $w_i = 0$. Montrer que $L(W_i) = W_1 \oplus \cdots \oplus W_r$ si et seulement si les W_i sont indépendants. (Ici $L(W_i)$ indique l'ensemble des générateurs des W_i).

10.44. Montrer que $V = W_1 \oplus \cdots \oplus W_r$ si et seulement si (i) $W = L(W_i)$ et (ii) $W_k \cap L(W_1, \ldots, W_{k-1}, W_{k+1}, \ldots, W_r) = \{0\}$ $k = 1, \ldots, r$.

10.45. Montrer que $L(W_i) = W_1 \oplus \cdots \oplus W_r$ si et seulement si $\dim L(W_i) = \dim W_1 + \cdots + \dim W_r$.

10.46. Supposons que le polynôme caractéristique de $T : V \to V$ soit $\Delta(t) = f_1(t)^{n_1} f_2(t)^{n_2} \ldots f_r(t)^{n_r}$ où les $f_i(t)$ sont des polynômes normalisés distincts et irréductibles. Soit $V = W_1 \oplus \cdots \oplus W_r$ la décomposition primaire de V en sous-espaces invariants par T. Montrer que $f_i(t)^{n_i}$ est le polynôme caractéristique de la restriction de T à W_i.

NILPOTENTS OPERATEURS

10.47. Supposons que S et T soient des opérateurs nilpotents qui commutent c'est-à-dire tels que $ST = TS$. Montrer que $S + T$ et ST sont aussi nilpotents.

10.48. Supposons que A soit une matrice supertriangulaire, c'est-à-dire que tous les éléments sur et dessous la diagonale principale sont nuls. Montrer que A est nilpotent

10.49. Soit V l'espace vectoriel des polynômes de degré $\leqslant n$. Montrer que l'opérateur différentiel sur V est nilpotent et d'indice $n + 1$.

10.50. Montrer que les matrices nilpotentes suivantes d'ordre n sont semblables :

$$\begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & 0 \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} 0 & 0 & \ldots & 0 & 0 \\ 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 & 0 \end{pmatrix}$$

10.51. Montrer que deux matrices nilpotentes d'ordre 3 sont semblables si et seulement si elles ont même indice de nilpotence. Montrer par exemple que ce résultat n'est pas vrai pour des matrices nilpotentes d'ordre 4.

FORME CANONIQUE DE JORDAN

10.52. Trouver toutes les formes canoniques de Jordan possibles pour les matrices dont le polynôme caractéristique $\Delta(t)$ et le polynôme minimal $m(t)$ sont les suivants

(i) $\Delta(t) = (t-2)^4(t-3)^2, \quad m(t) = (t-2)^2(t-3)^2$
(ii) $\Delta(t) = (t-7)^5, \quad m(t) = (t-7)^2$
(iii) $\Delta(t) = (t-2)^7, \quad m(t) = (t-2)^3$
(iv) $\Delta(t) = (t-3)^4(t-5)^4, \quad m(t) = (t-3)^2(t-5)^2$

10.53. Montrer que toute matrice complexe est semblable à sa transposée. (Utiliser la forme canonique de Jordan et le problème 10-50).

10.54. Montrer que toutes les matrices complexes A d'ordre n pour lesquelles $A^n = I$ sont semblables.

10.55. Supposons que A soit une matrice complexe avec seulement des valeurs propres réelles. Montrer que A est semblable à une matrice avec seulement des éléments réels.

SOUS-ESPACES CYCLIQUES

10.56. Supposons que $T : V \to V$ soit linéaire. Démontrer que $Z(v, T)$ est l'intersection de tous les sous-espaces invariants par T contenant v.

10.57. Soient $f(t)$ et $g(t)$ les T-innihilateurs de u et v respectivement. Montrer que si $f(t)$ et $g(t)$ sont premiers entre eux, alors $f(t)\, g(t)$ est le T-annihilateur de $u + v$.

10.58. Démontrer que $Z(u, T) = Z(v, T)$ si et seulement si $g(T)\,(u) = v$ où $g(t)$ est premier par rapport au T-annihilateur de u.

10.59. Soit $W = Z(v, T)$ et supposons que le T-annihilateur de v soit $f(t)^n$ où $f(t)$ est un polynôme normalisé irréductible de degré d. Montrer que $f(T)^s\,(W)$ est un sous-espace cyclique engendré par $f(T)^s\,(v)$ et admet pour dimension $d\,(n - s)$ si $n > s$ et de dimension 0 si $n \leqslant s$.

FORMES RATIONNELLES CANONIQUES

10.60. Trouver toutes les formes rationnelles canoniques possibles pour

(i) des matrices 6 x 6 ayant pour polynôme minimal $m(t) = (t^2 + 3)\,(t + 1)^2$.
(ii) des matrices 6 x 6 ayant pour polynôme minimal $m(t) = (t + 1)^3$.
(iii) des matrices 8 x 8 ayant pour polynôme minimal $m(t) = (t^2 + 2)^2\,(t + 3)^2$.

10.61. Soit A une matrice 4 x 4 ayant pour polynôme minimal $m(t) = (t^2 + 1)\,(t^2 - 3)$. Trouver la forme rationnelle canonique de A si A est une matrice sur (i) le corps des rationnels $\mathbf{Q}$, (ii) le corps des réels $\mathbf{R}$, (iii) le corps des complexes $\mathbf{C}$.

10.62. Trouver la forme canonique rationnelle pour la matrice de Jordan $\begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix}$.

10.63. Démontrer que le polynôme caractéristique d'un opérateur $T : V \to V$ est un produit de ses diviseurs élémentaires.

10.64. Démontrer que deux matrices 3×3 ayant les mêmes polynômes caractéristique et minimal sont semblables.

10.65. Soit $C(f(t)$ la matrice compagnon d'un polynôme arbitraire $f(t)$. Montrer que $f(t)$ est le polynôme caractéristique de $C(f(t))$.

PROJECTIONS

10.66. Supposons $V = W_1 \oplus \cdots \oplus W_r$. Soit E_i la projection de V sur W_i. Démontrer (1) $E_iE_j = 0$, $i \neq j$;(2) $I = E_1 + \cdots + E_r$.

10.67. Soit $E_1, \ldots, E_r$ des opérateurs linéaires sur V tels que (1) $E_i^2 = E_i$ c'est-à-dire que les E_i sont des projections ; (2) $E_iE_j = 0, i \neq j$;(3) $I = E_1 + \cdots + E_r$. Démontrer que $V = \operatorname{Im} E_1 \oplus \cdots \oplus \operatorname{Im} E_r$.

10.68. Supposons que $E : V \to V$ soit une projection c'est-à-dire que $E^2 = E$. Démontrer que E a une représentation matricielle de la forme $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ où r est le rang de E et I_r la matrice $r \times r$ identique.

10.69. Démontrer que deux projections quelconques de même rang sont semblables. (Utiliser le résultat du problème 10.68).

10.70. Supposons que $E : V \to V$ soit une projection. Démontrer que

(1) $I - E$ est une projection et $V = \operatorname{Im} E \oplus \operatorname{Im}(I - E)$. (2) $I + E$ est inversible (si $1 + 1 \neq 0$).

ESPACES QUOTIENTS

10.71. Soit W un sous-espace de V. Supposons que l'ensemble des classes $\{v_1 + W, v_2 + W, \ldots, v_n + W\}$ dans V/W soit linéairement indépendant. Montrer que l'ensemble des vecteurs $\{v_1, v_2, \ldots, v_n\}$ dans V est aussi linéairement indépendant.

10.72. Soit W un sous-espace de V. Supposons que l'ensemble des vecteurs $\{u_1, u_2, \ldots, u_n\}$ dans V soit linéairement indépendant et que $L(u_i) \cap W = \{0\}$. Montrer que l'ensemble des classes $\{u_1 + W, \ldots, u_n + W\}$ dans V/W et aussi linéairement indépendant.

10.73. Supposons $V = U \oplus W$ et que $\{u_1, \ldots, u_n\}$ soit une base de U. Montrer que $\{u_1 + W, \ldots, u_n + W\}$ est une base de l'espace quotient V/W. (Remarquer qu'il n'existe aucune condition sur la dimension de V ou W).

10.74. Soit W l'espace solution de l'équation linéaire

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0, \qquad a_i \in K$$

et soit $v = (b_1, b_2, \ldots, b_n) \in K^n$. Démontrer que la classe $v + W$ de W dans K^n est l'ensemble solution de l'équation linéaire

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b \qquad \text{où} \qquad b = a_1b_1 + \cdots + a_nb_n$$

10.75. Soit V l'espace vectoriel des polynômes sur $\mathbf{R}$, et soit W le sous-espace des polynômes divisibles par t^4, c'est-à-dire de la forme $a_0t^4 + a_1t^5 + \ldots + a_{n-4}t^n$. Montrer que l'espace quotient V/W est de dimension 4.

10.76. Soit U et W les sous-espaces de V tels que $W \subset U \subset V$. Remarquons que toute classe $u + W$ de W dans U peut être aussi considérée comme une classe de W dans V puisque $u \in U$ implique $u \in V$; donc U/W est un sous-ensemble de V/W. Démontrer que (1) U/W est un sous-espace de V/W, (2) $\dim(V/W) - \dim(U/W) = \dim(V/U)$.

10.77. Soit U et W des sous-espaces de V. Montrer que les classes de $U \cap W$ dans V peuvent être obtenues par intersection de chacune des classes de U dans V par chacun des classes de W dans V:

$$V/(U \cap W) = \{(v + U) \cap (v' + W) : v, v' \in V\}$$

10.78. Soit $T : V \to V'$ une application linéaire ayant pour noyau W et pour image U. Montrer que l'espace quotient V/W est isomorphe à U par l'application $\theta : V/W \to U$ définie par $\theta(v + W) = T(v)$. De plus, montrer que $T = i \circ \theta \circ \eta$ où $\eta : V \to V/W$ est l'application naturelle de V sur V/W c'est-à-dire $\eta(v) = v + W$ et $i : U \subset V'$ est l'application $i(u) = u$ (voir le diagramme).

$$\begin{tikzcd} V \arrow[d, "\eta"'] \arrow[drr, "T"] & & \\ V/W \arrow[r, "\theta"'] & U \arrow[r, "i"'] & V' \end{tikzcd}$$

REPONSES AUX PROBLEMES SUPPLEMENTAIRES.

10.41. (i) $\mathbf{R}^2$ et $\{0\}$ (ii) $\mathbf{C}^2$, $\{0\}$, $W_1 = L((2, 1-2i))$, $W_2 = L((2, 1+2i))$.

10.42. (i)

$$\begin{pmatrix} 2&1&&&&\\ &2&&&&\\ &&2&1&&\\ &&&2&&\\ &&&&3&1\\ &&&&&3 \end{pmatrix}, \quad \begin{pmatrix} 2&1&&&&\\ &2&&&&\\ &&2&&&\\ &&&2&&\\ &&&&3&1\\ &&&&&3 \end{pmatrix}$$

(ii)

$$\begin{pmatrix} 7&1&&&\\ &7&&&\\ &&7&1&\\ &&&7&\\ &&&&7 \end{pmatrix}, \quad \begin{pmatrix} 7&1&&&\\ &7&&&\\ &&7&&\\ &&&7&\\ &&&&7 \end{pmatrix}$$

(iii)

$$\begin{pmatrix} 2&1&&&&&\\ &2&1&&&&\\ &&2&&&&\\ &&&2&1&&\\ &&&&2&1&\\ &&&&&2&\\ &&&&&&2 \end{pmatrix}, \quad \begin{pmatrix} 2&1&&&&&\\ &2&1&&&&\\ &&2&&&&\\ &&&2&1&&\\ &&&&2&&\\ &&&&&2&1\\ &&&&&&2 \end{pmatrix}$$

$$\begin{pmatrix} 2&1&&&&&\\ &2&1&&&&\\ &&2&&&&\\ &&&2&1&&\\ &&&&2&&\\ &&&&&2&\\ &&&&&&2 \end{pmatrix}, \quad \begin{pmatrix} 2&1&&&&&\\ &2&1&&&&\\ &&2&&&&\\ &&&2&&&\\ &&&&2&&\\ &&&&&2&\\ &&&&&&2 \end{pmatrix}$$

(iv)

$$\begin{pmatrix} 3&1&&&&&&\\ &3&&&&&&\\ &&3&1&&&&\\ &&&3&&&&\\ &&&&5&1&&\\ &&&&&5&&\\ &&&&&&5&1\\ &&&&&&&5 \end{pmatrix}, \quad \begin{pmatrix} 3&1&&&&&&\\ &3&&&&&&\\ &&3&1&&&&\\ &&&3&&&&\\ &&&&5&1&&\\ &&&&&5&&\\ &&&&&&5&\\ &&&&&&&5 \end{pmatrix},$$

$$\begin{pmatrix} 3&1&&&&&&&\\ &3&&&&&&&\\ &&3&&&&&&\\ &&&3&&&&&\\ &&&&5&1&&&\\ &&&&&5&&&\\ &&&&&&5&1&\\ &&&&&&&5& \end{pmatrix}, \quad \begin{pmatrix} 3&1&&&&&&&\\ &3&&&&&&&\\ &&3&&&&&&\\ &&&3&&&&&\\ &&&&5&1&&&\\ &&&&&5&&&\\ &&&&&&5&&\\ &&&&&&&5& \end{pmatrix}$$

10.60. (i)
$$\begin{pmatrix} 0 & -3 & & & & \\ 1 & 0 & & & & \\ & & 0 & -3 & & \\ & & 1 & 0 & & \\ & & & & 0 & -1 \\ & & & & 1 & -2 \end{pmatrix}, \quad \begin{pmatrix} 0 & -3 & & & & \\ 1 & 0 & & & & \\ & & 0 & -1 & & \\ & & 1 & -2 & & \\ & & & & 0 & -1 \\ & & & & 1 & -2 \end{pmatrix}, \quad \begin{pmatrix} 0 & -3 & & & & \\ 1 & 0 & & & & \\ & & 0 & -1 & & \\ & & 1 & -2 & & \\ & & & & -1 & \\ & & & & & -1 \end{pmatrix}$$

(ii)
$$\begin{pmatrix} 0 & 0 & -1 & & & \\ 1 & 0 & -3 & & & \\ 0 & 1 & -3 & & & \\ & & & 0 & 0 & -1 \\ & & & 1 & 0 & -3 \\ & & & 0 & 1 & -3 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & -1 & & & \\ 1 & 0 & -3 & & & \\ 0 & 1 & -3 & & & \\ & & & 0 & -1 & \\ & & & 1 & -2 & \\ & & & & & -1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & -1 & & & \\ 1 & 0 & -3 & & & \\ 0 & 1 & -3 & & & \\ & & & -1 & & \\ & & & & -1 & \\ & & & & & -1 \end{pmatrix}$$

(iii)
$$\begin{pmatrix} 0 & 0 & 0 & 2 & & & & \\ 1 & 0 & 0 & 0 & & & & \\ 0 & 1 & 0 & -4 & & & & \\ 0 & 0 & 1 & 0 & & & & \\ & & & & 0 & -2 & & \\ & & & & 1 & 0 & & \\ & & & & & & 0 & -9 \\ & & & & & & 1 & -6 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 & 2 & & & & \\ 1 & 0 & 0 & 0 & & & & \\ 0 & 1 & 0 & -4 & & & & \\ 0 & 0 & 1 & 0 & & & & \\ & & & & 0 & -9 & & \\ & & & & 1 & -6 & & \\ & & & & & & 0 & -9 \\ & & & & & & 1 & -6 \end{pmatrix},$$

$$\begin{pmatrix} 0 & 0 & 0 & 2 & & & & \\ 1 & 0 & 0 & 0 & & & & \\ 0 & 1 & 0 & -4 & & & & \\ 0 & 0 & 1 & 0 & & & & \\ & & & & 0 & -9 & & \\ & & & & 1 & -6 & & \\ & & & & & & -3 & \\ & & & & & & & -3 \end{pmatrix}$$

10.61. (i) $\begin{pmatrix} 0 & -1 & & \\ 1 & 0 & & \\ & & 0 & \sqrt{3} \\ & & 1 & 0 \end{pmatrix}$ (ii) $\begin{pmatrix} 0 & -1 & & \\ 1 & 0 & & \\ & & \sqrt{3} & \\ & & & -\sqrt{3} \end{pmatrix}$ (iii) $\begin{pmatrix} i & & & \\ & -i & & \\ & & \sqrt{3} & \\ & & & -\sqrt{3} \end{pmatrix}$

10.62.
$$\begin{pmatrix} 0 & 0 & 0 & -\lambda^4 \\ 1 & 0 & 0 & 4\lambda^3 \\ 0 & 1 & 0 & -6\lambda^2 \\ 0 & 0 & 1 & 4\lambda \end{pmatrix}$$

CHAPITRE 11

Formes linéaires et espace dual

INTRODUCTION

Dans ce chapitre nous étudierons les applications linéaires d'un espace vectoriel dans son corps des scalaires K (à moins qu'il en soit spécifié autrement, nous considérons K comme un espace vectoriel sur lui-même). Naturellement tous les théorèmes et les résultats obtenus pour des applications linéaires quelconques sur V restent vrais pour ce cas particulier. Cependant nous traiterons ces applications séparément à cause de leur importance fondamentale, et parce que la relation particulière entre V et K donne lieu à de nouvelles notions et résultats qui ne s'appliquent pas dans le cas général.

FORMES LINEAIRES ET ESPACE DUAL

Soit V un espace vectoriel sur le corps K. Une application $\phi : V \to K$ est appelée fonctionnelle linéaire ou forme linéaire si, pour tout $u, v \in V$ et pour tout $a, b \in K$

$$\phi(au + bv) = a\phi(u) + b\phi(v)$$

En d'autres termes, une forme linéaire ou fonctionnelle linéaire sur V est une application linéaire de V dans K.

Exemple 11.1 : Soit $\pi_i : K^n \to K$ l'application $i^{\text{ème}}$ projection, c'est-à-dire $\pi_i(a_1, a_2, \ldots, a_n) = a_i$. Alors π_i est linéaire et donc c'est une forme linéaire sur K^n.

Exemple 11.2 : Soit V l'espace vectoriel des polynômes en t sur $\mathbf{R}$. Soit $\mathscr{J} : V \to \mathbf{R}$ l'opérateur d'intégration défini par $\mathscr{J}(p(t)) = \int_0^1 p(t)\,dt$. Rappelons que $\mathscr{J}$ est linéaire ; et donc il s'agit d'une forme linéaire sur V.

Exemple 11.3 : Soit V l'espace vectoriel des matrices carrées $n \times n$ sur K. Soit $T : V \to K$ la trace :

$$T(A) = a_{11} + a_{22} + \cdots + a_{nn} \quad \text{où} \quad A = (a_{ij})$$

C'est-à-dire, T associe à une matrice A la somme de ses éléments diagonaux. Cette application est linéaire. (Problème 11.27) et donc il s'agit d'une forme linéaire sur V.

D'après le théorème 6-6, l'ensemble des formes linéaires sur un espace vectoriel V sur un corps K est aussi un espace vectoriel sur K avec l'addition et la multiplication par un scalaire définis par

$$(\phi + \sigma)(v) = \phi(v) + \sigma(v) \qquad \text{et} \qquad (k\phi)(v) = k\phi(v)$$

où ϕ et σ sont des formes linéaires sur V et $k \in K$. Cet espace est appelé l'espace dual de V et on le note V^*.

Exemple 11.4 : Soit $V = K^n$, l'espace vectoriel des n-tuples que nous écrivons comme des vecteurs colonnes. Alors l'espace dual V^* peut être identifié avec l'espace des vecteurs lignes. En particulier, toute forme linéaire $\phi = (a_1, \ldots, a_n)$ dans V^* a pour représentation

$$\phi(x_1, \ldots, x_n) = (a_1, a_2, \ldots, a_n)\begin{pmatrix} x_1 \\ x_2 \\ \cdot\cdot \\ x_n \end{pmatrix}$$

ou simplement

$$\phi(x_1, \ldots, x_n) = a_1x_1 + a_2x_2 + \cdots + a_nx_n$$

Historiquement, la forme précédente était appelée forme linéaire.

BASE DUALE

Supposons V un espace vectoriel de dimension n sur K. D'après le théorème 6.7, la dimension de l'espace dual V^* est aussi n (puisque K est de dimension 1 sur lui-même). En fait, chaque base de V détermine une base de V^* comme suit :

Théorème 11.1 : Supposons que $\{v_1, \ldots, v_n\}$ soit une base de V sur K. Soient $\phi_1, \ldots, \phi_n \in V^*$ les formes linéaires définies par

$$\phi_i(v_j) = \delta_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

Alors $\{\phi_1, \ldots, \phi_n\}$ est une base de V^*.

La base précédente $\{\phi_i\}$ est appelée la base duale de $\{v_i\}$. Les formules précédentes utilisant le symbole de Kronecker δ_{ij} sont une écriture contractée pour

$$\begin{array}{l} \phi_1(v_1) = 1,\ \phi_1(v_2) = 0,\ \phi_1(v_3) = 0,\ \ldots,\ \phi_1(v_n) = 0 \\ \phi_2(v_1) = 0,\ \phi_2(v_2) = 1,\ \phi_2(v_3) = 0,\ \ldots,\ \phi_2(v_n) = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \phi_n(v_1) = 0,\ \phi_n(v_2) = 0,\ \ldots,\ \phi_n(v_{n-1}) = 0,\ \phi_n(v_n) = 1 \end{array}$$

D'après le théorème 6.2, ces applications linéaires ϕ_i sont uniques et bien définies.

Exemple 11.5 : Considérons la base suivante de $\mathbf{R}^2$: $\{v_1 = (2, 1),\ v_2 = (3, 1)\}$. Trouver la base duale $\{\phi_1, \phi_2\}$.

Nous cherchons les formes linéaires $\phi_1(x, y) = ax + by$ et $\phi_2(x, y) = cx + dy$ telles que

$$\phi_1(v_1) = 1,\ \phi_1(v_2) = 0,\ \phi_2(v_1) = 0,\ \phi_2(v_2) = 1$$

Ainsi

$$\left.\begin{array}{l} \phi_1(v_1) = \phi_1(2,1) = 2a + b = 1 \\ \phi_1(v_2) = \phi_1(3,1) = 3a + b = 0 \end{array}\right\} \quad \text{ou} \quad a = -1,\ b = 3$$

$$\left.\begin{array}{l} \phi_2(v_1) = \phi_2(2,1) = 2c + d = 0 \\ \phi_2(v_2) = \phi_2(3,1) = 3c + d = 1 \end{array}\right\} \quad \text{ou} \quad c = 1,\ d = -2$$

Donc la base duale est $\{\phi_1(x, y) = -x + 3y,\ \phi_2(x, y) = x - y\}$.

Le théorème suivant établit des relations fondamentales entre les bases et leurs duales.

Théorème 11.2 : Soit $\{v_1, \ldots, v_n\}$ une base de V et soit $\{\phi_1, \ldots, \phi_n\}$ la base duale de V^*. Alors pour tout vecteur $u \in V$

$$u = \phi_1(u)v_1 + \phi_2(u)v_2 + \cdots + \phi_n(u)v_n$$

et pour toute forme linéaire $\sigma \in V^*$

$$\sigma = \sigma(v_1)\phi_1 + \sigma(v_2)\phi_2 + \cdots + \sigma(v_n)\phi_n$$

Théorème 11.3 : Soient $\{v_1, \ldots, v_n\}$ et $\{w_1, \ldots, w_n\}$ des bases de V et soit $\{\phi_1, \ldots, \phi_n\}$ et $\{\sigma_1, \ldots, \sigma_n\}$ les bases duales de V^* correspondant respectivement aux $\{v_i\}$ et $\{w_i\}$. Supposons que P soit la matrice de passage de la base $\{v_i\}$ à la base $\{w_i\}$. Alors $(P^{-1})^t$ est la matrice de passage de $\{\phi_i\}$ à $\{\sigma_i\}$.

ESPACE DUAL D'UN ESPACE DUAL OU ESPACE BIDUAL

Rappelons que tout espace vectoriel V a un espace dual V^* qui est l'ensemble de toutes les formes linéaires sur V. Ainsi V^* lui-même a un espace dual V^{**}, appelé espace bidual de V qui est l'ensemble de toutes les formes linéaires sur V^*.

Montrons maintenant que chaque $v \in V$ détermine un élément particulier $\hat{v} \in V^{**}$. D'abord nous définissons pour tout $\phi \in V^*$

$$\hat{v}(\phi) = \phi(v)$$

Il reste à démontrer que cette application $\hat{v} : V^* \to K$ est linéaire. Pour tous scalaires $a, b \in K$ et pour toutes formes linéaires $\phi, \sigma \in V^*$, nous avons

$$\hat{v}(a\phi + b\sigma) = (a\phi + b\sigma)(v) = a\,\phi(v) + b\,\sigma(v) = a\,\hat{v}(\phi) + b\,\hat{v}(\sigma)$$

C'est-à-dire, $\hat{v}$ est linéaire et donc $\hat{v} \in V^{**}$. Le théorème suivant s'applique alors.

Théorème 11.4 : Si V est de dimension finie, alors l'application $v \mapsto \hat{v}$ est un isomorphisme de V sur V^{**}.

L'application précédente $v \mapsto \hat{v}$ est appelée l'application canonique de V dans V^{**}. Insistons sur le fait que cette application n'est jamais surjective sur V^{**} si V n'est pas de dimension finie. Cependant, elle est toujours linéaire et de plus elle est toujours injective.

Supposons maintenant que V soit de dimension finie. D'après le théorème précédent l'application canonique détermine un isomorphisme de V dans V^{**}. A moins qu'il en soit spécifié autrement nous identifierons V avec V^{**} par cette application. En conséquence nous considérerons V comme l'espace des formes linéaires sur V^* et nous écrirons $V = V^{**}$. Remarquons que si $\{\phi_i\}$ est la base de V^* duale de la base $\{v_i\}$ de V alors $\{v_i\}$ est la base de $V = V^{**}$ duale de $\{\phi_i\}$.

ANNIHILATEURS

Soit W un sous-ensemble (non nécessairement un sous-espace) d'un espace vectoriel V. Une forme linéaire $\phi \in V^*$ est appelée un annihilateur de W si $\phi(w) = 0$ pour tout $w \in W$, c'est-à-dire si $\phi(W) = \{0\}$. Nous démontrerons que l'ensemble de telles applications noté par W^0 qui est appelé l'annihilateur de W, est un sous-espace de V^*. Il est clair que $0 \in W^0$. Supposons maintenant $\phi, \sigma \in W^0$. Alors pour tous scalaires $a, b \in K$ et pour tout $w \in W$

$$(a\phi + b\sigma)(w) = a\,\phi(w) + b\,\sigma(w) = a0 + b0 = 0$$

Ainsi $a\phi + b\sigma \in W^0$ et donc W^0 est un sous-espace de V^*.

Dans le cas où W est un sous-espace de V, nous avons la relation fondamentale suivante entre W et son annihilateur W^0.

Théorème 11.5 : Supposons V de dimension finie et W un sous-espace de V. Alors (i) $\dim W + \dim W^0 = \dim V$ et (ii) $W^{00} = W$.

Ici $W^{00} = \{v \in V \,;\, \phi(v) = 0 \text{ pour tout } \phi \in W^0\}$, ou de façon équivalente $W^{00} = (W^0)^0$ où W^{00} est considéré comme un sous-espace de V par rapport à l'identification de V à V^{**}.

La notion d'annihilateur nous permet de donner une autre interprétation d'un système homogène d'équations linéaires,

$$\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0 \end{array} \qquad (*)$$

Ici chaque ligne $(a_{i1}, a_{i2}, \ldots, a_{in})$ de la matrice des coefficients $A = (a_{ij})$ est considérée comme un élément de K^n et chaque vecteur solution $\phi = (x_1, x_2, \ldots, x_n)$ est considéré comme un élément de l'espace dual. Dans ce contexte, l'espace solution S de (*) est l'annihilateur des lignes de A et donc de l'espace ligne de A. En conséquence, en utilisant le théorème 11-5, nous obtenons encore le résultat fondamental suivant sur la dimension de l'espace solution d'un système homogène d'équations linéaires :

$$\dim S = \dim K^n - \dim (\text{espace ligne de } A) = n - \text{rang}\ (A)$$

TRANSPOSEE D'UNE APPLICATION LINEAIRE

Soit $T : V \to U$ une application linéaire quelconque d'un espace vectoriel V dans l'espace vectoriel U. Pour une forme linéaire quelconque $\phi \in U^*$, l'application composée $\phi \circ T$ est une application linéaire de V dans K :

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[rr, bend right, "\phi \circ T"'] & U \arrow[r, "\phi"] & K \end{tikzcd}$$

C'est-à-dire $\phi \circ T \in V^*$. Donc la correspondance

$$\phi \mapsto \phi \circ T$$

est une application de U^* dans V^* ; nous la noterons par T^t et l'appellerons la transposée de T. En d'autres termes, $T^t : U^* \to V^*$ est définie par

$$T^t(\phi) = \phi \circ T$$

Ainsi $(T^t(\phi))\ (v) = \phi(T(v))$ pour tout $v \in V$.

Théorème 11.6 : L'application transposée T^t définie plus haut est linéaire.

Démonstration : Pour tous scalaires a, $b \in K$ et pour toutes formes linéaires ϕ, $\sigma \in U^*$,

$$\begin{aligned} T^t(a\phi + b\sigma) &= (a\phi + b\sigma) \circ T = a(\phi \circ T) + b(\sigma \circ T) \\ &= a\,T^t(\phi) + b\,T^t(\sigma) \end{aligned}$$

C'est-à-dire que T^t est linéaire, comme nous l'avions affirmé.

Insistons sur le fait que si T est une application linéaire de V dans U, alors T^t est une application linéaire de U^* dans V^* :

$$V \xrightarrow{T} U \qquad V^* \xleftarrow{T^t} U^*$$

Le nom de "transposée" pour l'application T^t provient du théorème suivant.

Théorème 11.7 : Soit $T : V \to U$ une application linéaire, et soit A la représentation matricielle de T relativement aux bases $\{v_i\}$ de V et $\{u_i\}$ de U. Alors la matrice transposée A^t est la représentation matricielle de $T^t : U^* \to V^*$ relativement aux bases duales de $\{u_i\}$ et $\{v_i\}$ respectivement.

PROBLEMES RESOLUS

ESPACES DUAUX ET BASES

11.1. Soit $\phi : \mathbf{R}^2 \to \mathbf{R}$ et $\sigma : \mathbf{R}^2 \to \mathbf{R}$ les formes linéaires définies par $\phi(x, y) = x + 2y$ et $\sigma(x, y) = 3x - y$. Trouver (i) $\phi + \sigma$, (ii) 4ϕ, (iii) $2\phi - 5\sigma$.

(i) $(\phi + \sigma)(x, y) = \phi(x, y) + \sigma(x, y) = x + 2y + 3x - y = 4x + y$

(ii) $(4\phi)(x, y) = 4\,\phi(x, y) = 4(x + 2y) = 4x + 8y$

(iii) $(2\phi - 5\sigma)(x, y) = 2\,\phi(x, y) - 5\,\sigma(x, y) = 2(x + 2y) - 5(3x - y) = -13x + 9y$

11.2. Considérons les bases suivantes de $\mathbf{R}^3$: $\{v_1 = (1, -1, 3),\ v_2 = (0, 1, -1),\ v_3 = (0, 3, -2)\}$. Trouver la base duale $\{\phi_1, \phi_2, \phi_3\}$.

Nous cherchons les formes linéaires

$$\phi_1(x, y, z) = a_1x + a_2y + a_3z, \quad \phi_2(x, y, z) = b_1x + b_2y + b_3z, \quad \phi_3(x, y, z) = c_1x + c_2y + c_3z$$

telles que

$$\begin{array}{lll} \phi_1(v_1) = 1 & \phi_1(v_2) = 0 & \phi_1(v_3) = 0 \\ \phi_2(v_1) = 0 & \phi_2(v_2) = 1 & \phi_2(v_3) = 0 \\ \phi_3(v_1) = 0 & \phi_3(v_2) = 0 & \phi_3(v_3) = 1 \end{array}$$

Nous trouvons ϕ_1 comme suit :

$$\begin{aligned} \phi_1(v_1) &= \phi_1(1, -1, 3) = a_1 - a_2 + 3a_3 = 1 \\ \phi_1(v_2) &= \phi_1(0, 1, -1) = a_2 - a_3 = 0 \\ \phi_1(v_3) &= \phi_1(0, 3, -2) = 3a_2 - 2a_3 = 0 \end{aligned}$$

En résolvant le système d'équations, nous obtenons $a_1 = 1$, $a_2 = 0$, $a_3 = 0$. Donc $\phi_1(x, y, z) = x$.

Nous trouvons ensuite ϕ_2 :

$$\begin{aligned} \phi_2(v_1) &= \phi_2(1, -1, 3) = b_1 - b_2 + 3b_3 = 0 \\ \phi_2(v_2) &= \phi_2(0, 1, -1) = b_2 - b_3 = 1 \\ \phi_2(v_3) &= \phi_2(0, 3, -2) = 3b_2 - 2b_3 = 0 \end{aligned}$$

En résolvant le système, nous obtenons $b_1 = 7$, $b_2 = -2$, $b_3 = -3$. Donc $\phi_2(x, y, z) = 7x - 2y - 3z$.

Finalement nous trouvons ϕ_3 :

$$\begin{aligned} \phi_3(v_1) &= \phi_3(1, -1, 3) = c_1 - c_2 + 3c_3 = 0 \\ \phi_3(v_2) &= \phi_3(0, 1, -1) = c_2 - c_3 = 0 \\ \phi_3(v_3) &= \phi_3(0, 3, -2) = 3c_2 - 2c_3 = 1 \end{aligned}$$

En résolvant le système, nous obtenons $c_1 = -2$, $c_2 = 1$, $c_3 = 1$. Donc $\phi_3(x, y, z) = -2x + y + z$.

11.3. Soit V l'espace vectoriel des polynômes sur $\mathbf{R}$ de degré $\leqslant 1$, c'est-à-dire $V = \{a + bt : a, b \in \mathbf{R}\}$. Soit $\phi_1 : V \to \mathbf{R}$ et $\phi_2 : V \to \mathbf{R}$ définies par

$$\phi_1(f(t)) = \int_0^1 f(t)\,dt \qquad \text{et} \qquad \phi_2(f(t)) = \int_0^2 f(t)\,dt$$

(Nous remarquons que ϕ_1 et ϕ_2 sont linéaires et donc appartiennent à l'espace dual V^*). Trouver la base $\{v_1, v_2\}$ de V qui est la base duale de $\{\phi_1, \phi_2\}$.

Soient $v_1 = a + bt$ et $v_2 = c + dt$. D'après la définition de la base duale

$$\phi_1(v_1) = 1, \ \phi_2(v_1) = 0 \quad \text{et} \quad \phi_1(v_2) = 0, \ \phi_2(v_2) = 1$$

Ainsi

$$\left.\begin{aligned} \phi_1(v_1) &= \int_0^1 (a + bt)\,dt = a + \tfrac{1}{2}b = 1 \\ \phi_2(v_1) &= \int_0^2 (a + bt)\,dt = 2a + 2b = 0 \end{aligned}\right\} \quad \text{ou} \quad a = 2,\ b = -2$$

$$\left.\begin{array}{rcl} \phi_1(v_2) &=& \int_0^1 (c+dt)\,dt \;=\; c+\frac{1}{2}d \;=\; 0 \\ \phi_2(v_2) &=& \int_0^2 (c+dt)\,dt \;=\; 2c+2d \;=\; 1 \end{array}\right\} \quad \text{ou} \quad c=-\tfrac{1}{2},\ d=1$$

En d'autres termes, $\{2-2t,\ -\frac{1}{2}+t\}$ est la base de V dont la base duale est $\{\phi_1, \phi_2\}$.

11.4. Démontrer le théorème 11.1 : Supposons que $\{v_1, v_2, \ldots, v_n\}$ soit une base de V sur K. Soit $\phi_1, \phi_2, \ldots, \phi_n \in V^*$ des formes linéaires définies par

$$\phi_i(v_j) \;=\; \delta_{ij} \;=\; \begin{cases} 1 & \text{si } i=j \\ 0 & \text{si } i\neq j \end{cases}$$

Alors $\{\phi_1, \phi_2, \ldots, \phi_n\}$ est une base de V^*.

Démontrons d'abord que $\{\phi_1, \phi_2, \ldots, \phi_n\}$ engendrent V^*. Soit ϕ un élément arbitraire de V^*, et supposons que :

$$\phi(v_1)=k_1,\ \phi(v_2)=k_2,\ \ldots,\ \phi(v_n)=k_n$$

Posons $\sigma = k_1\phi_1 + \cdots + k_n\phi_n$. Alors

$$\begin{aligned} \sigma(v_1) &= (k_1\phi_1+\cdots+k_n\phi_n)(v_1) \\ &= k_1\,\phi_1(v_1) + k_2\,\phi_2(v_1) + \cdots + k_n\,\phi_n(v_1) \\ &= k_1\cdot 1 + k_2\cdot 0 + \cdots + k_n\cdot 0 \;=\; k_1 \end{aligned}$$

De même pour $i=2,\ldots,n$,

$$\begin{aligned} \sigma(v_i) &= (k_1\phi_1+\cdots+k_n\phi_n)(v_i) \\ &= k_1\,\phi_1(v_i) + \cdots + k_i\,\phi_i(v_i) + \cdots + k_n\,\phi_n(v_i) \;=\; k_i \end{aligned}$$

Ainsi $\phi(v_i)=\sigma(v_i)$ pour $i=1,\ldots,n$. Puisque ϕ et σ appartiennent à la base des vecteurs $\phi=\sigma=k_1\phi_1+\cdots+k_n\phi_n$. En conséquence $\{\phi_1, \phi_2, \ldots, \phi_n\}$ engendre V^*.

Il reste à montrer que $\{\phi_1, \ldots, \phi_n\}$ est linéairement indépendant. Supposons

$$a_1\phi_1 + a_2\phi_2 + \cdots + a_n\phi_n \;=\; 0$$

Appliquons les deux membres à v_1, on obtient

$$\begin{aligned} 0 \;=\; 0(v_1) &= (a_1\phi_1+\cdots+a_n\phi_n)(v_1) \\ &= a_1\,\phi_1(v_1) + a_2\,\phi_2(v_1) + \cdots + a_n\,\phi_n(v_1) \\ &= a_1\cdot 1 + a_2\cdot 0 + \cdots + a_n\cdot 0 \;=\; a_1 \end{aligned}$$

De même, pour $i=2,\ldots,n$,

$$\begin{aligned} 0 \;=\; 0(v_i) &= (a_1\phi_1+\cdots+a_n\phi_n)(v_i) \\ &= a_1\,\phi_1(v_i) + \cdots + a_i\,\phi_i(v_i) + \cdots + a_n\,\phi_n(v_i) \;=\; a_i \end{aligned}$$

C'est-à-dire $a_1=0,\ldots,a_n=0$. Donc $\{\phi_1, \ldots, \phi_n\}$ est linéairement indépendant et donc est bien une base de V^*.

11.5. Démontrer le théorème 11-2 : Soit $\{v_1, v_2, \ldots, v_n\}$ une base de V et soit $\{\phi_1, \ldots, \phi_n\}$ la base duale de V^*. Alors pour tout vecteur $u \in V$

$$u \;=\; \phi_1(u)v_1 + \phi_2(u)v_2 + \cdots + \phi_n(u)v_n \tag{1}$$

et pour toute forme linéaire $\sigma \in V^*$

$$\sigma \;=\; \sigma(v_1)\phi_1 + \sigma(v_2)\phi_2 + \cdots + \sigma(v_n)\phi_n \tag{2}$$

Supposons $$u \;=\; a_1v_1 + a_2v_2 + \cdots + a_nv_n \tag{3}$$

Alors

$$\phi_1(u) \;=\; a_1\,\phi_1(v_1) + a_2\,\phi_1(v_2) + \cdots + a_n\,\phi_1(v_n) \;=\; a_1\cdot 1 + a_2\cdot 0 + \cdots + a_n\cdot 0 \;=\; a_1$$

De façon analogue, pour $i = 2, \ldots, n$,

$$\phi_i(u) = a_1\,\phi_i(v_1) + \cdots + a_i\,\phi_i(v_i) + \cdots + a_n\,\phi_i(v_n) = a_i$$

C'est-à-dire, $\phi_1(u) = a_1$, $\phi_2(u) = a_2, \ldots, \phi_n(u) = a_n$. En substituant ces résultats dans (*3*) nous obtenons (*1*).

Démontrons ensuite (*2*). Appliquons la forme linéaire σ aux deux membres de (*1*),

$$\begin{aligned}\sigma(u) &= \phi_1(u)\,\sigma(v_1) + \phi_2(u)\,\sigma(v_2) + \cdots + \phi_n(u)\,\sigma(v_n)\\ &= \sigma(v_1)\,\phi_1(u) + \sigma(v_2)\,\phi_2(u) + \cdots + \sigma(v_n)\,\phi_n(u)\\ &= (\sigma(v_1)\phi_1 + \sigma(v_2)\phi_2 + \cdots + \sigma(v_n)\phi_n)(u)\end{aligned}$$

Donc l'égalité précédente est vraie pour tout $u \in V$, $\sigma = \sigma(v_1)\phi_1 + \sigma(v_2)\phi_2 + \cdots + \sigma(v_n)\phi_n$ comme demandé.

11.6. Démontrer le théorème 11.3 : Soient $\{v_1, v_2, \ldots, v_n\}$ et $\{w_1, \ldots, w_n\}$ des bases de V et soient $\{\phi_1, \ldots, \phi_n\}$ et $\{\sigma_1, \ldots, \sigma_n\}$ les bases de V^* dual de $\{v_i\}$ et $\{w_i\}$ respectivement. Supposons que P soit la matrice de passage de $\{v_i\}$ à $\{w_i\}$. Alors $(P^{-1})^t$ est la matrice de passage de $\{\phi_i\}$ à $\{\sigma_i\}$.

Supposons

$$\begin{aligned} w_1 &= a_{11}v_1 + a_{12}v_2 + \cdots + a_{1n}v_n & \sigma_1 &= b_{11}\phi_1 + b_{12}\phi_2 + \cdots + b_{1n}\phi_n\\ w_2 &= a_{21}v_1 + a_{22}v_2 + \cdots + a_{2n}v_n & \sigma_2 &= b_{21}\phi_1 + b_{22}\phi_2 + \cdots + b_{2n}\phi_n\\ &\cdots\cdots\cdots\cdots\cdots\cdots & &\cdots\cdots\cdots\cdots\cdots\cdots\\ w_n &= a_{n1}v_1 + a_{n2}v_2 + \cdots + a_{nn}v_n & \sigma_n &= b_{n1}\phi_1 + b_{n2}\phi_2 + \cdots + b_{nn}\phi_n\end{aligned}$$

où $P = (a_{ij})$ et $Q = (b_{ij})$. Nous cherchons à démontrer que $Q = (P^{-1})^t$.

