MT09-A2017 – Examen médian – Questions de cours
Durée : 30mn. Sans documents ni outils électroniques – Rédiger sur l'énoncé

NOM PRÉNOM : Place n^o:

ATTENTION, il y a 3 exercices indépendants pour cette partie questions de cours!

Exercice 1 *(barème approximatif : 2 points)*

Soient C une matrice carrée de $\mathcal{M}_{nn}(\mathbb{R})$ ($n > 0$) et d un vecteur de $\mathbb{R}^n$. On étudie l'itération linéaire :

$$\begin{cases} x^{(k+1)} = Cx^{(k)} + d & \forall k = 0, 1, \dots \\ x^{(0)} \text{ donné.} \end{cases}$$

1. Donner une condition nécessaire sur C pour que la suite $(x^{(k)})_{k\in\mathbb{N}}$ converge.

   Réponse : cf. cours : s'il existe une norme subordonnée $|||\cdot|||$ telle que $|||C||| < 1$, alors $(x^{(k)})_{k\in\mathbb{N}}$ converge quel que soit le vecteur de départ $x^{(0)}$. ☐

2. On veut résoudre le système $Ax = b$. Donner la matrice C et le vecteur d dans le cas où on applique la méthode de Gauss-Seidel. On définira les matrices D, E et F du cours.

   Réponse : cf. cours : $C = (D - E)^{-1}F$ et $d = (D - E)^{-1}b$. ☐

3. Application : dire si pour la matrice $A = \begin{bmatrix} 1 & \frac{3}{4} \\ -1 & 1 \end{bmatrix}$, la méthode de Gauss-Seidel converge ou non.

   Réponse : il vient :

$$C = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 0 & -\frac{3}{4} \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & -\frac{3}{4} \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -\frac{3}{4} \\ 0 & 0 \end{bmatrix},$$

   donc $|||C|||_1 = |||C|||_\infty = \frac{3}{4} < 1$ donc Gauss-Seidel converge (et $|||C|||_2 = \sqrt{\rho(C^TC)} = \frac{3}{4}$ aussi). ☐

Exercice 2 *(barème approximatif : 1.5 points)*

Soit une matrice $A \in \mathcal{M}_{nn}(\mathbb{R})$ pour $n > 0$ une matrice symétrique définie positive.

1. Montrer que toute les sous-matrices principales de A, notées $[A]_k$ pour $k = 1, \dots, n$, sont symétriques définies positives.

   Réponse : revoir la définition de SDP. Soit k fixé dans $\{1, \dots, n\}$. Il est évident que $[A]_k$ est symétrique. De plus pour $y \in \mathbb{R}^k$, si $y \neq 0$, alors

$$y^T[A]_k y = \sum_{i=1}^{k}\sum_{j=1}^{k} y_i a_{ij} y_j = [y^T 0 \dots 0] A \begin{bmatrix} y \\ 0 \\ \vdots \\ 0 \end{bmatrix} > 0,$$

   car A est SDP et $x = \begin{bmatrix} y \\ 0 \\ \vdots \\ 0 \end{bmatrix} \in \mathbb{R}^n$ est non nul. ☐

2. Déterminer le noyau d'une matrice symétrique définie positive.

   Réponse : soit $x \in \mathrm{Ker}(A)$ où A est SDP. Alors $Ax = 0$, ce qui implique $x^TAx = 0$ et donc $x = 0$. Donc $\mathrm{Ker}(A) = \{0\}$ et A, carrée, est inversible. ☐

3. Conclure sur la faisabilité de la factorisation $A = LU$.

   Réponse : soit A SDP, donc toutes ses sous-matrices principales sont aussi SDP donc elles sont inversibles. C'est une CNS pour que la factorisation $A = LU$ soit faisable sans permutation. ☐

Exercice 3 *(barème approximatif : 2 points)*

1. Définir l'ensemble des flottants $\mathcal{F}_{10}$. On expliquera ce que signifie les constantes t, L et U (notations du cours).

   Réponse : cf. cours.

   $$\mathcal{F}_{10} = \Big\{ \pm 0.d_1 d_2 \dots d_t \ 10^e \mid d_i \in \{0,1,2,\dots,9\} \ \forall i = 1,\dots,t, \quad d_1 \neq 0, \quad L \leq e \leq U \Big\} \cup \{0\},$$

   où t est le nombre de chiffres significatifs, L et U constituent les bornes inférieure et supérieure de l'exposant e. Par convention, l'exposant e est choisi de façon que le premier chiffre d_1 soit toujours non-nul. Le nombre 0 est explicitement inséré dans $\mathcal{F}_{10}$ car 0 ne s'écrit pas comme un nombre flottant normal. $\square$

2. Dans le reste de cet exercice, on prend $t = 3$, $L = -1$, $U = 3$.

   (a) Donner l'écart absolu entre deux flottants successifs. Que vaut $\varepsilon_{\mathrm{mach}}$?

