Corrigé du partiel d'Analyse Numérique L3 Mathématiques, 23 octobre 2015.

Exercice 1 : Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice symétrique définie positive (SDP), et $b \in \mathbb{R}^n$ fixé, on note $\bar{x} \in \mathbb{R}^n$ la solution du système $Ax = b$. On note $A = D - E - F$ où D est la partie diagonale de A, $-E$ sa partie triangulaire inférieure stricte, et $-F$ sa partie triangulaire supérieure stricte. Précisément, on a pour tous $(i,j) \in \{1, \cdots, n\}^2$

$$D_{i,j} = \begin{cases} A_{i,i} & \text{si } i = j, \\ 0 & \text{si } i \neq j, \end{cases} \qquad E_{i,j} = \begin{cases} -A_{i,j} & \text{si } j < i, \\ 0 & \text{si } j \geq i, \end{cases} \qquad F_{i,j} = \begin{cases} -A_{i,j} & \text{si } j > i, \\ 0 & \text{si } j \leq i. \end{cases}$$

Soit $\omega > 0$, on considère la méthode itérative suivante (SSOR) : étant donné $x^{(0)} \in \mathbb{R}^n$, on calcule pour tous $k \in \mathbb{N}$ le vecteur $x^{(k+1)}$ selon la relation de récurrence

$$x^{(k+1)} = x^{(k)} + C^{-1}\,(b - Ax^{(k)}),$$

où le préconditionnement C est défini par la matrice

$$C = (\frac{D}{\omega} - E)D^{-1}(\frac{D}{\omega} - F).$$

(1) Montrer que la matrice C est symétrique définie positive.

Soit $e^{(i)} \in \mathbb{R}^n$ tel que $e^{(i)}_j = \delta_{i,j}$ pour tous $j = 1, \cdots, n$. La matrice A étant SDP on en déduit que $A_{i,i} = (Ae^{(i)}, e^{(i)}) > 0$ et donc la matrice diagonale D ayant tous ses éléments diagonaux strictement positifs est SDP.
Comme A est symétrique on a $E^t = F$ et $F^t = E$ donc aussi $(\frac{D}{\omega} - E)^t = (\frac{D}{\omega} - F)$ et $(\frac{D}{\omega} - F)^t = (\frac{D}{\omega} - E)$. On en déduit que

$$C^t = (\frac{D}{\omega} - F)^t D^{-t} (\frac{D}{\omega} - E)^t = (\frac{D}{\omega} - E)D^{-1}(\frac{D}{\omega} - F) = C,$$

donc que C est une matrice symétrique. Soit un vecteur $x \in \mathbb{R}^n$. On pose $y = (\frac{D}{\omega} - F)x$, on a donc

$$\begin{aligned}(Cx, x) &= \Big((\frac{D}{\omega} - E)D^{-1}(\frac{D}{\omega} - F)x, x\Big) = \Big(D^{-1}(\frac{D}{\omega} - F)x, (\frac{D}{\omega} - E)^t x\Big) \\ &= \Big(D^{-1}(\frac{D}{\omega} - F)x, (\frac{D}{\omega} - F)x\Big) = (D^{-1}y, y).\end{aligned}$$

Comme D^{-1} est SDP (du fait que D est SDP) on en déduit que $(D^{-1}y, y) \geq 0$ et donc que C est positive. Enfin si $(Cx, x) = 0$, il résulte que $(D^{-1}y, y)$, donc que $y = 0$ car D^{-1} est SDP et enfin que $x = 0$ car la matrice $(\frac{D}{\omega} - F)$ est inversible (du fait qu'il s'agit d'une matrice triangulaire avec termes diagonaux $A_{i,i}$ non nuls pour tous $i = 1, \cdots, n$).

(2) Calculer la matrice B telle que

$$(\bar{x} - x^{(k+1)}) = B(\bar{x} - x^{(k)}) \text{ pour tout } k \in \mathbb{N}.$$

En remarquant que $b = A\bar{x}$ on a
$$x^{(k+1)} - \bar{x} = x^{(k)} - \bar{x} + C^{-1}\ (A\bar{x} - Ax^{(k)}) = (I - C^{-1}A)(\bar{x} - x^{(k)})$$
d'où $B = (I - C^{-1}A)$.

(3) Calculer la matrice $G = C^t + C - A$ en l'exprimant comme combinaison linéaire des matrices D, A et $ED^{-1}F$.

$$\begin{aligned} G = C^t + C - A = 2C - A &= 2(\tfrac{D}{\omega} - E)D^{-1}(\tfrac{D}{\omega} - F) - A = 2(\tfrac{D}{\omega} - E)(\tfrac{I}{\omega} - D^{-1}F) \\ &= \tfrac{2}{\omega^2}D - \tfrac{2}{\omega}(E + F) + 2ED^{-1}F - A. \end{aligned}$$
En utilisant la définition $A = D - E - F$ soit $E + F = D - A$ on obtient que
$$G = \frac{2}{\omega}(\frac{1}{\omega} - 1)D + (\frac{2}{\omega} - 1)A + 2ED^{-1}F.$$

(4) Montrer que la matrice $ED^{-1}F$ est symétrique positive.

