ENIT -

Analyse Numérique   Corrigé Série 4

EXERCICE 1

Soient : $A \in \mathcal{M}_n(\mathbb{R})$ une matrice symétrique définie positive.

- $b \in \mathbb{R}^n$
- (S) $\quad Ax = b$
- $0 < \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$, les valeurs propres de A

<u>Algorithme du Gradient à pas constant :</u>

$$(1) \quad \begin{cases} x^{(0)} \text{ donné dans } \mathbb{R}^n \\ x^{(k+1)} = x^{(k)} - \alpha(Ax^{(k)} - b) \end{cases}$$

Si $J(x) = \frac{1}{2}(Ax,x) - (b,x)$ alors $\nabla J(x) = Ax - b$

Rappelons que : $\quad A\tilde{x} = b \iff J(\tilde{x}) \leq J(x) \quad \forall x \in \mathbb{R}^n$

1- Montrons que l'algorithme du gradient à pas constant α, pour la résolution du système (S), converge (pour tout choix de $x^{(0)}$) si et seulement si $0 < \alpha < \frac{2}{\lambda_n}$.

On a : $\quad x^{(k+1)} = (I - \alpha A)x^{(k)} + \alpha b \quad$ (méthode itérative dont la matrice est $I - \alpha A$)

la suite $(x^{(k)})_k$ est convergente $\iff \rho(I - \alpha A) < 1 \quad$ (cf cours méthodes itératives)

or $\rho(I - \alpha A) = \max_{1 \leq i \leq n} |1 - \alpha\lambda_i|$

$$\rho(I - \alpha A) < 1 \iff |1 - \alpha\lambda_i| < 1 \quad \forall i \iff 0 < \alpha < \frac{2}{\lambda_i} \quad \forall i$$

$$\underset{\lambda_1 \leq \dots \leq \lambda_n}{\iff} 0 < \alpha < \frac{2}{\lambda_n} \quad \forall i$$

Ainsi l'algorithme du gradient à pas constant α converge ssi $0 < \alpha < \frac{2}{\lambda_n}$

2- En déduire que si il $\exists M > 0$ tel que $\|Ax\|_2 \leq M\|x\|_2$, $\forall x \in \mathbb{R}^n$, alors l'algorithme (1) converge ($\forall x^{(0)}$) si $\alpha \in]0, \frac{2}{M}[$.

Supposons donc que $\alpha \in]0, \frac{2}{M}[$ et montrons que l'algorithme (1) converge.

(Suite exercice 1)

On a, par hypothèse, $\|Ax\|_2 \leq M \|x\|_2 \quad \forall x \in \mathbb{R}^n$, et par suite

$$\|A\|_2 = \sup_{\substack{x\in\mathbb{R}^n \\ x\neq 0}} \frac{\|Ax\|_2}{\|x\|_2} \leq M$$

or A étant symétrique, on a $\|A\|_2 = \rho(A) = \lambda_n$. D'où $\rho(A) = \lambda_n \leq M$.

Donc, $0 < \alpha < \frac{2}{M} \Longrightarrow 0 < \alpha < \frac{2}{M} \leq \frac{2}{\lambda_n}$. Ainsi $0 < \alpha < \frac{2}{\lambda_n}$, ce qui prouve, grâce à 1-), que l'algorithme du gradient à pas constant converge.

3- Déterminons la valeur optimale α^* de α qui assure la convergence la plus rapide de l'algorithme (1).

On a $\quad x^{(k+1)} = (I - \alpha A)x^{(k)} + \alpha b$

Soit $\bar{x}$ la solution du système $Ax = b$, alors $\bar{x}$ vérifie

$$\bar{x} = (I - \alpha A)\bar{x} + \alpha b$$

D'où $\quad x^{(k+1)} - \bar{x} = (I - \alpha A)(x^{(k)} - \bar{x})$. Et par suite

$$x^{(k+1)} - \bar{x} = (I - \alpha A)^{k+1}(x^{(0)} - \bar{x})$$

D'où $$\|x^{(k+1)} - \bar{x}\|_2 \leq \|I - \alpha A\|_2^{k+1} \|x^{(0)} - \bar{x}\|_2 \qquad \forall k$$

