MINES-PONTS 2003. Filière MP. MATHÉMATIQUES 2.

Corrigé de JL. Lamard (jean-louis.lamard@prepas.org)

Dans la suite on notera $\overrightarrow{x}$ le vecteur de $\mathbb{R}^n$ dont les composantes sont la matrice colonne x dans la base canonique.

1 Soit u l'endomorphisme canoniquement attaché à M. Comme M est symétrique réelle, u est orthodiagonalisable dans une base notée $\mathcal{B}$ en la matrice $\mathrm{Diag}(\lambda_1,\dots,\lambda_n)$. Notons (y_i) les composantes de $\overrightarrow{x}$ sur la base $\mathcal{B}$. Il vient alors $\big(Mx|x\big) = \big(u(\overrightarrow{x})|\overrightarrow{x}\big) = \sum\limits_{i=1}^{n} \lambda_i y_i^2$. Or $\sum\limits_{i=1}^{n} y_i^2 = ||\overrightarrow{x}||^2 = ||x||^2$.
Il en découle que, pour tout vecteur x, $p||x||^2 \leqslant (Mx|x) \leqslant q||x||^2$ avec $p = \inf\limits_{1\leqslant i\leqslant n} \lambda_i$ et $q = \sup\limits_{1\leqslant i\leqslant n} \lambda_i$. $\square$

Remarquons que cette égalité est la meilleure possible d'une manière générale car les deux égalités sont atteintes en prenant pour x un vecteur propre relatif à la plus petite ou à la plus grande des valeurs propres.

2 Avec les notations précédentes, on a vu que $(Mx|x) = \sum\limits_{i=1}^{n} \lambda_i y_i^2$. Il en découle immédiatement que M est définie positive c'est à dire positive inversible si et seulement si $\lambda_i > 0$ pour tout i. $\square$

3 Toujours avec les notations précédentes en supposant pour fixer les idées que $|\lambda_n| = \max|\lambda_i|$, si x est un vecteur unitaire, il vient que $||Mx||^2 = \sum\limits_{i=1}^{n}(\lambda_i y_i)^2 \leqslant \lambda_n^2 \sum\limits_{i=1}^{n} y_i^2 = \lambda_n^2$ donc $N(M) \leqslant |\lambda_n|$ par définition même du sup.
En outre si x est le vecteur unitaire de composantes toutes nulles sur la base $\mathcal{B}$ sauf la dernière égale à 1, il vient que $||Mx||^2 = \lambda_n^2$ donc $N(M) \geqslant |\lambda_n|$.
En conclusion, $N(M) = \sup\limits_{1\leqslant i\leqslant n} |\lambda_i|$. $\square$

4 Il convient dans cette question, afin de pouvoir parler ***du*** vecteur z solution de $Ax = b$, de supposer A en outre inversible c'est à dire définie positive c'est à dire $0 < \lambda_1 \leqslant \ldots \leqslant \lambda_n$.
Notons $S = I - \alpha A$. C'est une matrice évidemment symétrique dont les valeurs propres sont les $1 - \alpha\lambda_i$.
Or, comme $\lambda_i > 0$, il vient que $0 < \alpha\lambda_i < 2\dfrac{\lambda_i}{\lambda_n} \leqslant 2$ donc $-1 \leqslant 1 - 2\dfrac{\lambda_i}{\lambda_n} < 1 - \alpha\lambda_i < 1$.
Ainsi toutes les valeurs propres de S sont dans l'intervalle ouvert $]-1,1[$ de sorte que $N(S) < 1$ d'après la question précédente.
En notant z la solution de $Ax = b$, en remarquant que la suite (x^k) est définie par l'itération affine $x^{k+1} = Sx^k + \alpha b$ dont z est solution, il vient que $x^{k+1} - z = S(x^k - z)$ donc $||x^{k+1} - z|| \leqslant N(S)||x^k - z|| \leqslant \cdots \leqslant N(S)^{k+1}||x^0 - z||$ ce qui prouve bien que la suite (x^k) converge vers z puisque $N(S) < 1$. $\square$

