PARTIE I. Gradient et Hessienne.

Q1 Un premier exemple.

a) $\forall (x_1,x_2)\in\mathbb{R}^2$, $F(x_1,x_2)=\frac{1}{2}f_1^2(x_1,x_2)+\frac{1}{2}f_2^2(x_1,x_2)=\frac{1}{2}(x_1^2+x_2+1)^2+\frac{1}{2}(x_1+x_2^2+1)^2$.

Ainsi F est une fonction polynôme donc F est de classe C^2 sur $\mathbb{R}^2$.

Soit $(x_1,x_2)\in\mathbb{R}^2$. $\frac{\partial F}{\partial x_1}(x_1,x_2)=\frac{1}{2}\times 2(2x_1)(x_1^2+x_2+1)+\frac{1}{2}\times 2\times 1\times(x_1+x_2^2+1)$.

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

De même $\frac{\partial F}{\partial x_2}(x_1,x_2)=2x_2^3+2x_2x_1+x_1^2+3x_2+1$.

$\forall (x_1,x_2)\in\mathbb{R}^2$, $\nabla F(x_1,x_2)=(2x_1^3+2x_1x_2+x_2^2+3x_1+1,\ 2x_2^3+2x_2x_1+x_1^2+3x_2+1)$.

b) Soit $(x_1,x_2)\in\mathbb{R}^2$. Notons que :

$$\frac{\partial F}{\partial x_1}(x_1,x_2)-\frac{\partial F}{\partial x_2}(x_1,x_2)=2x_1^3+2x_1x_2+x_2^2+3x_1+1-2x_2^3-2x_2x_1-x_1^2-3x_2-1=2(x_1^3-x_2^3)-(x_1^2-x_2^2)+3(x_1-x_2)$$

$$\frac{\partial F}{\partial x_1}(x_1,x_2)-\frac{\partial F}{\partial x_2}(x_1,x_2)=(x_1-x_2)\left[2(x_1^2+x_1x_2+x_2^2)-(x_1+x_2)+3\right].$$

Alors $\nabla F(x_1,x_2)=0 \iff \frac{\partial F}{\partial x_1}(x_1,x_2)=\frac{\partial F}{\partial x_2}(x_1,x_2)=0 \iff \begin{cases}\frac{\partial F}{\partial x_1}(x_1,x_2)=0\\ \frac{\partial F}{\partial x_1}(x_1,x_2)-\frac{\partial F}{\partial x_2}(x_1,x_2)=0\end{cases}$.

$\nabla F(x_1,x_2)=0_{\mathbb{R}^2} \iff \begin{cases}2x_1^3+2x_1x_2+x_2^2+3x_1+1=0\\ (x_1-x_2)\left(2(x_1^2+x_1x_2+x_2^2)-(x_1+x_2)+3\right)=0\end{cases}$. Ainsi :

$\forall (x_1,x_2)\in\mathbb{R}^2$, $\nabla F(x_1,x_2)=0_{\mathbb{R}^2} \iff \begin{cases}2x_1^3+2x_1x_2+3x_1+x_2^2+1=0\\ (x_1-x_2)(2x_1^2+2x_1x_2+2x_2^2-x_1-x_2+3)=0\end{cases}$

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

$$2x_1^2+2x_1x_2+2x_2^2-x_1-x_2+3=(x_1+x_2)^2+x_1^2+x_2^2-x_1-x_2+3$$

$$2x_1^2+2x_1x_2+2x_2^2-x_1-x_2+3=(x_1+x_2)^2+\left(x_1-\tfrac{1}{2}\right)^2-\tfrac{1}{4}+\left(x_2-\tfrac{1}{2}\right)^2-\tfrac{1}{4}+3$$

$$2x_1^2+2x_1x_2+2x_2^2-x_1-x_2+3=(x_1+x_2)^2+\left(x_1-\tfrac{1}{2}\right)^2+\left(x_2-\tfrac{1}{2}\right)^2+\tfrac{5}{2}>0.$$

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

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

$$\nabla F(x_1,x_2)=0_{\mathbb{R}^2} \Leftrightarrow \begin{cases} 2x_1^3+2x_1x_2+3x_1+x_2^2+1=0 \\ (x_1-x_2)(2x_1^2+2x_1x_2+2x_2^2-x_1-x_2+3)=0 \end{cases} \Leftrightarrow \begin{cases} x_1=x_2 \\ 2x_1^3+2x_1^2+3x_1+x_1^2+1=0 \end{cases}$$

$$\nabla F(x_1,x_2)=0_{\mathbb{R}^2} \Leftrightarrow \begin{cases} x_1=x_2 \\ 2x_1^3+3x_1^2+3x_1+1=0 \end{cases}.$$

Or $2x_1^3+3x_1^2+3x_1+1=(2x_1+1)(x_1^2+x_1+1)=(2x_1+1)\left(\left(x_1+\frac{1}{2}\right)^2-\frac{1}{4}+1\right)$

$$\nabla F(x_1,x_2)=0_{\mathbb{R}^2} \Leftrightarrow \begin{cases} x_1=x_2 \\ (2x_1+1)\left(\left(x_1+\frac{1}{2}\right)^2+\frac{3}{4}\right)=0 \end{cases} \Leftrightarrow x_1=x_2=-\frac{1}{2}.$$

F admet un point critique et un seul : $\left(-\frac{1}{2},-\frac{1}{2}\right)$.

d) Soit $(x_1,x_2)\in\mathbb{R}^2$. $\frac{\partial F}{\partial x_1}(x_1,x_2)=2x_1^3+2x_1x_2+x_2^2+3x_1+1$.

Alors $\frac{\partial^2 F}{\partial x_1^2}(x_1,x_2)=6x_1^2+2x_2+3$ et $\frac{\partial^2 F}{\partial x_1\partial x_2}(x_1,x_2)=\frac{\partial^2 F}{\partial x_2\partial x_1}(x_1,x_2)=2x_1+2x_2$. De même :

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

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

Posons $r=\frac{\partial^2 F}{\partial x_1^2}\left(-\frac{1}{2},-\frac{1}{2}\right)$, $s=\frac{\partial^2 F}{\partial x_2\partial x_1}\left(-\frac{1}{2},-\frac{1}{2}\right)$ et $t=\frac{\partial^2 F}{\partial x_2^2}\left(-\frac{1}{2},-\frac{1}{2}\right)$.

$r=6\left(-\frac{1}{2}\right)^2+2\left(-\frac{1}{2}\right)+3=\frac{7}{2}$, $s=2\left(-\frac{1}{2}\right)+2\left(-\frac{1}{2}\right)=-2$ et $t=r=\frac{7}{2}$!

$r>0$ et $rt-s^2=\frac{49}{4}-4=\frac{33}{4}>0$. Ainsi :

F admet en $\left(-\frac{1}{2},-\frac{1}{2}\right)$ un minimum local.

c) f_1 et f_2 sont des fonctions polynômes donc f_1 et f_2 sont de classe $\mathcal{C}^2$ sur $\mathbb{R}^2$.

$\forall (x_1,x_2)\in\mathbb{R}^2$, $f_1(x_1,x_2)=x_1^2+x_2+1$ et $f_2(x_1,x_2)=x_1+x_2^2+1$.

$\forall (x_1,x_2)\in\mathbb{R}^2$, $\dfrac{\partial f_1}{\partial x_1}(x_1,x_2)=2x_1$, $\dfrac{\partial f_1}{\partial x_2}(x_1,x_2)=1$, $\dfrac{\partial f_2}{\partial x_1}(x_1,x_2)=1$ et $\dfrac{\partial f_2}{\partial x_2}(x_1,x_2)=2x_2$.

$\forall (x_1,x_2)\in\mathbb{R}^2$, $\dfrac{\partial^2 f_1}{\partial x_1^2}(x_1,x_2)=2$, $\dfrac{\partial^2 f_1}{\partial x_2\partial x_1}(x_1,x_2)=\dfrac{\partial^2 f_1}{\partial x_1\partial x_2}(x_1,x_2)=0$ et $\dfrac{\partial^2 f_1}{\partial x_2^2}(x_1,x_2)=0$.