La convergence la plus rapide de $x^{(k)}$ vers $\bar{x}$ est obtenue lorsque $\|I - \alpha A\|_2^{k+1}$ tend vers 0 le plus rapidement. (la convergence est assurée lorsque $\|I - \alpha A\|_2 < 1$). Ainsi le paramètre α^* qui assure la convergence la plus rapide de (1), vérifie

(2) $$\|I - \alpha^* A\|_2 \leq \|I - \alpha A\|_2 \qquad \forall \alpha$$

Or $I - \alpha^* A$ et $I - \alpha A$ étant deux matrices symétriques, alors l'inégalité (2) est équivalente à

(3) $$\rho(I - \alpha^* A) \leq \rho(I - \alpha A) \qquad \forall \alpha$$

Soit $f(\alpha) = \rho(I - \alpha A)$. Alors α^* est tel que $f(\alpha^*) = \min_{\alpha \geq 0} f(\alpha)$

Or $$f(\alpha) = \rho(I - \alpha A) = \max_{1 \leq i \leq n} |1 - \alpha\lambda_i|$$
$$= \max\{|1 - \alpha\lambda_1|, |1 - \alpha\lambda_n|\}$$

puisque $\quad 1 - \alpha\lambda_n \leq 1 - \alpha\lambda_{n-1} \leq \dots \leq 1 - \alpha\lambda_1, \quad$ car $\alpha \geq 0$ et $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$

(suite exercice 1)

Ainsi α^* est tel que

$$\rho(I-\alpha^* A) = f(\alpha^*) = \min_{\alpha \geq 0} \left\{ \max\left(|1-\alpha\lambda_1|, |1-\alpha\lambda_n| \right) \right\}$$

soit D_1 la droite d'équation

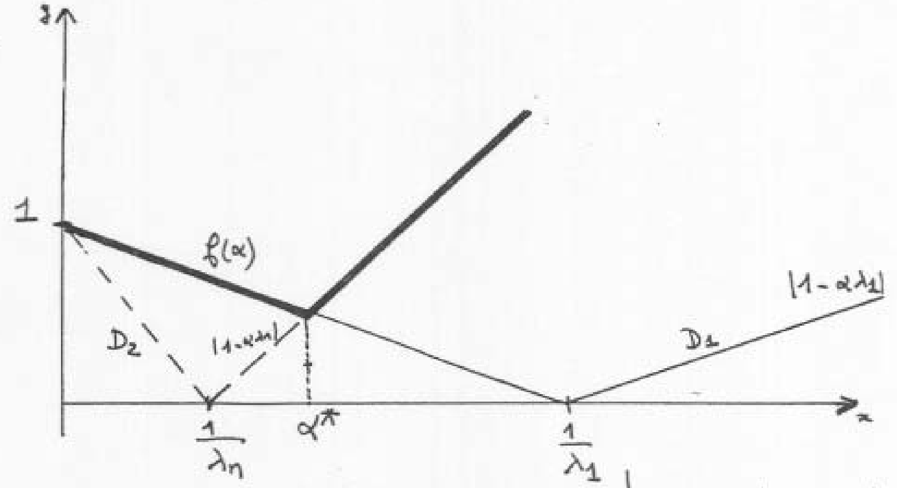
$$y_1^{(\alpha)} = |1-\alpha\lambda_1|$$

et

soit D_2 la droite d'équation

$$y_2^{(\alpha)} = |1-\alpha\lambda_n|$$

$$\left(f(\alpha) = \max\{ y_1^{(\alpha)}, y_2^{(\alpha)} \} \right)$$



α^* est tel que $y_1(\alpha^*) = y_2(\alpha^*)$

avec $\frac{1}{\lambda_n} < \alpha^* < \frac{1}{\lambda_1}$

or pour $\frac{1}{\lambda_n} < \alpha < \frac{1}{\lambda_1}$ on a : $y_1(\alpha) = 1-\alpha\lambda_1$; $y_2(\alpha) = \alpha\lambda_n - 1$

Ainsi α^* vérifie $1-\alpha^*\lambda_1 = \alpha^*\lambda_n - 1$

D'où $$\alpha^* = \frac{2}{\lambda_1+\lambda_n}$$

* Montrons que $\rho(I-\alpha^* A) = \frac{\text{cond}_2(A)-1}{\text{cond}_2(A)+1}$

