



PRÉPASIMPLE

prepasimple.fr

HEC

Mathématiques 1

2008

Sujet de mathématiques approfondies

Mathématiques approfondies — Prépa ECG

Annales de concours · Sujet

Préambule et notations

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

On identifie $\mathcal{M}_{n,1}(\mathbb{R})$ à $\mathbb{R}^n$ et $\mathcal{M}_{p,1}(\mathbb{R})$ à $\mathbb{R}^p$: un vecteur est identifié à la colonne de ses coordonnées dans la base canonique.

Ces espaces sont munis de leur structure euclidienne canonique. On note $\langle u, v \rangle = {}^tu v$ leur produit scalaire et $\|u\|$ la norme associée.

Pour $1 \leq i \leq n$, soit $f_i : \mathbb{R}^p \rightarrow \mathbb{R}$ une fonction de classe $\mathcal{C}^2$. On définit $F : \mathbb{R}^p \rightarrow \mathbb{R}$ par

$$F(x_1, \dots, x_p) = \frac{1}{2} \sum_{i=1}^n [f_i(x_1, \dots, x_p)]^2.$$

Ainsi, pour $X \in \mathbb{R}^p$,

$$F(X) = \frac{1}{2} \sum_{i=1}^n f_i(X)^2 = \frac{1}{2} \|f(X)\|^2,$$

où $f(X) = {}^t(f_1(X), \dots, f_n(X))$.

Le problème étudie la recherche du minimum de F .

Partie I. Gradient et hessienne

Pour $X = (x_1, \dots, x_p)$, le gradient et la hessienne de F sont

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

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

La hessienne est symétrique. On note $J(X) \in \mathcal{M}_{n,p}(\mathbb{R})$ la matrice

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

où i est l'indice de ligne et j celui de colonne. On pose $G(X) = {}^t(J(X))J(X)$.

Si $\nabla F(X) \neq 0$, on dit que $h \in \mathbb{R}^p$ est une direction de décroissance de F en X lorsque $\langle \nabla F(X), h \rangle < 0$.

Dans les trois exemples suivants, $p = 2$.

1. Un premier exemple.

On considère $f_1(x_1, x_2) = x_1^2 + x_2 + 1$ et $f_2(x_1, x_2) = x_1 + x_2^2 + 1$, et $F = \frac{1}{2}(f_1^2 + f_2^2)$.

(a) Justifier que F est de classe $\mathcal{C}^2$ sur $\mathbb{R}^2$. Calculer son gradient en tout point.

(b) Montrer que le système déterminant les points critiques de F peut s'écrire

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

(c) Établir, pour tout $(x_1, x_2) \in \mathbb{R}^2$, l'inégalité

$$2x_1^2 + 2x_1x_2 + 2x_2^2 - x_1 - x_2 + 3 > 0.$$

En déduire que l'unique point critique est $(-1/2, -1/2)$.

(d) Déterminer la hessienne $\nabla^2 F(x_1, x_2)$ en tout point. En déduire que F admet un minimum local en $(-1/2, -1/2)$.

(e) On note $\nabla^2 f_1(X)$ et $\nabla^2 f_2(X)$ les hessiennes de f_1, f_2 . Préciser $J(X)$ et 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.

Soient $a = (a_1, \dots, a_n)$, $b = (b_1, \dots, b_n)$, $c = (c_1, \dots, c_n)$ trois vecteurs non nuls de $\mathbb{R}^n$, tels que (a, b) soit une famille libre. Pour $1 \leq i \leq n$, on pose $f_i(x_1, x_2) = a_i x_1 + b_i x_2 - c_i$.

- (a) Exprimer $\nabla F(x_1, x_2)$ à l'aide de $x_1, x_2, \|a\|, \|b\|, \langle a, b \rangle, \langle a, c \rangle$ et $\langle b, c \rangle$.
- (b) Justifier $\|a\|^2 \|b\|^2 - \langle a, b \rangle^2 > 0$. En déduire que F possède un unique point critique $(\hat{x}_1, \hat{x}_2)$ et exprimer ses coordonnées en fonction des quantités précédentes.
- (c) Calculer la hessienne en tout point. En déduire que F admet un minimum local en $(\hat{x}_1, \hat{x}_2)$.
- (d) En utilisant la structure euclidienne de $\mathbb{R}^n$, montrer que F admet un minimum global en $(\hat{x}_1, \hat{x}_2)$.

3. Un troisième exemple.

Soient $c_1, \dots, c_n$ des réels non tous égaux. On note

$$\bar{c} = \frac{1}{n} \sum_{i=1}^n c_i$$

et

$$s^2 = \frac{1}{n} \sum_{i=1}^n (c_i - \bar{c})^2$$

la moyenne et la variance de cette série statistique. Pour $1 \leq i \leq n$, on pose

$$f_i(x_1, x_2) = x_1 + x_2 - c_i.$$

- (a) Déterminer les points critiques de F .
- (b) Soit $(\hat{x}_1, \hat{x}_2)$ un point critique. Exprimer $F(\hat{x}_1, \hat{x}_2)$ en fonction de s^2 et montrer que, pour tout (x_1, x_2) ,

$$F(x_1, x_2) - F(\hat{x}_1, \hat{x}_2) = \frac{n}{2} (x_1 + x_2 - \bar{c})^2.$$

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

4. Retour au cas général.

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

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

(b) En notant $\nabla^2 f_i(X)$ la hessienne de f_i , établir

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

Partie II. Une approximation de F

On conserve les définitions précédentes. Dans cette partie, $X \in \mathbb{R}^p$ est fixé et vérifie $\nabla F(X) \neq 0$. Pour $h \in \mathbb{R}^p$, on pose

$$\ell(h) = f(X) + J(X)h, \quad L(h) = \frac{1}{2} \|\ell(h)\|^2.$$

1. Établir, pour tout h ,

$$L(h) = F(X) + {}^t h \nabla F(X) + \frac{1}{2} {}^t h G(X) h.$$

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

(a) Justifier que P est diagonalisable.