Appelons R_i la *i*ème ligne de Q et C_{ij} la *j*ème colonne de P^t. Alors

$$R_i = (b_{i1}, b_{i2}, \ldots, b_{in}) \quad \text{et} \quad C_j = (a_{j1}, a_{j2}, \ldots, a_{jn})^t$$

D'après la définition de la base duale

$$\begin{aligned}\sigma_i(w_j) &= (b_{i1}\phi_1 + b_{i2}\phi_2 + \cdots + b_{in}\phi_n)(a_{j1}v_1 + a_{j2}v_2 + \cdots + a_{jn}v_n)\\ &= b_{i1}a_{j1} + b_{i2}a_{j2} + \cdots + b_{in}a_{jn} = R_iC_j = \delta_{ij}\end{aligned}$$

où δ_{ij} est le symbole de Kronecker. Ainsi

$$QP^t = \begin{pmatrix} R_1C_1 & R_1C_2 & \ldots & R_1C_n\\ R_2C_1 & R_2C_2 & \ldots & R_2C_n\\ \ldots & \ldots & \ldots & \ldots\\ R_nC_1 & R_nC_2 & \ldots & R_nC_n\end{pmatrix} = \begin{pmatrix} 1 & 0 & \ldots & 0\\ 0 & 1 & \ldots & 0\\ \ldots & \ldots & \ldots & \ldots\\ 0 & 0 & \ldots & 1\end{pmatrix} = I$$

et donc $Q = (P^t)^{-1} = (P^{-1})^t$ comme nous l'affirmions.

11.7. Supposons que V soit de dimension finie. Montrer que si $v \in V$, $v \neq 0$, alors il existe $\phi \in V^*$ tel que $\phi(v) \neq 0$.

Etendons $\{v\}$ à une base $\{v_1, v_2, \ldots, v_n\}$ de V. D'après le théorème 6.1, il existe une application linéaire unique $\phi : V \to K$ telle que $\phi(v) = 1$ et $\phi(v_i) = 0$, $i = 2, \ldots, n$. Donc ϕ a la propriété demandée.

11.8. Démontrer le théorème 11.4 : Si V est de dimension finie, alors l'application $v \mapsto \hat{v}$ est un isomorphisme de V sur V^{**}. (Ici $\hat{v} : V^* \to K$ est définie par $\hat{v}(\phi) = \phi(v)$).

Démontrons d'abord que l'application $v \to \hat{v}$ est linéaire, c'est-à-dire pour tous vecteurs v, $w \in V$ et pour tous scalaires a, $b \in K$. $\widehat{av + bw} = a\hat{v} + b\hat{w}$. Pour toute forme linéaire $\phi \in V^*$,

$$\begin{aligned}\widehat{av + bw}\,(\phi) &= \phi(av + bw) = a\,\phi(v) + b\,\phi(w)\\ &= a\,\hat{v}(\phi) + b\,\hat{w}(\phi) = (a\hat{v} + b\hat{w})(\phi)\end{aligned}$$

Puisque $\widehat{av + bw}\,(\phi) = (a\hat{v} + b\hat{w})\phi$ pour chaque $\phi \in V^*$, nous avons $\widehat{av + bw} = a\hat{v} + b\hat{w}$. Donc l'application $v \mapsto \hat{v}$ est linéaire.

Supposons maintenant que $v \in V$, $v \neq 0$. Alors d'après le problème précédent il existe $\phi \in V^*$ pour lequel $\phi(v) \neq 0$. Donc $\hat{v}(\phi) = \phi(v) \neq 0$ et ainsi $\hat{v} \neq 0$. Puisque $v \neq 0$ implique $\hat{v} \neq 0$, l'application $v \mapsto \hat{v}$ n'est pas singulière et donc est un isomorphisme (Théorème 6.5).

Nous avons maintenant $\dim V = \dim V^* = \dim V^{**}$ car V est de dimension finie. En conséquence l'application $v \mapsto \hat{v}$ est un isomorphisme de V sur V^{**}.

ANNIHILATEURS

11.9. Montrer que si $\phi \in V^*$ annule un sous-ensemble S de V, alors ϕ annule l'ensemble des combinaisons linéaires $L(S)$ engendré par S. Donc $S^0 = (L(S))^0$.

Supposons $v \in L(S)$. Alors il existe $w_1, \ldots, w_r \in S$ pour lesquels $v = a_1 w_1 + a_2 w_2 + \cdots + a_r w_r$.

$$\phi(v) = a_1\phi(w_1) + a_2\phi(w_2) + \cdots + a_r\phi(w_r) = a_1 0 + a_2 0 + \cdots + a_r 0 = 0$$

Puisque v est un élément quelconque de $L(S)$, ϕ annule $L(S)$ comme il était demandé.

11.10. Soit W le sous-espace de $\mathbf{R}^4$ engendré par $v_1 = (1, 2, -3, 4)$ et $v_2 = (0, 1, 4, -1)$. Trouver une base de l'annihilateur de W.

D'après le problème précédent, il suffit de trouver une base de l'ensemble des formes linéaires $\phi(x, y, z, w) = ax + by + cz + dw$ pour laquelle $\phi(v_1) = 0$ et $\phi(v_2) = 0$:

$$\begin{aligned} \phi(1,2,-3,4) &= a + 2b - 3c + 4d = 0 \\ \phi(0,1,4,-1) &= \quad\;\; b + 4c - d = 0 \end{aligned}$$

Le système d'équations ayant pour inconnues a, b, c, d est mis sous forme échelonnée.

Posons $c = 1$, $d = 0$ on obtient la solution $a = 11$, $b = -4$, $c = 1$, $d = 0$ et donc la forme linéaire $\phi_1(x, y, z, w) = 11x - 4y + z$.

Posons $c = 0$, $d = -1$ on obtient la solution $a = 6$, $b = -1$, $c = 0$, $d = -1$ et donc la forme linéaire $\phi_2(x, y, z, w) = 6x - y - w$.

L'ensemble des formes linéaires $\{\phi_1, \phi_2\}$ est une base de W^0, l'annihilateur de W.

11.11. Montrer que (1) pour tout sous-ensemble S de V, $S \subset S^{00}$; (2) si $S_1 \subset S_2$, alors $S_2^0 \subset S_1^0$.

(1) Soit $v \in S$. Alors pour chaque forme linéaire $\phi \in S^0$, $\hat{v}(\phi) = \phi(v) = 0$. Donc $\hat{v} \in (S^0)^0$. Donc d'après l'identification de V et de V^{**}, $v \in S^{00}$. En conséquence $S \subset S^{00}$.

(2) Soit $\phi \in S_2^0$. Alors $\phi(v) = 0$ pour chaque $v \in S_2$. Mais $S_1 \subset S_2$; donc ϕ annule chaque élément de S_1, c'est-à-dire $\phi \in S_1$. De plus $S_2^0 \subset S_1^0$.

11.12. Démontrer le théorème 11.5 : Supposons V de dimension finie et W un sous-espace de V. Alors (1) $\dim W + \dim W^0 = \dim V$ et (2) $W^{00} = W$.

(1) Supposons $\dim V = n$ et $\dim W = r \leqslant n$. Nous désirons montrer que $\dim W^0 = n - r$. Choisissons une base $\{w_1, \ldots, w_r\}$ de W et étendons-la à la base suivante de $V = \{w_1, \ldots, w_r, v_1, \ldots, v_{n-r}\}$. Considérons la base duale

$$\{\phi_1, \ldots, \phi_r, \sigma_1, \ldots, \sigma_{n-r}\}$$

D'après la définition de la base duale, chacun des σ précédents s'annule sur chacun des w_i ; donc $\sigma_1, \ldots, \sigma_{n-r} \in W^0$. On peut donc dire que $\{\sigma_j\}$ est un sous-ensemble d'une base de V^* et donc est linéairement indépendant.

Montrons maintenant que $\{\sigma_j\}$ engendre W^0. Soit $\sigma \in W^0$. D'après le théorème 11.2,

$$\begin{aligned} \sigma &= \sigma(w_1)\phi_1 + \cdots + \sigma(w_r)\phi_r + \sigma(v_1)\sigma_1 + \cdots + \sigma(v_{n-r})\sigma_{n-r} \\ &= 0\phi_1 + \cdots + 0\phi_r + \sigma(v_1)\sigma_1 + \cdots + \sigma(v_{n-r})\sigma_{n-r} \\ &= \sigma(v_1)\sigma_1 + \cdots + \sigma(v_{n-r})\sigma_{n-r} \end{aligned}$$

Ainsi $\{\sigma_1, \ldots, \sigma_{n-r}\}$ engendre W^0 et donc est une base de W^0. En conséquence $\dim W^0 = n - r = \dim V - \dim W$ comme demandé.

(2) Supposons $\dim V = n$ et $\dim W = r$. Alors $\dim V^* = n$ et d'après (1) $\dim W^0 = n - r$. Ainsi d'après (1) $\dim W^{00} = n - (n - r) = r$; donc $\dim W = \dim W^{00}$. D'après le problème précédent $W \subset W^{00}$. En conséquence $W = W^{00}$.

11.13. Soient U et W des sous-espaces de V. Démontrer que $(U + W)^0 = U^0 \cap W^0$.

Soit $\phi \in (U + W)^0$. Alors ϕ s'annule sur $U + W$, et aussi en particulier sur U et V. C'est-à-dire , $\phi \in U^0$ et $\phi \in W^0$; donc $\phi \in U^0 \cap W^0$. Ainsi $(U + W)^0 \subset U^0 \cap W^0$.

D'autre part, supposons $\sigma \in U^0 \cap W^0$. Alors σ s'annule sur U et aussi W. Si $v \in U + W$, alors $v = u + w$ où $u \in U$ et $w \in W$. Donc $\sigma(v) = \sigma(u) + \sigma(w) = 0 + 0 = 0$. Ainsi σ s'annule sur $U + W$, c'est-à-dire $\sigma \in (U + W)^0$. En conséquence $U^0 + W^0 \subset (U + W)^0$.

Les deux relations d'inclusion nous donnent la relation d'égalité demandée.

Remarque : Observons qu'aucun raisonnement de dimension n'est employé dans la démonstration ; en conséquence le résultat reste valable qu'il s'agisse d'espaces de dimension finie ou infinie.

TRANSPOSEE D'UNE APPLICATION LINEAIRE

11.14. Soit ϕ la forme linéaire de $\mathbf{R}^2$ définie par $\phi(x, y) = x - 2y$. Pour chacun des opérateurs linéaires suivants T de $\mathbf{R}^2$ trouver $(T^t(\phi)\ (x, y)$: (1) $T(x, y) = (x, 0)$; (2) $T(x, y) = (y, x+y)$; (3) $T(x, y) = (2x - 3y, 5x + 2y)$.

Par définition de l'application transposée $T^t(\phi) = \phi \circ T$ c'est-à-dire $(T^t(\phi))(v) = \phi(T(v))$ pour chaque vecteur v. Donc

(i) $(T^t(\phi))(x, y) = \phi(T(x, y)) = \phi(x, 0) = x$

(ii) $(T^t(\phi))(x, y) = \phi(T(x, y)) = \phi(y, x + y) = y - 2(x + y) = -2x - y$

(iii) $(T^t(\phi))(x, y) = \phi(T(x, y)) = \phi(2x - 3y, 5x + 2y) = (2x - 3y) - 2(5x + 2y) = -8x - 7y$.

11.15. Soit $T : V \to U$ une application linéaire et soit $T^t : U^* \to V^*$ sa transposée. Montrer que le noyau de T^t est l'annihilateur de l'image de T ; c'est-à-dire $\operatorname{Ker} T^t = (\operatorname{Im} T)^0$.

Supposons $\phi \in \operatorname{Ker} T^t$; c'est-à-dire $T^t(\phi) = \phi \circ T = 0$. Si $u \in \operatorname{Im} T$, alors $u = T(v)$ pour un $v \in V$; donc

$$\phi(u) = \phi(T(v)) = (\phi \circ T)(v) = 0(v) = 0$$

On a $\phi(u) = 0$ pour tout $u \in \operatorname{Im} T$; donc $\phi \in (\operatorname{Im} T)^0$. Donc $\operatorname{Ker} T^t \subset (\operatorname{Im} T)^0$.

D'autre part, supposons $\sigma \in (\operatorname{Im} T)^0$; c'est-à-dire $\sigma(\operatorname{Im} T) = \{0\}$. Alors pour tout $v \in V$,

$$(T^t(\sigma))(v) = (\sigma \circ T)(v) = \sigma(T(v)) = 0 = 0(v)$$

On a donc $(T^t(\sigma))(v) = 0(v)$ pour tout $v \in V$; donc $T^t(\sigma) = 0$. De plus $\sigma \in \operatorname{Ker} T^t$ et donc $(\operatorname{Im} T)^0 \subset \operatorname{Ker} T^t$.

Les deux relations d'inclusion nous donnent la relation d'égalité demandée.

11.16. Supposons V et U de dimension finie et supposons $T : V \to U$ linéaire. Démontrer que $\operatorname{rang}(T) = \operatorname{rang}(T^t)$.

Supposons $\dim V = n$ et $\dim U = m$. Supposons aussi $\operatorname{rang}(T) = r$. Alors, d'après le théorème 11.5,

$$\dim((\operatorname{Im} T)^0) = \dim U - \dim(\operatorname{Im} T) = m - \operatorname{rang}(T) = m - r$$

D'après le problème précédent, $\operatorname{Ker} T^t = (\operatorname{Im} T)^0$. Donc la nullité $(T^t) = m - r$. Il s'ensuit que, comme nous l'avions dit,

$$\operatorname{rang}(T^t) = \dim U^* - \text{nullité}\ (T^t) = m - (m - r) = r = \operatorname{rang}(T)$$

11.17. Démontrer le théorème 11.7 : soit $T : V \to U$ linéaire et soit A la représentation matricielle de T relativement aux bases $\{v_1, \ldots, v_m\}$ de V et $\{u_1, \ldots, u_n\}$ de U. Alors la matrice transposée A^t est la représentation matricielle de $T^t : U^* \to V^*$ relativement aux bases duales de $\{u_i\}$ et $\{v_j\}$.

Supposons

$$\begin{array}{rcl} T(v_1) &=& a_{11}u_1 + a_{12}u_2 + \cdots + a_{1n}u_n \\ T(v_2) &=& a_{21}u_1 + a_{22}u_2 + \cdots + a_{2n}u_n \\ \multicolumn{3}{c}{\dots\dots\dots\dots\dots\dots\dots\dots\dots} \\ T(v_m) &=& a_{m1}u_1 + a_{m2}u_2 + \cdots + a_{mn}u_n \end{array} \qquad (1)$$

Nous désirons montrer que

$$\begin{aligned} T^t(\sigma_1) &= a_{11}\phi_1 + a_{21}\phi_2 + \cdots + a_{m1}\phi_m \\ T^t(\sigma_2) &= a_{12}\phi_1 + a_{22}\phi_2 + \cdots + a_{m2}\phi_m \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ T^t(\sigma_n) &= a_{1n}\phi_1 + a_{2n}\phi_2 + \cdots + a_{mn}\phi_m \end{aligned} \tag{2}$$

où $\{\sigma_i\}$ et $\{\phi_j\}$ sont les bases duales de $\{u_i\}$ et $\{v_j\}$ respectivement.

Soit $v \in V$ et supposons que $v = k_1 v_1 + k_2 v_2 + \ldots + k_m v_m$. Alors d'après (1)

$$\begin{aligned} T(v) &= k_1 T(v_1) + k_2 T(v_2) + \cdots + k_m T(v_m) \\ &= k_1(a_{11}u_1 + \cdots + a_{1n}u_n) + k_2(a_{21}u_1 + \cdots + a_{2n}u_n) + \cdots + k_m(a_{m1}u_1 + \cdots + a_{mn}u_n) \\ &= (k_1a_{11} + k_2a_{21} + \cdots + k_ma_{m1})u_1 + \cdots + (k_1a_{1n} + k_2a_{2n} + \cdots + k_ma_{mn})u_n \\ &= \sum_{i=1}^{n} (k_1a_{1i} + k_2a_{2i} + \cdots + k_ma_{mi})u_i \end{aligned}$$

Donc pour $j = 1, \ldots, n$,

$$\begin{aligned} (T^t(\sigma_j)(v)) &= \sigma_j(T(v)) = \sigma_j\left(\sum_{i=1}^{n} (k_1a_{1i} + k_2a_{2i} + \cdots + k_ma_{mi})u_i\right) \\ &= k_1a_{1j} + k_2a_{2j} + \cdots + k_ma_{mj} \end{aligned} \tag{3}$$

D'autre part, pour $j = 1, \ldots, n$,

$$\begin{aligned} (a_{1j}\phi_1 + a_{2j}\phi_2 + \cdots + a_{mj}\phi_m)(v) &= (a_{1j}\phi_1 + a_{2j}\phi_2 + \cdots + a_{mj}\phi_m)(k_1v_1 + k_2v_2 + \cdots + k_mv_m) \\ &= k_1a_{1j} + k_2a_{2j} + \cdots + k_ma_{mj} \end{aligned} \tag{4}$$

Puisque $v \in V$ est arbitraire,(3) et (4) impliquent que

$$T^t(\sigma_j) = a_{1j}\phi_1 + a_{2j}\phi_2 + \cdots + a_{mj}\phi_m, \qquad j = 1, \ldots, n$$

qui est équivalent à (2). Le théorème est ainsi démontré.

11.18. Soit A une matrice $m \times n$ arbitraire sur un corps K. Démontrer que le rang ligne et le rang colonne de A sont égaux.

Soit $T : K^n \to K^m$ l'application linéaire définie par $T(v) = Av$, où les éléments de K^n et K^m sont écrits sous forme de vecteurs colonnes. Alors A est la représentation matricielle de T relativement aux bases usuelles de K^n et K^m, et l'image de T est l'espace colonne de A. Donc

$$\text{rang}(T) = \text{rang colonne de } A$$

D'après le théorème 11.7, A^t est la représentation matricielle de T^t relativement aux bases duales. Donc

$$\text{rang}(T^t) = \text{rang colonne de } A^t = \text{rang ligne de } A$$

Mais d'après le problème 11.16, rang (T) = rang (T^t) ; donc le rang ligne et le rang colonne de A sont égaux. (Ce résultat a été énoncé plus haut, Théorème 5.9, page 90, et a été démontré de manière directe dans le problème 5.21.)

PROBLEMES SUPPLEMENTAIRES

ESPACES DUAUX ET BASES DUALES

11.19. Soit ϕ ; $\mathbf{R}^3 \to \mathbf{R}$ et $\sigma : \mathbf{R}^3 \to \mathbf{R}$, les formes linéaires définies par $\phi(x, y, z) = 2x - 3y + z$ et $\sigma(x,y,z) = 4x - 2y + 3z$. Trouver (1) $\phi + \sigma$, (2) $3\ \phi$, (3) $2\phi - 5\sigma$.

11.20. Soit ϕ la forme linéaire de $\mathbf{R}^2$ définie par $\phi(2, 1) = 15$ et $\phi(1, -2) = -10$. Trouver $\phi(x, y)$ et en particulier trouver $\phi(-2, 7)$.

11.21. Trouver les bases duales de chacune des bases suivantes de $\mathbf{R}^3$:

(i) $\{(1,0,0), (0,1,0), (0,0,1)\}$, (ii) $\{(1,-2,3), (1,-1,1), (2,-4,7)\}$.

11.22. Soit V l'espace vectoriel des polynômes sur **R** de degré $\leqslant 2$. Soit ϕ_1, ϕ_2 et ϕ_3 les formes linéaires sur V définis par

$$\phi_1(f(t)) = \int_0^1 f(t)\,dt, \quad \phi_2(f(t)) = f'(1), \quad \phi_3(f(t)) = f(0)$$

Ici $f(t) = a + bt + ct^2 \in V$ et $f'(t)$ est la dérivée de $f(t)$. Trouver la base $\{f_1(t), f_2(t), f_3(t)\}$ de V qui est la base duale de $\{\phi_1, \phi_2, \phi_3\}$.

11.23. Supposons $u, v \in V$ et supposons que $\phi(u) = 0$ implique $\phi(v) = 0$ pour tout $\phi \in V^*$. Montrer que $v = ku$ pour un certain scalaire k.

11.24. Supposons $\phi, \sigma \in V^*$ et supposons que $\phi(v) = 0$ implique $\sigma(v) = 0$ pour tout $v \in V$. Montrer que $\sigma = k\phi$ pour un certain scalaire k.

11.25. Soit V l'espace vectoriel des polynômes sur K. Pour $a \in K$, nous définissons $\phi_a : V \to K$ par $\phi_a(f(t)) = f(a)$. Montrer que (1) ϕ_a est linéaire;(2) si $a \neq b$, alors $\phi_a \neq \phi_b$.

11.26. Soit V l'espace vectoriel des polynômes de degré $\leqslant 2$. Soient $a, b, c \in K$ des scalaires distincts. Soient ϕ_a, ϕ_b et ϕ_c les formes linéaires définies par $\phi_a(f(t)) = f(a))$; $\phi_b(f(t)) = f(b)$, $\phi_c(f(t)) = f(c)$. Montrer que $\{\phi_a, \phi_b, \phi_c\}$ est linéairement indépendant et trouver la base $\{f_1(t), f_2(t), f_3(t)\}$ de V qui est sa base duale.

11.27. Soit V l'espace vectoriel des matrices carrées d'ordre n. Soit $T : V \to K$ l'application trace : $T(A) = a_{11} + a_{22} + a_{nn}$ où $A = (a_{ij})$. Montrer que T est linéaire.

11.28. Soit W un sous-espace de V. Pour toute forme linéaire ϕ sur W, montrer qu'il existe une forme linéaire σ sur V telle que $\sigma(w) = \phi(w)$ pour tout $w \in W$, c'est-à-dire que ϕ est la restriction de σ à W.

11.29. Soit $\{e_1, \ldots, e_n\}$ la base usuelle de K^n. Montrer que la base duale est $\{\pi_1, \ldots, \pi_n\}$ où π_i est l'application $i^{\text{ème}}$ projection: $\pi_i(a_1, \ldots, a_n) = a_i$.

11.30. Soit V un espace vectoriel sur **R**. Soit $\phi_1, \phi_2 \in V^*$ et supposons que $\sigma : V \to \mathbf{R}$ définie par $\sigma(v) = \phi_1(v)\,\phi_2(v)$ appartienne aussi à V^*. Montrer que soit $\phi_1 = 0$, soit $\phi_2 = 0$.

ANNIHILATEURS

11.31. Soit W le sous-espace de $\mathbf{R}^4$ engendré par $(1, 2, -3, 4)$, $(1, 3, -2, 6)$ et $(1, 4, -1, 8)$. Trouver une base de l'annihilateur de W.

11.32. Soit W le sous-espace de $\mathbf{R}^3$ engendré par $(1, 1, 0)$ et $(0, 1, 1)$. Trouver une base de l'annihilateur de W.

11.33. Montrer que, quel que soit le sous-ensemble S de V, $L(S) = S^{00}$ où $L(S)$ est l'espace des combinaisons linéaires d'éléments de S. ($L(S)$ est le sous-espace engendré par S).

11.34. Soient U et W des sous-espaces d'un espace vectoriel V de dimension finie. Démontrer que $(U \cap W)^0 = U^0 + W^0$.

11.35. Supposons $V = U \oplus W$. Démontrer que $V^* = U^0 \oplus W^0$.

TRANSPOSEE D'UNE APPLICATION LINEAIRE

11.36. Soit ϕ la forme linéaire de $\mathbf{R}^2$ définie par $\phi(x, y) = 3x - 2y$. Pour toute application linéaire $T: \mathbf{R}^3 \to \mathbf{R}^2$, trouver $(T^t(\phi))\,(x, y, z)$:

(1) $T(x, y, z) = (x + y, y + z)$; (2) $T(x, y, z) = (x + y + z, 2x - y)$.

11.37. Supposons $S : U \to V$ et $T : V \to W$ des applications linéaires. Démontrer que $(T \circ S)^t = S^t \circ T^t$.

11.38. Supposons $T : V \to U$ une application linéaire, V étant de dimension finie. Démontrer que $\operatorname{Im} T^t = (\operatorname{Ker} T)^0$.

11.39. Supposons $T : V \to U$ une application linéaire et $u \in U$. Démontrer que $u \in \operatorname{Im} T$ ou qu'il existe $\phi \in V^*$ tel que $T^t(\phi) = 0$ et $\phi(u) = 1$.

11.40. Soit V de dimension finie. Montrer que l'application $T \mapsto T^t$ est un isomorphisme de $\operatorname{Hom}(V, V)$ sur $\operatorname{Hom}(V^*, V^*)$. (Ici T est l'opérateur linéaire générique sur V).

PROBLEMES DIVERS

11.41. Soit V un espace vectoriel sur $\mathbf{R}$. Le segment de droite $\overline{uv}$ joignant deux points $u, v \in V$ est défini par $\overline{uv} = \{tu + (1-t)v : 0 \leqslant t \leqslant 1\}$. Un sous-ensemble S de V est dit convexe si $u, v \in S$ implique $\overline{uv} \subset S$. Soit $\phi \in V^*$ et soit

$$W^+ = \{v \in V : \phi(v) > 0\}, \quad W = \{v \in V : \phi(v) = 0\}, \quad W^- = \{v \in V : \phi(v) < 0\}$$

Démontrer que W^+, W et W^- sont convexes.

11.42. Soit V un espace vectoriel de dimension finie. Un hyperplan H de V est défini comme le noyau d'une forme linéaire non nulle ϕ de V. Montrer que chaque sous-espace de V est l'intersection d'un nombre fini d'hyperplans.

REPONSES AUX PROBLEMES SUPPLEMENTAIRES

11.19. (i) $6x - 5y + 4z$, (ii) $6x - 9y + 3z$, (iii) $-16x + 4y - 13z$

11.20. $\phi(x, y) = 4x + 7y$, $\phi(-2, 7) = 41$

11.21. (i) $\{\phi_1(x, y, z) = x,\ \phi_2(x, y, z) = y,\ \phi_3(x, y, z) = z\}$
(ii) $\{\phi_1(x, y, z) = -3x - 5y - 2z,\ \phi_2(x, y, z) = 2x + y,\ \phi_3(x, y, z) = x + 2y + z\}$

11.25. (2) Posons $f(t) = t$. Alors $\phi_a(f(t)) = a \neq b = \phi_b(f(t))$ et donc $\phi_a \neq \phi_b$

11.26. $\left\{f_1(t) = \dfrac{t^2 - (b+c)t + bc}{(a-b)(a-c)},\ f_2(t) = \dfrac{t^2 - (a+c)t + ac}{(b-a)(b-c)},\ f_3(t) = \dfrac{t^2 - (a+b)t + ab}{(c-a)(c-b)}\right\}$

11.31. $\{\phi_1(x, y, z, t) = 5x - y + z,\ \phi_2(x, y, z, t) = 2y - t\}$

11.32. $\{\phi(x, y, z) = x - y + z\}$

11.36. (i) $(T^t(\phi))(x, y, z) = 3x + y - 2z$, (ii) $(T^t(\phi))(x, y, z) = -x + 5y + 3z$.

CHAPITRE 12

Formes bilinéaires — Formes quadratiques et hermitiennes

FORMES BILINEAIRES

Soit V un espace vectoriel de dimension finie sur un corps K. Une forme bilinéaire sur V est une application $f : V \times V \to K$ qui satisfait

(1) $f(au_1 + bu_2, v) = af(u_1, v) + bf(u_2, v)$

(2) $f(u, av_1 + bv_2) = af(u, v_1) + bf(u, v_2)$

pour tous $a, b \in K$ et tous $u_i, v_i \in V$. Nous exprimons la condition (1) en disant que f est linéaire par rapport à la première variable et la condition (2) en disant que f est linéaire par rapport à la seconde variable.

Exemple 12.1 : Soit ϕ et σ des formes linéaires quelconques sur V. Soit $f : V \times V \to K$ définie par $f(u, v) = \phi(u)\,\sigma(v)$. Alors f est bilinéaire puisque ϕ et σ sont chacune linéaires (Une telle forme bilinéaire f est en fait le "produit tensoriel" de ϕ et σ et est écrite quelquefois sous la forme $f = \phi \otimes \sigma$).

Exemple 12.2 : Soit f le produit scalaire dans $\mathbf{R}^n$, c'est-à-dire

$$f(u,v) = u \cdot v = a_1b_1 + a_2b_2 + \cdots + a_nb_n$$

où $u = (a_i)$ et $v = (b_i)$. Alors f est une forme bilinéaire sur $\mathbf{R}^n$.

Exemple 12.3 : Soit $A = (a_{ij})$ une matrice quelconque $n \times n$ sur K. Alors A peut être considérée comme une forme bilinéaire f sur K^n en définissant

$$f(X,Y) = X^tAY = (x_1, x_2, \ldots, x_n)\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ \ldots \\ y_n \end{pmatrix}$$

$$= \sum_{i,j=1}^{n} a_{ij}x_iy_j = a_{11}x_1y_1 + a_{12}x_1y_2 + \cdots + a_{nn}x_ny_n$$

L'expression formelle précédente des variables x_i, y_i est appelée le polynôme bilinéaire correspondant à la matrice A. La formule (1) plus bas montre que dans un certain sens, chaque forme bilinéaire est de ce type.

Nous noterons $B(V)$ l'ensemble des formes bilinéaires sur V. On peut munir $B(V)$ d'une structure d'espace vectoriel en définissant $f + g$ et kf par

$$(f+g)(u,v) = f(u,v) + g(u,v)$$
$$(kf)(u,v) = kf(u,v)$$

pour tous $f, g \in B(V)$ et pour tout $k \in K$. En fait,

Théorème 12.1 : Soit V un espace vectoriel de dimension n sur K. Soit $\{\phi_1, \ldots, \phi_n\}$ une base de l'espace dual V^*. Alors $\{f_{ij} : i, j = 1, \ldots, n\}$ est une base de $B(V)$ où f_{ij} est définie par $f_{ij}(u, v) = \phi_i(u) \cdot \phi_j(v)$. Ainsi en particulier, $\dim B(V) = n^2$.

FORMES BILINEAIRES ET MATRICES

Soit f une forme bilinéaire sur V et soit $\{e_1, \ldots, e_n\}$ une base de V. Supposons $u, v \in V$ et supposons

$$u = a_1e_1 + \cdots + a_ne_n, \quad v = b_1e_1 + \cdots + b_ne_n$$

Alors
$$\begin{aligned} f(u, v) &= f(a_1e_1 + \cdots + a_ne_n, b_1e_1 + \cdots + b_ne_n) \\ &= a_1b_1f(e_1, e_1) + a_1b_2f(e_1, e_2) + \cdots + a_nb_nf(e_n, e_n) = \sum_{i,j=1}^{n} a_ib_jf(e_i, e_j) \end{aligned}$$

Ainsi f est complètement déterminée par les n^2 valeurs $f(e_i, e_j)$.

La matrice $A = (a_{ij})$ où $a_{ij} = f(e_i, e_j)$ est appelée la représentation matricielle de f relativement à la base $\{e_i\}$ ou simplement la matrice de f dans $\{e_i\}$. Elle "représente" f dans le sens que

$$f(u, v) = \sum a_ib_jf(e_i, e_j) = (a_1, \ldots, a_n)A\begin{pmatrix} b_1 \\ b_2 \\ \cdot\cdot \\ b_n \end{pmatrix} = [u]_e^t A[v]_e \tag{1}$$

pour tous $u, v \in V$. (Comme à l'habitude, $[u]_e$ indique le vecteur coordonnée colonne de $u \in V$ dans la base $\{e_i\}$).

Demandons-nous ensuite comment une matrice représentant une forme bilinéaire est transformée lorsqu'on choisit une nouvelle base. La réponse à cette question est donnée dans le théorème suivant. (Rappelons d'après le Théorème 7.4 que la matrice de passage P d'une base $\{e_i\}$ à une autre base $\{e'_i\}$ est telle que $[u]_{e'} = P[u]_e$, pour tout $u \in V$).

Théorème 12.2 : Soit P la matrice de passage d'une base à une autre. Si A est la matrice de f dans la base initiale, alors

$$B = P^tAP$$

est la matrice de f dans la nouvelle base.

Le précédent théorème entraîne la définition suivante.

Définition : Une matrice B est dite congruente à une matrice A s'il existe une matrice inversible (ou non singulière) P telle que $B = P^tAP$.

Ainsi d'après le précédent théorème les matrices représentant la même forme bilinéaire dans diverses bases sont congruentes. Remarquons que les matrices congruentes ont le même rang car P et P^t sont non singulières ; donc la définition suivante est bien définie.

Définition : Le rang d'une forme bilinéaire f sur V, que l'on note par rang (f), est donc le rang de toute matrice représentant f. Nous disons que f est dégénérée ou non dégénérée suivant que le rang $(f) <$ dim V ou rang $(f) =$ dim V.

FORMES BILINEAIRES ALTERNEES

Une forme bilinéaire f sur V est dite alternée si

$$\text{(i)} \quad f(v, v) = 0$$

pour tout $v \in V$. Si f est alternée, alors

$$0 = f(u+v, u+v) = f(u, u) + f(u, v) + f(v, u) + f(v, v)$$

et donc
$$\text{(ii)} \quad f(u, v) = -f(v, u)$$

pour tout $u, v \in V$. Une forme bilinéaire qui satisfait la condition (ii) est dite anti-symétrique. Si $1 + 1 \neq 0$ dans K, la condition (ii) implique $f(v, v) = -f(v, v)$, ce qui implique la condition (i). En d'autres termes, les formes alternées ou anti-symétriques sont équivalentes lorsque $1 + 1 \neq 0$.

On en déduit le théorème principal suivant sur les formes bilinéaires alternées.

Théorème 12.3 : Soit f une forme bilinéaire alternée sur V. Il existe alors une base de V dans laquelle f est représentée par une matrice de la forme

$$\begin{pmatrix} \begin{matrix}0 & 1\\ -1 & 0\end{matrix} & & & & & & & \\ & \begin{matrix}0 & 1\\ -1 & 0\end{matrix} & & & & & & \\ & & \ddots & & & & & \\ & & & \begin{matrix}0 & 1\\ -1 & 0\end{matrix} & & & & \\ & & & & 0 & & & \\ & & & & & 0 & & \\ & & & & & & \ddots & \\ & & & & & & & 0 \end{pmatrix}$$

De plus le nombre des $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ est uniquement déterminé par f (car il est égal à $\frac{1}{2}$ rang (f)).

En particulier, le théorème précédent montre qu'une forme bilinéaire alternée doit avoir un rang pair.

FORMES BILINEAIRES SYMETRIQUES, FORMES QUADRATIQUES

Une forme bilinéaire f sur V est dite symétrique si

$$f(u, v) = f(v, u)$$

pour tout $u, v \in V$. Si A est une représentation matricielle de f, nous pouvons écrire

$$f(X, Y) = X^tAY = (X^tAY)^t = Y^tA^tX$$

(Utilisons le fait que X^tAY est un scalaire et donc égal à sa transposée). Donc si f est symétrique

$$Y^tA^tX = f(X, Y) = f(Y, X) = Y^tAX$$

et puisque ceci est vrai pour tous les vecteurs X, Y il s'ensuit que $A = A^t$ d'où A est symétrique. Réciproquement si A est symétrique, alors f est symétrique.

Le résultat principal concernant les formes bilinéaires symétriques est le suivant :

Théorème 12.4 : Soit f une forme bilinéaire symétrique de V sur K (dans lequel $1 + 1 \neq 0$). Alors V admet $\{v_1, \ldots, v_n\}$ dans laquelle f est représentée par une matrice diagonale, c'est-à-dire $f(v_i, v_j) = 0$ pour $i \neq j$.

Autre forme du Théorème 12.4 : Soit A une matrice symétrique sur K (dans lequel $1 + 1 \neq 0$). Alors il existe une matrice inversible (ou non singulière) P telle que P^tAP est diagonale. C'est-à-dire que A est congrue à une matrice diagonale.

Puisqu'une matrice inversible est un produit de matrices élémentaires (Problème 3.36), une façon d'obtenir la forme diagonale P^tAP est d'effectuer une suite d'opérations élémentaires sur les lignes et d'effectuer la même suite d'opérations élémentaires sur les colonnes. Ces deux suites de mêmes opérations élémentaires sur I nous donneront P^t. Cette méthode est illustrée dans l'exemple suivant.

Exemple 12.4 : Soit $A = \begin{pmatrix} 1 & 2 & -3 \\ 2 & 5 & -4 \\ -3 & -4 & 8 \end{pmatrix}$, une matrice symétrique. Il est commode de former la matrice bloc (A, I) :

$$(A, I) = \left(\begin{array}{ccc|ccc} 1 & 2 & -3 & 1 & 0 & 0 \\ 2 & 5 & -4 & 0 & 1 & 0 \\ -3 & -4 & 8 & 0 & 0 & 1 \end{array}\right)$$

Appliquons les opérations élémentaires $R_2 \to -2R_1 + R_2$ et $R_3 \to 3R_1 + R_3$ à (A, I) et ensuite les opérations correspondantes $C_2 \to -2C_1 + C_2$ et $C_3 \to 3C_1 + C_3$ à A ; on obtient

$$\left(\begin{array}{ccc|ccc} 1 & 2 & -3 & 1 & 0 & 0 \\ 0 & 1 & 2 & -2 & 1 & 0 \\ 0 & 2 & -1 & 3 & 0 & 1 \end{array}\right) \quad \text{et alors} \quad \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 2 & -2 & 1 & 0 \\ 0 & 2 & -1 & 3 & 0 & 1 \end{array}\right)$$

Appliquons ensuite l'opération élémentaire $R_3 \to -2R_2 + R_3$ et l'opération correspondante $C_3 \to -2C_2 + C_3$; on obtient

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 2 & -2 & 1 & 0 \\ 0 & 0 & -5 & 7 & -2 & 1 \end{array}\right) \quad \text{et alors} \quad \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -2 & 1 & 0 \\ 0 & 0 & -5 & 7 & -2 & 1 \end{array}\right)$$

Maintenant A a été diagonalisée. Posons

$$P = \begin{pmatrix} 1 & -2 & 7 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{et donc} \quad P^tAP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -5 \end{pmatrix}$$

Définition : Une application $q : V \to K$ est appelée forme quadratique si $q(v) = f(v, v)$ pour une certaine forme bilinéaire symétrique quelconque f sur V.

Nous appelons q la forme quadratique associée à la forme bilinéaire symétrique f. Si $1 + 1 \neq 0$ dans K, alors f peut être obtenue à partir de q grâce à l'identité

$$f(u, v) = \tfrac{1}{2}(q(u + v) - q(u) - q(v))$$

La formule précédente est appelée la forme polaire de f.

Si f est représentée maintenant par une matrice symétrique $A = (a_{ij})$ alors q a pour forme

$$q(X) = f(X, X) = X^tAX = (x_1, \ldots, x_n)\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix}$$

$$= \sum_{i,j} a_{ij}x_ix_j = a_{11}x_1^2 + a_{22}x_2^2 + \cdots + a_{nn}x_n^2 + 2\sum_{i<j} a_{ij}x_ix_j$$

L'expression formelle précédente des variables x_i est appelée le polynôme quadratique correspondant à la matrice symétrique A. Remarquons que si A est une matrice diagonale, alors q admet une représentation diagonale

$$q(X) = X^tAX = a_{11}x_1^2 + a_{22}x_2^2 + \cdots + a_{nn}x_n^2$$

c'est-à-dire, le polynôme quadratique représentant q ne contiendra aucun terme rectangle x_ix_j, $i \neq j$. D'après le théorème 12.4, toute forme quadratique a une telle représentation (lorsque $1 + 1 \neq 0$).

Exemple 12.5 : Considérons la forme quadratique suivante sur $\mathbf{R}^2$:

$$q(x, y) = 2x^2 - 12xy + 5y^2$$

Une manière de diagonaliser q est la méthode de Gauss de "complétion des carrés", qui est entièrement décrite dans le Problème 12.35. Dans ce cas, nous posons $x = s + 3t$, $y = t$ pour obtenir la forme diagonale

$$q(x, y) = 2(s + 3t)^2 - 12(s + 3t)t + 5t^2 = 2s^2 - 13t^2$$

FORMES BILINEAIRES SYMETRIQUES REELLES. LOI D'INERTIE

Dans cette partie nous traiterons des formes bilinéaires symétriques et des formes quadratiques sur des espaces vectoriels sur le corps des réels $\mathbf{R}$. Ces formes apparaissent dans diverses branches des mathématique et de la physique. La nature spéciale de $\mathbf{R}$ permet une théorie indépendante. En voici le résultat principal :

Théorème 12.5 : Soit f une forme bilinéaire symétrique sur V dans $\mathbf{R}$. Alors il y a une base de V dans laquelle f est représentée par une matrice diagonale ; toute autre représentation diagonale a le même nombre P d'éléments positifs et le même nombre N d'éléments négatifs. La différence $S = P - N$ est appelée la signature de f.

Une forme bilinéaire symétrique réelle f est dite positive si

$$q(v) = f(v, v) \geqslant 0$$

pour tout vecteur v ; et est dite définie positive si

$$q(v) = f(v, v) > 0$$

pour tout vecteur $v \neq 0$. D'après le théorème précédent,

(i) f est positive si et seulement si $S = \text{rang}\ (f)$

(ii) f est définie positive si et seulement si $S = \dim V$

où S est la signature de f.

Exemple 12.6 : Soit f le produit scalaire sur $\mathbf{R}^n$; c'est-à-dire

$$f(u, v) = u \cdot v = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$$

où $u = (a_i)$ et $v = (b_i)$. Remarquons que f est symétrique puisque

$$f(u, v) = u \cdot v = v \cdot u = f(v, u)$$

De plus, f est définie positive parce que

$$f(u, u) = a_1^2 + a_2^2 + \cdots + a_n^2 > 0$$

lorsque $u \neq 0$.

Dans le chapitre suivant, nous verrons comment une forme quadratique q se transforme lorsque la matrice de passage P est orthogonale. S'il n'existe aucune condition sur P, alors q peut être représentée sous une forme diagonale avec seulement 1 et -1 comme coefficients non nuls. En particulier

Corollaire 12.6 : Une forme quadratique réelle quelconque q admet une représentation unique de la forme

$$q(x_1, \ldots, x_n) = x_1^2 + \cdots + x_s^2 - x_{s+1}^2 - \cdots - x_r^2$$

Le résultat précédent pour les formes réelles quadratiques est quelquefois appelé Loi d'inertie de Sylvester.