On a $\rho(I-\alpha^* A) = f(\alpha^*) = 1-\alpha^*\lambda_1 \quad (= \alpha^*\lambda_n - 1)$

$$= 1 - \frac{2}{\lambda_1+\lambda_n}\lambda_1 = \frac{\lambda_n-\lambda_1}{\lambda_n+\lambda_1} = \frac{\frac{\lambda_n}{\lambda_1}-1}{\frac{\lambda_n}{\lambda_1}+1}$$

or A étant symétrique, on a $\text{cond}_2(A) = \frac{\lambda_n}{\lambda_1}$ (*), et par suite on obtient

$$\rho(I-\alpha^* A) = \frac{\text{cond}_2(A)-1}{\text{cond}_2(A)+1}$$

(*) A étant inversible, les valeurs propres de A sont toutes non nulles. ■

EXERCICE 2

Soit (S) : $Ax = b$, avec

$$A = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 1 & 2 \end{pmatrix} \; ; \; b = \begin{pmatrix} 3 \\ 2 \\ 3 \end{pmatrix}$$

1 – On considère l'algorithme du gradient à pas constant α appliqué au système linéaire (S).

1-a. Cherchons les valeurs de α pour lesquelles l'algorithme converge.

Une itération de la méthode du gradient à pas constant α est donnée par :

$$x^{(k+1)} = x^{(k)} - \alpha \left(Ax^{(k)} - b\right) = \left(I - \alpha A\right) x^{(k)} + \alpha b.$$

On a, d'après un résultat du cours (ou d'après l'exercice 1) :

l'algorithme du gradient à pas constant α converge

$$\Updownarrow$$

$$0 < \alpha < \frac{2}{\lambda_3}$$

où λ_3 est une valeur propre de A telle que $|\lambda_3| = \rho(A)$.

Calculons λ_3 : on a

$$P_\lambda(A) = \det(A - \lambda I) = \begin{vmatrix} 2-\lambda & 0 & 1 \\ 0 & 2-\lambda & 0 \\ 1 & 1 & 2-\lambda \end{vmatrix} = (2-\lambda)^3 - (2-\lambda).$$

Les valeurs propres de A sont alors données par :

$$\lambda_1 = 1 \quad ; \quad \lambda_2 = 2 \quad ; \quad \lambda_3 = 3.$$

Ainsi l'algorithme considéré converge, si et seulement si :

$$0 < \alpha < \frac{2}{3}.$$

1-b. Déterminons la suite $(x^{(k)})_k$ générée par l'algorithme du gradient à pas constant $\alpha = \frac{1}{2}$ $\left(< \frac{2}{3}\right)$, et choisissons $x^{(0)} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$.

La suite $x^{(k)}$ est alors donnée par :

$$x^{(k+1)} = -\frac{1}{2}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix} x^{(k)} + \frac{1}{2} b = -\frac{1}{2}\begin{pmatrix} x_3^{(k)} - b_1 \\ -b_2 \\ x_1^{(k)} + x_2^{(k)} - b_3 \end{pmatrix}$$

(suite exercice 2)

$$x^{(0)} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad b = \begin{pmatrix} 3 \\ 2 \\ 3 \end{pmatrix}$$

calcul !

2 – Algorithme du gradient à pas optimal :

$$\begin{cases} x^{(0)} \text{ donné} \\ g^{(k)} = A x^{(k)} - b \\ \rho_k = \|g^{(k)}\|_2^2 / (A g^{(k)}, g^{(k)}) \\ x^{(k+1)} = x^{(k)} - \rho_k g^{(k)} \end{cases}$$

$$A = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 2 \end{pmatrix} ; \; b = \begin{pmatrix} 3 \\ 2 \\ 3 \end{pmatrix}$$

<u>1ère itération</u> : $k = 0$

<u>2ème itération</u> : $k = 1$

(suite exercice 2)

3- Algorithme du gradient conjugué :

$x^{(0)}$ donné

$w^{(0)} = g^{(0)} = Ax^{(0)} - b$, si $g^{(0)} = 0$ Alors $x^{(0)}$ est solution , Stop.