$\forall (x_1,x_2)\in\mathbb{R}^2$, $\dfrac{\partial^2 f_2}{\partial x_1^2}(x_1,x_2)=0$, $\dfrac{\partial^2 f_2}{\partial x_2\partial x_1}(x_1,x_2)=\dfrac{\partial^2 f_2}{\partial x_1\partial x_2}(x_1,x_2)=0$ et $\dfrac{\partial^2 f_2}{\partial x_2^2}(x_1,x_2)=2$.

$\forall (x_1,x_2)\in\mathbb{R}^2$, $\nabla^2 f_1(x_1,x_2)=\begin{pmatrix}2&0\\0&0\end{pmatrix}$ et $\nabla^2 f_2(x_1,x_2)=\begin{pmatrix}0&0\\0&2\end{pmatrix}$.

Soit $X=(x_1,x_2)\in\mathbb{R}^2$. $J(X)=\begin{pmatrix}\dfrac{\partial f_1}{\partial x_1}(X) & \dfrac{\partial f_1}{\partial x_2}(X)\\ \dfrac{\partial f_2}{\partial x_1}(X) & \dfrac{\partial f_2}{\partial x_2}(X)\end{pmatrix}=\begin{pmatrix}2x_1&1\\1&2x_2\end{pmatrix}$.

$\forall X=(x_1,x_2)\in\mathbb{R}^2$, $J(X)=\begin{pmatrix}2x_1&1\\1&2x_2\end{pmatrix}$.

Soit $X=(x_1,x_2)\in\mathbb{R}^2$. ${}^tJ(X)f(X)=\begin{pmatrix}2x_1&1\\1&2x_2\end{pmatrix}\begin{pmatrix}f_1(X)\\f_2(X)\end{pmatrix}=\begin{pmatrix}2x_1&1\\1&2x_2\end{pmatrix}\begin{pmatrix}x_1^2+x_2+1\\x_1+x_2^2+1\end{pmatrix}$.

${}^tJ(X)f(X)=\begin{pmatrix}2x_1(x_1^2+x_2+1)+x_1+x_2^2+1\\ x_1^2+x_2+1+2x_2(x_1+x_2^2+1)\end{pmatrix}=\begin{pmatrix}\dfrac{\partial F}{\partial x_1}(X)\\ \dfrac{\partial F}{\partial x_2}(X)\end{pmatrix}$.

En tenant compte des identifications on a : $\forall X\in\mathbb{R}^2$, ${}^tJ(X)f(X)=\nabla F(X)$.

Soit $X=(x_1,x_2)\in\mathbb{R}^2$.

$G(X)={}^tJ(X)J(X)=\begin{pmatrix}2x_1&1\\1&2x_2\end{pmatrix}\begin{pmatrix}2x_1&1\\1&2x_2\end{pmatrix}=\begin{pmatrix}4x_1^2+1&2x_1+2x_2\\2x_1+2x_2&4x_2^2+1\end{pmatrix}$

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

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

$\forall X\in\mathbb{R}^2$, $G(X)+\sum_{i=1}^{2}f_i(X)\,\nabla^2 f_i(X)=\nabla^2F(X)$.

Q2 Un deuxième exemple.

a) $\forall (x_1,x_2)\in\mathbb{R}^2$, $F(x_1,x_2)=\frac{1}{2}\sum_{i=1}^{n}(f_i(x_1,x_2))^2=\frac{1}{2}\sum_{i=1}^{n}(a_ix_1+b_ix_2-c_i)^2$.

F est une fonction polynôme donc de classe C^1 dans $\mathbb{R}^2$ sur $\mathbb{R}^2$.

$\forall (x_1,x_2)\in\mathbb{R}^2$, $\frac{\partial F}{\partial x_1}(x_1,x_2)=\frac{1}{2}\sum_{i=1}^{n}2a_i(a_ix_1+b_ix_2-c_i)=x_1\sum_{i=1}^{n}a_i^2+x_2\sum_{i=1}^{n}a_ib_i-\sum_{i=1}^{n}a_ic_i$.

$\forall (x_1,x_2)\in\mathbb{R}^2$, $\frac{\partial F}{\partial x_1}(x_1,x_2)=x_1\|a\|^2+x_2\langle a,b\rangle-\langle a,c\rangle$. De même :

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

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

b) L'inégalité de Cauchy-Schwarz donne :

$(\langle a,b\rangle)^2\leq\|a\|^2\|b\|^2$.

Comme (a,b) est une famille libre : $(\langle a,b\rangle)^2<\|a\|^2\|b\|^2$.

Ainsi $\|a\|^2\|b\|^2-\langle a,b\rangle^2>0$.

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

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

$\|a\|^2\times\|b\|^2-\langle a,b\rangle\langle a,b\rangle>0$ donc la matrice $\begin{pmatrix}\|a\|^2 & \langle a,b\rangle\\ \langle a,b\rangle & \|b\|^2\end{pmatrix}$ est

inversible. Le cours indique que son inverse est :

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

$$\nabla F(x_1,x_2)=0\Leftrightarrow\begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}\|a\|^2 & \langle a,b\rangle\\ \langle a,b\rangle & \|b\|^2\end{pmatrix}^{-1}\begin{pmatrix}\langle a,c\rangle\\ \langle b,c\rangle\end{pmatrix}\Leftrightarrow\begin{pmatrix}x_1\\x_2\end{pmatrix}=\frac{1}{\alpha}\begin{pmatrix}\|b\|^2 & -\langle a,b\rangle\\ -\langle a,b\rangle & \|a\|^2\end{pmatrix}\begin{pmatrix}\langle a,c\rangle\\ \langle b,c\rangle\end{pmatrix}$$

$$\nabla F(x_1,x_2)=0\Leftrightarrow\begin{cases}x_1=\dfrac{\|b\|^2\langle a,c\rangle-\langle a,b\rangle\langle b,c\rangle}{\|a\|^2\|b\|^2-\langle a,b\rangle^2}\\[2ex] x_2=\dfrac{-\langle a,b\rangle\langle a,c\rangle+\|a\|^2\langle b,c\rangle}{\|a\|^2\|b\|^2-\langle a,b\rangle^2}\end{cases}.$$

F admet un point critique et un seul $(\hat{x}_1, \hat{x}_2)$ où $\begin{cases} \hat{x}_1 = \dfrac{\|b\|^2\langle a,c\rangle - \langle a,b\rangle\langle b,c\rangle}{\|a\|^2\|b\|^2 - \langle a,b\rangle^2} \\ \hat{x}_2 = \dfrac{\|a\|^2\langle b,c\rangle - \langle a,b\rangle\langle a,c\rangle}{\|a\|^2\|b\|^2 - \langle a,b\rangle^2} \end{cases}$.

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

Alors $\forall (x_1,x_2)\in\mathbb{R}^2$, $\dfrac{\partial^2 F}{\partial x_1^2}(x_1,x_2) = \|a\|^2$, $\dfrac{\partial^2 F}{\partial x_1\partial x_2}(x_1,x_2) = \dfrac{\partial^2 F}{\partial x_2\partial x_1}(x_1,x_2) = \langle a,b\rangle$ et $\dfrac{\partial^2 F}{\partial x_2^2}(x_1,x_2) = \|b\|^2$

Ainsi $\forall (x_1,x_2)\in\mathbb{R}^2$, $\nabla^2 F(x_1,x_2) = \begin{pmatrix} \|a\|^2 & \langle a,b\rangle \\ \langle a,b\rangle & \|b\|^2 \end{pmatrix}$.

Posons $r = \dfrac{\partial^2 F}{\partial x_1^2}(\hat{x}_1,\hat{x}_2)$, $s = \dfrac{\partial^2 F}{\partial x_2\partial x_1}(\hat{x}_1,\hat{x}_2)$ et $t = \dfrac{\partial^2 F}{\partial x_2^2}(\hat{x}_1,\hat{x}_2)$

$r = \|a\|^2 > 0$ (a n'est pas nul car (a,b) est libre) et $rt - s^2 = \|a\|^2\|b\|^2 - (\langle a,b\rangle)^2 > 0$.

