Correction du DS n°9

Sujet de type Centrale-Mines

Première partie

1°) On utilise le théorème spectral : M , symétrique réelle, admet une base $\mathcal{B}$ *orthonormée* de vecteurs propres . En écrivant $(x)_i$ les composantes d'un vecteur quelconque $x \in \mathbb{R}^n$ dans cette base, on a que l'expression du produit scalaire est l'expression canonique, et d'autre part, que $M.x$ s'écrit simplement en fonction des valeurs propres (λ_i) correspondantes de M : $(M.x)_i = \lambda_i.(x)_i$. On en déduit : $\left(Mx \mid x\right) = \sum_{i=1}^{n} \lambda_i.(x)_i^2$. D'où, en appelant p (respectivement q) la plus petite (respectivement la plus grande) des valeurs propres de M , on a l'inégalité :

$$p\|x\|^2 = \sum_{i=1}^{n} p.(x)_i^2 \leq \left(Mx \mid x\right) \leq \sum_{i=1}^{n} q(x)_i^2 = q\|x\|^2 \quad .$$

Remarquons que ces inégalités sont les meilleures possibles, puisqu'elles deviennent des égalités lorsque x est un vecteur propre associé à la plus petite (respectivement la plus grande) des valeurs propres .

2°) La condition suffisante découle de l'inégalité ci-dessus : si $p > 0$, on a pour tout $x \in \mathbb{R}^n$, $\left(Mx \mid x\right) \geq p.\|x\|^2 \geq 0$; d'autre part, 0 n'est alors pas valeur propre , donc M est inversible.
Mais la condition est aussi nécessaire, car l'inégalité $\left(Mx \mid x\right) \geq 0$ doit être en particulier vraie pour les vecteurs propres (non nuls) de M , d'où, pour toute valeur propre λ_i dont x_i est un vecteur propre, de M , $\lambda_i.\|x_i\|^2 \geq 0$, soit $\lambda_i \geq 0$; de plus, l'inversibilité de M interdit que 0 soit valeur propre. Donc les valeurs propres λ_i de M sont bien strictement positives.

3°) En se plaçant toujours dans une base orthonormée propre pour M , on a, pour tout vecteur $x \in \mathbb{R}^n$:

$$\|M.x\|^2 = \sum_{i=1}^{n} (Mx)_i^2 = \sum_{i=1}^{n} \lambda_i^2.x_i^2 \leq \left(\underset{1\leq i\leq n}{\text{Max}}\, \lambda_i^2\right).\sum_{i=1}^{n} (x)_i^2 = \left(\underset{1\leq i\leq n}{\text{Max}}\, \lambda_i^2\right).\|x\|^2 \quad .$$

En prenant la racine carrée des deux membres de l'inégalité ci-dessus, et en utilisant $\sqrt{\left(\underset{1\leq i\leq n}{\text{Max}}\, \lambda_i^2\right)} = \underset{1\leq i\leq n}{\text{Max}}\, |\lambda_i|$, on en déduit : $\forall\, x \in \mathbb{R}^n \quad \|M.x\| \leq \underset{1\leq i\leq n}{\text{Max}}\, |\lambda_i|.\|x\|$, d'où

$$N(M) \leq \underset{1\leq i\leq n}{\text{Max}}\, |\lambda_i|.$$

Soit $i_0 \in \{1, ...n\}$ tel que $\lambda_{i_0} = \underset{1\leq i\leq n}{\text{Max}}\, |\lambda_i|$ et x_{i_0} un vecteur propre associé. alors

$$\|M.x_{i_0}\|^2 = \lambda_{i_0}^2.$$

Donc

$$N(M) \geq \underset{1\leq i\leq n}{\text{Max}}\, |\lambda_i|.$$

Finalement $N(M) = \underset{1\leq i\leq n}{\text{Max}}\, |\lambda_i|$.

