

# HEC 2008

## Éléments de correction

### Première partie

1. a. La fonction

$$F : (x_1, x_2) \in \mathbb{R}^2 \longmapsto \frac{1}{2}(x_1^2 + x_2 + 1)^2 + \frac{1}{2}(x_1 + x_2^2 + 1)^2$$

est polynomiale donc de classe  $\mathcal{C}^2$  sur  $\mathbb{R}^2$  avec :

$$\forall (x_1, x_2) \in \mathbb{R}^2, \quad \nabla F(x_1, x_2) = (2x_1^3 + 2x_1x_2 + 3x_1 + x_2^2 + 1, 2x_2^3 + 2x_1x_2 + 3x_2 + x_1^2 + 1).$$

b. Pour  $(x_1, x_2) \in \mathbb{R}^2$ ,

$$\nabla F(x_1, x_2) = 0 \iff \begin{cases} \partial_1 F(x_1, x_2) = 0 \\ \partial_2 F(x_1, x_2) = 0 \end{cases} \iff \begin{cases} \partial_1 F(x_1, x_2) = 0 \\ \partial_1 F(x_1, x_2) - \partial_2 F(x_1, x_2) = 0 \end{cases}$$

où

$$\begin{aligned} \partial_1 F(x_1, x_2) - \partial_2 F(x_1, x_2) &= 2(x_1^3 - x_2^3) + 3(x_1 - x_2) - (x_1^2 - x_2^2) \\ &= (x_1 - x_2)(2x_1^2 + 2x_1x_2 + 2x_2^2 + 3 - x_1 - x_2), \end{aligned}$$

d'où le résultat.

c. Pour  $(x_1, x_2) \in \mathbb{R}^2$ ,

$$\begin{aligned} 2x_1^2 + 2x_1x_2 + 2x_2^2 - x_1 - x_2 + 3 &= (x_1^2 + 2x_1x_2 + x_2^2) + (x_1^2 - x_1) + (x_2^2 - x_2) + 3 \\ &= (x_1 + x_2)^2 + \left(x_1 - \frac{1}{2}\right)^2 + \left(x_2 - \frac{1}{2}\right)^2 + \frac{5}{2} \geq \frac{5}{2} > 0 \end{aligned}$$

si bien que dans le système obtenu en **b.**, la deuxième équation devient  $x_1 = x_2$  puis la première  $2x_1^3 + 3x_1^2 + 3x_1 + 1 = 0$ . Inspiré par l'énoncé, on vérifie que  $-\frac{1}{2}$  est solution de cette dernière équation polynomiale, puis la factorisation

$$2x_1^3 + 3x_1^2 + 3x_1 + 1 = 2\left(x_1 + \frac{1}{2}\right)(x_1^2 + x_1 + 1)$$

met en évidence que c'est la seule dans  $\mathbb{R}$ . La fonction  $F$  admet donc  $A = \left(-\frac{1}{2}, -\frac{1}{2}\right)$  pour seul point critique.

d. On a :

$$\forall (x_1, x_2) \in \mathbb{R}^2, \quad \nabla^2 F(x, y) = \begin{pmatrix} 6x_1^2 + 2x_2 + 3 & 2x_1 + 2x_2 \\ 2x_1 + 2x_2 & 6x_2^2 + 2x_1 + 3 \end{pmatrix}.$$

En particulier au point  $A$ ,

$$\nabla^2 F(A) = \begin{pmatrix} \frac{7}{2} & -2 \\ -2 & \frac{7}{2} \end{pmatrix} = \begin{pmatrix} r & s \\ s & t \end{pmatrix}$$

où  $rt - s^2 = \frac{33}{4} > 0$  et  $r > 0$ . La fonction  $F$  présente donc au point  $A$  un minimum local.

e. Pour  $X = (x_1, x_2) \in \mathbb{R}^2$ , il vient :

