HEC ESCP-EAP 2008
Math. I

Durée : 4 heures

Dans tout le problème, n et p désignent deux entiers vérifiant $1 \leqslant p \leqslant n$. On note $\mathcal{M}_{n,p}(\mathbb{R})$ l'espace vectoriel des matrices à n lignes et p colonnes, à coefficient réels.

La transposée d'une matrice A de $\mathcal{M}_{n,p}(\mathbb{R})$ est notée ${}^tA$. Lorsqu'une matrice 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 $\langle u, v\rangle = {}^tuv$ 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 $X = (x_1, x_2, \dots, x_p)$ 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 $\big(f_1(X), f_2(X), \dots, f_n(X)\big)$.

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 suivant :

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

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

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

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 :

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

dans laquelle i désigne l'indice de ligne et j l'indice de colonne.

On pose $G(X) = {}^tJ(X)J(X)$.

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 :

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

a) Justifier que F est de classe $\mathcal{C}^2$ 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 :

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

c) Etablir, 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 ${}^tJ(X)f(X)$ et $G(X) + \sum\limits_{i=1}^{2} f_i(X)\nabla^2 f_i(X)$ en fonction de $\nabla F(X)$ et $\nabla^2 F(X)$ respectivement.

2°) *Un deuxième exemple.*

Soit $a = (a_1, \ldots, a_n)$, $b = (b_1, \ldots, b_n)$ et $c = (c_1, \ldots, c_n)$ 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 $(\widehat{x}_1, \widehat{x}_2)$.

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

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 $(\widehat{x}_1, \widehat{x}_2)$.

3°) *Un troisième exemple.*

On suppose que $c_1, c_2, \ldots, c_n$ sont n réels donnés non tous égaux. On note $\overline{c}$ 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 $(\widehat{x}_1, \widehat{x}_2)$ 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 :
$$F(x_1, x_2) - F(\widehat{x}_1, \widehat{x}_2) = \frac{n}{2}(x_1 + x_2 - \overline{c})^2.$$

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

4°) *Retour au cas général.*
Soit $X = (x_1, \ldots, x_p)$ un point de $\mathbb{R}^p$.

a) Exprimer $\nabla F(X)$ en fonction de ${}^tJ(X)$ 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.
Etablir la formule : $\nabla^2 F(X) = G(X) + \sum\limits_{i=1}^{n} f_i(X)\nabla^2 f_i(X)$.

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

1°) Etablir, pour tout h de $\mathbb{R}^p$, l'égalité :
$$L(h) = F(X) + {}^th\nabla F(X) + \frac{1}{2}{}^thG(X)h.$$

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

a) Justifier que P est diagonalisable.

b) On note $\theta_1, \ldots, \theta_p$ les valeurs propres de P, et on pose $\theta = \max\limits_{1\leq j\leq p} |\theta_j|$.
Montrer, pour tout vecteur h de $\mathbb{R}^p$, l'inégalité suivante : $|{}^thPh| \leqslant \theta\|h\|^2$.

3°) a) Ecrire 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 :
$$\lim_{\|h\|\to 0} \frac{F(X+h) - L(h)}{\|h\|} = 0$$
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 : $G(X) = \big(g_{i,j}(X)\big)_{1\leq i,j\leq p}$. Soit φ_1 et φ_2 deux fonctions définies sur $\mathbb{R}^p$ par : $\varphi_1(h) = {}^th\nabla F(X)$ et $\varphi_2(h) = {}^thG(X)h$.

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

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. Etablir 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 ${}^tJJ$ est diagonalisable et que ses valeurs propres sont positives ou nulles.

b) Montrer que lorsque la matrice ${}^tJJ$ 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 $\langle \widehat{h}, \nabla F(X) \rangle \leqslant 0$.

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

a) Montrer que L admet un unique point critique $\widehat{h}$ donné par

$$\widehat{h} = -(G(X))^{-1} \times {}^tJ(X) f(X)$$

b) Etablir 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 $1 \leqslant q \leqslant p$, et des réels $\lambda_1, \lambda_2, \ldots, \lambda_q$ tels que $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_q > 0$, qui vérifient l'égalité ${}^tV\,{}^tJJV = D$, où $D = (d_{i,j})_{1 \le i,j \le p}$ est définie par : $d_{i,i} = \lambda_i$ si $1 \leqslant i \leqslant q$, et $d_{i,j} = 0$ sinon. Si $q < p$, on pose $\lambda_{q+1} = \cdots = \lambda_p = 0$.
Pour tout i de $[\![1, p]\!]$, on note V_i la $i^{\text{ème}}$ colonne de V.

