A 2003 Math MP 2

ÉCOLE NATIONALE DES PONTS ET CHAUSSÉES.
ÉCOLES NATIONALES SUPÉRIEURES DE L'AÉRONAUTIQUE ET DE L'ESPACE,
DE TECHNIQUES AVANCÉES, DES TÉLÉCOMMUNICATIONS,
DES MINES DE PARIS, DES MINES DE SAINT-ÉTIENNE, DES MINES DE NANCY,
DES TÉLÉCOMMUNICATIONS DE BRETAGNE.
ÉCOLE POLYTECHNIQUE (Filière TSI).

CONCOURS D'ADMISSION 2003

ÉPREUVE DE MATHÉMATIQUES
DEUXIÈME ÉPREUVE
Filière MP

(Durée de l'épreuve : 4 heures)
(L'usage d'ordinateur ou de calculette est interdit).

Sujet mis à la disposition des concours :
Cycle International, ENSTIM, ENSAE (Statistique), INT, TPE-EIVP.

Les candidats sont priés de mentionner de façon apparente sur la première page de la copie :
MATHÉMATIQUES 2-Filière MP.

Cet énoncé comporte 6 pages de texte.

Si, au cours de l'épreuve, un candidat repère ce qui lui semble être une erreur d'énoncé, il le signale sur sa copie et poursuit sa composition en expliquant les raisons des initiatives qu'il est amené à prendre.

L'objet du problème est l'étude de méthodes analytiques (méthodes du gradient, du Lagrangien) pour résoudre l'équation linéaire $A.x = b$ où A est une matrice symétrique positive, inversible, b un vecteur donné de $\mathbb{R}^n$ et x un vecteur inconnu de $\mathbb{R}^n$ ou d'un sous-espace vectoriel F de $\mathbb{R}^n$.

Dans tout le problème, l'entier n est un entier naturel supérieur ou égal à 2 ($n \geq 2$) ; la base canonique de $\mathbb{R}^n$ est notée $e_1,\ e_2, ...,\ e_n$; le produit scalaire de deux vecteurs x et y de $\mathbb{R}^n$ est noté $(x \mid y)$. La norme d'un vecteur x est notée $\|x\|$.

Les matrices considérées sont réelles ; l'espace vectoriel des matrices carrées réelles d'ordre n est noté $M_n(\mathbb{R})$. Il est admis que l'application qui, à une matrice M de $M_n(\mathbb{R})$, associe la borne supérieure $N(M)$ des normes des images par M des vecteurs unitaires de $\mathbb{R}^n$ est une norme :

$$N(M) = \sup_{\|x\|=1} \|M.x\| .$$

Une matrice symétrique A est dite positive lorsque, pour tout vecteur x de $\mathbb{R}^n$, le produit scalaire des vecteurs $A.x$ et x est positif ou nul $(A.x \mid x) \geq 0$.

Première partie

Le but de cette partie est la résolution de l'équation $A.x = b$ où A est une matrice carrée d'ordre n symétrique positive et inversible, b un vecteur donné de $\mathbb{R}^n$ et x un vecteur inconnu.

Résultats préliminaires :

Soit M une matrice carrée symétrique d'ordre n.

1. Démontrer qu'il existe un plus grand réel p et un plus petit réel q tels que, pour tout vecteur x de $\mathbb{R}^n$, le produit scalaire $(M.x \mid x)$ vérifie l'encadrement suivant :

$$p \ \|x\|^2 \leq (M.x \mid x) \leq q \ \|x\|^2 .$$

Préciser ces deux réels p et q en fonction des valeurs propres de la matrice M.

2. Montrer que, pour que cette matrice M soit inversible et positive, il faut et il suffit que toutes ses valeurs propres soient strictement positives.