Ainsi <u>F admet en $(\hat{x}_1,\hat{x}_2)$ un minimum local.</u>

<u>Remarque</u>.. Pour tout (x_1,x_2) dans $\mathbb{R}^2$ les valeurs propres de $\nabla^2 F(x_1,x_2)$ sont (strictement) positive donc F est convexe. $(\hat{x}_1,\hat{x}_2)$ étant un point critique de F, F admet un minimum global en $(\hat{x}_1,\hat{x}_2)$ non ?

d) Notons H le sous-espace vectoriel de $\mathbb{R}^n$ engendré par (a,b).
Le théorème de meilleure approximation montre que $\min\limits_{h\in H}\|h-c\|$ existe et est atteint en un point (t_1,t_2) et un seul qui est la projection orthogonale de c sur H.

Alors $\min\limits_{h\in H}\|h-c\|^2$ existe et est atteint en un point et un seul : (t_1,t_2).

Donc $\min\limits_{(x_1,x_2)\in\mathbb{R}^2}\|x_1a + x_2b - c\|^2$ existe et est atteint en un point et un seul : (t_1,t_2).

Or $\forall (x_1,x_2)\in\mathbb{R}^2$, $\|x_1a + x_2b - c\|^2 = \sum\limits_{i=1}^n (x_1a_i + x_2b_i - c_i)^2 = \sum\limits_{i=1}^n (f_i(x_1,x_2))^2 = 2F(x_1,x_2)$.

$\forall (x_1,x_2)\in\mathbb{R}^2$, $F(x_1,x_2) = \frac{1}{2}\|x_1a + x_2b - c\|^2$.

Ainsi F admet un minimum sur $\mathbb{R}^2$ atteint en un point seul (t_1, t_2).

$\mathbb{R}^2$ étant un ouvert (et F étant de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$), (t_1, t_2) est un point critique de F

Alors $(t_1, t_2) = (\hat{x}_1, \hat{x}_2)$.

Finalement <u>F admet un minimum global en $(\hat{x}_1, \hat{x}_2)$.</u>

(Q3) <u>Un troisième exemple.</u>

<u>a)</u> $\forall (x_1, x_2) \in \mathbb{R}^2$, $F(x_1, x_2) = \frac{1}{2} \sum_{i=1}^{n} (x_1 + x_2 - c_i)^2$. F est une fonction polynôme donc

F est de classe $\mathcal{C}^2$ sur $\mathbb{R}^2$.

$\forall (x_1, x_2) \in \mathbb{R}^2$, $\frac{\partial F}{\partial x_1}(x_1, x_2) = \frac{1}{2} \sum_{i=1}^{n} 2 \times 1 \times (x_1 + x_2 - c_i) = n(x_1 + x_2) - \sum_{i=1}^{n} c_i = n(x_1 + x_2 - \bar{c})$.

$\forall (x_1, x_2) \in \mathbb{R}^2$, $\frac{\partial F}{\partial x_2}(x_1, x_2) = \frac{1}{2} \sum_{i=1}^{n} 2 \times 1 (x_1 + x_2 - c_i) = n(x_1 + x_2) - \sum_{i=1}^{n} c_i = n(x_1 + x_2 - \bar{c})$

$\forall (x_1, x_2) \in \mathbb{R}^2$, $\nabla F(x_1, x_2) = 0 \iff \frac{\partial F}{\partial x_1}(x_1, x_2) = \frac{\partial F}{\partial x_2}(x_1, x_2) = 0 \iff x_1 + x_2 = \bar{c} \iff x_2 = x_1 - \bar{c}$

<u>L'ensemble des points critiques de F est $\{(x_1, x_1 - \bar{c}) \; ; \; x_1 \in \mathbb{R}\}$.</u>

<u>b)</u> Soit $(\hat{x}_1, \hat{x}_2)$ un point critique de F. $\hat{x}_1 + \hat{x}_2 = \bar{c}$.

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

<u>$F(\hat{x}_1, \hat{x}_2) = \frac{n}{2} s^2$.</u>

Soit $(x_1, x_2) \in \mathbb{R}^2$. $2F(x_1, x_2) = \sum_{i=1}^{n} (x_1 + x_2 - c_i)^2 = \sum_{i=1}^{n} \left((x_1 + x_2 - \bar{c}) + (\bar{c} - c_i) \right)^2$

$2F(x_1, x_2) = \sum_{i=1}^{n} (x_1 + x_2 - \bar{c})^2 + 2 \sum_{i=1}^{n} (x_1 + x_2 - \bar{c})(\bar{c} - c_i) + \sum_{i=1}^{n} (\bar{c} - c_i)^2$

$2F(x_1, x_2) = n(x_1 + x_2 - \bar{c})^2 + 2(x_1 + x_2 - \bar{c}) \sum_{i=1}^{n} (\bar{c} - c_i) + \sum_{i=1}^{n} (c_i - \bar{c})^2$

Notons que $\sum_{i=1}^{n} (\bar{c} - c_i) = n\bar{c} - \sum_{i=1}^{n} c_i = n\bar{c} - n\bar{c} = 0$ et $\sum_{i=1}^{n} (c_i - \bar{c})^2 = n s^2$

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

Finalement <u>$\forall (x_1, x_2) \in \mathbb{R}^2$, $F(x_1, x_2) - F(\hat{x}_1, \hat{x}_2) = \frac{n}{2} (x_1 + x_2 - \bar{c})^2$.</u>

c) Dans ces conditions $\forall (x_1,x_2)\in\mathbb{R}^2$, $F(x_1,x_2)-F(\hat{x}_1,\hat{x}_2)\geq 0$ si $(\hat{x}_1,\hat{x}_2)$ est un point critique de F.

Donc <u>F admet un minimum global sur $\mathbb{R}^2$ qui vaut $\frac{n}{2}s^2$ et qui est atteint à chaque point critique de F.</u>

Observons que $\{x_1+x_2 \,;\, (x_1,x_2)\in\mathbb{R}^2\}=\mathbb{R}$!

Rappelons que $\forall (x_1,x_2)\in\mathbb{R}^2$, $F(x_1,x_2)=\frac{1}{2}\sum_{i=1}^{n}(x_1+x_2-c_i)^2$.

Donc F admet un minimum global sur $\mathbb{R}^2$ si $\varphi : x\mapsto \frac{1}{2}\sum_{i=1}^{n}(x-c_i)^2$ admet un minimum sur $\mathbb{R}$. En cas d'existence les deux minimums sont égaux.

φ est dérivable sur $\mathbb{R}$ et $\forall x\in\mathbb{R}$, $\varphi'(x)=\sum_{i=1}^{n}(x-c_i)=n(x-\bar{c})$.

φ est strictement croissante sur $[\bar{c},+\infty[$ et strictement décroissante sur $]-\infty,\bar{c}]$.

Ainsi φ admet un minimum sur $\mathbb{R}$ atteint en le seul point $\bar{c}$ qui vaut $\varphi(\bar{c})$. Or $\varphi(\bar{c})=\frac{1}{2}\sum_{i=1}^{n}(\bar{c}-c_i)^2=\frac{n}{2}\frac{1}{n}\sum_{i=1}^{n}(\bar{c}-c_i)^2=\frac{n}{2}s^2$.

Dès lors on peut dire que F admet un minimum global sur $\mathbb{R}^2$, qui vaut $\frac{n}{2}s^2$ et atteint en tous les points (x_1,x_2) de $\mathbb{R}^2$ tels que $x_1+x_2=\bar{c}$ c'est à dire en tous les points critiques de F.

<u>Exercice 1.</u> Soit X une variable aléatoire admettant une variance.

Montrer que $\min_{a\in\mathbb{R}} E((X-a)^2)$ existe et vaut $V(X)$.

<u>Exercice 2.</u> Traiter $\min_{x\in\mathbb{R}}\left(\frac{1}{2}\sum_{i=1}^{n}(x-c_i)^2\right)$ par projection orthogonale dans $\mathbb{R}^n$ (... sur $\mathrm{Vect}(c)$, où $c=(c_1,c_2,\dots,c_n)$) !!