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

puis

$${}^t J(X) f(X) = \begin{pmatrix} 2x_1 & 1 \\ 1 & 2x_2 \end{pmatrix} \begin{pmatrix} x_1^2 + x_2 + 1 \\ x_2^2 + x_1 + 1 \end{pmatrix} = \begin{pmatrix} 2x_1^3 + 2x_1x_2 + 3x_1 + x_2^2 + 1 \\ 2x_2^3 + 2x_1x_2 + 3x_2 + x_1^2 + 1 \end{pmatrix} = \nabla F(X).$$

Enfin,

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

2. a. Ici encore, la fonction

$$F : (x_1, x_2) \in \mathbb{R}^2 \mapsto \frac{1}{2} \sum_{i=1}^n (a_i x_1 + b_i x_2 - c_i)^2$$

est polynomiale donc de classe  $\mathcal{C}^2$  sur  $\mathbb{R}^n$ , avec par exemple

$$\begin{aligned} \forall (x_1, x_2) \in \mathbb{R}^2, \quad \partial_1 F(x_1, x_2) &= \sum_{i=1}^n a_i (a_i x_1 + b_i x_2 - c_i) = \langle a, ax_1 + bx_2 - c \rangle \\ &= \|a\|^2 x_1 + \langle a, b \rangle x_2 - \langle a, c \rangle, \end{aligned}$$

et plus généralement

$$\forall (x_1, x_2) \in \mathbb{R}^2, \quad \nabla F(x_1, x_2) = (\|a\|^2 x_1 + \langle a, b \rangle x_2 - \langle a, c \rangle, \langle a, b \rangle x_1 + \|b\|^2 x_2 - \langle b, c \rangle).$$

b. La famille  $(a, b)$  étant libre par hypothèse, l'inégalité de Cauchy-Schwarz dans  $\mathbb{R}^n$  euclidien canonique assure que  $\|a\|^2 \|b\|^2 - \langle a, b \rangle^2 > 0$ . Le système

$$\nabla F(x_1, x_2) = 0 \iff \begin{cases} \|a\|^2 x_1 + \langle a, b \rangle x_2 = \langle a, c \rangle \\ \langle a, b \rangle x_1 + \|b\|^2 x_2 = \langle b, c \rangle \end{cases}$$

d'inconnue  $(x_1, x_2)$  est donc de Cramer ; on le résout aisément pour obtenir l'unique point critique de F :

$$(\hat{x}_1, \hat{x}_2) = \frac{1}{\|a\|^2 \|b\|^2 - \langle a, b \rangle^2} (\|b\|^2 \langle a, c \rangle - \langle a, b \rangle \langle a, c \rangle, \|a\|^2 \langle b, c \rangle - \langle a, b \rangle \langle a, c \rangle).$$

c. On a :

$$\forall X \in \mathbb{R}^2, \quad \nabla^2 F(X) = \begin{pmatrix} \|a\|^2 & \langle a, b \rangle \\ \langle a, b \rangle & \|b\|^2 \end{pmatrix} = \begin{pmatrix} r & s \\ s & t \end{pmatrix}$$

où  $rt - s^2 = \|a\|^2 \|b\|^2 - \langle a, b \rangle^2 > 0$  et  $r > 0$ . Ainsi la hessienne de F est-elle définie-positive au point critique  $(\hat{x}_1, \hat{x}_2)$ , si bien que F y présente un minimum local.

d. Par théorème et par croissance de la fonction carré sur  $\mathbb{R}_+$ , la quantité

$$F(x_1, x_2) = \frac{1}{2} \|ax_1 + bx_2 - c\|^2 = \frac{1}{2} \|u - c\|^2$$

est minimale lorsque  $u = x_1 a + x_2 b$  est égal au projeté orthogonal du vecteur  $c$  sur le sous-espace vectoriel  $\text{Vect}(a, b)$ . Sachant que la fonction F admet un unique point critique, seul extremum éventuel, ce projeté orthogonal est donc égal à  $\hat{x}_1 a + \hat{x}_2 b$  et la fonction F présente en  $(\hat{x}_1, \hat{x}_2)$  un minimum global.

