



Corrigés des épreuves du concours HEC 2008

Jean Mallet

Professeur en classes préparatoires, lycée Montaigne (Paris)



Michel Mitermique

Professeur en classes préparatoires, lycée Jean-Baptiste Corot (Savigny sur Orge) et IPESUP (Paris)

Voie scientifique



BANQUE COMMUNE D'ÉPREUVES

Concepteur : H.E.C.

OPTION SCIENTIFIQUE

MATHÉMATIQUES I

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.

Il ne doit être fait usage d'aucun document : 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 $\mathcal{M}_{p,1}(\mathbb{R})$ (respectivement $\mathcal{M}_{p,1}(\mathbb{R})$ et \mathbb{R}^p (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 $\langle u, v \rangle = {}^t u v$ 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 $f_2(x_1, x_2) = x_1 + x_2^2 + 1$.
a) 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 $\nabla F(x_1, x_2)$.

b) Montrer que le système d'équations qui permet de déterminer les éventuels points critiques de F , peut se mettre sous la forme suivante :

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

c) Établir, pour tout (x_1, x_2) 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 $(-1/2, -1/2)$.

d) 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 $(-1/2, -1/2)$.

e) 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 ${}^t J(X)J(X)$ 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 $a = (a_1, \dots, a_m)$, $b = (b_1, \dots, b_n)$ 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 $[1, n]$, la fonction f_i est définie sur \mathbb{R}^2 par : $f_i(x_1, x_2) = a_i x_1 + b_i x_2 - c_i$.

a) 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 $\langle b, c \rangle$.

b) 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 (\hat{x}_1, \hat{x}_2) .

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

c) 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 (\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.

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 $[1, n]$, la fonction f_i est définie sur \mathbb{R}^2 par : $f_i(x_1, x_2) = x_1 + x_2 - c_i$.

a) Déterminer les points critiques de F .

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

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 $[1, n]$, 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 : $|\langle h, Ph \rangle| \leq \theta \|h\|^2$.

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 $[1, p]$, 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 JJ est diagonalisable et que ses valeurs propres sont positives ou nulles.

b) Montrer que lorsque la matrice JJ est inversible, le rang de la matrice J est égal à p .

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

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

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

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

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 : ${}^tV^tJ JV = D$, 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 $[1, p]$, on note V_i la i -ième colonne de V .

2. a) Montrer que le rang de ${}^tJ J$ est égal à q .

b) Montrer que, pour tout i de $[1, q]$, $J V_i$ est un vecteur propre de la matrice $J^t J$ associé à la valeur propre λ_i . En déduire que les matrices ${}^tJ J$ et $J^t J$ ont les mêmes valeurs propres non nulles.

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

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

3. On pose, pour tout i de $[1, q]$: $U_i = \frac{1}{\sqrt{\lambda_i}} J V_i$.

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

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 $J^t J$.

4. On note U la matrice de $\mathcal{M}_n(\mathbb{R})$ telle que, pour tout i de $[1, n]$, 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 : $S = {}^tU J V$. En déduire l'égalité : $J = U S {}^tV$.

5. a) Montrer que la matrice $({}^tJ J + \mu I)$ 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 : $({}^tJ J + \mu I)^{-1} \times {}^tJ = V R U$.

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

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

Exemple 1

a)

Les applications f_1 et f_2 sont des applications polynomiales définies sur \mathbb{R}^2 , elles sont donc de classe C^2 sur \mathbb{R}^2 .

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

$$\begin{aligned} \frac{\partial F}{\partial x_1}(x_1, x_2) &= \frac{\partial f_1}{\partial x_1} f_1(x_1, x_2) + \frac{\partial f_2}{\partial x_1} f_2(x_1, x_2) \\ &= 2x_1(x_1^2 + x_2 + 1) + (x_1 + x_2^2 + 1) \\ &= 2x_1^3 + 2x_1x_2 + 3x_1 + x_2^2 + 1 \end{aligned}$$

$$\begin{aligned} \frac{\partial F}{\partial x_2}(x_1, x_2) &= \frac{\partial f_1}{\partial x_2} f_1(x_1, x_2) + \frac{\partial f_2}{\partial x_2} f_2(x_1, x_2) \\ &= (x_1^2 + x_2 + 1) + 2x_2(x_1 + x_2^2 + 1) \\ &= x_1^2 + 2x_1x_2 + 2x_2^3 + 3x_2 + 1 \end{aligned}$$

$$\nabla F(x_1, x_2) = (2x_1^3 + 2x_1x_2 + 3x_1 + x_2^2 + 1, x_1^2 + 2x_1x_2 + 2x_2^3 + 3x_2 + 1)$$

b)

(x_1, x_2) est un point critique de F si et seulement si $\nabla F(x_1, x_2) = 0$

$$\nabla F(x_1, x_2) = 0 \iff \begin{cases} 2x_1^3 + 2x_1x_2 + 3x_1 + x_2^2 + 1 = 0 \\ x_1^2 + 2x_1x_2 + 2x_2^3 + 3x_2 + 1 = 0 \end{cases} \quad L_2 \leftarrow L_2 - L_1$$

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

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

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

c)

Considérons $2x_1^3 + 2x_1x_2 + 2x_2^2 - x_1 - x_2 + 3 = 2x_1^3 + x_1(2x_2 - 1) + 2x_2^2 - x_2 + 3$ comme un trinôme en x_1 . Son discriminant Δ vaut $(2x_2 - 1)^2 - 8(2x_2^2 - x_2 + 3) = -12x_2^2 + 4x_2 - 23$. Il est alors immédiat que le discriminant de ce nouveau trinôme est négatif ; ce trinôme reste strictement négatif, $\Delta < 0$, d'où $(2x_2 - 1)^2 - 8(2x_2^2 - x_2 + 3) = -12x_2^2 + 4x_2 - 23 = 2x_1^3 + 2x_1x_2 + 2x_2^2 - x_1 - x_2 + 3 > 0$

Le trinôme $2x_1^2 + x_1(2x_2 - 1) + 2x_2^2 - x_2 + 3$ n'a pas de racines réelles ; il est donc constamment du signe du coefficient de x_1^2 :

$$\forall (x_1, x_2) \in \mathbb{R}^2, 2x_1^2 + 2x_1x_2 + 2x_2^2 - x_1 - x_2 + 3 > 0.$$

L'équation (2) donne $x_1 = x_2$ et l'équation (1) s'écrit alors $2x_1^3 + 3x_1^2 + 3x_1 + 1 = 0$.