3. Démontrer que la norme $N(M)$ d'une matrice M symétrique est égale à la plus grande valeur absolue des valeurs propres λ_i $(1 \leq i \leq n)$ de la matrice M :

$$N(M) = \sup_{1 \leq i \leq n} |\lambda_i| .$$

Étant donnés la matrice carrée, d'ordre n, symétrique positive A et le vecteur b, soit α un réel strictement positif strictement majoré par $2/\lambda_n$ $(0 < \alpha < 2/\lambda_n)$ où λ_n est la plus grande valeur propre de la matrice A ; soit $(x^k)_{k \in \mathbb{N}}$ la suite définie par un premier vecteur x^0 choisi arbitrairement dans $\mathbb{R}^n$ et par la relation de récurrence suivante : pour tout entier naturel k,

$$x^{k+1} = x^k + \alpha \ \left(b - A.x^k\right) .$$

Étude de la suite $(x^k)_{k \in \mathbb{N}}$:

4. Démontrer que la suite $(x^k)_{k \in \mathbb{N}}$ est une suite convergente de limite le vecteur z de l'espace $\mathbb{R}^n$, solution de l'équation $A.x = b$.

Soit f la fonction réelle, définie dans $\mathbb{R}^n$, par la relation :

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

Minimum de f :

5. Calcul préparatoire : démontrer que l'expression $f(x+u) - f(x)$ se calcule en fonction des expressions $(A.u \mid u)$, $(A.x \mid u)$ et $(b \mid u)$.

6. Démontrer que la fonction $f : x \longmapsto f(x)$ admet des dérivées partielles $\frac{\partial f}{\partial x_k}$ $(1 \leq k \leq n)$:

$$x \longmapsto \frac{\partial f}{\partial x_k}(x).$$

Étant donné un vecteur x de $\mathbb{R}^n$, soit $g(x)$ le vecteur de $\mathbb{R}^n$ dont les coordonnées, dans la base canonique de $\mathbb{R}^n$, sont égales aux valeurs des dérivées partielles de la fonction f en ce point x :

$$g(x) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(x) \; e_k.$$

7. Exprimer ce vecteur $g(x)$ au moyen de la matrice A et des vecteurs x et b.

Étant donnés deux vecteurs x et u de $\mathbb{R}^n$, soit $I(x,u)$ l'expression suivante :

$$I(x,u) = f(x+u) - f(x) - (g(x) \mid u).$$

8. Démontrer que, pour tout vecteur x donné, il existe deux constantes positives ou nulles r et s telles que, pour tout vecteur u, $I(x,u)$ vérifie la relation suivante :

$$r \; \|u\|^2 \leq I(x,u) \leq s \; \|u\|^2.$$

9. Démontrer que, pour que la fonction f admette en z un minimum, il faut et il suffit que le vecteur z vérifie la relation $A.z = b$.

Recherche du minimum de f :

Soit α un réel compris strictement entre 0 et $2/\lambda_n$ $(0 < \alpha < 2/\lambda_n)$.

10. Étant donné un vecteur x de $\mathbb{R}^n$, déterminer le signe de l'expression suivante

$$f(x - \alpha\, g(x)) - f(x).$$

11. Proposer, à partir de ce résultat, une méthode pour construire une suite de vecteurs $(y^k)_{k\in\mathbb{N}}$ qui converge vers le vecteur z en lequel la fonction f atteint son minimum ; la justification de la convergence n'est pas demandée.

Seconde partie

Le but de cette partie est de rechercher un vecteur x appartenant à un sous-espace vectoriel F de $\mathbb{R}^n$ qui vérifie l'équation $A.x = b$ où A est une matrice carrée d'ordre n symétrique positive et inversible. Le sous-espace vectoriel F de $\mathbb{R}^n$ est supposé être le noyau d'une matrice B appartenant à $M_n(\mathbb{R})$; ce noyau est supposé différent de tout l'espace $\mathbb{R}^n$ ($\ker B \neq \mathbb{R}^n$).

L'équivalence, établie dans la première partie, entre d'une part résoudre l'équation $A.x = b$ et d'autre part chercher le vecteur z rendant minimum la fonction f définie sur $\mathbb{R}^n$ par la relation suivante

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

conduit à se poser le problème suivant :

Soit B une matrice appartenant à $M_n(\mathbb{R})$ dont le noyau F est différent de $\mathbb{R}^n$; rechercher un vecteur $\overline{x}$ appartenant à F rendant minimum la restriction de la fonction f au sous-espace vectoriel F.