4°) On suppose ici, comme indiqué dans le préambule de cette partie de l'énoncé, que A est inversible, donc (cf **2.**) que les valeurs propres de A sont *strictement* positives.
La formule de récurrence s'écrit $x^{k+1} = S.x^k + \alpha.b$ où on a posé $S = Id - \alpha.A$. S est encore une matrice symétrique, dont les valeurs propres sont les $1 - \alpha.\lambda_i$, λ_1 , $\ldots$, λ_n étant les valeurs propres (distinctes ou non) de A . Vu

la condition imposée à α , les valeurs propres de S sont dans $]-1,1[$, et en particulier $N(S)<1$. L'unique point z vérifiant $A.z=b$ vérifie aussi $z=S.z+\alpha.b$, donc la suite de vecteurs $y^k=z-x^k$ vérifie la relation de récurrence $y^{k+1}=S.y^k$, soit $y^k=S^k.y^0$. Comme définition de la norme N , $\|S.y\|\le N(S).\|y\|$, on a donc pour tout entier naturel k

$$\|y^k\|\le\big(N(S)\big)^k.\|y^0\|.$$

De $N(S)\in]0,1[$ on conclut alors que $y^k\to 0_{\mathbb{R}^n}$, soit $x^k\to z$.

5°) Par bilinéarité et symétrie du produit scalaire, on trouve :

$$f(x+u)-f(x)=\big(Ax\mid u\big)-\big(b\mid u\big)+\frac{1}{2}\big(Au\mid u\big)=\big(Ax-b\mid u\big)+\frac{1}{2}\big(Au\mid u\big)\quad .$$

6°) Soit x élément de $\mathbb{R}^n$.
Comme pour tout $u\in\mathbb{R}^n$, $|\frac{1}{2}(Au|u)|\le N(A)\|u\|^2$ (merci à Cauchy et Schwarz)

$$\frac{1}{2}(Au|u)|=o(\|u\|)\,,\ u\to O_{\mathbb{R}^n}.$$

Donc l'application *linéaire* $(AX-b|\cdot)$ est la différentielle de f an x (et $Ax-b$ son gradient en x). Donc f admet en x n dérivées partielles et :

$$\partial_i f(x)=\delta f(x)\cdot e_k=(AX-b|e_k).$$

7°) Soit $x\in\mathbb{R}^n$. On a vu que $g(x)$, gradient de f en x vaut : $g(x)=A.x-b$.

8°) Cela a été vu en **1.** : par **5.**, $I(x,u)=\frac{1}{2}\big(Au\mid u\big)$ où la matrice $\frac{1}{2}A$ est symétrique définie positive, donc l'encadrement $r\|u\|^2\le I(x,u)\le s\|x\|^2$ a lieu pour r,s respectivement égaux à l'inf et au sup des valeurs propres de $\frac{1}{2}A$.

9°) CONDITION NÉCESSAIRE : un minimum absolu sur $\mathbb{R}^n$ est *a fortiori* un minimum local et par théorème ($\mathbb{R}^n$ est un ouvert et f polynomiale est de classe $\mathcal{C}^1$!), c'est nécessairement un point critique.
CONDITION SUFFISANTE : si $g(x)=0$, alors on a pour tout $u\in\mathbb{R}^n$, puisque $I(x,u)\ge0$, $f(x+u)=f(x)+I(x,u)\ge f(x)$, donc $f(x)$ est le minimum absolu de f.
Remarque. *La question 5. donne la formule de Taylor à l'ordre 2, qui est exacte puisque f est polynomiale de degré deux ou moins et $\frac{1}{2}A$ est en tout point la matrice Hessienne de f ?*

10°) En appliquant le calcul du **5.** au vecteur $u=-\alpha.g(x)=-\alpha(Ax-b)$, on obtient (par **1.** et dæinition de λ_n) :

$$f\big(x-\alpha g(x)\big)-f(x)=-\alpha\big(g(x)\mid g(x)\big)+\frac{\alpha^2}{2}\big(Ag(x)\mid g(x)\big)\le\Big(-\alpha+\frac{\alpha^2}{2}\lambda_n\Big).\|g(x)\|^2\quad .$$

L'hypothèse $0<\alpha<\dfrac{2}{\lambda_n}$ entraîne $-\alpha+\frac{\alpha^2}{2}\lambda_n<0$, donc $f\big(x-\alpha g(x)\big)-f(x)\le0$ (l'inégalité est même stricte sauf si $g(x)=0$).