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

b) Montrer que, pour tout i de $[\![1, q]\!]$, 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 $(Y_1, \ldots, Y_r)$ une base du sous-espace propre de ${}^tJJ$ associée à une valeur propre non nulle. Montrer que la famille $(JY_1, \ldots, JY_r)$ est une famille libre de $\mathcal{M}_{n,1}(\mathbb{R})$.

d) En déduire que les sous-espaces propres de ${}^tJJ$ et $J\,{}^tJ$ associés à la même valeur propre non nullesont de même dimension, et que le rang de $J\,{}^tJ$ vaut q.

3°) On pose, pour tout i de $[\![1, q]\!]$: $U_i = \dfrac{1}{\sqrt{\lambda_i}} JV_i$.

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

b) En déduire qu'il existe une base orthonormée $(U_1, \ldots, U_q, U_{q+1}, \ldots, U_n)$ 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 $i^{\text{ème}}$ colonne de U est la matrice colonne U_i de $\mathcal{M}_{n,1}(\mathbb{R})$.
Soit $S = (s_{i,j})_{1 \leq i \leq n, 1 \leq j \leq p}$ la matrice de $\mathcal{M}_{n,p}(\mathbb{R})$ définie par $s_{i,i} = \sqrt{\lambda_i}$ si $1 \leqslant i \leqslant p$ et $s_{i,j} = 0$ sinon.
Etablir 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 $R = (r_{i,j})_{1 \leq i \leq p, 1 \leq j \leq n}$ la matrice de $\mathcal{M}_{p,n}(\mathbb{R})$ définie par : $r_{i,i} = \dfrac{\sqrt{\lambda_i}}{\lambda_i + \mu}$ si $1 \leqslant i \leqslant p$ et $r_{i,j} = 0$ sinon.
Etablir la formule suivante : $({}^tJJ + \mu I)^{-1} \times {}^tJ = VR^tU$.

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

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 $\lim\limits_{\|h\| \to 0} \dfrac{F(X+h) - M(h)}{\|h\|} = 0$.

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^*.

A partir de ce minimum local h^* de M (ou du minimum local $\widehat{h}$ 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.

Solution

Partie I

Question 1.

a) Pour $(x_1, x_2) \in \mathbb{R}^2$, on a ici :

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

Par conséquent F est polynomiale et à ce titre est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}^2$. Comme pour toute fonction u dérivable, on a $(u^2)' = 2u'u$, on a facilement :

$$\frac{\partial F}{\partial x_1}(x_1,x_2) = 2x_1(x_1^2+x_2+1)+(x_1+x_2^2+1)$$
$$= 2x_1^3+2x_1x_2+3x_1+x_2^2+1$$

et de la même façon (ou mieux en invoquant la symétrie des écritures) :

$$\frac{\partial F}{\partial x_2}(x_1,x_2) = (x_1^2+x_2+1)+2x_2(x_1+x_2^2+1)$$
$$= 2x_2^3+2x_1x_2+3x_2+x_1^2+1$$

et, en écrivant en colonne :

$$\boxed{\nabla F(X) = \begin{pmatrix} 2x_1^3+2x_1x_2+3x_1+x_2^2+1 \\ 2x_2^3+2x_1x_2+3x_2+x_1^2+1 \end{pmatrix}}$$

b) Un point critique de F est un point dont les coordonnées (x_1,x_2) vérifient le système $\big(\frac{\partial F}{\partial x_1}(x_1,x_2) = 0, \frac{\partial F}{\partial x_2}(x_1,x_2) = 0\big)$, ce qui équivaut au système $\big(\frac{\partial F}{\partial x_1}(x_1,x_2) = 0, \frac{\partial F}{\partial x_1}(x_1,x_2) - \frac{\partial F}{\partial x_2}(x_1,x_2) = 0\big)$ et s'écrit :

