Université de Marseille
Licence de Mathématiques, 3eme année, analyse numérique et optimisation
Examen du 11 janvier 2017

L'examen contient 3 exercices. Le barème est sur 28 points, il n'est donc pas demandé de tout faire pour avoir 20.

Exercice 1 (Résolution d'un système sous forme particulière, barème 10 points)**.**
Soit $n \geq 1$, $p \geq 1$, $A \in \mathcal{M}_n(\mathbb{R})$ et $B \in \mathcal{M}_{n,p}(\mathbb{R})$.
On suppose que A est une matrice symétrique définie positive et que $\mathrm{rang}(B) = p$ (ceci implique que $p \leq n$).
Pour $i \in \{1, \ldots, p\}$, on pose $z_i = A^{-1}c_i(B)$ où $c_i(B)$ désigne la i-ieme colonne de B.

1. Montrer que $\{c_i(B), i \in \{1, \ldots, p\}\}$ est une base de $\mathrm{Im}(B)$.
   En déduire que $\{z_i, i \in \{1, \ldots, p\}\}$ est une famille libre de R^n.

   Corrigé – $\mathrm{Im}(B) = \mathrm{vect}\{c_i(B), i = 1, \ldots, p\}$. *Comme* $\mathrm{rang}(B) = \dim(\mathrm{Im}(B)) = p$*, la famille* $\{c_i(B), i = 1, \ldots, p\}$ *est une base de* $\mathrm{Im}(B)$ *et c'est donc une famille libre.*
   On en déduit que $\{z_i, i \in \{1, \ldots, p\}\}$ *est une famille libre. En effet, Soient* $\alpha_1, \ldots, \alpha_p \in \mathbb{R}$ *tels que* $\sum_{i=1}^p \alpha_i z_i = 0$*, on a donc* $A(\sum_{i=1}^p \alpha_i z_i) = \sum_{i=1}^p \alpha_i A z_i = \sum_{i=1}^p \alpha_i c_i(B) = 0$ *et donc* $\alpha_i = 0$ *pour tout* $i \in \{1, \ldots, p\}$ *(car la famille* $\{c_i(B), i = 1, \ldots, p\}$ *est libre).*

2. Montrer que $\{B^t z_i, i \in \{1, \ldots, p\}\}$ est une base de R^p.

   Corrigé – Soient $\alpha_1, \ldots, \alpha_p \in \mathbb{R}$ *tels que* $\sum_{i=1}^p \alpha_i B^t z_i = 0$*, c'est-à-dire* $B^t A^{-1}(\sum_{i=1}^p \alpha_i c_i(B)) = 0$*. En notant* $e_1, \ldots, e_p$ *la base canonique de* $\mathbb{R}^p$ *on a* $c_i(B) = Be_i$ *et donc*
   $$0 = B^t A^{-1}(\sum_{i=1}^p \alpha_i B e_i) \cdot (\sum_{i=}^p \alpha_i e_i) = A^{-1}(\sum_{i=1}^p \alpha_i B e_i) \cdot (\sum_{i=1}^p \alpha_i B e_i).$$
   Comme A^{-1} *est s.d.p., ceci donne* $B(\sum_{i=1}^p \alpha_i e_i) = \sum_{i=1}^p \alpha_i B e_i = 0$.
   Or, le théorème du rang donne $p = \mathrm{rang}(B) + \dim(\mathrm{Ker}(B))$ *et donc* $\mathrm{Ker}(B) = \{0\}$ *(car* $\mathrm{rang}(B) = p$*). On a donc* $\sum_{i=1}^p \alpha_i e_i = 0$*. On en déduit que* $\alpha_i = 0$ *pour tout* $i \in \{1, \ldots, p\}$ *et donc que* $\{B^t z_i, i \in \{1, \ldots, p\}\}$ *est une famille libre de* R^p*, elle a p éléments, c'est donc une base de* $\mathbb{R}^p$.

Soient $b \in \mathbb{R}^n$ et $c \in \mathbb{R}^p$. On cherche le couple (x, y), avec $x \in \mathbb{R}^n$ et $y \in \mathbb{R}^p$, solution du système suivant (écrit sous forme de blocs) :
$$\begin{bmatrix} A & B \\ B^t & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} b \\ c \end{bmatrix}. \tag{1}$$
On pose $z = A^{-1}b$ et on note $y_1, \ldots, y_p$ les composantes de y.