5 Un développement immédiat, joint au fait que A est symétrique et donc que $(Ax|u) = (Au|x)$, fournit :
$f(x+u) - f(x) = \dfrac{1}{2}(Au|u) - (b|u) + (Ax|u) = f(u) + (Ax|u)$. $\square$

6 Il est immédiat que $f(x)$ est une expression polynomiale de degré 2 des composantes canoniques de x et donc que f est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}^n$. $\square$
Pour le calcul notons que $f(x + he_k) - f(x) = \dfrac{1}{2}(Ahe_k|he_k) - (b|he_k) + (Ax|he_k)$.
Donc $\dfrac{1}{h}\left(f(x + he_k) - f(x)\right) = \dfrac{h}{2}(Ae_k|e_k) - (b|e_k) + (Ax|e_k)$.
Il en découle que $\dfrac{\partial f}{\partial x_k}(x) = (Ax - b|e_k)$. $\square$

7 La base canonique étant orthonormée pour le produit scalaire canonique (!), il résulte immédiatement de la question précédente que $g(x) = Ax - b$. $\square$

8 Un développement immédiat montre, compte tenu des questions 5 et 6, que $I(x,u) = \dfrac{1}{2}(Au|u)$ qui ne dépend donc pas de x.
D'après la première question, $r||u||^2 \leqslant I(x,u) \leqslant s||u||^2$ avec $s = \dfrac{1}{2}\lambda_1 = \dfrac{1}{2}\min\limits_{1\leqslant i\leqslant n} \lambda_i$ et $s = \dfrac{1}{2}\lambda_n = \dfrac{1}{2}\max\limits_{1\leqslant i\leqslant n} \lambda_i$. $\square$

9 Si f admet un minimum en z alors z est un point critique puisque les dérivées partielles existent. Il en découle que $g(x) = 0$ c'est à dire que $Az = b$.
Réciproquement si $Az = b$ alors alors $I(z,u) = f(z+u) - f(z)$ puisque $g(z) = 0$. De la question précédente on en déduit que $f(z+u) - f(z) \geqslant \dfrac{\lambda_1}{2}$ et donc que f admet un minimum en z puisque $\lambda_1 > 0$ car A est définie positive.
En conclusion, f admet un unique minimum en la solution z du système $Az = b$. $\square$

10 Nous avons $f(x+u)-f(x)=I(x,u)+(g(x)|u)$ donc :
$f(x-\alpha g(x))-f(x)=I(x,-\alpha g(x))-\alpha(g(x)|g(x))\leqslant\dfrac{\lambda_n}{2}\alpha^2||g(x)||^2-\alpha||g(x)||^2=\alpha\Big(\dfrac{\lambda_n}{2}\alpha-1\Big)||g(x)||^2\leqslant 0$ compte-tenu du fait que $0<\alpha<\dfrac{2}{\lambda_n}$.
Ainsi $f\big(x-\alpha g(x)\big)\leqslant f(x)$ pour tout vecteur x dès lors que $0<\alpha<\dfrac{2}{\lambda_n}$. $\square$

11 Partant d'une semence x^0 quelconque dans $\mathbb{R}^n$, le résultat précédent nous incite à envisager la suite itérée définie par $x^{k+1}=x^k-\alpha g(x^k)$. c'est à dire par $x^{k+1}=x^k+\alpha\big(b-Ax^k\big)$ compte-tenu de la question 6.
Ainsi cette "nouvelle" méthode n'est autre que celle définie dans la question 4 qui converge on le sait vers la solution z du système $Ax=b$! $\square$

12 On a $f(x)=\dfrac{1}{2}(Ax|x)-(b|x)\geqslant\dfrac{1}{2}(Ax|x)-\big|(b|x)\big|\geqslant\dfrac{\lambda_1}{2}||x||^2-||b||.||x||$ d'après la question 1 et l'inégalité de Cauchy-Schwarz. Il en résulte que $\lim\limits_{||x||\to+\infty}f(x)=+\infty$ puisque $\lambda_1>0$. D'où la propriété demandée. $\square$