$$\rho_0 = \frac{-\|g^{(0)}\|_2^2}{(Aw^{(0)}, w^0)}$$

$$x^{(1)} = x^{(0)} + \rho_0 w^{(0)}$$

$$g^{(1)} = g^{(0)} + \rho_0 Aw^{(0)}$$

$k = 1$

$$\alpha_k = \frac{\|g^{(k)}\|^2}{\|g^{(k-1)}\|^2}$$

$$w^{(k)} = g^{(k)} + \alpha_k w^{(k-1)}$$

$$\rho_k = -\frac{(g^{(k)}, w^{(k)})}{(Aw^{(k)}, w^{(k)})}$$

$$x^{(k+1)} = x^{(k)} + \rho_k w^{(k)}$$

$$g^{(k+1)} = g^{(k)} + \rho_k Aw^{(k)}$$

si $g^{(k+1)} = 0$ Alors $x^{(k+1)}$ est solution

Sinon

$k = k+1$

Nombre d'opérations élémentaires : Pour 1 itération $\sim 2n^2$ o.E

$A = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 2 \end{pmatrix}$; $b = \begin{pmatrix} 3 \\ 2 \\ 3 \end{pmatrix}$; $x^{(0)} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$

$g^{(0)} = \begin{pmatrix} -3 \\ -2 \\ -3 \end{pmatrix} = w^{(0)}$; $\rho_0 = \frac{-\|g^{(0)}\|_2^2}{(Ag^{(0)}, g^{(0)})}$; $\|g^{(0)}\|_2^2 =$; $(Ag^{(0)}, g^{(0)}) =$

(suite exercice 2)

Calcul!

EXERCICE 3

Soient $(g^{(k)})_k$, $(w^{(k)})_k$, $(x^{(k)})_k$ les suites de vecteurs générées par l'algorithme du gradient conjugué.

$$\begin{cases} g^{(0)} = w^{(0)} \; ; \quad g^{(0)} = Ax^{(0)} - b \; ; \quad x^{(1)} = x^{(0)} + \rho_0 w^{(0)} ; \quad g^{(1)} = g^{(0)} + \rho_0 A w^{0}. \qquad \rho_0 = -\|g^{(0)}\|^2 / (Ag^{(0)}, g^{(0)}) \\ k \geq 1 ; \quad \alpha_k = \dfrac{\|g^{(k)}\|}{\|g^{(k-1)}\|^2} \; ; \quad w^{(k)} = g^{(k)} + \alpha_k w^{(k-1)} \; ; \quad \rho_k = -\dfrac{(g^{(k)}, w^{(k)})}{(Aw^{(k)}, w^{(k)})} \\ x^{(k+1)} = x^{(k)} + \rho_k \, w^{(k)} \; ; \quad g^{(k+1)} = g^{(k)} + \rho_k A w^{(k)}. \end{cases}$$

On suppose que $g^{(i)} \neq 0$, $\forall i = 0, 1, \ldots, k-1$ (alors $w^{(i)} \neq 0$, $\forall i = 0, 1, \ldots, k-1$, c.f. cours).

Notation : $V(y^{(1)}, \ldots, y^{(p)})$ = l'espace vectoriel engendré par $y^{(1)}, \ldots, y^{(p)}$.

1 – Montrons que :

$$V(g^{(0)}, g^{(1)}, \ldots, g^{(k-1)}) = V(w^{(0)}, w^{(1)}, \ldots, w^{(k-1)}) = V(g^{(0)}, Ag^{(0)}, \ldots, A^{(k-1)} g^{(0)}) \equiv E_k$$

Notons.
$$A_k = V(g^{(0)}, g^{(1)}, \ldots, g^{(k-1)})$$
$$B_k = V(w^{(0)}, w^{(1)}, \ldots, w^{(k-1)})$$
$$E_k = V(g^{(0)}, Ag^{(0)}, \ldots, A^{(k-1)} g^{(0)}).$$

$A^{(i)}$ désigne la puissance ième de A : $\underbrace{A \times A \times \cdots \times A}_{i \text{ fois}}$

Montrons alors que $A_k = B_k = E_k$ (par récurrence sur k).

Pour $k = 0$; on a $A_0 = V(g^{(0)}) = V(w^{(0)}) = B_0$ puisque $w^{(0)} = g^{(0)}$

d'où $A_0 = B_0 = E_0$

Supposons que le résultat soit vrai jusqu'à l'ordre k :

$$A_\ell = B_\ell = E_\ell \qquad \forall \; 0 \leq \ell \leq k$$

et montrons que $A_{k+1} = B_{k+1} = E_{k+1}$.