Étudions le polynôme P défini par $\forall x \in \mathbb{R}, P(x) = 2x^3 + 3x^2 + 3x + 1$

$P'(x) = 3(2x^2 + 2x + 1)$; $\forall x \in \mathbb{R}, P'(x) > 0$ car son discriminant $\delta = -4$.

On en déduit que P est une bijection de \mathbb{R} dans \mathbb{R} (car P est continue, strictement croissante).

Conclusion : P admet une unique racine réelle ; on constate facilement que $P(-\frac{1}{2}) = 0$

Finalement l'équation $\nabla F(x_1, x_2) = 0$ admet une unique solution $(x_1, x_2) = (-\frac{1}{2}, -\frac{1}{2})$

La fonction F admet un point critique et un seul $(-\frac{1}{2}, -\frac{1}{2})$.

d)

La fonction F est de classe C^2 sur \mathbb{R}^2 , on peut donc appliquer le théorème Schwarz.

$$\begin{aligned} \frac{\partial^2 F}{\partial x_1^2}(x_1, x_2) &= 6x_1^2 + 2x_2 + 3 \\ \frac{\partial^2 F}{\partial x_1 \partial x_2} &= \frac{\partial^2 F}{\partial x_2 \partial x_1}(x_1, x_2) = 2x_1 + 2x_2 \\ \frac{\partial^2 F}{\partial x_2^2}(x_1, x_2) &= 2x_1 + 6x_2^2 + 3 \end{aligned}$$

On a alors $\nabla^2 F(x_1, x_2) = \begin{pmatrix} 6x_1^2 + 2x_2 + 3 & 2x_1 + 2x_2 \\ 2x_1 + 2x_2 & 2x_1 + 6x_2^2 + 3 \end{pmatrix}$.

Au point critique, $\nabla^2 F(-\frac{1}{2}, -\frac{1}{2}) = \begin{pmatrix} \frac{7}{2} & -2 \\ -2 & \frac{7}{2} \end{pmatrix}$

$(s^2 - rt)(-\frac{1}{2}, -\frac{1}{2}) = 4 - \frac{49}{4} = -\frac{33}{4} < 0$ et $r > 0$.

La fonction F admet au point $(-\frac{1}{2}, -\frac{1}{2})$ un minimum local

e)

$J(X) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(X) & \frac{\partial f_1}{\partial x_2}(X) \\ \frac{\partial f_2}{\partial x_1}(X) & \frac{\partial f_2}{\partial x_2}(X) \end{pmatrix} = \begin{pmatrix} 2x_1 & 1 \\ 1 & 2x_2 \end{pmatrix}$

On remarque que ${}^t J(X) = J(X)$.

$$\begin{aligned} {}^t J(X)f(X) &= \begin{pmatrix} 2x_1 & 1 \\ 1 & 2x_2 \end{pmatrix} \begin{pmatrix} f_1(X) \\ f_2(X) \end{pmatrix} \\ &= \begin{pmatrix} 2x_1 & 1 \\ 1 & 2x_2 \end{pmatrix} \begin{pmatrix} x_1^2 + x_2 + 1 \\ x_1 + x_2^2 + 1 \end{pmatrix} \\ &= \begin{pmatrix} 2x_1^3 + 2x_1x_2 + 3x_1 + x_2^3 + 1 \\ 2x_2^3 + 2x_1x_2 + 3x_2 + x_1^3 + 1 \end{pmatrix} \end{aligned}$$

On remarque que ${}^t J(X)f(X) = \nabla F(X)$ puisqu'on a identifié \mathbb{R}^2 et $\mathcal{M}_{2,1}(\mathbb{R})$

$G(X) = {}^t J(X)J(X) = \begin{pmatrix} 4x_1^2 + 1 & 2x_1 + 2x_2 \\ 2x_1 + 2x_2 & 4x_2^2 + 1 \end{pmatrix}$
 $\frac{\partial^2 f_1}{\partial x_1^2}(X) = 2$; $\frac{\partial^2 f_1}{\partial x_1 \partial x_2} = \frac{\partial^2 f_1}{\partial x_2 \partial x_1}(X) = 0$; $\frac{\partial^2 f_1}{\partial x_2^2}(X) = 0$; donc

$$\nabla^2 f_1(X) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$$

$\frac{\partial^2 f_2}{\partial x_1^2}(X) = 0$; $\frac{\partial^2 f_2}{\partial x_1 \partial x_2} = \frac{\partial^2 f_2}{\partial x_2 \partial x_1}(X) = 0$; $\frac{\partial^2 f_2}{\partial x_2^2}(X) = 2$; donc

$$\nabla^2 f_2(X) = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$$

$$\begin{aligned} G(X) + \sum_{i=1}^2 f_i(X) \nabla^2 f_i(X) &= \begin{pmatrix} 4x_1^2 + 1 & 2x_1 + 2x_2 \\ 2x_1 + 2x_2 & 4x_2^2 + 1 \end{pmatrix} + (x_1^2 + x_2 + 1) \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \\ &\quad + (x_1 + x_2^2 + 1) \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 4x_1^2 + 1 + 2x_1^2 + 2x_2 + 2 & 2x_1 + 2x_2 \\ 2x_1 + 2x_2 & 4x_2^2 + 1 + 2x_1 + 2x_2^2 + 2 \end{pmatrix} \\ &= \begin{pmatrix} 6x_1^2 + 2x_2 + 3 & 2x_1 + 2x_2 \\ 2x_1 + 2x_2 & 6x_2^2 + 2x_1 + 3 \end{pmatrix} \end{aligned}$$