11°) L'algorithme consiste à prendre un $y^0\in\mathbb{R}^n$ quelconque (par exemple $y^0=0$) puis de définir (y^k) par la relation $y^{k+1}=y^k-\alpha.g(y^k)$, α choisi comme précédemment. On obtient ainsi une suite (y^k) telle que $f(y^k)$ soit strictement décroissante. On remarque que cette suite est en fait exactement celle définie au **4.**, puisque $g(x)=Ax-b$. Elle converge donc vers $z=A^{-1}b$, qui vérifie $g(z)=0$. Par **9.**, f admet bien un minimum en $z=\lim y^k$ dans ce cas.

Seconde partie

12°) Par **1.** on a une inégalité de la forme : $\forall\,x\in\mathbb{R}^n\quad r.\|x\|^2\le\frac{1}{2}\big(Ax\mid x\big)$, où r est le réel strictement positif $\frac{1}{2}\min\limits_{\lambda\in sp(A)}\lambda$. De plus, par Cauchy-Schwarz, $\big(b\mid x\big)\le\|b\|.\|x\|$. D'où finalement :

$$\forall\,x\in\mathbb{R}^n\quad f(x)\ge\|x\|.\big(r.\|x\|-\|b\|\big)\quad .$$

Comme la fonction au membre de droite $z\mapsto h(z)=z(rz-\|b\|)$ tend vers $+\infty$ quand $z=\|x\|\to\infty$, la propriété demandée est clairement vérifiée.

13°) C'est la propriété précédente, appliquée à $c=f(y)$, et restreinte aux vecteurs x de F .

14°) Un sous-espace vectoriel étant non vide, choisissons $y\in F$. la question précédente fournit un réel $r>0$ tel que pour tout élément x de F si $\|x\|\ge r$ alors $f(x)\ge f(y)$.
Posons $K=F\bigcap\bar{B}(0,r)$, où $\bar{B}(0,r)$ désigne lma boule fermé de rayon r et de centre 0..
— K est non vide il admet 0 commme élément ;
— K est un fermé comme intersection de deux fermés (et même un fermé de F) ;
— K est borné car inclus dans $\bar{B}(0,r)$.
Donc comme E est de dimension finie, K ets un compact. La restriction de f, application continue, et même — nous le vismes dans la partie I — de classe $\mathcal{C}^1$, atteint sa borne inférieure en un point $\bar{x}$. Comme $f(\bar{x})\le f(y)$ et que pour tout point x de F qui n'est pas dans la boule $\bar{B}(0,r)$ on a $f(x)>f(y)$ la restriction de f admet un minimum atteint en $\bar{x}$.

15°) Admettons que f soit convexe, et supposons l'existence de 2 points $\bar{x} \neq \bar{y}$ où à $f_{|F}$ atteigne son minimum . Déjà, par double inégalité, $f(\bar{x}) = f(\bar{y})$. Puis, pour $\lambda = 1/2$, par exemple, on a :

$$f\Big(\frac{1}{2}\bar{x} + \frac{1}{2}\bar{y}\Big) < \frac{1}{2}f(\bar{x}) + \frac{1}{2}f(\bar{y}) = f(\bar{x}) \quad ,$$

ce qui, puisque $\frac{1}{2}\bar{x} + \frac{1}{2}\bar{y}$ reste dans F , contredit qu'en $\bar{x}$ l'application $f_{|F}$ a tteigne son minimum. Donc le minimum, qui existe par **14.**, est unique. *Remarque : f est la somme d'une fonction quadratique $x \mapsto \frac{1}{2}\big(Ax \mid x\big)$ et d'une fonction linéaire $x \mapsto -\big(b \mid x\big)$. Une fonction linéaire étant évidemment convexe (au sens large), la stricte convexité de f résulte de celle de $x \mapsto \big(Ax \mid x\big)$. Or celle-ci découle simplement de la définie positivité de la matrice symétrique A (exercice).*

