



BANQUE COMMUNE D'ÉPREUVES

CODE ÉPREUVE :

280

HEC_M1_S

Concepteur : H.E.C.

OPTION SCIENTIFIQUE

MATHÉMATIQUES I

Mercredi 30 avril 2008, de 8 h. à 12 h.

La présentation, la lisibilité, l'orthographe, la qualité de la rédaction, la clarté et la précision des raisonnements entreront pour une part importante dans l'appréciation des copies.

Les candidats sont invités à encadrer dans la mesure du possible les résultats de leurs calculs.

Ils ne doivent faire usage d'aucun document : l'utilisation de toute calculatrice et de tout matériel électronique est interdite.

Seule l'utilisation d'une règle graduée est autorisée.

Dans tout le problème, n et p désignent deux entiers vérifiant $1 \leq p \leq n$. On note $\mathcal{M}_{n,p}(\mathbb{R})$ l'espace vectoriel des matrices à n lignes et p colonnes, à coefficients réels. La transposée d'une matrice A de $\mathcal{M}_{n,p}(\mathbb{R})$ est notée tA . Lorsqu'une matrice A est inversible, on note A^{-1} son inverse.

Dans tout le problème, on identifie les deux espaces vectoriels $\mathcal{M}_{n,1}(\mathbb{R})$ (respectivement $\mathcal{M}_{p,1}(\mathbb{R})$) et \mathbb{R}^n (resp. \mathbb{R}^p), c'est-à-dire qu'on identifie un vecteur (point) de \mathbb{R}^n (resp. \mathbb{R}^p) avec le vecteur-colonne de ses coordonnées dans la base canonique de \mathbb{R}^n (resp. \mathbb{R}^p).

On munit \mathbb{R}^n (resp. \mathbb{R}^p) de sa structure euclidienne canonique, et pour tous vecteurs u et v de \mathbb{R}^n (resp. \mathbb{R}^p), on note $\langle u, v \rangle = {}^t uv$ leur produit scalaire, et $\|u\|$ la norme de u associée.

Pour tout i de $[[1, n]]$, on note f_i une fonction définie sur \mathbb{R}^p à valeurs réelles, et de classe C^2 sur \mathbb{R}^p . Soit F la

fonction définie sur \mathbb{R}^p , à valeurs réelles, par : $F(x_1, x_2, \dots, x_p) = \frac{1}{2} \sum_{i=1}^n [f_i(x_1, x_2, \dots, x_p)]^2$.

Autrement dit, si $X = (x_1, \dots, x_p)$ est un point de \mathbb{R}^p , on a : $F(X) = \frac{1}{2} \sum_{i=1}^n f_i^2(X) = \frac{1}{2} \|f(X)\|^2$, en notant $f(X)$ le vecteur $(f_1(X), \dots, f_n(X))$.

Le problème a pour objet l'étude de quelques aspects mathématiques liés à la recherche du minimum de la fonction F .

Partie I. Gradient et hessienne

Pour tout point $X = (x_1, x_2, \dots, x_p)$ de \mathbb{R}^p , on rappelle que :

- le gradient de F au point X , noté $\nabla F(X)$, est le vecteur de \mathbb{R}^p suivant :

$$\nabla F(X) = \left(\frac{\partial F}{\partial x_1}(X), \dots, \frac{\partial F}{\partial x_p}(X) \right)$$

- la *matrice hessienne* de F au point X , notée $\nabla^2 F(X)$, est la matrice symétrique de $\mathcal{M}_p(\mathbb{R})$ suivante :

$$\nabla^2 F(X) = \left(\frac{\partial^2 F}{\partial x_k \partial x_j}(X) \right)_{1 \leq k, j \leq p}$$

Pour tout point $X = (x_1, \dots, x_p)$ de \mathbb{R}^p , on note $J(X)$ la matrice de $\mathcal{M}_{n,p}(\mathbb{R})$ définie par :

$$J(X) = \left(\frac{\partial f_i}{\partial x_j}(X) \right)_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$$

dans laquelle i désigne l'indice de ligne et j l'indice de colonne. On pose : $G(X) = {}^t J(X) J(X)$.

Si X est un point de \mathbb{R}^p vérifiant $\nabla F(X) \neq 0$, on dit qu'un vecteur h de \mathbb{R}^p est une *direction de décroissance* de F en X , si on a : $\langle \nabla F(X), h \rangle < 0$.

Dans les trois exemples suivants, on suppose que p est égal à 2.