FORMES HERMITIENNES

Soit V un espace vectoriel de dimension finie sur le corps des complexes C. Soit $f : V \times V \to C$ telle que

$$\text{(i)} \quad f(au_1 + bu_2, v) = a\,f(u_1, v) + b\,f(u_2, v)$$

$$\text{(ii)} \quad f(u, v) = \overline{f(v, u)}$$

où $a, b \in c$ et $u_i, v \in V$. Alors f est appelée une forme hermitienne sur V. (Comme à l'ordinaire, $\bar{k}$ représente le nombre complexe conjugué de $k \in$ C). D'après (i) et (ii).

$$\begin{aligned} f(u, av_1 + bv_2) &= \overline{f(av_1 + bv_2, u)} = \overline{a\,f(v_1, u) + b\,f(v_2, u)} \\ &= \bar{a}\,\overline{f(v_1, u)} + \bar{b}\,\overline{f(v_2, u)} = \bar{a}\,f(u, v_1) + \bar{b}\,f(u, v_2) \end{aligned}$$

C'est-à-dire

$$\text{(iii)} \quad f(u, av_1 + bv_2) = \bar{a}\,f(u, v_1) + \bar{b}\,f(u, v_2)$$

Comme précédemment, on exprime la condition (i) en disant que f est linéaire par rapport à la première variable. D'autre part on exprime la condition (iii) en disant que f est antilinéaire par rapport à la seconde variable. Remarquons que, d'après (ii) $f(v, v) = \overline{f(v, v)}$ et donc que $f(v, v)$ est réelle pour tout $v \in V$.

Exemple 12.7 : Soit $A = (a_{ij})$ une $n \times n$ matrice sur **C**. Ecrivons $\bar{A}$ la matrice obtenue en prenant les éléments complexes conjugués de chaque élément de A, d'où $\bar{A} = (\overline{a_{ij}})$. Nous écrivons aussi A^* pour $\bar{A}^t = \overline{A^t}$. La matrice A est dite hermitienne si $A^* = A$ c'est-à-dire si $a_{ij} = \overline{a_{ji}}$. Si A est hermitienne alors $f(X, Y) = X^t A \bar{Y}$ définit une forme hermitienne sur $\mathbf{C}^n$. (Problème 12.16).

L'application $q : V \to \mathbf{R}$ définie par $q(v) = f(v, v)$ est appelée la forme quadratique hermitienne ou forme complexe quadratique associée à la forme hermitienne f. Nous pouvons obtenir f à partir de q grâce à l'identité suivante appelée forme polaire de f:

$$f(u, v) = \tfrac{1}{4}(q(u+v) - q(u-v)) + \tfrac{i}{4}(q(u+iv) - q(u-iv))$$

Supposons maintenant que $\{e_1, \ldots, e_n\}$ soit une base de V. La matrice $H = (h_{ij})$ où $h_{ij} = f(e_i, e_j)$ est appelée la représentation matricielle de f dans la base $\{e_i\}$. D'après (ii), $f(e_i, e_j) = \overline{f(e_j, e_i)}$; donc H est hermitienne et, en particulier, les éléments diagonaux de H sont réels. Ainsi toute représentation diagonale de f contient seulement des éléments réels. Le théorème suivant est le théorème analogue pour les complexes au théorème 12.5 sur les formes bilinéaires symétriques réelles.

Théorème 12.7 : Soit f une forme hermitienne sur V. Alors il existe une base $\{e_1, \ldots, e_n\}$ de V dans laquelle f est représentée par une matrice diagonale c'est-à-dire $f(e_i, e_j) = 0$ pour $i \neq j$. De plus, toute représentation diagonale de f a le même nombre P d'éléments positifs et le même nombre N d'éléments négatifs. La différence $S = P - N$ est appelée la signature de f.

De façon analogue, une forme hermitienne f est dite positive si

$$q(v) = f(v, v) \geqslant 0$$

pour tout $v \in V$, et est dite définie positive si

$$q(v) = f(v, v) > 0$$

pour tout $v \neq 0$.

Exemple 12.8 : Soit f le produit scalaire sur $\mathbf{C}^n$; c'est-à-dire

$$f(u, v) = u \cdot v = z_1\bar{w}_1 + z_2\bar{w}_2 + \cdots + z_n\bar{w}_n$$

où $u = (z_i)$ et $v = (w_i)$. Alors f est une forme hermitienne sur $\mathbf{C}^n$. De plus f est définie positive puisque pour tout $v \neq 0$,

$$f(u, u) = z_1\bar{z}_1 + z_2\bar{z}_2 + \cdots + z_n\bar{z}_n = |z_1|^2 + |z_2|^2 + \cdots + |z_n|^2 > 0$$

PROBLEMES RESOLUS

FORMES BILINEAIRES

12.1. Soit $u = (x_1, x_2, x_3)$ et $v = (y_1, y_2, y_3)$ et soit

$$f(u,v) = 3x_1y_1 - 2x_1y_2 + 5x_2y_1 + 7x_2y_2 - 8x_2y_3 + 4x_3y_2 - x_3y_3$$

Exprimer f en notation matricielle.

Soit A la matrice 3×3 dont le ij élément est le coefficient de x_iy_j. Alors

$$f(u,v) = X^tAY = (x_1, x_2, x_3)\begin{pmatrix} 3 & -2 & 0 \\ 5 & 7 & -8 \\ 0 & 4 & -1 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

12.2. Soit A une matrice $n \times n$ sur K. Montrer que l'application suivante f est une forme bilinéaire sur K^n : $f(X, Y) = X^tAY$.

Pour tout $a, b \in K$ et pour tout $X_i, Y_i \in K^n$,

$$\begin{aligned} f(aX_1 + bX_2, Y) &= (aX_1 + bX_2)^tAY = (aX_1^t + bX_2^t)AY \\ &= aX_1^tAY + bX_2^tAY = af(X_1, Y) + bf(X_2, Y) \end{aligned}$$

Donc f est linéaire par rapport à la première variable. On a aussi

$$f(X, aY_1 + bY_2) = X^tA(aY_1 + bY_2) = aX^tAY_1 + bX^tAY_2 = af(X, Y_1) + bf(X, Y_2)$$

Donc f est linéaire par rapport à la seconde variable, et donc f est une forme bilinéaire sur K^n.

12.3. Soit f la forme bilinéaire de $\mathbf{R}^2$ définie par

$$f((x_1, x_2), (y_1, y_2)) = 2x_1y_1 - 3x_1y_2 + x_2y_2$$

(i) Trouver la matrice A de f dans la base $\{u_1 = (1, 0),\ u_2 = (1, 1)\}$.

(ii) Trouver la matrice B de f dans la base $\{v_1 = (2, 1),\ v_2 = (1, -1)\}$.

(iii) Trouver la matrice de passage P de la base $\{u_i\}$ à la base $\{v_i\}$ et vérifier que $B = P^tAP$.

(i) Posons $A = (a_{ij})$ où $a_{ij} = f(u_i, u_j)$:

$$\begin{aligned} a_{11} &= f(u_1, u_1) = f((1,0),(1,0)) = 2 - 0 + 0 = 2 \\ a_{12} &= f(u_1, u_2) = f((1,0),(1,1)) = 2 - 3 + 0 = -1 \\ a_{21} &= f(u_2, u_1) = f((1,1),(1,0)) = 2 - 0 + 0 = 2 \\ a_{22} &= f(u_2, u_2) = f((1,1),(1,1)) = 2 - 3 + 1 = 0 \end{aligned}$$

Ainsi $A = \begin{pmatrix} 2 & -1 \\ 2 & 0 \end{pmatrix}$ est la matrice de f dans la base $\{u_1, u_2\}$

(ii) Posons $B = (b_{ij})$ où $b_{ij} = f_i(v_i, v_j)$:

$$\begin{aligned} b_{11} &= f(v_1, v_1) = f((2,1),(2,1)) = 8 - 6 + 1 = 3 \\ b_{12} &= f(v_1, v_2) = f((2,1),(1,-1)) = 4 + 6 - 1 = 9 \\ b_{21} &= f(v_2, v_1) = f((1,-1),(2,1)) = 4 - 3 - 1 = 0 \\ b_{22} &= f(v_2, v_2) = f((1,-1),(1,-1)) = 2 + 3 + 1 = 6 \end{aligned}$$

Ainsi $B = \begin{pmatrix} 3 & 9 \\ 0 & 6 \end{pmatrix}$ est la matrice de f dans la base $\{v_1, v_2\}$.

(iii) Nous devons écrire v_1 et v_2 en fonction des u_i:

$$\begin{aligned} v_1 &= (2,1) = (1,0) + (1,1) = u_1 + u_2 \\ v_2 &= (1,-1) = 2(1,0) - (1,1) = 2u_1 - u_2 \end{aligned}$$

Alors $P = \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}$ et donc $P^t = \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix}$. Ainsi

$$P^t A P = \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix}\begin{pmatrix} 2 & -1 \\ 2 & 0 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 3 & 9 \\ 0 & 6 \end{pmatrix} = B$$

12.4. Démontrer le théorème 12.1 : Soit V un espace vectoriel de dimension n sur K. Soit $\{\phi_1, \ldots, \phi_n$ une base de l'espace dual V^*. Alors $\{f_{ij} : i, j = 1, \ldots, n\}$ est une base de $B(V)$ où f_{ij} est définie par $f_{ij}(u, v) = \phi_i(u)\,\phi_j(v)$. Ainsi en particulier dim $B(V) = n^2$.

Soit $\{e_1, \ldots, e_n\}$ la base de V duale de $\{\phi_i\}$. Montrons d'abord que $\{f_{ij}\}$ engendre $B(V)$. Soit $f \in B(V)$ et supposons $f(e_i, e_j) = a_{ij}$. Nous avons $f = \sum a_{ij} f_{ij}$. Il suffit de montrer que $f(e_s, e_t) = (\sum a_{ij} f_{ij})(e_s, e_t)$ pour $s, t = 1, \ldots, n$. Nous avons

$$\begin{aligned}\left(\sum a_{ij} f_{ij}\right)(e_s, e_t) &= \sum a_{ij} f_{ij}(e_s, e_t) = \sum a_{ij}\,\phi_i(e_s)\,\phi_j(e_t)\\ &= \sum a_{ij}\,\delta_{is}\,\delta_{jt} = a_{st} = f(e_s, e_t)\end{aligned}$$

Donc $\{f_{ij}\}$ engendre $B(V)$.

Il reste à montrer que $\{f_{ij}\}$ est linéairement indépendant. Supposons $\sum a_{ij} f_{ij} = 0$. Alors pour $s, t = 1, \ldots, n$,

$$0 = 0(e_s, e_t) = \left(\sum a_{ij} f_{ij}\right)(e_s, e_t) = a_{rs}$$

Il s'ensuit que $\{f_{ij}\}$ est indépendant et donc est une base de $B(V)$.

12.5. Soit $[f]$ la représentation matricielle d'une forme bilinéaire f sur V relativement à une base $\{e_1, \ldots, e_n\}$ de V. Montrer que l'application $f \mapsto [f]$ est un isomorphisme de $B(V)$ dans l'espace vectoriel des matrices carrées $n \times n$.

Puisque f est complètement déterminée par les scalaires $f(e_i, e_j)$ l'application $f \to [f]$ est bijective. Il suffit de montrer que l'application $f \mapsto [f]$ est un homomorphisme ; c'est-à-dire, que

$$[af + bg] = a[f] + b[g] \qquad (*)$$

Cependant, pour $i, j = 1, \ldots, n$,

$$(af + bg)(e_i, e_j) = af(e_i, e_j) + bg(e_i, e_j)$$

qui est une autre forme de (*). Ainsi le résultat est démontré.

12.6. Démontrer le théorème 12.2 : Soit P la matrice de passage d'une base $\{e_i\}$ à une autre base $\{e'_i\}$. Si A est la matrice de f dans la base initiale $\{e_i\}$, alors $B = P^tAP$ est la matrice de f dans la nouvelle base $\{e'_i\}$.

Soit $u, v \in V$. Puisque P est la matrice de passage de $\{e_i\}$ à $\{e'_i\}$, nous avons $P[u]_{e'} = [u]_e$ et $P[v]_{e'} = [v]_e$; donc $[u]_e^t = [u]_{e'}^t P^t$. Ainsi

$$f(u, v) = [u]_e^t\, A\, [v]_e = [u]_{e'}^t\, P^t A P [v]_{e'}$$

Puisque u et v sont des éléments arbitraires de V, P^tAP est la matrice de f dans la base $\{e'_i\}$.

FORMES BILINEAIRES SYMETRIQUES' FORMES QUADRATIQUES

12.7. Trouver la matrice symétrique qui correspond à chacun des polynômes homogènes en x et y suivants, ou en x, y, z et de degré 2 :

(i) $q(x, y) = 4x^2 - 6xy - 7y^2$ (iii) $q(x, y, z) = 3x^2 + 4xy - y^2 + 8xy - 6yz + z^2$

(ii) $q(x, y) = xy + y^2$ (iv) $q(x, y, z) = x^2 - 2yz + xz$

La matrice symétrique $A = (a_{ij})$ représentant $q(x_1, \ldots, x_n)$ a l'élément diagonal a_{ii} égal au coefficient de x_i^2, et à ses éléments a_{ij} et a_{ji} chacun égaux à la moitié du coefficient de $x_i x_j$. Ainsi

$$\begin{pmatrix} 4 & -3 \\ -3 & -7 \end{pmatrix} \qquad \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{pmatrix} \qquad \begin{pmatrix} 3 & 2 & 4 \\ 2 & -1 & -3 \\ 4 & -3 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 0 & -1 \\ \frac{1}{2} & -1 & 0 \end{pmatrix}$$

(1) (2) (3) (4)

12.8. Pour chacune des matrices réelles symétriques suivantes A, trouver une matrice non singulière P telle que P^tAP soit diagonale et trouver aussi sa signature :

$$(1)\quad A = \begin{pmatrix} 1 & -3 & 2 \\ -3 & 7 & -5 \\ 2 & -5 & 8 \end{pmatrix} \qquad (2)\quad A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & -2 & 2 \\ 1 & 2 & -1 \end{pmatrix}$$

(i) Formons d'abord la matrice bloc (A, I):

$$(A, I) = \left(\begin{array}{ccc|ccc} 1 & -3 & 2 & 1 & 0 & 0 \\ -3 & 7 & -5 & 0 & 1 & 0 \\ 2 & -5 & 8 & 0 & 0 & 1 \end{array}\right)$$

Appliquons les opérations élémentaires sur les lignes suivantes : $R_2 \to 3R_1 + R_2$ et $R_3 \to -2R_1 + R_3$ à (A, I) et ensuite les opérations élémentaires correspondantes sur les colonnes $C_2 \to 3C_1 + C_2$ et $C_3 \to -2C_1 + C_3$ à A ; on obtient

$$\left(\begin{array}{ccc|ccc} 1 & -3 & 2 & 1 & 0 & 0 \\ 0 & -2 & 1 & 3 & 1 & 0 \\ 0 & 1 & 4 & -2 & 0 & 1 \end{array}\right) \quad \text{et ensuite} \quad \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -2 & 1 & 3 & 1 & 0 \\ 0 & 1 & 4 & -2 & 0 & 1 \end{array}\right)$$

Appliquons maintenant l'opération élémentaire sur les lignes $R_3 \to R_2 + 2R_3$ et ensuite l'opération élémentaire correspondante sur les colonnes $C_3 \to C_2 + 2C_3$; on obtient

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -2 & 1 & 3 & 1 & 0 \\ 0 & 0 & 9 & -1 & 1 & 2 \end{array}\right) \quad \text{et ensuite} \quad \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -2 & 0 & 3 & 1 & 0 \\ 0 & 0 & 18 & -1 & 1 & 2 \end{array}\right)$$

A a été maintenant diagonalisée. Posons $P = \begin{pmatrix} 1 & 3 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix}$; alors $P^tAP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 18 \end{pmatrix}$.

La signature S de A est $S = 2 - 1 = 1$.

(ii) Formons d'abord la matrice bloc (A, I):

$$(A, I) = \left(\begin{array}{ccc|ccc} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & -2 & 2 & 0 & 1 & 0 \\ 1 & 2 & -1 & 0 & 0 & 1 \end{array}\right)$$

De façon à amener l'élément non nul diagonal -1 à la première position sur la diagonale principale, appliquons l'opération élémentaire sur les lignes $R_1 \leftrightarrow R_3$ et ensuite l'opération élémentaire sur les colonnes correspondantes, $C_1 \leftrightarrow C_3$; on obtient

$$\left(\begin{array}{ccc|ccc} 1 & 2 & -1 & 0 & 0 & 1 \\ 1 & -2 & 2 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \end{array}\right) \quad \text{et ensuite} \quad \left(\begin{array}{ccc|ccc} -1 & 2 & 1 & 0 & 0 & 1 \\ 2 & -2 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \end{array}\right)$$

Appliquons les opérations élémentaires sur les lignes suivantes : $R_2 \to 2R_1 + R_2$ et $R_3 \to R_1 + R_3$ et ensuite les opérations élémentaires correspondantes sur les colonnes $C_2 \to 2C_1 + C_2$ et $C_3 \to C_1 + C_3$; on obtient

$$\left(\begin{array}{ccc|ccc} -1 & 2 & 1 & 0 & 0 & 1 \\ 0 & 2 & 3 & 0 & 1 & 2 \\ 0 & 3 & 1 & 1 & 0 & 1 \end{array}\right) \quad \text{et ensuite} \quad \left(\begin{array}{ccc|ccc} -1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 2 & 3 & 0 & 1 & 2 \\ 0 & 3 & 1 & 1 & 0 & 1 \end{array}\right)$$

Appliquons l'opération ligne $R_3 \to -3R_2 + 2R_3$ et ensuite l'opération colonne correspondante $C_3 \to -3C_2 + 2C_3$; on obtient

$$\begin{pmatrix} -1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 2 & 3 & 0 & 1 & 2 \\ 0 & 0 & -7 & 2 & -3 & -4 \end{pmatrix} \quad \text{et ensuite} \quad \begin{pmatrix} -1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 2 & 0 & 0 & 1 & 2 \\ 0 & 0 & -14 & 2 & -3 & -4 \end{pmatrix}$$

Maintenant A a été diagonalisée. Posons $P = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 1 & -3 \\ 1 & 2 & -4 \end{pmatrix}$; alors $P^tAP = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -14 \end{pmatrix}$.

La signature S de A est la différence $S = 1 - 2 = -1$.

12.9. Supposons $1 + 1 \neq 0$ dans K. Donner un algorithme formel pour diagonaliser une matrice symétrique $A = (a_{ij})$ sur K.

Cas I : $a_{11} \neq 0$. Appliquons les opérations élémentaires sur les lignes $R_i \to -a_{i1}R_1 + a_{11}R_i$, $i = 2, \ldots, n$, et ensuite les opérations correspondantes sur les colonnes $C_i \to -a_{i1}C_1 + a_{11}C_i$ pour réduire A à la forme $\begin{pmatrix} a_{11} & 0 \\ 0 & B \end{pmatrix}$.

Cas II : $a_{11} = 0$ mais $a_{ii} \neq 0$ pour un certain $i > 1$. Appliquons l'opération élémentaire ligne $R_1 \leftrightarrow R_i$ puis l'opération colonne correspondante $C_1 \leftrightarrow C_i$ pour trouver a_{ii} dans la première position sur la diagonale. Ceci réduit la matrice au cas *I*.

Cas III : Tous les éléments diagonaux $a_{ii} = 0$. Choisissons i, j tel que $a_{ij} \neq 0$ et appliquons l'opération élémentaire ligne $R_i \to R_j + R_i$ et l'opération élémentaire colonne correspondante $C_i \to C_j + C_i$ pour trouver $2a_{ij} \neq 0$ dans la *i*ème position diagonale. Ceci réduit la matrice au Cas II.

Dans chacun des cas, nous pouvons finalement réduire A à la forme $\begin{pmatrix} a_{11} & 0 \\ 0 & B \end{pmatrix}$ où B est une matrice symétrique d'ordre inférieur à A. Par récurrence nous pouvons finalement trouver A sous forme diagonale.

Remarque : L'hypothèse que $1 + 1 \neq 0$ dans K est utilisée dans le cas III où nous supposons que $2a_{ij} \neq 0$.

12.10. Soit q la forme quadratique associée à la forme bilinéaire symétrique f. Vérifier la formule polaire pour f : $f(u, v) = \frac{1}{2}(q(u + v) - q(u) - q(v))$. (On suppose $1 + 1 \neq 0$).

$$\begin{aligned} q(u+v) - q(u) - q(v) &= f(u+v, u+v) - f(u,u) - f(v,v) \\ &= f(u,u) + f(u,v) + f(v,u) + f(v,v) - f(u,u) - f(v,v) \\ &= 2f(u,v) \end{aligned}$$

Si $1 + 1 \neq 0$, nous pouvons diviser par 2 et nous obtenons l'identité demandée.

12.11. Démontrer le théorème 12.4 : Soit f une forme bilinéaire symétrique de V sur K (dans lequel $1 + 1 \neq 0$). Alors V admet une base $\{v_1, \ldots, v_n\}$ dans laquelle f est représentée par une matrice diagonale c'est-à-dire $f(v_i, v_j) = 0$ pour $i \neq j$.

Méthode 1.

Si $f = 0$ ou si $\dim V = 1$ alors le théorème est évident. Donc nous pouvons supposer $f \neq 0$ et $\dim V = n > 1$. Si $q(v) = f(v, v) = 0$ pour tout $v \in V$, alors la forme polaire de f (voir Problème 12.10) implique que $f = 0$. Ainsi nous pourrons supposer qu'il y a un vecteur $v_1 \in V$ tel que $f(v_1, v_1) \neq 0$. Soit U le sous-espace engendré par V_1 et soit W le sous-espace constitué des vecteurs $v \in V$ pour lesquels $f(v_1, v) = 0$. Nous allons montrer que $V = U \oplus W$.

(i) Démontrons que $U \cap W = \{0\}$. Supposons $u \in U \cap W$. Puisque $u \in U$, $u = kv_1$ pour un certain scalaire $k \in K$. Puisque $u \in W$, $0 = f(u, u) = f(kv_1, kv_1) = k^2 f(v_1, v_1)$. Mais $f(v_1, v_1) \neq 0$; donc $k = 0$ et de plus $u = kv_1 = 0$. Ainsi $U \cap W = \{0\}$.

(ii) Démontrons que $V = U + W$: Soit $v \in V$. Posons

$$w = v - \frac{f(v_1, v)}{f(v_1, v_1)} v_1 \qquad (1)$$

Alors
$$f(v_1, w) = f(v_1, v) - \frac{f(v_1, v)}{f(v_1, v_1)} f(v_1, v_1) = 0$$

Ainsi $w \in W$. D'après (*1*), v est la somme d'un élément de U et d'un élément de W. Donc $V = U + W$. D'après (i) et (ii), $V = U \oplus W$.

Maintenant f restreint à W est une forme bilinéaire symétrique sur W. Mais $\dim W = n - 1$; donc par récurrence il y a une base $\{v_2, \ldots, v_n\}$ de W telle que $f(v_i, v_j) = 0$ pour $i \neq j$ et $2 \leqslant i, j \leqslant n$. Mais par définition de W, $f(v_1, v_j) = 0$ pour $j = 2, \ldots, n$. Donc la base $\{v_1, \ldots, v_n\}$ de V a la propriété demandée c'est-à-dire $f(v_i, v_j) = 0$ pour $i \neq j$.

Méthode 2.

L'algorithme dans le problème 12.9 montre que chaque matrice symétrique sur K est congrue à la matrice diagonale. Ceci est équivalent au fait que f a une représentation matricielle diagonale.

12.12. Soit $A = \begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{pmatrix}$ une matrice diagonale sur K. Montrer que

(i) pour tous les scalaires non nuls $k_1, \ldots, k_n \in K$, A est congrue à une matrice diagonale, dont les éléments diagonaux sont $a_i k_i^2$.

(ii) si K est le corps des complexes $\mathbf{C}$, alors A est congrue à une matrice diagonale avec seulement des 1 ou des 0 pour éléments diagonaux.

(iii) si K est le corps des réels $\mathbf{R}$, alors A est congrue à une matrice diagonale avec seulement des 1, des -1 et des 0 comme éléments diagonaux.

(i) Soit P la matrice diagonale avec pour éléments diagonaux k_i. Alors

$$P^t A P = \begin{pmatrix} k_1 & & & \\ & k_2 & & \\ & & \ddots & \\ & & & k_n \end{pmatrix}\begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{pmatrix}\begin{pmatrix} k_1 & & & \\ & k_2 & & \\ & & \ddots & \\ & & & k_n \end{pmatrix} = \begin{pmatrix} a_1k_1^2 & & & \\ & a_2k_2^2 & & \\ & & \ddots & \\ & & & a_nk_n^2 \end{pmatrix}$$

(ii) Soit P la matrice diagonale avec des éléments diagonaux $b_i = \begin{cases} 1/\sqrt{a_i} & \text{si } a_i \neq 0 \\ 1 & \text{si } a_i = 0 \end{cases}$. Alors P^tAP a la forme demandée.

(iii) Soit P la matrice diagonale avec des éléments diagonaux $b_i = \begin{cases} 1/\sqrt{|a_i|} & \text{si } a_i \neq 0 \\ 1 & \text{si } a_i = 0 \end{cases}$. Alors P^tAP a la forme demandée.

Remarque : Remarquons que (ii) n'est plus vraie si la congruence est remplacée par une congruence hermitienne (voir Problèmes 12.40 et 12.41).

12.13. Démontrer le théorème 12.5 : Soit f une forme bilinéaire symétrique de V sur R. Alors il y a une base de V dans laquelle f est représentée par une matrice diagonale, et toute autre représentation diagonale de f a le même nombre d'éléments positifs et le même nombre d'éléments négatifs.

D'après le théorème 12.4, il y a une base $\{u_1, \ldots, u_n\}$ de V, dans laquelle f est représentée par une matrice diagonale avec P éléments positifs et N éléments négatifs. Supposons maintenant que $\{w_1, \ldots, w_n\}$ soit une autre base de V dans laquelle f est représentée par une matrice diagonale, avec P' éléments positifs et N' éléments négatifs. Nous pouvons supposer sans supprimer la généralité du problème que les éléments positifs dans chaque matrice apparaissent en première position. Puisque rang $(f) = P + N = P' + N'$ il suffit de montrer que $P = P'$.

Soit U l'ensemble des générateurs des $u_1, \ldots, u_p$ et soit W l'ensemble des générateurs des $w_{P'+1}, \ldots, w_n$. Alors $f(v, v) > 0$ pour tout vecteur non nul $v \in U$ et $f(v, v) \leqslant 0$ pour tout vecteur non nul $v \in W$. Donc $U \cap W = \{0\}$. Remarquons que $\dim U = P$ et $\dim W = n - P'$. Ainsi

$$\dim(U+W) \;=\; \dim U + \dim W - \dim(U\cap W) \;=\; P + (n-P') - 0 \;=\; P - P' + n$$

Mais $\dim(U + W) \leqslant \dim V = n$; donc $P - P' + n \leqslant n$ ou $P \leqslant P'$. De façon analogue $P' \leqslant P$ et donc $P = P'$ comme il était demandé.

Remarque. Le théorème précédent et sa démonstration dépendent seulement du concept de positivité. Ainsi le théorème est vrai pour tout sous-corps K du corps **R**.

12.14. Une matrice $n \times n$ symétrique réelle A est dite positive définie si $X^tAX > 0$ pour chaque vecteur (colonne) non nul $X \in \mathbf{R}^n$; c'est-à-dire, si A est définie positive on peut la considérer comme une forme bilinéaire. Soit B une matrice quelconque réelle non singulière. Montrer que (i) B^tB est symétrique et (ii) B^tB est définie positive.

(i) $(B^tB)^t = B^tB^{tt} = B^tB$; donc B^tB est symétrique

(ii) Puisque B est non singulière, $BX \neq 0$ pour un vecteur quelconque non nul $X \in \mathbf{R}^n$. Donc le produit scalaire de BX par lui-même $BX \cdot BX = (BX)^t\,(BX)$ est positif. Ainsi $X^t(B^tB)X = (X^tB^t)(BX) = (BX)^t\,(BX) > 0$ comme demandé.

FORMES HERMITIENNES

12.15. Déterminer laquelle des matrices suivantes est hermitienne :

$$\begin{pmatrix} 2 & 2+3i & 4-5i \\ 2-3i & 5 & 6+2i \\ 4+5i & 6-2i & -7 \end{pmatrix} \qquad \begin{pmatrix} 3 & 2-i & 4+i \\ 2-i & 6 & i \\ 4+i & i & 3 \end{pmatrix} \qquad \begin{pmatrix} 4 & -3 & 5 \\ -3 & 2 & 1 \\ 5 & 1 & -6 \end{pmatrix}$$

(1) (2) (3)

Une matrice $A = (a_{ij})$ est hermitienne ssi $A = A^*$ c'est-à-dire ssi $a_{ij} = \overline{a_{ji}}$.

(i) La matrice est hermitienne, puisqu'elle est égale à la conjuguée de sa transposée.

(ii) La matrice n'est pas hermitienne, elle est symétrique.

(iii) La matrice est hermitienne. En fait une matrice réelle est hermitienne si et seulement si elle est symétrique.

12.16. Soit A une matrice hermitienne. Montrer que f est une forme hermitienne sur $\mathbf{C}^n$ où f est définie par $f(X, Y) = X^tA\overline{Y}$.

Pour tout $a, b \in \mathbf{C}$, et tout $X_1, X_2, Y \in \mathbf{C}^n$,

$$\begin{aligned} f(aX_1 + bX_2, Y) &= (aX_1 + bX_2)^t A\,\bar{Y} \;=\; (aX_1^t + bX_2^t)A\,\bar{Y} \\ &= a\,X_1^tA\,\bar{Y} + b\,X_2^tA\,\bar{Y} \;=\; a\,f(X_1, Y) + b\,f(X_2, Y) \end{aligned}$$

Donc f est linéaire par rapport à la première variable. Aussi

$$\overline{f(X,Y)} \;=\; \overline{X^tA\bar{Y}} \;=\; (\overline{X^tA\bar{Y}})^t \;=\; \overline{\bar{Y}^tA^tX} \;=\; Y^tA^*\bar{X} \;=\; Y^tA\bar{X} \;=\; f(Y,X)$$

Ainsi f est une forme hermitienne sur $\mathbf{C}^n$. (*Remarque* : Utilisons le fait que $X^tA\overline{Y}$ est un scalaire et donc qu'il est égal à son transposé).

12.17. Soit f une forme hermitienne sur V. Soit H la matrice de f dans une base $\{e_1, \ldots, e_n\}$ de V. Montrer que

(i) $f(u, v) = [u]_e^t\; H\overline{[v]}_e$ pour tous $u, v \in V$.

(ii) Si P est la matrice de passage de $\{e_i\}$ à une nouvelle base $\{e_i'\}$ de V alors $B = P^tH\overline{P}$ (ou $B = Q^*HQ$ où $Q = \overline{P}$) est la matrice de f dans la nouvelle base $\{e_i'\}$.

Remarquons que (ii) est l'analogue en complexe du Théorème 12.2.

(i) Soit $u, v \in V$ et supposons

$$u \;=\; a_1e_1 + a_2e_2 + \cdots + a_ne_n \qquad \text{et} \qquad v \;=\; b_1e_1 + b_2e_2 + \cdots + b_ne_n$$

Alors $$f(u,v) = f(a_1e_1 + \cdots + a_ne_n,\ b_1e_1 + \cdots + b_ne_n)$$

$$= \sum_{i,j} a_i\bar{b}_j f(e_i,e_j) = (a_1,\ldots,a_n)\, H \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \\ \cdot\cdot \\ \bar{b}_n \end{pmatrix} = [u]_e^t H \overline{[v]_e}$$

comme il est demandé

(2) Puisque P est la matrice de passage de $\{e_i\}$ à $\{e_i'\}$, alors

$$P[u]_{e'} = [u]_e,\quad P[v]_{e'} = [v]_e \quad \text{et aussi} \quad [u]_e^t = [u]_{e'}^t P^t,\quad \overline{[v]_e} = \overline{P}\,\overline{[v]_{e'}}$$

Ainsi d'après (1) $f(u,v) = [u]_e^t H\overline{[v]_e} = [u]_{e'}^t P^t H \overline{P}\,\overline{[v]_{e'}}$. Mais u et v sont des éléments arbitraires de V ; ainsi $P^tH\overline{P}$ est la matrice de f dans la base $\{e_i'\}$.

12.18. Soit $H = \begin{pmatrix} 1 & 1+i & 2i \\ 1-i & 4 & 2-3i \\ -2i & 2+3i & 7 \end{pmatrix}$ une matrice hermitienne. Trouver une matrice non singulière P telle que $P^tH\overline{P}$ soit diagonale.

Formons d'abord la matrice bloc (H, I)

$$\left(\begin{array}{ccc|ccc} 1 & 1+i & 2i & 1 & 0 & 0 \\ 1-i & 4 & 2-3i & 0 & 1 & 0 \\ -2i & 2+3i & 7 & 0 & 0 & 1 \end{array}\right)$$

Appliquons les opérations élémentaires sur les lignes $R_2 \to (-1+i)\,R_1 + R_2$ et $R_3 \to 2i\,R_1 + R_3$ à (A, I) et ensuite les "les opérations élémentaires colonnes hermitiennes" correspondantes (voir Problème 12.42) $C_2 \to (-1-i)\,C_1 + C_2$ et $C_3 \to -2i\,C_1 + C_3$ à A ; on obtient

$$\left(\begin{array}{ccc|ccc} 1 & 1+i & 2i & 1 & 0 & 0 \\ 0 & 2 & -5i & -1+i & 1 & 0 \\ 0 & 5i & 3 & 2i & 0 & 1 \end{array}\right) \quad \text{et ensuite} \quad \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 2 & -5i & -1+i & 1 & 0 \\ 0 & 5i & 3 & 2i & 0 & 1 \end{array}\right)$$

Appliquons ensuite l'opération élémentaire ligne $R_3 \to -5iR_2 + 2R_3$ et l'opération élémentaire colonne hermitienne correspondante $C_3 \to 5iC_2 + 2C_3$, on obtient

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 2 & -5i & -1+i & 1 & 0 \\ 0 & 0 & -19 & 5+9i & -5i & 2 \end{array}\right) \quad \text{et ensuite} \quad \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 2 & 0 & -1+i & 1 & 0 \\ 0 & 0 & -38 & 5+9i & -5i & 2 \end{array}\right)$$

Maintenant H a été diagonalisée. Posons

$$P = \begin{pmatrix} 1 & -1+i & 5+9i \\ 0 & 1 & -5i \\ 0 & 0 & 2 \end{pmatrix} \quad \text{et ensuite} \quad P^tH\overline{P} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -38 \end{pmatrix}$$

Remarquons que la signature S de H est $S = 2 - 1 = 1$.

PROBLEMES DIVERS

12.19. Montrer qu'une forme bilinéaire quelconque f sur V est la somme d'une forme bilinéaire symétrique et d'une forme bilinéaire antisymétrique.

Posons $g(u,v) = \frac{1}{2}[f(u,v) + f(v,u)]$ and $h(u,v) = \frac{1}{2}[f(u,v) - f(v,u)]$. Alors g est symétrique car

$$g(u,v) = \tfrac{1}{2}[f(u,v) + f(v,u)] = \tfrac{1}{2}[f(v,u) + f(u,v)] = g(v,u)$$

et h est antisymétrique car

$$h(u,v) = \tfrac{1}{2}[f(u,v) - f(v,u)] = -\tfrac{1}{2}[f(v,u) - f(u,v)] = -h(v,u)$$

De plus $f = g + h$.

12.20. Démontrer le théorème 12.3 : Soit f une forme bilinéaire alternée sur V. Alors il existe une base de V dans laquelle f est représentée par une matrice de la forme

$$\begin{pmatrix} \begin{array}{cc|} 0 & 1 \\ -1 & 0 \\ \hline \end{array} & & & & & & \\ & \begin{array}{|cc|} \hline 0 & 1 \\ -1 & 0 \\ \hline \end{array} & & & & & \\ & & \ddots & & & & \\ & & & \begin{array}{|cc|} \hline 0 & 1 \\ -1 & 0 \\ \hline \end{array} & & & \\ & & & & \boxed{0} & & \\ & & & & & \boxed{0} & \\ & & & & & & \ddots \quad \boxed{0} \end{pmatrix}$$

De plus le nombre des $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ est uniquement déterminé par f (car il est égal à $\frac{1}{2}$ [rang (f)]).

Si $f = 0$, alors le théorème est évident. Aussi, si dim $V = 1$, alors $f(k_1 u, k_2 u) = k_1 k_2 f(u, u) = 0$ et donc $f = 0$. En conséquence nous pouvons supposer que dim $V > 1$ et $f \neq 0$.

Puisque $f \neq 0$, il existe $u_1, u_2 \in V$ non nuls tels que $f(u_1, u_2) \neq 0$. En fait, en multipliant u_1 par un facteur approprié, nous pouvons supposer que $f(u_1, u_2) = 1$ et donc $f(u_2, u_1) = -1$. Donc maintenant u_1 et u_2 sont linéairement indépendants, car si on pose $u_2 = ku_1$, alors $f(u_1, u_2) = f(u_1, ku_1) = kf(u_1, u_1) = 0$. Soit U le sous-espace engendré par u_1 et u_2 c'est-à-dire $U = L(u_1, u_2)$. Remarquons que :

(i) La représentation matricielle de la restriction de f à U dans la base $\{u_1, u_2\}$ est $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

(ii) si $u \in U$, en posant $u = au_1 + bu_2$, alors

$$f(u, u_1) = f(au_1 + bu_2, u_1) = -b$$
$$f(u, u_2) = f(au_1 + bu_2, u_2) = a$$

Soit W le sous-espace des vecteurs $w \in V$ tels que $f(w, u_1) = 0$ et $f(w, u_2) = 0$. De manière équivalente

$$W = \{w \in V :\ f(w, u) = 0 \text{ pour chaque } u \in U\}$$

Démontrons que $V = U \oplus W$. Il est clair que $U \cap W = \{0\}$ et donc il reste à montrer que $V = U + W$. Soit $v \in V$. Posons

$$u = f(v, u_2)u_1 - f(v, u_1)u_2 \qquad \text{et} \qquad w = v - u \tag{1}$$

Puisque u est une combinaison linéaire de u_1 et u_2, $u \in U$. Montrons que $w \in W$. D'après (*1*) et (ii) $f(u, u_1) = f(v, u_1)$ donc

$$f(w, u_1) = f(v - u, u_1) = f(v, u_1) - f(u, u_1) = 0$$

De façon analogue $f(u, u_2) = f(v, u_2)$ et donc

$$f(w, u_2) = f(v - u, u_2) = f(v, u_2) - f(u, u_2) = 0$$

Alors $w \in W$ et donc d'après (*1*), $v = u + w$, où $u \in U$ et $w \in W$. Ceci montre que $V = U + W$ et donc $V = U \oplus W$.

Maintenant la restriction de f à W est une forme bilinéaire alternée sur W. Par récurrence, il existe une base $u_3, \ldots, u_n$ de W dans laquelle la matrice représentant f restreinte à W a la forme demandée. Ainsi $u_1, u_2, u_3, \ldots, u_n$ est une base de V dans laquelle la matrice représentant f a la forme désirée.

PROBLEMES SUPPLEMENTAIRES

FORMES BILINEAIRES

12.21. Soit $u = (x_1, x_2)$ et $v = (y_1, y_2)$. Déterminer si les formes suivantes sont des formes bilinéaires sur $\mathbf{R}^2$.

(i) $f(u,v) = 2x_1y_2 - 3x_2y_1$ (iv) $f(u,v) = x_1x_2 + y_1y_2$

(ii) $f(u,v) = x_1 + y_2$ (v) $f(u,v) = 1$

(iii) $f(u,v) = 3x_2y_2$ (vi) $f(u,v) = 0$.

12.22. Soit f la forme bilinéaire de $\mathbf{R}^2$ définie par

$$f((x_1,x_2),(y_1,y_2)) = 3x_1y_1 - 2x_1y_2 + 4x_2y_1 - x_2y_2$$

(i) Trouver la matrice A de f dans la base $\{u_1 = (1, 1), u_2 = (1, 2)\}$.

(ii) Trouver la matrice B de f dans la base $\{v_1 = (1, -1), v_2 = (3, 1)\}$

(iii) Trouver la matrice de passage P de $\{u_i\}$ à $\{v_i\}$ et vérifier que $B = P^tAP$.

12.23. Soit V l'espace vectoriel des matrices 2×2 sur $\mathbf{R}$. Soit $M = \begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix}$ et soit $f(A, B) = \text{tr}\,(A^tMB)$ où $A, B \in V$ et "tr" notant la trace. (i) Montrer que f est une forme bilinéaire sur V. (ii) Trouver la matrice de f dans la base $\left\{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right\}$.

12.24. Soit $B(V)$ l'ensemble des formes bilinéaires V sur K. Démontrer :

(i) si $f, g \in B(V)$, alors $f + g$ et kf, pour $k \in K$, appartiennent aussi à $B(V)$ et donc $B(V)$ est un sous-espace de l'espace vectoriel des fonctions de $V \times V$ dans K.

(ii) si ϕ et σ sont des formes linéaires sur V, alors $f(u, v) = \phi(u)\,\sigma(v)$ appartient à $B(V)$.

12.25. Soit f une forme bilinéaire sur V. Pour tout sous-ensemble S de V, nous écrivons

$$S^\perp = \{v \in V \,;\, f(u, v) = 0 \text{ pour chaque } u \in S\}, \quad S^\top = \{v \in V \,;\, f(v, u) = 0 \text{ pour chaque } u \in S\}.$$

Montrer que : (i) $S^\perp$ et $S^\top$ sont des sous-espaces de V, (ii) $S_1 \subset S_2$ implique $S_2^\perp \subset S_1^\perp$ et $S_2^\top \subset S_1^\top$, (iii) $\{0\}^\perp = \{0\}^\top = V$.

12.26. Démontrer que si f est une forme bilinéaire sur V, alors rang $(f) = \dim V - \dim V^\perp = \dim V - \dim V^\top$ et donc $\dim V^\perp = \dim V^\top$.

12.27. Soit f une forme bilinéaire sur V. Pour chaque $u \in V$, soit $\hat{u} : V \to K$ et $\tilde{u} : V \to K$ définies par $\hat{u}(x) = f(x, u)$ et $\tilde{u}(x) = f(u, x)$. Démontrer que

(i) $\hat{u}$ et $\tilde{u}$ sont chacune linéaires, c'est-à-dire $\hat{u}, \tilde{u} \in V^*$.

(ii) $u \mapsto \hat{u}$ et $u \mapsto \tilde{u}$ sont chacune des applications linéaires de V dans V^*.

(iii) rang (f) = rang $(u \mapsto \hat{u})$ = rang $(u \mapsto \tilde{u})$.

12.28. Montrer que la congruence des matrices est une relation d'équivalence, c'est-à-dire que (i) A est congruente à A, (ii) si A est congruente à B, B est congruente à A, (iii) si A, est congruente B et B congruente à C, alors A est congruente à C.

FORMES BILINEAIRES SYMETRIQUES. FORMES QUADRATIQUES

12.29. Trouver la matrice symétrique appartenant à chacun des polynômes quadratiques suivants (polynômes homogènes par rapport à l'ensemble des variables x, y, z et de degré 2) :

(i) $q(x,y,z) = 2x^2 - 8xy + y^2 - 16xz + 14yz + 5z^2$

(ii) $q(x,y,z) = x^2 - xz + y^2$

(iii) $q(x,y,z) = xy + y^2 + 4xz + z^2$

(iv) $q(x,y,z) = xy + yz$.