Q1 a) $\forall X \in \mathbb{R}^p,\quad F(X) = \frac{1}{2}\sum_{j=1}^{n} (f_j(X))^2.$

Pour tout $j \in [\![1,n]\!]$, f_j est de classe $\mathcal{C}^2$ sur $\mathbb{R}^p$ donc pour tout $j \in [\![1,n]\!]$, f_j^2 est de classe $\mathcal{C}^2$ sur $\mathbb{R}^p$.

Alors F est de classe $\mathcal{C}^2$ sur $\mathbb{R}^p$ comme combinaison linéaire de n fonctions de classe $\mathcal{C}^2$ sur $\mathbb{R}^p$.

$\forall i \in [\![1,p]\!]$, $\frac{\partial F}{\partial x_i}(X) = \frac{1}{2}\sum_{j=1}^{n} 2 f_j(X)\frac{\partial f_j}{\partial x_i}(X) = \sum_{j=1}^{n}\frac{\partial f_j}{\partial x_i}(X) f_j(X)$ pour tout $X \in \mathbb{R}^p$.

Soit $X \in \mathbb{R}^p$. Posons un instant (!) ${}^tJ(X) = (a_{ij})_{\substack{1\le i\le p\\ 1\le j\le n}}$ et $\begin{pmatrix} t_1\\ t_2\\ \vdots\\ t_p\end{pmatrix} = {}^tJ(X)\begin{pmatrix} f_1(X)\\ f_2(X)\\ \vdots\\ f_n(X)\end{pmatrix}$.

Alors $\forall (i,j) \in [\![1,p]\!]\times[\![1,n]\!]$, $a_{ij} = \frac{\partial f_j}{\partial x_i}(X)$ car $J(X) = \left(\frac{\partial f_i}{\partial x_j}(X)\right)_{\substack{1\le i\le n\\ 1\le j\le p}}$.

Ainsi $\forall i \in [\![1,p]\!]$, $t_i = \sum_{j=1}^{n} a_{ij} f_j(X) = \sum_{j=1}^{n}\frac{\partial f_j}{\partial x_i}(X) f_j(X) = \frac{\partial F}{\partial x_i}(X)$.

Alors $\begin{pmatrix}\frac{\partial F}{\partial x_1}(X)\\ \vdots\\ \frac{\partial F}{\partial x_p}(X)\end{pmatrix} = {}^tJ(X)\begin{pmatrix} f_1(X)\\ \vdots\\ f_p(X)\end{pmatrix}$. Avec les identifications proposées il vient :

$\forall X \in \mathbb{R}^p,\quad \nabla F(X) = {}^tJ(X)\, f(X).$

b) Nous avons vu que $\forall i \in [\![1,p]\!]$, $\forall X \in \mathbb{R}^p$, $\frac{\partial F}{\partial x_i}(X) = \sum_{\ell=1}^{n}\frac{\partial f_\ell}{\partial x_i}(X) f_\ell(X)$

Alors $\forall i \in [\![1,p]\!]$, $\forall j \in [\![1,p]\!]$, $\forall X \in \mathbb{R}^p$, $\frac{\partial^2 F}{\partial x_j \partial x_i}(X) = \sum_{\ell=1}^{n}\left(\frac{\partial^2 f_\ell}{\partial x_j \partial x_i}(X) f_\ell(X) + \frac{\partial f_\ell}{\partial x_i}(X)\frac{\partial f_\ell}{\partial x_j}(X)\right)$

$\forall i \in [\![1,p]\!]$, $\forall j \in [\![1,p]\!]$, $\forall X \in \mathbb{R}^p$, $\frac{\partial^2 F}{\partial x_j \partial x_i}(X) = \sum_{\ell=1}^{n}\frac{\partial^2 f_\ell}{\partial x_j \partial x_i}(X) f_\ell(X) + \sum_{\ell=1}^{n}\frac{\partial f_\ell}{\partial x_i}(X)\frac{\partial f_\ell}{\partial x_j}(X)$

Soit $X \in \mathbb{R}^p$. Un instant encore posons $G(X) + \sum_{\ell=1}^{n} f_\ell(X)\,\nabla^2 f_\ell(X) = (\alpha_{ij})_{\substack{1\le i\le p\\ 1\le j\le p}}$

$J(X) = (\beta_{ij})_{\substack{1\le i\le n\\ 1\le j\le p}}$ et ${}^tJ(X) = (\gamma_{ij})_{\substack{1\le i\le p\\ 1\le j\le n}}$.

$$G(X) = {}^tJ(X)J(X) = \left(\sum_{k=1}^{n} \alpha_{ik}\,\beta_{kj} \right)_{\substack{1\le i\le p\\1\le j\le p}} = \left(\sum_{\ell=1}^{n} \frac{\partial f_\ell}{\partial x_i}(X)\,\frac{\partial f_\ell}{\partial x_j}(X) \right)_{\substack{1\le i\le p\\1\le j\le p}}.$$

rappelons que $\forall k \in [\![1,n]\!]$, $\nabla^2 f_k(X) = \left(\frac{\partial^2 f_k}{\partial x_j \partial x_i}(X) \right)_{\substack{1\le i\le p\\1\le j\le p}}$ et que

$(\alpha_{ij})_{\substack{1\le i\le p\\1\le j\le p}} = {}^tJ(X)J(X) + \sum_{k=1}^{n} f_k(X)\,\nabla^2 f_k(X)$. Alors :

$$\forall (i,j) \in [\![1,p]\!]^2,\ \alpha_{ij} = \sum_{\ell=1}^{n} \frac{\partial f_\ell}{\partial x_i}(X)\,\frac{\partial f_\ell}{\partial x_j}(X) + \sum_{\ell=1}^{n} f_\ell(X)\,\frac{\partial^2 f_\ell}{\partial x_j \partial x_i}(X)$$

Donc $\forall (i,j) \in [\![1,p]\!]^2$, $\alpha_{ij} = \sum_{\ell=1}^{n} \frac{\partial f_\ell}{\partial x_i}(X)\,\frac{\partial f_\ell}{\partial x_j}(X) + \sum_{\ell=1}^{n} \frac{\partial^2 f_\ell}{\partial x_j \partial x_i}(X)\,f_\ell(X) = \frac{\partial^2 F}{\partial x_j \partial x_i}(X)$.

plus de doute : $\forall X \in \mathbb{R}^p$, $\nabla^2 F(X) = G(X) + \sum_{k=1}^{n} f_k(X)\,\nabla^2 f_k(X)$ ou (!!)

$$\underline{\underline{\forall X \in \mathbb{R}^p,\ \nabla^2 F(X) = G(X) + \sum_{i=1}^{n} f_i(X)\,\nabla^2 f_i(X).}}$$

Partie II. Une approximation de F

Q1 Soit $h \in \mathbb{R}^p$. $2L(h) = \|\ell(h)\|^2 = {}^t\ell(h)\ell(h) = {}^t(f(X) + J(X)h)(f(X) + J(X)h)$

$$2L(h) = ({}^tf(X) + {}^th\,{}^tJ(X))(f(X) + J(X)h) = \underbrace{{}^tf(X)f(X)}_{\|f(X)\|^2} + \underbrace{{}^tf(X)J(X)}_{{}^t(\nabla F(X))}h + {}^th\,\underbrace{{}^tJ(X)f(X)}_{\nabla F(X)} + {}^th\,\underbrace{{}^tJ(X)J(X)}_{G(X)}h$$

$2L(h) = \|f(X)\|^2 + \langle \nabla F(X), h\rangle + \langle h, \nabla F(X)\rangle + {}^th\,G(X)h$.

$2L(h) = 2F(X) + 2\langle h, \nabla F(X)\rangle + {}^th\,G(X)h = 2F(X) + 2\,{}^th\,\nabla F(X) + {}^th\,G(X)h$.

Donc $\forall h \in \mathbb{R}^p,\ L(h) = F(X) + {}^th\,\nabla F(X) + \frac{1}{2}{}^th\,G(X)h$.

Q2 a) P est une matrice symétrique et à coefficients réels donc **P est diagonalisable !!**

b) Soit $(h_1, h_2, \dots, h_p)$ une base orthonormée de $\mathbb{R}^p$ (!!) constituée de vecteurs propres de P respectivement associées aux valeurs propres $\lambda_1, \lambda_2, \dots, \lambda_p$.