3. Montrer que (x, y) est solution de (1) si et seulement si
   $$\sum_{i=1}^p y_i B^t z_i = B^t z - c, \tag{2}$$
   $$x = z - \sum_{i=1}^p y_i z_i. \tag{3}$$
   En déduire que le système (1) a une unique solution.

   Corrigé – (x, y) *solution de (1) signifie que* $Ax + By = b$ *et* $B^t x = c$.
   Comme A *est inversible et que* $By = \sum_{i=1}^p y_i c_i(B)$*, l'équation* $Ax + By = b$ *est équivalente à* $x = A^{-1}b - \sum_{i=1}^p y_i A^{-1}(c_i(B))$ *et donc équivalente à* $x = z - \sum_{i=1}^p y_i z_i$.
   (x, y) *est donc solution de (1) si et seulement si* $x = z - \sum_{i=1}^p y_i z_i$ *et* $B^t(z - \sum_{i=1}^p y_i z_i) = c$*. ce qui est bien équivalent à (2)-(3).*

L'équation (2) est une équation dans $\mathbb{R}^p$. Comme $\{B^t z_i,\, i \in \{1,\ldots,p\}\}$ est une base de R^p, cette équation admet une et une seule solution. On obtient ensuite x de manière unique avec l'équation (3). Il existe donc une et une seule solution au système (2)-(3), ce qui prouve bien que le système (1) a une et une seule solution.

NB : Une manière de résoudre le système (1) (de $(n+p)$ équations à $(n+p)$ inconnues) consiste donc à calculer z, $z_1,\ldots,z_p$ (avec une factorisation de A) puis à résoudre (2) (qui est un système de p équations à p inconnues) et enfin à calculer x.

4. Montrer que la matrice (symétrique) $\begin{bmatrix} A & B \\ B^t & 0 \end{bmatrix}$ est inversible mais n'est pas symétrique définie positive.

   Corrigé – La question précédente montre que la matrice $\begin{bmatrix} A & B \\ B^t & 0 \end{bmatrix}$ est inversible. Elle est symétrique. Elle n'est pas s.d.p. car sa diagonale contient des termes nuls.

5. Soit $C \in \mathcal{M}_{n,p}(\mathbb{R})$.

   Donner une condition sur C (ne dépendant ni de A ni de B) impliquant que la matrice $\begin{bmatrix} A & B \\ C & 0 \end{bmatrix}$ est inversible.

   Corrigé – On peut refaire les 3 premières questions de l'exercice en remplaçant B^t par C pourvu que $\{Cz_i,\, i \in \{1,\ldots,p\}\}$ soit une base de R^p.

   Une condition suffisante pour cela est, par exemple, que $\mathrm{Ker}(C) = \{0\}$. Cette condition est donc suffisante pour montrer que la matrice $\begin{bmatrix} A & B \\ C & 0 \end{bmatrix}$ est inversible. Elle a l'inconvénient d'imposer $n \le p$ (et donc $n = p$ car l'hypothèse sur B impose $p \le n$).

Exercice 2 (Méthode de Newton pour calculer une valeur propre, barème 6 points)**.**
Soit $A \in \mathcal{M}_n(\mathbb{R})$ $(n \ge 1)$ une matrice symétrique. Soient $\overline{\lambda}$ une valeur propre simple de A et $a \in \mathbb{R}^n$ tel que $a \not\perp \mathrm{Ker}(A - \overline{\lambda} I)$ (où I est la matrice “identité”).

1. Montrer qu'il existe un unique $\overline{u} \in \mathrm{Ker}(A - \bar{\lambda} I)$ tel que $\overline{u} \cdot a = 1$.

   Corrigé – $\bar{\lambda}$ est une valeur propre simple de A, il existe donc $u \ne 0$ tel que $\mathrm{Ker}(A - \bar{\lambda} I) = \mathbb{R}u$. Comme $a \not\perp \mathrm{Ker}(A - \bar{\lambda} I)$, on a $u \cdot a \ne 0$ et $\overline{u} = u/(u \cdot a)$.

Dans la suite, on cherche à calculer le couple $(\overline{u}, \overline{\lambda})$, $\overline{u}$ donné par la question 1, avec l'algorithme de Newton. On définit F de $\mathbb{R}^{n+1}$ dans $\mathbb{R}^{n+1}$ par $F(\begin{bmatrix} u \\ \lambda \end{bmatrix}) = \begin{bmatrix} Au - \lambda u \\ u \cdot a - 1 \end{bmatrix}$ (pour $u \in \mathbb{R}^n$ et $\lambda \in \mathbb{R}$).