12.30. Pour chacune des matrices suivantes A, trouver une matrice non singulière P telle que P^tAP soit diagonale.

$$\text{(i)}\quad A = \begin{pmatrix} 2 & 3 \\ 3 & 4 \end{pmatrix}, \qquad \text{(ii)}\quad A = \begin{pmatrix} 1 & -2 & 3 \\ -2 & 6 & -9 \\ 3 & -9 & 4 \end{pmatrix}, \qquad \text{(iii)}\quad A = \begin{pmatrix} 1 & 1 & -2 & -3 \\ 1 & 2 & -5 & -1 \\ -2 & -5 & 6 & 9 \\ -3 & -1 & 9 & 11 \end{pmatrix}$$

Dans chacun des cas, trouver le rang et la signature.

12.31. Soit q la forme quadratique associée à une forme bilinéaire symétrique f. Vérifier une autre forme polaire pour f : $f(u, v) = \frac{1}{4}[q(u + v) - q(u - v)]$.

12.32. Soit $S(V)$ l'ensemble des formes bilinéaires symétriques sur V. Montrer que :
(i) $S(V)$ est un sous-espace de $B(V)$; (ii) si $\dim V = n$ alors $\dim S(V) = \frac{1}{2}n(n + 1)$.

12.33. Soit f la forme bilinéaire symétrique associée à la forme réelle quadratique $q(x, y) = ax^2 + bxy + cy^2$. Montrer que :

(i) f est non dégénérée si et seulement si $b^2 - 4ac \neq 0$;

(ii) f est définie positive si et seulement si $a > 0$ et $b^2 - 4ac < 0$.

12.34. Supposons que A soit une matrice définie positive réelle et symétrique. Montrer qu'il existe une matrice non singulière P telle que $A = P^tP$.

12.35. Considérons un polynôme réel quadratique $q(x_1, \ldots, x_n) = \sum_{i,j=1}^{n} a_{ij}x_ix_j$, où $a_{ij} = a_{ji}$.

(i) Si $a_{11} \neq 0$, montrer que la substitution

$$x_1 = y_1 - \frac{1}{a_{11}}(a_{12}y_2 + \cdots + a_{1n}y_n), \quad x_2 = y_2, \quad \ldots, \quad x_n = y_n$$

donne l'équation $q(x_1, x_2, \ldots, x_n) = a_{11}y_1^2 + q'(y_2, \ldots, y_n)$ où q' est aussi un polynôme quadratique.

(ii) Si $a_{11} = 0$, mais $a_{12} \neq 0$, montrer que la substitution

$$x_1 = y_1 + y_2,\ x_2 = y_1 - y_2,\ x_3 = y_3, \ldots, x_n = y_n$$

donne l'équation $q(x_1, \ldots, x_n) = \sum b_{ij}\, y_iy_j$ où $b_{11} \neq 0$, c'est-à-dire réduit ce cas au cas (i).

Cette méthode de diagonalisation de q est connue sous le nom "méthode des carrés de Gauss".

12.36. Utiliser des raisonnements du type précédent pour réduire chaque polynôme quadratique du Problème 12.29 à la forme diagonale. Trouver le rang et la signature dans chaque cas.

FORMES HERMITIENNES

12.37. Pour toute matrice complexe A, B et pour tout $k \in C$, montrer que :
(i) $\overline{A + B} = \overline{A} + \overline{B}$, (ii) $\overline{kA} = \overline{k}\overline{A}$, (iii) $\overline{AB} = \overline{A}\,\overline{B}$, (iv) $\overline{A}^t - \overline{A^t}$

12.38. Pour chacune des matrices hermitiennes suivantes H, trouver une matrice non singulière P telle que $P^tH\overline{P}$ soit diagonale :

$$\text{(i)}\quad H = \begin{pmatrix} 1 & i \\ -i & 2 \end{pmatrix}, \qquad \text{(ii)}\quad H = \begin{pmatrix} 1 & 2 + 3i \\ 2 - 3i & -1 \end{pmatrix}, \qquad \text{(iii)}\quad H = \begin{pmatrix} 1 & i & 2 + i \\ -i & 2 & 1 - i \\ 2 - i & 1 + i & 2 \end{pmatrix}$$

Trouver le rang et la signature de chacune de ces matrices.

12.39. Soit A une matrice quelconque complexe non singulière. Montrer que $H = A^*A$ est hermitienne définie positive.

12.40. Nous savons que B est congrue hermitienne à A s'il existe une matrice non singulière Q telle que $B = Q^*AQ$. Montrer que la congruence de matrice hermitienne est une relation d'équivalence.

12.41. Démontrer le théorème 12.7 : Soit f une forme hermitienne sur V. Alors il existe une base $\{e_1, e_2, \ldots, e_n\}$ de V dans laquelle f est représentée par une matrice diagonale, c'est-à-dire $f(e_i, e_j) = 0$ pour $i \neq j$. De plus, chaque représentation diagonale de f a le même nombre P d'éléments positifs et le même nombre N d'éléments négatifs. (Remarquons que la seconde partie du théorème n'est pas vraie pour les formes complexes symétriques bilinéaires, comme on l'a vu dans le Problème 12.12 (ii). Cependant, la démonstration; du théorème 12.5 dans le problème 12-13 s'étend pour le cas des formes hermitiennes.

PROBLEMES DIVERS

12.42. Considérons les opérations élémentaires sur les lignes suivantes :

$$[a_1]\ R_i \leftrightarrow R_j, \qquad [a_2]\ R_i \to kR_i,\ k \neq 0, \qquad [a_3]\ R_i \to kR_j + R_i$$

Les opérations élémentaires correspondantes sur les colonnes sont respectivement

$$[b_1]\ C_i \leftrightarrow C_j, \qquad [b_2]\ C_i \to kC_i,\ k \neq 0, \qquad [b_3]\ C_i \to kC_j + C_i$$

Si K est le corps des complexes $\mathbf{C}$, alors les opérations colonnes hermitiennes correspondantes sont respectivement

$$[c_1]\ C_i \leftrightarrow C_j, \qquad [c_2]\ C_i \to \bar{k}C_i,\ \bar{k} \neq 0, \qquad [c_3]\ C_i \to \bar{k}C_j + C_i$$

(i) Montrer que la matrice élémentaire correspondant à $[b_i]$ est la transposée de la matrice élémentaire correspondant à $[a_i]$.

(ii) Montrer que la matrice élémentaire correspondant à $[c_i]$ est la conjuguée transposée de la matrice élémentaire correspondant à $[a_i]$.

12.43. Soient V et W deux espaces vectoriels sur K. Une application $f : V \times W \to K$ est appelée forme bilinéaire V et W si

(i) $f(av_1 + bv_2, w) = af(v_1, w) + bf(v_2, w)$

(ii) $f(v, aw_1 + bw_2) = af(v, w_1) + bf(v, w_2)$

pour chaque $a, b \in K$, $v_i \in V$, $w_j \in W$. Démontrer que

(i) L'ensemble $B(V, W)$ des formes bilinéaires sur V et W est un sous-espace de l'espace vectoriel des fonctions de $V \times W$ sur K.

(ii) Si $\{\phi_1, \ldots, \phi_m\}$ est une base de V^* et $\{\sigma_1, \sigma_2, \ldots, \sigma_n\}$ une base de W^* alors $\{f_{ij} : i = 1, \ldots, m, j = 1, \ldots, n\}$ est une base de $B(V, W)$ où f_{ij} est définie par $f_{ij}(v, w) = \phi_i(v)\ \sigma_j(w)$. Ainsi $\dim B(V, W) = \dim V \cdot \dim W$.

(*Remarque* : Observons que si $V = W$, alors nous obtenons l'espace $B(V)$ traité dans ce chapitre).

12.44. Soit V un espace vectoriel sur K. Une application $f : \overbrace{V \times V \ldots \times V}^{m \text{ fois}} \to K$ est appelée une forme multilinéaire sur V si f est linéaire par rapport à chaque variable : c'est-à-dire, pour $i = 1, \ldots, m$,

$$f(\ldots, \widehat{au + bv}, \ldots) = af(\ldots, \hat{u}, \ldots) + bf(\ldots, \hat{v}, \ldots)$$

où $\wedge$ note la $i$ème composante, les autres composantes étant fixées. Une forme multilinéaire f est dite alternée si

$$f(v_1, \ldots, v_m) = 0 \qquad \text{pour} \qquad v_i = v_k,\ i \neq k$$

Démontrer que :

(i) L'ensemble $B_m(V)$ des formes m-linéaires sur V est un sous-espace de l'espace vectoriel des fonctions de $V \times V \times \cdots \times V$ dans K.

(ii) L'ensemble $A_m(V)$ des formes m-linéaires alternées sur V est un sous-espace de $B_m(V)$.

Remarque 1. Si $m = 2$, alors nous obtenons l'espace $B(V)$ traité dans ce chapitre.

Remarque 2. Si $V = K^m$, alors le déterminant est une forme particulière m-linéaire alternée sur V.

REPONSES AUX PROBLEMES SUPPLEMENTAIRES

12.21. (i) Oui, (ii) Non, (iii) Oui, (iv) Non, (v) Non, (vi) Oui.

12.22. (i) $A = \begin{pmatrix} 4 & 1 \\ 7 & 3 \end{pmatrix}$ (ii) $B = \begin{pmatrix} 0 & -4 \\ 20 & 32 \end{pmatrix}$ (iii) $P = \begin{pmatrix} 3 & 5 \\ -2 & -2 \end{pmatrix}$

12.23. (ii) $\begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 2 \\ 3 & 0 & 4 & 0 \\ 0 & 3 & 0 & 4 \end{pmatrix}$

12.29. (i) $\begin{pmatrix} 2 & -4 & -8 \\ -4 & 1 & 7 \\ -8 & 7 & 5 \end{pmatrix}$ (ii) $\begin{pmatrix} 1 & 0 & -\frac{1}{2} \\ 0 & 1 & 0 \\ -\frac{1}{2} & 0 & 0 \end{pmatrix}$ (iii) $\begin{pmatrix} 0 & \frac{1}{2} & 2 \\ \frac{1}{2} & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}$ (iv) $\begin{pmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \end{pmatrix}$

12.30. (i) $P = \begin{pmatrix} 1 & -3 \\ 0 & 2 \end{pmatrix}$, $P^tAP = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$, $S = 0$.

(ii) $P = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 2 \end{pmatrix}$, $P^tAP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -38 \end{pmatrix}$, $S = 1$.

(iii) $P = \begin{pmatrix} 1 & -1 & -1 & 26 \\ 0 & 1 & 3 & 13 \\ 0 & 0 & 1 & 9 \\ 0 & 0 & 0 & 7 \end{pmatrix}$, $P^tAP = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -7 & 0 \\ 0 & 0 & 0 & 469 \end{pmatrix}$, $S = 2$.

12.38. (i) $P = \begin{pmatrix} 1 & i \\ 0 & 1 \end{pmatrix}$, $P^tH\bar{P} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $S = 2$.

(ii) $P = \begin{pmatrix} 1 & -2+3i \\ 0 & 1 \end{pmatrix}$, $P^tH\bar{P} = \begin{pmatrix} 1 & 0 \\ 0 & -14 \end{pmatrix}$, $S = 0$.

(iii) $P = \begin{pmatrix} 1 & i & -3+i \\ 0 & 1 & i \\ 0 & 0 & 1 \end{pmatrix}$, $P^tH\bar{P} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -4 \end{pmatrix}$, $S = 1$.

CHAPITRE 13

Espaces préhilbertiens

INTRODUCTION

La définition d'un espace vectoriel V suppose donné un corps quelconque K. Dans ce chapitre nous restreindrons K soit au corps des réels $\mathbf{R}$, soit au corps des complexes $\mathbf{C}$. Dans le premier cas, V est appelé un espace vectoriel réel, dans le second cas V est un espace vectoriel complexe.

Rappelons que les concepts de "longueur" et "d'orthogonalité" n'apparaissent pas dans l'étude des espaces vectoriels arbitraires (quoiqu'ils soient apparus dans le chapitre I pour les espaces $\mathbf{R}^n$ et $\mathbf{C}^n$). Dans ce chapitre, nous ajoutons une nouvelle structure à un espace vectoriel V pour obtenir un espace préhilbertien, et nous définirons ces concepts dans ce cadre.

Nous supposerons que V reste un espace vectoriel de dimension finie à moins qu'il en soit spécifié autrement. De fait, de nombreux théorèmes de ce chapitre ne sont pas valables pour des espaces vectoriels de dimension infinie. Nous illustrerons cela par quelques exemples et problèmes.

ESPACES PREHILBERTIENS

Donnons d'abord une définition.

Définition : Soit V un espace vectoriel (réel ou complexe) sur K. Supposons que pour chaque paire de vecteurs $u, v \in V$ on associe un scalaire $<u, v> \in K$. Cette application est appelée produit scalaire dans V, si elle satisfait aux axiomes suivants :

$[I_1]$ $\langle au_1 + bu_2, v\rangle = a\langle u_1, v\rangle + b\langle u_2, v\rangle$

$[I_2]$ $\langle u, v\rangle = \overline{\langle v, u\rangle}$

$[I_3]$ $\langle u, u\rangle \geqslant 0$ et $\langle u, u\rangle = 0$ si et seulement si $u = 0$

L'espace vectoriel muni d'un produit scalaire est appelé espace préhilbertien.

Remarquons que $\langle u, u\rangle$ est toujours réel d'après $[I_2]$ et donc l'inégalité donnée dans $[I_3]$ a un sens. On utilise aussi la notation

$$\|u\| = \sqrt{\langle u, u\rangle}$$

Ce nombre réel non négatif $\|u\|$ est appelé la norme ou la longueur de u. En utilisant $[I_1]$ et $[I_2]$ on obtient (Problème 13.1) la relation

$$\langle u, av_1 + bv_2\rangle = \bar{a}\langle u, v_1\rangle + \bar{b}\langle u, v_2\rangle$$

Si le corps initial K est réel, le signe conjugué apparaissant précédemment et dans $[I_2]$ ne doit pas être considéré.

D'après le chapitre précédent, un produit scalaire est une forme bilinéaire positive définie et symétrique si le corps de base est réel, et est une forme hermitienne définie positive si le corps de base est complexe.

Un espace préhilbertien réel est appelé quelquefois espace euclidien, et un espace préhilbertien complexe est quelquefois appelé espace unitaire.

Exemple 13.1 : Considérons le produit scalaire dans $\mathbf{R}^n$:

$$u \cdot v = a_1b_1 + a_2b_2 + \cdots + a_nb_n$$

où $u = (a_i)$ et $v = (b_i)$. Il s'agit d'un produit scalaire sur $\mathbf{R}^n$, et $\mathbf{R}^n$ muni de ce produit scalaire est appelé habituellement espace euclidien de dimension n. Quoiqu'il existe plusieurs manières de définir un produit scalaire sur $\mathbf{R}^n$ (voir Problème 13-2) nous supposerons qu'il s'agit de ce produit scalaire sur $\mathbf{R}^n$ à moins qu'il en soit spécifié autrement.

Exemple 13.2 : Considérons le produit scalaire sur $\mathbf{C}^n$:

$$u \cdot v = z_1\overline{w}_1 + z_2\overline{w}_2 + \cdots + z_n\overline{w}_n$$

où $u = (z_i)$ et $v = (w_i)$. Comme dans le cas réel, il s'agit d'un produit scalaire sur $\mathbf{C}^n$, et nous supposerons qu'il s'agit du produit scalaire de $\mathbf{C}^n$, à moins qu'il en soit spécifié autrement.

Exemple 13.3 : Soit V l'espace vectoriel des $m \times n$ matrices sur $\mathbf{R}$. L'expression suivante est un produit scalaire sur V :

$$\langle A, B\rangle = \operatorname{tr}(B^tA)$$

où tr désigne la trace de la matrice, c'est-à-dire la somme des éléments diagonaux.

De façon analogue, si U est l'espace vectoriel des $m \times n$ matrices sur $\mathbf{C}$, alors l'expression suivante est un produit scalaire sur U :

$$\langle A, B\rangle = \operatorname{tr}(B^*A)$$

Comme d'habitude, B^*, indique la matrice transposée conjuguée de la matrice B.

Exemple 13.4 : Soit V l'espace vectoriel des fonctions réelles continues sur l'intervalle $a \leqslant t \leqslant b$. L'expression suivante est alors un produit scalaire sur V

$$\langle f, g\rangle = \int_a^b f(t)\, g(t)\, dt$$

De façon analogue, si U est l'espace vectoriel des fonctions complexes continues sur l'intervalle réel $a \leqslant t \leqslant b$, alors l'expression suivante est un produit scalaire sur U :

$$\langle f, g\rangle = \int_a^b f(t)\, \overline{g(t)}\, dt$$

Exemple 13.5 : Soit V l'espace vectoriel des suites infinies de nombres réels $(a_1, a_2, \ldots)$ satisfaisant à

$$\sum_{i=1}^{\infty} a_i^2 = a_1^2 + a_2^2 + \cdots < \infty$$

c'est-à-dire la somme précédente converge. L'addition et la multiplication scalaire des suites sont définies par :

$$(a_1, a_2, \ldots) + (b_1, b_2, \ldots) = (a_1 + b_1, a_2 + b_2, \ldots)$$

$$k(a_1, a_2, \ldots) = (ka_1, ka_2, \ldots)$$

Un produit scalaire est défini dans V par

$$\langle (a_1, a_2, \ldots), (b_1, b_2, \ldots)\rangle = a_1b_1 + a_2b_2 + \cdots$$

La somme précédente converge absolument pour tout couple de points de V (Problème 13-44) ; donc le produit scalaire est bien défini. Cet espace préhilbertien est appelé l_2.

Remarque 1 : Si $\|v\| = 1$, c'est-à-dire si $\langle v, v\rangle = 1$, alors v est appelé vecteur unitaire, on dit qu'il a été normé. Remarquons que tout vecteur non nul $u \in V$ peut être normé en posant $v = u/\|u\|$.

Remarque 2 : Le nombre réel non négatif $d(u, v) = \|v - u\|$ est appelé distance de u à v ; cette fonction satisfait les axiomes d'un espace métrique (voir Problème 13-51).

INEGALITE DE CAUCHY-SCHWARZ

La formule suivante, appelée inégalité de Cauchy-Schwarz est utilisée dans de nombreuses branches des mathématiques.

Théorème 13.1 : (Cauchy-Schwarz) : Pour tous vecteurs $u, v \in V$

$$|\langle u, v\rangle| \leqslant ||u|| \; ||v||$$

Examinons maintenant cette inégalité dans des cas bien déterminés.

Exemple 13.6 : Considérons des nombres complexes quelconques $a_1, a_2, \ldots, a_n, b_1, \ldots, b_n \in \mathbf{C}$. Alors d'après l'inégalité de Cauchy-Schwarz

$$(a_1\bar{b}_1 + \cdots + a_n\bar{b}_n)^2 \leq (|a_1|^2 + \cdots + |a_n|^2)(|b_1|^2 + \cdots + |b_n|^2)$$

c'est-à-dire $$(u \cdot v)^2 \leq ||u||^2 \, ||v||^2$$

où $u = (a_i)$ et $v = (b_i)$.

Exemple 13.7 : Soient f et g deux fonctions quelconques réelles et continues définies sur un intervalle unité $0 \leqslant t \leqslant 1$. D'après l'inégalité de Cauchy-Schwarz,

$$(\langle f, g\rangle)^2 = \left(\int_0^1 f(t)\, g(t)\, dt\right)^2 \leq \int_0^1 f^2(t)\, dt \int_0^1 g^2(t)\, dt = ||f||^2 \, ||g||^2$$

Ici V est l'espace préhilbertien de l'exemple 13-4.

ORTHOGONALITE

Soit V un espace préhilbertien. Les vecteurs $u, v \in V$ sont dits orthogonaux si $\langle u, v\rangle = 0$. Cette relation est évidemment symétrique ; c'est-à-dire si u est orthogonal à v, alors $\langle v, u\rangle = \overline{\langle u, v\rangle} = \overline{0} = 0$ et donc v est orthogonal à u. Remarquons que $0 \in V$ est orthogonal à tout vecteur $v \in V$, car

$$\langle 0, v\rangle = \langle 0v, v\rangle = 0\langle v, v\rangle = 0$$

Réciproquement, si u est orthogonal à chaque $v \in V$, alors $\langle u, u\rangle = 0$ et donc $u = 0$ d'après $[I_3]$.

Supposons maintenant que W soit un sous-espace quelconque de V. L'orthogonal de W, noté par $W^\perp$, contient les vecteurs de V qui sont orthogonaux à chaque $w \in W$:

$$W^\perp = \{v \in V : \langle v, w\rangle = 0 \text{ pour tout } w \in W\}$$

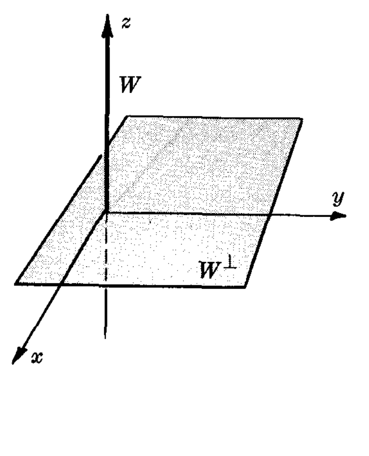
Montrons que $W^\perp$ est un sous-espace de V. On a évidemment $0 \in W^\perp$. Supposons maintenant $u, v \in W^\perp$. Alors pour tout $a, b \in K$ et pour tout $w \in W$

$$\langle au + bv, w\rangle = a\langle u, w\rangle + b\langle v, w\rangle = a \cdot 0 + b \cdot 0 = 0$$

Ainsi $au + bw \in W^\perp$ et donc W est un sous-espace de V.

Théorème 13.2 : Soit W un sous-espace de V. Alors V est la somme directe de W et $W^\perp$, c'est-à-dire $V = W \oplus W^\perp$.

Maintenant si W est un sous-espace de V, alors $V = W \oplus W^\perp$ d'après le théorème précédent ; il existe donc une projection unique $E_W : V \to V$ ayant pour image W et pour noyau $W^\perp$. C'est-à-dire, si $v \in V$ et $v = w + w'$ où $w \in W$, $w' \in W^\perp$, alors E_W est définie par $E_W(v) = w$. Cette application E_W est appelée projection orthogonale de V sur W.



Exemple 13.8 : Soit W l'axe des z dans $\mathbf{R}^3$, c'est-à-dire

$W = \{(0, 0, c) : c \in \mathbf{R}\}$

Alors $W^\perp$ est le plan xy, c'est-à-dire

$W^\perp = \{(a, b, 0) : a, b \in \mathbf{R}\}$

Comme nous l'avions prévu, $\mathbf{R}^3 = W \oplus W^\perp$. La projection orthogonale E de $\mathbf{R}^3$ sur W est donnée par $E(x, y, z) = (0, 0, z)$.

Exemple 13.9 : Considérons un système homogène d'équations linéaires sur **R** :

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0 \end{aligned}$$

ou en notation matricielle $AX = 0$. Rappelons que l'espace solution W peut être considéré comme le noyau de l'opérateur linéaire A. Nous pouvons aussi considérer W comme l'ensemble de tous les vecteurs $v = (x_1, \ldots, x_n)$ qui sont orthogonaux à chaque ligne de A. Ainsi W est l'espace orthogonal à l'espace ligne de A. Le théorème 13-2 nous donne alors une autre démonstration du résultat fondamental $\dim W = n - \text{rang}(A)$.

Remarque : Si V est un espace préhilbertien réel, alors l'angle que font entre eux deux vecteurs non nuls $u, v \in V$ est défini par

$$\cos\theta = \frac{\langle u, v\rangle}{\|u\|\,\|v\|}$$

D'après l'inégalité de Cauchy-Schwarz, $-1 \leqslant \cos\theta \leqslant 1$ et donc l'angle θ existe toujours. Observons que u et v sont orthogonaux si et seulement si ils sont "perpendiculaires" c'est-à-dire $\theta = \pi/2$.

ENSEMBLES ORTHONORMES

Un ensemble $\{u_i\}$ de vecteurs dans V est dit orthogonal si ses éléments distincts sont orthogonaux, c'est-à-dire si $\langle u_i, u_j\rangle = 0$ pour $i \neq j$. En particulier l'ensemble $\{u_i\}$ est dit orthonormé s'il est orthogonal et si chaque u_i a pour longueur 1, c'est-à-dire si

$$\langle u_i, u_j\rangle = \delta_{ij} = \begin{cases} 0 \text{ pour } i \neq j \\ 1 \text{ pour } i = j \end{cases}$$

Un ensemble orthonormé peut toujours être obtenu à partir d'un ensemble orthogonal de vecteurs non nuls en normant chacun des vecteurs.

Exemple 13.10 : Considérons la base usuelle de l'espace euclidien à 3 dimensions $\mathbf{R}^3$:

$$\{e_1 = (1,0,0),\ e_2 = (0,1,0),\ e_3 = (0,0,1)\}$$

Il est clair que

$$\langle e_1, e_1\rangle = \langle e_2, e_2\rangle = \langle e_3, e_3\rangle = 1 \quad \text{et} \quad \langle e_i, e_j\rangle = 0 \text{ pour } i \neq j$$

c'est-à-dire, $\{e_1, e_2, e_3\}$ est une base orthonormée de $\mathbf{R}^3$. Plus généralement, la base usuelle de $\mathbf{R}^n$ ou de $\mathbf{C}^n$ est orthonormée quel que soit n.

Exemple 13.11 : Soit V l'espace vectoriel des fonctions réelles continues sur l'intervalle $-\pi \leqslant t \leqslant \pi$ avec le produit scalaire défini par $\langle f, g\rangle = \int_{-\pi}^{+\pi} f(t)\, g(t)\, dt$. L'ensemble suivant est un exemple classique de sous-ensemble orthogonal de V:

$$\{1, \cos t, \cos 2t, \ldots, \sin t, \sin 2t, \ldots\}$$

L'ensemble orthogonal précédent joue un rôle fondamental dans la théorie des séries de Fourier.

Les propriétés suivantes d'un ensemble orthonormé seront utilisées dans cette partie.

Lemme 13.3 : Un ensemble orthonormé $\{u_1, u_2, \ldots, u_r\}$ est linéairement indépendant, et pour tout $v \in V$ le vecteur

$$w = v - \langle v, u_1\rangle u_1 - \langle v, u_2\rangle u_2 - \cdots - \langle v, u_r\rangle u_r$$

est orthogonal à chacun des u_i.

PROCEDE D'ORTHOGONALISATION DE GRAM-SCHMIDT

Les bases orthonormées jouent un rôle primordial dans les espaces préhilbertiens. Le théorème suivant montre qu'une telle base existe toujours ; sa démonstration utilise le célèbre procédé d'orthogonalisation de Gram-Schmidt.

Théorème 13.4 : Soit $\{v_1, \ldots, v_n\}$ une base arbitraire d'un espace préhilbertien V. Il existe alors une base orthonormée $\{u_1, u_2, \ldots, u_n\}$ de V telle que la matrice de passage de $\{v_i\}$ à $\{u_i\}$ soit triangulaire, c'est-à-dire, pour $i = 1, \ldots, n$,

$$u_i = a_{i1}v_1 + a_{i2}v_2 + \cdots + a_{ii}v_i$$

Démonstration : Posons $u_1 = v_1/\|v_1\|$, alors $\{u_1\}$ est orthonormé. Posons ensuite

$$w_2 = v_2 - \langle v_2, u_1\rangle u_1 \quad \text{et} \quad u_2 = w_2/\|w_2\|$$

D'après le lemme 13.3 w_2 (et donc u_2) est orthogonal à u_1 ; d'où $\{u_1, u_2\}$ est orthonormé. Posons ensuite

$$w_3 = v_3 - \langle v_3, u_1\rangle u_1 - \langle v_3, u_2\rangle u_2 \quad \text{et} \quad u_3 = w_3/\|w_3\|$$

De nouveau d'après le lemme 13.3 w_3 (et donc u_3) est orthogonal à u_1 et u_2 donc $\{u_1, u_2, u_3\}$ est orthonormé. En général, après avoir obtenu $\{u_1, u_2, \ldots, u_i\}$ nous posons

$$w_{i+1} = v_{i+1} - \langle v_{i+1}, u_1\rangle u_1 - \cdots - \langle v_{i+1}, u_i\rangle u_i \quad \text{et} \quad u_{i+1} = w_{i+1}/\|w_{i+1}\|$$

Remarquons que $w_{i+1} \neq 0$ car $v_{i+1} \notin L(v_1, \ldots, v_i) = L(u_1, \ldots, u_i)$). Comme plus haut $\{u_1, \ldots, u_{i+1}\}$ est aussi orthonormé. Par récurrence, nous obtenons un ensemble orthonormé $\{u_1, \ldots, u_n\}$ qui est indépendant et donc qui est une base de V. La construction de cet ensemble montre bien que la matrice de passage est triangulaire.

Exemple 13.12 : Considérons la base suivante de l'espace euclidien $\mathbf{R}^3$:

$$\{v_1 = (1,1,1),\ v_2 = (0,1,1),\ v_3 = (0,0,1)\}$$

Utilisons le procédé d'orthogonalisation de Gram-Schmidt pour transformer $\{v_i\}$ en une base orthonormée $\{u_i\}$. Normalisons d'abord v_1, posons.

$$u_1 = \frac{v_1}{\|v_1\|} = \frac{(1,1,1)}{\sqrt{3}} = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$$

Posons ensuite

$$w_2 = v_2 - \langle v_2, u_1\rangle u_1 = (0,1,1) - \frac{2}{\sqrt{3}}\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right) = \left(-\frac{2}{3}, \frac{1}{3}, \frac{1}{3}\right)$$

et normalisons w_2, on pose alors

$$u_2 = \frac{w_2}{\|w_2\|} = \left(-\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right)$$

Finalement posons

$$\begin{aligned} w_3 &= v_3 - \langle v_3, u_1\rangle u_1 - \langle v_3, u_2\rangle u_2 \\ &= (0,0,1) - \frac{1}{\sqrt{3}}\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right) - \frac{1}{\sqrt{6}}\left(-\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right) = \left(0, -\frac{1}{2}, \frac{1}{2}\right) \end{aligned}$$

puis normalisons w_3 :

$$u_3 = \frac{w_3}{\|w_3\|} = \left(0, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$$

La base orthonormée de $\mathbf{R}^3$ demandée est donc

$$\left\{u_1 = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right),\ u_2 = \left(-\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right),\ u_3 = \left(0, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)\right\}$$

FORMES LINEAIRES ET OPERATEURS ADJOINTS

Soit V un espace préhilbertien. Chaque $u \in V$ détermine une application $\hat{u} : V \to K$ définie par

$$\hat{u}(v) = \langle v, u\rangle$$

Donc pour tout $a, b \in K$ et pour tout $v_1, v_2 \in V$,

$$\hat{u}(av_1 + bv_2) = \langle av_1 + bv_2, u\rangle = a\langle v_1, u\rangle + b\langle v_2, u\rangle = a\hat{u}(v_1) + b\hat{u}(v_2)$$

C'est-à-dire, $\hat{u}$ est une forme linéaire sur V. La réciproque est aussi vraie pour des espaces de dimension finie et constitue un théorème important.

Théorème 13.5 : Soit ϕ une forme linéaire sur un espace préhilbertien de dimension finie V. Il existe alors un vecteur unique $u \in V$ tel que $\phi(v) = \langle v, u\rangle$, pour chaque $v \in V$.

Remarquons que le théorème précédent n'est pas vrai pour des espaces de dimension infinie (Problème 13-45), quoiqu'on ait quelques résultats dans cette direction. L'un d'entre eux est connu sous le nom de théorème de représentation de Riesz).

Utilisons le précédent théorème pour démontrer le théorème suivant

Théorème 13.6 : Soit T un opérateur linéaire sur un espace préhilbertien V de dimension finie. Il existe alors un opérateur linéaire unique T^* sur V tel que

$$\langle T(u), v\rangle = \langle u, T^*(v)\rangle$$

pour tout $(u, v) \in V$. De plus, si A est la matrice de T relativement à une base orthonormée $\{e_i\}$ de V, alors la conjuguée transposée A^* de A est la matrice de T^* dans la base $\{e_i\}$.

Insistons sur le fait qu'il n'existe pas de telle relation simple entre les matrices représentant T et T^* si la base n'est pas orthonormée. Nous voyons ici une propriété très utile des bases orthonormées.

Définition : Un opérateur linéaire T sur un espace préhilbertien V est dit admettre un opérateur adjoint T^* sur V si $\langle T(u), v\rangle = \langle u, T^*(v)\rangle$ pour chaque $u, v \in V$.

Le théorème 13.6 établit ainsi que chaque opérateur T a un adjoint si V est de dimension finie. Ce théorème n'est pas valable si V est de dimension infinie (Problème 13.78).

Exemple 13.13 : Soit T l'opérateur linéaire de $\mathbf{C}^3$ défini par

$$T(x, y, z) = (2x + iy, y - 5iz, x + (1-i)y + 3z)$$

On trouve une formule analogue pour l'adjoint T^* de T. Remarquons (Problème 7.3) que la matrice de T dans la base usuelle de $\mathbf{C}^3$ est

$$[T] = \begin{pmatrix} 2 & i & 0 \\ 0 & 1 & -5i \\ 1 & 1-i & 3 \end{pmatrix}$$

Rappelons que la base usuelle est orthonormée. Ainsi d'après le théorème 13.6, la matrice de T^* dans cette base est la transposée conjuguée de $[T]$:

$$[T^*] = \begin{pmatrix} 2 & 0 & 1 \\ -i & 1 & 1+i \\ 0 & 5i & 3 \end{pmatrix}$$

En conséquence,

$$T^*(x, y, z) = (2x + z, -ix + y + (1+i)z, 5iy + 3z)$$

Le théorème suivant résume quelques propriétés de l'adjoint.

Théorème 13.7 : Soient S et T des opérateurs linéaires sur V et soit $k \in K$. Alors

(i) $(S+T)^* = S^* + T^*$ (iii) $(ST)^* = T^*S^*$

(ii) $(kT)^* = \bar{k}T^*$ (iv) $(T^*)^* = T$

ANALOGIE ENTRE $A(V)$ et C ; OPERATEURS SPECIAUX

Appelons $A(V)$ l'algèbre des opérateurs linéaires sur un espace préhilbertien V de dimension finie. L'application adjointe $T \mapsto T^*$ sur $A(V)$ est tout à fait analogue à l'application conjuguée $z \mapsto \bar{z}$ sur le corps des complexes C. Pour illustrer cette analogie, nous identifierons dans le tableau suivant certaines classes d'opérateurs $T \in A(V)$ dont le comportement par l'application adjointe est analogue au comportement par conjugaison de certaines classes familières de nombres complexes.

Classe de nombres complexes	Comportemant par conjugaison	Classe d'opérateurs dans $A(v)$	Comportement par l'application adjointe
Cercle unité ($\lvert z\rvert = 1$)	$\bar{z} = 1/z$	Opérateurs orthogonaux (cas réel) Opérateurs unitaires (cas complexe)	$T^* = T^{-1}$
Axe réel	$\bar{z} = z$	Opérateurs auto adjoints Appelés aussi symétriques (cas réel) hermitiens (cas complexe)	$T^* = T$
Axe imaginaire	$\bar{z} = -z$	Opérateur anti-adjoints – antisymétriques (cas réel) – antihermitiens (cas complexe)	$T^* = -T$
Demi-axe positif $(0, +\infty)$	$z = \bar{w}.w, w \neq 0$	Opérateurs définis positifs	$T = S^*S$ avec S non singulier

L'analogie entre ces classes d'opérateurs T et les nombres complexes z est donnée dans le théorème suivant

Théorème 13.8 : Soit λ une valeur propre d'un opérateur linéaire T sur V.

(1) Si $T^* = T^{-1}$ alors $|\lambda| = 1$.

(2) Si $T^* = T$, alors λ est réel.

(3) Si $T^* = -T$, alors λ est imaginaire pur.

(4) Si $T = S^*S$ avec S non singulière, alors λ est réel et positif.

Nous démontrons maintenant le théorème ci-dessus. Dans chaque cas, soit v un vecteur propre non nul de T appartenant à λ, c'est-à-dire $T(v) = \lambda v$ avec $v \neq 0$; donc $\langle v, v\rangle$ est positif.

Démonstration de (1) : Montrons que $\lambda\bar{\lambda}\langle v, v\rangle = \langle v, v\rangle$.

$$\lambda\bar{\lambda}\langle v,v\rangle = \langle \lambda v, \lambda v\rangle = \langle T(v), T(v)\rangle = \langle v, T^*T(v)\rangle = \langle v, I(v)\rangle = \langle v,v\rangle$$

Mais $\langle v, v\rangle \neq 0$: donc $\lambda\bar{\lambda} = 1$ et donc $|\lambda| = 1$.

Démonstration de (2) : Montrons que $\lambda\langle v, v\rangle = \bar{\lambda}\langle v, v\rangle$

$$\lambda\langle v,v\rangle = \langle \lambda v, v\rangle = \langle T(v), v\rangle = \langle v, T^*(v)\rangle = \langle v, T(v)\rangle = \langle v, \lambda v\rangle = \bar{\lambda}\langle v,v\rangle$$

Mais $\langle v, v\rangle \neq 0$; donc $\lambda = \bar{\lambda}$ et donc λ est réel.

Démonstration de (3) : Montrons que $\lambda\langle v, v\rangle = -\bar{\lambda}\langle v, v\rangle$.

$$\lambda\langle v,v\rangle = \langle \lambda v, v\rangle = \langle T(v), v\rangle = \langle v, T^*(v)\rangle = \langle v, -T(v)\rangle = \langle v, -\lambda v\rangle = -\bar{\lambda}\langle v,v\rangle$$

Mais $\langle v, v\rangle \neq 0$; donc $\lambda = -\bar{\lambda}$ ou $\bar{\lambda} = -\lambda$ et donc λ est un imaginaire pur.

Démonstration de (4) : Remarquons d'abord que $S(v) \neq 0$, car S est non singulière ; donc $\langle S(v), S(v)\rangle$ est positif. Montrons que $\lambda \langle v, v\rangle = \langle S(v), S(v)\rangle$.

$$\lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle T(v), v\rangle = \langle S^*S(v), v\rangle = \langle S(v), S(v)\rangle$$

Mais $\langle v, v\rangle$ et $\langle S(v), S(v)\rangle$ sont positifs ; donc λ est positif.

Remarquons que tous les opérateurs précédents T commutent avec leur adjoint, donc $TT^* = T^*T$. De tels opérateurs sont appelés opérateurs normaux.

OPERATEURS UNITAIRES ET ORTHOGONAUX

Soit U un opérateur linéaire sur un espace préhilbertien V de dimension finie. Comme on l'a défini plus haut, si

$$U^* = U^{-1} \text{ ou de façon équivalente } UU^* = U^*U = I$$

alors U est dit orthogonal ou unitaire suivant que le corps sous-jacent est réel ou complexe. Le théorème suivant donne diverses caractérisations de ces opérateurs.

Théorème 13.9 : Les conditions suivantes sur un opérateur U sont équivalentes.

(i) $U^* = U^{-1}$, c'est-à-dire $UU^* = U^*U = I$.

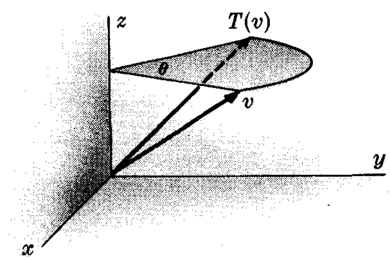
(ii) U conserve les produits scalaires, c'est-à-dire pour chaque $v, w \in V$ $\langle U(v), U(w)\rangle = \langle v, w\rangle$.

(iii) U conserve les longueurs, c'est-à-dire pour chaque $v \in V$ $\|U(v)\| = \|v\|$.

Exemple 13.14 : Soit $T : \mathbf{R}^3 \to \mathbf{R}^3$ l'opérateur linéaire qui fait subir à chaque vecteur une rotation autour de l'axe des z d'un angle fixe θ:

$$T(x, y, z) = (x \cos\theta - y \sin\theta,\ x \sin\theta + y \cos\theta,\ z)$$

Observons que les longueurs (les distances à l'origine) sont conservées par T. Donc T est un opérateur orthogonal.



Exemple 13.15 : Soit V l'espace l_2 de l'exemple 13.5. Soit $T : V \to V$ l'opérateur linéaire défini par $T(a_1, a_2, \ldots) = (0, a_1, a_2, \ldots)$. Il est clair que T conserve les produits scalaires et les longueurs. Cependant T n'est pas surjectif puisque, par exemple, $(1, 0, 0, \ldots)$ n'appartient pas à l'image de T ; donc T n'est pas inversible. Ainsi nous remarquons que le théorème 13.9 n'est pas valable pour des espaces de dimension infinie.

Un isomorphisme d'un espace préhilbertien dans un autre est une application bijective qui conserve les trois opérations de base d'un espace préhilbertien, addition vectorielle, multiplication scalaire, et les produits scalaires. Ainsi les applications précédentes (orthogonale et unitaire) peuvent être aussi caractérisées comme des isomorphismes de V dans lui-même. Remarquons qu'une telle application U conserve aussi les distances, puisque

$$\|U(v) - U(w)\| = \|U(v - w)\| = \|v - w\|$$

et donc U est aussi appelé une isométrie.

MATRICES ORTHOGONALES ET UNITAIRES

Soit U un opérateur linéaire sur un espace préhilbertien V. D'après le théorème 13.6 nous obtenons le résultat suivant quand le corps de base K est complexe.

Théorème 13.10A : Une matrice A à éléments complexes représente un opérateur unitaire U (relativement à une base orthonormée) si et seulement si $A^* = A^{-1}$.

D'autre part, si le corps de base est réel alors $A^* = A^t$; donc nous obtenons ainsi le théorème suivant pour les espaces préhilbertiens réels.

Théorème 13.10B : Une matrice A à éléments réels représente un opérateur orthogonal U (relativement à une base orthonormée) si et seulement si $A^t = A^{-1}$.

Les théorèmes précédents nous conduisent aux définitions.

Définition : Une matrice complexe A pour laquelle $A^* = A^{-1}$, ou de façon équivalente $AA^* = A^*A = I$ est appelée une matrice unitaire.

Définition : Une matrice réelle A pour laquelle $A^t = A^{-1}$, ou de façon équivalente $AA^t = A^tA = I$ est appelée une matrice orthogonale.

Remarquons qu'une matrice unitaire à éléments réels est orthogonale.