3. L'énoncé aurait pu préciser la définition de  $s^2$  pour lever l'ambiguïté (facteur  $\frac{1}{n}$  ou  $\frac{1}{n-1}$  devant la somme) :

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

a. D'après 2. avec  $a = b = (1, \dots, 1)$  (l'hypothèse  $(a, b)$  libre n'intervient qu'à partir de la question b.), la fonction F est de classe  $\mathcal{C}^2$  sur  $\mathbb{R}^2$ , avec :

$$\forall (x_1, x_2) \in \mathbb{R}^2, \quad \nabla F(x_1, x_2) = n(x_1 + x_2 - \bar{c}, x_1 + x_2 - \bar{c}).$$

Les points critiques de F sont donc les points de la droite affine d'équation  $x_1 + x_2 = \bar{c}$ .

b. Étant donné un point  $(\hat{x}_1, \hat{x}_2)$  critique pour F, i.e. tel que  $\hat{x}_1 + \hat{x}_2 = \bar{c}$ , on a :

$$F(\hat{x}_1, \hat{x}_2) = \frac{1}{2} \sum_{i=1}^n (\bar{c} - c_i)^2 = \frac{n}{2} s^2.$$

Pour  $(x_1, x_2) \in \mathbb{R}^2$ , il vient alors :

$$\begin{aligned} F(x_1, x_2) - F(\hat{x}_1, \hat{x}_2) &= \frac{1}{2} \sum_{i=1}^n ((x_1 + x_2 - c_i)^2 - (\bar{c} - c_i)^2) = \frac{(x_1 + x_2 - \bar{c})}{2} \sum_{i=1}^n (x_1 + x_2 + \bar{c} - 2c_i) \\ &= \frac{(x_1 + x_2 - \bar{c})}{2} (n(x_1 + x_2 + \bar{c}) - 2n\bar{c}) = \frac{n}{2} (x_1 + x_2 - \bar{c})^2. \end{aligned}$$

c. Il ressort immédiatement de b. que pour tout point critique  $(\hat{x}_1, \hat{x}_2)$  de F,

$$\forall (x_1, x_2) \in \mathbb{R}^2, \quad F(x_1, x_2) - F(\hat{x}_1, \hat{x}_2) \geq 0$$

si bien que F y présente un minimum global.

www.rbllo.fr

Ce résultat peut être obtenu plus directement en étudiant les variations de la fonction d'une variable  $\psi : x \mapsto \frac{1}{2} \sum_{i=1}^n (x - c_i)^2$  : il apparaît sur l'expression de  $\psi' : x \mapsto n(x - \bar{c})$  que  $\psi$  est strictement décroissante sur  $]-\infty, \bar{c}]$  puis strictement croissante sur  $[\bar{c}, +\infty[$ . Elle présente donc un minimum strict égal à  $\frac{n}{2}s^2$ , atteint au point  $\bar{c}$ . Puisque  $x_1 + x_2$  prend toutes les valeurs réelles lorsque  $(x_1, x_2)$  décrit  $\mathbb{R}^2$ , il en ressort que F admet un minimum global égal à  $\frac{n}{2}s^2$ , atteint en tout point tel que  $x_1 + x_2 = \bar{c}$ .

4. a. On a :

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

d'où  $\nabla F(X) = {}^t J(X) f(X)$ .

b. Pour  $k, j \in \llbracket 1, p \rrbracket$ ,

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

d'où le résultat :

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

### Deuxième partie

1. Pour  $h \in \mathbb{R}^p$ , on a :

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

d'après I.4.a. et par définition de  $G(X)$ .

2. a. La matrice P est diagonalisable en tant que matrice symétrique réelle.

b. Plus précisément, il existe des vecteurs  $v_1, \dots, v_p$  propres pour P, respectivement associés aux valeurs propres  $\theta_1, \dots, \theta_p$ , et formant une base orthonormale de  $\mathbb{R}^p$ . On a alors  $P = \sum_{j=1}^p \theta_j v_j v_j^t$ .