   Réponse : pour $e \in \{-1,0,1,2,3\}$, sur chaque intervalle $[0.1,\ 1] \times 10^e$, l'écart entre deux flottants successifs f_1 et f_2, noté $\delta_e = |f_2 - f_1|$, vaut : $\delta_e = 0.001 \times 10^e = 10^{-t+e}$. Dans le cours, on définit $\varepsilon_{\mathbf{mach}} = \frac{1}{2}10^{-t+1} = 0.5 \times 10^{-2}$. $\square$

3. (a) Calculer en addition flottante : $x = 100 \oplus 0.6$. Que vaut l'erreur relative?

   Réponse : il faut aligner les nombres sur l'exposant le plus grand. Note: on prend ici la notation scientifique standard $d_1.d_2d_3 \times 10^{e'}$ ($d_1 \neq 0$, $e' = e - 1 \in \{-2,-1,0,1,2\}$) que les humains sont en général plus habitués à manipuler.

   $$\begin{aligned} \tilde{x} &= 100 \oplus 0.6 = \mathrm{fl}(1.00 \times 10^2 + 6.00 \times 10^{-1}) \\ &= \mathrm{fl}(1.00 \times 10^2 + 0.006 \times 10^2) = \mathrm{fl}(1.006 \times 10^2) \\ &= 1.01 \times 10^2 \end{aligned}$$

   car on arrondit au plus proche flottant de $\mathcal{F}_{10}$. La valeur exacte vaut $x = 100.6$. L'erreur relative vaut

   $$\frac{|x - \tilde{x}|}{|x|} = \frac{101 - 100.6}{100.6} = \frac{0.4}{100.6} \approx 4 \times 10^{-3} \leq \varepsilon_{\mathbf{mach}} = 5 \times 10^{-3}.$$

   $\square$

   (b) Calculer en addition flottante : $y = (100 \oplus 0.6) \ominus 100$. Que vaut l'erreur relative? Commenter.

   Réponse : il vient :

   $$\begin{aligned} \tilde{y} &= (100 \oplus 0.6) \ominus 100 = \tilde{x} \ominus 100 \\ &= \mathrm{fl}(1.01 \times 10^2 - 1.00 \times 10^2) \\ &= 1.00 = 1. \end{aligned}$$

   Comme le calcul exact donne $y = 0.6$, l'erreur relative vaut $\frac{|y - \tilde{y}|}{|y|} = \frac{1-0.6}{0.6} = \frac{0.4}{0.6} = 66\%$. Il y a eu une erreur catastrophique lors du calcul de x (presque toute l'information sur 0.6 a été perdue lors de cette addition de 2 termes d'ordre de grandeur très différents), qui a été révélée lors du calcul de $\tilde{y}$ (qui est très inexact.) $\square$

MT09-A2017- Examen médian

Durée : 1h30.

Polycopiés de cours et scilab autorisés - pas d'outils numériques

Questions de cours déjà traitées : environ *5.5 points.*

RÉDIGER LES EXERCICES 2 ET 3 SUR LA MÊME COPIE!

EXERCICE 1 SUR COPIE SÉPARÉE.

Exercice 1 : *(barème approximatif : 6 points)* **CHANGEZ DE COPIE**

Il est possible de traiter une question en admettant les résultats précédents.

Toutes les matrices de cet exercice sont carrées à n lignes et n colonnes ($n > 0$). Soit la matrice $A \in \mathcal{M}_{n,n}$ définie par

$$A = \frac{1}{h}\begin{bmatrix} 4 & 1 & 0 & \cdots & \cdots & 0 \\ 1 & 4 & 1 & 0 & \cdots & 0 \\ 0 & 1 & 4 & 1 & 0 & \cdots \\ & & \ddots & \ddots & \ddots & \\ 0 & \cdots & 0 & 1 & 4 & 1 \\ 0 & \cdots & \cdots & 0 & 1 & 4 \end{bmatrix}, \quad \text{où } h = \frac{1}{n}.$$

1. Calculer $|||A|||_\infty$.

   Réponse : $|||A|||_\infty = \max_{i=1,\ldots,n} \sum_{j=1}^{n} |a_{ij}| = \frac{6}{h}$**.** ☐

2. On pose $A = \frac{4}{h}(I + N)$, où I est la matrice identité. Déterminer N et calculer $|||N|||_\infty$.

   Réponse : il vient

$$N = \frac{1}{4}\begin{bmatrix} 0 & 1 & 0 & \cdots & \cdots & 0 \\ 1 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 1 & 0 & \cdots \\ & & \ddots & \ddots & \ddots & \\ 0 & \cdots & 0 & 1 & 0 & 1 \\ 0 & \cdots & \cdots & 0 & 1 & 0 \end{bmatrix}, \text{ **et** } |||N|||_\infty = \frac{2}{4} = \frac{1}{2}.$$

☐

3. Soit $|||\cdot|||$ une norme matricielle subordonnée.

   (a) Montrer que $|||I||| = 1$.