Exemple 13.16 : Supposons que $A = \begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix}$ soit une matrice unitaire. Alors $AA^* = I$ et donc

$$AA^* = \begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix}\begin{pmatrix} \bar{a}_1 & \bar{b}_1 \\ \bar{a}_2 & \bar{b}_2 \end{pmatrix} = \begin{pmatrix} |a_1|^2 + |a_2|^2 & a_1\bar{b}_1 + a_2\bar{b}_2 \\ \bar{a}_1 b_1 + \bar{a}_2 b_2 & |b_1|^2 + |b_2|^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$$

Ainsi

$$|a_1|^2 + |a_2|^2 = 1, \quad |b_1|^2 + |b_2|^2 = 1 \quad \text{et} \quad a_1\bar{b}_1 + a_2\bar{b}_2 = 0$$

En conséquence, les lignes de A forment un ensemble orthonormé. De façon analogue $A^*A = I$ signifie que les colonnes de A forment un ensemble orthonormé.

Le résultat de l'exemple précédent reste vrai en général, ce qui donne

Théorème 13.11 : Les conditions suivantes pour une matrice A sont équivalentes.

(i) A est unitaire (orthogonale).

(ii) Les lignes de A forment un ensemble orthonormé.

(iii) Les colonnes de A forment un ensemble orthonormé.

Exemple 13.17 : La matrice A représentant la rotation T dans l'exemple 13.14 relativement à la base usuelle de $\mathbf{R}^3$ est

$$A = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

On vérifie bien que les lignes et les colonnes de A forment chacun un ensemble orthonormé ; c'est-à-dire que A est une matrice orthogonale.

CHANGEMENT DE BASES ORTHONORMEES

Si l'on considère le rôle spécial joué par les bases orthonormées dans la théorie des espaces préhilbertiens, on est amené naturellement à considérer les propriétés de la matrice de passage d'une base orthonormée à une autre base orthonormée. Le théorème suivant s'applique alors:

Théorème 13.12 : Soit $\{e_1, \ldots, e_n\}$ une base orthonormée d'un espace préhilbertien V. La matrice de passage de $\{e_i\}$ à une autre base orthonormée est unitaire (orthogonale). Réciproquement, si $P = (a_{ij})$ est une matrice unitaire (orthogonale) alors la base suivante est orthonormée :

$$\{e_i' = a_{1i}e_1 + a_{2i}e_2 + \cdots + a_{ni}e_n : \; i = 1, \ldots, n\}$$

Rappelons que les matrices A et B représentant le même opérateur linéaire T sont semblables, c'est-à-dire $B = P^{-1}AP$ où P est la matrice de passage (non singulière). D'autre part, si V est un espace préhilbertien nous sommes intéressés d'habitude par le cas où P est unitaire (ou orthogonale) comme le théorème précédent le suggère. (Rappelons que P est unitaire si $P^* = P^{-1}$ et P est orthogonale si $P^t = P^{-1}$). Ceci nous conduit à la définition suivante :

Définition : Des matrices complexes A et B sont unitairement équivalentes s'il existe une matrice unitaire P pour laquelle $B = P^*AP$. De façon analogue, les matrices réelles A et B sont orthogonalement équivalentes s'il existe une matrice orthogonale P pour laquelle $B = P^tAP$.

Remarquons que des matrices orthogonalement équivalentes sont nécessairement congruentes (voir page 262).

OPERATEURS POSITIFS

Soit P un opérateur linéaire sur un espace préhilbertien V ; P est dit positif (ou semi-défini) si

$$P = S^*S \text{ pour un certain opérateur } S$$

et il est dit défini positif si S est aussi non singulier. Les théorèmes suivants donnent diverses caractérisations de ces opérateurs.

Théorème 13.13A : Les conditions suivantes sur un opérateur P sont équivalentes :

(1) $P = T^2$ pour un certain opérateur auto-adjoint T.
(2) $P = S^*S$ pour un certain opérateur S.
(3) P est auto-adjoint et $\langle P(u), u\rangle \geqslant 0$ pour chaque $u \in V$.

Le théorème correspondant pour les opérateurs définis positifs est alors

Théorème 13.13B : Les conditions suivantes sur un opérateur P sont équivalentes :

(1) $P = T^2$ pour un certain opérateur non singulier auto-adjoint T.
(2) $P = S^*S$ pour un certain opérateur non singulier S.
(3) P est auto-adjoint et $\langle P(u), u\rangle > 0$ pour chaque $u \neq 0$ dans V.

DIAGONALISATION ET FORMES CANONIQUES DANS LES ESPACES EUCLIDIENS

Soit T un opérateur linéaire sur un espace préhilbertien V de dimension finie sur K. La possibilité de représenter T par une matrice diagonale dépend des vecteurs propres et des valeurs propres de T, et donc dépend des racines du polynôme caractéristique $\Delta(t)$ de T(Théorème 9.6). Mais $\Delta(t)$ peut toujours se décomposer en un produit de polynômes linéaires sur le corps des complexes **C**, mais ne peut pas toujours se décomposer en produit de facteurs linéaires sur le corps des réels **R**. Ainsi la situation pour les espaces euclidiens (où $K = \mathbf{R}$) est complètement différente de celle pour les espaces unitaires (où $K = \mathbf{C}$) ; nous traiterons donc ceux-ci à part. Nous étudierons les espaces euclidiens d'abord, et les espaces unitaires dans le prochain paragraphe.

Théorème 13.14 : Soit T un opérateur symétrique (auto-adjoint réel) sur un espace préhilbertien V réel de dimension finie. Il existe alors une base orthonormée de V contenant uniquement les vecteurs propres de T ; c'est-à-dire que T peut être représenté par une matrice diagonale relativement à une base orthonormée.

Donnons les théorèmes correspondants pour les matrices.

Autre forme du théorème 13.14 : Soit A une matrice réelle symétrique. Il existe alors une matrice orthogonale P telle que $B = P^{-1}AP = P^tAP$ soit diagonale.

Les colonnes de la matrice précédente P forment un système orthonormal de vecteurs propres de A ; et les éléments diagonaux de B sont les valeurs propres correspondantes.

Exemple 13.18 : Soit $A = \begin{pmatrix} 2 & -2 \\ -2 & 5 \end{pmatrix}$. Trouvons une matrice orthogonale P telle que P^tAP soit diagonale. Le polynôme caractéristique $\Delta(t)$ de A est

$$\Delta(t) = |tI - A| = \begin{vmatrix} t-2 & 2 \\ 2 & t-5 \end{vmatrix} = (t-6)(t-1)$$

Les valeurs propres de A sont 6 et 1. Remplaçons t par 6 dans la matrice $tI - A$, on obtient le système homogène correspondant d'équations linéaires

$$4x + 2y = 0, \quad 2x + y = 0$$

Une solution non nulle est $v_1 = (1, -2)$. Remplaçons maintenant t par 1 dans la matrice $tI - A$, on trouve le système homogène correspondant

$$-x + 2y = 0, \quad 2x - 4y = 0$$

Une solution non nulle est (2, 1). Comme dans le Problème 13.31, v_1 et v_2 sont orthogonaux. Normalisons v_1 et v_2, on obtient la base orthonormée

$$\{u_1 = (1/\sqrt{5}, -2/\sqrt{5}), \; u_2 = (2/\sqrt{5}, 1/\sqrt{5})\}$$

Finalement soit P la matrice dont les colonnes sont u_1 et u_2 respectivement. Alors

$$P = \begin{pmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ -2/\sqrt{5} & 1/\sqrt{5} \end{pmatrix} \quad \text{et} \quad P^{-1}AP = P^tAP = \begin{pmatrix} 6 & 0 \\ 0 & 1 \end{pmatrix}$$

Comme prévu, les éléments diagonaux de P^tAP sont les valeurs propres correspondant aux colonnes de P.

Nous remarquons que la matrice $B = P^{-1}AP = P^tAP$ est aussi congruente avec la matrice A. Si q est une forme réelle quadratique représentée par la matrice A, alors la méthode précédente peut être utilisée pour diagonaliser q par un changement orthogonal de coordonnées. Ceci est illustré par l'exemple suivant :

Exemple 13.19 : Trouver une transformation orthogonale de coordonnées qui diagonalise la forme quadratique $q(x, y) = 2x^2 - 4xy + 5y^2$.

La matrice symétrique représentant q est $A = \begin{pmatrix} 2 & -2 \\ -2 & 5 \end{pmatrix}$. Dans le précédent exemple nous avons obtenu la matrice orthogonale

$$P = \begin{pmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ -2/\sqrt{5} & 1/\sqrt{5} \end{pmatrix} \quad \text{pour laquelle} \quad P^tAP = \begin{pmatrix} 6 & 0 \\ 0 & 1 \end{pmatrix}$$

(Ici 6 et 1 sont les valeurs propres de A). Ainsi la transformation orthogonale des coordonnées demandée est

$$\begin{pmatrix} x \\ y \end{pmatrix} = P\begin{pmatrix} x' \\ y' \end{pmatrix} \quad \text{c'est-à-dire} \quad \begin{aligned} x &= x'/\sqrt{5} + 2y'\sqrt{5} \\ y &= -2x'/\sqrt{5} + y'/\sqrt{5} \end{aligned}$$

Par ce changement de coordonnées q est transformée en la forme diagonale

$$q(x', y') = 6x'^2 + y'^2$$

Remarquons que les éléments diagonaux de q sont les valeurs propres de A.

Un opérateur orthogonal T n'est pas nécessairement symétrique et donc ne peut être représenté par une matrice diagonale relativement à une base orthonormée. Cependant, un tel opérateur T peut avoir une représentation canonique simple, comme il est dit dans le théorème suivant :

Théorème 13.15 : Soit T un opérateur orthogonal sur un espace préhilbertien réel V. Il existe alors une base orthonormée par rapport à laquelle T a la forme suivante :

$$\begin{pmatrix}
1 & & & & & & & & & & & & \\
& 1 & & & & & & & & & & & \\
& & \ddots & & & & & & & & & & \\
& & & 1 & & & & & & & & & \\
& & & & -1 & & & & & & & & \\
& & & & & -1 & & & & & & & \\
& & & & & & \ddots & & & & & & \\
& & & & & & & -1 & & & & & \\
& & & & & & & & \cos\theta_1 & -\sin\theta_1 & & & \\
& & & & & & & & \sin\theta_1 & \sin\theta_1 & & & \\
& & & & & & & & & & \ddots & & \\
& & & & & & & & & & & \cos\theta_r & -\sin\theta_r \\
& & & & & & & & & & & \sin\theta_r & \cos\theta_r
\end{pmatrix}$$

Le lecteur reconnaîtra que les blocs diagonaux 2 à 2 représentent des rotations dans les sous-espaces à 2 dimensions correspondants.

DIAGONALISATION ET FORMES CANONIQUES DANS DES ESPACES UNITAIRES

Présentons maintenant le théorème fondamental de diagonalisation pour les espaces préhilbertiens complexes c'est-à-dire pour les espaces unitaires. Rappelons qu'un opérateur T est dit normal s'il commute avec son adjoint, c'est-à-dire si $TT^* = T^*T$. De façon analogue, une matrice complexe A est dite normale si elle commute avec sa transposée conjuguée, c'est-à-dire si $AA^* = A^*A$.

Exemple 13.20 : Soit $A = \begin{pmatrix} 1 & 1 \\ i & 3+2i \end{pmatrix}$. Alors

$$AA^* = \begin{pmatrix} 1 & 1 \\ i & 3+2i \end{pmatrix}\begin{pmatrix} 1 & -i \\ 1 & 3-2i \end{pmatrix} = \begin{pmatrix} 2 & 3-3i \\ 3+3i & 14 \end{pmatrix}$$

$$A^*A = \begin{pmatrix} 1 & -i \\ 1 & 3-2i \end{pmatrix}\begin{pmatrix} 1 & 1 \\ i & 3+2i \end{pmatrix} = \begin{pmatrix} 2 & 3-3i \\ 3+3i & 14 \end{pmatrix}$$

Ainsi A est une matrice normale.

Le théorème suivant s'applique alors :

Théorème 13.16 : Soit T un opérateur normal sur un espace préhilbertien complexe de dimension finie V. Il existe alors une base orthonormée de V contenant uniquement les vecteurs propres de T ; c'est-à-dire que T peut être représenté par une matrice diagonale relativement à une base orthonormée.

Donnons les théorèmes correspondants pour les matrices.

Autre forme du théorème 13.16 : Soit A une matrice normale. Il existe alors une matrice unitaire P telle que $B = P^{-1}AP = P^*AP$ soit diagonale.

Le théorème suivant montre que même des opérateurs non normaux sur des espaces unitaires ont une forme relativement simple.

Théorème 13.17 : Soit T un opérateur quelconque sur un espace préhilbertien complexe de dimension finie V. Alors T peut être représenté par une matrice triangulaire relativement à une base orthonormée de V.

Autre forme du théorème 13.17 : Soit A une matrice quelconque complexe. Il existe alors une matrice unitaire P telle que $B = P^{-1}AP = P^*AP$ soit triangulaire.

THEOREME SPECTRAL

Le théorème spectral est une autre formulation des théorèmes 13.14 et 13.16 de diagonalisation.

Théorème 13.18 (Théorème spectral) : Soit T un opérateur normal (symétrique) sur un espace préhilbertien complexe (réel) de dimension finie V. Il existe alors des projections orthogonales $E_1, E_2, \ldots, E_r$ sur V et des scalaires $\lambda_1, \lambda_2, \ldots, \lambda_r$ tels que

(1) $T = \lambda_1 E_1 + \lambda_2 E_2 + \cdots + \lambda_r E_r$

(2) $E_1 + E_2 + \cdots E_r = I$

(3) $E_i E_j = 0$ pour $i \neq j$

L'exemple suivant nous montre une relation importante entre une représentation matricielle diagonale et les projections orthogonales correspondantes.

Exemple 13.21 : Considérons une matrice diagonale, c'est-à-dire $A = \begin{pmatrix} 2 & & & \\ & 3 & & \\ & & 3 & \\ & & & 5 \end{pmatrix}$. Soit

$$E_1 = \begin{pmatrix} 1 & & & \\ & 0 & & \\ & & 0 & \\ & & & 0 \end{pmatrix}, \quad E_2 = \begin{pmatrix} 0 & & & \\ & 1 & & \\ & & 1 & \\ & & & 0 \end{pmatrix}, \quad E_3 = \begin{pmatrix} 0 & & & \\ & 0 & & \\ & & 0 & \\ & & & 1 \end{pmatrix}$$

Le lecteur peut vérifier que les E_i sont des projections, c'est-à-dire que $E_i^2 = E_i$ et que (i) $A = 2E_1 + 3E_2 + 5E_3$, (ii) $E_1 + E_2 + E_3 = I$, (iii) $E_iE_j = 0$ pour $i \neq j$.

PROBLEMES RESOLUS

PRODUITS SCALAIRES

13.1. Vérifier la relation $\langle u, av_1 + bv_2\rangle = \bar{a}\langle u, v_1\rangle + \bar{b}\langle u, v_2\rangle$.

En utilisant $[I_2]$, $[I_1]$ et enfin $[I_2]$, nous trouvons

$$\begin{aligned}\langle u, av_1 + bv_2\rangle &= \overline{\langle av_1 + bv_2, u\rangle} = \overline{a\langle v_1,u\rangle + b\langle v_2, u\rangle} \\ &= \bar{a}\,\overline{\langle v_1, u\rangle} + \bar{b}\,\overline{\langle v_2, u\rangle} = \bar{a}\langle u, v_1\rangle + \bar{b}\langle u, v_2\rangle\end{aligned}$$

13.2. Vérifier que l'expression suivante est un produit scalaire sur $\mathbf{R}^2$:

$\langle u, v\rangle = x_1y_1 - x_1y_2 - x_2y_1 + 3x_2y_2$, où $u = (x_1, x_2)$, $v = (y_1, y_2)$.

Méthode 1.

Vérifions les trois axiomes d'un produit scalaire. En posant $w = (z_1, z_2)$, nous trouvons

$$au + bw = a(x_1, x_2) + b(z_1, z_2) = (ax_1 + bz_1, ax_2 + bz_2)$$

Ainsi

$$\begin{aligned}\langle au+bw,\, v\rangle &= \langle (ax_1+bz_1,\, ax_2+bz_2),\, (y_1, y_2)\rangle \\ &= (ax_1+bz_1)y_1 - (ax_1+bz_1)y_2 - (ax_2+bz_2)y_1 + 3(ax_2+bz_2)y_2 \\ &= a(x_1y_1 - x_1y_2 - x_2y_1 + 3x_2y_2) + b(z_1y_1 - z_1y_2 - z_2y_1 + 3z_2y_2) \\ &= a\langle u, v\rangle + b\langle w, v\rangle\end{aligned}$$

et donc l'axiome $[I_1]$ est satisfait. On a aussi

$$\langle v, u\rangle = y_1x_1 - y_1x_2 - y_2x_1 + 3y_2x_2 = x_1y_1 - x_1y_2 - x_2y_1 + 3x_2y_2 = \langle u, v\rangle$$

et donc l'axiome $[I_2]$ est satisfait. Finalement,

$$\langle u, u\rangle = x_1^2 - 2x_1x_2 + 3x_2^2 = x_1^2 - 2x_1x_2 + x_2^2 + 2x_2^2 = (x_1-x_2)^2 + 2x_2^2 \geq 0$$

Donc $\langle u, u\rangle = 0$ si et seulement si $x_1 = 0$, $x_2 = 0$, c'est-à-dire si $u = 0$. Donc le dernier axiome $[I_3]$ est satisfait.

Méthode 2 :

Raisonnons à l'aide de matrices. Nous pouvons écrire $\langle u, v\rangle$ sous forme matricielle :

$$\langle u, v\rangle = u^tAv = (x_1, x_2)\begin{pmatrix} 1 & -1 \\ -1 & 3 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

donc $[I_1]$ est vrai. Puisque A est symétrique, $[I_2]$ est vrai. Il nous reste donc uniquement à montrer que A est positive définie. Appliquons l'opération élémentaire ligne $R_2 \to R_1 + R_2$ puis l'opération élémentaire colonne correspondante $C_2 \to C_1 + C_2$, on transforme A en la forme diagonale $\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$. Ainsi A est positive et $[I_3]$ est vrai.

13.3. Trouver la norme de $v = (3, 4) \in \mathbf{R}^2$ par rapport (i) au produit scalaire ordinaire, (ii) au produit scalaire du Problème 13.2.

(i) $\|v\|^2 = \langle v, v\rangle = \langle (3,4), (3,4)\rangle = 9 + 16 = 25$; donc $\|v\| = 5$.

(ii) $\|v\|^2 = \langle v, v\rangle = \langle (3,4), (3,4)\rangle = 9 - 12 - 12 + 48 = 33$; donc $\|v\| = \sqrt{33}$.

13.4. Normaliser chacun des vecteurs suivants dans l'espace euclidien $\mathbf{R}^3$.
(i) $u = (2, 1, -1)$, (ii) $v = (\frac{1}{2}, \frac{2}{3}, -\frac{1}{4})$.

(i) Appelons $\langle u, u\rangle$ la somme des carrés des éléments de u ; c'est-à-dire $\langle u, u\rangle = 2^2 + 1^2 + (-1)^2 = 6$. Divisons u par $\|u\| = \sqrt{\langle u, u\rangle} = \sqrt{6}$ on obtient le vecteur unitaire désiré

$$u/\|u\| = (2/\sqrt{6},\, 1/\sqrt{6},\, -1/\sqrt{6})$$

(ii) Multiplions d'abord v par 12 pour éliminer les fractions ; on a $12v = (6, 8, -3)$. Nous avons $\langle 12v, 12v\rangle = 6^2 + 8^2 + (-3)^2 = 109$; le vecteur unitaire désiré est donc

$$12v/\|12v\| = (6/\sqrt{109},\, 8/\sqrt{109},\, -3/\sqrt{109})$$

13.5. Soit V l'espace vectoriel des polynômes muni du produit scalaire donné par $\langle f, g\rangle = \int_0^1 f(t)\, g(t)\, dt$. Soit $f(t) = t + 2$ et $g(t) = t^2 - 2t - 3$. Trouver (i) $\langle f, g\rangle$ et (ii) $\|f\|$.

(i) $\langle f, g\rangle = \int_0^1 (t+2)(t^2-2t-3)\, dt = \Big[t^4/4 - 7t^2/2 - 6t\Big]_0^1 = -37/4$

(ii) $\langle f, f\rangle = \int_0^1 (t+2)(t+2)\, dt = 19/3$ et $\|f\| = \sqrt{\langle f, f\rangle} = \sqrt{19/3}$

13.6. Démontrer le théorème 13.1 (Cauchy-Schwarz) : $|\langle u\,,\,v\rangle| \leqslant \|u\|\,.\,\|v\|$.

Si $v = 0$, l'inégalité se réduit à $0 \leqslant 0$ qui est vraie. Supposons maintenant $v \neq 0$, et utilisons $z\bar{z} = |z|^2$ (qui est valable pour tout nombre complexe z) et $\langle v\,,\,u\rangle = \overline{\langle u\,,\,v\rangle}$, nous savons que $\|u - \langle u\,,\,v\rangle\, tv\|^2 \geqslant$ où t est une valeur réelle quelconque :

$$\begin{aligned} 0 &\leq \|u - \langle u,v\rangle tv\|^2 = \langle u - \langle u,v\rangle tv,\ u - \langle u,v\rangle tv\rangle \\ &= \langle u,u\rangle - \overline{\langle u,v\rangle} t\langle u,v\rangle - \langle u,v\rangle t\langle v,u\rangle + \langle u,v\rangle\overline{\langle u,v\rangle} t^2\langle v,v\rangle \\ &= \|u\|^2 - 2t\,|\langle u,v\rangle|^2 + |\langle u,v\rangle|^2\, t^2\,\|v\|^2 \end{aligned}$$

Posons $t = 1/\|v\|^2$, on trouve $0 \leqslant \|u\|^2 - \dfrac{|\langle u\,,\,v\rangle|^2}{\|v\|^2}$, d'où l'on tire $|\langle u\,,\,v\rangle|^2 \leqslant \|u\|^2\,\|v\|^2$. En prenant la racine carrée des deux membres, on obtient l'inégalité demandée.

13.7. Démontrer que la norme dans un espace préhilbertien satisfait les axiomes suivants :

$[N_1]$ $\|v\| \geqslant 0$ et $\|v\| = 0$ si et seulement si $v = 0$.

$[N_2]$ $\|kv\| = |k|\ \|v\|$.

$[N_3]$ $\|u + v\| \leqslant \|u\| + \|v\|$.

D'après $[I_3]$ $\langle v\,,\,v\rangle \geqslant 0$; donc $\|v\| = \sqrt{\langle v\,,\,v\rangle} \geqslant 0$. De plus, $\|v\| = 0$ si et seulement si $\langle v\,,\,v\rangle = 0$ ce qui est vrai si et seulement si $v = 0$. Ainsi $[N_1]$ est démontré.

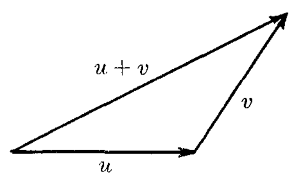
On trouve $\|kv\|^2 = \langle kv, kv\rangle = k\bar{k}\ \langle v\,,\,v\rangle = |k|^2\ \|v\|^2$. En prenant la racine carrée des deux membres, on trouve $[N_2]$.

En utilisant l'inégalité de Cauchy-Schwarz, on obtient

$$\begin{aligned} \|u+v\|^2 &= \langle u+v, u+v\rangle = \langle u,u\rangle + \langle u,v\rangle + \overline{\langle u,v\rangle} + \langle v,v\rangle \\ &\leq \|u\|^2 + 2\|u\|\,\|v\| + \|v\|^2 = (\|u\| + \|v\|)^2 \end{aligned}$$

En prenant la racine carrée des deux membres on trouve $[N_3]$.

Remarque : $[N_3]$ est fréquemment appelé l'inégalité triangulaire, car si nous considérons $u + v$ comme le côté d'un triangle formé par u et par v (comme sur la figure de droite), alors $[N_3]$ indique que la longueur d'un côté d'un triangle est inférieure ou égale à la somme des longueurs des deux autres côtés.



ORTHOGONALITE

13.8. Montrer que si u est orthogonal à v, alors tout multiple scalaire de u est aussi orthogonal à v. Trouver un vecteur unitaire orthogonal à $v_1 = (1\,,1\,,2)$ et $v_2 = (0\,,1\,,3)$ dans $\mathbf{R}^3$.

Si $\langle u\,,\,v\rangle = 0$ alors $\langle ku\,,\,v\rangle = k\,\langle u\,,\,v\rangle = k\,.\,0 = 0$ comme demandé. Soit $w = (x\,,y\,,z)$. Nous désirons avoir

$$0 = \langle w, v_1\rangle = x + y + 2z \quad \text{et} \quad 0 = \langle w, v_2\rangle = y + 3z$$

Nous obtenons ainsi le système homogène

$$x + y + 2z = 0 \quad , \quad y + 3z = 0$$

Posons $z = 1$, on trouve $y = -3$ et $x = 1$; alors $w = (1, -3, 1)$. En normalisant w, on obtient le vecteur unitaire w' orthogonal à v_1 et v_2 : $w' = w/\|w\| = (1/\sqrt{11},\ -3/\sqrt{11},\ 1/\sqrt{11})$.

13.9. Soit W le sous-espace de $\mathbf{R}^5$ engendré par $u = (1, 2, 3, -1, 2)$ et $v = (2, 4, 7, 2, -1)$. Trouver une base de l'orthogonal $W^\perp$ de W.

Nous cherchons tous les vecteurs $w = (x\,,y\,,z\,,s\,,t)$ tels que

$$\begin{aligned} \langle w,u\rangle &= x + 2y + 3z - s + 2t = 0 \\ \langle w,v\rangle &= 2x + 4y + 7z + 2s - t = 0 \end{aligned}$$

Eliminons x dans la seconde équation, nous trouvons le système équivalent

$$\begin{aligned} x + 2y + 3z - s + 2t &= 0 \\ z + 4s - 5t &= 0 \end{aligned}$$

Les inconnues libres sont y, s et t. Posons $y = -1$, $s = 0$, $t = 0$, on obtient la solution $w_1 = (2, -1, 0, 0, 0)$. Posons $y = 0$, $s = 1$, $t = 0$, on obtient la solution $w_2 = (13, 0, -4, 1, 0)$. Posons $y = 0$, $s = 0$, $t = 1$, on obtient la solution $w_3 = (-17, 0, 5, 0, 1)$. L'ensemble $\{w_1, w_2, w_3\}$ est une base de $W^\perp$.

13.10. Trouver une base orthonormée du sous-espace W de $\mathbf{C}^3$ engendré par $v_1 = (1, i, 0)$ et $v_2 = (1, 2, 1 - i)$.

Appliquons le procédé d'orthogonalisation de Gram-Schmidt. Normalisons tout d'abord v_1. On trouve

$$\|v_1\|^2 = \langle v_1, v_1\rangle = 1 \cdot 1 + i \cdot (-i) + 0 \cdot 0 = 2 \quad \text{et donc} \quad \|v_1\| = \sqrt{2}$$

Ainsi $u_1 = v_1/\|v_1\| = (1/\sqrt{2}, i/\sqrt{2}, 0)$.

Pour former $w_2 = v_2 - \langle v_2, u_1\rangle u_1$, calculons d'abord

$$\langle v_2, u_1\rangle = \langle (1, 2, 1-i), (1/\sqrt{2}, i/\sqrt{2}, 0)\rangle = 1/\sqrt{2} - 2i/\sqrt{2} = (1-2i)/\sqrt{2}$$

Alors
$$w_2 = (1, 2, 1-i) - \frac{1-2i}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}, \frac{i}{\sqrt{2}}, 0\right) = \left(\frac{1+2i}{2}, \frac{2-i}{2}, 1-i\right)$$

Normalisons maintenant w_2, ou de façon équivalence $2w_2 = (1 + 2i, 2 - i, 2 - 2i)$. Nous avons

$$\|2w_1\|^2 = \langle 2w_1, 2w_1\rangle = (1+2i)(1-2i) + (2-i)(2+i) + (2-2i)(2+2i) = 18$$

et $\|2w_1\| = \sqrt{18}$. Ainsi la base orthonormée de W est

$$\left\{u_1 = \left(\frac{1}{\sqrt{2}}, \frac{i}{\sqrt{2}}, 0\right),\ u_2 = \frac{2w_1}{\|2w_1\|} = \left(\frac{1+2i}{\sqrt{18}}, \frac{2-i}{\sqrt{18}}, \frac{2-2i}{\sqrt{18}}\right)\right\}$$

13.11. Démontrer le lemme 13.3 : Un ensemble orthonormé $\{u_1, \ldots, u_2\}$ est linéairement indépendant et pour un vecteur quelconque $v \in V$ le vecteur

$$w = v - \langle v, u_1\rangle u_1 - \langle v, u_2\rangle u_2 - \cdots - \langle v, u_r\rangle u_r$$

est orthogonal à chaque u_i.

Supposons $a_1 u_1 + \ldots + a_r u_r = 0$. Calculons le produit scalaire des deux membres par u_1,

$$\begin{aligned} 0 = \langle 0, u_1\rangle &= \langle a_1 u_1 + \cdots + a_r u_r, u_1\rangle \\ &= a_1\langle u_1, u_1\rangle + a_2\langle u_2, u_1\rangle + \cdots + a_r\langle u_r, u_1\rangle \\ &= a_1 \cdot 1 + a_2 \cdot 0 + \cdots + a_r \cdot 0 = a_1 \end{aligned}$$

ou $a_1 = 0$. De façon analogue, pour $i = 2, \ldots, r$,

$$\begin{aligned} 0 = \langle 0, u_i\rangle &= \langle a_1 u_1 + \cdots + a_r u_r, u_i\rangle \\ &= a_1\langle u_1, u_i\rangle + \cdots + a_i\langle u_i, u_i\rangle + \cdots + a_r\langle u_r, u_i\rangle = a_i \end{aligned}$$

En conséquence, $\{u_1, \ldots, u_r\}$ est linéairement indépendant.

Il reste à démontrer que w est orthogonal à chacun des u_i. Faisons le produit de w par u_1,

$$\begin{aligned} \langle w, u_1\rangle &= \langle v, u_1\rangle - \langle v, u_1\rangle\langle u_1, u_1\rangle - \langle v, u_2\rangle\langle u_2, u_1\rangle - \cdots - \langle v, u_r\rangle\langle u_r, u_1\rangle \\ &= \langle v, u_1\rangle - \langle v, u_1\rangle \cdot 1 - \langle v, u_2\rangle \cdot 0 - \cdots - \langle v, u_r\rangle \cdot 0 = 0 \end{aligned}$$

C'est-à-dire, w est orthogonal à u_1. De façon analogue pour $i = 2, \ldots, r$.

$$\langle w, u_i\rangle = \langle v, u_i\rangle - \langle v, u_1\rangle\langle u_1, u_i\rangle - \cdots - \langle v, u_i\rangle\langle u_i, u_i\rangle - \cdots - \langle v, u_r\rangle\langle u_r, u_i\rangle = 0$$

Ainsi w est orthogonal à u_i pour $i = 1, \ldots, r$ comme demandé.

13.12. Soit W un sous-espace d'un espace préhilbertien V. Montrer qu'il existe une base orthonormée de W qui est incluse dans une base orthonormée de V.

Choisissons une base $\{v_1, \ldots, v_r\}$ de W et étendons-la à une base $\{v_1, v_2, \ldots, v_n\}$ de V. Appliquons alors le procédé d'orthogonalisation de Gram-Schmidt à $\{v_1, \ldots, v_n\}$ pour obtenir une base orthonormée $\{u_1, \ldots, u_n\}$ de V où pour $i = 1, \ldots, n$, $u_i = a_{i1} v_1 + \ldots + a_{ii} v_i$. Ainsi $u_1, \ldots, u_r \in W$ et donc $\{u_1, \ldots, u_r\}$ est une base orthonormée de W.

13.13. Démontrer le théorème 13.2. Soit W un sous-espace de V ; on a alors $V = W \oplus W^{\perp}$.

D'après le problème 13.12, il existe une base orthonormée $\{u_1, \ldots, u_r\}$ de W qui est une partie d'une base orthonormée $\{u_1, \ldots, u_n\}$ de V. Puisque $\{u_1, u_2, \ldots, u_n\}$ est orthonormée, $u_{r+1}, \ldots, u_n \in W^{\perp}$. Si $v \in V$,

$$v = a_1u_1 + \cdots + a_nu_n \quad \text{où} \quad a_1u_1 + \cdots + a_ru_r \in W, \; a_{r+1}u_{r+1} + \cdots + a_nu_n \in W^{\perp}$$

En conséquence $V = W \oplus W^{\perp}$.

D'autre part, si $w \in W \cap W^{\perp}$, alors $\langle w, w\rangle = 0$. Ceci implique $w = 0$; donc $W \cap W^{\perp} = \{0\}$.

Les deux conditions $V = W + W^{\perp}$ et $W \cap W^{\perp} = \{0\}$ donnent le résultat désiré $V = W \oplus W^{\perp}$.

Remarquons que nous avons démontré le théorème uniquement dans le cas où V est de dimension finie ; remarquons qu'un théorème analogue est valable pour des espaces de dimension quelconque.

13.14. Soit W un sous-espace de W. Montrer que $W \subset W^{\perp\perp}$ et que $W = W^{\perp\perp}$ lorsque V est de dimension finie.

Soit $w \in W$. Alors $\langle w, v\rangle = 0$ pour chaque $v \in W^{\perp}$; donc $w \in W^{\perp\perp}$. En conséquence $W \subset W^{\perp\perp}$. Supposons maintenant V de dimension finie. D'après le théorème 13.2 $V = W \oplus W^{\perp}$ et aussi $V = W^{\perp} \oplus W^{\perp\perp}$. Donc

$$\dim W = \dim V - \dim W^{\perp} \quad \text{et} \quad \dim W^{\perp\perp} = \dim V - \dim W^{\perp}$$

Ceci donne $\dim W = \dim W^{\perp\perp}$. Mais $W \subset W^{\perp\perp}$, donc $W = W^{\perp\perp}$, comme il est demandé

13.15. Soit $\{e_1, \ldots, e_n\}$ une base orthonormée de V. Démontrer

(i) pour tout $u \in V$, $u = \langle u, e_1\rangle e_1 + \langle u, e_2\rangle e_2 + \cdots + \langle u, e_n\rangle e_n$;

(ii) $\langle a_1e_1 + \cdots + a_ne_n, b_1e_1 + \cdots + b_ne_n\rangle = a_1\overline{b_1} + a_2\overline{b_2} + \cdots + a_n\overline{b_n}$;

(iii) pour tout $u, v \in V$, $\langle u, v\rangle = \langle u, e_1\rangle\overline{\langle v, e_1\rangle} + \cdots + \langle u, e_n\rangle\overline{\langle v, e_n\rangle}$;

(iv) Si $T : V \to V$ est linéaire, alors $\langle T(e_j), e_i\rangle$ est le ij-élément de la matrice A représentant T dans la base donnée $\{e_i\}$.

(1) Supposons $u = k_1e_1 + k_2e_2 + \cdots + k_ne_n$. Considérons le produit scalaire de u par e_1.

(i)
$$\begin{aligned}\langle u, e_1\rangle &= \langle k_1e_1 + k_2e_2 + \cdots + k_ne_n, e_1\rangle \\ &= k_1\langle e_1, e_1\rangle + k_2\langle e_2, e_1\rangle + \cdots + k_n\langle e_n, e_1\rangle \\ &= k_1 \cdot 1 + k_2 \cdot 0 + \cdots + k_n \cdot 0 = k_1\end{aligned}$$

De façon analogue pour $i = 2, \ldots, n$

$$\begin{aligned}\langle u, e_i\rangle &= \langle k_1e_1 + \cdots + k_ie_i + \cdots + k_ne_n, e_i\rangle \\ &= k_1\langle e_1, e_i\rangle + \cdots + k_i\langle e_i, e_i\rangle + \cdots + k_n\langle e_n, e_i\rangle \\ &= k_1 \cdot 0 + \cdots + k_i \cdot 1 + \cdots + k_n \cdot 0 = k_i\end{aligned}$$

Remplaçons $\langle u, e_i\rangle$ par k_i dans l'équation $u = k_1e_1 + \cdots + k_ne_n$, on obtient le résultat désiré.

(2) Nous avons
$$\left\langle \sum_{i=1}^n a_ie_i, \sum_{j=1}^n b_je_j \right\rangle = \sum_{i,j=1}^n a_i\overline{b_j}\langle e_i, e_j\rangle$$

Mais $\langle e_i, e_j\rangle = 0$ pour $i \neq j$ et $\langle e_i, e_j\rangle = 1$ pour $i = j$; donc comme il est demandé

$$\left\langle \sum_{i=1}^n a_ie_i, \sum_{j=1}^n b_je_j \right\rangle = \sum_{i=1}^n a_i\bar{b}_i = a_1\bar{b}_1 + a_2\bar{b}_2 + \cdots + a_n\bar{b}_n$$

(3) D'après (1) $u = \langle u, e_1\rangle e_1 + \cdots + \langle u, e_n\rangle e_n$ et $v = \langle v, e_1\rangle e_1 + \cdots + \langle v, e_n\rangle e_n$

D'après (2), $\langle u, v\rangle = \langle u, e_1\rangle\overline{\langle v, e_1\rangle} + \langle u, e_2\rangle\overline{\langle v, e_2\rangle} + \cdots + \langle u, e_n\rangle\overline{\langle v, e_n\rangle}$

(4) D'après (1)

$$\begin{aligned} T(e_1) &= \langle T(e_1), e_1\rangle e_1 + \langle T(e_1), e_2\rangle e_2 + \cdots + \langle T(e_1), e_n\rangle e_n \\ T(e_2) &= \langle T(e_2), e_1\rangle e_1 + \langle T(e_2), e_2\rangle e_2 + \cdots + \langle T(e_2), e_n\rangle e_n \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ T(e_n) &= \langle T(e_n), e_1\rangle e_1 + \langle T(e_n), e_2\rangle e_2 + \cdots + \langle T(e_n), e_n\rangle e_n \end{aligned}$$

La matrice A représentant T dans la base $\{e_i\}$ est la transposée de la matrice précédente des coefficients ; donc le ij élément de A est $\langle T(e_j), e_i\rangle$.

ADJOINTS

13.16. Soit T un opérateur linéaire sur $\mathbf{C}^3$ défini par

$$T(x, y, z) = (2x + (1-i)y,\ (3+2i)x - 4iz,\ 2ix + (4-3i)y - 3z)$$

Trouver $T^*(x, y, z)$.

Trouvons d'abord la matrice A représentant T dans la base usuelle de $\mathbf{C}^3$ (voir Problème 7.3):

$$A = \begin{pmatrix} 2 & 1-i & 0 \\ 3+2i & 0 & -4i \\ 2i & 4-3i & -3 \end{pmatrix}$$

Formons la transposée conjuguée A^* de A:

$$A^* = \begin{pmatrix} 2 & 3-2i & -2i \\ 1+i & 0 & 4+3i \\ 0 & 4i & -3 \end{pmatrix}$$

Ainsi $$T^*(x, y, z) = (2x + (3-2i)y - 2iz,\ (1+i)x + (4+3i)z,\ 4iy - 3z)$$

13.17. Démontrer le théorème 13.5 : Soit ϕ une forme linéaire sur un espace préhilbertien de dimension finie V. Alors il existe un vecteur unique $u \in V$ tel que $\phi(v) = \langle v, u\rangle$ pour chaque $v \in V$.

Soit $\{e_1, \ldots, e_n\}$ une base orthonormée de V. Posons

$$u = \overline{\phi(e_1)}e_1 + \overline{\phi(e_2)}e_2 + \cdots + \overline{\phi(e_n)}e_n$$

Soit $\hat{u}$ la forme linéaire de V définie par $\hat{u}(v) = \langle v, u\rangle$ pour chaque $v \in V$. Alors pour $i = 1, \ldots, n$,

$$\hat{u}(e_i) = \langle e_i, u\rangle = \langle e_i, \overline{\phi(e_1)}e_1 + \cdots + \overline{\phi(e_n)}e_n\rangle = \phi(e_i)$$

Puisque $\hat{u}$ et ϕ ont la même action sur chaque vecteur de base $\hat{u} = \phi$.

Supposons maintenant que u' soit un autre vecteur de V pour lequel $\phi(v) = \langle v, u'\rangle$ pour chaque $v \in V$. Alors $\langle v, u\rangle = \langle v, u'\rangle$, ou $\langle v, u - u'\rangle = 0$. En particulier ceci est vrai pour $v = u - u'$ et donc $\langle u - u', u - u'\rangle = 0$. Ceci entraîne $u - u' = 0$ donc $u = u'$. Ainsi un tel vecteur u est unique comme nous l'avions dit.

13.18. Démontrer le théorème 13.6 : Soit T un opérateur linéaire sur un espace préhilbertien de dimension finie V. Alors il existe un opérateur linéaire unique T^* sur V tel que $\langle T(u), v\rangle = \langle u, T^*(v)\rangle$ pour chaque $u, v \in V$. De plus, si A est la matrice représentant T dans une base orthonormée $\{e_i\}$ de V ; alors la transposée conjuguée A^* de A est la matrice représentant T^* dans $\{e_i\}$.

Définissons d'abord l'application T^*. Soit v fixé dans V. L'application $u \mapsto \langle T(u), v\rangle$ est une forme linéaire sur V. Donc d'après le théorème 13.5, il existe un élément unique $v' \in V$ tel que $\langle T(u), v\rangle = \langle u, v'\rangle$ pour chaque $u \in V$. Nous définissons $T^*\ V \to V$ par $T^*(v) = v'$. D'où $\langle T(u), v\rangle = \langle u, T^*(v)\rangle$, pour chaque $u, v \in V$.

Montrons maintenant que T^* est linéaire. Pour tout u, $v_i \in V$ et pour tout a, $b \in K$,

$$\begin{aligned}\langle u, T^*(av_1 + bv_2)\rangle &= \langle T(u), av_1 + bv_2\rangle = \bar{a}\langle T(u), v_1\rangle + \bar{b}\langle T(u), v_2\rangle \\ &= \bar{a}\langle u, T^*(v_1)\rangle + \bar{b}\langle u, T^*(v_2)\rangle = \langle u, aT^*(v_1) + bT^*(v_2)\rangle\end{aligned}$$

Mais ceci est vrai pour chaque $u \in V$; donc $T^*(av_1 + bv_2) = aT^*(v_1) + bT^*(v_2)$. Ainsi T^* est linéaire.

D'après le problème 13.15 (iv) les matrices $A = (a_{ij})$ et $B = (b_{ij})$ représentant T et T^* respectivement dans la base $\{e_i\}$ sont données par $a_{ij} = \langle T(e_j), e_i\rangle$ et $b_{ij} = \langle T^*(e_j), e_i\rangle$. Donc.

$$b_{ij} = \langle T^*(e_j), e_i\rangle = \overline{\langle e_i, T^*(e_j)\rangle} = \overline{\langle T(e_i), e_j\rangle} = \overline{a_{ji}}$$

Ainsi on a bien $B = A^*$.