$\{\lambda_1, \lambda_2, \dots, \lambda_p\} = \{\theta_1, \theta_2, \dots, \theta_r\}$ donc $\theta = \max\limits_{1 \le j \le p} |\theta_j| = \max\limits_{1 \le j \le r} |\lambda_j|$.

Soit $h \in \mathbb{R}^p$.

$\exists (\alpha_1, \alpha_2, \dots, \alpha_p) \in \mathbb{R}^p$, $h = \sum_{i=1}^p \alpha_i h_i$. $Ph = \sum_{i=1}^p \alpha_i P h_i = \sum_{i=1}^p \alpha_i \lambda_i h_i$.

Alors $|{}^th P h| = |\langle h, Ph\rangle| = \left|\sum_{i=1}^p \alpha_i(\alpha_i \lambda_i)\right| \le \sum_{i=1}^p |\lambda_i| \alpha_i^2 \le \theta \sum_{i=1}^p \alpha_i^2 = \theta \|h\|^2$.

$(h_1, h_2, \dots, h_p)$ est orthonormée

$|\lambda_i| \le \theta$ et $\alpha_i^2 \ge 0$.

$\forall h \in \mathbb{R}^p,\ |{}^th P h| \le \theta \|h\|^2$.

Q3 a) forme quadratique associée à $\nabla^2 F(X)$

$$F(X+h) \underset{h \to 0}{=} F(X) + \langle \nabla F(X), h\rangle + \frac{1}{2} q_X(h) + o(\|h\|^2) \text{ ou}$$

$$F(X+h) \underset{h \to 0}{=} F(X) + \langle \nabla F(X), h\rangle + \frac{1}{2}{}^th\,\nabla^2 F(X)h + o(\|h\|^2).$$

b) $\forall h \in \mathbb{R}^p$, $L(h) = F(X) + {}^th\,\nabla F(X) + \frac{1}{2}{}^th\,G(X)h = F(X) + \langle \nabla F(X), h\rangle + \frac{1}{2}{}^th\,G(X)h$.

Alors $F(X+h) - L(h) \underset{h \to 0}{=} \frac{1}{2}{}^th\,(G(X) - \nabla^2 F(X))h + o(\|h\|^2)$

Alors $\dfrac{F(X+h)-L(h)}{\|h\|} \underset{h\to 0}{=} \dfrac{1}{2}\dfrac{{}^th(G(X)-\nabla^2F(X))h}{\|h\|} + o(\|h\|)$.

Ainsi pour montrer que $\lim\limits_{h\to 0}\dfrac{F(X+h)-L(h)}{\|h\|}=0$ il suffit de montrer que

$\lim\limits_{h\to 0}\dfrac{{}^th(G(X)-\nabla^2F(X))h}{\|h\|}=0$.

$G(X)={}^tJ(X)J(X)$ est une matrice symétrique de $M_p(\mathbb{R})$. $\nabla^2F(X)$ est également une matrice symétrique de $M_p(\mathbb{R})$. Posons $\hat{P}=G(X)-\nabla^2F(X)$.

Par différence $\hat{P}$ est une matrice symétrique de $M_p(\mathbb{R})$. Posons $\hat{\theta}=\max\limits_{\lambda\in S_p\hat{P}}|\lambda|$.

d'après a), $\forall h\in\mathbb{R}^p$, $|{}^th\hat{P}h|\leq\hat{\theta}\|h\|^2$.

$\forall h\in\mathbb{R}^p-\{0\}$, $\left|\dfrac{{}^th\hat{P}h}{\|h\|}\right|\leq\hat{\theta}\|h\|$ et $\lim\limits_{h\to 0}(\hat{\theta}\|h\|)=0$.

Alors par encadrement $\lim\limits_{h\to 0}\dfrac{{}^th\hat{P}h}{\|h\|}=0$. Donc $\lim\limits_{h\to 0}\dfrac{{}^th(G(X)-\nabla^2F(X))h}{\|h\|}=0$

ceci achève de montrer que : $\lim\limits_{h\to 0}\dfrac{F(X+h)-L(h)}{\|h\|}=0$.

Notons que φ_1 et φ_2 sont de classe C^1 comme fonctions polynômes.

Q4) a) $\forall h=(h_1,h_2,\dots,h_p)\in\mathbb{R}^p$, $\varphi_1(h)={}^th\,\nabla F(X)=\sum\limits_{k=1}^{p}h_k\dfrac{\partial F}{\partial x_k}(X)$.

Alors $\forall h\in\mathbb{R}^p$, $\dfrac{\partial\varphi_1}{\partial h_i}(h)=\dfrac{\partial F}{\partial x_i}(X)$ et ceci pour tout $i\in[\![1,p]\!]$.

$\forall i\in[\![1,p]\!],\ \forall h\in\mathbb{R}^p,\ \dfrac{\partial\varphi_1}{\partial h_i}(h)=\dfrac{\partial F}{\partial x_i}(X)$ ou $\forall h\in\mathbb{R}^p$, $\nabla\varphi_1(h)=\nabla F(X)$.

Soit $h=(h_1,h_2,\dots,h_p)\in\mathbb{R}^p$. $\varphi_2(h)={}^thG(X)h=\sum\limits_{i=1}^{p}\sum\limits_{j=1}^{p}g_{i,j}(X)h_ih_j$

Fixons k dans $[\![1,p]\!]$.

$\varphi_2(h)=g_{k,k}(X)h_k^2+\sum\limits_{\substack{i=1\\i\neq k}}^{p}g_{i,k}(X)h_ih_k+\sum\limits_{\substack{j=1\\j\neq k}}^{p}g_{k,j}(X)h_kh_j+\sum\limits_{\substack{i=1\\i\neq k}}^{p}\sum\limits_{\substack{j=1\\j\neq k}}^{p}g_{i,j}(X)h_ih_j$.

Alors $\dfrac{\partial\varphi_2}{\partial h_k}(h)=2h_kg_{k,k}(X)+\sum\limits_{\substack{i=1\\i\neq k}}^{p}g_{i,k}(X)h_i+\sum\limits_{\substack{j=1\\j\neq k}}^{p}g_{k,j}(X)h_j=\sum\limits_{i=1}^{p}g_{i,k}(X)h_i+\sum\limits_{j=1}^{p}g_{k,j}(X)h_j$

On a alors $\frac{\partial \varphi_2}{\partial h_\ell}(h) = \sum_{i=1}^{p} g_{i,\ell}(x) h_i + \sum_{i=1}^{p} g_{\ell,i}(x) h_i \underset{\substack{\uparrow \\ G(x) \text{ est symétrique}}}{=} 2\sum_{i=1}^{p} g_{\ell,i}(x) h_i$.

$$\forall \ell \in [\![1,p]\!],\ \forall h \in \mathbb{R}^p,\quad \frac{\partial \varphi_2}{\partial h_\ell}(h) = 2\sum_{i=1}^{p} g_{\ell,i}(x) h_i = 2\sum_{i=1}^{p} g_{i,\ell}(x) h_i.$$

Ou $\forall j \in [\![1,p]\!],\ \forall h \in \mathbb{R}^p,\quad \frac{\partial \varphi_2}{\partial h_j}(h) = 2\sum_{i=1}^{p} g_{j,i}(x) h_i = 2\sum_{i=1}^{p} g_{i,j}(x) h_i.$

Notons que ceci signifie aussi que $\underline{\forall h \in \mathbb{R}^p,\ \nabla \varphi_2(h) = 2\,G(x)\,h}$

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

L est de classe C^1 sur $\mathbb{R}^p$ comme fonction polynôme.

$\forall h \in \mathbb{R}^p,\ \nabla L(h) = 0_{\mathbb{R}^p} + \nabla \varphi_1(h) + \frac{1}{2}\nabla \varphi_2(h)$ ($h \mapsto F(x)$ est constante !).

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

c) L est de classe C^2 sur $\mathbb{R}^p$ comme fonction polynôme. Soit $(i,j) \in [\![1,p]\!]^2$.