Pour  $h \in \mathbb{R}^p$ , on a alors, vu l'expression du produit scalaire en base orthonormale :

$$|{}^t h P h| = \left| \sum_{j=1}^p \theta_j {}^t h v_j v_j^t h \right| = \left| \sum_{j=1}^p \theta_j \langle v_j, h \rangle^2 \right| \leq \sum_{j=1}^p |\theta_j| \langle v_j, h \rangle^2 \leq \theta \sum_{j=1}^p \langle v_j, h \rangle^2 = \theta \|h\|^2.$$

3. a. La fonction F étant de classe  $\mathcal{C}^2$  sur  $\mathbb{R}^p$  par opérations sur les fonctions  $\mathcal{C}^2$ , elle admet au point X le développement limité à l'ordre 2 :

$$F(X + h) = F(X) + {}^t h \nabla F(X) + \frac{1}{2} {}^t h \nabla^2 F(X) h + o(\|h\|^2), \quad h \rightarrow 0.$$

b. D'après a.,

$$\frac{F(X + h) - L(h)}{\|h\|} = \frac{1}{2 \|h\|} {}^t h (\nabla^2 F(X) - G(X)) h + o(\|h\|), \quad h \rightarrow 0.$$

Or, en notant  $\theta$  le réel associé à la matrice symétrique réelle  $\nabla^2 F(X) - G(X)$ , on a :

$$\left| \frac{F(X + h) - L(h)}{\|h\|} \right| \leq \frac{1}{2 \|h\|} |{}^t h (\nabla^2 F(X) - G(X)) h| + o(\|h\|) \leq \frac{\theta}{2} \|h\| + o(\|h\|), \quad h \rightarrow 0$$

si bien que

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

par encadrement.

4. a. On a :

$$\forall h = (h_1, \dots, h_p) \in \mathbb{R}^p, \quad \varphi_1(h) = \sum_{k=1}^p \frac{\partial F}{\partial x_k}(X) h_k$$

si bien que :

$$\forall j \in \llbracket 1, p \rrbracket, \quad \forall h \in \mathbb{R}^p, \quad \frac{\partial \varphi_1}{\partial h_j}(h) = \frac{\partial F}{\partial x_j}(X).$$

www.rblld.fr

Quant à  $\varphi_2$  : pour  $j \in \llbracket 1, p \rrbracket$  et  $h = (h_1, \dots, h_p) \in \mathbb{R}^p$ , on a :

$$\begin{aligned} \varphi_2(h) &= \sum_{1 \leq k, \ell \leq p} g_{k,\ell}(X) h_k h_\ell = g_{j,j}(X) h_j^2 + \sum_{\substack{1 \leq k, \ell \leq p \\ k=j, \ell \neq j}} g_{k,\ell}(X) h_k h_\ell + \\ &\quad + \sum_{\substack{1 \leq k, \ell \leq p \\ k \neq j, \ell=j}} g_{k,\ell}(X) h_k h_\ell + \sum_{\substack{1 \leq k, \ell \leq p \\ k \neq j, \ell \neq j}} g_{k,\ell}(X) h_k h_\ell \\ &= g_{j,j}(X) h_j^2 + 2 \sum_{\substack{1 \leq k \leq p \\ k \neq j}} g_{j,k}(X) h_j h_k + \sum_{\substack{1 \leq k, \ell \leq p \\ k \neq j, \ell \neq j}} g_{k,\ell}(X) h_k h_\ell, \end{aligned}$$

car  $G(X)$  est symétrique, d'où :

$$\frac{\partial \varphi_2}{\partial h_j}(h) = 2g_{j,j}(X) h_j + 2 \sum_{\substack{1 \leq k \leq p \\ k \neq j}} g_{j,k}(X) h_k = 2 \sum_{k=1}^p g_{j,k}(X) h_k.$$