13.19. Démontrer le théorème 13.7 : Soient S et T des opérateurs linéaires sur un espace préhilbertien V de dimension finie et soit $k \in K$. On a alors

(i) $(S+T)^* = S^* + T^*$ (iii) $(ST)^* = T^*S^*$

(ii) $(kT)^* = \bar{k}T^*$ (iv) $(T^*)^* = T$

(1) Pour tout u, $v \in V$

$$\begin{aligned}\langle (S+T)(u), v\rangle &= \langle S(u) + T(u), v\rangle = \langle S(u), v\rangle + \langle T(u), v\rangle = \langle u, S^*(v)\rangle + \langle u, T^*(v)\rangle \\ &= \langle u, S^*(v) + T^*(v)\rangle = \langle u, (S^* + T^*)(v)\rangle\end{aligned}$$

L'unicité de l'adjoint implique $(S+T)^* = S^* + T^*$.

(2) Pour tout u, $v \in V$

$$\langle (kT)(u), v\rangle = \langle kT(u), v\rangle = k\langle T(u), v\rangle = k\langle u, T^*(v)\rangle = \langle u, \bar{k}T^*(v)\rangle = \langle u, (\bar{k}T^*)(v)\rangle$$

L'unicité de l'adjoint implique $(kT)^* = \bar{k}T^*$

(3) Pour tout u, $v \in V$

$$\langle (ST)(u), v\rangle = \langle S(T(u)), v\rangle = \langle T(u), S^*(v)\rangle = \langle u, T^*(S^*(v))\rangle = \langle u, (T^*S^*)(v)\rangle$$

L'unicité de l'adjoint implique $(ST)^* = T^*S^*$.

(4) Pour tout u, $v \in V$ $\quad \langle T^*(u), v\rangle = \overline{\langle v, T^*(u)\rangle} = \overline{\langle T(v), u\rangle} = \langle u, T(v)\rangle$

L'unicité de l'adjoint implique $(T^*)^* = T$

13.20. Montrer que : (1) $I^* = I$ (2) $= 0^* = 0$ (3) si T est inversible alors $(T^{-1})^* = T^{*-1}$.

(1) Pour chaque u, $v \in V$ $\langle I(u), v\rangle = \langle u, v\rangle = \langle u, I(v)\rangle$; donc $I^* = I$.

(2) Pour chaque u, $v \in V$ $\langle 0(u), v\rangle = \langle 0, v\rangle = 0 = \langle u, 0\rangle = \langle u, 0(v)\rangle$ donc $0^* = 0$.

(3) $I = I^* = (TT^{-1})^* = (T^{-1})^* T^*$; donc $(T^{-1})^* = T^{*-1}$.

13.21. Soit T un opérateur linéaire sur V, et soit W un sous-espace invariant par T de V. Montrer que $W^\perp$ est invariant par T^*.

Soit $u \in W^\perp$. Si $v \in W$, alors $T(w) \in W$ et donc $\langle w, T^*(u)\rangle = \langle T(w), u\rangle = 0$. Ainsi $T^*(u) \in W^\perp$ puisqu'il est orthogonal pour chaque $w \in W$. Donc $W^\perp$ est invariant par T^*.

13.22. Soit T un opérateur linéaire sur V. Montrer que chacune des conditions suivantes implique $T = 0$.

(1) $\langle T(u), v\rangle = 0$ pour chaque u, $v \in V$.

(2) V est un espace complexe et $\langle T(u), u\rangle = 0$ pour chaque $u \in V$.

(3) T est un auto-adjoint et $\langle T(u), u\rangle = 0$ pour chaque $u \in V$.

Donner un exemple d'un opérateur T sur un espace réel V pour lequel $\langle T(u), u\rangle = 0$ pour chaque $u \in V$ mais avec $T \neq 0$.

(1) Posons $v = T(u)$. Alors on a $\langle T(u), T(u)\rangle = 0$ et donc $T(u) = 0$ pour tout $u \in V$. En conséquence $T = 0$.

(2) D'après l'hypothèse $\langle T(v+w), v+w\rangle = 0$ pour tout $v, w \in V$. En posant $\langle T(v), v\rangle = 0$ et $\langle T(w), w\rangle = 0$,

$$\langle T(v), w\rangle + \langle T(w), v\rangle = 0 \qquad (1)$$

Remarquons que w est quelconque dans (1). Remplaçons w par iw et utilisons $\langle T(w), iw\rangle = \bar{i}\langle T(v), w\rangle = -i\langle T(v), w\rangle$ et $\langle T(iw), v\rangle = \langle iT(w), v\rangle = i\langle T(w), v\rangle$,

$$-i\langle T(v), w\rangle + i\langle T(w), v\rangle = 0$$

Divisons par i et ajoutons (1), on obtient $\langle T(w), v\rangle = 0$ pour un couple quelconque $v, w \in V$. D'après (1), $T = 0$.

(3) D'après (2), le résultat est valable pour le cas des nombres complexes ; nous considérerons uniquement le cas réel. En considérant $\langle T(v+w), v+w\rangle = 0$, nous obtenons encore (1). Puisque T est auto-adjoint et puisque l'on est dans un espace réel, $\langle T(w), v\rangle = \langle w, T(v)\rangle = \langle T(v), w\rangle$. En remplaçant dans (1), nous obtenons $\langle T(v), w\rangle = 0$ pour un couple quelconque $v, w \in V$. D'après (1) $T = 0$.

Pour notre exemple, considérons l'opérateur linéaire T de $\mathbf{R}^2$ défini par $T(x, y) = (y, -x)$. Alors $\langle T(u), u\rangle = 0$ pour chaque $u \in V$ mais $T \neq 0$.

OPERATEURS UNITAIRES ET ORTHOGONAUX ET MATRICES

13.23. Démontrer le théorème 13.9 : Les conditions suivantes sur un opérateur U sont équivalentes (1) $U^* = U^{-1}$; (2) $\langle U(v), U(w)\rangle = \langle v, w\rangle$ pour chaque $v, w \in V$; (3) $\|U(v)\| = \|v\|$ pour chaque $v \in V$.

Supposons que (1) soit vrai. Alors pour chaque $v, w \in V$,

$$\langle U(v), U(w)\rangle = \langle v, U^*U(w)\rangle = \langle v, I(w)\rangle = \langle v, w\rangle$$

Ainsi (1) implique (2). Si (2) est maintenant vrai, on a alors

$$\|U(v)\| = \sqrt{\langle U(v), U(v)\rangle} = \sqrt{\langle v, v\rangle} = \|v\|$$

Donc (2) implique (3). Il reste à démontrer que (3) implique (1).

Supposons (3) vrai. Alors pour chaque $v \in V$,

$$\langle U^*U(v), v\rangle = \langle U(v), U(v)\rangle = \langle v, v\rangle = \langle I(v), v\rangle$$

Donc $\langle (U^*U - I)(v), v\rangle = 0$ pour chaque $v \in V$. Mais $U^*U - I$ est auto-adjoint (à démontrer !) ; alors d'après le problème 13.22, nous avons $U^*U - I = 0$ et donc $U^*U = I$. Ainsi $U^* = U^{-1}$ comme il est demandé.

13.24. Soit U un opérateur unitaire (orthogonal) sur V, et soit W un sous-espace invariant par U. Montrer que $W^\perp$ est aussi invariant par U.

Puisque U est non singulier, $U(W) = W$; c'est-à-dire pour un quelconque $w \in W$, il existe $w' \in W$ tel que $U(w') = w$. Posons maintenant $v \in W^\perp$. Alors pour tout $w \in W$,

$$\langle U(v), w\rangle = \langle U(v), U(w')\rangle = \langle v, w'\rangle = 0$$

Ainsi $U(v)$ appartient à $W^\perp$. Donc $W^\perp$ est invariant par U.

13.25. Soit A une matrice ayant pour lignes R_i et pour colonnes C_i. Montrer que (i) le ij-élément de AA^* est $\langle R_i, R_j\rangle$, (ii) le ij-élément de A^*A est $\langle C_j, C_i\rangle$.

Si $A = (a_{ij})$ alors $A^* = (b_{ij})$ où $b_{ij} = \overline{a_{ji}}$. Ainsi $AA^* = (c_{ij})$ où

$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} = \sum_{k=1}^{n} a_{ik}\overline{a_{jk}} = a_{i1}\overline{a_{j1}} + a_{i2}\overline{a_{j2}} + \cdots + a_{in}\overline{a_{jn}}$$
$$= \langle (a_{i1}, \ldots, a_{in}), (a_{j1}, \ldots, a_{jn})\rangle = \langle R_i, R_j\rangle$$

comme il était demandé. On a aussi $A^*A = (d_{ij})$ où

$$d_{ij} = \sum_{k=1}^{n} b_{ik}a_{kj} = \sum_{k=1}^{n} a_{kj}\overline{a_{ki}} = a_{1j}\overline{a_{1i}} + a_{2j}\overline{a_{2i}} + \cdots + a_{nj}\overline{a_{ni}}$$
$$= \langle (a_{1j}, \ldots, a_{nj}), (a_{1i}, \ldots, a_{ni})\rangle = \langle C_j, C_i\rangle$$

13.26. Démontrer le théorème 13.11 : Les conditions suivantes pour une matrice A sont équivalentes. (i) A est unitaire (orthogonale). (ii) Les lignes de A forment un ensemble orthonormé. (iii) Les colonnes de A forment un ensemble orthonormé.

Soient R_i et C_i les lignes et les colonnes de A respectivement. D'après le problème précédent $AA^* = (c_{ij})$ où $c_{ij} = \langle R_i, R_j \rangle$. Ainsi $AA^* = I$ si et seulement si $\langle R_i, R_j \rangle = \delta_{ij}$. C'est-à-dire (i) est équivalent à (ii).

D'après le précédent problème aussi $A^*A = (d_{ij})$ où $d_{ij} = \langle C_j, C_i \rangle$. Ainsi $A^*A = I$ si et seulement si $\langle C_j, C_i \rangle = \delta_{ij}$. C'est-à-dire que (i) est équivalent à (iii).

Remarque : Puisque (ii) et (iii) sont équivalents, A est unitaire (orthogonale) si et seulement si la transposée de A est unitaire (orthogonale).

13.27. Trouver une matrice orthogonale A dont la première ligne est $u_1 = (1/3,\ 2/3,\ 2/3)$.

Il s'agit de trouver d'abord un vecteur non nul $w_2 = (x, y, z)$ qui soit orthogonal à u_1 c'est-à-dire pour lequel

$$0 = \langle u_1, w_2 \rangle = x/3 + 2y/3 + 2z/3 = 0 \quad \text{ou} \quad x + 2y + 2z = 0$$

Une telle solution est $w_2 = (0,\ 1,\ -1)$. Normalisons w_2; on obtient la seconde ligne de A, c'est-à-dire $u_2 = (0,\ 1/\sqrt{2},\ -1/\sqrt{2})$.

Trouvons ensuite un vecteur non nul $w_3 = (x, y, z)$ qui soit orthogonal à la fois à u_1 et u_2 c'est-à-dire pour lequel

$$\begin{aligned} 0 &= \langle u_1, w_3 \rangle = x/3 + 2y/3 + 2z/3 = 0 \quad &\text{ou}\quad x + 2y + 2z = 0 \\ 0 &= \langle u_2, w_3 \rangle = y/\sqrt{2} - z/\sqrt{2} = 0 \quad &\text{ou}\quad y - z = 0 \end{aligned}$$

Posons $z = -1$ et trouvons la solution qui est $w_3 = (4,\ -1,\ -1)$. Normalisons w_3; on obtient ainsi la troisième ligne de A, c'est-à-dire $u_z = (4/\sqrt{18}, -1/\sqrt{18},\ -1/\sqrt{18})$. Ainsi

$$A = \begin{pmatrix} 1/3 & 2/3 & 2/3 \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} \\ 4/3\sqrt{2} & -1/3\sqrt{2} & -1/3\sqrt{2} \end{pmatrix}$$

La matrice A précédente n'est cependant pas unique.

13.28. Démontrer le théorème 13.12 : Soit $\{e_1, \ldots, e_n\}$ une base orthonormée d'un espace préhilbertien V. La matrice de passage de $\{e_i\}$ à une autre base orthonormée est unitaire (orthogonale). Réciproquement, si $P = (a_{ij})$ est une matrice unitaire (orthogonale), alors la base suivante est orthonormée :

$$\{e_i' = a_{1i}e_1 + a_{2i}e_2 + \cdots + a_{ni}e_n : \ i = 1, \ldots, n\}$$

Supposons $\{f_i\}$ une autre base orthonormée et supposons

$$f_i = b_{i1}e_1 + b_{i2}e_2 + \cdots + b_{in}e_n, \quad i = 1, \ldots, n \tag{1}$$

D'après le problème 13.15 et puisque $\{f_i\}$ est orthonormée

$$\delta_{ij} = \langle f_i, f_j \rangle = b_{i1}\overline{b_{j1}} + b_{i2}\overline{b_{j2}} + \cdots + b_{in}\overline{b_{jn}} \tag{2}$$

Soit $B = (b_{ij})$ la matrice des coefficients de (*1*). (B^t est alors la matrice de passage de $\{e_i\}$ à $\{f_i\}$. D'après le problème 13.25, $BB^* = (c_{ij})$ où $c_{ij} = b_{i1}\overline{b_{j1}} + b_{i2}\overline{b_{j2}} + \cdots + b_{in}\overline{b_{jn}}$. D'après (*2*) $c_{ij} = \delta_{ij}$ et de plus $BB^* = I$. En conséquence B et donc B^t sont unitaires.

Il reste à démontrer que $\{e_i'\}$ est orthonormée. D'après le problème 13.15,

$$\langle e_i', e_j' \rangle = a_{1i}\overline{a_{1j}} + a_{2i}\overline{a_{2j}} + \cdots + a_{ni}\overline{a_{nj}} = \langle C_i, C_j \rangle$$

où C_i est la $i$ème colonne de la matrice (unitaire) orthogonale $P = (a_{ij})$. D'après le théorème 13.11, les colonnes de P sont orthonormées, donc $\langle e_i', e_j' \rangle = \langle C_i, C_j \rangle = \delta_{ij}$. Ainsi $\{e_i'\}$ est une base orthonormée.

13.29. Supposons que A soit orthogonale. Montrer que $\det(A) = 1$ ou -1.

Puisque A est orthogonale, $AA^t = I$. En utilisant $|A| = |A^t|$

$$1 = |I| = |AA^t| = |A|\,|A^t| = |A|^2$$

Donc $|A| = 1$ ou -1.

13.30. Montrer que chaque matrice orthogonale 2×2, pour laquelle $\det(A) = 1$, est de la forme $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ pour un certain réel θ donné.

Supposons $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Puisque A est orthogonale, ses lignes forment un ensemble orthonormé ; donc

$$a^2 + b^2 = 1, \quad c^2 + d^2 = 1, \quad ac + bd = 0, \quad ad - bc = 1$$

La dernière équation se déduit de $\det(A) = 1$. Considérons séparément les cas $a = 0$ et $a \neq 0$.

Si $a = 0$, la première équation donne $b^2 = 1$ et donc $b = \pm 1$. La quatrième équation donne $c = -b = \mp 1$ et la seconde équation donne $1 + d^2 = 1$ ou $d = 0$. Ainsi

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \text{ou} \quad \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

La première matrice a la forme demandée avec $\theta = -\pi/2$, et la seconde a la forme demandée avec $\theta = \pi/2$.

Si $a \neq 0$, la troisième équation peut être résolue et donne $c = -bd/a$. Remplaçons c par cette valeur dans la seconde équation,

$$b^2d^2/a^2 + d^2 = 1 \quad \text{ou} \quad b^2d^2 + a^2d^2 = a^2 \quad \text{ou} \quad (b^2 + a^2)d^2 = a^2 \quad \text{ou} \quad a^2 = d^2$$

et donc $a = d$ ou $a = -d$. Si $a = -d$, alors la troisième équation donne $c = b$ et donc la quatrième équation donne $-a^2 - c^2 = 1$ qui est impossible. D'où $a = d$. Mais la troisième équation donne $b = -c$ et donc

$$A = \begin{pmatrix} a & -c \\ c & a \end{pmatrix}$$

Puisque $a^2 + c^2 = 1$, il existe un nombre réel θ tel que $a = \cos\theta$, $c = \sin\theta$, et donc A a la forme demandée dans ce cas aussi.

OPERATEURS SYMETRIQUES ET FORMES CANONIQUES DANS DES ESPACES EUCLIDIENS

13.31. Soit T un opérateur symétrique. Montrer que (1) le polynôme caractéristique $\Delta(t)$ de T est un produit de polynômes linéaires (sur **R**) (2) T a un vecteur propre non nul (2) les vecteurs propres de T appartenant à des valeurs propres distinctes sont orthogonaux.

(1) Soit A une matrice représentant T relativement à une base orthonormée de V ; on a alors $A = A^t$. Soit $\Delta(t)$ le polynôme caractéristique de A. En considérant A comme un opérateur complexe auto-adjoint, A a seulement des valeurs propres réelles d'après le théorème 13-8. Ainsi

$$\Delta(t) = (t - \lambda_1)(t - \lambda_2)\cdots(t - \lambda_n)$$

où les λ_i sont tous réels. En d'autres termes, $\Delta(t)$ est un produit de polynômes linéaires sur **R**.

(2) D'après (1), T a au moins une valeur propre (réelle). Donc T a au moins un vecteur propre non nul.

(3) Supposons $T(v) = \lambda v$ et $T(w) = \mu w$ où $\lambda \neq \mu$. Montrons que $\lambda\langle v, w\rangle = \mu\langle v, w\rangle$:

$$\lambda\langle v, w\rangle = \langle \lambda v, w\rangle = \langle T(v), w\rangle = \langle v, T(w)\rangle = \langle v, \mu w\rangle = \mu\langle v, w\rangle$$

Mais on a $\lambda \neq \mu$; donc $\langle v, w\rangle = 0$, ce que l'on voulait montrer.

13.32. Démontrer le théorème 13.14 : Soit T un opérateur symétrique sur un espace préhilbertien réel V. Il existe alors une base orthonormée de V contenant les vecteurs propres de T ; c'est-à-dire que T peut être représentée par une matrice diagonale relativement à une base orthonormée.

La démonstration se fait par récurrence sur la dimension de V. Si $\dim V = 1$, le théorème est trivial. Supposons maintenant $\dim V = n > 1$. D'après le problème précédent, il existe un vecteur propre non nul v_1 de T. Soit W l'espace engendré par v_1 et soit u_1 un vecteur unitaire de W ; posons $u_1 = v_1/\|v_1\|$.

Puisque v_1 est un vecteur propre de T, le sous-espace W de V est invariant par T. D'après le problème 13.21, $W^\perp$ est invariant par $T^* = T$. Ainsi la restriction $\hat{T}$ de T à $W^\perp$ est un opérateur symétrique. D'après le théorème 13.2 $V = W \oplus W^\perp$. Donc dim $W^\perp = n - 1$ puisque dim $W = 1$. Par l'hypothèse de récurrence, il existe une base orthonormée $\{u_2, \ldots, u_n\}$ de $W^\perp$ contenant des vecteurs propres de $\hat{T}$ et donc T. Mais $\langle u_1, u_i \rangle = 0$ pour $i = 2, \ldots, n$, par $u_i \in W^\perp$. En conséquence $\{u_1, u_2, \ldots, u_n\}$ est un ensemble orthonormé et contient les vecteurs propres de T. Le théorème est ainsi démontré.

13.33. Soit $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$. Trouver une matrice orthogonale (réelle) P pour laquelle P^tAP soit diagonale.

Le polynôme caractéristique $\Delta(t)$ de A est

$$\Delta(t) = |tI - A| = \begin{vmatrix} t-1 & -2 \\ -2 & t-1 \end{vmatrix} = t^2 - 2t - 3 = (t-3)(t+1)$$

et ainsi les valeurs propres de A sont 3 et -1. Remplaçons t par 3 dans la matrice $tI - A$; on obtient le système homogène correspondant d'équations linéaires

$$2x - 2y = 0, \quad -2x + 2y = 0$$

Une solution non nulle est $v_1 = (1, 1)$. Normalisons v_1, on trouve la solution unitaire $u_1 = (1/\sqrt{2}, 1/\sqrt{2})$.

Remplaçons ensuite t par -1 dans la matrice $tI - A$, on obtient le système homogène correspondant d'équations linéaires

$$-2x - 2y = 0, \quad -2x - 2y = 0$$

Une solution non nulle est donnée par $v_2 = (1, -1)$. Normalisons v_2, on trouve alors la solution unitaire $u_2 = (1/\sqrt{2}, -1/\sqrt{2})$.

Finalement soit P la matrice dont les colonnes sont u_1 et u_2 respectivement, alors on a

$$P = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} \quad \text{et} \quad P^tAP = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix}$$

Comme il était prévu, les éléments diagonaux de P^tAP sont les valeurs propres de A.

13.34. Soit $A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$. Trouver une matrice orthogonale réelle P pour laquelle P^tAP soit diagonale.

Trouvons d'abord le polynôme caractéristique $\Delta(t)$ de A :

$$\Delta(t) = |tI - A| = \begin{vmatrix} t-2 & -1 & -1 \\ -1 & t-2 & -1 \\ -1 & -1 & t-2 \end{vmatrix} = (t-1)^2(t-4)$$

Ainsi les valeurs propres de A sont 1 (avec pour multiplicité 2) et 4 (avec pour multiplicité un). Remplaçons t par 1 dans la matrice $tI - A$, on obtient le système homogène correspondant:

$$-x - y - z = 0, \quad -x - y - z = 0, \quad -x - y - z = 0$$

c'est-à-dire, $x + y + z = 0$. Le système a deux solutions indépendantes. Une de ces solutions est $v_1 = (1, -1, 0)$. Cherchons une seconde solution $v_2 = (a, b, c)$ qui soit orthogonale aussi à v_1 ; c'est-à-dire telle que

$$a + b + c = 0 \quad \text{et aussi} \quad a - b = 0$$

Par exemple, $v_2 = (1, 1, -2)$. Normalisons ensuite v_1 et v_2 de manière à obtenir des solutions unitaires et orthogonales

$$u_1 = (1/\sqrt{2}, -1/\sqrt{2}, 0), \quad u_2 = (1/\sqrt{6}, 1/\sqrt{6}, -2/\sqrt{6})$$

Remplaçons maintenant t par 4 dans la matrice $tI - A$, on trouve ainsi le système homogène correspondant

$$2x - y - z = 0, \quad -x + 2y - z = 0, \quad -x - y + 2z = 0$$

On trouve une solution non nulle telle que $v_3 = (1, 1, 1)$ et normalisons v_3 de façon à obtenir une solution unitaire $u_3 = (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})$. Finalement si P est la matrice dont les colonnes sont les u_i respectifs on trouve

$$P = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{3} \\ -1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{3} \\ 0 & -2/\sqrt{6} & 1/\sqrt{3} \end{pmatrix} \quad \text{et} \quad P^tAP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

13.35. Trouver un changement orthogonal de coordonnées qui diagonalise la forme réelle quadratique $q(x, y) = 2x^2 + 2xy + 2y^2$.

Trouvons d'abord la matrice symétrique A représentant q et ensuite son polynôme caractéristique $\Delta(t)$:

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \quad \text{et} \quad \Delta(t) = |tI - A| = \begin{vmatrix} t-2 & -1 \\ -1 & t-2 \end{vmatrix} = (t-1)(t-3)$$

Les valeurs propres de A sont 1 et 3 ; la forme diagonale de q est donc

$$q(x', y') = x'^2 + 3y'^2$$

Trouvons la transformation correspondante des coordonnées, en obtenant un ensemble de valeurs propres de A orthonormées.

Posons $t = 1$ dans la matrice $tI - A$, on obtient le système homogène correspondant

$$-x - y = 0, \quad -x - y = 0$$

Une solution non nulle est donnée par $v_1 = (1, -1)$. Posons maintenant $t = 3$ dans la matrice $tI - A$, on obtient le système homogène correspondant

$$x - y = 0, \quad -x + y = 0$$

Une solution non nulle est donnée par $v_2 = (1, 1)$. Comme dans le Problème 13.31, v_1 et v_2 sont orthogonaux. Normalisons v_1 et v_2 de manière à obtenir une base orthonormée

$$\{u_1 = (1/\sqrt{2}, -1/\sqrt{2}), \; u_2 = (1/\sqrt{2}, 1/\sqrt{2})\}$$

La matrice de passage P et la transformation des coordonnées sont donc

$$P = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} x \\ y \end{pmatrix} = P\begin{pmatrix} x' \\ y' \end{pmatrix} \quad \text{ou} \quad \begin{array}{l} x = (x' + y')/\sqrt{2} \\ y = (-x' + y')/\sqrt{2} \end{array}$$

Remarquons que les colonnes de P sont u_1 et u_2. On peut aussi exprimer x' et y' en fonction de x et y en utilisant la matrice $P^{-1} = P^t$; c'est-à-dire

$$x' = (x - y)/\sqrt{2}, \quad y' = (x + y)/\sqrt{2}$$

13.36. Démontrer le théorème 13.15 : Soit T un opérateur orthogonal sur un espace préhilbertien réel V. Alors il existe une base orthonormée par rapport à laquelle T a la forme suivante :

$$\begin{pmatrix}
\begin{array}{cccc|cccc|cc|c|cc}
1 &&&&&&&&&&&& \\
& 1 &&&&&&&&&&& \\
&& \ddots &&&&&&&&&& \\
&&& 1 &&&&&&&&& \\ \hline
&&&& -1 &&&&&&&& \\
&&&&& -1 &&&&&&& \\
&&&&&& \ddots &&&&&& \\
&&&&&&& -1 &&&&& \\ \hline
&&&&&&&& \cos\theta_1 & -\sin\theta_1 &&& \\
&&&&&&&& \sin\theta_1 & \cos\theta_1 &&& \\ \hline
&&&&&&&&&& \ddots && \\ \hline
&&&&&&&&&&& \cos\theta_r & -\sin\theta_r \\
&&&&&&&&&&& \sin\theta_r & \cos\theta_r
\end{array}
\end{pmatrix}$$

Soit $S = T + T^{-1} = T + T^*$. Donc $S^* = (T + T^*)^* = T^* + T = S$. Ainsi S est un opérateur symétrique sur V. D'après le théorème 13-14, il existe une base orthonormée de V contenant les vecteurs propres de S. Si $\lambda_1, \ldots, \lambda_m$ sont les valeurs propres distinctes de S, alors V peut être décomposé en somme directe $V = V_1 \oplus V_2 \oplus \cdots \oplus V_m$ où les V_i contiennent les vecteurs propres de S correspondant à la valeur propre λ_i. On peut affirmer que chacun des V_i est invariant par T. Supposons en effet que $v \in V_i$, alors $S(v) = \lambda_i v$ et

$$S(T(v)) = (T+T^{-1})T(v) = T(T+T^{-1})(v) = TS(v) = T(\lambda_i v) = \lambda_i T(v)$$

C'est-à-dire, $T(v) \in V_i$. Donc V_i est invariant par T. Puisque les V_i sont orthogonaux les uns par rapport aux autres, nous pouvons restreindre notre étude au cas où T agit sur un seul V_i.

Sur un V_i donné $(T + T^{-1})v = S(v) = \lambda_i v$. Multiplions par T.

$$(T^2 - \lambda_i T + I)(v) = 0$$

Considérons séparément les cas $\lambda_i = \pm 2$ et $\lambda_i \neq \pm 2$. Si $\lambda_i = \pm 2$ alors $(T \pm I)^2(v) = 0$ ce qui donne $(T \pm I)(v) = 0$ ou $T(v) = \pm v$. Ainsi T restreint à ce V_i est soit I soit $-I$.

Si $\lambda_i \neq \pm 2$, alors T n'a pas de vecteurs propres dans V_i puisque d'après le théorème 13.8 les seules valeurs propres de T sont 1 ou -1. En conséquence, pour $v \neq 0$, les vecteurs v et $T(v)$ sont linéairement indépendants. Soit W le sous-espace engendré par v et $T(v)$. Alors W est invariant par T, puisque

$$T(T(v)) = T^2(v) = \lambda_i T(v) - v$$

D'après le théorème 13.2, $V_i = W \oplus W^\perp$. De plus, d'après le problème 13.24, $W^\perp$ est aussi invariant par T. Ainsi nous pouvons décomposer V_i en somme directe de deux sous-espaces W_j de dimension finie, où les W_j sont orthogonaux les uns aux autres et où chaque W_j est invariant par T. Nous pouvons ainsi de nouveau restreindre notre étude au cas où T n'agit que sur un seul W_j.

Puisque $T^2 - \lambda_i T + I = 0$ le polynôme caractéristique $\Delta(t)$ de T agissant sur W_j est $\Delta(t) = t^2 - \lambda_i t + 1$. Le déterminant de T est ainsi 1, qui est le terme constant dans $\Delta(t)$. D'après le problème 13.30, la matrice A représentant T agissant sur W_j relativement à une base orthonormée quelconque de W_j doit être de la forme

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

La réunion des bases des W_j donne une base orthonormée des V_i et l'union des bases des V_i donne une base orthonormée de V dans laquelle la matrice représentant T a la forme demandée.

OPERATEURS NORMAUX ET FORMES CANONIQUES DANS DES ESPACES UNITAIRES

13.37. Déterminer laquelle des matrices suivantes est normale : (i) $A = \begin{pmatrix} 1 & i \\ 0 & 1 \end{pmatrix}$, (ii) $B = \begin{pmatrix} 1 & i \\ 1 & 2+i \end{pmatrix}$.

(i) $$AA^* = \begin{pmatrix} 1 & i \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -i & 1 \end{pmatrix} = \begin{pmatrix} 2 & i \\ -i & 1 \end{pmatrix} \qquad A^*A = \begin{pmatrix} 1 & 0 \\ -i & 1 \end{pmatrix}\begin{pmatrix} 1 & i \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & i \\ -i & 2 \end{pmatrix}$$

Puisque $AA^* \neq A^*A$ la matrice A n'est pas normale.

(ii) $$BB^* = \begin{pmatrix} 1 & i \\ 1 & 2+i \end{pmatrix}\begin{pmatrix} 1 & 1 \\ -i & 2-i \end{pmatrix} = \begin{pmatrix} 2 & 2+2i \\ 2-2i & 6 \end{pmatrix}$$

$$B^*B = \begin{pmatrix} 1 & 1 \\ -i & 2-i \end{pmatrix}\begin{pmatrix} 1 & i \\ 1 & 2+i \end{pmatrix} = \begin{pmatrix} 2 & 2+2i \\ 2-2i & 6 \end{pmatrix}$$

Puisque $BB^* = B^*B$ la matrice B est normale.

13.38. Soit T un opérateur normal. Démontrer que

(i) $T(v) = 0$ si et seulement si $T^*(v) = 0$.

(ii) $T - \lambda I$ est normal.

(iii) Si $T(v) = \lambda v$ alors $T^*(v) = \bar{\lambda} v$; donc tout vecteur propre de T est aussi un vecteur propre de T^*.

(iv) Si $T(v) = \lambda_1 v$ et $T(w) = \lambda_2 w$ où $\lambda_1 \neq \lambda_2$ alors $\langle v, w\rangle = 0$; c'est-à-dire les vecteurs propres de T correspondant aux valeurs propres distinctes de T sont orthogonaux.

(i) Montrons que $\langle T(v), T(v)\rangle = \langle T^*(v), T^*(v)\rangle$:

$$\langle T(v), T(v)\rangle = \langle v, T^*T(v)\rangle = \langle v, TT^*(v)\rangle = \langle T^*(v), T^*(v)\rangle$$

Donc d'après $[I_3]$, $T(v) = 0$ si et seulement si $T^*(v) = 0$.

(ii) Montrons que $T - \lambda I$ commute avec son adjoint :

$$\begin{aligned}(T-\lambda I)(T-\lambda I)^* &= (T-\lambda I)(T^*-\bar\lambda I) = TT^* - \lambda T^* - \bar\lambda T + \lambda\bar\lambda I\\ &= T^*T - \bar\lambda T - \lambda T^* + \bar\lambda\lambda I = (T^*-\bar\lambda I)(T-\lambda I)\\ &= (T-\lambda I)^*(T-\lambda I)\end{aligned}$$

Ainsi $T - \lambda I$ est normal.

(iii) Si $T(v) = \lambda v$, alors $(T - \lambda I)(v) = 0$. Donc $T - \lambda I$ est normal d'après (ii) ; donc d'après (i) $(T - \lambda I)^*(v) = 0$. C'est-à-dire $(T^* - \bar\lambda I)(v) = 0$, donc $T^*(v) = \bar\lambda v$.

(iv) Montrons que $\lambda_1 \langle v, w\rangle = \lambda_2 \langle v, w\rangle$:

$$\lambda_1\langle v, w\rangle = \langle \lambda_1 v, w\rangle = \langle T(v), w\rangle = \langle v, T^*(w)\rangle = \langle v, \bar\lambda_2 w\rangle = \lambda_2\langle v, w\rangle$$

Mais $\lambda_1 \neq \lambda_2$ donc $\langle v, w\rangle = 0$.

13.39. Démontrer le théorème 13.16 : Soit T un opérateur normal sur un espace préhilbertien complexe V de dimension finie. Il existe alors une base orthonormée de V contenant les vecteurs propres de T ; c'est-à-dire que T peut être représentée par une matrice diagonale relativement à une base orthonormée.

La démonstration se fait par récurrence sur la dimension de V. Si $\dim V = 1$, le théorème est alors trivialement vrai. Supposons maintenant que $\dim V = n > 1$. Puisque V est un espace vectoriel complexe, T a au moins une valeur propre et donc un vecteur propre non nul. Soit W le sous-espace de V engendré par v et soit u_1 un vecteur unitaire de W.

Puisque v est un vecteur propre de T, le sous-espace W est invariant par T. Cependant, v est aussi un vecteur propre de T^* d'après le problème précédent ; donc W est aussi invariant par T^*. D'après le problème 13.21, $W^\perp$ est invariant par $T^{**} = T$. La suite de la démonstration est identique à la dernière partie de la démonstration du théorème 13.14. (Problème 13.22).

13.40. Démontrer le théorème 13.17 : Soit T un opérateur arbitraire sur un espace préhilbertien complexe de dimension finie V. Alors T peut être représenté par une matrice triangulaire relativement à une base orthonormé $\{u_1, u_2, \ldots, u_n\}$ c'est-à-dire, pour $i = 1, \ldots, n$,

$$T(u_i) = a_{i1}u_1 + a_{i2}u_2 + \cdots + a_{ii}u_i$$

La démonstration se fait par récurrence sur la dimension de V. Si $\dim V = 1$, le théorème est trivialement vérifié. Supposons maintenant $\dim V = n > 1$. Puisque V est un espace vectoriel complexe, T a au moins une valeur propre et donc au moins un vecteur propre non nul v. Soit W le sous-espace de V engendré par v et soit u_1 un vecteur unitaire de W. Alors u_1 est un vecteur propre de T et donc $T(u_1) = a_{11}u_1$.

D'après le théorème 13.2, $V = W \oplus W^\perp$. Soit E la projection orthogonale de V sur $W^\perp$. Il est clair que $W^\perp$ est invariant par l'opérateur ET. Par récurrence, il existe une base orthonormée $\{u_2, \ldots, u_n\}$ de $W^\perp$ telle que, pour $i = 2, \ldots, n$,

$$ET(u_i) = a_{i2}u_2 + a_{i3}u_3 + \cdots + a_{ii}u_i$$

(Remarquons que $\{u_1, u_2, \ldots, u_n\}$ est une base orthonormée de V). Mais E est la projection orthogonale de V sur $W^\perp$, donc nous devons avoir

$$T(u_i) = a_{i1}u_1 + a_{i2}u_2 + \cdots + a_{ii}u_i$$

pour $i = 2, \ldots, n$. Ceci avec $T(u_1) = a_{11}u_1$ nous donne le résultat demandé.

PROBLEMES DIVERS

13.41. Démontrer le Théorème 13.13A : Les conditions suivantes s'appliquant à un opérateur P sont équivalentes :

(i) $P = T^2$ pour un certain opérateur auto-adjoint T.

(ii) $P = S^*S$ pour un certain opérateur S.

(iii) P est auto-adjoint et $\langle P(u), u\rangle \geqslant 0$ pour chaque $u \in V$.

Supposons que (i) soit vrai, c'est-à-dire $P = T^2$ où $T = T^*$. Alors $P = T \cdot T = T^*T$ et donc (i) implique (ii). Supposons maintenant que (ii) soit vrai. On a alors $P^* = (S^*S)^* = S^*S^{**} = S^*S = P$ et donc P est auto-adjoint. De plus

$$\langle P(u), u\rangle = \langle S^*S(u), u\rangle = \langle S(u), S(u)\rangle \geqslant 0.$$

Donc (ii) implique (iii) et donc il reste à montrer que (iii) implique (i).

Supposons donc maintenant que (iii) soit vrai. Puisque P est auto-adjoint, il existe une base orthonormée $\{u_1, \ldots, u_n\}$ de V contenant les vecteurs propres de P ; c'est-à-dire $P(u_i) = \lambda_i u_i$. D'après le théorème 13.8 les λ_i sont réels. En utilisant (iii), montrons que les λ_i sont non négatifs. Nous avons, pour chaque i,

$$0 \leq \langle P(u_i), u_i\rangle = \langle \lambda_i u_i, u_i\rangle = \lambda_i\langle u_i, u_i\rangle$$

Donc $\langle u_i, u_i\rangle \geqslant 0$ entraîne $\lambda_i \geqslant 0$ comme nous l'avions affirmé. En conséquence, $\sqrt{\lambda_i}$ est un nombre réel. Appelons T l'opérateur linéaire défini par

$$T(u_i) = \sqrt{\lambda_i}\, u_i, \quad \text{pour } i = 1, \ldots, n$$

Puisque T est représenté par une matrice réelle diagonale relativement à la base orthonormée $\{u_i\}$, T est auto-adjoint. De plus, pour chaque i,

$$T^2(u_i) = T(\sqrt{\lambda_i}\, u_i) = \sqrt{\lambda_i}\, T(u_i) = \sqrt{\lambda_i}\sqrt{\lambda_i}\, u_i = \lambda_i u_i = P(u_i)$$

Puisque T^2 et P ont la même représentation sur une base de V, $P = T^2$. Le théorème est ainsi démontré.

Remarque : L'opérateur précédent T est l'unique opérateur positif tel que $P = T^2$ (Problème 13.93) ; on l'appelle la racine carrée positive de P.

13.42. Montrer qu'un opérateur quelconque T est la somme d'un opérateur qui est auto-adjoint et d'un opérateur anti-adjoint.

Posons $S = \frac{1}{2}(T + T^*)$ et $U = \frac{1}{2}(T - T^*)$. Alors $T = S + U$ où

$$S^* = (\tfrac{1}{2}(T + T^*))^* = \tfrac{1}{2}(T^* + T^{**}) = \tfrac{1}{2}(T^* + T) = S$$

et

$$U^* = (\tfrac{1}{2}(T - T^*))^* = \tfrac{1}{2}(T^* - T) = -\tfrac{1}{2}(T - T^*) = -U$$

donc S est auto-adjoint et U est anti-adjoint.

13.43. Démontrer : Soit T un opérateur linéaire arbitraire sur un espace préhilbertien de dimension finie V. Alors T est le produit d'un opérateur unitaire (orthogonal) U et d'un opérateur unique positif P, c'est-à-dire $T = UP$. De plus, si T est inversible, alors U est déterminé de manière unique.

D'après le théorème 13.13, T^*T est un opérateur positif et donc il existe un opérateur positif (unique) P tel que $P^2 = T^*T$ (Problème 13.93). Remarquons que

$$\|P(v)\|^2 = \langle P(v), P(v)\rangle = \langle P^2(v), v\rangle = \langle T^*T(v), v\rangle = \langle T(v), T(v)\rangle = \|T(v)\|^2 \qquad (1)$$

Considérons maintenant séparément les cas où T est inversible et non inversible.

Si T est inversible, posons alors $\hat{U} = PT^{-1}$. Montrons que $\hat{U}$ est unitaire.

$$\hat{U}^* = (PT^{-1})^* = T^{-1*}P^* = (T^*)^{-1}P \quad \text{et} \quad \hat{U}^*\hat{U} = (T^*)^{-1}PPT^{-1} = (T^*)^{-1}T^*TT^{-1} = I$$

Ainsi $\hat{U}$ est unitaire. Posons maintenant $U = \hat{U}^{-1}$. Alors U est aussi unitaire et $T = UP$ comme demandé.

Pour démontrer l'unicité, nous supposons que $T = U_0P_0$, où U_0 est unitaire et P_0 est positif. Alors

$$T^*T = P_0^*U_0^*U_0P_0 = P_0IP_0 = P_0^2$$

Mais la racine carrée positive de T^*T est unique (Problème 13.93), donc $P_0 = P$. (Remarquons que l'inversibilité de T n'est pas utilisée pour démontrer l'unicité de P). Si T est maintenant inversible, alors P est aussi inversible par (*1*). Multiplions $U_0P = UP$ à gauche par P^{-1}; on obtient $U_0 = U$. Ainsi U est aussi unique lorsque T est inversible.

Supposons maintenant que T ne soit pas inversible. Soit W l'image de P, c'est-à-dire $W = \operatorname{Im} P$. Définissons $U_1 : W \to V$ par

$$U_1(w) = T(v) \quad \text{où} \quad P(v) = w. \tag{2}$$

Nous devons démontrer que U_1 est bien défini, c'est-à-dire que $P(v) = P(v')$ implique $T(v) = T(v')$. Ceci découle du fait que $P(v - v') = 0$ est équivalent à $\|P(v - v')\| = 0$ qui entraîne $\|T(v - v')\| = 0$ d'après (*1*). Ainsi U_1 est bien défini. Définissons ensuite $U_2 : W \to V$. Remarquons que d'après (*1*) P et T ont les mêmes noyaux. Donc les images de P et T ont la même dimension, c'est-à-dire $\dim(\operatorname{Im} P) = \dim W = \dim(\operatorname{Im} T)$. En conséquence, $W^\perp$ et $(\operatorname{Im} T)^\perp$ ont aussi la même dimension. Prenons pour U_2 un isomorphisme quelconque entre $W^\perp$ et $(\operatorname{Im} T)^\perp$.

Posons ensuite $U = U_1 \oplus U_2$. (où U est défini comme suit : si $v \in V$ et $v = w + w'$ où $w \in W, w' \in W^\perp$, alors $U(v) = U_1(w) + U_2(w')$. Donc U est linéaire (Problème 13.121) et si $v \in V$ et $P(v) = w$, alors d'après (*2*)

$$T(v) = U_1(w) = U(w) = UP(v)$$

Ainsi $T = UP$ comme demandé.