16°) **Version longue.**
Supposons que $f_{|F}$ atteigne en F un minimum absolue en un point y. Soit alors u un vecteur de f, l'application

$$\phi \;:\; \mathbb{R} \to \mathbb{R}\, t \mapsto f(y + tu)$$

admet un minimum en 0 (pour tout t réel, $y + tu \in F$). Donc

$$0 = \phi'(0) = (\nabla f(y + 0u)|u) = (Ay - b|u).$$

Comme u est quelconque $Ay - n \perp F$.
Version courte. Supposons que $f_{|F}$ atteigne en F un minimum absolue en un point y. Alors le cours affirme que $\mathrm{d}f(y)$ est nulle sur l'espace tangent à F en B, c'est-à-dire à F lui-même (l'espace tangent à un espace affine est l'espace vectoriel qui le dirige). Donc le gradient de f en y qui vaut $Ay - B$ est orthogonal à F.
Réciproquement, si $Ay - b$ est orthogonal à F , on a comme à la question **9.** que pour tout élément u de F, $f(y + u) - f(y) = I(y, u) = \frac{1}{2}\big(Au \mid u\big) \geq 0$, et $f(y)$ est bien un minimum de $f_{|F}$.

17°) On a par Q16. $\big(A\bar{x} - b \mid \bar{x}\big) = 0$, soit $\big(A\bar{x} \mid \bar{x}\big) = \big(b \mid \bar{x}\big)$. De l'expression de f on déduit alors

$$f(\bar{x}) = -\frac{1}{2}\big(A\bar{x} \mid \bar{x}\big) = -\frac{1}{2}\big(b \mid \bar{x}\big) \quad .$$

18°) Petite diffuculté non signalée ici : les *inf* et *sup* sont à prendre dans $\bar{\mathbb{R}}$: par exemple, si $Bx \neq 0$, comme la fonction $y \mapsto L(x, y)$ est linéaire, elle est non bornée car non identiquement nulle !
L'inégalité demandée est une propriété générale des fonctions de $\mathbb{R}^n \times \mathbb{R}^n$ dans $\bar{\mathbb{R}}$. En effet, pour tout $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$, en utilisant que le sup (respectivement l'inf) est un majorant (respectivement un minorant), on a :

$$\operatorname*{Inf}_{x \in \mathbb{R}^n} f(x, y) \leq f(x, y) \leq \operatorname*{Sup}_{y \in \mathbb{R}^n} f(x, y) \quad ,$$

inégalité qu'on peut noter $m(y) \leq M(x)$. A x fixé, $M(x)$ est un majorant des $m(y)$, donc, puisque le sup est le plus petit des majorants, $\operatorname*{Sup}_{y \in \mathbb{R}^n} m(y) \leq M(x)$; puis, par le même raisonnement, quand x varie dans $\mathbb{R}^n$, l'inf étant le plus grand des majorants, $\operatorname*{Sup}_{y \in \mathbb{R}^n} m(y) \leq \operatorname*{Inf}_{x \in \mathbb{R}^n} M(x)$, ce qui est l'inégalité demandée.

19°) En un point selle (x^*, y^*) on a les inégalités

$$\forall\, (x, y) \in \mathbb{R}^n \times \mathbb{R}^n \quad L(x^*, y) \overset{(1)}{\leq} L(x^*, y^*) \overset{(2)}{\leq} L(x, y^*) \quad .$$

L'inégalité (1) donne, puisque le sup est plus petit qu'un majorant, $\operatorname*{Sup}_{y \in \mathbb{R}^n} L(x^*, y) \leq L(x^*, y^*)$. Puis, comme l'inf minore une instance particulière, $\operatorname*{Inf}_{x \in \mathbb{R}^n} \operatorname*{Sup}_{y \in \mathbb{R}^n} L(x, y) \leq L(x^*, y^*)$. De même, l'inégalité (2) donne $L(x^*, y^*) \leq \operatorname*{Inf}_{x \in \mathbb{R}^n} L(x, y^*)$. puis $L(x^*, y^*) \leq \operatorname*{Sup}_{y \in \mathbb{R}^n} \operatorname*{Inf}_{x \in \mathbb{R}^n} L(x, y)$. A cause de l'inégalité « inverse »démontrée en **18.**, les inégalités ci-dessus sont des égalités, cqfd.