2. Pour tout $\begin{bmatrix} u \\ \lambda \end{bmatrix} \in \mathbb{R}^{n+1}$, calculer la matrice jacobienne de F au point $\begin{bmatrix} u \\ \lambda \end{bmatrix}$.

   Corrigé – La matrice jacobienne de F au point $\begin{bmatrix} u \\ \lambda \end{bmatrix}$ appartient à $\mathcal{M}_{n+1}(\mathbb{R})$, on la note $J_F(u,\lambda)$ et un calcul simple donne (sous forme de blocs) $J_F(u,\lambda) = \begin{bmatrix} A - \lambda I & -u \\ a^t & 0 \end{bmatrix}$.

3. Ecrire l'algorithme de Newton pour trouver les points où F s'annule.

   Corrigé – L'algorithme de Newton pour F s'écrit, avec la matrice jacobienne trouvée à la question précédente :
   Initialisation : *$u^{(0)} \in \mathbb{R}^n$, $\lambda_0 \in \mathbb{R}$,*
   Itérations : *pour tout $k \ge 0$, $J_F(u^{(k)}, \lambda_k) \begin{bmatrix} x^{(k+1)} - x^{(k)} \\ \lambda_{k+1} - \lambda_k \end{bmatrix} = -\begin{bmatrix} Au^{(k)} - \lambda_k u^{(k)} \\ u^{(k)} \cdot a - 1 \end{bmatrix}$.*

4. On note $(u^{(0)}, \lambda_0)$ l'initialisation de l'algorithme de Newton appliqué à la recherche d'un point où F s'annule. Monter qu'il existe $\varepsilon > 0$ tel que, si $|u^{(0)} - \overline{u}| \le \varepsilon$ et $|\lambda_0 - \overline{\lambda}| \le \varepsilon$, la suite donnée par l'algorithme de Newton existe et converge vers $(\overline{u}, \overline{\lambda})$.

   Corrigé – *La fonction F est de classe C^2 (et même de classe C^∞). Pour répondre à la question, il suffit de montrer que la matrice $J_F(\overline{u}, \overline{\lambda})$ est inversible. Soit donc $v \in \mathbb{R}^n$ et $\mu \in \mathbb{R}$ tels que $J_F(\overline{u}, \overline{\lambda}) \begin{bmatrix} v \\ \mu \end{bmatrix} = 0$. Il s'agit de montrer que $v = 0$ et $\mu = 0$.*

   Pour cela on remarque que $Av - \overline{\lambda} v - \mu \overline{u} = 0$ et $a \cdot v = 0$. En prenant le produit scalaire de la première équation par $\overline{u}$, on a donc

   $$Av \cdot \overline{u} - \overline{\lambda} v \cdot \overline{u} - \mu \overline{u} \cdot \overline{u} = 0.$$

   Comme A est symétrique et $\overline{u} \in \mathrm{Ker}(A - \overline{\lambda} I)$, on a $Av \cdot \overline{u} - \overline{\lambda} v \cdot \overline{u} = A\overline{u} \cdot v - \overline{\lambda}\overline{u} \cdot v = 0$ et donc $\mu \overline{u} \cdot \overline{u} = 0$ Comme $\overline{u} \neq 0$, on a donc $\mu = 0$. On en déduit que $Av - \overline{\lambda} v = 0$ et donc qu'il existe $\beta \in \mathbb{R}$ tel que $v = \beta \overline{u}$ (car $\overline{\lambda}$ est une valeur propre simple). Enfin, comme $a \cdot v = 0$, on a $\beta = a \cdot \beta \overline{u} = a \cdot v = 0$ et donc $v = 0$.

Exercice 3 (Gauss-Seidel et optimisation, barème 12 points)**.** Soit A une matrice carrée d'ordre n, symétrique définie positive, et $b \in \mathbb{R}^n$, $n > 1$. On pose $\bar{x} = A^{-1} b$. On note D la partie diagonale de A, $-E$ la partie triangulaire inférieure stricte de A et $-F$ la partie triangulaire supérieure stricte de A, c'est-à-dire :

$$\begin{aligned}
&D = (d_{i,j})_{i,j=1,\dots,n},\ d_{i,j} = 0 \text{ si } i \neq j,\ d_{i,i} = a_{i,i},\\
&E = (e_{i,j})_{i,j=1,\dots,n},\ e_{i,j} = 0 \text{ si } i \le j,\ e_{i,j} = -a_{i,j} \text{ si } i > j,\\
&F = (f_{i,j})_{i,j=1,\dots,n},\ f_{i,j} = 0 \text{ si } i \ge j,\ f_{i,j} = -a_{i,j} \text{ si } i < j.
\end{aligned}$$

Dans toute la suite, on considère la fonction f de $\mathbb{R}^n$ dans $\mathbb{R}$ définie par $f(x) = \frac{1}{2} Ax \cdot x - b \cdot x$.