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

Exemple 2

a)

$$\begin{aligned} F(x_1, x_2) &= \frac{1}{2} \sum_{i=1}^n (a_i x_1 + b_i x_2 - c_i)^2 \\ \frac{\partial F}{\partial x_1}(x_1, x_2) &= \sum_{i=1}^n a_i (a_i x_1 + b_i x_2 - c_i) = x_1 \sum_{i=1}^n a_i^2 + x_2 \sum_{i=1}^n a_i b_i - \sum_{i=1}^n a_i c_i \\ &= \|a\|^2 x_1 + \langle a, b \rangle x_2 - \langle a, c \rangle \\ \frac{\partial F}{\partial x_2}(x_1, x_2) &= \sum_{i=1}^n b_i (a_i x_1 + b_i x_2 - c_i) = x_1 \sum_{i=1}^n a_i b_i + x_2 \sum_{i=1}^n b_i^2 - \sum_{i=1}^n b_i c_i \\ &= \langle a, b \rangle x_1 + \|b\|^2 x_2 - \langle b, c \rangle \end{aligned}$$

$$\nabla F(x_1, x_2) = (\|a\|^2 x_1 + \langle a, b \rangle x_2 - \langle a, c \rangle, \langle a, b \rangle x_1 + \|b\|^2 x_2 - \langle b, c \rangle)$$

b)

On sait, d'après l'inégalité de Cauchy-Schwarz, que $|\langle a, b \rangle| \leq \|a\| \|b\|$. On sait aussi qu'il y a égalité si et seulement si les vecteurs a et b sont liés. D'après les hypothèses, les vecteurs a et b sont libres, donc l'inégalité est stricte :

$$|\langle a, b \rangle| < \|a\| \|b\|, \text{ ce qui équivaut à } \langle a, b \rangle^2 < \|a\|^2 \|b\|^2$$

Le couple (x_1, x_2) est un point critique de F si et seulement si $\nabla F(x_1, x_2) = 0$

$$\begin{aligned} \nabla F(x_1, x_2) = 0 &\iff \begin{cases} \|a\|^2 x_1 + \langle a, b \rangle x_2 = \langle a, c \rangle \\ \langle a, b \rangle x_1 + \|b\|^2 x_2 = \langle b, c \rangle \end{cases} \quad L_2 \leftarrow \|a\|^2 L_2 - \langle a, b \rangle L_1 \\ &\quad \text{opération permise car } \|a\|^2 \neq 0 \\ &\iff \begin{cases} \|a\|^2 x_1 & + \langle a, b \rangle x_2 = \langle a, c \rangle \\ (\|a\|^2 \|b\|^2 - \langle a, b \rangle^2) x_2 = \|a\|^2 \langle b, c \rangle - \langle a, b \rangle \langle a, c \rangle \end{cases} \end{aligned}$$

D'après le point précédent, $(\|a\|^2 \|b\|^2 - \langle a, b \rangle^2) > 0$.

Le système est de Cramer : il admet donc une unique solution notée $(\widehat{x}_1, \widehat{x}_2)$

Calcul de la solution

On a tout de suite $\widehat{x}_2 = \frac{\|a\|^2 \langle b, c \rangle - \langle a, b \rangle \langle a, c \rangle}{\|a\|^2 \|b\|^2 - \langle a, b \rangle^2}$.

La première ligne du système va donner \widehat{x}_1

$$\widehat{x}_1 = \frac{1}{\|a\|^2} (\langle a, c \rangle - \langle a, b \rangle \frac{\|a\|^2 \langle b, c \rangle - \langle a, b \rangle \langle a, c \rangle}{\|a\|^2 \|b\|^2 - \langle a, b \rangle^2})$$

opération permise car $\|a\|^2 \neq 0$

$$= \frac{1}{\|a\|^2} \frac{1}{\|a\|^2 \|b\|^2 - \langle a, b \rangle^2} (\langle a, c \rangle \|a\|^2 \|b\|^2 - \langle a, c \rangle \langle a, b \rangle^2 - \|a\|^2 \langle a, b \rangle \langle b, c \rangle + \langle a, b \rangle^2 \langle a, c \rangle)$$

$$= \frac{1}{\|a\|^2} \frac{1}{\|a\|^2 \|b\|^2 - \langle a, b \rangle^2} (\|a\|^2 (\|b\|^2 \langle a, c \rangle - \langle a, b \rangle \langle b, c \rangle))$$

D'où les résultats :

$$\widehat{x}_1 = \frac{\|b\|^2 \langle a, c \rangle - \langle a, b \rangle \langle b, c \rangle}{\|a\|^2 \|b\|^2 - \langle a, b \rangle^2} \text{ et } \widehat{x}_2 = \frac{\|a\|^2 \langle b, c \rangle - \langle a, b \rangle \langle a, c \rangle}{\|a\|^2 \|b\|^2 - \langle a, b \rangle^2}$$

c) _____

$$\nabla^2 F(x_1, x_2) = \begin{pmatrix} \frac{\partial^2 F}{\partial x_1^2}(x_1, x_2) & \frac{\partial^2 F}{\partial x_1 \partial x_2}(x_1, x_2) \\ \frac{\partial^2 F}{\partial x_2 \partial x_1}(x_1, x_2) & \frac{\partial^2 F}{\partial x_2^2}(x_1, x_2) \end{pmatrix}$$

$$= \begin{pmatrix} \|a\|^2 & \langle a, b \rangle \\ \langle a, b \rangle & \|b\|^2 \end{pmatrix}$$

$$(s^2 - rt)(\widehat{x}_1, \widehat{x}_2) = \langle a, b \rangle^2 - \|a\|^2 \|b\|^2 < 0$$

$$(s^2 - rt)(\widehat{x}_1, \widehat{x}_2) < 0 \text{ et } r(\widehat{x}_1, \widehat{x}_2) > 0 : F \text{ possède en } (\widehat{x}_1, \widehat{x}_2) \text{ un minimum local}$$

d) _____

Considérons le vecteur $Y = (Y_1, \dots, Y_i, \dots, Y_n)$ de \mathbb{R}^n où $\forall i \in [1, n]$, $Y_i = (a_i x_1 + b_i x_2 - c_i)$. On constate qu'alors $F(X) = \frac{1}{2} \|Y\|^2$.