Il reste à montrer que U est unitaire. Chaque vecteur $x \in V$ peut être écrit sous la forme $x = P(v) + w'$ où $w' \in W^\perp$. Alors on a $U(x) = UP(v) + U_2(w') = T(v) + U_2(w')$ où $\langle T(v), U_2(w')\rangle = 0$ d'après la définition de U_2. On a aussi $\langle T(v), T(v)\rangle = \langle P(v), P(v)\rangle$ d'après (*1*). Ainsi

$$\begin{aligned}\langle U(x), U(x)\rangle &= \langle T(v) + U_2(w'), T(v) + U_2(w')\rangle \\ &= \langle T(v), T(v)\rangle + \langle U_2(w'), U_2(w')\rangle \\ &= \langle P(v), P(v)\rangle + \langle w', w'\rangle = \langle P(v) + w', P(v) + w'\rangle \\ &= \langle x, x\rangle\end{aligned}$$

(Nous utilisons aussi le fait que $\langle P(v), w'\rangle = 0$). Donc U est unitaire et le théorème est démontré.

13.44. Soient $(a_1, a_2, \ldots)$ et $(b_1, b_2, \ldots)$ un couple quelconque de points de l'espace l_2 de l'exemple 13.5. Montrer que la somme $\sum_{i=1}^{\infty} a_i b_i = a_1 b_1 + a_2 b_2 + \cdots$ converge absolument.

D'après le problème 1.16 (inégalité de Cauchy-Schwarz)

$$|a_1 b_1| + \cdots + |a_n b_n| \le \sqrt{\sum_{i=1}^{n} a_i^2}\sqrt{\sum_{i=1}^{n} b_i^2} \le \sqrt{\sum_{i=1}^{\infty} a_i^2}\sqrt{\sum_{i=1}^{\infty} b_i^2}$$

qui est vraie quel que soit n. Ainsi la suite des sommes $S_n = |a_1 b_1| + \cdots + |a_n b_n|$ est bornée et donc converge. Donc la somme infinie converge absolument.

13.45. Soit V l'espace vectoriel des polynômes sur $\mathbf{R}$ muni d'un produit scalaire défini par $\langle f, g\rangle = \int_0^1 f(t)\, g(t)\, dt$. Donner un exemple d'une forme linéaire ϕ sur V pour laquelle le théorème 13.5 n'est pas vrai, c'est-à-dire qu'il n'existe pas un polynôme $h(t)$ pour lequel $\phi(f) = \langle f, h\rangle$ quel que soit $f \in V$.

Soit $\phi : V \to \mathbf{R}$ défini par $\phi(f) = f(0)$, c'est-à-dire ϕ donne à $f(t)$ la valeur 0 et donc transforme $f(t)$ en son terme constant. Supposons qu'un polynôme $h(t)$ existe pour lequel

$$\phi(f) = f(0) = \int_0^1 f(t)\, h(t)\, dt \tag{1}$$

pour chaque polynôme $f(t)$. Remarquons que ϕ transforme le polynôme $tf(t)$ en 0 ; donc d'après (*1*),

$$\int_0^1 tf(t)\, h(t)\, dt = 0 \tag{2}$$

pour chaque polynôme $f(t)$. En particulier (*2*) doit être vérifié pour $f(t) = th(t)$ c'est-à-dire

$$\int_0^1 t^2 h^2(t)\, dt = 0$$

Cette intégrale implique que $h(t)$ est le polynôme nul ; donc $\phi(f) = \langle f, h\rangle = \langle f, 0\rangle = 0$ pour chaque polynôme $f(t)$. Ceci est en contradiction avec le fait que ϕ est non nulle ; donc le polynôme $h(t)$ n'existe pas.

PROBLEMES SUPPLEMENTAIRES

PRODUITS SCALAIRES

13.46. Vérifier que

$$\langle a_1u_1 + a_2u_2,\ b_1v_1 + b_2v_2\rangle = a_1\bar{b}_1\langle u_1, v_1\rangle + a_1\bar{b}_2\langle u_1, v_2\rangle + a_2\bar{b}_1\langle u_2, v_1\rangle + a_2\bar{b}_2\langle u_2, v_2\rangle$$

Plus généralement démontrer que

$$\left\langle \sum_{i=1}^{m} a_iu_i,\ \sum_{j=1}^{n} b_jv_j \right\rangle = \sum_{i,j} a_i\bar{b}_j\langle u_i, v_j\rangle$$

13.47. Soient $u = (x_1, x_2)$ et $v = (y_1, y_2)$ appartenant à $\mathbf{R}^2$.

(i) Vérifier que l'expression suivante est un produit scalaire sur $\mathbf{R}^2$:

$$f(u,v) = x_1y_1 - 2x_1y_2 - 2x_2y_1 + 5x_2y_2$$

(ii) Pour quelles valeurs de k l'expression suivante est-elle un produit scalaire sur $\mathbf{R}^2$?

$$f(u,v) = x_1y_1 - 3x_1y_2 - 3x_2y_1 + kx_2y_2$$

(iii) Pour quelles valeurs de $a, b, c, d \in \mathbf{R}$ l'expression suivante est-elle un produit scalaire sur $\mathbf{R}^2$?

$$f(u,v) = ax_1y_1 + bx_1y_2 + cx_2y_1 + dx_2y_2$$

13.48. Trouver la norme de $v = (1, 2) \in \mathbf{R}^2$ par rapport au (i) produit scalaire usuel,(ii) au produit scalaire défini dans le problème 13.47 (i).

13.49. Soient $u = (z_1, z_2)$ et $v = (w_1, w_2)$ appartenant à $\mathbf{C}^2$.

(i) Vérifier que l'expression suivante est un produit scalaire sur $\mathbf{C}^2$:

$$f(u,v) = z_1\bar{w}_1 + (1+i)z_1\bar{w}_2 + (1-i)z_2\bar{w}_1 + 3z_2\bar{w}_2$$

(ii) Pour quelles valeurs de $a, b, c, d \in \mathbf{C}$ l'expression suivante est-elle un produit scalaire sur $\mathbf{C}^2$?

$$f(u,v) = az_1\bar{w}_1 + bz_1\bar{w}_2 + cz_2\bar{w}_1 + dz_2\bar{w}_2$$

13.50. Trouver la norme de $v = (1 - 2i,\ 2 + 3i) \in \mathbf{C}^2$ en utilisant (i) le produit scalaire usuel, (ii) le produit scalaire du problème 13.49 (i).

13.51. Montrer que la fonction distance $d(u, v) = \| v - u\|$ où $u, v \in V$ satisfait aux axiomes suivantes d'un espace métrique :

$[D_1]$ $d(u, v) \geqslant 0$ et $d(u, v) = 0$ si et seulement si $u = v$.

$[D_2]$ $d(u,v) = d(v,u)$.

$[D_3]$ $d(u,v) \leq d(u,w) + d(w,v)$.

13.52. Vérifier la règle du parallélogramme: $\|u+v\| + \|u-v\| = 2\|u\| + 2\|v\|$.

13.53. Vérifier les formes polaires suivantes pour $\langle u, v\rangle$:

(i) $\langle u,v\rangle = \frac{1}{4}\|u+v\|^2 - \frac{1}{4}\|u-v\|^2$ (cas réel);

(ii) $\langle u,v\rangle = \frac{1}{4}\|u+v\|^2 - \frac{1}{4}\|u-v\|^2 + \frac{i}{4}\|u+iv\|^2 - \frac{i}{4}\|u-iv\|^2$ (cas complexe).

13.54. Soit V l'espace vectoriel des $m \times n$ matrices de $\mathbf{R}$. Montrer que $\langle A, B\rangle = \operatorname{tr}(B^tA)$ définit un produit scalaire dans V.

13.55. Soit V l'espace vectoriel des polynômes sur $\mathbf{R}$. Montrer que $\langle f, g\rangle = \int_0^1 f(t)\, g(t)\, dt$ définit un produit scalaire dans V.

13.56. Trouver la norme de chacun des vecteurs suivants :

(i) $u = (\frac{1}{2}, -\frac{1}{4}, \frac{1}{3}, \frac{1}{6}) \in \mathbf{R}^4$.

(ii) $v = (1 - 2i, 3 + i, 2 - 5i) \in \mathbf{C}^3$.

(iii) $f(t) = t^2 - 2t + 3$ dans l'espace défini dans le problème 13.55.

(iv) $A = \begin{pmatrix} 1 & 2 \\ 3 & -4 \end{pmatrix}$ dans l'espace défini dans le problème 13.54.

13.57. Montrer que : (i) la somme de deux produits scalaires est un produit scalaire ; (ii) un multiple positif d'un produit scalaire est un produit scalaire.

13.58. Soient a, b, $c \in \mathbf{R}$ tels que $at^2 + bt + c \geqslant 0$ pour chaque t. Montrer que $b^2 - 4ac \leqslant 0$. Utiliser ce résultat pour démontrer l'inégalité de Cauchy-Schwarz pour les espaces préhilbertiens réels en faisant une extension de la formule $\|tu + v\|^2 \geqslant 0$.

13.59. Supposons $|\langle u, v\rangle| = \|u\| \, \|v\|$. (C'est-à-dire que l'inégalité de Cauchy-Schwarz se réduit alors à une égalité). Montrer que u et v sont linéairement dépendants.

13.60. Trouver le cosinus de l'angle θ que font u, v si

(i) $u = (1, -3, 2)$, $v = (2, 1, 5)$ dans $\mathbf{R}^3$.

(ii) $u = 2t - 1$, $v = t^2$ dans l'espace défini dans le Problème 13.55.

(iii) $u = \begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix}$, $v = \begin{pmatrix} 0 & -1 \\ 2 & 3 \end{pmatrix}$ dans l'espace du Problème 13.54.

ORTHOGONALITE

13.61. Trouver une base du sous-espace W de $\mathbf{R}^4$ orthogonal à $u_1 = (1, -2, 3, 4)$ et $u_2 = (3, -5, 7, 8)$.

13.62. Trouver une base orthonormé du sous-espace W de $\mathbf{C}^3$ engendré par $u_1 = (1, i, 1)$ et $u_2 = (1 + i, 0, 2)$.

13.63. Soit V l'espace vectoriel des polynômes sur $\mathbf{R}$ de degré $\leqslant 2$ muni du produit scalaire $\langle f, g\rangle = \int_0^1 f(t)\, g(t)\, dt$.

(i) Trouver une base du sous-espace W orthogonal à $h(t) = 2t + 1$.

(ii) Appliquer le procédé d'orthogonalisation de Gram-Schmidt à la base $\{1, t, t^2\}$ pour obtenir une base orthonormée $\{u_1(t), u_2(t), u_3(t)\}$ de V.

13.64. Soit V l'espace vectoriel des matrices 2 x 2 sur $\mathbf{R}$ avec le produit scalaire défini par $\langle A, B\rangle = \operatorname{tr}(B^tA)$.

(i) Montrer que le système suivant est une base orthonormée de V:

$$\left\{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right\}$$

(ii) Trouver une base de l'espace orthogonal (a) des matrices diagonales,(b) des matrices symétriques.

13.65. Soit W un sous-ensemble (non nécessairement un sous-espace) de V. Démontrer que (i) $W^{\perp\perp} \supset L(W)$; (ii) si V est de dimension finie alors $W^{\perp\perp} = L(W)$, (où $L(W)$ est l'espace engendré par W).

13.66. Soit W le sous-espace engendré par un vecteur non nul w de V et soit E la projection orthogonale de V sur W. Démontrer que $E(v) = \dfrac{\langle v, w\rangle}{\|w\|^2}\, w$. On appelle $E(v)$ la projection de v sur w.

13.67. Trouver la projection de v sur w si

(i) $v = (1, -1, 2)$, $w = (0, 1, 1)$ dans $\mathbf{R}^3$.

(ii) $v = (1 - i, 2 + 3i)$, $w = (2 - i, 3)$ dans $\mathbf{C}^2$.

(iii) $v = 2t - 1$, $w = t^2$ dans l'espace défini dans le problème 13.55.

(iv) $v = \begin{pmatrix} 1 & 2 \\ 1 & -3 \end{pmatrix}$, $w = \begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix}$ dans l'espace défini dans le problème 13.54.

13.68. Supposons que $\{u_1, u_2, \ldots, u_r\}$ soit une base d'un sous-espace W de V avec $\dim V = n$. Supposons que $\{v_1, \ldots, v_{n-r}\}$ soit un système de $n - r$ vecteurs indépendants tels que $\langle u_i, v_j\rangle = 0$ pour chaque i et chaque j. Montrer que $\{v_1, \ldots, v_{n-r}\}$ est une base de l'orthogonal de W, $W^{\perp}$.

13.69. Supposons que $\{u_1, \ldots, u_r\}$ soit une base orthonormée d'un sous-espace W de V. Soit $E : V \to V$ une application linéaire définie par

$$E(v) = \langle v, u_1\rangle u_1 + \langle v; u_2\rangle u_2 + \cdots + \langle v, u_r\rangle u_r$$

Montrer que E est la projection orthogonale de V sur W.

13.70. Soit $\{u_1, \ldots, u_r\}$ un sous-ensemble orthonormé de V. Montrer que, pour tout $v \in V$, $\sum_{i=1}^{r} |\langle v, u_i\rangle|^2 \leq ||v||^2$. Cette inégalité est connue sous le nom d'inégalité de Bessel.

13.71. Soit V un espace préhilbertien réel. Montrer que :

(1) $\|u\| = \|v\|$ si et seulement si $\langle u + v, u - v\rangle = 0$;

(2) $\|u + v\|^2 = \|u\|^2 + \|v\|^2$ si et seulement si $\langle u, v\rangle = 0$.

Montrer à l'aide de contre-exemples que les théorèmes précédents ne sont pas vrais pour $\mathbf{C}^2$.

13.72. Soient U et W des sous-espaces d'un espace préhilbertien V de dimension finie. Montrer que (1) $(U + W)^{\perp} = U^{\perp} \cap W^{\perp}$; (2) $(U \cap W)^{\perp} = U^{\perp} + W^{\perp}$.

OPERATEUR ADJOINT

13.73. Soit $T : \mathbf{R}^3 \to \mathbf{R}^3$ défini par $T(x, y, z) = (x + 2y,\ 3x - 4z,\ y)$. Trouver $T^*(x, y, z)$.

13.74. Soit $T : \mathbf{C}^3 \to \mathbf{C}^3$ défini par

$$T(x, y, z) = (ix + (2 + 3i)y,\ 3x + (3 - i)z,\ (2 - 5i)y + iz)$$

Trouver $T^*(x, y, z)$.

13.75. Pour chacune des formes linéaires suivantes ϕ sur V trouver un vecteur $u \in V$ tel que $\phi(v) = \langle v, u\rangle$ pour chaque $v \in V$:

(1) $\phi : \mathbf{R}^3 \to \mathbf{R}$ définie par $\phi(x, y, z) = x + 2y - 3z$.

(2) $\phi : \mathbf{C}^3 \to \mathbf{C}$ définie par $\phi(x, y, z) = ix + (2 + 3i)y + (1 - 2i)z$.

(3) $\phi : V \to \mathbf{R}$ définie par $\phi(f) = f(1)$ où V est l'espace vectoriel du problème 13.63.

13.76. Supposons V de dimension finie. Démontrer que l'image de T^* est l'orthogonal du noyau de T, c'est-à-dire que $\operatorname{Im} T^* = (\operatorname{Ker} T)^{\perp}$. Donc $\operatorname{rang}(T) = \operatorname{rang}(T^*)$.

13.77. Montrer que $T^*T = 0$ implique $T = 0$.

13.78. Soit V l'espace vectoriel des polynômes sur $\mathbf{R}$ muni du produit scalaire défini par $\langle f, g\rangle = \int_0^1 f(t)g(t)\,dt$.

Soit D l'opérateur de dérivation sur V tel que $D(f) = df/dt$. Montrer qu'il n'y a pas d'opérateur D^* sur V tel que $\langle D(f), g\rangle = \langle f, D^*(g)\rangle$ pour chaque couple $f, g \in V$, c'est-à-dire que D n'a pas d'adjoint.

OPERATEURS ET MATRICES UNITAIRES ET ORTHOGONALES

13.79. Trouver une matrice orthogonale dont la première ligne est (1) $(1/\sqrt{5},\ 2/\sqrt{5})$; (2) un multiple de $(1, 1, 1)$.

13.80. Trouver une matrice symétrique orthogonale dont la première ligne est $(1/3,\ 2/3,\ 2/3)$. (Comparer avec le problème 13.27).

13.81. Trouver une matrice unitaire dont la première ligne est (1) un multiple de $(1, 1 - i)$; (2) $\left(\frac{1}{2}, \frac{1}{2}i, \frac{1}{2} - \frac{1}{2}i\right)$.

13.82. Démontrer que : Le produit et l'inverse de matrices orthogonales sont des matrices orthogonales. (Les matrices orthogonales forment un groupe par rapport à la multiplication appelé groupe orthogonal).

13.83. Démontrer que : Le produit et l'inverse de matrices unitaires sont des matrices unitaires. (Les matrices unitaires forment un groupe par rapport à la multiplication appelé groupe unitaire).

13.84. Montrer que si une matrice orthogonale (unitaire) est triangulaire alors elle est diagonale.

13.85. Rappelons que les matrices complexes A et B sont unitairement équivalentes s'il existe une matrice unitaire P telle que $B = P^*AP$. Montrer que cette relation est une relation d'équivalence.

13.86. Rappelons que les matrices réelles A et B sont orthogonalement équivalentes s'il existe une matrice orthogonale P telle que $B = P^tAP$. Montrer que cette relation est une relation d'équivalence.

13.87. Soit W un sous-espace de V. Pour tout $v \in V$ posons $v = w + w'$ où $w \in W$ et $w' \in W^\perp$. (Une telle somme est unique car $V = W \oplus W^\perp$). Posons $T : V \to V$ défini par $T(v) = w - w'$. Montrer que T est un opérateur unitaire auto-adjoint de V.

13.88. Soit V un espace préhilbertien, et supposons $U : V \to V$ (non nécessairement linéaire) soit surjective et conserve les produits scalaires, c'est-à-dire que $\langle U(v), U(w)\rangle = \langle u, w\rangle$ pour chaque couple $v, w \in V$. Démontrer que U est linéaire et donc unitaire.

OPERATEURS POSITIFS ET DEFINIS POSITIFS

13.89. Montrer que la somme de deux opérateurs positifs (définis positifs) est positive (définie positive).

13.90. Soit T un opérateur linéaire sur V et soit $f : V \times V \to K$ définie par $f(u, v) = \langle T(u), v\rangle$. Montrer que f est elle-même un produit scalaire sur V si et seulement si T est définie positive.

13.91. Supposons E une projection orthogonale sur un sous-espace donné W de V. Démontrer que $kI + E$ est positif (défini positif) si $k \geqslant 0$ ($k > 0$).

13.92. Démontrer le théorème 13.13 B, page 288, sur les opérateurs définis positifs. (Le théorème correspondant 13.13A pour les opérateurs positifs est démontré dans le problème 13.41).

13.93. Considérons l'opérateur T défini par $T(u_i) = \sqrt{\lambda_i}\, u_i, i = 1, \ldots, n$ dans la démonstration du théorème 13.13A (Problème 13.41). Montrer que T est positif et que c'est le seul opérateur positif pour lequel $T^2 = P$.

13.94. Supposons P à la fois positif et unitaire. Démontrer que $P = I$.

13.95. Une matrice $n \times n$ (réelle ou complexe) $A = (a_{ij})$ est dite positive, si A considéré comme un opérateur linéaire de K^n est positif. (Une définition analogue pour les matrices définies positives). Démontrer que A est positive (définie positive) si et seulement si $a_{ij} = \overline{a_{ji}}$ et

$$\sum_{i,j=1}^{n} a_{ij}x_i\bar{x}_j \geqq 0 \qquad (>0)$$

pour chaque $(x_1, \ldots, x_n)$ de K^n.

13.96. Déterminer laquelle des matrices suivantes est positive (ou définie positive) :

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \quad \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$
$$\text{(i)} \qquad \text{(ii)} \qquad \text{(iii)} \qquad \text{(iv)} \qquad \text{(v)} \qquad \text{(vi)}$$

13.97. Démontrer qu'une matrice complexe 2×2 $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ est positive si et seulement si (1) $A = A^*$ et (2) a, d, et $ad\text{-}bc$ sont des nombres réels non négatifs.

13.98. Démontrer qu'une matrice diagonale A est positive (définie positive) si et seulement si chaque élément diagonal est un nombre réel non négatif (positif).

OPERATEURS AUTO-ADJOINTS - OPERATEURS SYMETRIQUES

13.99. Pour un opérateur quelconque T, montrer que $T + T^*$ est auto-adjoint et que $T - T^*$ est un anti-adjoint.

13.100. Supposons que T soit auto-adjoint. Montrer que $T^2(v) = 0$ implique $T(v) = 0$. Utiliser ceci pour démontrer que $T^n(v) = 0$ implique aussi $T(v) = 0$ pour $n > 0$.

13.101. Soit V un espace préhilbertien complexe. Supposons que $\langle T(v), v\rangle$ soit réel pour chaque $v \in V$. Montrer que T est auto-adjoint.

13.102. Supposons que S et T soient auto-adjoints. Montrer que ST est auto-adjoint si et seulement si S et T commutent, c'est-à-dire si $ST = TS$.

13.103. Pour chacune des matrices symétriques A suivantes, trouver une matrice orthogonale P pour laquelle P^tAP soit diagonale :

(i) $A = \begin{pmatrix} 1 & 2 \\ 2 & -2 \end{pmatrix}$, (ii) $A = \begin{pmatrix} 5 & 4 \\ 4 & -1 \end{pmatrix}$, (iii) $A = \begin{pmatrix} 7 & 3 \\ 3 & -1 \end{pmatrix}$

13.104. Trouver une transformation orthogonale des coordonnées qui diagonalise chacune des formes quadratiques:

(1) $q(x, y) = 2x^2 - 6xy + 10y^2$, (2) $q(x, y) = x^2 + 8xy - 5y^2$

13.105. Trouver une transformation orthogonale des coordonnées qui diagonalise la forme quadratique $q(x, y, z) = 2xy + 2xz + 2yz$.

OPERATEURS NORMAUX ET MATRICES

13.106. Vérifier que $A = \begin{pmatrix} 2 & i \\ i & 2 \end{pmatrix}$ est normale. Trouver une matrice unitaire P telle que P^*AP soit diagonale et trouver P^*AP.

13.107. Montrer qu'une matrice triangulaire est normale si et seulement si elle est diagonale.

13.108. Démontrer que si T est normale sur V, alors $\|T(v)\| = \|T^*(v)\|$ pour chaque $v \in V$. Démontrer que la réciproque est vraie dans des espaces préhilbertiens complexes.

13.109. Montrer que les opérateurs auto-adjoints, anti-adjoints et unitaires (orthogonaux) sont normaux.

13.110. Supposons que T soit normal. Démontrer que:

(1) T est anto-adjoint si et seulement si ses valeurs propres sont réelles.

(2) T est unitaire si et seulement si ses valeurs propres ont pour valeur absolue 1.

(3) T est positive si et seulement si ses valeurs propres sont des nombres réels non négatifs.

13.111. Montrer que si T est normal, alors T et T^* ont le même noyau et la même image.

13.112. Supposons que S et T soient normaux et commutent. Montrer que $S + T$ et ST sont aussi normaux.

13.113. Supposons T normal et commutant avec S. Montrer que T commute aussi avec S^*.

13.114. Démontrer : Soient S et T des opérateurs normaux sur un espace vectoriel complexe de dimension finie. Il existe alors une base orthonormée de V contenant les vecteurs propres à la fois de S et T (c'est-à-dire que S et T peuvent être diagonalisés simultanément).

PROBLEMES D'ISOMORPHISMES

13.115. Soit $\{e_1, \ldots, e_n\}$ une base orthonormée d'un espace préhilbertien V sur K. Montrer que l'application $v \mapsto [v]_e$ est un isomorphisme d'espace préhilbertien entre V et K^n (où $[v]_e$ indique les coordonnées d'un vecteur v dans la base $\{e_i\}$.).

13.116. Montrer que les espaces préhilbertiens V et W sur K sont isomorphes si et seulement si V et W ont la même dimension.

13.117. Supposons que $\{e_1, \ldots, e_n\}$ et $\{e'_1, \ldots, e'_n\}$ soient des bases orthonormées de V et W respectivement. Soit $T : V \to W$ une application linéaire définie par $T(e_i) = e'_i$ pour chaque i. Montrer que T est un isomorphisme.

13.118. Soit V un espace préhilbertien. Rappelons (page 283) que chaque $u \in V$ détermine une forme linéaire $\hat{u}$ dans l'espace dual V^* par définition $\hat{u}(v) = \langle v, u \rangle$ pour chaque $v \in V$. Montrer que l'application $u \mapsto \hat{u}$ est linéaire et non singulière et donc est un isomorphisme de V dans V^* (sur V^* si la dimension est finie).

13.119. Considérons l'espace préhilbertien V du problème 13.54. Montrer que V est isomorphe à $\mathbf{R}^{mn}$ par l'application

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \mapsto (R_1, R_2, \dots, R_m)$$

où $R_i = (a_{i1}, a_{i2}, \dots, a_{in})$ est la $i^{\text{ème}}$ ligne de A.

PROBLEMES DIVERS

13.120. Montrer qu'il existe une base orthonormée $\{u_1, \dots, u_n\}$ de Y constituée de vecteurs propres de T si et seulement si il existe des projections orthogonales $E_1, \dots, E_r$ et des scalaires $\lambda_1, \dots, \lambda_r$, tels que (1) $T = \lambda_1 E_1 + \dots + \lambda_r E_r$; (2) $E_1 + \dots + E_r = I$; (3) $E_i E_j = 0$ pour $i \neq j$.

13.121. Supposons que $V = U \oplus W$ et supposons $T_1 : U \to V$ et $T_2 : W \to V$ linéaires. Montrer que $T = T_1 \oplus T_2$ est aussi linéaire (T étant définie par : si $v \in V$ et $v = u + w$ où $u \in U$, $w \in W$ alors $T(v) = T_1(u) + T_2(w)$).

13.122. Supposons que U soit un opérateur orthogonal sur $\mathbf{R}^3$ avec déterminant positif. Montrer que U est soit une rotation soit une symétrie par rapport à un plan.

REPONSES AUX PROBLEMES SUPPLEMENTAIRES

13.47. (ii) $k > 9$; (iii) $a > 0,\ d > 0,\ ad - bc > 0$

13.48. (i) $\sqrt{5}$, (ii) $\sqrt{13}$

13.50. (i) $3\sqrt{2}$, (ii) $5\sqrt{2}$

13.56. (i) $\|u\| = \sqrt{65}/12$, (ii) $\|v\| = 2\sqrt{11}$, (iii) $\|f(t)\| = \sqrt{83/15}$, (iv) $\|A\| = \sqrt{30}$

13.60. (i) $\cos\theta = 9/\sqrt{420}$, (ii) $\cos\theta = \sqrt{15}/6$, (iii) $\cos\theta = 2/\sqrt{210}$

13.61. $\{v_1 = (1, 2, 1, 0),\ v_2 = (4, 4, 0, 1)\}$

13.62. $\{v_1 = (1, i, 1)/\sqrt{3},\ v_2 = (2i, 1-3i, 3-i)/\sqrt{24}\}$

13.63. (i) $\{f_1(t) = 7t^2 - 5t,\ f_2(t) = 12t^2 - 5\}$
(ii) $\{u_1(t) = 1,\ u_2(t) = (2t-1)/\sqrt{3},\ u_3(t) = (6t^2 - 6t + 1)/\sqrt{5}\}$

13.64. (ii) (*a*) $\left\{\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\right\}$, (*b*) $\left\{\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\right\}$

13.67. (i) $(0, 1/\sqrt{2}, 1/\sqrt{2})$, (ii) $(26 + 7i, 27 + 24i)/\sqrt{14}$, (iii) $\sqrt{5}\,t^2/6$, (iv) $\begin{pmatrix} 0 & 7/\sqrt{6} \\ -7/\sqrt{6} & -14/\sqrt{6} \end{pmatrix}$

13.73. $T^*(x, y, z) = (x + 3y, 2x + z, -4y)$

13.74. $T^*(x, y, z) = (-ix + 3y, (2-3i)x + (2+5i)z, (3+i)y - iz)$

13.75. Posons $u = \overline{\phi(e_1)}\, e_1 + \cdots + \overline{\phi(e_n)}\, e_n$ où $\{e_i\}$ est une base orthonormée.

(i) $u = (1, 2, -3)$, (ii) $u = (-i, 2-3i, 1+2i)$, (iii) $u = (18t^2 - 8t + 13)/15$

13.79. (i) $\begin{pmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ 2/\sqrt{5} & -1/\sqrt{5} \end{pmatrix}$, (ii) $\begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} \\ 2/\sqrt{6} & -1/\sqrt{6} & -1/\sqrt{6} \end{pmatrix}$

13.80. $\begin{pmatrix} 1/3 & 2/3 & 2/3 \\ 2/3 & -2/3 & 1/3 \\ 2/3 & 1/3 & -2/3 \end{pmatrix}$

13.81. (i) $\begin{pmatrix} 1/\sqrt{3} & (1-i)/\sqrt{3} \\ (1+i)/\sqrt{3} & -1/\sqrt{3} \end{pmatrix}$, (ii) $\begin{pmatrix} \frac{1}{2} & \frac{1}{2}i & \frac{1}{2} - \frac{1}{2}i \\ i/\sqrt{2} & -1/\sqrt{2} & 0 \\ \frac{1}{2} & -\frac{1}{2}i & -\frac{1}{2} + \frac{1}{2}i \end{pmatrix}$

13.96. Seulement (1) et (5) sont positives. De plus (5) est définie positive.

13.103. (i) $P = \begin{pmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{pmatrix}$, (ii) $P = \begin{pmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{pmatrix}$, (iii) $P = \begin{pmatrix} 3/\sqrt{10} & -1/\sqrt{10} \\ -1/\sqrt{10} & 3/\sqrt{10} \end{pmatrix}$

13.104. (i) $x = (3x' - y')/\sqrt{10},\ y = (x' + 3y')/\sqrt{10}$, (ii) $x = (2x' - y')/\sqrt{5},\ y = (x' + 2y')/\sqrt{5}$

13.105. $x = x'/\sqrt{3} + y'/\sqrt{2} + z'/\sqrt{6},\ y = x'/\sqrt{3} - y'/\sqrt{2} + z'/\sqrt{6},\ z = x'/\sqrt{3} - 2z'/\sqrt{6}$

13.106. $P = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$, $P^*AP = \begin{pmatrix} 2+i & 0 \\ 0 & 2-i \end{pmatrix}$

APPENDICE A

Ensembles et relations

ENSEMBLES, ELEMENTS

Une liste ou une collection d'objets bien définie est appelée un ensemble ; les objets formant cet ensemble sont appelés ses éléments. Nous écrivons

$$p \in A \quad \text{(lu : } p \text{ appartient à } A\text{) si } p \text{ est un élément de l'ensemble } A$$

Si chaque élément de A appartient à un ensemble B, c'est-à-dire si $x \in A$ implique $x \in B$, alors A est un sous-ensemble de B; on dit aussi que A est inclus dans B, ce que l'on note

$$A \subset B \quad \text{ou} \quad B \supset A$$

Deux ensembles sont égaux s'ils possèdent tous les deux les mêmes éléments ; c'est-à-dire,

$$A = B \quad \text{si et seulement si} \quad A \subset B \quad \text{et} \quad B \subset A$$

Les négations de $p \in A$, $A \subset B$ et $A = B$ s'écrivent $p \notin A$, $A \not\subset B$ et $A \neq B$ respectivement.

Nous caractériserons un ensemble particulier, soit en écrivant tous ses éléments, soit en donnant une propriété caractéristique des éléments de l'ensemble. Par exemple

$$A = \{1, 3, 5, 7, 9\}$$

ce qui veut dire que A est l'ensemble des nombres 1, 3, 5, 7 et 9 et

$$B = \{x : x \text{ est nombre premier}, x < 15\}$$

ce qui veut dire que B est l'ensemble des nombres premiers inférieurs à 15. Nous utiliserons aussi des symboles spéciaux pour noter des ensembles qui sont couramment employés dans le texte. A moins qu'il en soit spécifié autrement on notera :

$\mathbf{N}$ = l'ensemble des nombres entiers positifs : $1, 2, 3, \ldots$;

$\mathbf{Z}$ = l'ensemble des entiers relatifs : $\ldots, -2, -1, 0, 1, 2, \ldots$;

$\mathbf{Q}$ = l'ensemble des nombres rationnels ;

$\mathbf{R}$ = l'ensemble des nombres réels ;

$\mathbf{C}$ = l'ensemble des nombres complexes.

Nous utiliserons aussi le symbole ϕ qui représente l'ensemble vide, c'est-à-dire l'ensemble ne possédant aucun élément ; cet ensemble est un sous-ensemble de tout autre ensemble.

Fréquemment les éléments d'un ensemble sont eux-mêmes des ensembles. Par exemple chaque droite dans un ensemble de droites est elle-même un ensemble de points. Pour aider à clarifier ces situations, nous utiliserons les mots classe, collection et famille de façon synonyme à ensemble. Les mots sous-classe, sous-collection et sous-famille seront utilisés de manière analogue à sous-ensemble.

Exemple A.1 : Les ensembles A et B précédents peuvent être aussi écrits

$A = \{x \in \mathbf{N} : x \text{ est impair}, x < 10\}$ et $B = \{2, 3, 5, 7, 11, 13\}$

Remarquons que $9 \in A$ mais $9 \notin B$ et $11 \in B$ mais $11 \notin A$; tandis que $3 \in A$ et $3 \in B$ et $6 \notin A$ et $6 \notin B$.

Exemple A.2 : Les ensembles de nombres sont tels que $\mathbf{N} \subset \mathbf{Z} \subset \mathbf{Q} \subset \mathbf{R} \subset \mathbf{C}$.

Exemple A.3 : Soit $C = \{x : x^2 = 4 ; x \text{ est impair}\}$. Alors $C = \phi$, C est l'ensemble vide.

Exemple A.4 : Les éléments de la classe $\{\{2, 3\}, \{2\}, \{5, 6\}\}$ sont les ensembles $\{2, 3\}$, $\{2\}$ et $\{5, 6\}$

Le théorème suivant s'applique.

Théorème A.1 : Soient A, B, C des ensembles quelconques. Alors (i) $A \subset A$; (ii) si $A \subset B$ et $B \subset A$, alors $A = B$; (iii) si $A \subset B$ et $B \subset C$ alors $A \subset C$.

La relation $A \subset B$ n'exclut pas que $A = B$. Cependant si $A \subset B$, mais $A \neq B$ alors on dit que A est un sous-ensemble propre de B (quelques auteurs utilisent le symbole $\subseteq$ pour un sous-ensemble et le symbole $\subset$ seulement pour un sous-ensemble propre).

Lorsque nous parlons d'un ensemble indexé $\{a_i : i \in I\}$ ou simplement $\{a_i\}$, nous indiquons qu'il y a une application ϕ de l'ensemble I dans l'ensemble A et que l'image $\phi(i)$ de $i \in I$ est a_i. L'ensemble I est appelé l'ensemble des indices et les éléments a_i sont dits indexés par I. Un ensemble $\{a_1, a_2, \ldots\}$ indexé par les entiers positifs N est appelé une suite. Une classe indexée d'ensemble $\{A_i : i \in I\}$ ou simplement $\{A_i\}$ a une signification analogue sauf que maintenant l'application Φ associe à chaque $i \in I$ un ensemble A_i et non plus un élément a_i.

OPERATIONS SUR LES ENSEMBLES

Soient A et B deux ensembles arbitraires. L'union de A et B, écrite $A \cup B$, est l'ensemble des éléments appartenant à A ou à B ; et l'intersection de A et B, écrite $A \cap B$, est l'ensemble des éléments appartenant à la fois à A et à B :

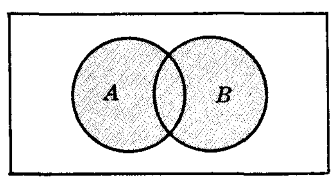
$$A \cup B = \{x : x \in A \text{ ou } x \in B\} \text{ et } A \cap B = \{x : x \in A \text{ et } x \in B\}$$

Si $A \cap B = \emptyset$, c'est-à-dire si A et B n'ont pas d'éléments communs on dit alors que A et B sont disjoints.

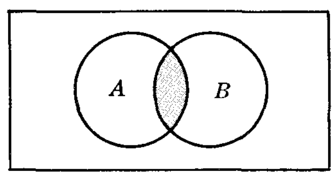
Nous supposons que tous ces ensembles sont eux-mêmes des sous-ensembles d'un ensemble universel (noté ici U) appelé ensemble référentiel. Alors le complémentaire de A, que l'on écrit $\complement A$, est l'ensemble des éléments qui n'appartiennent pas à A:

$$\complement A = \{x \in U : x \notin A\}$$

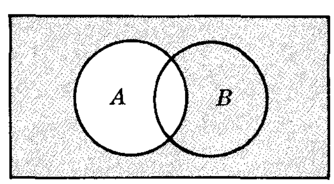
Exemple A.5 : Les diagrammes suivants, appelés diagrammes de Venn, illustrent les opérations précédentes sur les ensembles. Ici les ensembles sont représentés par des aires planes simples et U, l'ensemble universel, par l'aire d'un rectangle.



$A \cup B$ est ombrée



$A \cap B$ est ombrée



$\complement A$ est ombrée

Les ensembles munis des opérations précédentes satisfont à certaines lois ou identités qui sont rassemblées dans le tableau ci-dessous.

Théorème A.2 : Les ensembles satisfont aux lois du tableau 1.

ALGEBRE DES ENSEMBLES	
Lois idempotentes	
1a. $A \cup A = A$	1b. $A \cap A = A$
Lois associatives	
2a. $(A \cup B) \cup C = A \cup (B \cup C)$	2b. $(A \cap B) \cap C = A \cap (B \cap C)$
Lois commutatives	
3a. $A \cup B = B \cup A$	3b. $A \cap B = B \cap A$
Lois distributives	
4a. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$	4b. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
Lois d'identité	
5a. $A \cup \emptyset = A$	5b. $A \cap U = A$
6a. $A \cup U = U$	6b. $A \cap \emptyset = \emptyset$
Lois sur les complémentaires	
7a. $A \cup \complement A$	7b. $A \cap \complement A = \emptyset$
8a. $\complement A = A$	8b. $\complement U = \emptyset$ $\complement \emptyset = U$
Lois de Morgan	
9a. $\complement(A \cup B) = \complement A \cap \complement B$	9b. $\complement(A \cap B) = \complement A \cup \complement B$

Table 1

Remarque : Chacune des lois précédentes découle d'une loi de logique analogue. Par exemple

$A \cap B = \{x : x \in A \text{ et } x \in B\} = \{x : x \in B \text{ et } x \in A\} = B \cap A$

Ici nous utilisons le fait que "p et q" écrit $p \wedge q$ est logiquement équivalent à "q et p" écrit $q \wedge p$.

Il s'ensuit la relation fondamentale entre l'inclusion et les opérations précédentes sur les ensembles.

Théorème A.3 : Chacune des conditions suivantes est équivalente à $A \subset B$.

(i) $A \cap B = A$ (iii) $\complement B \subset \complement A$ (v) $B \cup \complement A = U$

(ii) $A \cup B = B$ (iv) $A \cap \complement B = \emptyset$

On peut généraliser les opérations sur les ensembles précédemment définies comme suit. Soit $\{A_i : i \in I\}$ une famille quelconque d'ensembles. L'union des A_i, que l'on écrit $\cup_{i \in I} A_i$ (ou simplement $\cup_i A_i$) est l'ensemble des éléments appartenant à l'un au moins des A_i ; et l'intersection des A_i, écrite $\cap_{i \in I} A_i$ ou plus simplement $\cap_i A_i$ est l'ensemble des éléments appartenant à chacun des A_i.

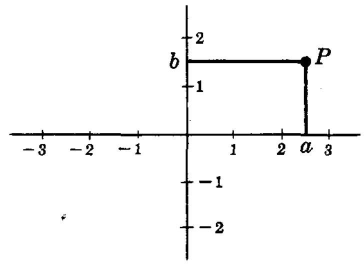
PRODUIT CARTESIEN D'ENSEMBLES

Soient A et B deux ensembles ; leur produit cartésien, noté $A \times B$, est l'ensemble constitué de tous les couples ordonnés (a, b) où $a \in A$ et $b \in B$:

$$A \times B = \{(a, b) : a \in A, \ b \in B\}$$

Le produit cartésien d'un ensemble par lui-même, $A \times A$, est noté A^2.

Exemple A.6 : Le lecteur est familiarisé avec le plan cartésien $\mathbf{R}^2 = \mathbf{R} \times \mathbf{R}$ dessiné ci-dessous. Chaque point P est représenté par un couple ordonné (a, b) de nombres réels et vice versa.



Exemple A.7 : Soient $A = \{1, 2, 3\}$ et $B = \{a, b\}$. Alors
$A \times B = \{(1, a), (1, b), (2, a), (2, b), (3, a), (3, b)\}$

Remarque : Le couple ordonné (a, b) est défini rigoureusement par $(a, b) = \{\{a\}, \{a, b\}\}$. D'après cette définition on peut démontrer la propriété d'ordre, c'est-à-dire que $(a, b) = (c, d)$ si et seulement si $a = c$ et $b = d$.

La définition de produit cartésien d'ensembles peut être étendue à un nombre quelconque fini d'ensembles de manière naturelle. L'ensemble produit cartésien des ensembles $A_1, \ldots, A_m$ que l'on écrit $A_1 \times A_2 \times \ldots \times A_m$ est l'ensemble constitué de tous les m-tuples $(a_1, a_2, \ldots, a_m)$ où $a_i \in A_i$ pour chaque i.

RELATIONS

Une *relation binaire* ou simplement une *relation* R d'un ensemble A dans un ensemble B est la donnée, pour tout couple ordonné $(a, b) \in A \times B$, d'une et d'une seule des deux phrases suivantes :

(1) "a est liée à b" que l'on écrit $a\,R\,b$.

(2) "a n'est pas lié à b" que l'on écrit $a \not R\, b$.

Une relation d'un ensemble A dans lui-même est appelée une relation dans A.

Exemple A.8 : La relation d'inclusion est une relation dans toute classe d'ensembles. Pour un couple donné d'ensembles A et B on a soit $A \subset B$ soit $A \not\subset B$

Remarquons qu'une relation R quelconque de A vers B définit d'une manière unique un sous-ensemble $\hat{R}$ de $A \times B$ comme il suit :

$$\hat{R} = \{(a, b) : a\,R\,b\}$$

Réciproquement, un sous-ensemble quelconque $\hat{R}$ de $A \times B$ définit une relation de A vers B, comme il suit :

$$a\,R\,b \text{ si et seulement si } (a, b) \in \hat{R}$$

En considérant la correspondance précédente entre relations de A vers B et les sous-ensembles de $A \times B$, on peut donner une nouvelle définition d'une relation :

Définition : Une relation R de A vers B est un sous-ensemble de $A \times B$.