(b) On note $\theta_1, \dots, \theta_p$ ses valeurs propres et $\theta = \max_{1 \leq j \leq p} |\theta_j|$. Montrer que, pour tout $h \in \mathbb{R}^p$,

$$|{}^t h P h| \leq \theta \|h\|^2.$$

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

(b) À l'aide de 2(b), en déduire

$$\lim_{\|h\| \rightarrow 0} \frac{F(X+h) - L(h)}{\|h\|} = 0.$$

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, au sens de la limite précédente.

4. On note $G(X) = (g_{i,j}(X))_{1 \leq i,j \leq p}$ et on définit $\varphi_1(h) = {}^t h \nabla F(X)$, $\varphi_2(h) = {}^t h G(X) h$.

(a) Montrer que, pour $1 \leq j \leq p$,

$$\frac{\partial \varphi_1}{\partial h_j}(h) = \frac{\partial F}{\partial x_j}(X), \quad \frac{\partial \varphi_2}{\partial h_j}(h) = 2 \sum_{i=1}^p g_{i,j}(X) h_i.$$

(b) En déduire $\nabla L(h) = \nabla F(X) + G(X)h$.

(c) Établir $\nabla^2 L(h) = G(X)$.

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

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

(b) Montrer que, si ${}^t J J$ est inversible, alors J est de rang p .

6. Montrer que, si L admet des points critiques $\hat{h}$, ils vérifient $\langle \hat{h}, \nabla F(X) \rangle \leq 0$.

7. On suppose $G(X)$ inversible.

(a) Montrer que L admet un unique point critique

$$\hat{h} = -G(X)^{-1} {}^t (J(X)) f(X).$$

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

Partie III. Une décomposition d'une matrice rectangulaire

Pour éviter certains inconvénients liés à l'inversion de $G(X)$, on la remplace par $G(X) + \mu I_p$, où $\mu > 0$. La décomposition étudiée permettra de remplacer le calcul d'un inverse suivi d'un produit par une somme de matrices.

Soit $J \neq 0$ une matrice de $\mathcal{M}_{n,p}(\mathbb{R})$.

1. Montrer qu'il existe $V \in \mathcal{M}_p(\mathbb{R})$ vérifiant $V^{-1} = {}^tV$, un entier q tel que $1 \leq q \leq p$ et des réels $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_q > 0$, tels que

$${}^tV^t J J V = D,$$

où $D = (d_{i,j})$ est définie par $d_{i,i} = \lambda_i$ pour $1 \leq i \leq q$, et $d_{i,j} = 0$ sinon. Si $q < p$, on pose $\lambda_{q+1} = \dots = \lambda_p = 0$. On note V_i la i -ème colonne de V .

2. (a) Montrer que $\text{rg}({}^t J J) = q$.

(b) Pour $1 \leq i \leq q$, montrer que $J V_i$ est un vecteur propre de $J^t J$ associé à λ_i . En déduire que ${}^t J J$ et $J^t J$ ont les mêmes valeurs propres non nulles.

- (c) Soit $(Y_1, \dots, Y_r)$ une base de l'espace propre de ${}^t J J$ associé à une valeur propre non nulle λ . Montrer que $(J Y_1, \dots, J Y_r)$ est une famille libre de $\mathcal{M}_{n,1}(\mathbb{R})$.

- (d) En déduire que les espaces propres de ${}^t J J$ et $J^t J$ associés à une même valeur propre non nulle ont la même dimension, et que $\text{rg}(J^t J) = q$.

3. Pour $1 \leq i \leq q$, on pose $U_i = J V_i / \sqrt{\lambda_i}$.

(a) Montrer que $(U_1, \dots, U_q)$ est une famille orthonormale de vecteurs propres de $J^t J$.

(b) En déduire qu'il existe une base orthonormale $(U_1, \dots, U_q, U_{q+1}, \dots, U_n)$ de $\mathcal{M}_{n,1}(\mathbb{R})$, formée de vecteurs propres de $J^t J$.

4. Soit $U \in \mathcal{M}_n(\mathbb{R})$ de colonnes U_i . On note $S = (s_{i,j}) \in \mathcal{M}_{n,p}(\mathbb{R})$ la matrice définie par $s_{i,i} = \sqrt{\lambda_i}$ pour $1 \leq i \leq p$, et $s_{i,j} = 0$ sinon.

Établir $S = {}^tUJV$, puis en déduire $J = US^tV$.

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

(b) Soit $R = (r_{i,j}) \in \mathcal{M}_{p,n}(\mathbb{R})$ définie par $r_{i,i} = \sqrt{\lambda_i}/(\lambda_i + \mu)$ pour $1 \leq i \leq p$, et $r_{i,j} = 0$ sinon. Établir

$$({}^tJJ + \mu I_p)^{-1} {}^tJ = VR^tU.$$

(c) En déduire

$$({}^tJJ + \mu I_p)^{-1} {}^tJ = \sum_{i=1}^q \frac{\sqrt{\lambda_i}}{\lambda_i + \mu} V_i {}^t(U_i).$$

6. Soit $X \in \mathbb{R}^p$ fixé vérifiant $\nabla F(X) \neq 0$. Pour $h \in \mathbb{R}^p$, 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 le gradient $\nabla M(h)$ et la hessienne $\nabla^2 M(h)$.

(c) En appliquant les résultats précédents à $J(X)$, montrer que M admet un unique point critique h^* . Donner une expression de h^* utilisant 5(c).

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