   Réponse : on appelle $\|\cdot\|$ la norme vectorielle associée à la norme $|||\cdot|||$. On a

$$|||I||| = \sup_{x \in \mathbf{R}^n \setminus \{0\}} \frac{\|Ix\|}{\|x\|} = \sup_{x \in \mathbf{R}^n \setminus \{0\}} \frac{\|x\|}{\|x\|} = 1.$$

☐

   (b) Soit E une matrice carrée. Montrer que si $|||E||| < 1$, alors la matrice $I + E$ est inversible.

   Réponse : soit $x \in \mathrm{Ker}(I+E)$. Alors $(I+E)x = 0$ ce qui est équivalent à $x = -Ex$. On en déduit $\|x\| = \|-Ex\| = \|Ex\| \le |||E|||\ \|x\|$. Cela donne $(1 - |||E|||)\|x\| \le 0$ avec $1 - |||E||| > 0$ et $\|x\| \ge 0$. Ceci n'est possible que si $\|x\| = 0$, soit $x = 0$ et donc $(I+E)$ est injective donc inversible car carrée. ☐

   (c) Vérifier alors que $(I+E)^{-1} = I - (I+E)^{-1}E$ et en déduire

$$|||(I+E)^{-1}||| \le \frac{1}{1 - |||E|||}.$$

   Réponse : on vérifie

$$(I+E)(I - (I+E)^{-1}E) = I + E - (I+E)(I+E)^{-1}E = I + E - IE = I$$

   donc $I - (I+E)^{-1}E$ est bien l'inverse de $I+E$. On en déduit

$$|||(I+E)^{-1}||| = |||(I - (I+E)^{-1}E)||| \le |||I||| + |||(I+E)^{-1}E||| \le 1 + |||(I+E)^{-1}|||\ |||E|||,$$

   ce qui donne

$$(1 - |||E|||)|||(I+E)^{-1}||| \le 1,$$

   et on conclut en remarquant que $(1 - |||E|||) > 0$. ☐

4. Utiliser le résultat précédent pour obtenir une majoration de $|||A^{-1}|||_\infty$.

   Réponse : d'abord, on remarque que $A = \frac{4}{h}(I+N)$ avec $|||N|||_\infty = \frac{1}{2} < 1$, donc A est inversible et $A^{-1} = \frac{h}{4}(I+N)^{-1}$. On déduit de l'inégalité précédente que

$$|||A^{-1}|||_\infty = \frac{h}{4}|||(I+N)^{-1}|||_\infty \leq \frac{h}{4}\frac{1}{1-|||N|||_\infty} = \frac{h}{2}.$$

□

5. En déduire une majoration du conditionnement de A pour la norme $|||\cdot|||_\infty$.

   Réponse : on en déduit

$$\kappa_\infty(A) = |||A|||_\infty \; |||A^{-1}|||_\infty \leq \frac{6}{h}\frac{h}{2} = 3.$$

□

6. Que peut-on dire de ce comportement quand $h \to 0$?

   Réponse : le conditionnement de A est majoré de façon indépendante de h et il reste petit. C'est un excellent conditionnement. □

Exercice 2 : *(barème approximatif : 9 points)* **CHANGEZ DE COPIE**

Les questions 4, 5 et 6 sont partiellement indépendantes des précédentes.

Soit un entier $n \geq 1$. On rappelle que la factorisation de Cholesky $A = CC^T$ d'une matrice de taille n nécessite de l'ordre de $\dfrac{n^3}{6}$ multiplications. Le coût de la résolution d'un système triangulaire est de l'ordre de $\dfrac{n^2}{2}$.

Soient $n > 0$ et $p > 0$ deux entiers. Soient $A_1, A_2, \ldots, A_n$, n matrices symétriques et inversibles de $\mathcal{M}_{p,p}$. Soient également $B_1, B_2, \ldots, B_{n-1}$, $n-1$ matrices de $\mathcal{M}_{p,p}$. On définit la matrice $K \in \mathcal{M}_{np,np}$ par

$$K = \begin{bmatrix} A_1 & B_1^T & 0 & \cdots & \cdots & 0 \\ B_1 & A_2 & B_2^T & 0 & \cdots & 0 \\ 0 & B_2 & A_3 & B_3^T & 0 & \cdots \\ & & \ddots & \ddots & \ddots & \\ 0 & \cdots & 0 & B_{n-2} & A_{n-1} & B_{n-1}^T \\ 0 & \cdots & \cdots & 0 & B_{n-1} & A_n \end{bmatrix},$$

où $0 \in \mathcal{M}_{p,p}$ est la matrice nulle. On veut résoudre $K\mathbf{x} = \mathbf{f}$, où $\mathbf{f} \in \mathbb{R}^{np}$ est donné. Pour ce faire, on décompose les vecteurs $\mathbf{x}$ et $\mathbf{f}$ en blocs :