**b.** D'après la question **1.**,

$$\forall h \in \mathbb{R}^p, \quad L(h) = F(X) + \varphi_1(h) + \frac{1}{2} \varphi_2(h)$$

d'où, d'après **a.** :

$$\forall h \in \mathbb{R}^p, \quad \forall j \in \llbracket 1, p \rrbracket, \quad \frac{\partial L}{\partial h_j}(h) = \frac{\partial F}{\partial x_j}(X) + \sum_{k=1}^p g_{j,k}(X) h_k$$

et, par suite :

$$\forall h \in \mathbb{R}^p, \quad \nabla L(h) = \nabla F(X) + G(X)h.$$

**c.** À partir de **a.** et **b.**, il vient :

$$\forall k, j \in \llbracket 1, p \rrbracket, \quad \forall h \in \mathbb{R}^p, \quad \frac{\partial^2 L}{\partial h_k \partial h_j}(h) = \frac{\partial^2 \varphi_1}{\partial h_k \partial h_j}(h) + \frac{1}{2} \frac{\partial^2 \varphi_2}{\partial h_k \partial h_j}(h) = g_{k,j}(X),$$

si bien que  $\nabla^2 L(h) = G(X)$  pour tout  $h \in \mathbb{R}^p$ .

**5. a.** La matrice  ${}^t J J$  est carrée d'ordre  $p$ , symétrique réelle donc diagonalisable. Si  $\lambda$  en est une valeur propre et  $V$  un vecteur propre associé, alors  ${}^t V ({}^t J J V) = \lambda {}^t V V = \lambda \|V\|^2$  d'où, en comparant à  ${}^t V ({}^t J J V) = ({}^t J V) J V = \|J V\|^2$  avec  $\|V\| > 0$  car  $V \neq 0$ ,  $\lambda = \frac{\|J V\|^2}{\|V\|^2} \geq 0$ .

**b.** Si  ${}^t J J$  est inversible alors, pour tout  $X \in \mathbf{M}_{p,1}(\mathbb{R})$ ,  $J X = 0$  implique  ${}^t J J X = 0$  et donc  $X = 0$ . Ainsi  $\text{Ker } J = \{0\}$  puis, d'après le théorème du rang,  $\text{rg } J = p - \dim \text{Ker } J = p$ .

**6.** Si  $\hat{h}$  est un point critique de  $L$ , alors  $\nabla L(\hat{h}) = 0$  si bien, d'après **4.b.**, que :

$$\langle \hat{h}, \nabla F(X) \rangle = {}^t \hat{h} \nabla F(X) = -{}^t \hat{h} G(X) \hat{h} = -{}^t \hat{h} {}^t J(X) J(X) \hat{h} = -\|J(X) \hat{h}\|^2 \leq 0. \tag{1}$$

**7. a.** Si  $G(X)$  est inversible alors, toujours d'après **5.b.**,

$$\nabla L(h) = 0 \iff \nabla F(X) = -G(X)h \iff h = -G(X)^{-1} \nabla F(X)$$

si bien que  $L$  admet un unique point critique :  $\hat{h} = -G(X)^{-1} \nabla F(X) = -G(X)^{-1} {}^t J(X) f(X)$  d'après **I.4.a.**

**b.** Pour justifier que  $\hat{h}$  est une direction de décroissance pour  $F$ , il suffit de vérifier que l'inégalité  $\langle \hat{h}, \nabla F(X) \rangle \leq 0$  établie en **6.** est stricte. Or, en cas d'égalité, on aurait  $J(X) \hat{h} = 0$  d'après (1) puis  $\hat{h} = 0$  d'après **5.b.** puisque  $G(X)$  est inversible par hypothèse. Mais cela impliquerait d'après **a.**  $\nabla F(X) = {}^t J(X) f(X) = 0$ , en contradiction avec les hypothèses.

Par ailleurs, la matrice  $\nabla^2 L(\hat{h}) = G(X) = {}^t J(X) J(X)$  est définie-positive : ses valeurs propres sont positives d'après **5.a.**, et non nulles car  $G(X)$  est inversible par hypothèse. La fonction  $L$  présente donc au point critique  $\hat{h}$  un minimum local.