$\forall h = (h_1, \dots, h_p) \in \mathbb{R}^p,\quad \frac{\partial L}{\partial x_i}(h) = \frac{\partial F}{\partial x_i}(x) + \sum_{k=1}^{p} g_{i,k}(x) h_k$.

donc $\forall h \in \mathbb{R}^p,\quad \frac{\partial^2 L}{\partial x_j \partial x_i}(h) = 0 + \sum_{\substack{\ell=1 \\ \ell \neq j}}^{p} g_{i,\ell}(x) \times 0 + g_{ij}(x) = g_{i,j}(x)$.

Ainsi $\underline{\forall h \in \mathbb{R}^p,\ \nabla^2 L(h) = G(x)}$.

Q5) a) $J \in \mathcal{M}_{n,p}(\mathbb{R})$ et ${}^tJ \in \mathcal{M}_{p,n}(\mathbb{R})$ donc ${}^tJJ \in \mathcal{M}_p(\mathbb{R})$.

${}^t({}^tJJ) = {}^tJ\,{}^t({}^tJ) = {}^tJJ$ donc ${}^tJJ$ est symétrique.

${}^tJJ$ est une matrice symétrique de $\mathcal{M}_p(\underline{\mathbb{R}})$; ${}^tJJ$ est diagonalisable.

Soit λ une valeur propre de ${}^tJJ$. $\exists x \in \mathbb{R}^p$ (!!), ${}^tJJx = \lambda x$ et $x \neq 0_{\mathbb{R}^p}$.

$\lambda \|x\|^2 = \lambda\,{}^tx x = {}^tx(\lambda x) = {}^tx\,{}^tJJx = {}^t(Jx) Jx = \|Jx\|^2$. $\lambda \underset{\substack{\uparrow \\ x \neq 0}}{=} \frac{\|Jx\|^2}{\|x\|^2} \geqslant 0$.

Ainsi $\underline{\text{les valeurs propres de } {}^tJJ \text{ sont positives ou nulles}}$.

b) Soit $\mathcal{B}_p$ (resp. $\mathcal{B}_n$) la base canonique de $\mathbb{R}^p$ (resp. $\mathbb{R}^n$).

Soit φ l'application linéaire de $\mathbb{R}^p$ dans $\mathbb{R}^n$ telle que $M(\varphi, \mathcal{B}_p, \mathcal{B}_n) = J$.

Supposons ${}^tJJ$ inversible. Soit $x \in \operatorname{Ker} \varphi$.

Comme on a fait tout $Jx = 0$; alors ${}^tJJx = 0$.

Supposons ${}^tJJ$ inversible. Alors $x = 0_{\mathbb{R}^p}$. Ainsi $\operatorname{Ker} \varphi = \{0_{\mathbb{R}^p}\}$.

Alors $p = \dim \operatorname{Ker} \varphi + \operatorname{rg} \varphi = \operatorname{rg} \varphi$. Donc $\operatorname{rg} J = \operatorname{rg} \varphi = p$.

<u>Si ${}^tJJ$ est inversible : $\operatorname{rg} J = p$.</u>

(Q6) Soit $\hat{h}$ un point critique de L. $\nabla L(\hat{h}) = 0$

Ainsi $\nabla F(X) + G(X)\hat{h} = 0$.

$\langle \hat{h}, \nabla F(X) \rangle = -\langle \hat{h}, G(X)\hat{h} \rangle = -{}^t\hat{h}\, G(X)\hat{h} = -{}^t\hat{h}\, {}^tJ(X)J(X)\hat{h} = -\|J(X)\hat{h}\|^2$

Donc <u>si $\hat{h}$ est un point critique de L : $\langle \hat{h}, \nabla F(X) \rangle \leqslant 0$.</u>

(Q7) a) On suppose que $G(X)$ est inversible. Soit $h \in \mathbb{R}^p$.

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

<u>L a donc un point critique et un seul : $\hat{h} = -(G(X))^{-1}\nabla F(X) \underset{\uparrow}{=} -(G(X))^{-1}\,{}^tJ(X)f(X)$</u>

IQ4a

b) Nous savons déjà que $\langle \hat{h}, \nabla F(X) \rangle \leqslant 0$.

$\boxed{V1}$ Supposons que $\langle \hat{h}, \nabla F(X) \rangle = 0$.

Alors ${}^t(\nabla F(X))\hat{h} = 0$. ${}^t(\nabla F(X))(-(G(X))^{-1})\nabla F(X) = 0$.

Posons $u = \nabla F(X)$. Alors $-{}^tu\,(G(X))^{-1}u = 0$ ou ${}^tu\,(G(X))^{-1}u = 0$.

$G(X) = {}^tJ(X)J(X)$ est une matrice symétrique à valeurs propres positives ou nulles

et $G(X)$ est inversible. Donc $G(X)$ est une matrice symétrique à valeurs propres

strictement positives. Il en est de même de $(G(X))^{-1}$.

Soit $(u_1, \dots, u_p)$ une base orthonormée de $\mathbb{R}^p$ constituée de vecteurs propres

de $(G(X))^{-1}$ associés aux valeurs propres $\beta_1, \dots, \beta_p$.

$\exists(\gamma_1,\dots,\gamma_p)\in\mathbb{R}^p$, $u=\sum_{i=1}^{p}\gamma_i u_i$. $(G(X))^{-1}u=\sum_{i=1}^{p}\gamma_i\beta_i u_i$

Alors $0={}^t u\,(G(X))^{-1}u=\sum_{i=1}^{p}\gamma_i\gamma_i\beta_i=\sum_{i=1}^{p}\gamma_i^2\beta_i$

Or $\forall i\in[\![1,p]\!]$, $\gamma_i^2\beta_i\geq 0$.

Ainsi $\forall i\in[\![1,p]\!]$, $\gamma_i^2\beta_i=0$ et $\beta_i>0$. Alors $\forall i\in[\![1,p]\!]$, $\gamma_i^2=0$.

$\forall i\in[\![1,p]\!]$, $\gamma_i=0$. $u=0$. Donc $\nabla F(X)=0$ ce qui est contraire à l'hypothèse.

Donc $\langle\hat{h},\nabla F(X)\rangle\neq 0$ et $\langle\hat{h},\nabla F(X)\rangle\leqslant 0$ $\boxed{V2}$

Alors $\langle\hat{h},\nabla F(X)\rangle<0$. $\hat{h}$ est une direction de décroissance de F en X.

$\nabla^2 L(\hat{h})=G(X)$ et $G(X)$ a toutes ses valeurs propres strictement positives.

Ainsi L admet en $\hat{h}$ un minimum local (stricte).

$\boxed{V2}$ Supposons que $\langle\hat{h},\nabla F(X)\rangle=0$.

Alors d'après ce que nous avons vu dans Q6 $\|J(X)\hat{h}\|^2=0$.

Donc $J(X)\hat{h}=0$. ${}^tJ(X)J(X)\hat{h}=0$. Alors $G(X)\hat{h}=0$.

Alors $\nabla F(X)=-G(X)\hat{h}=0$. Ceci est impossible.

Partie III Une décomposition d'une matrice rectangulaire

(Q1) Notons $\alpha_1, \alpha_2, \dots, \alpha_{r'}$ les valeurs propres distinctes de la matrice ${}^tJJ$, données dans l'ordre décroissant. Rappelons que ${}^tJJ$ est une matrice symétrique de $M_p(\mathbb{R})$ dont les valeurs propres sont positives ou nulles. Ainsi $\alpha_1 > \alpha_2 > \dots > \alpha_{r'} \geq 0$.

Pour tout $i \in [\![1, r']\!]$ considérons une base orthonormée $\mathcal{B}_i$ du sous-espace propre de ${}^tJJ$ associé à la valeur propre α_i. Alors $\mathcal{B} = $ "$\mathcal{B}_1 \cup \mathcal{B}_2 \cup \dots \cup \mathcal{B}_{r'}$" est une base orthonormée de $M_{p,1}(\mathbb{R})$ constituée de vecteurs propres de ${}^tJJ$ respectivement associés aux valeurs propres $\lambda'_1, \lambda'_2, \dots, \lambda'_p$ car $SEP({}^tJJ, \alpha_1), SEP({}^tJJ, \alpha_2), \dots, SEP({}^tJJ, \alpha_{r'})$ sont deux à deux orthogonaux et $SEP({}^tJJ, \alpha_1) \oplus SEP({}^tJJ, \alpha_2) \oplus \dots \oplus SEP({}^tJJ, \alpha_{r'}) = M_{p,1}(\mathbb{R})$.