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_n \end{bmatrix}, \quad \mathbf{f} = \begin{bmatrix} \mathbf{f}_1 \\ \mathbf{f}_2 \\ \vdots \\ \mathbf{f}_n \end{bmatrix}, \quad \text{où } \mathbf{x}_i \in \mathbb{R}^p, \text{ et } \mathbf{f}_i \in \mathbb{R}^p, \quad \forall i = 1, \ldots, n.$$

1. On suppose que A_1 est symétrique définie positive.

   (a) Montrer que les p premières lignes de $K\mathbf{x} = \mathbf{f}$ sont équivalentes à

$$C_1^T\mathbf{x}_1 + D_1\mathbf{x}_2 = \mathbf{g}_1, \tag{1}$$

   où C_1 est triangulaire inférieure avec des termes diagonaux strictement positifs. On donnera l'expression de la matrice D_1 et du vecteur $\mathbf{g}_1$ en fonction de C_1, de B_1 et de $\mathbf{f}_1$.

   Réponse : On remarque que K est symétrique. Comme la matrice A_1 est SDP, elle admet une factorisation de Cholesky : il existe C_1 triangulaire inférieure avec des termes diagonaux strictement positifs telle que $A_1 = C_1C_1^T$. Les matrices C_1 et C_1^T sont inversibles car triangulaires avec des termes diagonaux tous non nuls. Le premier bloc de lignes s'écrit donc

$$\begin{aligned} A_1\mathbf{x}_1 + B_1^T\mathbf{x}_2 = \mathbf{f}_1 \quad &\Longleftrightarrow \quad C_1^T\mathbf{x}_1 + C_1^{-1}B_1^T\mathbf{x}_2 = C_1^{-1}\mathbf{f}_1 \\ &\Longleftrightarrow \quad C_1^T\mathbf{x}_1 + D_1\mathbf{x}_2 = \mathbf{g}_1 \qquad (2) \\ \textbf{en posant} \quad &\left\{ \begin{array}{l} A_1 = C_1C_1^T \\ D_1 = C_1^{-1}B_1^T \\ \mathbf{g}_1 = C_1^{-1}\mathbf{f}_1 \end{array} \right. . \qquad (3) \end{aligned}$$

□

(b) Exprimer B_1 en fonction de C_1 et de D_1.

Réponse : il vient

$$
\begin{aligned}
D_1 = C_1^{-1} B_1^T \quad &\Longleftrightarrow \quad C_1 D_1 = B_1^T \\
&\Longleftrightarrow \quad B_1 = D_1^T C_1^T .
\end{aligned} \tag{4}
$$

$\square$

2. (a) En utilisant (1), montrer que le deuxième bloc d'équations dans $K\mathbf{x} = \mathbf{f}$ (lignes $i = p+1$ à $2p$) est équivalent à

$$
\widetilde{A}_2 \mathbf{x}_2 + B_2^T \mathbf{x}_3 = \widetilde{\mathbf{f}}_2, \tag{5}
$$

où $\widetilde{A}_2$ et $\widetilde{\mathbf{f}}_2$ sont à exprimer en fonction de A_2, D_1, de $\mathbf{f}_2$ et de $\mathbf{g}_1$.

Réponse : en utilisant (2) et (4), le deuxième bloc d'équations dans $K\mathbf{x} = \mathbf{f}$ s'écrit

$$
\begin{aligned}
B_1\mathbf{x}_1 + A_2\mathbf{x}_2 + B_2^T\mathbf{x}_2 = \mathbf{f}_2 \quad &\Longleftrightarrow \quad D_1^T C_1^T \mathbf{x}_1 + A_2\mathbf{x}_2 + B_2^T\mathbf{x}_3 = \mathbf{f}_2 \\
&\Longleftrightarrow \quad (A_2 - D_1^T D_1)\mathbf{x}_2 + B_2^T\mathbf{x}_3 = \mathbf{f}_2 - D_1^T\mathbf{g}_1 \\
&\Longleftrightarrow \quad \widetilde{A}_2\mathbf{x}_2 + B_2^T\mathbf{x}_3 = \widetilde{\mathbf{f}}_2,
\end{aligned}
$$

en posant

$$
\widetilde{A}_2 = A_2 - D_1^T D_1 \ \textbf{et}\ \widetilde{\mathbf{f}}_2 = \mathbf{f}_2 - D_1^T \mathbf{g}_1 .
$$

$\square$

(b) Montrer que $\widetilde{A}_2$ est symétrique.