### Troisième partie

**1.** La matrice  ${}^t J J \in \mathbf{M}_p(\mathbb{R})$  étant symétrique réelle, il existe une base orthonormale  $(V_1, \dots, V_p)$  de  $\mathbf{M}_{p,1}(\mathbb{R})$  formée de colonnes propres de  ${}^t J J$ , respectivement associées à des valeurs propres  $\lambda_1, \dots, \lambda_p$ . Ces valeurs propres étant toutes positives ou nulles d'après **II.5.a.**, on peut supposer que  $\lambda_1 \geq \dots \geq \lambda_q > \lambda_{q+1} = \dots = \lambda_p = 0$ , quitte à réordonner les vecteurs  $V_1, \dots, V_p$ .

En notant  $V$  la matrice de passage de la base canonique de  $\mathbf{M}_{p,1}(\mathbb{R})$  à la base  $(V_1, \dots, V_p)$ , orthogonale car ces deux bases sont orthonormales, la formule de changement de base assure que  ${}^tV{}^tJJV = V^{-1}({}^tJJ)V = D = \text{diag}(\lambda_1, \dots, \lambda_p)$ . Il reste à remarquer qu'à remarquer que par définition,  $V_1, \dots, V_p$  sont en fait les colonnes de la matrice  $V$ .

2. **a.** Puisque  $V$  est orthogonale et sachant que deux matrices semblables ont même rang, on a  $\text{rg } {}^tJJ = \text{rg } D = q$ .
- b.** Pour  $i \in \llbracket 1, q \rrbracket$ , on a  $J{}^tJ(JV_i) = J({}^tJJV_i) = \lambda_i JV_i$  où  $JV_i \neq 0$  puisque  ${}^tJJV_i = \lambda_i V_i \neq 0$ , si bien que  $JV_i$  est vecteur propre de  $J{}^tJ$  pour la valeur propre  $\lambda_i$ .  
On a ainsi établi l'inclusion  $(\text{Sp } {}^tJJ) \setminus \{0\} \subset (\text{Sp } J{}^tJ) \setminus \{0\}$ . En appliquant ce résultat à la matrice  ${}^tJ$  (en lieu et place de  $J$ ), on en déduit l'inclusion réciproque pour finalement conclure que  $(\text{Sp } {}^tJJ) \setminus \{0\} = (\text{Sp } J{}^tJ) \setminus \{0\}$  : les deux matrices  ${}^tJJ$  et  $J{}^tJ$  ont les mêmes valeurs propres non nulles.
- c.** Étant donné des réels  $\alpha_1, \dots, \alpha_r$  tels que  $\sum_i \alpha_i JY_i = 0$ , il vient  $\sum_i \alpha_i {}^tJJY_i = 0$  i.e.  $\sum_i \alpha_i \lambda Y_i = 0$ . La famille  $(Y_1, \dots, Y_r)$  étant libre par hypothèse, on en déduit que  $\alpha_i \lambda = 0$  et donc  $\alpha_i = 0$  pour tout  $i \in \llbracket 1, r \rrbracket$  car  $\lambda \neq 0$ . Ainsi la famille  $(JY_1, \dots, JY_r)$  est libre dans  $\mathbf{M}_{n,1}(\mathbb{R})$ .
- d.** Soit  $\lambda \neq 0$  une valeur propre de  ${}^tJJ$ . À partir des résultats des questions **b.** et **c.**, on observe que  $\dim E_\lambda({}^tJJ) \leq \dim E_\lambda(J{}^tJ)$ . En appliquant ce résultat à la matrice  ${}^tJ$ , sachant que  $\lambda$  est valeur propre de  $J{}^tJ$  d'après **b.**, on obtient l'inégalité réciproque, pour finalement conclure que  $\dim E_\lambda({}^tJJ) = \dim E_\lambda(J{}^tJ)$  : les sous-espaces propres des deux matrices  ${}^tJJ$  et  $J{}^tJ$  associés à une même valeur propre sont de même dimension.