Si l'on considère la matrice $A \in \mathcal{M}_{n,2}(\mathbb{R})$ suivante : $A = \begin{pmatrix} a_1 & b_1 \\ \vdots & \vdots \\ a_i & b_i \\ \vdots & \vdots \\ a_n & b_n \end{pmatrix}$ et $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathcal{M}_{2,1}(\mathbb{R})$,

on a immédiatement $AX = \begin{pmatrix} a_1 x_1 + b_1 x_2 \\ \vdots \\ a_i x_1 + b_i x_2 \\ \vdots \\ a_n x_1 + b_n x_2 \end{pmatrix}$

Introduisons maintenant le vecteur $C = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \in \mathcal{M}_{n,1}(\mathbb{R})$, on obtient

$$\begin{pmatrix} a_1 x_1 + b_1 x_2 - c_1 \\ \vdots \\ a_i x_1 + b_i x_2 - c_i \\ \vdots \\ a_n x_1 + b_n x_2 - c_n \end{pmatrix} = AX - C.$$

$$F(X) = \frac{1}{2} \|AX - C\|^2$$

- La matrice A est de rang 2 puisque les vecteurs a et b sont libres. D'après le cours sur les moindres carrés, $\|AX - C\|$ sera minimale pour l'unique vecteur \widehat{X} de \mathbb{R}^2 donné par : $\widehat{X} = ({}^t AA)^{-1} \times {}^t AC$.

Vérifions que $\widehat{X} = (\widehat{x}_1, \widehat{x}_2)$: cela prouvera que le minimum est global.

$${}^t AA = \begin{pmatrix} a_1 & \dots & a_n \\ b_1 & \dots & b_n \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \|a\|^2 & \langle a, b \rangle \\ \langle a, b \rangle & \|b\|^2 \end{pmatrix}$$

- Calcul de $({}^t AA)^{-1}$ (on sait que cette matrice existe car A , donc ${}^t A$ sont inversibles) . Pour cela résolvons le système : ${}^t AA \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ d'inconnues x et y .

$${}^t AA \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \iff \begin{cases} \|a\|^2 x + \langle a, b \rangle y = \alpha \\ \langle a, b \rangle x + \|b\|^2 y = \beta \end{cases}$$

On a déjà résolu ce système à la question **b)** : il suffit de changer $\langle a, c \rangle$ en α , $\langle b, c \rangle$ en β et x_1 en x , x_2 en y : les formules

$$x_1 = \frac{\|b\|^2 \langle a, c \rangle - \langle a, b \rangle \langle b, c \rangle}{\|a\|^2 \|b\|^2 - \langle a, b \rangle^2} \text{ et } x_2 = \frac{\|a\|^2 \langle b, c \rangle - \langle a, b \rangle \langle a, c \rangle}{\|a\|^2 \|b\|^2 - \langle a, b \rangle^2} \text{ donnent}$$

$$x = \frac{\|b\|^2 \alpha - \langle a, b \rangle \beta}{\|a\|^2 \|b\|^2 - \langle a, b \rangle^2} \text{ et } y = \frac{\|a\|^2 \beta - \langle a, b \rangle \alpha}{\|a\|^2 \|b\|^2 - \langle a, b \rangle^2}$$

Matriciellement

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{\|a\|^2 \|b\|^2 - \langle a, b \rangle^2} \begin{pmatrix} \|b\|^2 & -\langle a, b \rangle \\ -\langle a, b \rangle & \|a\|^2 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

$$({}^t AA)^{-1} = \frac{1}{\|a\|^2 \|b\|^2 - \langle a, b \rangle^2} \begin{pmatrix} \|b\|^2 & -\langle a, b \rangle \\ -\langle a, b \rangle & \|a\|^2 \end{pmatrix}$$

$${}^t AC = \begin{pmatrix} a_1 & \dots & a_n \\ b_1 & \dots & b_n \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} \langle a, c \rangle \\ \langle b, c \rangle \end{pmatrix}$$

$$\widehat{X} = \frac{1}{\|a\|^2 \|b\|^2 - \langle a, b \rangle^2} \begin{pmatrix} \|b\|^2 & -\langle a, b \rangle \\ -\langle a, b \rangle & \|a\|^2 \end{pmatrix} \begin{pmatrix} \langle a, c \rangle \\ \langle b, c \rangle \end{pmatrix}$$

$$= \frac{1}{\|a\|^2 \|b\|^2 - \langle a, b \rangle^2} \begin{pmatrix} \|b\|^2 \langle a, c \rangle - \langle a, b \rangle \langle b, c \rangle \\ \|a\|^2 \langle b, c \rangle - \langle a, b \rangle \langle a, c \rangle \end{pmatrix}$$

$$\widehat{X} = \begin{pmatrix} \widehat{x}_1 \\ \widehat{x}_2 \end{pmatrix} : F \text{ admet en } (\widehat{x}_1, \widehat{x}_2) \text{ un minimum global}$$

Exemple 3

a) _____

$$F(X) = \frac{1}{2} \sum_{i=1}^n (x_1 + x_2 - c_i)^2$$

$$\frac{\partial F}{\partial x_1}(X) = \sum_{i=1}^n (x_1 + x_2 - c_i) = n(x_1 + x_2 - \bar{c}) \quad \text{car } \bar{c} = \frac{1}{n} \sum_{i=1}^n c_i$$

$$\frac{\partial F}{\partial x_2}(X) = \sum_{i=1}^n (x_1 + x_2 - c_i) = n(x_1 + x_2 - \bar{c})$$

$$\nabla F(X) = (n(x_1 + x_2 - \bar{c}), n(x_1 + x_2 - \bar{c}))$$

Les points critiques de F vérifient : $x_1 + x_2 = \bar{c}$

b)

Soit $(\widehat{x}_1, \widehat{x}_2)$ un point critique de F .