1. Montrer que la méthode de Gauss-Seidel pour calculer la solution du système $Ax = b$ peut s'écrire sous la forme suivante :
   Initialisation : $x^{(0)} \in \mathbb{R}^n$,
   Itérations : pour tout $k \ge 0$, $x^{(k+1)} = x^{(k)} + w^{(k)}$, avec $w^{(k)} = (D - E)^{-1} r^{(k)}$, $r^{(k)} = b - Ax^{(k)}$.

   Corrigé –
   Les itérations de Gauss-Seidel sécrivent $(D - E)x^{(k+1)} = Fx^{(k)} + b$, c'est-à-dire , comme $D - E - F = A$,

   $$\begin{aligned}
   x^{(k+1)} &= (D - E)^{-1} F x^{(k)} + (D - E)^{-1} b = x^{(k)} + (D - E)^{-1}(Fx^{(k)} - (D - E)x^{(k)}) + (D - E)^{-1} b\\
   &= x^{(k)} - (D - E)^{-1}(Ax^{(k)}) + (D - E)^{-1} b = x^{(k)} + (D - E)^{-1}(b - Ax^{(k)}).
   \end{aligned}$$

2. Dans cette question, on étudie la méthode de Gauss-Seidel écrite sous la forme donnée à la question 1.
   (a) Soit $k \ge 0$. On suppose que $w^{(k)} = 0$. Montrer que $x^{(p)} = \bar{x}$ pour tout $p \ge k$.

   Corrigé – *Si $w^{(k)} = 0$, on a $r^{(k)} = (D - E)w^{(k)} = 0$ et donc $x^{(k)} = \bar{x}$. Par récurrence, on a alors $x^{(p)} = \bar{x}$ pour tout $p \ge k$.*

   (b) Soit $k \ge 0$. On suppose que $w^{(k)} \neq 0$. Montrer que $2r^{(k)} \cdot w^{(k)} = Aw^{(k)} \cdot w^{(k)} + Dw^{(k)} \cdot w^{(k)}$. [On pourra utiliser le fait que $F = E^t$.]

   Corrigé – *Comme $F = E^t$, on a $Ew^{(k)} \cdot w^{(k)} = w^{(k)} \cdot Fw^{(k)} = Fw^{(k)} \cdot w^{(k)}$ et donc*
   $$2r^{(k)} \cdot w^{(k)} = 2(D - E)w^{(k)} \cdot w^{(k)} = (2D - E - F)w^{(k)} \cdot w^{(k)} = Aw^{(k)} \cdot w^{(k)} + Dw^{(k)} \cdot w^{(k)}.$$

   (c) Déduire des questions (2a) et (2b) que la méthode de Gauss-Seidel est une méthode de descente pour la minimisation de f et que $w^{(k)}$ est une direction de descente stricte en $x^{(k)}$ dès que $w^{(k)} \neq 0$.

Corrigé – *Soit $k \geq 0$.*
Si $w^{(k)} = 0$, on obtient, à partir de l'itération k, la solution recherchée.
Si $w^{(k)} \neq 0$, comme A est s.d.p. la matrice D est aussi s.d.p (car $a_{i,i} > 0$ pour tout i), on a donc $r^{(k)} \cdot w^{(k)} > 0$ (par la question précédente). Comme $r^{(k)} = -\nabla f(x^{(k)}$, ceci montre que $w^{(k)}$ est alors une direction de descente stricte en $x^{(k)}$.
Ceci montre bien que la méthode de Gauss-Seidel est une méthode de descente. La direction de descente est $w^{(k)}$ et le pas dans cette direction est 1.

3. On cherche maintenant à améliorer la méthode de Gauss-Seidel en prenant non plus un pas fixe dans l'algorithme de descente de la question 1, mais un pas optimal qui est défini à l'itération k par

$$f(x^{(k)} + \alpha_k w^{(k)}) = \min_{\alpha > 0} f(x^{(k)} + \alpha w^{(k)}), \tag{4}$$

où $w^{(k)}$ est défini à la question 1. On définit alors une méthode de descente à pas optimal par :

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

On appelle cette nouvelle méthode "méthode de Gauss-Seidel à pas optimal".