Il en ressort, avec le résultat de la question **a.** appliqué à  $J$  et à  ${}^tJ$ , que :

$$\text{rg } J{}^tJ = \sum_{\substack{\lambda \in \text{Sp } J{}^tJ \\ \lambda \neq 0}} \dim E_\lambda(J{}^tJ) = \sum_{\substack{\lambda \in \text{Sp } {}^tJJ \\ \lambda \neq 0}} \dim E_\lambda({}^tJJ) = \text{rg } {}^tJJ = q. \tag{2}$$

3. **a.** Pour  $i, j \in \llbracket 1, q \rrbracket$ ,

$${}^tU_i U_j = \frac{1}{\sqrt{\lambda_i \lambda_j}} {}^tV_i {}^tJJV_j = \sqrt{\frac{\lambda_j}{\lambda_i}} {}^tV_i V_j = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{sinon} \end{cases}$$

d'après **1.**, si bien que  $(U_1, \dots, U_q)$  est une famille orthonormée de  $\mathbf{M}_{n,1}(\mathbb{R})$ , dont les vecteurs sont propres pour  $J{}^tJ$  d'après **2.b.**

- b.** D'après **2.b.** et **a.**,  $(U_1, \dots, U_q)$  est une famille orthonormale donc libre formée de  $q$  vecteurs du sous-espace  $\bigoplus_{\lambda \in (\text{Sp } J{}^tJ) \setminus \{0\}} E_\lambda(J{}^tJ)$ . Comme ce sous-espace est de dimension  $q$  d'après (2), c'en est donc une base orthonormale. Enfin, comme  $E_0(J{}^tJ)$  en est le supplémentaire orthogonal dans  $\mathbf{M}_{n,1}(\mathbb{R})$  puisque  $J{}^tJ$  est symétrique réelle, cette famille peut être complétée par des éléments de  $E_0(J{}^tJ)$  en une base orthonormale de  $\mathbf{M}_{n,1}(\mathbb{R})$ .

4. On peut remarquer pour commencer que la matrice  $U$  est orthogonale d'après **3.b.**

Il est alors possible de justifier la relation  $S = {}^tUJV$  grâce à la formule de changement de base, mais le plus simple est de faire appel à des produits par blocs : la matrice  $JV$  a pour colonnes  $JV_j$ ,  $1 \leq j \leq p$ , et la matrice  ${}^tU$  pour lignes  $U_i$ ,  $1 \leq i \leq n$ . Le produit  ${}^tUJV \in \mathbf{M}_{n,p}(\mathbb{R})$  admet donc pour coefficient générique  ${}^tU_i J V_j$ ,  $1 \leq i \leq n$ ,  $1 \leq j \leq p$ , que l'on va calculer en distinguant trois cas :

> Si  $i, j \leq q$ ,

$${}^tU_i J V_j = \frac{1}{\sqrt{\lambda_i}} {}^tV_i {}^tJJV_j = \frac{\lambda_j}{\sqrt{\lambda_i}} {}^tV_i V_j = \begin{cases} \sqrt{\lambda_i} & \text{si } i = j \\ 0 & \text{sinon} \end{cases}.$$

> Si  $i > q$  et  $j \leq q$ ,  $U_i \in E_0(J{}^tJ)$  alors que  $JV_j \in E_{\lambda_j}(J{}^tJ) \subset E_0(J{}^tJ)^\perp$  si bien que  ${}^tU_i J V_j = 0$ .

> Si  $i, j > q$ , alors  $\|JV_j\|^2 = {}^tV_j({}^tJJV_j) = 0$  car  $V_j \in E_0(J{}^tJ)$ , et par suite  ${}^tU_i J V_j = 0$ .

Dans tous les cas,  ${}^tU_i J V_j = s_{i,j}$ , si bien que  $S = {}^tUJV$  c'est-à-dire, comme  $U$  et  $V$  sont orthogonales,  $J = US{}^tV$ .