1. Un premier exemple.

On considère les deux fonctions f_1 et f_2 définies sur \mathbb{R}^2 par : $f_1(x_1, x_2) = x_1^2 + x_2 + 1$, et $f_2(x_1, x_2) = x_1 + x_2^2 + 1$.

- Justifier que F est de classe C^2 sur \mathbb{R}^2 . Calculer, en tout point (x_1, x_2) de \mathbb{R}^2 , le gradient $\nabla F(x_1, x_2)$.
- Montrer que le système d'équations qui permet de déterminer les éventuels points critiques de F , peut se mettre sous la forme suivante :

$$\begin{cases} 2x_1^3 + 2x_1x_2 + 3x_1 + x_2^2 + 1 = 0 \\ (x_1 - x_2)(2x_1^2 + 2x_1x_2 + 2x_2^2 - x_1 - x_2 + 3) = 0 \end{cases}$$

- Établir, pour tout (x_1, x_2) de \mathbb{R}^2 , l'inégalité : $2x_1^2 + 2x_1x_2 + 2x_2^2 - x_1 - x_2 + 3 > 0$. En déduire que l'unique point critique de F est $(-1/2, -1/2)$.
- Déterminer, en tout point (x_1, x_2) de \mathbb{R}^2 , la matrice hessienne $\nabla^2 F(x_1, x_2)$. En déduire que F admet un minimum local en $(-1/2, -1/2)$.
- On note pour tout point X de \mathbb{R}^2 , $\nabla^2 f_1(X)$ et $\nabla^2 f_2(X)$ respectivement, les matrices hessiennes de f_1 et f_2 au point X . Préciser la matrice $J(X)$. Exprimer ${}^t J(X) f(X)$ et $G(X) + \sum_{i=1}^2 f_i(X) \nabla^2 f_i(X)$ en fonction de $\nabla F(X)$ et $\nabla^2 F(X)$ respectivement.

2. Un deuxième exemple.

Soit $a = (a_1, \dots, a_n)$, $b = (b_1, \dots, b_n)$ et $c = (c_1, \dots, c_n)$ trois vecteurs non nuls donnés de \mathbb{R}^n , tels que la famille (a, b) soit libre.

Pour tout i de $\llbracket 1, n \rrbracket$, la fonction f_i est définie sur \mathbb{R}^2 par : $f_i(x_1, x_2) = a_i x_1 + b_i x_2 - c_i$.

- Exprimer, pour tout point (x_1, x_2) de \mathbb{R}^2 , le gradient $\nabla F(x_1, x_2)$ à l'aide de $x_1, x_2, \|a\|, \|b\|, \langle a, b \rangle, \langle a, c \rangle$ et $\langle b, c \rangle$.
- Justifier l'inégalité : $\|a\|^2 \times \|b\|^2 - \langle a, b \rangle^2 > 0$. En déduire que la fonction F possède un unique point critique $(\widehat{x}_1, \widehat{x}_2)$.

Exprimer \widehat{x}_1 et \widehat{x}_2 en fonction de $\|a\|, \|b\|, \langle a, b \rangle, \langle a, c \rangle$ et $\langle b, c \rangle$.

- Calculer, en tout point (x_1, x_2) de \mathbb{R}^2 , la matrice hessienne $\nabla^2 F(x_1, x_2)$; en déduire que F admet un minimum local en $(\widehat{x}_1, \widehat{x}_2)$.
- En utilisant la structure euclidienne de \mathbb{R}^n , montrer que F admet un minimum global en $(\widehat{x}_1, \widehat{x}_2)$.

3. Un troisième exemple.

On suppose que c_1, c_2, \dots, c_n sont n réels donnés non tous égaux. On note \bar{c} et s^2 respectivement, la moyenne arithmétique et la variance de la série statistique $(c_i)_{1 \leq i \leq n}$.

Pour tout i de $\llbracket 1, n \rrbracket$, la fonction f_i est définie sur \mathbb{R}^2 par : $f_i(x_1, x_2) = x_1 + x_2 - c_i$.

- Déterminer les points critiques de F .

b) Soit $(\widehat{x}_1, \widehat{x}_2)$ un point critique de F . Exprimer $F(\widehat{x}_1, \widehat{x}_2)$ en fonction de s^2 . Montrer, pour tout (x_1, x_2) de \mathbb{R}^2 , l'égalité : $F(x_1, x_2) - F(\widehat{x}_1, \widehat{x}_2) = \frac{n}{2}(x_1 + x_2 - \bar{c})^2$.

c) En déduire la nature des points critiques de F . Ce résultat était-il prévisible ?