13 Il suffit d'appliquer la question précédente avec $c=f(y)$. $\square$

14 Soit y_0 fixé quelconque dans F. D'après la question précédente, il existe $r>0$ tel que pour $x\in F$ et $||x||\geqslant r$ on ait $f(x)\geqslant f(y_0)$. Il en découle que $\inf\limits_{x\in F}f(x)=\inf\limits_{x\in K}f(x)$ où $K=F\bigcap\overline{B}(0,r)$. Or K est compact (fermé borné en dimension finie) donc cet inf est atteint puisque f étant de classe $\mathcal{C}^\infty$ comme noté précédemment est a fortiori continue. $\square$

15 Montrons que f est strictement convexe :
Dans une base orthonormée de vecteurs propres de A, $f(x)$ s'écrit $\sum\limits_{i=1}^{n}\dfrac{\lambda_i}{2}x_i^2-b_ix_i$. Pour prouver que f set strictement convexe, il suffit clairement de prouver que les fonctions $\varphi_i(t)=\dfrac{\lambda_i}{2}t^2-b_it$ sont strictement convexes sur $\mathbb{R}$ ce qui est clair car $\varphi_i''(t)=\lambda_i>0$. $\square$

- Montrons que f atteint son minimum sur F en un seul point :
En effet en supposant qu'il soit atteint en x_1 et x_2 différents, on aurait par convexité stricte $f(x)<f(x_1)$ pour tout x appartenant au segment ouvert $]x_1,x_2[$ bien inclus dans F ce qui est contradiction avec le fait que le minimum de f sur F soit atteint en x_1. $\square$

16 Notons que f atteint son minimum en y si et seulement si $\Delta(u)=f(y+u)-f(y)\geqslant 0$ pour tout $u\in F$ c'est à dire si et seulement si $\Delta(u)=I(y,u)+\big(g(y)|u\big)=\dfrac{1}{2}(Au|u)+\big((Ay-b)|u\big)\geqslant 0$ pour tout $u\in F$.

- Si $Ay-b$ est orthogoanal à F l'inégalité est bien vérifiée car A est positive. Donc le minimum sur F de f est bien atteint en y.
- Réciproquement si le minimum est atteint en y alors $\Delta(u)\geqslant 0$ pour tout u de F. Supposons qu'il existe $u_0\in F$ tel que $\big((Ay-b|u_0\big)$ soit non nul. Quitte à changer u_0 en son opposé (qui appartient bien encore à F), on peut supposer $\big((Ay-b|u_0\big)<0$. Comme $\Delta(u)\geqslant 0$ pour tout $u\in F$, on a en particulier $\Delta(\alpha u_0)\geqslant 0$ pour tout $\alpha>0$. Or $\Delta(\alpha u_0)=\dfrac{1}{2}(Au_0|u_0).\alpha^2+(b|u_0).\alpha\sim(b|u_0).\alpha<0$ lorsque $\alpha\to 0^+$. Contradiction.
- En conclusion f atteint son minimum sur F en l'unique point $\overline{x}$ tel que $A\overline{x}-b$ soit orthogonal à F. $\square$

17 Comme $A\overline{x}-b$ est orthogoanl à F, on en particulier $\big((A\overline{x}-b)|\overline{x}\big)=0$ donc $(A\overline{x}|\overline{x})=(b|\overline{x})$ de sorte que :
$f(\overline{x})\underset{\text{DEF}}{=}\dfrac{1}{2}\big((A\overline{x})|\overline{x}\big)-(b|\overline{x})=-\dfrac{1}{2}(A\overline{x}|\overline{x})=-\dfrac{1}{2}(b|\overline{x})$. $\square$