Comme $E^t = F$ et $F^t = E$ il résulte que $ED^{-1}F$ est symétrique. Ensuite pour tous $x \in \mathbb{R}^n$, on pose $y = Fx$ et on a
$$(ED^{-1}Fx, x) = (D^{-1}Fx, E^t x) = (D^{-1}Fx, Fx) = (D^{-1}y, y).$$
Comme D^{-1} est SDP on en déduit que $(ED^{-1}Fx, x) \geq 0$ donc que $ED^{-1}F$ est positive. En revanche $ED^{-1}F$ n'est pas SDP car F n'est pas inversible.

(5) En déduire que G est symétrique définie positive si $\omega \in]0, 1]$. D'après un lemme du cours il en résulte que $\rho(B) < 1$ si $\omega \in]0, 1]$. En déduire que la suite $(x^{(k)})_{k\in\mathbb{N}}$ converge vers $\bar{x}$ si $\omega \in]0, 1]$.

Si $\omega \in]0, 1[$ on a $\frac{2}{\omega}(\frac{1}{\omega} - 1) > 0$ et $(\frac{2}{\omega} - 1) > 0$. Pour $\omega = 1$ on a $\frac{2}{\omega}(\frac{1}{\omega} - 1) = 0$ et $(\frac{2}{\omega} - 1) = 1$. Comme la somme d'une matrice symétrique positive et d'une matrice SDP est SDP, on en déduit que G est SDP pour $\omega \in]0, 1]$. D'après le cours il en résulte que $\rho(B) < 1$ et donc que la suite $(x^{(k)})_{k\in\mathbb{N}}$ converge vers $\bar{x}$.

Exercice 2 : soit $A \in \mathcal{M}_n$ une matrice symétrique définie positive (SDP) et $C \in \mathcal{M}_n$ une matrice SDP définissant le préconditionnement. On suppose qu'il existe deux constantes $0 < c_1 \leq c_2$ telles que pour tous $x \in \mathbb{R}^n$ on ait
$$c_1(Cx, x) \leq (Ax, x) \leq c_2(Cx, x).$$
On rappelle que pour tout $s \in \mathbb{R}$, C^s est une matrice SDP et que l'on a pour tous $s \in \mathbb{R}$, $s' \in \mathbb{R}$ la relation $C^s C^{s'} = C^{s+s'}$.

(1) Montrer que $G = C^{-1/2}AC^{-1/2}$ est une matrice SDP.

Comme $C^{-1/2}$ et A sont symétriques, on déduit que G est symétrique. Ensuite soit $x \in \mathbb{R}^n$ on calcule (Gx, x) en posant $y = C^{-1/2}x$. En tenant compte de la symétrie de $C^{-1/2}$ on a
$$(Gx, x) = (C^{-1/2}AC^{-1/2}x, x) = (AC^{-1/2}x, C^{-1/2}x) = (Ay, y).$$
Comme A est SDP, on en déduit que $(Ay, y) = (Gx, x) \geq 0$ et donc que G est positive. Enfin $(Gx, x) = 0$ implique $y = 0$ car A est SDP puis que $x = 0$ car $C^{-1/2}$ est inversible.

(2) On note $0 < \lambda_1 \leq \cdots \leq \lambda_i \cdots \leq \lambda_n$ les valeurs propres de G. Montrer que

$$\lambda_n \leq c_2$$

et

$$\lambda_1 \geq c_1.$$

On pourra considérer un vecteur propre $u \in \mathbb{R}^n$ de G de valeur propre λ et faire le changement de variable $x = C^{-\frac{1}{2}}u$ dans la relation $Gu = \lambda u$. En déduire que le conditionnement de la matrice G défini par $Cond_2(G) = \frac{\lambda_n}{\lambda_1}$ verifie

$$Cond_2(G) \leq \frac{c_2}{c_1}.$$

Soit un vecteur propre $u \in \mathbb{R}^n$ de G de valeur propre λ. On considère le changement de variable $x = C^{-\frac{1}{2}}u$ dans la relation $Gu = \lambda u$. On obtient

$$C^{-1/2}AC^{-1/2}C^{1/2}x = \lambda C^{1/2}x,$$

soit

$$Ax = \lambda Cx.$$

En prenant le produit scalaire de chacun des membres de cette égalité par le vecteur x on obtient que

$$(Ax, x) = \lambda(Cx, x)$$

soit comme x est non nul $\lambda = \frac{(Ax,x)}{(Cx,x)}$. On déduit de l'hypothèse que

$$c_1 \leq \lambda \leq c_2,$$

donc que $\lambda_n \leq c_2$ et que $\lambda_1 \geq c_1$. Il vient donc que $Cond_2(G) \leq \frac{c_2}{c_1}$