5. **a.** D'après **1.**,  ${}^tV({}^tJJ + \mu I_p)V = D + \mu I_p$  si bien que  ${}^tJJ + \mu I_p$  admet pour valeurs propres les réels de la forme  $\lambda_i + \mu \geq \mu > 0$ ,  $1 \leq i \leq p$ . Puisqu'ainsi  $0$  n'en est pas valeur propre, la matrice  ${}^tJJ + \mu I_p$  est inversible.

**b.** À partir de  $J = US{}^tV$ , il vient  ${}^tJ = V{}^tS{}^tU$ . Par ailleurs, on a vu en **a.** que  ${}^tJJ + \mu I_p = V(D + \mu I_p){}^tV$  d'où  $({}^tJJ + \mu I_p)^{-1} = V(D + \mu I_p)^{-1}{}^tV$  puis finalement  $({}^tJJ + \mu I_p)^{-1}{}^tJ = V(D + \mu I_p)^{-1}{}^tS{}^tU$ . Il ne reste plus qu'à vérifier facilement à partir des définitions de  $R$  et  $S$  que  $(D + \mu I_p)^{-1}{}^tS = R$ .

c. Par un nouveau calcul par blocs,

$$({}^t\mathbf{J}\mathbf{J} + \mu\mathbf{I}_p)^{-1}{}^t\mathbf{J} = \mathbf{V}(\mathbf{R}{}^t\mathbf{U}) = \mathbf{V}{}^t(r_{1,1}\mathbf{U}_1 \cdots r_{p,p}\mathbf{U}_p) = \sum_{i=1}^p r_{i,i}\mathbf{V}_i{}^t\mathbf{U}_i = \sum_{i=1}^q \frac{\sqrt{\lambda_i}}{\lambda_i + \mu} \mathbf{V}_i{}^t\mathbf{U}_i.$$

6. a. Immédiat à partir de 3.a. et d'un théorème opératoire évident.

b. On obtient classiquement, d'après II.4.,

$$\forall h \in \mathbb{R}^p, \quad \nabla M(h) = \nabla L(h) + \mu h = \nabla F(\mathbf{X}) + (\mathbf{G}(\mathbf{X}) + \mu\mathbf{I}_p)h$$

et

$$\forall h \in \mathbb{R}^p, \quad \nabla^2 M(h) = \nabla^2 L(h) + \mu\mathbf{I}_p = \mathbf{G}(\mathbf{X}) + \mu\mathbf{I}_p.$$

c. D'après 5.a. appliqué à  $J(\mathbf{X})$ , la matrice  $\mathbf{G}(\mathbf{X}) + \mu\mathbf{I}_p = {}^tJ(\mathbf{X})J(\mathbf{X}) + \mu\mathbf{I}_p$  est inversible, si bien que :

$$\nabla M(h) = 0 \iff \nabla F(\mathbf{X}) + (\mathbf{G}(\mathbf{X}) + \mu\mathbf{I}_p)h = 0 \iff h = -(\mathbf{G}(\mathbf{X}) + \mu\mathbf{I}_p)^{-1}\nabla F(\mathbf{X}).$$

D'après la question 5.c. (et en conservant pour  $J(\mathbf{X})$  les notations introduites pour  $J$  tout au long de cette partie), la fonction  $M$  admet donc pour unique point critique :

$$h^* = -(\mathbf{G}(\mathbf{X}) + \mu\mathbf{I}_p)^{-1}{}^tJ(\mathbf{X})f(\mathbf{X}) = \sum_{i=1}^q \frac{\sqrt{\lambda_i}}{\lambda_i + \mu} \mathbf{V}_i{}^t\mathbf{U}_i f(\mathbf{X}).$$

d. La fonction  $M$  présente au point critique  $h^*$  un minimum local : en effet, sa matrice hessienne y est définie-positive puisque toutes ses valeurs propres sont strictement positives d'après 5.a..