Réponse :

$$
\widetilde{A}_2^T = (A_2 - D_1^T D_1)^T = A_2^T - (D_1^T D_1)^T = A_2^T - D_1^T (D_1^T)^T = A_2 - D_1^T D_1 = \widetilde{A}_2,
$$

car A_2 est symétrique et $(D_1^T)^T = D_1$. $\square$

(c) On suppose que $\widetilde{A}_2$ est définie positive. Montrer que (5) est équivalent à

$$
C_2^T \mathbf{x}_2 + D_2 \mathbf{x}_3 = \mathbf{g}_2, \tag{6}
$$

où C_2 est triangulaire inférieure avec des termes diagonaux strictement positifs. On donnera l'expression de la matrice D_2 et du vecteur $\mathbf{g}_2$ en fonction de C_2, de B_2, de D_1, de $\mathbf{f}_2$ et de $\mathbf{g}_1$.

Réponse : on refait comme pour le bloc 1. Comme la matrice $\widetilde{A}_2$ est SDP, elle admet une factorisation de Cholesky : il existe C_2 triangulaire inférieure avec des termes diagonaux strictement positifs telle que $\widetilde{A}_2 = C_2 C_2^T$. Les matrices C_2 et C_2^T sont inversibles. On en déduit

$$
\begin{aligned}
B_1\mathbf{x}_1 + A_2\mathbf{x}_2 + B_2^T\mathbf{x}_2 = \mathbf{f}_2 \quad &\Longleftrightarrow \quad \widetilde{A}_2\mathbf{x}_2 + B_2^T\mathbf{x}_3 = \widetilde{\mathbf{f}}_2, \\
&\Longleftrightarrow \quad C_2^T\mathbf{x}_2 + C_2^{-1}B_2^T\mathbf{x}_3 = C_2^{-1}\widetilde{\mathbf{f}}_2, \\
&\Longleftrightarrow \quad C_2^T\mathbf{x}_2 + D_2\mathbf{x}_3 = \mathbf{g}_2,
\end{aligned} \tag{7}
$$

$$
\textbf{en posant} \left\{ \begin{array}{l} A_2 - D_1^T D_1 = C_2 C_2^T \\ D_2 = C_2^{-1} B_2^T \\ \mathbf{g}_2 = C_2^{-1}(\mathbf{f}_2 - D_1^T \mathbf{g}_1) \end{array} \right. . \tag{8}
$$

$\square$

3. On suppose par la suite que toute les matrices $\widetilde{A}_i$ sont symétriques définies positives.

Montrer que le système $K\mathbf{x} = \mathbf{f}$ est équivalent à

$$
\left\{ \begin{array}{rcl} C_i^T \mathbf{x}_i + D_i \mathbf{x}_{i+1} &=& \mathbf{g}_i, \quad i = 1, \dots, n-1 \\ C_n^T \mathbf{x}_n &=& \mathbf{g}_n \end{array} \right. , \tag{9}
$$

où $C_i \in \mathcal{M}_{p,p}$ est triangulaire inférieure avec des termes diagonaux strictement positifs. On donnera l'équation vérifiée par C_i en fonction de A_i et de D_{i-1} et on donnera l'expression de la matrice D_i et du vecteur $\mathbf{g}_i$ en fonction de C_i, de B_i, de D_{i-1}, de $\mathbf{f}_i$ et de $\mathbf{g}_{i-1}$.

Réponse : on travaille par récurrence. On suppose que pour $i \geq 2$, on a

$$
C_i^T \mathbf{x}_i + D_i \mathbf{x}_{i+1} = \mathbf{g}_i, \qquad \textbf{en posant} \left\{ \begin{array}{l} A_i - D_{i-1}^T D_{i-1} = C_i C_i^T \\ D_i = C_i^{-1} B_i^T \\ \mathbf{g}_i = C_i^{-1}(\mathbf{f}_i - D_{i-1}^T \mathbf{g}_{i-1}) \end{array} \right. . \tag{10}
$$