(a) Soit $k \geq 0$ tel que $w^{(k)} \neq 0$. Justifier l'existence et l'unicité du pas optimal défini par (4) et donner son expression.

Corrigé – *On pose $\varphi(\alpha) = f(x^{(k)} + \alpha w^{(k)})$. L'existence de α_k découle du fait que $\varphi'(0) < 0$ et $\lim_{\alpha \to +\infty} \varphi(\alpha) = +\infty$. L'unicité de α_k découle du fait que φ est strictement convexe (car $\varphi''(\alpha) = Aw^{(k)}.w^{(k)} > 0$).*
Le nombre α_k est caractérisé par le fait que $\nabla f(x^{(k)} + \alpha_k w^{(k)}) \cdot w^{(k)} = 0$, c'est-à-dire $(A(x^{(k)} + \alpha_k w^{(k)}) - b) \cdot w^{(k)} = 0$. Ceci donne $(\alpha_k A w^{(k)} - r^{(k)}) \cdot w^{(k)} = 0$ et donc

$$\alpha_k = \frac{r^{(k)} \cdot w^{(k)}}{Aw^{(k)} \cdot w^{(k)}}.$$

(b) Soit $k \geq 0$ tel que $w^{(k)} \neq 0$. Montrer que $f(x^{(k)}) - f(x^{(k+1)}) = \dfrac{|r^{(k)} \cdot w^{(k)}|^2}{2Aw^{(k)} \cdot w^{(k)}}$.

Corrigé – *En utilisant la symétrie de A, on a*

$$f(x^{(k+1)}) = f(x^{(k)}) + \alpha_k Ax^{(k)} \cdot w^{(k)} + \frac{\alpha_k^2}{2} Aw^{(k)} \cdot w^{(k)} - \alpha_k b \cdot w^{(k)} = f(x^{(k)}) - \alpha_k r^{(k)} \cdot w^{(k)} + \frac{\alpha_k^2}{2} Aw^{(k)} \cdot w^{(k)}.$$

En remplaçant α_k par sa valeur, on obtient bien $f(x^{k+1}) = f(x^{(k)}) - \dfrac{|r^{(k)} \cdot w^{(k)}|^2}{2Aw^{(k)} \cdot w^{(k)}}$, ce qui donne la formule demandée.

(c) Montrer que $r^{(k)} \to 0$ lorsque $k \to +\infty$, et en déduire que la suite donnée par la méthode de Gauss-Seidel à pas optimal converge vers la solution $\bar{x}$ du système linéaire $Ax = b$.

Corrigé – *Il suffit de considérer le cas où $w^{(k)} \neq 0$ pour tout k. On utilise tout d'abord la question 3b. Comme f est bornée inférieurement et que $f(x^{(k+1)}) < f(x^{(k)})$ pour tout k, la suite $(f(x^{(k)}))_{k \in \mathbb{N}}$ est convergente (dans $\mathbb{R}$). On a donc $\lim_{k \to +\infty} \dfrac{|r^{(k)} \cdot w^{(k)}|^2}{2Aw^{(k)} \cdot w^{(k)}} = 0$.*
On utilise maintenant la question 2b. Elle donne $2r^{(k)} \cdot w^{(k)} = Aw^{(k)} \cdot w^{(k)} + Dw^{(k)} \cdot w^{(k)} > Aw^{(k)} \cdot w^{(k)}$ (car D est s.d.p) et donc $|r^{(k)} \cdot w^{(k)}|^2 > (1/4)(Aw^{(k)} \cdot w^{(k)})^2$. En notant λ_1 la plus petite valeur propre de A et λ_n la plus grande valeur propre de A, on a donc

$$\frac{1}{8} \frac{\lambda_1^2}{\lambda_n} |w^{(k)}|^2 = \frac{1}{8} \frac{\lambda_1^2}{\lambda_n} \frac{|w^{(k)}|^4}{|w^{(k)}|^2} \leq \frac{|r^{(k)} \cdot w^{(k)}|^2}{2Aw^{(k)} \cdot w^{(k)}}.$$

On en déduit que $\lim_{k \to +\infty} w^{(k)} = 0$ et donc aussi $\lim_{k \to +\infty} r^{(k)} = 0$ (car $r^{(k)} = (D - E)w^{(k)}$). Ce qui donne bien $\lim_{k \to +\infty} x^{(k)} = \overline{x}$.