$$F(\widehat{x}_1, \widehat{x}_2) = \frac{1}{2} \sum_{i=1}^n (\widehat{x}_1 + \widehat{x}_2 - c_i)^2 = \frac{1}{2} \sum_{i=1}^n (\bar{c} - c_i)^2. \text{ Or par définition, } s^2 = \frac{1}{n} \sum_{i=1}^n (\bar{c} - c_i)^2. \text{ Donc}$$

$$F(\widehat{x}_1, \widehat{x}_2) = \frac{ns^2}{2}$$

$$\begin{aligned} \bullet \quad F(x_1, x_2) - F(\widehat{x}_1, \widehat{x}_2) &= \frac{1}{2} \sum_{i=1}^n (x_1 + x_2 - c_i)^2 - \frac{ns^2}{2} \\ &= \frac{1}{2} \sum_{i=1}^n (x_1 + x_2 - \bar{c} + \bar{c} - c_i)^2 - \frac{ns^2}{2} \\ &= \frac{1}{2} \sum_{i=1}^n (x_1 + x_2 - \bar{c})^2 + \frac{1}{2} \sum_{i=1}^n (\bar{c} - c_i)^2 + \frac{2}{2} \sum_{i=1}^n (x_1 + x_2 - \bar{c})(\bar{c} - c_i) - \frac{ns^2}{2} \\ &= \frac{n}{2} (x_1 + x_2 - \bar{c})^2 + \frac{ns^2}{2} + (x_1 + x_2 - \bar{c}) \underbrace{\sum_{i=1}^n (\bar{c} - c_i)}_{=0} - \frac{ns^2}{2} \\ &= \frac{n}{2} (x_1 + x_2 - \bar{c})^2 + \underbrace{(x_1 + x_2 - \bar{c}) \left(n\bar{c} - \sum_{i=1}^n c_i \right)}_{=0} \end{aligned}$$

$$F(x_1, x_2) - F(\widehat{x}_1, \widehat{x}_2) = \frac{n}{2} (x_1 + x_2 - \bar{c})^2 \geq 0$$

c)

On vient de voir que $\forall (x_1, x_2) \in \mathbb{R}^2, F(x_1, x_2) - F(\widehat{x}_1, \widehat{x}_2) \geq 0$

Aux points critiques, F orésente des minima globaux

Il était prévisible que F ne pouvait pas présenter de maximum globaux car $\lim_{x_1 \rightarrow +\infty} F(x_1, x_2) = +\infty$.

Cas général

a)

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

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

$$\forall j \in [1, p], \frac{\partial F}{\partial x_j}(X) = \sum_{i=1}^n \frac{\partial f_i}{\partial x_j}(X) f_i(X)$$

$$\text{Or } J(X) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(X) & \dots & \frac{\partial f_1}{\partial x_j}(X) & \dots & \frac{\partial f_1}{\partial x_p}(X) \\ \vdots & & \vdots & & \vdots \\ \frac{\partial f_i}{\partial x_1}(X) & \dots & \frac{\partial f_i}{\partial x_j}(X) & \dots & \frac{\partial f_i}{\partial x_p}(X) \\ \vdots & & \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1}(X) & \dots & \frac{\partial f_n}{\partial x_j}(X) & \dots & \frac{\partial f_n}{\partial x_p}(X) \end{pmatrix}, \text{ donc}$$

$${}^t J(X) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(X) & \dots & \frac{\partial f_i}{\partial x_1}(X) & \dots & \frac{\partial f_n}{\partial x_1}(X) \\ \vdots & & \vdots & & \vdots \\ \frac{\partial f_1}{\partial x_i}(X) & \dots & \frac{\partial f_j}{\partial x_i}(X) & \dots & \frac{\partial f_n}{\partial x_i}(X) \\ \vdots & & \vdots & & \vdots \\ \frac{\partial f_1}{\partial x_p}(X) & \dots & \frac{\partial f_j}{\partial x_p}(X) & \dots & \frac{\partial f_n}{\partial x_p}(X) \end{pmatrix}$$

$${}^t J(X) f(X) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(X) & \dots & \frac{\partial f_j}{\partial x_1}(X) & \dots & \frac{\partial f_n}{\partial x_1}(X) \\ \vdots & & \vdots & & \vdots \\ \frac{\partial f_1}{\partial x_i}(X) & \dots & \frac{\partial f_j}{\partial x_i}(X) & \dots & \frac{\partial f_n}{\partial x_i}(X) \\ \vdots & & \vdots & & \vdots \\ \frac{\partial f_1}{\partial x_p}(X) & \dots & \frac{\partial f_j}{\partial x_p}(X) & \dots & \frac{\partial f_n}{\partial x_p}(X) \end{pmatrix} \begin{pmatrix} f_1(X) \\ \vdots \\ f_j(X) \\ \vdots \\ f_n(X) \end{pmatrix}$$

${}^t J(X) f(X)$ est une colonne dont la i ème ligne est $\sum_{j=1}^n \frac{\partial f_j}{\partial x_i}(X) f_j(X)$. Comme on assimile les colonnes aux listes, on peut dire que ${}^t J(X) f(X)$ est une n -liste dont le i ème terme est $\sum_{j=1}^n \frac{\partial f_j}{\partial x_i}(X) f_j(X)$. On reconnaît le i ème terme de $\nabla F(X)$

$$\nabla F(X) = {}^t J(X) f(X)$$

b)

Par définition, $\forall X \in \mathbb{R}^p, \nabla^2 F(X) = \left(\frac{\partial^2 F}{\partial x_k \partial x_j}(X) \right) \in \mathcal{M}_p(\mathbb{R})$. Cette matrice est symétrique d'après le théorème de Schwarz puisque F est de classe C^2 sur \mathbb{R}^p . Explicitons un peu :

$$\text{La } k \text{ème ligne de } \nabla^2 F(X) \text{ est } \left(\frac{\partial^2 F}{\partial x_k \partial x_1}(X), \dots, \frac{\partial^2 F}{\partial x_k \partial x_j}(X), \dots, \frac{\partial^2 F}{\partial x_k \partial x_p}(X) \right)$$

D'autre part, pour tout $j \in [1; p]$

$$\frac{\partial F}{\partial x_j}(X) = \sum_{i=1}^n \frac{\partial f_i(X)}{\partial x_j} f_i(X), \text{ donc } \forall (k, j) \in ([1; p])^2,$$

$$\frac{\partial^2 F}{\partial x_k \partial x_j}(X) = \sum_{i=1}^n \frac{\partial f_i(X)}{\partial x_j} \frac{\partial f_i(X)}{\partial x_k} + \sum_{i=1}^n \frac{\partial^2 f_i(X)}{\partial x_k \partial x_j} f_i(X) \quad (1)$$

$$\bullet \quad \forall i \in [1; n], \nabla^2 f_i(X) = \begin{pmatrix} \frac{\partial^2 f_i(X)}{\partial x_1^2} & \dots & \dots & \frac{\partial^2 f_i(X)}{\partial x_1 \partial x_p} \\ \vdots & & & \vdots \\ \frac{\partial^2 f_i(X)}{\partial x_p \partial x_1} & \dots & \dots & \frac{\partial^2 f_i(X)}{\partial x_p^2} \end{pmatrix} \in \mathcal{M}_p(\mathbb{R})$$