(3) On considère l'algorithme suivant de Richardson à pas variable préconditionné par C pour résoudre le système $Ax = b$: étant donné $x^{(0)} \in \mathbb{R}^n$ on calcule la suite

$$x^{(k+1)} = x^{(k)} + \alpha^{(k)}C^{-1}(b - Ax^{(k)}),\ k \in \mathbb{N}.$$

On note $d = C^{-1/2}b$ et on considère le changement de variable $y = C^{1/2}x$. Montrer que l'algorithme de Richardson précédent équivaut à $y^{(0)} = C^{1/2}x^{(0)}$ et

$$y^{(k+1)} = y^{(k)} + \alpha^{(k)}(d - Gy^{(k)}),\ k \in \mathbb{N}.$$

En multipliant les deux membre de l'équation de récurrence par $C^{1/2}$ et en utilisant le changement de variable $y^{(k)} = C^{1/2}x^{(k)}$, on obtient que

$$y^{(k+1)} = y^{(k)} + \alpha^{(k)}C^{-1/2}(b - AC^{-1/2}y^{(k)}) = y^{(k)} + \alpha^{(k)}(d - Gy^{(k)}).$$

(4) D'après le cours un bon choix pour $\alpha^{(k)}$ minimisant l'erreur à l'itération $k+1$ est le suivant

$$\alpha^{(k)} = \frac{(s^{(k)}, s^{(k)})}{(Gs^{(k)}, s^{(k)})}.$$

avec $s^{(k)} = d - Gy^{(k)}$. Montrer que $\alpha^{(k)}$ s'écrit aussi en posant $r^{(k)} = b - Ax^{(k)}$ comme suit :

$$\alpha^{(k)} = \frac{(C^{-1}r^{(k)}, r^{(k)})}{(C^{-1}r^{(k)}, AC^{-1}r^{(k)})}.$$

On a

$$s^{(k)} = d - Gy^{(k)} = C^{-1/2}b - C^{-1/2}AC^{-1/2}C^{1/2}x^{(k)} = C^{-1/2}r^{(k)},$$

d'où

$$\alpha^{(k)} = \frac{(s^{(k)}, s^{(k)})}{(Gs^{(k)}, s^{(k)})} = \frac{(C^{-1/2}r^{(k)}, C^{-1/2}r^{(k)})}{(C^{-1/2}AC^{-1/2}C^{-1/2}r^{(k)}, C^{-1/2}r^{(k)})}$$

Comme $C^{-1/2}$ est symétrique et $C^{-1/2}C^{-1/2} = C^{-1}$ on a

$$\alpha^{(k)} = \frac{(r^{(k)}, C^{-1}r^{(k)})}{(AC^{-1}r^{(k)}, C^{-1}r^{(k)})} = \frac{(C^{-1}r^{(k)}, r^{(k)})}{(C^{-1}r^{(k)}, AC^{-1}r^{(k)})}.$$

(5) Ecrire l'algorithme avec l'initialisation et la boucle itérative pour le calcul de la suite $x^{(k)}$, $\kappa \in \mathbb{N}$, en faisant apparaître un seul produit matrice vecteur avec la matrice A et un seul préconditionnement (résolution d'un système $Cp = r$) par itération de l'algorithme.

— Choix de la précision ϵ sur le résidu relatif
— Initialisation : $x^{(0)}, r^{(0)} = b - Ax^{(0)}, nr = nr^{(0)} = \|r^{(0)}\|$
— Itérer tant que $\frac{nr}{nr^{(0)}} \geq \epsilon$
  — résoudre le système $Cq^{(k)} = r^{(k)}$
  — $p^{(k)} = Aq^{(k)}$
  — $\alpha^{(k)} = \frac{(q^{(k)}, r^{(k)})}{(p^{(k)}, q^{(k)})}$
  — $x^{(k+1)} = x^{(k)} + \alpha^{(k)}q^{(k)}$
  — $r^{(k+1)} = r^{(k)} - \alpha^{(k)}p^{(k)}$
  — $nr = \|r^{(k+1)}\|$

(6) On note $\bar{y}$ la solution du système $Gy = d$, $f^{(k)} = \bar{y} - y^{(k)}$ les erreurs pour $k \in \mathbb{N}$. On a montré en cours que

$$(Gf^{(k)}, f^{(k)}) \leq (1 - \frac{1}{Cond_2(G)})^k (Gf^{(0)}, f^{(0)}).$$

En déduire une condition suffisante sur le nombre d'itérations k de la méthode itérative pour que l'erreur relative vérifie

$$\frac{(Gf^{(k)}, f^{(k)})^{\frac{1}{2}}}{(Gf^{(0)}, f^{(0)})^{\frac{1}{2}}} \leq \epsilon$$

pour un $\epsilon > 0$ fixé.

La condition $\frac{(Gf^{(k)}, f^{(k)})^{\frac{1}{2}}}{(Gf^{(0)}, f^{(0)})^{\frac{1}{2}}} \leq \epsilon$ est vérifiée dès que

$$(1 - \frac{1}{Cond_2(G)})^k \leq \epsilon^2$$

ce qui équivaut à

$$k \;\; ln((1 - \frac{1}{Cond_2(G)})) \leq 2 \; ln(\epsilon),$$

soit

$$k \geq 2 \frac{ln(\epsilon)}{ln((1 - \frac{1}{Cond_2(G)}))}.$$