• Utilisant les relations $g^{(p)} = w^{(p)} - \alpha_p w^{(p-1)}$, $\forall p$, on obtient

$$A_{k+1} = V(g^{(0)}, g^{(1)}, \ldots, g^{(k)}) \subset V(w^{(0)}, w^{(1)}, \ldots, w^{(k)}) = B_{k+1}$$

• On a $g^{(k)} = g^{(k-1)} + \rho_{k-1} A \underbrace{w^{(k-1)}} =$

combinaison linéaire de $g^{(0)}, Ag^{(0)}, \ldots, A^{(k-1)} g^{(0)}$ par hypothèse de récurrence.

$$= \sum_{i=0}^{k-1} \beta_i \, A^{(i)} g^{(0)} + \rho_{k-1} A \sum_{i=0}^{k-1} \gamma_i A^{(i)} g^{(0)} = \sum_{i=0}^{k-1} \beta_i A^{(i)} g^{(0)} + \rho_{k-1} \gamma_i A^{(i+1)} g^{(0)}$$

(suite exercice 3)

Il s'en suit que

$$A_{k+1} = V\left(g^{(0)}, g^{(1)}, \ldots, g^{(k-1)}, g^{(k)}\right) \subset E_{k+1} = V\left(g^{(0)}, Ag^{(0)}, \ldots, A^{(k-1)}g^{(0)}, A^{(k)}g^{(0)}\right)$$

puisque $g^{(\ell)}$, $0 \leq \ell \leq k-1$ est combinaison linéaire de $g^{(0)}, Ag^{(0)}, \ldots, A^{(k-1)}g^{(0)}$ par hypothèse de récurrence.

Réciproquement, montrons que $B_{k+1} \subset A_{k+1}$ et que $E_{k+1} \subset A_{k+1}$.

$B_{k+1} = V\left(w^{(0)}, w^{(1)}, \ldots, w^{(k-1)}, w^{(k)}\right)$.

Par hypothèse de récurrence, $w^{(\ell)}$, $0 \leq \ell \leq k-1$ st combinaison linéaire de $g^{(0)}, \ldots, g^{(k-1)}$.

On a $w^{(k)} = g^{(k)} + \alpha_k w^{(k-1)}$. Et par suite $w^{(k)}$ st combinaison linéaire de $g^{(0)}, g^{(1)}, \ldots, g^{(k)}$. Ainsi $B_{k+1} \subset A_{k+1}$.

$E_{k+1} = V\left(g^{(0)}, Ag^{(0)}, \ldots, A^{(k-1)}g^{(0)}, A^{(k)}g^{(0)}\right)$

$A_{k+1} = V\left(g^{(0)}, g^{(1)}, \ldots, g^{(k-1)}, g^{(k)}\right)$.

Par hypothèse de récurrence: $E_k = A_k$. Pour montrer que $E_{k+1} \subset A_{k+1}$, il suffit de montrer que $A^{(k)}g^{(0)}$ est combinaison linéaire de $g^{(0)}, g^{(1)}, \ldots, g^{(k)}$.

Puisque $g^{(k)} \in A_{k+1} \subset E_{k+1}$, alors

$$g^{(k)} = \sum_{i=0}^{k} \delta_i A^{(i)} g^{(0)} = \underbrace{\sum_{i=0}^{k-1} \delta_i A^{(i)} g^{(0)}}_{\substack{= \sum_{i=0}^{k-1} \beta_i g^{(i)} \text{ par hypothèse de récurrence.} \\ (\in A_k)}} + \delta_k A^{(k)} g^{(0)}$$

D'où $A^{(k)}g^{(0)} = \frac{1}{\delta_k}\left(g^{(k)} - \sum_{i=0}^{k-1} \beta_i g^{(i)}\right)$. Ce qui prouve que $A^{(k)}g^{(0)} \in A_{k+1}$.

Ainsi $E_{k+1} \subset A_{k+1}$. finalement on a $A_{k+1} = B_{k+1} = E_{k+1}$

Conclusion: $A_k = B_k = E_k$

2 – Montrons que : $J(x^{(k)}) \leq J(x)$, $\forall x \in x^{(0)} + E_k$. Où $J(x) = \frac{1}{2}(Ax, x) - (b, x)$

On a $x \in x^{(0)} + E_k \iff x = x^{(0)} + \sum_{i=0}^{k-1} \gamma_i A^{(i)} g^{(0)}$

Or
$$x^{(k)} = x^{(k-1)} + \rho_{k-1} w^{(k-1)} = x^{(k-2)} + \rho_{k-2} w^{(k-2)} + \rho_{k-1} w^{(k-1)} = \ldots \ldots =$$
$$= x^{(0)} + \underbrace{\rho_0 w^{(0)} + \rho_1 w^{(1)} + \cdots + \rho_{k-1} w^{(k-1)}}_{\in B_k = E_k}$$

$\Rightarrow x^{(k)} \in E_k + x^{(0)}$.