$\forall (k, j) \in ([1; p])^2, \frac{\partial^2 F}{\partial x_k \partial x_j}(X)$ est le terme général de $\nabla^2 F(X)$

$\sum_{i=1}^n \frac{\partial f_i(X)}{\partial x_j} \frac{\partial f_i(X)}{\partial x_k}$ est le terme général de ${}^t J(X) J(X)$

$\sum_{i=1}^n \frac{\partial^2 f_i(X)}{\partial x_k \partial x_j} f_i(X)$ est le terme général de $f_i(X) \nabla^2 f_i(X)$,

donc, d'après l'égalité (1), le terme général de $\nabla^2(F(X))$ est égal au terme général de ${}^tJ(X)J(X)$ + la somme, pour $i \in \llbracket 1; n \rrbracket$, du terme général de $f_i(X)\nabla^2 f_i(X)$

$$\text{Conclusion : } \nabla^2(F(X)) = {}^tJ(X)J(X) + \sum_{i=1}^n f_i(X)\nabla^2 f_i(X) = G(X) + \sum_{i=1}^n f_i(X)\nabla^2 f_i(X)$$

Partie II

1) _____

$$\begin{aligned} L(h) &= \frac{1}{2} \|l(h)\|^2 = \frac{1}{2} {}^t l(h) l(h) \\ &= \frac{1}{2} ({}^t f(X) + J(X)h) (f(X) + J(X)h) \\ &= \frac{1}{2} ({}^t f(X) + {}^t h {}^t J(X)) (f(X) + J(X)h) \\ &= \frac{1}{2} ({}^t f(X) f(X) + {}^t f(X) J(X)h + {}^t h {}^t J(X) f(X) + {}^t h {}^t J(X) J(X)h) \\ &= \frac{1}{2} (\|f(X)\|^2 + {}^t f(X) J(X)h + {}^t h {}^t J(X) f(X) + {}^t h G(X)h) \end{aligned}$$

$J(X) \in \mathcal{M}_{n,p}(\mathbb{R})$, $h \in \mathcal{M}_{p,1}(\mathbb{R})$, donc $J(X)h \in \mathcal{M}_{n,1}(\mathbb{R})$

$f(X) = (f_1(X), \dots, f_n(X)) \in \mathbb{R}^n$ que l'on assimile à $\mathcal{M}_{n,1}(\mathbb{R})$; il en résulte que ${}^t f(X) J(X)h \in \mathcal{M}_1(\mathbb{R})$, donc ${}^t f(X) J(X)h \in \mathbb{R}$.

Même raisonnement pour ${}^t h {}^t J(X) f(X)$. Donc ${}^t h {}^t J(X) f(X) = ({}^t h {}^t J(X) f(X)) = {}^t f(X) J(X)h$

Il en résulte que $L(h) = \frac{1}{2} (\|f(X)\|^2 + 2 {}^t h {}^t J(X) f(X) + {}^t h G(X)h)$

D'après I-4-b), ${}^t J(X) f(X) = \nabla F(X)$ et on sait que $\frac{1}{2} \|f(X)\|^2 = F(X)$, donc

$$\begin{aligned} L(h) &= F(X) + {}^t h {}^t J(X) f(X) + \frac{1}{2} {}^t h G(X)h \\ &= F(X) + {}^t h \nabla F(X) + \frac{1}{2} {}^t h G(X)h \quad \text{d'après I-4-a)} \end{aligned}$$

2-a) _____

La matrice P est symétrique réelle de $\mathcal{M}_p(\mathbb{R})$, elle est donc diagonalisable en base orthonormée.

2-b) _____

Il existe donc une matrice orthogonale Q appartenant à $\mathcal{M}_p(\mathbb{R})$, il existe une matrice diagonale D appartenant à $\mathcal{M}_p(\mathbb{R})$ telles que $P = QDQ^{-1} = QD^t Q$ (puisque Q est orthogonale).

$${}^t h P h = {}^t h Q D^t Q h = ({}^t Q h) D^t Q h$$

Posons $Y = {}^t Q h$; Y appartient à $\mathcal{M}_{p,1}(\mathbb{R})$. On peut écrire $Y = (y_1, \dots, y_p)$ en identifiant, comme le dit l'énoncé, les éléments de \mathbb{R}^p , les matrices colonnes de $\mathcal{M}_{p,1}(\mathbb{R})$ et les matrices lignes de $\mathcal{M}_{1,p}(\mathbb{R})$.

L'égalité précédente devient ${}^t h P h = {}^t Y D Y = \sum_{k=1}^p \theta_k y_k^2$.

L'inégalité triangulaire donne $|{}^t h P h| \leq \sum_{k=1}^p |\theta_k| y_k^2$

Or par hypothèse $\forall k \in \llbracket 1; p \rrbracket$, $|\theta_k| \leq \theta$. On multiplie ces inégalités par $y_k^2 \geq 0$ et on les ajoute; il vient

$$\sum_{k=1}^p |\theta_k| y_k^2 \leq \sum_{k=1}^p \theta y_k^2 \leq \theta \sum_{k=1}^p y_k^2, \text{ soit finalement}$$

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

3-a) _____

F est de classe C^2 sur \mathbb{R}^p , son développement à l'ordre 2 existe et on a $\forall h \in \mathbb{R}^p$,

$F(X+h) = F(X) + \langle \nabla F(X), h \rangle + \frac{1}{2} q_X(h) + \|h\|^2 \varepsilon(h)$ avec $\lim_{h \rightarrow 0} \varepsilon(h) = 0$ et où q_X est la forme quadratique associée à la hessienne de F au point X .