C'est vrai pour $i = 2$. On montre que c'est encore vrai pour $i+1$. En utilisant $D_i = C_i^{-1}B_i^T \iff B_i = D_i^T C_i^T$ et $C_i^T\mathbf{x}_i = -D_i\mathbf{x}_{i+1} + \mathbf{g}_i$, le bloc $i+1$ s'écrit

$$\begin{aligned}
B_i\mathbf{x}_i + A_{i+1}\mathbf{x}_{i+1} + B_{i+1}^T\mathbf{x}_{i+2} = \mathbf{f}_{i+1} \quad &\iff \quad D_i^T C_i^T\mathbf{x}_i + A_{i+1}\mathbf{x}_{i+1} + B_{i+1}^T\mathbf{x}_{i+2} = \mathbf{f}_{i+1},\\
&\iff \quad (A_{i+1}^T - D_i^T D_i)\mathbf{x}_{i+1} + B_{i+1}^T\mathbf{x}_{i+2} = \mathbf{f}_{i+1} - D_i^T\mathbf{g}_i,\\
&\iff \quad C_{i+1}^T\mathbf{x}_{i+1} + C_{i+1}^{-1}B_{i+1}^T\mathbf{x}_{i+2} = C_{i+1}^{-1}(\mathbf{f}_{i+1} - D_i^T\mathbf{g}_i,),\\
&\iff \quad C_{i+1}^T\mathbf{x}_{i+1} + D_{i+1}\mathbf{x}_{i+2} = \mathbf{g}_{i+1},\\
&\qquad \textbf{en posant} \left\{\begin{array}{l} A_{i+1} - D_i^T D_i = C_{i+1}C_{i+1}^T\\ D_{i+1} = C_{i+1}^{-1}B_{i+1}^T\\ \mathbf{g}_{i+1} = C_{i+1}^{-1}(\mathbf{f}_{i+1} - D_i^T\mathbf{g}_i)\end{array}\right. ,
\end{aligned}$$

en supposant que $\widetilde{A}_{i+1} = A_{i+1} - D_i^T D_i$ est SDP et donc qu'elle admet une factorisation de Cholesky $\widetilde{A}_{i+1} = C_{i+1}C_{i+1}^T$. Ceci achève la récurrence.
Note pour le bloc $i = n$, on fait la même chose, mais avec $B_n = D_n = 0$, soit

$$\begin{aligned}
B_{n-1}\mathbf{x}_{n-1} + A_n\mathbf{x}_n = \mathbf{f}_n \quad &\iff \quad D_{n-1}^T C_{n-1}^T\mathbf{x}_{n-1} + A_n\mathbf{x}_n = \mathbf{f}_n,\\
&\iff \quad (A_n^T - D_{n-1}^T D_{n-1})\mathbf{x}_n = \mathbf{f}_n - D_{n-1}^T\mathbf{g}_{n-1},\\
&\iff \quad C_n^T\mathbf{x}_n = C_n^{-1}(\mathbf{f}_n - D_{n-1}^T\mathbf{g}_{n-1},),\\
&\iff \quad C_n^T\mathbf{x}_n = \mathbf{g}_n,\\
&\qquad \textbf{en posant} \left\{\begin{array}{l} A_n - D_{n-1}^T D_{n-1} = C_n C_n^T\\ (D_n = C_n^{-1}B_n^T = 0)\\ \mathbf{g}_n = C_n^{-1}(\mathbf{f}_n - D_{n-1}^T\mathbf{g}_{n-1})\end{array}\right. .
\end{aligned} \tag{11}$$

□

4. On suppose qu'on dispose des fonctions scilab : `[x]=solsup(U,b)`, `[x]=solinf(L,b)`, `[C]=cholesky(A)`, `[L,U]=LU(A)`, `[B]=inverse(A)` (on ne demande pas de les réécrire ici).

   (a) Écrire une fonction scilab : `[M]=solinfMat(L, N)`, qui étant données une matrice triangulaire inférieure inversible L et une matrice N, calcule $M = L^{-1}N$

   Réponse : on travaille par colonnes de M:

```
=================================
function [M] = solinfMat(L, N)
exec("solinf.sci", 0);
n = size(L, 1);  m = size(N, 2);
if n ~= size(N, 1) then
      error('Incompatible sizes of matrices.');
end
M = zeros(n, m);
for jj = 1:m
      M(: , jj) = solinf(L, N(: , jj) );
end
endfunction
=================================
```

□

   (b) Donner le coût en nombre de multiplications d'une telle fonction.