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

Puisque $x_1^3-x_2^3 = (x_1-x_2)(x_1^2+x_1x_2+x_2^2)$ et $x_1^2-x_2^2 = (x_1-x_2)(x_1+x_2)$, il reste bien :

$$\boxed{(x_1,x_2) \text{ critique} \iff \begin{cases} 2x_1^3+2x_1x_2+3x_1+x_2^2+1=0 \\ (x_1-x_2)(2x_1^2+2x_1x_2+2x_2^2-x_1-x_2+3)=0 \end{cases} \quad (*)}$$

c) Pour x_2 fixé dans $\mathbb{R}$, posons :

$$\begin{aligned}\varphi(x_1) &= 2x_1^2+2x_1x_2+2x_2^2-x_1-x_2+3 \\ &= 2x_1^2+(2x_2-1)x_1+2x_2^2-x_2+3\end{aligned}$$

Ce trinôme du second degré en x_1 a pour discriminant :

$$\begin{aligned}\Delta &= (2x_2-1)^2-8\times(2x_2^2-x_2+3) = -12x_2^2+4x_2-23 \\ &= -12\big(x_2-\frac{1}{6}\big)^2+\frac{1}{3}-23<0\end{aligned}$$

Ainsi $\varphi(x_1)$ est toujours du signe de son coefficient dominant, c'est-à-dire :

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

Le système $(*)$ précédent se réduit donc à :

$$\begin{cases} 2x_1^3+2x_1x_2+3x_1+x_2^2+1=0 \\ x_1-x_2=0 \end{cases}, \text{ soit : } \begin{cases} x_2=x_1 \\ 2x_1^3+3x_1^2+3x_1+1=0 \end{cases}$$

L'énoncé nous dit que $-1/2$ doit être racine de l'équation du 3$^{\text{ème}}$ degré précédente. En effet en effectuant la division par $2x_1+1$, on obtient :

$$2x_1^3+3x_1^2+3x_1+1 = (2x_1+1)(x_1^2+x_1+1)$$

et comme l'équation $x_1^2+x_1+1=0$ n'a pas de racine réelle (les racines dans $\mathbb{C}$ sont j et j^2) la seule solution de $(*)$ correspond à $x_1=-1/2$ et donc également à $x_2=-1/2$:

$$\boxed{\big(-\frac{1}{2},-\frac{1}{2}\big) \text{ est } \textbf{LE} \text{ point critique de } F}$$

d) C'est parti, les expressions obtenues en **1°) a)** donnent :

$$\frac{\partial^2 F}{\partial x_1^2}(x_1,x_2) = 6x_1^2 + 2x_2 + 3\,;\ \frac{\partial^2 F}{\partial x_2^2}(x_1,x_2) = 6x_2^2 + 2x_1 + 3$$

$$\frac{\partial^2 F}{\partial x_1 \partial x_2}(x_1,x_2) = \frac{\partial^2 F}{\partial x_2 \partial x_1}(x_1,x_2) = 2(x_1 + x_2)$$

Soit :

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

En particulier $\nabla^2 F(-1/2,-1/2) = \begin{pmatrix} 7/2 & -2 \\ -2 & 7/2 \end{pmatrix}$. Ainsi avec les notations de Monge :

$$\det(\Delta^2 F(-1/2,-1/2)) = rt - s^2 = \frac{49}{4} - 4 > 0, \text{ avec } r \text{ et } t \text{ positifs.}$$

Le premier point indique qu'il y a en ce point critique extremum local et le second point précise qu'il s'agit d'un minimum.

e) $\star$ $f(X) = \begin{pmatrix} x_1^2 + x_2 + 1 \\ x_1 + x_2^2 + 1 \end{pmatrix}$, donc

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

$\star$ Et :

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

Soit :

$$\boxed{{}^tJ(X)f(X) = \nabla F(X)}$$

$\star$ Avec $f_1(X) = x_1^2 + x_2 + 1$ et $f_2(X) = x_1 + x_2^2 + 1$, il vient :

$$\nabla^2 f_1(X) = \begin{pmatrix} \dfrac{\partial^2 f_1}{\partial x_1^2}(X) & \dfrac{\partial^2 f_1}{\partial x_1 \partial x_2}(X) \\ \dfrac{\partial^2 f_1}{\partial x_2 \partial x_1}(X) & \dfrac{\partial^2 f_1}{\partial x_2^2}(X) \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$$

et de même $\nabla^2 f_2(X) = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$, d'où :