3-b) _____

$$\langle \nabla F(X), h \rangle = {}^t h \nabla F(X) \text{ et } q_X(h) = {}^t h \nabla^2 F(X) h.$$

D'après I-4-b), $q_X(h) = {}^t h (G(X) + \sum_{i=1}^n f_i(X)\nabla^2 f_i(X)) h$, donc

$$\begin{aligned} F(X+h) &= F(X) + {}^t h \nabla F(X) + \frac{1}{2} {}^t h (G(X) + \sum_{i=1}^n f_i(X)\nabla^2 f_i(X)) h + \|h\|^2 \varepsilon(h) \\ &= F(X) + {}^t h \nabla F(X) + \underbrace{\frac{1}{2} {}^t h G(X) h}_{=L(h)} + \frac{1}{2} {}^t h \left(\sum_{i=1}^n f_i(X)\nabla^2 f_i(X) \right) h + \|h\|^2 \varepsilon(h) \\ &= F(X+h) = L(h) + \frac{1}{2} {}^t h \left(\sum_{i=1}^n f_i(X)\nabla^2 f_i(X) \right) h + \|h\|^2 \varepsilon(h) \quad (2) \end{aligned}$$

Pour tout indice $i \in \llbracket 1; n \rrbracket$, la matrice $f_i(X)\nabla^2 f_i(X)$ est symétrique réelle à cause du théorème de Schwarz, donc la matrice $\frac{1}{2} \left(\sum_{i=1}^n f_i(X)\nabla^2 f_i(X) \right)$ est aussi symétrique réelle.

Posons $P = \frac{1}{2} \left(\sum_{i=1}^n f_i(X)\nabla^2 f_i(X) \right)$; l'égalité (2) s'écrit

$$\begin{aligned} F(X+h) &= L(h) + {}^t h P h + \|h\|^2 \varepsilon(h) \\ \forall h \neq 0, \left| \frac{F(X+h) - L(h)}{\|h\|} \right| &= \left| \frac{{}^t h P h + \|h\|^2 \varepsilon(h)}{\|h\|} \right| \\ &\leq \frac{1}{\|h\|} (|{}^t h P h| + \|h\|^2 |\varepsilon(h)|) \\ &\leq \frac{1}{\|h\|} |{}^t h P h| + \|h\| |\varepsilon(h)| \\ &\leq \frac{\theta \|h\|^2}{\|h\|} + \|h\| |\varepsilon(h)| \quad \text{d'après 2-b)} \\ &\leq \theta \|h\| + \|h\| |\varepsilon(h)| \end{aligned}$$

$\lim_{h \rightarrow 0} (\theta \|h\| + \|h\| |\varepsilon(h)|) = 0$ donc par encadrement

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

4-a) _____

$$\varphi_1(h) = {}^t h \nabla F(X) = \sum_{i=1}^p h_i \frac{\partial F}{\partial x_i}(X); \text{ donc } \frac{\partial \varphi_1}{\partial h_j}(h) = \frac{\partial F}{\partial x_j}(X)$$

$\varphi_2(h) = \sum_{i=1}^p g_{i,i}(X) h_i^2 + 2 \sum_{1 \leq i < k \leq p} g_{i,k}(X) h_i h_k$ d'après la formule classique du développement d'une forme quadratique lorsque l'on connaît sa matrice.

Soit $j \in \llbracket 1; p \rrbracket$; réécrivons la deuxième somme de ce développement en faisant apparaître les termes en h_j

avec $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_q > 0$ et $\lambda_{q+1} = \dots = \lambda_p = 0$

2-a) _____

Le rang de la matrice tJJ est égal au rang de D puisque les deux matrices sont semblables.

$$\text{rang}({}^tJJ) = q$$

2-b) _____

On sait que $\forall i \in [1; q], V_i$ est un vecteur propre de tJJ associé à la valeur propre $\lambda_i > 0 : {}^tJJV_i = \lambda_i V_i$.

Donc $(J^tJ)V_i = \lambda_i JV_i$.

Si $JV_i = 0$, alors $V_i \in \text{Ker } J$, donc $V_i \in \text{Ker } {}^tJJ$ d'après II-5-b) ; le vecteur V_i serait un vecteur propre de tJJ associé à la valeur propre 0, **ce qui est faux**.

Donc JV_i est un vecteur propre de J^tJ associé à la valeur propre non nulle λ_i .

Toute valeur propre non nulle de tJJ est une valeur propre non nulle de J^tJ . Par symétrie des rôles joués par J et tJ , on en déduit que toute valeur propre non nulle de J^tJ est une valeur propre non nulle de tJJ .

Les matrices $J^tJ \in \mathcal{M}_n(\mathbb{R})$ et ${}^tJJ \in \mathcal{M}_p(\mathbb{R})$ ont les mêmes valeurs propres non nulles

2-c) _____

Nous supposons que la famille $((Y_1, \dots, Y_r))$ est une base orthogonale du sous-espace propre de tJJ associé à une valeur propre $\lambda > 0$, ce qui semble avoir été omis dans l'énoncé.

Soit (Y_1, \dots, Y_r) une base orthogonale de vecteurs propres du sous-espace propre de tJJ associé à une valeur propre $\lambda > 0$.

Soit $(\alpha_1, \dots, \alpha_r) \in \mathbb{R}^r / \sum_{k=1}^r \alpha_k JY_k = 0$. Prenons le produit scalaire de ces deux membres avec JY_i pour $i \in [1; r]$. On obtient, après avoir utilisé la linéarité par rapport à la première variable du produit scalaire,

$$\begin{aligned} \sum_{k=1}^r \alpha_k \langle JY_k, JY_i \rangle = 0 &\iff \sum_{k=1}^r \alpha_k {}^t(JY_k)JY_i = 0 \\ &\iff \sum_{k=1}^r \alpha_k {}^tY_k {}^tJJY_i = 0 \\ &\iff \sum_{k=1}^r \alpha_k \lambda {}^tY_k Y_i = 0 \quad \text{puisque } Y_i \text{ est un vecteur propre de } {}^tJJ \text{ associé à } \lambda \\ &\iff \alpha_i \lambda \|Y_i\|^2 = 0 \end{aligned}$$

puisque les vecteurs Y_k sont deux à deux orthogonaux.