Existence du minimum de la fonction f dans F :

12. Démontrer que la fonction f possède la propriété suivante : pour tout réel c, il existe un réel ρ, tel que, pour tout vecteur x de F de norme supérieure ou égale à ρ ($\|x\| \geq \rho$), le réel $f(x)$ est supérieur ou égal à c ($f(x) \geq c$).

13. En déduire que, si y est un point de F, il existe un réel r tel que pour tout vecteur x de F de norme supérieure ou égale à r ($\|x\| \geq r$), $f(x)$ est supérieur ou égal à $f(y)$.

14. Démontrer à l'aide du résultat précédent qu'il existe au moins un vecteur $\overline{x}$ du sous-espace vectoriel F en lequel la restriction de la fonction f à ce sous-espace F atteint un minimum.

15. Démontrer qu'il existe un seul vecteur $\overline{x}$ en lequel la fonction f atteint son minimum dans F, en admettant que la fonction f est convexe ; c'est-à-dire : pour tout couple $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ de vecteurs et tout réel λ appartenant à l'intervalle ouvert $]0, 1[$, les valeurs prises par la fonction f vérifient la relation suivante :

$$f(\lambda\, x + (1-\lambda)\; y) \leq \lambda\, f(x) + (1-\lambda)\; f(y),$$

où l'inégalité est stricte si et seulement si les vecteurs x et y sont différents.

Propriétés du point $\overline{x}$:

16. Démontrer que, pour qu'un vecteur y de F rende minimum la restriction de la fonction f au sous-espace vectoriel F, il faut et il suffit que le vecteur $Ay-b$ soit orthogonal à ce sous-espace F de $\mathbb{R}^n$.

17. Démontrer que la valeur prise par la fonction f au point $\overline{x}$, en lequel elle atteint son minimum dans F, est donnée par la relation suivante :

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

Le Lagrangien L :

Soit L la fonction définie sur l'espace produit $\mathbb{R}^n \times \mathbb{R}^n$ par la relation suivante :

$$L(x,\ y) = f(x) + (y \mid Bx).$$

Un point $(x^*,\ y^*)$ de l'espace produit $\mathbb{R}^n \times \mathbb{R}^n$ est dit point selle de la fonction L, s'il possède la propriété suivante : quel que soit le point $(x,\ y)$ de l'espace produit $\mathbb{R}^n \times \mathbb{R}^n$, les valeurs prises par la fonction L aux points $(x^*,\ y)$, $(x^*,\ y^*)$ et $(x,\ y^*)$ vérifient la double inégalité suivante :

$$L(x^*,\ y) \leq L(x^*,\ y^*) \leq L(x,\ y^*).$$

Propriétés du Lagrangien et de ses points selles :

18. Établir l'inégalité suivante :

$$\sup_{y \in \mathbb{R}^n} \left(\inf_{x \in \mathbb{R}^n} L(x,\ y) \right) \leq \inf_{x \in \mathbb{R}^n} \left(\sup_{y \in \mathbb{R}^n} L(x,\ y) \right).$$

Il est supposé dans toute la suite qu'il existe un point selle $(x^*,\ y^*)$ de la fonction L.

19. Démontrer que la valeur prise par la fonction L en un point selle $(x^*,\ y^*)$ vérifie les égalités suivantes :

$$L(x^*,\ y^*) = \sup_{y \in \mathbb{R}^n} \left(\inf_{x \in \mathbb{R}^n} L(x,\ y) \right) = \inf_{x \in \mathbb{R}^n} \left(\sup_{y \in \mathbb{R}^n} L(x,\ y) \right).$$