$$\begin{aligned} K &= G(X) + f_1(X)\nabla^2 f_1(X) + f_2(X)\nabla^2 f_2(X) \\ &= \begin{pmatrix} 2x_1 & 1 \\ 1 & 2x_2 \end{pmatrix}^2 + (x_1^2 + x_2 + 1)\begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} + (x_1 + x_2^2 + 1)\begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 4x_1^2 + 1 & 2(x_1+x_2) \\ 2(x_1+x_2) & 4x_2^2 + 1 \end{pmatrix} + \begin{pmatrix} 2x_1^2 + 2x_2 + 2 & 0 \\ 0 & 2x_1 + 2x_2^2 + 2 \end{pmatrix} \\ &= \begin{pmatrix} 6x_1^2 + 2x_2 + 3 & 2(x_1+x_2) \\ 2(x_1+x_2) & 6x_2^2 + 2x_1 + 3 \end{pmatrix} \text{ et on constate que :} \end{aligned}$$

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

Question 2.

a) Ici $F(X) = \frac{1}{2}\sum_{i=1}^{n} f_i^2(X) = \frac{1}{2}\sum_{i=1}^{n}(a_i x_1 + b_i x_2 - c_i)^2$, donc :

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

et de même $\dfrac{\partial F}{\partial x_2}(X) = \langle a, b\rangle x_1 + \|b\|^2 x_2 - \langle b, c\rangle$:

$$\boxed{\nabla F(X) = \begin{pmatrix} \|a\|^2 x_1 + \langle a, b\rangle x_2 - \langle a, c\rangle \\ \langle a, b\rangle x_1 + \|b\|^2 x_2 - \langle b, c\rangle \end{pmatrix}}$$

b) On sait que $\langle a, b\rangle^2 \leqslant \|a\|^2\|b\|^2$, avec égalité si et seulement si la famille (a, b) est liée, donc ici l'inégalité est stricte et le système

$$\begin{cases} \|a\|^2 x_1 + \langle a, b\rangle x_2 = \langle a, c\rangle \\ \langle a, b\rangle x_1 + \|b\|^2 x_2 = \langle b, c\rangle \end{cases}$$

est de déterminant $\|a\|^2\|b\|^2 - \langle a, b\rangle^2$ non nul. Il admet une solution et une seule $(\widehat{x}_1, \widehat{x}_2)$, qui est donc le seul point critique de F, et on obtient aisément :

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

c) Le calcul des dérivées partielles secondes de F est banal et donne :

$$\nabla^2 F(X) = \begin{pmatrix} \|a\|^2 & \langle a, b\rangle \\ \langle a, b\rangle & \|b\|^2 \end{pmatrix} \quad \text{(indépendant du point } X\text{)}$$

et comme $\|a\|^2\|b\|^2 - \langle a, b\rangle^2 > 0$ avec $\|a\|^2 > 0$, on en conclut que F admet un minimum local en $(\widehat{x}_1, \widehat{x}_2)$.
(D'ailleurs, par « canonisation » de F, il serait facile de voir qu'il s'agit en fait d'un minimum global.)

Question 3.

Vu la fin de la question, on fait semblant de ne rien voir pour l'instant . . .

a) $\nabla F(x_1, x_2) = \begin{pmatrix} \sum_{i=1}^{n}(x_1 + x_2 - c_i) \\ \sum_{i=1}^{n}(x_1 + x_2 - c_i) \end{pmatrix} = n\begin{pmatrix} x_1 + x_2 - \overline{c} \\ x_1 + x_2 - \overline{c} \end{pmatrix}$

Donc les points critiques de F, qui sont les points où $\nabla F(x_1, x_2) = 0$, sont tous les points $(\widehat{x}_1, \widehat{x}_2)$ tels que $\widehat{x}_1 + \widehat{x}_2 = \overline{c}$.

b) $\star$ $F(\widehat{x}_1, \widehat{x}_2) = \frac{1}{2}\sum_{i=1}^{n}(\widehat{x}_1 + \widehat{x}_2 - c_i)^2 = \frac{1}{2}\sum_{i=1}^{n}(\overline{c} - c_i)^2 = \frac{n}{2}\times\frac{1}{n}\sum_{i=1}^{n}(\overline{c} - c_i)^2$

Soit :

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