Comme $\alpha_1 > \alpha_2 > \dots > \alpha_{r'} \geq 0$, par construction de $\mathcal{B}$: $\lambda'_1 \geq \lambda'_2 \geq \dots \geq \lambda'_p$.

Notons V la matrice de passage de la base canonique de $M_{p,1}(\mathbb{R})$ à $\mathcal{B}$.

V est une matrice orthogonale et ${}^tV({}^tJJ)V$ est la matrice diagonale

$D = \mathrm{Diag}(\lambda'_1, \lambda'_2, \dots, \lambda'_p)$.

1er Cas.. $\alpha_{r'} > 0$. Posons alors $q = p$ et $\forall i \in [\![1, r']\!], \lambda_i = \lambda'_i$

Alors $1 \leq q \leq p$ et $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_q > 0$.

V, q, $\lambda_1, \lambda_2, \dots, \lambda_q$ répondent à la question !

2e Cas.. $\alpha_{r'} = 0$. Notons que ${}^tJJ$ ne peut pas avoir comme seule valeur propre 0

(*) (sinon ${}^tJJ$ étant diagonalisable, ${}^tJJ$ serait la matrice nulle et ainsi J serait nulle). Ainsi $r' \geq 2$.

Notons q le nombre d'éléments de "$\mathcal{B}_1 \cup \mathcal{B}_2 \cup \dots \cup \mathcal{B}_{r'-1}$"

Alors $\lambda'_1 \geq \lambda'_2 \geq \dots \geq \lambda'_q > 0$ et $\lambda'_{q+1} = \lambda'_{q+2} = \dots = \lambda'_p = 0$.

Posons $\forall i \in [\![1, q]\!], \lambda_i = \lambda'_i$. $1 \leq q \leq p$ et $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_q > 0$.

V, q, $\lambda_1, \lambda_2, \dots, \lambda_q$ répondent à la question et ceci achève la première question.

<u>Remarque</u> .. Dans la suite on pose $\lambda_{q+1} = \lambda_{q+2} = \dots = \lambda_p = 0$. Ainsi $\forall i \in [\![1, p]\!], \lambda_i = \lambda'_i$!!

(*) Supposons ${}^tJJ = 0$. $\forall x \in \mathbb{R}^p$, $\|Jx\|^2 = {}^tx{}^tJJx = 0$; $\forall x \in \mathbb{R}^p$, $Jx = 0$; $J = 0$

Q2 a) 1er cas. $q=p$. Alors toutes les valeurs propres de ${}^tJJ$ sont strictement positives.

${}^tJJ$ est inversible. Ainsi $rg\,{}^tJJ = p = q$.

2e cas... $q<p$. Le nombre d'éléments de B est p et celui de $B_1 \cup \dots \cup B_{r-1}$ est q. Alors card $B_r = p-q$ donc dim $SEP({}^tJJ, 0) = p-q$.

Alors $rg\,{}^tJJ = p - \dim SEP({}^tJJ,0) = p-(p-q) = q$.

Dans les deux cas $\underline{rg\,{}^tJJ = q}$.

b) • Soit $i \in [\![1,q]\!]$. ${}^tJJ\,V_i = \lambda_i V_i$.

Supposons $JV_i = 0$. Alors $\lambda_i V_i = 0$. V_i est alors nul car λ_i est non nul.

Ceci est impossible car la matrice V est inversible.

Ainsi $JV_i \neq 0$ et $(J\,{}^tJ)(JV_i) = J({}^tJJ\,V_i) = J(\lambda_i V_i) = \lambda_i JV_i$.

Donc λ_i est une valeur propre de $J\,{}^tJ$.

• Réciproquement soit λ une valeur propre non nulle de $J\,{}^tJ$ $(J\,{}^tJ \in \mathcal{M}_n(\mathbb{R}))$.

$\exists Z \in \mathcal{M}_{n,1}(\mathbb{R})$, $Z \neq 0_{\mathcal{M}_{n,1}(\mathbb{R})}$ et $J\,{}^tJZ = \lambda Z$.

Supposons ${}^tJZ = 0$. Alors $\lambda Z = 0$ donc $Z = 0$! car $\lambda \neq 0$. Ceci est impossible.

Donc ${}^tJZ \neq 0$ et $({}^tJJ)({}^tJZ) = {}^tJ(J\,{}^tJZ) = {}^tJ(\lambda Z) = \lambda\,{}^tJZ$.

Alors λ est une valeur propre de ${}^tJJ$.

Finalement $\underline{{}^tJJ \text{ et } J\,{}^tJ \text{ ont les mêmes valeurs propres non nulles.}}$

c) Soit $(z_1, z_2, \dots, z_r) \in \mathbb{R}^r$ tel que $\sum_{i=1}^r z_i JY_i = 0_{\mathcal{M}_{n,1}(\mathbb{R})}$.

Alors $0_{\mathcal{M}_{p,1}(\mathbb{R})} = {}^tJ\,0_{\mathcal{M}_{n,1}(\mathbb{R})} = {}^tJ\left(\sum_{i=1}^r z_i JY_i\right) = \sum_{i=1}^r z_i\,{}^tJJY_i = \sum_{i=1}^r z_i \lambda Y_i$.

Ainsi $\lambda \sum_{i=1}^r z_i Y_i = 0_{\mathcal{M}_{p,1}(\mathbb{R})}$ et $\lambda \neq 0$ donc $\sum_{i=1}^r z_i Y_i = 0_{\mathcal{M}_{p,1}(\mathbb{R})}$.

Or $(Y_1, Y_2, \dots, Y_r)$ est libre donc $z_1 = z_2 = \dots = z_r = 0$.

Ceci achève de montrer que $\underline{(JY_1, JY_2, \dots, JY_r) \text{ est libre.}}$

d] Soit λ une valeur propre non nulle de ${}^tJJ$ donc de $J{}^tJ$.

Soit $(Y_1, Y_2, \dots, Y_r)$ une base de $SEP({}^tJJ, \lambda)$.

Alors 1°/ $(JY_1, JY_2, \dots, JY_r)$ est une famille libre

2°/ $(JY_1, JY_2, \dots, JY_r)$ sont des éléments de $SEP(J{}^tJ, \lambda)$.

Ainsi $\dim SEP(J{}^tJ, \lambda) \geq r = \dim SEP({}^tJJ, \lambda)$.

On montre rigoureusement de la même manière que $\dim SEP({}^tJJ, \lambda) \geq \dim SEP(J{}^tJ, \lambda)$.

Finalement $\dim SEP({}^tJJ, \lambda) = \dim SEP(J{}^tJ, \lambda)$.

Les sous-espaces propres de ${}^tJJ$ et de $J{}^tJ$ associés à une même valeur propre non nulle sont de même dimension.

Rappelons que : $q = \operatorname{rg} {}^tJJ = \sum_{i=1}^{r'-1} \dim SEP({}^tJJ, \alpha_i)$ et

$\alpha_1, \alpha_2, \dots, \alpha_{r'-1}$ sont les valeurs propres non nulles de ${}^tJJ$ donc de $J{}^tJ$.

De même alors $\operatorname{rg} J{}^tJ = \sum_{i=1}^{r'-1} \dim SEP(J{}^tJ, \alpha_i) = \sum_{i=1}^{r'-1} \dim SEP({}^tJJ, \alpha_i) = \operatorname{rg} {}^tJJ = q$

Finalement $\operatorname{rg} J{}^tJ = \operatorname{rg} {}^tJJ = q$.

(Q3) a] Pour tout i dans $[\![1, q]\!]$, JV_i est un vecteur propre de $J{}^tJ$ associé à la valeur propre λ_i.

Donc pour tout i dans $[\![1, q]\!]$, $U_i = \frac{1}{\sqrt{\lambda_i}} JV_i$ est un vecteur propre de $J{}^tJ$ associé à la valeur propre λ_i

Soit $(i,j) \in [\![1, q]\!]^2$. $\langle U_i, U_j \rangle = \frac{1}{\sqrt{\lambda_i}} \frac{1}{\sqrt{\lambda_j}} \langle JV_i, JV_j \rangle = \frac{1}{\sqrt{\lambda_i}\sqrt{\lambda_j}} {}^t(JV_i) JV_j$.