18 • <u>Démonstration 1 dans le cas particulier</u> :
Commençons par remarquer que pour $x\notin F=\operatorname{Ker}B$ on a $\sup\limits_{y\in\mathbb{R}^n}L(x,y)=f(x)+\sup\limits_{y\in\mathbb{R}^n}(y|Bx)=+\infty$ clairement comme on le voit en considérant $y=\alpha.Bx$ avec $\alpha\to+\infty$.
Par contre pour $x\in F$ on a $L(x,y)=f(x)$ donc $\sup\limits_{y\in\mathbb{R}^n}L(x,y)=f(x)$.
Ainsi $\inf\limits_{x\in\mathbb{R}^n}\Big(\sup\limits_{y\in\mathbb{R}^n}L(x,y)\Big)=\inf\limits_{x\in F}f(x)=f(\overline{x})$.
Pour établir l'inégalité demandée, il suffit donc d'établir, par définition même du sup, que $\inf\limits_{x\in\mathbb{R}^n}L(x,y_0)\leqslant f(\overline{x})$ pour tout $y_0\in\mathbb{R}^n$.
Or $L(\overline{x},y_0)=f(\overline{x})$ puisque $\overline{x}\in F$. Donc $\inf\limits_{x\in\mathbb{R}^n}L(x,y)\leqslant f(\overline{x})$.
En conclusion $\sup\limits_{y\in\mathbb{R}^n}\Big(\inf\limits_{x\in\mathbb{R}^n}L(x,y)\Big)\leqslant\inf\limits_{x\in\mathbb{R}^n}\Big(\sup\limits_{y\in\mathbb{R}^n}L(x,y)\Big)$. $\square$

- <u>Démonstration 2 dans le cas général</u> :
Quitte bien sûr à se placer dans $\overline{R}$, l'inégalité demandée est valable pour toute application de φ de $\mathrm{E}\times F$ dans $\mathbb{R}$.

En effet pour établir l'inégalité, par définition même du sup, il suffit de prouver que, pour tout $y_0 \in F$:
$\inf\limits_{x\in E} \varphi(x,y_0)\Big) \leqslant \inf\limits_{x\in E}\Big(\sup\limits_{y\in F}\varphi(x,y)\Big)$ ce qui est clair car, pour tout $x \in E$, $\varphi(x,y_0) \leqslant \sup\limits_{y\in F}\varphi(x,y)$. $\square$

19 Soit (x^*, y^*) un point selle du Lagrangien (que l'on suppose exister). D'après la définition d'un point selle et des bornes supérieures et inférieures, il vient $\sup\limits_{y\in\mathbb{R}^n} L(x^*,y) \leqslant L(x^*,y^*) \leqslant \inf\limits_{x\in\mathbb{R}^n} L(x,y^*)$ donc a fortiori :

$\inf\limits_{x\in\mathbb{R}^n}\Big(\sup\limits_{y\in\mathbb{R}^n} L(x,y)\Big) \leqslant L(x^*,y^*) \leqslant \sup\limits_{y\in\mathbb{R}^n}\Big(\inf\limits_{x\in\mathbb{R}^n} L(x,y)\Big)$ d'où l'égalité cherchée compte tenu de la question 18.

Comte-tenu en outre du fait (Cf démonstration 1 question 18) que $\inf\limits_{x\in\mathbb{R}^n}\Big(\sup\limits_{y\in\mathbb{R}^n} L(x,y)\Big) = f(\overline{x})$ on a donc :

Si (x^*, y^*) est un point selle alors $L(x^*,y^*) = \inf\limits_{x\in\mathbb{R}^n}\Big(\sup\limits_{y\in\mathbb{R}^n} L(x,y)\Big) = \sup\limits_{y\in\mathbb{R}^n}\Big(\inf\limits_{x\in\mathbb{R}^n} L(x,y)\Big) = f(\overline{x})$. $\square$

20 • On a $L(x_1,y) \leqslant L(x_1,y_1)$ pour tout y de $\mathbb{R}^n$ si et seulement si $\sup\limits_{y\in\mathbb{R}^n} L(x_1,y) \leqslant L(x_1,y_1)$. Or on a noté dans la démonstration 1 de la question 18 que $\sup\limits_{y\in\mathbb{R}^n} L(x_1,y)$ vaut $+\infty$ si $x_1 \notin F$ et vaut $f(x_1) = L(x_1,y_1)$ si $x_1 \in F$.