$\lambda > 0$ et $\|Y_i\|^2 > 0$; l'égalité précédente implique $\alpha_i = 0$

$$\forall (\alpha_1, \dots, \alpha_r) \in \mathbb{R}^r / \sum_{k=1}^r \alpha_k JY_k = 0 \implies \alpha_k = 0 \text{ pour tout } k \in [1; r] : \text{ la famille } (JY_k) \text{ est libre}$$

2-d) _____

Soit E_λ le sous-espace propre de tJJ associé à la valeur propre $\lambda > 0$ et notons r sa dimension. Soit $(V_1, \dots, V_r) \in (\mathcal{M}_{p,1}(\mathbb{R}))^r$ une base orthogonale de E_λ .

On vient de voir que la famille $(JV_1, \dots, JV_r) \in (\mathcal{M}_{n,1}(\mathbb{R}))^r$ est une famille libre de E'_λ , sous-espace propre de J^tJ associé à $\lambda > 0$. Si l'on note r' sa dimension, on en déduit $r \leq r'$. Par la symétrie de rôles joués par tJ et J , on en déduit que $r' \leq r$.

Conclusion $r = r'$.

Puisque tJJ est diagonalisable, son image est égale à la somme des sous-espaces propres associés aux valeurs propres non nulles. De même pour J^tJ . D'après le résultat précédent, et puisque ces deux matrices ont les mêmes valeurs propres non nulles, on conclut

Les matrices ${}^tJJ \in \mathcal{M}_p(\mathbb{R})$ et $J^tJ \in \mathcal{M}_n(\mathbb{R})$ ont le même rang q

3-a) _____

$\forall i \in [1; q], \lambda_i > 0$, donc $U_i = \frac{1}{\sqrt{\lambda_i}} JV_i$ existe et appartient à \mathbb{R}^n ou $\mathcal{M}_{n,1}(\mathbb{R})$. Soit $(i, j) \in ([1; q])^2$.

$$\begin{aligned} \langle U_i, U_j \rangle &= \frac{1}{\sqrt{\lambda_i} \sqrt{\lambda_j}} \langle JV_i, JV_j \rangle \\ &= \frac{1}{\sqrt{\lambda_i} \sqrt{\lambda_j}} {}^t(JV_i)JV_j = \frac{1}{\sqrt{\lambda_i} \sqrt{\lambda_j}} {}^tV_i {}^tJJV_j \\ &= \frac{1}{\sqrt{\lambda_i} \sqrt{\lambda_j}} {}^tV_i ({}^tJJ)V_j \\ &= \frac{1}{\sqrt{\lambda_i} \sqrt{\lambda_j}} \lambda_j {}^tV_i V_j \quad \text{car } V_j \text{ est un vecteur propre de } {}^tJJ \text{ associé à } \lambda_j \end{aligned}$$

Les vecteurs V_i et V_j sont des vecteurs colonnes de la matrice V , qui est une matrice orthogonale, donc ces deux vecteurs sont normés et orthogonaux dès que $i \neq j$. D'où le résultat :

$$\langle U_i, U_j \rangle = \begin{cases} 0 & \text{si } i \neq j \\ \frac{\lambda_j}{\sqrt{\lambda_j} \sqrt{\lambda_j}} = 1 & \text{si } j = i \end{cases}$$

La famille $(U_k)_{1 \leq k \leq q}$ est une famille orthonormale de \mathbb{R}^n , formée de vecteurs propres de J^tJ car V_k est un vecteur propre de tJJ , donc JV_k est un vecteur propre de J^tJ , donc U_k aussi.

3-b) _____

(U_1, \dots, U_q) est une famille libre (orthonormale) de q vecteurs propres de J^tJ , associés aux valeurs propres non nulles de J^tJ . Donc $\forall j \in [1; q], U_j \in \text{Im}(J^tJ)$.

La dimension de cette image est q d'après 2-c), donc (U_1, \dots, U_q) est une base orthonormée de $\text{Im } J^tJ$; c'est aussi une base orthonormée de la somme des sous-espaces propres associés aux valeurs propres non nulles de J^tJ .

Puisque la matrice J^tJ est diagonalisable (en base orthonormée d'ailleurs) les sous-espaces propres sont supplémentaires. Le sous-espace propre associé à la valeur 0 est $\text{Ker}(J^tJ)$.

$\mathbb{R}^n = \text{Im}(J^tJ) \oplus \text{Ker}(J^tJ)$; d'ailleurs ces deux sous-espaces sont orthogonaux.

Donc si l'on prend une base orthonormée (U_{q+1}, \dots, U_n) de $\text{Ker}(J^tJ)$, la famille $(U_1, \dots, U_q, U_{q+1}, \dots, U_n)$ est une base orthonormée de \mathbb{R}^n

4) _____

Soit $S' = {}^tUJV$, on remarque que $S' \in \mathcal{M}_{n,p}(\mathbb{R})$ car $V \in \mathcal{M}_p(\mathbb{R}), J \in \mathcal{M}_{n,p}(\mathbb{R}), U \in \mathcal{M}_n(\mathbb{R})$. Ecrivons $s'_{i,j}$ son terme général (on a $1 \leq i \leq n, 1 \leq j \leq p$).

$s'_{i,j}$ est le résultat du produit de la i ème ligne de tU , qui est tU_i , et de la j ème colonne de JV , qui est JV_j . Donc $s'_{i,j} = {}^tU_i JV_j = \langle U_i, JV_j \rangle$

- Pour $1 \leq j \leq q, s'_{i,j} = \langle U_i, \sqrt{\lambda_j} U_j \rangle$ d'après la définition des U_j , donc $s'_{i,j} = \sqrt{\lambda_j} \langle U_i, U_j \rangle$, ce qui donne

$\forall j \in [1; q], s'_{i,j} = 0$ pour $i \neq j$ et $s'_{i,j} = \sqrt{\lambda_j}$ puisque U_j est normé.

- Pour $j \geq q+1$, le vecteur JV_j est nul ; en effet, V_j est un vecteur propre de tJJ associé à la valeur propre 0, donc c'est un vecteur de $\text{Ker } {}^tJJ$ d'après la question 1) ; d'après la question II-5-a) on sait que V_j est aussi un vecteur de $\text{Ker } J$, donc $JV_j = 0$ et par conséquent $\langle U_i, JV_j \rangle = 0$