   Réponse : comme le coût d'une résolution de système triangulaire est de $\frac{n^2}{2}$, si $N \in \mathcal{M}_{n,m}$ alors le coût est $m\frac{n^2}{2}$ (et vaut donc $\frac{n^3}{2}$ si $n = m$). □

5. On suppose qu'on dispose en plus de fonctions scilab :
   `[Ai]=getBlockMat(A, i)`, `[A]=setBlockMat(A, i, Bi)`,
   `[xi]=getBlockVec(x, i)`, `[x]=setBlockVec(x, i, yi)`.
   Étant donné $A = [A_1, A_2, \dots, A_n] \in \mathcal{M}_{p,np}$, la fonction `getBlockMat` extrait de A le bloc $A_i \in \mathcal{M}_{p,p}$ pour $i = 1, \dots, n$. Inversement, la fonction `setBlockMat` insère dans le bloc i de A la matrice $B_i \in \mathcal{M}_{p,p}$ ($A_i \leftarrow B_i$ et tous les autres blocs restent inchangés). La fonction `getBlockVec` extrait du vecteur bloc $\mathbf{x} = [\mathbf{x}_1, \dots, \mathbf{x}_n]^T \in \mathcal{M}_{pn,1}$ le vecteur $\mathbf{x}_i \in \mathcal{M}_{p,1}$. Inversement la fonction `setBlockVec` insère dans le bloc i de $\mathbf{x}$ le vecteur $\mathbf{y}_i \in \mathcal{M}_{p,1}$ ($\mathbf{x}_i \leftarrow \mathbf{y}_i$ et tous les autres blocs restent inchangés).

   En utilisant certaines fonctions scilab disponibles, écrire une fonction scilab : `[x]=resout(A, B, f, n)` qui étant donnés $A = [A_1, A_2, \dots, A_n] \in \mathcal{M}_{p,np}$, $B = [B_1, B_2, \dots, B_{n-1}] \in \mathcal{M}_{p,(n-1)p}$ et $\mathbf{f} \in \mathcal{M}_{np,1}$ résout $K\mathbf{x} = \mathbf{f}$ par la méthode décrite ci-dessus.

Si vous n'avez pas trouvé l'algorithme : à défaut, vous pouvez programmer la fonction : `[x]=resout(C, D, g, n)` qui résout (9), en supposant connus $C = [C_1, C_2, \ldots, C_n] \in \mathcal{M}_{p,np}$, $D = [D_1, D_2, \ldots, D_{n-1}] \in \mathcal{M}_{p,(n-1)p}$ et $\mathbf{g} \in \mathcal{M}_{np,1}$.

Réponse : il faut implémenter l'algorithme donné par (2), (3), (10) et (11). Cet algorithme est une adaptation de celui de Richtmayer pour les matrices tridiagonales par blocs et symétriques définies positives, qui exploite la forme particulière de K. Voici une implémentation possible.

```
==================================
function [x]=resout(A, B, f, n)
exec("cholesky.sci", 0);   exec("solinf.sci", 0);   exec("solinfMat.sci", 0);   exec("solsup.sci", 0);
p = size(A, 1);

if p*n ~= size(A, 2) then
      error('Incompatible sizes of matrices.');
end

C = zeros(p, n*p); // Stocke dans C : blocs C_1 à C_n
D = zeros(p, (n-1)*p); // Stocke dans D : blocs D_1 à D_{n-1}

x = zeros(n*p, 1);   g = zeros(n*p, 1);

// tmp matrices and vectors (to ease the use):
Ai = zeros(p, 1); Bi = zeros(p, 1);   Ci = zeros(p, 1);   Di = zeros(p, 1);
xi = zeros(p, 1);   gi = zeros(n, 1);

// bloc 1
// on extrait le bloc, on fait le calcul puis on stocke :
Ai = getBlockMat( A, 1 );   Ci = cholesky( Ai );   C = setBlockMat( C, 1, Ci );
fi = getBlockVec( f, 1 );   gi = solinf( Ci, fi );   g = setBlockVec( g, 1, gi );
Bi = getBlockMat( B, 1 );   Di = solinfMat( Ci, Bi' );   D = setBlockMat( D, 1, Di );

for ii = 2:n
      // à ce stade, "Di" vaut D_{i-1} et "gi" vaut g_{i-1} :
      Ai = getBlockMat( A, ii );   Ai = Ai - Di' * Di;   Ci = cholesky( Ai );   C = setBlockMat( C, ii, Ci );
      fi = getBlockVec( f, ii );   fi = fi - Di' * gi;   gi = solinf( Ci, fi );   g = setBlockVec( g, ii, gi );
      if ii < n then  // les blocs B_n et D_n n'existent pas et sont inutiles.
            Bi = getBlockMat( B, ii );   Di = solinfMat( Ci, Bi' );   D = setBlockMat( D, ii, Di );
      end
end

// bloc n : résolution de C_n^T x_n = g_n
// à ce stade, "gi" vaut g_n et "Ci" vaut C_n :
xi = solsup( Ci' , gi );   x = setBlockVec( x, n, xi );

for ii = n-1:-1:1
      // Résolution de C_i^T x_i = -D_i x_{i+1} + g_i :
      // à ce stade, "xi" vaut x_{i+1} :
      Ci = getBlockMat( C, ii );   Di = getBlockMat( D, ii );   gi = getBlockVec( g, ii );
      gi = gi - Di * xi ;   xi = solsup( Ci' , gi );   x = setBlockVec( x, ii, xi );
end
endfunction
==================================
```

La remontée (dernière boucle et initialisation au bloc n) était programmable sans avoir trouvé l'algorithme.