20°) • On cherche une proposition équivalente à la proposition

$$\forall\, y \in \mathbb{R}^n \quad L(x_1, y) - L(x_1, y_1) = \big(y - y_1 \mid Bx_1\big) \leq 0 \quad .$$

Si $Bx_1 = 0$, l'inégalité est vraie car c'est une égalité, si $Bx_1 \neq 0$, l'inégalité est fausse pour $y = y_1 + Bx_1$, donc la condition nécessaire et suffisante est bien $Bx_1 = 0$.
• Par propriété du produit scalaire canonique, on a

$$\big(y_1 \mid Bx\big) = {}^t y_1.Bx = {}^t\big({}^t y_1.Bx\big) = {}^t x\, {}^t B y_1 = \big({}^t B y_1 \mid x\big) \quad ,$$

On voit donc qu'si y_1 fix $L(x, y_1) = \frac{1}{2}\big(Ax \mid x\big) - \big(b \mid x\big) + \big({}^t By_1 \mid x\big)$ est une fonction f où on a remplace b par $b - {}^t By_1$. La question **9.** dit alors que $x \mapsto L(x, y_1)$ admet un minimum en x_1 si et seulement si $Ax_1 = b - {}^t By_1$, soit $Ax_1 + {}^t By_1 = b$, ce qui est bien la propriété demandandé.

21°) La définition du point selle donne par les équivalences ci-dessus où on a remplacé (x_1, y_1) par (x^*, y^*) : si (x^*, y^*) est un point selle alors $Bx^* = 0$ et $Ax^* + {}^tBy^* = b$.
Mais de plus, quand $Bx = 0$, i.e. quand $x \in F$, $L(x, y)$ se réduit à $f(x)$. Donc, l'inégalité de définition $\forall x \in \mathbb{R}^n \;\; L(x^*, y^*) \leq L(x, y^*)$ implique bien que $f(x^*)$ est un minimum de $f_{|F}$.
Réciproquement, si $x_1 \in F$ réalise un minimum pour $f_{|F}$, alors on a on a $Bx_1 = 0$, et si de plus $Ax_1 + {}^tBy_1 = b$ alors la question **20.** assure que (x_1, y_1) est un point selle par définition du point selle.

22°) • Supposons y^m construit. L'application $x \mapsto L(x, y^m) = \frac{1}{2}\big(Ax \mid x\big) - \big(b - {}^tBy^m \mid x\big)$ est du même type que la fonction f dans les questions **14.** et **15.**, elle admet donc un minimum et un seul. Ainsi x^m est bien défini, ainsi que $y^{m+1} = y^m + \rho_m.Bx^m$.
• Par **20.** on a $Ax^m + {}^tBy^m = b$, par **21.**, $Ax^* + tBy^* = b$, donc par différence, $A(x^m - x^*) + {}^tB(y^m - y^*) = 0$.
• Toujours par **21.**, on sait que $x^* \in F$, donc $Bx^* = 0$. Ainsi la relation $y^{m+1} - y^* = y^m - y^* + \rho_m.B(x^m - x^*)$ découle-t-elle de la relation de définition de y^{m+1} .