Ainsi $J(x^{(k)}) \leq J(x)$, $\forall x \in x^{(0)} + E_k$ $\quad \boxed{J(x^{(k)}) = \inf_{x \in x^{(0)} + E_k} J(x)}$

$\iff J(x^{(k)}) \leq J(x^{(k)} + y)$ $\forall y \in E_k$

puisque $x^{(k)} \in E_k + x^{(0)}$.

(suite exercice 3)

or $J(x^{(k)}) \leq J(x^{(k)}+y) \quad \forall y \in E_k$

Soit $f(t) = J(x^{(k)}+ty)$; $t \in \mathbb{R}$, $y \in E_k$ $\qquad (y \in E_k,\ ty \in E_k)$

Ainsi $f(0) \leq f(t) \quad \forall t \quad (\forall y \in E_k)$

Ainsi $f'(0) = 0$ or $f'(0) = (\nabla J(x^{(k)}), y)$

Ainsi $\left[J(x^{(k)}) \leq J(x^{(k)}+y) \quad \forall y \in E_k\right]$

$$\Updownarrow$$

$$(\underbrace{\nabla J(x^{(k)})}_{g^k}, y) = 0 \quad \forall y \in E_k$$

or, d'après 1-, $E_k = A_k = V(g^{(0)}, g^{(1)}, \dots, g^{(k-1)})$

Ainsi $\left(J(x^{(k)}) \leq J(x^{(k)}+y) \quad \forall y \in E_k\right) \iff (g^k, y) = 0 \quad \forall y \in A_k$

$\iff (g^k, g^j) = 0 \quad \forall\ 0 \leq j \leq k-1$

(ce qui est vrai, c.f. cours).

Ainsi $J(x^{(k)}) \leq J(x)$, $\forall x \in x^{(0)} + E_k$.

EXERCICE 4

Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice inversible non nécessairement symétrique.

On a :

$$Ax = b \iff A^T A x = A^T b \ .$$

La matrice $A^T A$ est définie positive, puisque A est inversible.

Algorithme du GC appliqué à $Bx = c$	Algorithme du GC appliqué à $A^TAx = A^Tb$ (sans avoir à calculer A^TA).
$x^{(0)}$; $g^{(0)} = Bx^{(0)} - c$; $w^{(0)} = g^{(0)}$	$x^{(0)}$, $g^{(0)} = Ax^{(0)} - b$; $w^{(0)} = A^T g^{(0)}$
$\rho_k = -\Vert g^{(k)}\Vert_2^2 / (Bw^{(k)}, w^{(k)})$	$\rho_k = \dfrac{-\Vert A^T g^{(k)}\Vert_2^2}{(A^TAw^{(k)}, w^{(k)})} = \dfrac{\Vert A^T g^{(k)}\Vert_2^2}{\Vert Aw^{(k)}\Vert_2^2}$
$x^{(k+1)} = x^{(k)} + \rho_k w^{(k)}$	$x^{(k+1)} = x^{(k)} + \rho_k w^{(k)}$
$g^{(k+1)} = g^{(k)} + \rho_k B w^{(k)}$	$g^{(k+1)} = g^{(k)} + \rho_k A w^{(k)}$
$\alpha_{k+1} = \dfrac{\Vert g^{(k+1)}\Vert_2^2}{\Vert g^{(k)}\Vert_2^2}$	$\alpha_{k+1} = \Vert A^T g^{(k+1)}\Vert_2^2 / \Vert A^T g^{(k)}\Vert_2^2$
$w^{(k+1)} = g^{(k+1)} + \alpha_{k+1} w^{(k)}$	$w^{(k+1)} = A^T g^{(k+1)} + \alpha_{k+1} w^{(k)}$.
Test d'arrêt sur $\Vert g^{(k)}\Vert_2$	Test d'arrêt sur $\Vert A^T g^{(k)}\Vert_2$

Fin Série 4