RELATIONS D'EQUIVALENCE

Une relation dans un ensemble A est appelée relation d'équivalence, si elle satisfait aux axiomes suivants :

$[E_1]$ Pour tout $a \in A$ a est en relation avec lui-même

$[E_2]$ Si a est en relation avec b, alors b est en relation avec a.

$[E_3]$ Si a est en relation avec b, et si b est en relation avec c, alors a est en relation avec c.

En général une relation est dite réflexive si elle satisfait à $[E_1]$, elle est dite symétrique si elle satisfait à $[E_2]$ et elle est transitive si elle satisfait à $[E_3]$. En d'autres termes, une relation est une relation d'équivalence si elle est réflexive, symétrique et transitive.

Exemple A.9 : Considérons la relation $\subset$ inclusion d'ensemble. D'après le théorème A.1 $A \subset A$ pour tout ensemble A ; et si $A \subset B$ et $B \subset C$, alors $A \subset C$. Donc, $\subset$ est à la fois réflexive et transitive. D'autre part, $\subset$ n'est pas symétrique, puisque $A \subset B$ et $A \neq B$ implique $B \not\subset A$.

Exemple A.10 : En géométrie euclidienne, la similitude des triangles est une relation d'équivalence, car si α, β et γ sont des triangles quelconques, alors (1) α est semblable à lui-même, (2) si α est semblable à β, alors β est semblable à α et (3) si α est semblable à β et β semblable à γ alors α est semblable à γ.

Si R est une relation d'équivalence dans A, la classe d'équivalence d'un élément quelconque $a \in A$, notée par $[a]$, est l'ensemble des éléments qui sont liés à a par la relation R :

$$[a] = \{x : a\,R\,x\}$$

L'ensemble des classes d'équivalence, noté A/R, est appelé ensemble quotient de A par R.

$$A/R = \{[a] \ : \ a \in A\}$$

Voici la propriété fondamentale des relations d'équivalence :

Théorème A.4 : Soit R une relation d'équivalence dans A. L'ensemble quotient A/R est une partition de A, c'est-à-dire chaque $a \in A$ appartient à un élément de A/R et les éléments de A/R sont deux à deux disjoints

Exemple A.11 : Soit R_5 la relation définit dans **Z** ensemble des entiers relatifs par

$$x \equiv y \pmod 5$$

qui se lit "x est congru à y modulo 5" ce qui veut dire "$x - y$ est divisible par 5". Alors R_5 est une relation d'équivalence dans **Z**. Il y a exactement cinq classes d'équivalence distinctes dans $\mathbf{Z}/R_5$:

$$\begin{aligned}
A_0 &= \{\ldots, -10, -5, 0, 5, 10\}\\
A_1 &= \{\ldots, -9, -4, 1, 6, 11\}\\
A_2 &= \{\ldots, -8, -3, 2, 7, 12\}\\
A_3 &= \{\ldots, -7, -2, 3, 8, 13\}\\
A_4 &= \{\ldots, -6, -1, 4, 9, 14\}
\end{aligned}$$

Chaque entier x peut s'écrire de manière unique sous la forme $x = 5q + r$ avec $0 \leqslant r < 5$; remarquons que $x \in E_r$ où r est le reste. Remarquons que les classes d'équivalence sont disjointes deux à deux et que $\mathbf{Z} = A_0 \cup A_1 \cup A_2 \cup A_3 \cup A_4$

APPENDICE B

Structures algébriques

INTRODUCTION

Nous définissons ici les structures algébriques que l'on rencontre dans presque toutes les branches des mathématiques. En particulier, nous définirons la structure de corps qui apparaît dans la définition d'un espace vectoriel. Nous commencerons par donner la définition d'un groupe, qui est une structure algébrique relativement simple puisqu'elle ne contient qu'une seule opération et qu'elle est utilisée dans beaucoup d'autres structures algébriques.

GROUPES

Soit G un ensemble non vide muni d'une opération binaire (aussi appelée : loi de composition interne) c'est-à-dire qui, à chaque paire d'éléments $a, b \in G$, associe un élément $a\,b \in G$. G est alors appelé un groupe si les axiomes suivants sont vérifiés :

[G_1] Quels que soient $a, b, c \in G$, nous avons $(ab)\,c = a(bc)$ (loi associative).

[G_2] Il existe un élément $e \in G$ appelé élément identité ou neutre tel que pour tout $a \in G$ on ait $ae = ea = a$.

[G_3] Pour chaque $a \in G$ il existe un élément $a^{-1} \in G$ appelé l'inverse de a, tel que $a\,.\,a^{-1} = a^{-1}a = e$.

Un groupe G est dit abélien (ou commutatif) si la loi interne précédente est commutative, c'est-à-dire si $ab = ba$ pour tout $a, b \in G$.

Lorsque la loi de composition est notée par une juxtaposition comme précédemment, le groupe G est dit multiplicatif. Quelquefois lorsque G est abélien, l'opération binaire est notée par $+$ et G est dit additif. Dans un tel cas l'élément neutre est noté par 0 et est appelé l'élément nul, et l'inverse est noté par $(-a)$ et est appelé l'opposé de a.

Si A et B sont des sous-ensembles d'un groupe G, on écrit alors

$$AB = \{ab : a \in A\,,\; b \in B\} \quad \text{ou} \quad A + B = \{a + b : a \in A,\; b \in B\}$$

Nous écrirons aussi a pour $\{a\}$.

Un sous-ensemble H d'un groupe G est appelé un sous-groupe de G, si H est lui-même un groupe par rapport à l'opération de G. Si H est un sous-groupe de G et $a \in G$, alors l'ensemble Ha est appelé une classe à droite de H et l'ensemble aH une classe à gauche de A.

Définition : Un sous-groupe H de G est appelé sous-groupe normal (ou : sous-groupe distingué) si $a^{-1}\,Ha \subset H$ pour chaque $a \in G$. De façon équivalente, H est normal si $aH = Ha$ pour chaque $a \in G$, c'est-à-dire si les classes à droite et à gauche de H coïncident.

Remarquons que tout sous-groupe d'un groupe abélien est normal.

Théorème B.1 : Soit H un sous-groupe normal de G. Alors les classes de H dans G forment un groupe par rapport à la multiplication des classes. Ce groupe est appelé groupe quotient et est noté par G/H.

Exemple B.1 : L'ensemble **Z** des entiers relatifs forme un groupe abélien par rapport à l'addition. (Remarquons que les entiers relatifs pairs forment un sous-groupe de **Z**, mais non les entiers relatifs impairs). Soit H l'ensemble des multiples de 5 c'est-à-dire $H = \{\ldots -10, -5, 0, 5, 10\ldots\}$. Alors H est un sous-groupe (nécessairement normal) de **Z**. Donnons les classes H dans **Z** :

$$\begin{aligned}
\bar{0} &= 0 + H = H = \{\ldots, -10, -5, 0, 5, 10, \ldots\} \\
\bar{1} &= 1 + H = \{\ldots, -9, -4, 1, 6, 11, \ldots\} \\
\bar{2} &= 2 + H = \{\ldots, -8, -3, 2, 7, 12, \ldots\} \\
\bar{3} &= 3 + H = \{\ldots, -7, -2, 3, 8, 13, \ldots\} \\
\bar{4} &= 4 + H = \{\ldots, -6, -1, 4, 9, 14, \ldots\}
\end{aligned}$$

Pour tout autre entier relatif $n \in \mathbf{Z}$, $\bar{n} = n + H$ coïncide avec l'une des classes précédentes. Ainsi d'après le théorème précédent $Z/H = \{\bar{0}, \bar{1}, \bar{2}, \bar{3}, \bar{4}\}$ forme un groupe par rapport à l'addition dans les classes, sa table d'addition étant

$+$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$
$\bar{0}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$
$\bar{1}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{0}$
$\bar{2}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{0}$	$\bar{1}$
$\bar{3}$	$\bar{3}$	$\bar{4}$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{4}$	$\bar{4}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$

Ce groupe quotient $\mathbf{Z}/H$ est le même que celui des entiers relatifs modulo 5 et est noté par $\mathbf{Z}_5$. De façon analogue, pour un entier positif quelconque n, il existe un groupe quotient $\mathbf{Z}_n$ appelé groupe des entiers modulo n.

Exemple B.2 : Les permutations de n symboles (voir page 171) forment un groupe par rapport à la loi de composition des applications ; on l'appelle le groupe symétrique de degré n et on le note par S_n. Etudions ici S_3 dont les éléments sont

$$\epsilon = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} \quad \sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \quad \phi_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$$

$$\sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \quad \sigma_3 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \quad \phi_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$$

Ici $\begin{pmatrix} 1 & 2 & 3 \\ i & j & k \end{pmatrix}$ est la permutation qui transforme $1 \mapsto i$, $2 \mapsto j$, $3 \mapsto k$. La table de multiplication de S_3 est

	ϵ	σ_1	σ_2	σ_3	ϕ_1	ϕ_2
ϵ	ϵ	σ_1	σ_2	σ_3	ϕ_1	ϕ_2
σ_1	σ_1	ϵ	ϕ_1	ϕ_2	σ_2	σ_3
σ_2	σ_2	ϕ_2	ϵ	ϕ_1	σ_3	σ_1
σ_3	σ_3	ϕ_1	ϕ_2	ϵ	σ_1	σ_2
ϕ_1	ϕ_1	σ_3	σ_1	σ_2	ϕ_2	ϵ
ϕ_2	ϕ_2	σ_2	σ_3	σ_1	ϵ	ϕ_1

L'élément de la $a^{\text{ième}}$ ligne et la $b^{\text{ième}}$ colonne est ab. L'ensemble $H = \{\epsilon, \sigma_1\}$ est un sous-groupe de S_3, ses classes à droite et à gauche étant

Classes à droite	Classes à gauche
$H = \{\epsilon, \sigma_1\}$	$H = \{\epsilon, \sigma_1\}$
$H\phi_1 = \{\phi_1, \sigma_2\}$	$\phi_1 H = \{\phi_1, \sigma_3\}$
$H\phi_2 = \{\phi_2, \sigma_3\}$	$\phi_2 H = \{\phi_2, \sigma_2\}$

Observons que les classes à droite et à gauche sont distinctes ; donc H n'est pas un sous-groupe normal de S_3.

Une application f d'un groupe G dans un groupe G' est appelée un homomorphisme si $f(ab) = f(a)\,f(b)$ pour tout $a, b \in G$. (Si f est aussi bijective, c'est-à-dire à la fois injective et surjective, alors f est appelée un isomorphisme et G et G' sont isomorphes). Si $f : G \to G'$ est un homomorphisme, alors le noyau de f est l'ensemble des éléments de G qui ont pour image l'élément neutre $e' \in G'$:

$$\text{noyau de } f = \{a \in G : f(a) = e'\}$$

(Comme d'habitude, $f(G)$ est appelé l'image de l'application $f : G \to G'$). Le théorème suivant s'applique alors.

Théorème B.2 : Soit $f : G \to G'$ un homomorphisme ayant pour noyau K. Alors K est un sous-groupe normal de G, et le groupe quotient G/K est isomorphe à l'image de f.

Exemple B.3 : Soit G le groupe des nombres réels muni de l'opération d'addition et soit G' le groupe des nombres réels positifs muni de la multiplication. L'application $f : G \to G'$ définie par $f(a) = 2^a$ est un homomorphisme car

$$f(a + b) = 2^{a+b} = 2^a \,.\, 2^b = f(a)\, f(b)$$

En particulier, f est bijective ; donc G et G' sont isomorphes.

Exemple B.4 : Soit G le groupe des nombres complexes non nuls par rapport à la multiplication et soit G' le groupe des nombres réels non nuls par rapport à la multiplication. L'application $f : G \to G'$ définie par $f(z) = |z|$ est un homomorphisme car

$$f(z_1 z_2) = |z_1 z_2| = |z_1|\,|z_2| = f(z_1)\, f(z_2)$$

Le noyau K de f est constitué des nombres complexes z situés sur le cercle unité, c'est-à-dire pour lesquels $|z| = 1$. Ainsi G/K est isomorphe à l'image de f, c'est-à-dire au groupe des nombres réels positifs par rapport à la multiplication.

ANNEAUX. ANNEAUX INTEGRES ET CORPS

Soit R un ensemble non vide avec deux opérations binaires, une opération d'addition (notée par $+$) et une opération de multiplication (notée par la juxtaposition). R est alors appelé un anneau si les axiomes suivants sont satisfaits:

[R_1] Pour tout $a, b, c \in R$, nous avons $(a + b) + c = a + (b + c)$.

[R_2] Il existe un élément $0 \in R$, appelé élément zéro, tel que pour tout $a \in R, a + 0 = 0 + a = a$.

[R_3] Pour tout $a \in R$, il existe un élément $-a \in R$, appelé l'opposé de a, tel que $a + (-a) = (-a) + a = 0$.

[R_4] Pour tout $a, b \in R$ nous avons $a + b = b + a$.

[R_5] Pour tout $a, b, c \in R$, nous avons $(ab)\,c = a(bc)$.

[R_6] Pour tout $a, b, c \in R$, nous avons

(1) $a(b + c) = ab + ac$ et (2) $(b + c)\,a = ba + ca$.

Remarquons que les axiomes [R_1] jusqu'à [R_4] peuvent être résumés en disant que R est un groupe abélien par rapport à l'addition.

La soustraction est définie dans R par $a - b = a + (-b)$.

On peut montrer (voir Problème B.25) que $a \,.\, 0 = 0 \,.\, a = 0$ pour tout $a \in R$.

R est appelé un anneau commutatif si $ab = ba$ pour tout $a, b \in R$. On dit aussi que R est un anneau unitaire s'il existe un élément non nul $1 \in R$ tel que $a \,.\, 1 = 1 \,.\, a = a$ pour tout $a \in R$.

Un sous-ensemble non vide S de R est appelé un sous-anneau de R, si S est lui-même un anneau pour les opérations définies dans R. Remarquons que S est un sous-anneau de R si et seulement si $a, b \in S$ implique $a - b \in S$ et $a\,b \in S$.

Un sous-ensemble non vide I de R est appelé un idéal à gauche dans R si (1) $a - b \in I$ quels que soient $a, b \in I$, et (2) $ra \in I$ pour tout $r \in R$, $a \in I$. Remarquons qu'un idéal à gauche I dans R est aussi un sous-anneau de R. De façon analogue nous pouvons définir un idéal à droite et un idéal bilatère. Il est clair que tous les idéaux dans un anneau commutatif sont des idéaux bilatères. Le terme idéal signifiera désormais idéal bilatère à moins qu'il en soit spécifié autrement.

Théorème B.3 : Soit I un idéal bilatère d'un anneau R. Les classes $\{a + I : a \in R\}$ forment un anneau par rapport à l'addition et à la multiplication des classes. Cet anneau est noté par R/I et est appelé anneau quotient.

Soit R un anneau commutatif et unitaire. Pour tout $a \in R$, l'ensemble $(a) = \{ra : r \in R\}$ est un idéal et est appelé l'idéal principal engendré par a. Si chaque idéal dans R est un idéal principal alors R est appelé un anneau principal.

Définition : Un anneau commutatif R ayant un élément unité est appelé un domaine d'intégrité si R n'a pas de diviseurs de zéro ; c'est-à-dire si $ab = 0$ implique $a = 0$ ou $b = 0$. (On dit aussi que l'anneau R est intègre).

Supposons maintenant que p divise $f_1, f_2, \ldots, f_n$. Si p divise f_1, alors le lemme est démontré. Si non, d'après le résultat précédent p divise le produit $f_2 \ldots . f_n$, Par récurrence sur n, p divise l'un des polynômes $f_2, \ldots, f_n$. Le lemme est ainsi démontré.

Théorème C.7 : (Théorème de la factorisation) : Soit f un polynôme non nul de $K(t)$. Alors f peut être écrit de façon unique comme un produit

$$f = kp_1 p_2 \ldots p_n$$

où $k \in K$ et où les p_i sont des polynômes normalisés irréductibles de $K(t)$.

Démonstration : Démontrons l'existence d'un tel produit d'abord. Si f est irréductible, ou si $f \in K$, alors un tel produit existe de façon évidente. D'autre part, supposons $f = gh$ où f et g ne sont pas des scalaires. Alors g et h sont de degré inférieur à celui de f. Par récurrence, nous pouvons supposer

$$g = k_1 g_1 g_2 \ldots g_r \quad \text{et} \quad h = k_2 h_1 h_2 \ldots h_s$$

où $k_1, k_2 \in K$ et où les g_i et h_j sont des polynômes normalisés irréductibles. En conséquence

$$f = (k_1 k_2) g_1 g_2 \ldots g_r h_1 h_2 \ldots h_s$$

est la représentation demandée.

Prouvons ensuite l'unicité à l'ordre près d'un tel produit pour f. Supposons

$$f = kp_1 p_2 \ldots p_n = k' q_1 q_2 \ldots q_m$$

où k et $k' \in K$ et où les $p_1, \ldots, p_n, q_1, \ldots, q_m$ sont des polynômes normalisés irréductibles. Donc p_1 divise maintenant $k' q_1 \ldots . q_m$. Puisque p_1 est irréductible il doit donc diviser l'un des q_i, d'après le lemme précédent. Par exemple p_1 divise q_1. Puisque p_1 et q_1 sont tous deux irréductibles et normalisés, $p_1 = q_1$. En conséquence

$$kp_2 \ldots p_n = k' q_2 \ldots q_m$$

Par récurrence nous avons $n = m$ et $p_2 = q_2, \ldots, p_n = q_m$ à un réarrangement près des q_i. Nous avons aussi $k = k'$. Le théorème est donc démontré.

Si le corps K est le corps des complexes C, nous avons alors le résultat suivant qui est connu sous le nom de théorème fondamental de l'algèbre ou théorème de d'Alembert ; sa démonstration dépasse le contenu de cet ouvrage.

Théorème C.8 (Théorème fondamental de l'Algèbre) : Soit $f(t)$ un polynôme non nul sur le corps des complexes C. Alors $f(t)$ peut être écrit de manière unique (à l'ordre près) sous forme d'un produit

$$f(t) = k(t - r_1)(t - r_2) \cdots (t - r_n)$$

où $k, r_i \in \mathbf{C}$, c'est-à-dire comme un produit de polynômes linéaires.

Dans le cas du corps des réels $\mathbf{R}$, nous avons le résultat suivant.

Théorème C.9 : Soit $f(t)$ un polynôme non nul sur le corps des réels $\mathbf{R}$. Alors $f(t)$ peut être écrit de manière unique (à l'ordre près) sous la forme du produit

$$f(t) = kp_1(t)\, p_2(t) \cdots p_m(t)$$

où $k \in \mathbf{R}$ et où les $p_i(t)$ sont des polynômes irréductibles normalisés de degré un ou 2.

où soit $r = 0$ soit $\deg r < \deg g$. Substituons ceci dans (*1*) et résolvons en f ; on obtient

$$f = \left(q_1 + \frac{a_n}{b_m} t^{n-m}\right) g + r$$

qui est la représentation désirée.

Théorème C.3 : L'anneau $K[t]$ des polynômes sur un corps K est un anneau principal. Si I est un idéal de $K[t]$, il existe alors un polynôme normalisé unique d qui engendre I, c'est-à-dire, tel que d divise tout polynôme $f \in I$.

Démonstration. Soit d un polynôme de plus bas degré dans I. Puisque nous pouvons multiplier d par un scalaire non nul, et qu'il appartient toujours à I, nous pouvons supposer sans perte de généralité que d est un polynôme normalisé. Supposons maintenant $f \in I$. D'après le théorème C.2 il existe des polynômes q et r tels que

$$f = qd + r \quad \text{où soit } r = 0, \text{soit } \deg r < \deg d$$

De plus f, $d \in I$ implique $qd \in I$ et donc $r = f - qd \in I$. Mais d est un polynôme de plus bas degré dans I. En conséquence $r = 0$ et $f = qd$, c'est-à-dire d divise f. Il reste à montrer que d est unique. Si d' est un autre polynôme normalisé qui engendre I, alors d divise d' et d' divise d. Ceci implique que $d = d'$ car d et d' sont normalisés. Ainsi le théorème est démontré.

Théorème C.4 : Soient f et g deux polynômes non nuls dans $K(t)$. Il existe alors un polynôme normalisé unique d tel que (1) d divise f et g, (2) si d' divise f et g, alors d' divise d.

Définition : Le précédent polynôme d est appelé le plus grand commun diviseur de f et g. Si $d = 1$. alors f et g sont dits premiers entre eux.

Démonstration du théorème C.4. L'ensemble $I = \{mf + ng : m, n \in K(t)\}$est un idéal. Soit d le polynôme normalisé engendrant I. Comme f et $g \in I$; donc d divise f et g. Supposons maintenant que d' divise f et g. Soit J l'idéal engendré par d'. Alors $f, g \in J$ et donc $I \subset J$. En conséquence, $d \in J$ et donc d' divise d, comme il a été affirmé. Il reste à montrer que d est unique. Si d_1 est un autre plus grand commun diviseur normalisé de f et g, alors d divise d_1 et d_1 divise d, ce qui implique que $d = d_1$ car d et d_1 sont normalisés. Le théorème est démontré.

Corollaire C.5 : Soit d le plus grand commun diviseur des polynômes f et g. Alors il existe des polynômes m et n tels que $d = mf + ng$. En particulier si f et g sont premiers entre eux il existe des polynômes m et n tels que $mf + ng = 1$.

Ce corollaire est issu directement du fait que g engendre l'idéal

$$I = \{mf + ng : m, n \in K(t)\}$$

FACTORISATION

Un polynôme $p \in K(t)$ de degré positif est dit irréductible si $p = fg$ implique que f ou g est un scalaire.

Lemme C.6 : Supposons $p \in K(t)$ irréductible. Si p divise le produit fg de deux polynômes $f, g \in K(t)$, alors p divise f ou p divise g. Plus généralement, si p divise le produit de n polynômes $f_1, f_2, \ldots, f_n$, alors p divise l'un d'entre eux.

Démonstration. Supposons que p divise fg mais non f. Puisque p est irréductible, les polynômes f et p doivent être premiers entre eux. Il existe ainsi des polynômes $m, n \in K(t)$ tels que $mf + np = 1$. En multipliant cette équation par g, on obtient $mfg + npg = g$, mais p divise fg et donc mfg, et p divise npg ; donc p divise la somme $g = mfg + npg$.

NOTATION

Identifions le scalaire $a_0 \in K$ avec le polynôme

$$a_0 = (\ldots, 0, a_0)$$

Choisissons un symbole, par exemple t, pour noter le polynôme

$$t = (\ldots, 0, 1, 0)$$

On appelle le symbole t une variable. Multiplions t par lui-même, on obtient

$$t^2 = (\ldots, 0, 1, 0, 0), \quad t^3 = (\ldots, 0, 1, 0, 0, 0), \quad \ldots$$

Le polynôme précédent f peut alors s'écrire de manière unique sous la forme usuelle

$$f = a_n t^n + \cdots + a_1 t + a_0$$

Lorsque le symbole t est pris comme variable, l'anneau des polynômes sur K est noté par

$$K[t]$$

et on note fréquemment un polynôme f sous la forme $f(t)$.

Nous considérerons aussi le corps K comme un sous-ensemble de $K[t]$ grâce à l'identification précédente. Ceci est possible puisque les opérations d'addition et de multiplication d'éléments de K sont conservées par cette identification :

$$(\ldots, 0, a_0) + (\ldots, 0, b_0) = (\ldots, 0, a_0 + b_0)$$

$$(\ldots, 0, a_0) \cdot (\ldots, 0, b_0) = (\ldots, 0, a_0 b_0)$$

Remarquons que les éléments non nuls de K sont les unités de l'anneau $K[t]$.

Remarquons aussi que chaque polynôme non nul est l'associé d'un polynôme normalisé unique. Donc si d et d' sont des polynômes normalisés tels que d divise d' et d' divise d, alors $d = d'$. (Un polynôme g divise un polynôme f, s'il existe un polynôme h tel que $f = hg$).

DIVISIBILITE

Le théorème suivant formalise le processus appelé "division euclidienne".

Théorème C.2 (Algorithme de la division) : Soient f et g deux polynômes sur un corps K avec $g \neq 0$. Il existe alors des polynômes q et r tels que

$$f = qg + r$$

où soit $r = 0$ soit $\deg r < \deg g$.

Démonstration : Si $f = 0$, ou si $\deg f < \deg g$, nous avons alors l'égalité demandée

$$f = 0g + f$$

Supposons maintenant que $\deg f \geqslant \deg g$, c'est-à-dire

$$f = a_n t^n + \cdots + a_1 t + a_0 \quad \text{et} \quad g = b_m t^m + \cdots + b_1 t + b_0$$

où a_n, $b_m \neq 0$ et $n \geqslant m$. Formons le polynôme

$$f_1 = f - \frac{a_n}{b_m} t^{n-m} g \tag{1}$$

On a $\deg f_1 < \deg f$. Par récurrence, il existe des polynômes q_1 et r tels que

$$f_1 = q_1 g + r$$

APPENDICE C

Polynômes sur un corps

INTRODUCTION

Nous étudierons dans ce chapitre l'ensemble des polynômes sur un corps K et nous montrerons qu'ils ont plusieurs propriétés qui sont analogues aux propriétés des entiers relatifs. Ces résultats jouent un rôle important pour obtenir les formes canoniques d'un opérateur linéaire T sur un espace vectoriel V sur K.

ANNEAU DES POLYNOMES

Soit K un corps. De manière formelle, un polynôme f sur K peut être considéré comme une suite infinie d'éléments de K dans laquelle tous les nombres, sauf un nombre fini d'entre eux, sont nuls :

$$f = (\ldots, 0, a_n, \ldots, a_1, a_0)$$

(Nous écrivons la suite de cette manière pour qu'elle puisse être prolongée à gauche et non pas à droite). L'élément a_k est appelé le k^{ieme} coefficient de f). Si n est le plus grand entier pour lequel $a_n \neq 0$, on dit alors que le degré de f est n, on écrit :

$$\deg f = n$$

a_n est appelé aussi le coefficient initial de f, et si $a_n = 1, f$ est appelé un polynôme normalisé. D'autre part si chaque coefficient de f est nul, alors f est appelé le polynôme nul et on l'écrit $f = 0$. Le degré du polynôme nul n'est pas défini.

Maintenant si g est un autre polynôme sur K, par exemple

$$g = (\ldots, 0, b_m, \ldots, b_1, b_0)$$

alors la somme $f + g$ est le polynôme obtenu en additionnant les coefficients correspondants c'est-à-dire, si $m \leqslant n$ alors

$$f + g = (\ldots, 0, a_n, \ldots, a_m + b_m, \ldots, a_1 + b_1, a_0 + b_0)$$

De plus, le produit fg est le polynôme

$$fg = (\ldots, 0, a_n b_m, \ldots, a_1 b_0 + a_0 b_1, a_0 b_0)$$

c'est-à-dire que le k^{ieme} coefficient c_k de fg est

$$c_k = \sum_{i=0}^{k} a_i b_{k-i} = a_0 b_k + a_1 b_{k-1} + \cdots + a_k b_0$$

On a alors le théorème suivant :

Théorème C.1 : L'ensemble P des polynômes sur un corps K par rapport aux opérations d'addition et de multiplication forme un anneau commutatif unitaire intègre c'est-à-dire un domaine d'intégrité. Si f et g sont des polynômes non nuls de P, alors $\deg(fg) = (\deg f) \cdot (\deg g)$.

B.27. Supposons $a^2 = a$ pour tout $a \in R$. Démontrer que R est un anneau commutatif. (Un tel anneau est appelé anneau de Boole).

B.28. Soit R un anneau unitaire. Transformons R en un anneau $\hat{R}$ défini par $a \oplus b = a + b + 1$ et $a \cdot b = ab + a + b$. (i) Vérifier que $\hat{R}$ est un anneau. (ii) Déterminer l'élément nul et l'élément unité de $\hat{R}$.

B.29. Soit G un groupe abélien additif quelconque. Définissons une multiplication dans G par $a \cdot b = 0$. Montrer que G devient alors un anneau.

B.30. Démontrer le théorème B.3 : Soit I un idéal (bilatère) dans un anneau R. Les classes $\{a + I : a \in R\}$ forment un anneau pour l'addition et la multiplication des classes.

B.31. Soient I_1 et I_2 deux idéaux de R. Démontrer que $I_1 + I_2$ et $I_1 \cap I_2$ sont aussi deux idéaux de R.

B.32. Soient R et R' deux anneaux. Une application $f : R \to R'$ est appelée un homomorphisme (ou un homomorphisme d'anneau) si

$$\text{(i) } f(a + b) = f(a) + f(b) \text{ et (ii) } f(ab) = f(a) \cdot f(b)$$

pour tout $a, b \in R$. Démontrer que si $f : R \to R'$ est un homomorphisme alors l'ensemble $K = \{r \in R : f(r) = 0\}$ est un idéal de R. (L'ensemble K est appelé le noyau de f).

DOMAINES D'INTEGRITE ET CORPS

B.33. Démontrer que dans un anneau intègre, si $ab = ac$, $a \neq 0$ alors $b = c$.

B.34. Démontrer que $F = \{a + b\sqrt{2} : a, b \text{ rationnels}\}$ est un corps.

B.35. Démontrer que $D = \{a + b\sqrt{2} : a \text{ et } b \text{ entiers relatifs}\}$ est un anneau intègre, mais non un corps.

B.36. Démontrer qu'un anneau intègre fini D est un corps.

B.37. Montrer que les seuls idéaux d'un corps K sont $\{0\}$ et K.

B.38. Un nombre complexe $a + bi$ où a et b sont des entiers relatifs est appelé un entier de Gauss. Montrer que l'ensemble G des entiers de Gauss est un anneau intègre. Montrer aussi que les unités de G sont ± 1 et $\pm i$.

B.39. Soit D un anneau intègre et soit I un idéal de D. Démontrer que l'anneau quotient D/I est intègre si et seulement si I est un idéal premier. (Un idéal I est premier si $ab \in I$ implique $a \in I$ ou $b \in I$).

B.40. Considérons l'anneau intègre $D = \{a + b\sqrt{13} : a, b \text{ entiers relatifs}\}$ (voir Exemple B.11). Si $\alpha = a + b\sqrt{13}$, nous définissons $N(\alpha) = a^2 - 13b^2$. Démontrer que (i) $N(\alpha\beta) = N(\alpha)\, N(\beta)$; (ii) α est une unité si et seulement si $N(\alpha) = \pm 1$. (iii) les unités de D sont ± 1 ; $18 \pm 5\sqrt{13}$ et $-18 \pm 5\sqrt{13}$. (iv) les nombres 2, $3 - \sqrt{13}$ et $-3 - \sqrt{13}$ sont irréductibles.

MODULES

B.41. Soit M un R-module et soient A et B deux sous-modules de M. Montrer que $A + B$ et $A \cap B$ sont aussi des sous-modules de M.

B.42. Soit M un R-module avec un sous-module N. Montrer que les classes $\{u + N ; u \in M\}$ forment un R-module pour l'addition des classes et la multiplication scalaire définie par $r(u + N) = ru + N$. (Ce module noté par M/N est appelé le module quotient).

B.43. Soient M et M' deux R-modules et soit $f : M \to M'$ un homomorphisme de R. Montrer que l'ensemble $K = \{u \in M : f(u) = 0\}$ est un sous-module de f. (L'ensemble K est appelé le noyau de f).

B.44. Soit M un R-module et soit $E(M)$ l'ensemble de tous les R-homomorphismes de M dans lui-même. Définissez les opérations appropriées d'addition et de multiplication dans $E(M)$ de sorte que $E(M)$ devienne un anneau.

B.4. Montrer que si G est un groupe abélien, alors $(ab)^n = a^n b^n$ quels que soient a, $b \in G$ et $n \in \mathbf{Z}$.

B.5. Supposons que G soit un groupe tel $(ab)^2 = a^2 b^2$ pour tout a, $b \in G$. Montrer que G est abélien.

B.6. Supposons que H soit un sous-ensemble d'un groupe G. Montrer que H est un sous-groupe de G si et seulement si (i) H n'est pas vide et (ii) a, $b \in H$ implique $ab^{-1} \in H$.

B.7. Démontrer que l'intersection d'un nombre quelconque de sous-groupes de G est aussi un sous-groupe de G.

B.8. Montrer que l'ensemble de toutes les puissances de $a \in G$ est un sous-groupe de G ; on l'appelle le groupe cyclique engendré par a.

B.9. Un groupe G est dit cyclique si G est engendré par l'un de ses éléments $a \in G$; c'est-à-dire $G = \{a^n, n \in \mathbf{Z}\}$. Montrer que chaque sous-groupe d'un groupe cyclique est cyclique.

B.10. Supposons que G soit un sous-groupe cyclique. Montrer que G est isomorphe à l'ensemble $\mathbf{Z}$ (des entiers relatifs pour l'addition ou à l'ensemble $\mathbf{Z}_n$ (des entiers modulo n) pour l'addition.

B.11. Soit H un sous-groupe de G. Montrer que les classes à droite où à gauche de H réalisent une partition de H en ensembles disjoints deux à deux.

B.12. L'ordre d'un groupe G, noté par $|G|$, est le nombre d'éléments de G. Démontrer le théorème de Lagrange : Si H est un sous-groupe d'un groupe fini G, alors $|H|$ divise $|G|$.

B.13. On suppose $|G| = p$ où p est premier. Montrer que G est cyclique.

B.14. On suppose que H et N sont des sous-groupes de G avec N normal. Montrer que (i) HN est un sous-groupe de G et (ii) $H \cap N$ est un sous-groupe normal de G.

B.15. Soit H un sous-groupe de G avec seulement deux classes à droite ou gauche. Montrer que H est un sous-groupe normal de G.

B.16. Démontrer le théorème B.1 : Soit H un sous-groupe normal de G. Les classes de H dans G forment un groupe G/H pour la multiplication des classes.

B.17. Supposons que G soit un groupe abélien. Montrer que tout groupe G/H est aussi abélien.

B.18. Soit $f : G \to G'$ un homomorphisme de groupe. Montrer que :

(i) $f(e) = e'$ où e et e' sont les éléments neutres de G et G' respectivement ;

(ii) $f(a^{-1}) = f(a)^{-1}$ pour tout $a \in G$.

B.19. Démontrer le théorème B.2 : Soit $f : G \to G'$ un homomorphisme de groupe admettant pour noyau K. K est alors un sous-groupe normal de G, et le groupe quotient G/K est isomorphe à l'image de f.

B.20. Soit G le groupe multiplicatif des nombres complexes z tels que $|z| = 1$, et soit $\mathbf{R}$ le groupe additif des nombres réels. Montrer que G est isomorphe à $\mathbf{R}/\mathbf{Z}$.

B.21. Pour un élément fixé $g \in G$; soit $\hat{g} : G \to G$ définie par $\hat{g}(a) = g^{-1}ag$. Montrer que G est un isomorphisme de G sur G.

B.22. Soit G le groupe multiplicatif des $n \times n$ matrices non singulières sur $\mathbf{R}$. Montrer que l'application $A \mapsto |A|$ est un homomorphisme de G dans le groupe multiplicatif des nombres réels non nuls.

B.23. Soit G un groupe abélien. Pour un n fixé $n \in \mathbf{Z}$, montrer que l'application $a \mapsto a^n$ est un homomorphisme de G dans G.

B.24. Supposons que H et N soient des sous-groupes de G avec N normal. Montrer que $H \cap N$ est normal dans H et $H/(H \cap N)$ est isomorphe à HN/N.

ANNEAUX

B.25. Montrer que dans un anneau R :

(i) $a \cdot 0 = 0 \cdot a = 0$, (ii) $a(-b) = (-a)b = -ab$, (iii) $(-a)(-b) = ab$.

B.26. Montrer que dans un anneau avec élément unité on a (i) $(-1)a = -a$. (ii) $(-1)(-1) = 1$.

$[M_1]$ $r(m_1+m_2) = rm_1+rm_2$

$[M_2]$ $(r+s)m = rm+sm$

$[M_3]$ $(rs)m = r(sm)$

$[M_4]$ $1\cdot m = m$

pour r, $s \in R$ et $m_i \in M$.

La notion de R-module est une généralisation de celle d'espace vectoriel lorsque les scalaires appartiennent à un anneau au lieu d'appartenir à un corps.

Exemple B.12 : Soit G un groupe abélien additif quelconque. On peut toujours considérer G comme un module sur l'anneau Z des entiers relatifs en posant

$$ng = \overbrace{g+g+\cdots+g}^{n \text{ termes}}, \qquad 0g = 0, \qquad (-n)g = -ng$$

Exemple B.13 : Soit R un anneau et soit I un idéal de R. Alors I peut être considéré comme un module sur R.

Exemple B.14 : Soit V un espace vectoriel sur un corps K et soit $T : V \to V$ une application linéaire. Transformons V en un module sur l'anneau $K(x)$ des polynômes sur K en définissant $f(x)v = f(T)v$. Le lecteur vérifiera qu'une multiplication scalaire a bien été définie.

Soit M un module sur R. Un sous-groupe additif N de M est appelé un sous-module de M si $u \in N$ et $k \in R$ impliquent $ku \in N$. (Remarquons que N est alors un module sur R).

Soient M et M' deux R-modules. Une application $T : M \to M'$ est appelée un homomorphisme (ou R-homomorphisme) si

$$\text{(i)}\ T(u+v) = T(u)+T(v) \quad \text{et} \quad \text{(ii)}\ T(ku) = kT(u)$$

pour $u, v \in M$ et $k \in R$.

PROBLEMES

GROUPES

B.1. Déterminer lequel des ensembles suivants a une structure de groupe G :

(i) G = ensemble des entiers, muni de l'opération soustraction.

(ii) $G = \{1, -1\}$, muni de l'opération multiplication.

(iii) G = ensemble des nombres rationnels non nuls, muni de l'opération division.

(iv) G = ensemble des matrices non singulières $n \times n$, muni de la multiplication matricielle.

(v) $G = \{a+bi : a, b \in \mathbf{Z}\}$ muni de l'opération addition.

B.2. Montrer que dans un groupe G :

(i) l'élément neutre de G est unique.

(ii) chaque $a \in G$ a un inverse unique $a^{-1} \in G$.

(iii) $(a^{-1})^{-1} = a$ et $(ab)^{-1} = b^{-1}a^{-1}$.

(iv) $ab = ac$ implique $b = c$, et $ba = ca$ implique $b = c$.

B.3. Dans un groupe G, les puissances de a sont définies par

$$a^0 = e,\ a^n = aa^{n-1},\ a^{-n} = (a^n)^{-1} \text{ où } n \in N$$

Montrer que les formules suivantes sont vraies pour des entiers quelconques $r, s, t \in \mathbf{Z}$: (i) $a^ra^s = a^{r+s}$, (ii) $(a^r)^s = a^{rs}$, (iii) $(a^{r+s})^t = a^{rt+st}$.

Définition : Un anneau commutatif R ayant un élément unité est appelé un corps si tout $a \in R$ non nul a un inverse par rapport à la multiplication c'est-à-dire qu'il existe un élément $a^{-1} \in R$ tel que $aa^{-1} = a^{-1}a = 1$.

Un corps est nécessairement intègre car si $ab = 0$ et $a \neq 0$ alors

$$b = 1 \cdot b = a^{-1}ab = a^{-1} \cdot 0 = 0$$

Remarquons qu'un corps peut être aussi considéré comme un anneau commutatif dans lequel les éléments non nuls forment un groupe par rapport à la multiplication.

Exemple B.5 : L'ensemble **Z** des entiers relatifs muni de l'addition et de la multiplication habituelles est un exemple classique d'anneau intègre avec unité. Chaque idéal I dans **Z** est un idéal principal c'est-à-dire $I = (n)$ pour un entier n. L'anneau quotient $\mathbf{Z}_n = \mathbf{Z}/(n)$ est appelé l'anneau des entiers modulo n. Si n est premier, alors Z_n est un corps. D'autre part, si n n'est pas premier alors $\mathbf{Z}_n$ admet des diviseurs de zéro. Par exemple dans l'anneau $\mathbf{Z}_6$, $\overline{2}.\overline{3} = \overline{0}$ et $\overline{2} \neq \overline{0}$ et $\overline{3} \neq \overline{0}$.

Exemple B.6 : Les nombres rationnels **Q** et les nombres réels **R** forment chacun un corps pour les opérations habituelles d'addition et de multiplication.

Exemple B.7 : Soit **C** l'ensemble des couples de nombres réels muni de l'addition et de la multiplication définies par

$$(a, b) + (c, d) = (a + c, b + d)$$

$$(a, b) \cdot (c, d) = (ac - bd, ad + bc)$$

Alors **C** satisfait à toutes les propriétés d'un corps. En fait **C** est le corps des nombres complexes (voir page 4).

Exemple B.8 : L'ensemble M de toutes les matrices 2×2 à éléments réels forme un anneau non commutatif admettant des diviseurs de zéro, pour les deux opérations addition et multiplication des matrices.

Exemple B.9 : Soit **R** un anneau quelconque. Alors l'ensemble $R[x]$ de tous les polynômes sur R est un anneau pour l'addition et la multiplication des polynômes. De plus, si R est intègre, alors $R[x]$ est aussi intègre.

Soit D un anneau intègre. Par définition on dira que b divise a (a et $b \in D$), s'il existe c dans D avec $a = bc$. Un élément u de D est appelée une "unité" de D si u divise 1, c'est-à-dire si u admet un inverse pour la multiplication dans D. Deux éléments a et b de D sont dit "associés" s'il existe une unité u de D avec $a = bu$ (on a aussi $b = a(u^{-1})$. Un élément p qui n'est pas une unité est dit "irréductible" si $p = ab$ implique que l'un des éléments a ou b est une unité.

Un anneau intègre est appelé un anneau factoriel (ou anneau à factorisation unique) si tout élément a qui n'est pas une unité admet une représentation unique (à l'ordre près, et aux éléments associés près) en un produit d'éléments irréductibles.

Exemple B.10 : L'anneau **Z** des entiers est l'exemple classique d'un anneau factoriel. Les unités de **Z** sont 1 et -1. Les seuls associés de n sont n et $-n$. Les éléments irréductibles sont les nombres premiers.

Exemple B.11 : L'ensemble $D = \{a + b\sqrt{13}\ ;\ a, b \in \mathbf{Z}\}$ est un anneau intègre. Ses unités sont ± 1, $18 \pm 5\sqrt{13}$, $-18 \pm 5\sqrt{13}$. Les éléments 2, $3 - \sqrt{13}$ et $-3 - \sqrt{13}$ sont irréductibles dans D. Mais $4 = 2 \cdot 2 = (3 - \sqrt{13})(-3 - \sqrt{13})$. Donc D n'est pas factoriel (voir Problème B.40).

MODULES

Soit M un ensemble non vide et soit R un anneau avec élément unité, Alors M est appelé un R-module (à gauche) si M est un groupe abélien et il existe une application $R \times M \to M$ qui satisfait aux axiomes suivants :

Index

S

T

V

Z

IMPRIMERIE LOUIS-JEAN
Publications scientifiques et littéraires
TYPO - OFFSET
05002 GAP - Téléphone 51-35-23 +
Dépôt légal 118-1977