20. Démontrer, pour tout point $(x_1,\ y_1)$ de $\mathbb{R}^n \times \mathbb{R}^n$, les équivalences suivantes :

$$\forall y \in \mathbb{R}^n,\quad L(x_1,\ y) \leq L(x_1,\ y_1) \Longleftrightarrow Bx_1 = 0.$$
$$\forall x \in \mathbb{R}^n,\quad L(x_1,\ y_1) \leq L(x,\ y_1) \Longleftrightarrow Ax_1 + {}^tBy_1 = b.$$

21. Soient x_1 un vecteur du sous-espace vectoriel F et y_1 un vecteur de $\mathbb{R}^n$. Démontrer qu'une condition nécessaire et suffisante pour que le couple $(x_1,\ y_1)$

soit un point selle du Lagrangien L est que le vecteur x_1 réalise le minimum de la restriction de la fonction f à F et que les vecteurs x_1 et y_1 vérifient la relation suivante :

$$Ax_1 + {}^tBy_1 = b.$$

La suite logique est la recherche d'un point selle du Lagrangien L.

Algorithme d'Uzawa : soit toujours $(x^*,\ y^*)$ un point selle, supposé exister ; étant donnés un vecteur y^0 arbitraire de $\mathbb{R}^n$, une suite $(\rho_m)_{m\in\mathbf{N}}$ de réels, qui seront précisés plus loin, soient $(x^m)_{m\in\mathbf{N}}$ et $(y^m)_{m\in\mathbf{N}}$ les deux suites de vecteurs définies par les conditions suivantes :

• Pour tout entier naturel m, le vecteur x^m est le vecteur qui rend minimum la fonction $x \longmapsto L(x,\ y^m)$.

• Pour tout entier naturel m, le vecteur y^{m+1} est défini par la relation suivante :

$$y^{m+1} = y^m + \rho_m\ Bx^m.$$

Existence des deux suites $(x^m)_{m\in\mathbf{N}}$ et $(y^m)_{m\in\mathbf{N}}$:

22. Démontrer que les conditions énoncées permettent de déterminer tous les termes de ces deux suites $(x^m)_{m\in\mathbf{N}}$ et $(y^m)_{m\in\mathbf{N}}$ et que les vecteurs de ces suites vérifient, pour tout entier naturel m, les relations suivantes :

$$A(x^m - x^*) + {}^tB(y^m - y^*) = 0,$$

$$y^{m+1} - y^* = y^m - y^* + \rho_m\ B(x^m - x^*).$$

où x^* et y^* sont les deux vecteurs d'un point selle de L.

23. En déduire l'égalité ci-dessous :

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

Convergence de la suite numérique de terme général $\|y^m - y^*\|^2$, $m \in \mathbb{N}$:

24. Un résultat préliminaire : démontrer l'existence d'une matrice carrée d'ordre n symétrique positive inversible, notée $A^{1/2}$, telle que :

$$\left(A^{1/2}\right)^2 = A.$$

Soit C la matrice définie par la relation suivante :

$$C = A^{-1/2}.{}^tB.B.A^{-1/2},$$

où la matrice $A^{-1/2}$ est la matrice inverse de la matrice $A^{1/2}$.

25. Démontrer que la matrice C est une matrice symétrique positive. Établir qu'il existe une constante ν telle que, pour tout vecteur u de $\mathbb{R}^n$, l'inégalité ci-dessous soit vraie :

$$\|Bu\|^2 \leq \nu \ (Au \mid u).$$

Soient α et β deux réels tels que le segment $[\alpha, \ \beta]$ soit contenu dans l'intervalle ouvert $]0, \ 2/\nu[$, $(0 < \alpha < \beta < 2/\nu)$. La suite des réels ρ_m est supposée vérifier pour tout entier naturel m l'inégalité suivante :

$$\alpha \leq \rho_m \leq \beta.$$

26. Démontrer que la suite de terme général $\|y^m - y^*\|^2$, $m \in N$ est monotone décroissante ; utiliser, pour simplifier, la suite $(u^m)_{m\in\mathbf{N}}$ dont le terme général est définie par la relation suivante :

$$u^m = x^m - x^*.$$

Convergence de la suite $(x^m)_{m\in\mathbf{N}}$:

27. En déduire la convergence et la limite de la suite $(x^m)_{m\in\mathbf{N}}$.

FIN DU PROBLÈME

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 : $\big(Mx \mid x\big) = \sum_{i=1}^{n} \lambda_i.(x)_i^2$. D'où, en appelant p (respectivement q) la plus petite (respectivement la plus grande) des valeurs propres de M , on a l'inégalité :

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

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 $\forall x \in \mathbb{R}^n \quad \big(Mx \mid x\big) \geqslant p.\|x\|^2 \geqslant 0$; d'autre part, 0 n'est alors pas valeur propre , donc $M.x = 0 \Rightarrow x = 0$, et M est injective, donc inversible.
Mais la condition est aussi nécessaire, car l'inégalité $\big(Mx \mid x\big) \geqslant 0$ doit être en particulier vraie pour les vecteurs propres (non nuls) de M , d'où, pour toute valeur propre λ_i de M , $\lambda_i.\|x\|^2 \geqslant 0$, soit $\lambda_i \geqslant 0$; de plus, l'inversibilité de M interdit la valeur propre 0 , donc les valeurs propres λ_i de M sont bien > 0 .

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

$$\|M.x\|^2 = \sum_{i=1}^{n} (Mx)_i^2 = \sum_{i=1}^{n} \lambda_i^2.x_i^2 \leqslant \Big(\operatorname*{Max}_{1\leqslant i\leqslant n} \lambda_i^2\Big).\sum_{i=1}^{n} (x)_i^2 = \Big(\operatorname*{Max}_{1\leqslant i\leqslant n} \lambda_i^2\Big).\|x\|^2 \quad .$$

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

$$(1) \qquad \forall x \in \mathbb{R}^n \quad \|x\| = 1 \Rightarrow \|M.x\| \leqslant \operatorname*{Max}_{1\leqslant i\leqslant n} |\lambda_i| \quad ;$$

Mais la première inégalité ci-dessus devient une égalité si x est un vecteur propre associé à la valeur propre λ_{i_0} telle que $|\lambda_{i_0}| = \operatorname*{Max}_{1\leqslant i\leqslant n} |\lambda_i|$. Comme un vecteur propre, non nul, peut toujours être normalisé, on voit que dans l'implication (1) on peut obtenir l'égalité à droite. Donc le nombre $\operatorname*{Max}_{1\leqslant i\leqslant n} |\lambda_i|$ est le plus petit des majorants possibles, soit par définition, $N(M) = \operatorname*{Max}_{1\leqslant i\leqslant n} |\lambda_i|$.

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 , ... , 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 $y^k = S^k.y^0$. Par propriété de la norme N , $\|S.y\| \leqslant N(S).\|y\|$, on a donc $\|y^k\| \leqslant \big(N(S)\big)^k.\|y^0\|$.
De $N(S) \in]0,1[$ on conclut alors que $y^k \to 0$, 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°) Par définition, une dérivée partielle s'obtient en cherchant la limite de $\frac{1}{t}\big(f(x+te_k) - f(x)\big)$ lorsque $t \to 0$. Par le calcul précédent, on a

$$\frac{1}{t}\big(f(x+te_k) - f(x)\big) = \big(Ax - b \mid e_k\big) + \frac{t}{2}\big(Ae_k \mid e_k\big) \quad ,$$

et il en résulte clairement l'existence de la dérivée partielle $\dfrac{\partial f}{\partial x_k}$, égale à $\big(Ax - b \mid e_k\big)$.

7°) On sait qu'en base orthonormée les composantes d'un vecteurs s'expriment par les produits scalaires avec les vecteurs de base, donc, vu la définition de g et le calcul ci-dessus, on a $g(x) = A.x - b$.