23°) Par bilinéarité du produit scalaire, la dernière relation ci-dessus implique

$$\|y^{m+1} - y^*\|^2 = \|y^m - y^*\|^2 + \rho_m^2.\|B(x^m - x^*)\|^2 + 2\rho_m.\big(y^m - y \mid B(x^m - x)\big) \quad .$$

Mais par **22.**

$$\big(y^m - y \mid B(x^m - x)\big) = \big({}^tB(y^m - y) \mid x^m - x\big) = -\big(A(x^m - x^*) \mid x^m - x^*\big) \quad ,$$

et donc finalement

$$\|y^{m+1} - y^*\|^2 = \|y^m - y^*\|^2 + \rho_m^2.\|B(x^m - x^*)\|^2 - 2\rho_m.\big(A(x^m - x^*) \mid x^m - x^*\big) \quad .$$

24°) Classique; les valeurs propres λ_i de A sont > 0 , donc admettent des racines carrées réelles. En écrivant une diagonalisation de A sous la forme $A = P.diag(\lambda_1, \ldots, \lambda_n).{}^tP$, P matrice carrée orthogonale, on voit que la matrice $A^{1/2} = P.diag(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n}).{}^tP$ convient : elle est bien symétrique réelle, à valeurs propres positives donc positive, et $\big(A^{1/2}\big)^2 = A$.
Notons qu'en fait, puisque les λ_i sont strictement positives, A est définie positive, donc $A^{-1/2}$ existe bien.

25°) L'inverse d'une matrice symétrique inversible est symétrique, donc $A^{-1/2}$ est symétrique. Il en résulte par la règle de transposition d'un produit, ${}^tC = {}^tA^{-1/2}.{}^tB.B.{}^tA^{-1/2} = C$: C est symétrique.
Pour tout $x \in \mathbb{R}^n$, $\big(Cx \mid x\big) = \big(A^{-1/2}.{}^tB.B.A^{-1/2}.x \mid x\big) = \big(B.A^{-1/2}.x \mid B.A^{-1/2}.x\big) = \|B.A^{-1/2}.x\|^2 \geq 0$, donc C est positive.
Pour tout $u \in \mathbb{R}^n$, posons $x = A^{1/2}u$; ainsi, $u = A^{-1/2}x$. Par le calcul ci-dessus et par **1.**,

$$\|Bu\|^2 = \big(Cx \mid x\big) \leq \Big(\sup_{\mu \in sp(C)} \mu\Big).\big(x \mid x\big) = \nu.\big(A^{1/2}u \mid A^{1/2}u\big) = \nu.\big(Au \mid u\big) \quad ,$$

où on a posé $\nu = \sup_{\mu \in sp(C)} \mu$.

26°) L'égalité de **23.** donne alors, puisque $\rho_m \geq 0$:

$$\|y^{m+1} - y^*\|^2 - \|y^m - y^*\|^2 = \rho_m\Big(\rho_m\|Bu^m\|^2 - 2\big(Au^m \mid u^m\big)\Big) \leq \rho_m(\rho_m\nu - 2)\big(Au^m \mid u^m\big) \quad .$$

Par les hypothèses sur ρ_m , $\rho_m\nu - 2 < 0$, et comme A est positive, on a $\|y^{m+1} - y^*\|^2 - \|y^m - y^*\|^2 \leq 0$: la suite $\|y^m - y^*\|$ est décroissante.

27°) La suite positive décroissante $\|y^m - y^*\|^2$ est donc convergente, ce qui implique que $\|y^{m+1} - y^*\|^2 - \|y^m - y^*\|^2 \to 0$. Compte tenu des hypothèses sur ρ_m , l'inégalité ci-dessus donne

$$\|y^{m+1} - y^*\|^2 - \|y^m - y^*\|^2 \leq \alpha(\beta\nu - 2)\big(Au^m \mid u^m\big) \leq 0 \quad , \quad \text{avec } \alpha(\beta\nu - 2) \neq 0 \quad ,$$

d'où par théorème d'encadrement des limites, $\big(Au^m \mid u^m\big) \to 0$. Mais alors, de l'encadrement, donné par **1.**, $0 \leq \|u^m\|^2 \leq \frac{1}{\rho}\big(Au^m \mid u^m\big)$, on déduit que $\|u^m\| \to 0$, soit que $x^m \to x^*$.