Donc $\left(L(x_1,y) \leqslant L(x_1,y_1) \quad \forall y \in \mathbb{R}^n\right) <=> \left(Bx_1 = 0\right)$ (et alors $L(x_1,y) = f(x_1) \quad \forall y \in \mathbb{R}^n$) $\square$

$$\begin{aligned}
\bullet \quad \left(L(x_1,y_1) \leqslant L(x,y_1) \quad \forall x \in \mathbb{R}^n\right) &\iff \left(L(x_1,y_1) \leqslant L(x_1+u,y_1) \quad \forall u \in \mathbb{R}^n\right)\\
&\iff \left(f(x_1) \leqslant f(x_1+u) + (y_1|Bu) \quad \forall u \in \mathbb{R}^n\right)\\
&\iff \left((Ax_1-b|u) + \frac{1}{2}(Au|u) + (y_1|Bu) \geqslant 0 \quad \forall u \in \mathbb{R}^n\right) \text{ (questions 7 et 8)}\\
&\iff \left((X_1|u) + \frac{1}{2}(Au|u) \geqslant 0 \quad \forall u \in \mathbb{R}^n\right) \text{ avec } X_1 = Ax_1 + {}^tBy_1 - b
\end{aligned}$$

Si $X_1 = 0$ l'inégalité est bien satisfaite puisque A est positive.
Réciproquement si $X_1 \neq 0$ elle n'est pas satisfaite car avec $u = -\alpha X_1$ avec $\alpha > 0$ on a :
$(X_1|u) + \dfrac{1}{2}(Au|u) = -\alpha||X_1||^2 + \dfrac{\alpha^2}{2}(AX_1|X_1) \sim -\alpha||X_1||^2 < 0$ lorsque $\alpha \to 0^+$.

En conclusion $\left(L(x_1,y_1) \leqslant L(x,y_1) \quad \forall x \in \mathbb{R}^n\right) \iff \left(Ax_1 + {}^tBy_1 = b\right)$ $\square$

21 • Si $x_1 = \overline{x}$ (ce qui implique que $x_1 \in F$) et $Ax_1 + {}^tBy_1 = b$ alors les deux inégalités de la question précédente sont satisfaites ce qui signifie exactement que (x_1,y_1) est un point selle.
• Réciproquement si (x_1,y_1) est un point selle alors, d'après la question précédente, on a $x_1 \in F$ et $Ax_1 + {}^tBy_1 = b$.
En outre d'après la question 19 on a $L(x_1,y_1) = f(\overline{x})$ mais aussi $L(x_1,y_1) = f(x_1)$ puisque $x_1 \in F$. Ainsi $x_1 = \overline{x}$ d'après l'unicité du point x de F minimisant $f(x)$.
• En conclusion (x_1,y_1) est un point selle si et ssi $x_1 = \overline{x}$ et $Ax_1 + {}^tBy_1 = b$. $\square$

22 • Remarque liminaire : Pour y fixé quelconque dans $\mathbb{R}^n$, la seconde équivalence de la question 20 prouve que la fonction $x \longmapsto L(x,y)$ atteint son minimum en la solution du système $Ax = b - {}^tBy$.
• En particulier x^0 est bien défini par $Ax^0 = b - {}^tBy^0$, puis y^1 par $y^1 = y^0 + \rho_0 Bx^0$, puis x^1 par $Ax^1 = b - {}^tBy^1$ puis y_2 ...
Étant donnée une suite $(\rho_m)_{m\in\mathbb{N}}$ les deux suites $(x^m)_{m\in\mathbb{N}}$ et $(y^m)_{m\in\mathbb{N}}$ sont parfaitement définies. $\square$
• Comme noté ci-dessus, on a $Ax^m = b - {}^tBy^m$. Par ailleurs $Ax^* = b - {}^tBy^*$ d'après la question 21.
Donc $A(x^m - x^*) + {}^tB(y^m - y^*) = 0 \quad (1)$ $\square$
• Comme $x^* = \overline{x} \in F$ on a $Bx^* = 0$ et de $y^{m+1} = y^m + \rho_m Bx^m$ on tire :
$y^{m+1} - y^* = y^m - y^* + \rho_m B(x^m - x^*) \quad (2)$ $\square$