8°) Cela a été vu en **1.** : par **5.**, $I(x,u) = \frac{1}{2}(Au \mid u)$ où la matrice $\frac{1}{2}A$ est symétrique définie positive, donc l'encadrement $r\|u\|^2 \leqslant I(x,u) \leqslant s\|x\|^2$ 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 !), c'est nécessairement un point critique.
Condition suffisante : si $g(x) = 0$, on a pour tout $u \in \mathbb{R}^n$, puisque $I(x,u) \geqslant 0$, $f(x+u) = f(x) + I(x,u) \geqslant f(x)$, donc $f(x)$ est un minimum absolu.

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

$$f(x - \alpha g(x)) - f(x) = -\alpha(g(x) \mid g(x)) + \frac{\alpha^2}{2}(Ag(x) \mid g(x)) \leqslant \left(-\alpha + \frac{\alpha^2}{2}\lambda_n\right).\|g(x)\|^2 \quad .$$

L'hypothèse $0 < \alpha < \frac{2}{\lambda_n}$ entraîne $-\alpha + \frac{\alpha^2}{2}\lambda_n < 0$, donc $f(x - \alpha g(x)) - f(x) \leqslant 0$ (l'ingalit est mme 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.

Deuxième partie

12°) Par **1.** on a une inégalité de la forme : $\forall x \in \mathbb{R}^n \;\; r.\|x\|^2 \leqslant \frac{1}{2}(Ax \mid x)$, $r = \frac{1}{2} \underset{\lambda \in sp(A)}{\mathrm{Min}}$ λ étant strictement positif. De plus, par Cauchy-Schwarz, $(b \mid x) \leqslant \|b\|.\|x\|$. D'où finalement :

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

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, soit $y \in F$ fixé. Par ci-dessus, r étant choisi tel que $\|x\| \geqslant r \Rightarrow f(x) \geqslant f(y)$, le minimum de $f_{|F}$ est à chercher dans $K = F \bigcap \bar{B}(0,r)$. Or, en dimension finie, un sous-espace vectoriel est un fermé, ainsi qu'une boule fermée, et de même pour leur intersection. Donc K , fermé et borné d'un espace vectoriel de dimension finie, est compact. f est une application continue (car polynômiale en les composantes), donc elle admet un minimum $\bar{x}$ atteint sur K , cqfd.

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

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

ce qui, puisque $\frac{1}{2}\bar{x} + \frac{1}{2}\bar{y}$ reste dans F , contredit que $\bar{x}$ est un minimum pur $f_{|F}$. Donc le minimum, qui existe par **14.**, est unique.

Remarque : f est la somme d'une fonction quadratique $x \mapsto \frac{1}{2}(Ax \mid x)$ et d'une fonction linéaire $x \mapsto -(b \mid x)$. Une fonction linéaire étant évidemment convexe (au sens large), la stricte convexité de f résulte de celle de $x \mapsto (Ax \mid x)$. Or celle-ci découle simplement de la définie positivité de la matrice symétrique A . On a en effet, pour tout $\lambda \in]0,1[$ et pour tout x, y de $\mathbb{R}^n$ tels que $x \neq y$:

$$(A((1-\lambda)x + \lambda y) \mid (1-\lambda)x + \lambda y) \overset{?}{<} (1-\lambda)(Ax \mid x) + \lambda(Ay \mid y)$$

$$\Longleftrightarrow \lambda(1-\lambda)(Ax \mid x) - 2\lambda(1-\lambda)(Ax \mid y) + \lambda(1-\lambda)(Ay \mid y) \overset{?}{>} 0$$

$$\Longleftrightarrow \lambda(1-\lambda)(A(x-y) \mid x-y) \overset{?}{>} 0 \quad ,$$

et cette dernière inégalité est acquise par hypothèse sur λ et définie- positivité de A .