□

6. Calculer le nombre de multiplications de la fonction que vous avez programmée.

Réponse : Dans la première boucle, il y a :

- **n appels à `solinf` : coût $n\frac{p^2}{2}$ (négligeable),**
- **$n-1$ appels à `solinfMat` (matrices de taille p) : coût $n\frac{p^3}{2}$ (on néglige les termes d'ordre p^3),**

- n **appels à** `cholesky` **(matrices de taille** p**) : coût** $n\frac{p^3}{6}$**,**
- n **multiplications matrice** $\times$ **matrice (**$D_{i-1}^T D_{i-1}$**, matrices de taille** p**) : coût** np^3**,**
- n **multiplications matrice** $\times$ **vecteur (**$D_{i-1}^T \mathbf{g}_{i-1}$**, matrice de taille** p**) : coût** np^2 **(négligeable),**

et dans la deuxième boucle, il y a :

- n **appels à** `solsup` **: coût** $n\frac{p^2}{2}$ **(négligeable),**
- $n-1$ **multiplications matrice** $\times$ **vecteur (**$D_i \mathbf{x}_i$**, matrice de taille** p**) : coût** np^2 **(négligeable).**

Donc au final, le coût de l'algorithme en négligeant les termes d'ordre np^2 **et** p^3**, est**

$$n\frac{p^3}{2} + n\frac{p^3}{6} + np^3 = \frac{5np^3}{3}.$$

Par comparaison, l'algorithme de Cholesky sur K **tout entier sans réflechir aurait coûté** $\frac{(np)^3}{6} = \frac{n^3p^3}{6}$**, soit un facteur** n^2 **plus cher.** $\square$

Exercice 3 : *(barème approximatif : 2 points)*

RÉDIGER SUR LA MÊME COPIE QUE L'EXERCICE 2

Soit un réel $\varepsilon > 0$. On étudie le système

$$\begin{bmatrix} \varepsilon & 1 \\ 1 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1-\varepsilon \\ 6 \end{bmatrix}.$$

1. Effectuer l'élimination de Gauss en arithmétique exacte.

   Réponse : le déterminant de la matrice vaut $7\varepsilon - 1$**, donc la matrice est inversible si** $\varepsilon < 1/7$**, ce qu'on suppose.**

$$\begin{aligned} Ax = b \quad & \Longleftrightarrow \quad \left\{ \begin{array}{rcl} \varepsilon x_1 + x_2 & = & 1-\varepsilon \\ (7-\frac{1}{\varepsilon})x_2 & = & 6 - \frac{1-\varepsilon}{\varepsilon} \end{array} \right. \\ & \Longleftrightarrow \quad \left\{ \begin{array}{rcl} \varepsilon x_1 + x_2 & = & 1-\varepsilon \\ (7-\frac{1}{\varepsilon})x_2 & = & 7 - \frac{1}{\varepsilon} \end{array} \right. \\ & \Longleftrightarrow \quad x = \begin{bmatrix} -1 \\ 1 \end{bmatrix}. \end{aligned}$$

$\square$

2. On suppose que l'on travaille en flottant en base 10 avec 3 chiffres significatifs et on prend $\varepsilon = 10^{-4}$. Refaire les calculs en arithmétique flottante.

   Réponse : on fait les calculs flottants :

$$\begin{aligned} 1 \ominus \varepsilon &= \mathrm{fl}(1.00 - 0.0001) = \mathrm{fl}(0.9999) = 1.00 \\ 7 \ominus \frac{1}{\varepsilon} &= \mathrm{fl}(7.00 - 10000) = \mathrm{fl}(-9993) = -9.99 \times 10^3 \\ 6 \ominus \left(\frac{1 \ominus \varepsilon}{\varepsilon}\right) &= \mathrm{fl}(6.00 - \frac{1.00}{\varepsilon}) = \mathrm{fl}(6.00 - 10000) = \mathrm{fl}(-9994) = -9.99 \times 10^3, \end{aligned}$$

donc on trouve en arithmétique flottante

$$\begin{aligned} \widetilde{A}\widetilde{x} = \widetilde{b} \quad & \Longleftrightarrow \quad \left\{ \begin{array}{rcl} \varepsilon \widetilde{x}_1 + \widetilde{x}_2 & = & 1.00 \\ -9.99 \times 10^3 \widetilde{x}_2 & = & -9.99 \times 10^3 \end{array} \right. \\ & \Longleftrightarrow \quad \widetilde{x} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \end{aligned}$$

Il y a une erreur catastrophique. $\square$