23 Nous avons $\big((y^m - y^*)|B(x^m - x^*)\big) = \big({}^tB(y^m - y^*)|(x^m - x^*)\big) = -\big(A(x^m - x^*)|(x^m - x^*)\big)$ d'après (1).
En élevant (2) au carré on obtient alors l'égalité demandée. $\square$

24 Il existe P orthogonale telle que $A = PD\,{}^tP$ avec $D = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n)$. Alors $P\Delta\,{}^tP$ avec $\Delta = (\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n})$ convient clairement : en effet elle est bien symétrique et comme ses valeurs propres sont strictement positives, elle est définie positive. $\square$
On peut d'ailleurs démontrer l'unicité de cette solution mais c'st moins simple.

25 • Notons $M = B.A^{-1/2}$. La matrice $A^{-1/2} = P.\Delta^{-1}.{}^tP$ étant symétrique, il en découle que ${}^tM = A^{-1/2}.{}^tB$. Ainsi $C = {}^tM.M$ est symétrique positive classiquement. $\square$

• Il en découle d'après la question 1 que $(Cx|x) \leqslant \nu ||x||^2$ pour tout $x \in \mathbb{R}^n$ avec $\nu = \max\limits_{\lambda \in \mathrm{Sp}(C)} \lambda$.

Or $(Cx|x) = ||BA^{-1/2}x||^2$. Pour $x = A^{1/2}u$ l'inégalité ci-dessus s'écrit donc $||Bu||^2 \leqslant \nu ||A^{1/2}u||^2$.

Par ailleurs $||A^{1/2}u||^2 = (A^{1/2}u|A^{1/2}u) = (Au|u)$ car $A^{1/2}$ est symétrique et $A^{1/2}.A^{1/2} = A$.
Ainsi $||Bu||^2 \leqslant \nu(Au|u)$ pour tout $u \in \mathbb{R}^n$ avec $\nu = \max\limits_{\lambda \in \mathrm{Sp}(C)} \lambda$. $\square$

26 Posons $a^m = ||y^m - y^*||^2$ et $u^m = x^m - x^*$. D'après la question 23, il vient :
$a^{m+1} - a^m = -\rho_m \Big(2(Au^m|u^m) - \rho_m ||Bu^m||^2 \Big)$.
Or $2(Au^m|u^m) - \rho_m ||Bu^m||^2 \geqslant 2(Au^m|u^m) - \rho_m \nu (Au^m|u^m) = (2 - \nu\rho_m)(Au^m|u^m) \geqslant 0$ car $\nu\rho_m < 2$ et A positive.
Ainsi $a^{m+1} - a^m \leqslant 0$ car en outre $\rho_m \geqslant 0$. $\square$

27 Ainsi la suite $(a^m)_{m \in \mathbb{N}}$ positive décroissante converge dans $\mathbb{R}$ et par passage à la limite dans l'égalité de la question 23 et en utilisant le fait que $0 < \alpha \leqslant \rho_m$ il vient que $2(Au^m|u^m) - \rho_m ||Bu^m||^2 \xrightarrow[m \to +\infty]{} 0$.

Or dans la question précédente on a établi l'inégalité $0 \leqslant (2 - \nu\rho_m)(Au^m|u^m) \leqslant 2(Au^m|u^m) - \rho_m ||Bu^m||^2$.
Par le théorème des gendarmes on en déduit que $(2 - \nu\rho_m)(Au^m|u^m) \xrightarrow[m \to +\infty]{} 0$.

Or $2 - \nu\rho_m \geqslant 2 - \nu\beta > 0$ donc $(Au^m|u^m) \xrightarrow[m \to +\infty]{} 0$.

Comme $0 \leqslant ||u^m||^2 \leqslant \dfrac{1}{\lambda_1}(Au^m|u^m)$ on en déduit que $u^m \xrightarrow[m \to +\infty]{} 0$ c'est à dire que $x^m \xrightarrow[m \to +\infty]{} x^*$. $\square$

FIN