16°) g , calculé à la question **7.**, est le gradient de f , et on sait que, pour tout vecteur u , $\frac{\partial f}{\partial u}(y) = df_y(u) = (g(y) \mid u)$.
Or, $\frac{\partial f}{\partial u}(y)$ représente la dérivée en $t = 0$ de la fonction $\varphi : t \mapsto f(y + t.u)$, donc si $y \in F$ est un minimum pour $f_{|F}$, et si $u \in F$, φ admet un minimum absolu en 0 . Il en résulte bien $0 = \varphi'(0) = (g(y) \mid u) = (Ay - b \mid u)$.
Réciproquement, si $Ay - b$ est orthogonal à F , on a comme à la question **9.** que $\forall u \in F \ \ f(y+u) - f(y) = I(y,u) = \frac{1}{2}(Au \mid u) \geqslant 0$, et $f(y)$ est bien un minimum de $f_{|F}$.

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

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

18°) Petite diffuculté non signalée ici : les inf et sup sont à prendre dans $\bar{R}$: 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 $\bar{R}$. 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 :

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

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

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

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

L'inégalité (1) donne, puisque le sup est plus petit qu'un majorant, $\operatorname*{Sup}_{y\in\mathbb{R}^n} L(x^*,y) \leqslant L(x^*,y^*)$. Puis, comme l'inf minore une instance particulière, $\operatorname*{Inf}_{x\in\mathbb{R}^n} \operatorname*{Sup}_{y\in\mathbb{R}^n} L(x,y) \leqslant L(x^*,y^*)$. De même, l'inégalité (2) donne $L(x^*,y^*) \leqslant \operatorname*{Inf}_{x\in\mathbb{R}^n} L(x,y^*)$. puis $L(x^*,y^*) \leqslant \operatorname*{Sup}_{y\in\mathbb{R}^n} \operatorname*{Inf}_{x\in\mathbb{R}^n} L(x,y)$. 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

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

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 CNS est bien $Bx_1 = 0$.
• Par propriété du produit scalaire canonique, on a

$$(y_1 \mid Bx) = {}^t y_1.Bx = {}^t({}^t y_1.Bx) = {}^t x\,{}^t By_1 = ({}^t By_1 \mid x) \ ,$$

On voit donc qu' y_1 fix $L(x,y_1) = \frac{1}{2}(Ax \mid x) - (b \mid x) + ({}^t By_1 \mid x)$ est une fonction f o on a remplac 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 proprit demande.

21°) La définition du point selle donne par les équivalences ci-dessus où on a remplacé (x_1,y_1) par (x^*,y^*) :

$$(x^*,y^*) \text{ point selle } \Rightarrow Bx^* = 0 \ et \ Ax^* + {}^t By^* = b \ .$$

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 $\forall x \in \mathbb{R}^n \ \ L(x^*,y^*) \leqslant L(x,y^*)$ implique bien que $f(x^*)$ est un minimum de $f_{|F}$.
Réciproquement, si $x_1 \in F$ réalise un minimum pour $f_{|F}$, on a $Bx_1 = 0$, et si de plus $Ax_1 + {}^t By_1 = b$ la question **20.** assure que (x_1,y_1) est un point selle par définition du point selle.

22°) • A $y = y^m$ fixé, l'application $x \mapsto L(x, y^m) = \frac{1}{2}(Ax \mid x) - (b - {}^tBy^m \mid x)$ 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

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

Mais par **22.**

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

et donc finalement

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

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 $(A^{1/2})^2 = A$.
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 $x \in \mathbb{R}^n$., $(Cx \mid x) = (A^{-1/2}.{}^tB.B.A^{-1/2}.x \mid x) = (B.A^{-1/2}.x \mid B.A^{-1/2}.x) = \|B.A^{-1/2}.x\|^2 \geqslant 0$, donc C est positive.
Pour tout $u \in \mathbb{R}^n$, posons $x = A^{1/2}.u \Leftrightarrow u = A^{-1/2}x$. Par le calcul ci-dessus et par **1.**,

$$\|Bu\|^2 = (Cx \mid x) \leqslant \Big(\operatorname*{Sup}_{\mu \in sp(C)} \mu\Big).(x \mid x) = \nu.(A^{1/2}u \mid A^{1/2}u) = \nu.(Au \mid u) \ ,$$

où on a posé $\nu = \operatorname*{Sup}_{\mu \in sp(C)} \mu$.

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

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

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

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

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

*
* *