4. Retour au cas général.

Soit $X = (x_1, \dots, x_p)$ un point de \mathbb{R}^p .

a) Exprimer $\nabla F(X)$ en fonction de ${}^t J(X)$ et de $f(X)$.

b) Pour tout i de $\llbracket 1, n \rrbracket$, on note $\nabla^2 f_i(X)$ la matrice hessienne de f_i au point X .

Établir la formule : $\nabla^2 F(X) = G(X) + \sum_{i=1}^n f_i(X) \nabla^2 f_i(X)$.

Partie II. Une approximation de F

Dans cette partie, on conserve les définitions et les notations de la partie I, et on suppose que X est un vecteur fixé de \mathbb{R}^p vérifiant : $\nabla F(X) \neq 0$.

Pour tout vecteur $h = (h_1, h_2, \dots, h_p)$ de \mathbb{R}^p , on pose : $\ell(h) = f(X) + J(X)h$ et $L(h) = \frac{1}{2} \|\ell(h)\|^2$.

1. Établir, pour tout h de \mathbb{R}^p , l'égalité : $L(h) = F(X) + {}^t h \nabla F(X) + \frac{1}{2} {}^t h G(X) h$.

2. Soit P une matrice symétrique de $\mathcal{M}_p(\mathbb{R})$.

a) Justifier que P est diagonalisable.

b) On note $\theta_1, \dots, \theta_p$ les valeurs propres de P , et on pose : $\theta = \max_{1 \leq j \leq p} |\theta_j|$. Montrer, pour tout vecteur h de \mathbb{R}^p , l'inégalité suivante : $|{}^t h P h| \leq \theta \|h\|^2$.

3. a) Écrire un développement limité à l'ordre 2 de la fonction F au point X .

b) En déduire, à l'aide de la question 2.b, que l'on a : $\lim_{\|h\| \rightarrow 0} \frac{F(X+h) - L(h)}{\|h\|} = 0$.

Pour X fixé de \mathbb{R}^p , on dit que $L(h)$ est une approximation à l'ordre 2 de $F(X+h)$ lorsque $\|h\|$ tend vers 0.

4. On note : $G(X) = (g_{i,j}(X))_{1 \leq i,j \leq p}$. Soit φ_1 et φ_2 deux fonctions définies sur \mathbb{R}^p par : $\varphi_1(h) = {}^t h \nabla F(X)$ et $\varphi_2(h) = {}^t h G(X) h$.

a) Montrer que pour tout j de $\llbracket 1, p \rrbracket$, on a : $\frac{\partial \varphi_1}{\partial h_j}(h) = \frac{\partial F}{\partial x_j}(X)$ et $\frac{\partial \varphi_2}{\partial h_j}(h) = 2 \sum_{i=1}^p g_{i,j}(X) h_i$.

b) En déduire que le gradient $\nabla L(h)$ de L en h , est donné par : $\nabla L(h) = \nabla F(X) + G(X)h$.

c) Soit $\nabla^2 L(h)$ la matrice hessienne de L en h . Établir la formule : $\nabla^2 L(h) = G(X)$.

5. Soit J une matrice de $\mathcal{M}_{n,p}(\mathbb{R})$.

a) Montrer que la matrice ${}^t J J$ est diagonalisable et que ses valeurs propres sont positives ou nulles.

b) Montrer que lorsque la matrice ${}^t J J$ est inversible, le rang de la matrice J est égal à p .

6. Montrer que si la fonction L admet des points critiques \widehat{h} , alors ceux-ci vérifient l'inéquation : $\langle \widehat{h}, \nabla F(X) \rangle \leq 0$.

7. On suppose que la matrice $G(X)$ est inversible.

a) Montrer que L admet un unique point critique \widehat{h} donné par : $\widehat{h} = -(G(X))^{-1} \times {}^t J(X) f(X)$.

b) Établir que \widehat{h} est une direction de décroissance de F en X . En déduire que L admet un minimum local en \widehat{h} .

Partie III. Une décomposition d'une matrice rectangulaire

Afin de réduire les inconvénients liés à l'inversion de la matrice $G(X)$, on remplace celle-ci par la matrice $G(X) + \mu I$, où μ désigne un paramètre réel strictement positif, et I la matrice identité d'ordre p . Certains résultats d'algèbre linéaire permettent alors de substituer à l'inversion d'une matrice, le calcul plus simple d'une somme de matrices.

Soit J une matrice non nulle de $\mathcal{M}_{n,p}(\mathbb{R})$.

1. Montrer qu'il existe une matrice V orthogonale de $\mathcal{M}_p(\mathbb{R})$, un entier q tel que $1 \leq q \leq p$, et des réels $\lambda_1, \lambda_2, \dots, \lambda_q$ tels que $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_q > 0$, qui vérifient l'égalité : ${}^tV^tJJV = D$, où $D = (d_{i,j})_{1 \leq i,j \leq p}$ est définie par : $d_{i,i} = \lambda_i$ si $1 \leq i \leq q$, et $d_{i,j} = 0$ sinon. Si $q < p$, on pose : $\lambda_{q+1} = \dots = \lambda_p = 0$.

Pour tout i de $\llbracket 1, p \rrbracket$, on note V_i la i -ième colonne de V .

2. a) Montrer que le rang de tJJ est égal à q .

b) Montrer que, pour tout i de $\llbracket 1, q \rrbracket$, JV_i est un vecteur propre de la matrice J^tJ associé à la valeur propre λ_i . En déduire que les matrices tJJ et J^tJ ont les mêmes valeurs propres non nulles.

c) Soit (Y_1, \dots, Y_r) une base du sous-espace propre de tJJ associée à une valeur propre λ non nulle. Montrer que la famille (JY_1, \dots, JY_r) est une famille libre de $\mathcal{M}_{n,1}(\mathbb{R})$.

d) En déduire que les sous-espaces propres de tJJ et de J^tJ associés à la même valeur propre non nulle sont de même dimension, et que le rang de J^tJ est égal à q .

3. On pose, pour tout i de $\llbracket 1, q \rrbracket$: $U_i = \frac{1}{\sqrt{\lambda_i}}JV_i$.

a) Montrer que la famille (U_1, \dots, U_q) est une famille orthonormée de vecteurs propres de J^tJ .

b) En déduire qu'il existe une base orthonormée $(U_1, \dots, U_q, U_{q+1}, \dots, U_n)$ de $\mathcal{M}_{n,1}(\mathbb{R})$, formée de vecteurs propres de J^tJ .

4. On note U la matrice de $\mathcal{M}_n(\mathbb{R})$ telle que, pour tout i de $\llbracket 1, n \rrbracket$, la i -ième colonne de U est la matrice-colonne U_i de $\mathcal{M}_{n,1}(\mathbb{R})$.

Soit $S = (s_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$ la matrice de $\mathcal{M}_{n,p}(\mathbb{R})$ définie par : $s_{i,i} = \sqrt{\lambda_i}$ si $1 \leq i \leq p$ et $s_{i,j} = 0$ sinon.

Établir l'égalité matricielle suivante : $S = {}^tUJV$. En déduire l'égalité : $J = US^tV$.

5. a) Montrer que la matrice $({}^tJJ + \mu I)$ est inversible.

b) On note $R = (r_{i,j})_{\substack{1 \leq i \leq p \\ 1 \leq j \leq n}}$ la matrice de $\mathcal{M}_{p,n}(\mathbb{R})$ définie par : $r_{i,i} = \frac{\sqrt{\lambda_i}}{\lambda_i + \mu}$ si $1 \leq i \leq p$ et $r_{i,j} = 0$ sinon.

Établir la formule suivante : $({}^tJJ + \mu I)^{-1} \times {}^tJ = VR^tU$.

c) En déduire l'égalité : $({}^tJJ + \mu I)^{-1} \times {}^tJ = \sum_{i=1}^q \frac{\sqrt{\lambda_i}}{\lambda_i + \mu} V_i {}^tU_i$

6. Soit X un vecteur **fixé** de \mathbb{R}^p vérifiant : $\nabla F(X) \neq 0$.

Pour tout vecteur h de \mathbb{R}^p , on pose : $M(h) = L(h) + \frac{\mu}{2} \|h\|^2$.

a) Montrer que : $\lim_{\|h\| \rightarrow 0} \frac{F(X+h) - M(h)}{\|h\|} = 0$.

b) Calculer, pour tout h de \mathbb{R}^p , le gradient $\nabla M(h)$ et la matrice hessienne $\nabla^2 M(h)$ de M en h .

c) En appliquant les résultats des questions précédentes à la matrice $J(X)$, montrer que M admet un unique point critique h^* . Donner une expression de h^* qui utilise les résultats de la question 5.c.

d) Montrer que M admet un minimum local en h^* .

À partir de ce minimum local h^* de M (ou du minimum local \hat{h} de L), on pourrait utiliser une méthode algorithmique permettant, sous certaines conditions, d'approcher avec une précision donnée un minimum local de la fonction F