$\langle U_i, U_j \rangle = \frac{1}{\sqrt{\lambda_i}\sqrt{\lambda_j}} {}^tV_i {}^tJJ V_j = \frac{1}{\sqrt{\lambda_i}\sqrt{\lambda_j}} {}^tV_i (\lambda_j V_j) = \frac{\lambda_j}{\sqrt{\lambda_i}\sqrt{\lambda_j}} \langle V_i, V_j \rangle$

Or $(V_1, V_2, \dots, V_q)$ est une famille orthonormée.

Donc si $i \neq j$: $\langle U_i, U_j \rangle = \frac{\lambda_i}{\sqrt{\lambda_i}\sqrt{\lambda_j}} \times 0 = 0$;

si $i = j$: $\langle U_i, U_j \rangle = \frac{\lambda_i}{\sqrt{\lambda_i}\sqrt{\lambda_i}} \times 1 = 1$.

ceci achève de montrer que $(U_1, U_2, \dots, U_q)$ est une famille orthonormée de vecteurs propres de J^tJ

b) 1er cas.. $q = p$. $(U_1, U_2, \dots, U_q)$ est une famille orthonormée de $M_{p,1}(\mathbb{R})$, de cardinal $q = p$ et $\dim M_{p,1}(\mathbb{R}) = p$. Donc $(U_1, U_2, \dots, U_p)$ est une base orthonormée de $M_{p,1}(\mathbb{R})$ ou de $\mathbb{R}^p$ constituée de vecteurs propres de J^tJ.

2e cas.. $q < p$. $rg\, J^tJ < p$. Ainsi 0 est valeur propre de J^tJ.

Rappelons que les sous-espaces propres de J^tJ sont deux à deux orthogonaux et que $(U_1, U_2, \dots, U_q)$ est une famille orthonormée de $M_{p,1}(\mathbb{R})$ ou de $\mathbb{R}^p$ constituée de vecteurs propres de J^tJ associés à des valeurs propres non nulles. Alors $(U_1, \dots, U_q)$ est une famille orthonormée de $(SEP(J^tJ, 0))^{\perp}$.

Or $\dim (SEP(J^tJ, 0))^{\perp} = n - \dim SEP(J^tJ, 0) = n - [n - rg(J^tJ)] = rg(J^tJ) = q$.

Alors $(U_1, U_2, \dots, U_q)$ est une famille orthonormée de $(SEP(J^tJ, 0))^{\perp}$ de cardinal q et $\dim (SEP(J^tJ, 0))^{\perp} = q$.

Ainsi $(U_1, U_2, \dots, U_q)$ est une base orthonormée de $(SEP(J^tJ, 0))^{\perp}$.

Soit alors $(U_{q+1}, \dots, U_n)$ une base orthonormée de $SEP(J^tJ, 0)$.

Par concaténation $(U_1, U_2, \dots, U_q, U_{q+1}, \dots, U_n)$ est une base orthonormée de $M_{n,1}(\mathbb{R})$ constituée de vecteurs propres de J^tJ.

<u>Dans les deux cas il existe une base orthonormée $(U_1, U_2, \dots, U_q, U_{q+1}, \dots, U_n)$ de $M_{n,1}(\mathbb{R})$ formée de vecteurs propres de J^tJ.</u>

Q4 Notons que 1) V est une matrice orthogonale de $M_n(\mathbb{R})$

2) On peut dire que : $\forall (i,j) \in [\![1,n]\!] \times [\![1,p]\!],\ \Delta_{ij} = \begin{cases} \sqrt{\lambda_i} & \text{si } i = j \\ 0 & \text{sinon} \end{cases}$

(en effet $\lambda_{q+1} = \lambda_{q+2} = \dots = \lambda_p = 0$).

Soient $(E_1,\dots,E_p)$ la base canonique de $M_{p,1}(\mathbb{R})$ et $(E'_1,\dots,E'_n)$ la base canonique de $M_{n,1}(\mathbb{R})$. Soit $(i,j)\in[\![1,n]\!]\times[\![1,p]\!]$.

SE_j est la j^e colonne de S donc ${}^tE'_i S E_j = s_{ij}$ (ok ??).

Posons $T={}^tUJV=(t_{ij})_{\substack{1\le i\le n\\ 1\le j\le p}}$. Rappelons que $p\le n$.

De même $t_{ij}={}^tE'_i T E_j$.

Alors $t_{ij}={}^tE'_i\,{}^tUJVE_j={}^t(UE'_i)JV_j={}^tU_iJV_j$.

Si $j\in[\![1,q]\!]$, $JV_j=\sqrt{\lambda_j}\,U_j$.

Supposons $j\notin[\![1,q]\!]$. $\|JV_j\|^2={}^t(JV_j)JV_j={}^tV_j\,{}^tJJV_j=0$

V_j est un vecteur propre de ${}^tJJ$ associé à la valeur propre 0.

$\|JV_j\|=0$.

$JV_j=0$. Or $\sqrt{\lambda_j}=0$ car $j\notin[\![1,q]\!]$.

On a donc encore $JV_j=\sqrt{\lambda_j}\,U_j$.

Ainsi $t_{ij}={}^tU_i\sqrt{\lambda_j}\,U_j=\sqrt{\lambda_j}\,\langle U_i,U_j\rangle=\begin{cases}\sqrt{\lambda_i} & \text{si } i=j\\ 0 & \text{si } i\neq j\end{cases}$ car

$(U_1,U_2,\dots,U_n)$ est une famille orthonormée.

Par conséquent $t_{ij}=s_{ij}$ et ceci pour tout $(i,j)\in[\![1,n]\!]\times[\![1,p]\!]$.

Donc $\underline{\underline{S={}^tUJV}}$. Alors $US\,{}^tV=U\,{}^tUJV\,{}^tV=J$ car U et V sont des matrices orthogonales. Par conséquent : $\underline{\underline{J=US\,{}^tV}}$

(Q5) a) ${}^tV\,{}^tJJV=\mathrm{Diag}(\lambda_1,\lambda_2,\dots,\lambda_p)$.

${}^tV({}^tJJ+\mu I)V={}^tV\,{}^tJJV+\mu I=\mathrm{Diag}(\lambda_1,\dots,\lambda_p)+\mu I$.

Alors $V^{-1}({}^tJJ+\mu I)V=\mathrm{Diag}(\lambda_1+\mu,\dots,\lambda_p+\mu)$.

${}^tJJ+\mu I$ est semblable à $\mathrm{Diag}(\lambda_1+\mu,\dots,\lambda_p+\mu)$.

Alors le spectre de ${}^tJJ+\mu I$ est $\{\lambda_1+\mu,\dots,\lambda_p+\mu\}$.

$\forall i \in [\![1,p]\!]$, $\lambda_i \geqslant 0$ et $\mu > 0$, donc $\forall i \in [\![1,p]\!]$, $\lambda_i + \mu > 0$.

On les a par valeurs propres de ${}^tJJ + \mu I$.

Ainsi $\underline{\underline{{}^tJJ + \mu I \text{ est inversible}}}$.

b) ${}^tV({}^tJJ + \mu I)V = \mathrm{Diag}(\lambda_1+\mu, \dots, \lambda_p+\mu)$.

Alors ${}^tJJ + \mu I = V\,\mathrm{Diag}(\lambda_1+\mu, \dots, \lambda_p+\mu)\,{}^tV = V\,\mathrm{Diag}(\lambda_1+\mu, \dots, \lambda_p+\mu)\,V^{-1}$

Notons que ${}^tJJ + \mu I$, V, $\mathrm{Diag}(\lambda_1+\mu, \dots, \lambda_p+\mu)$ et V^{-1} sont inversibles.

Alors $({}^tJJ + \mu I)^{-1} = (V^{-1})^{-1}(\mathrm{Diag}(\lambda_1+\mu, \dots, \lambda_p+\mu))^{-1}\,V^{-1}$

$({}^tJJ + \mu I)^{-1} = V\,\mathrm{Diag}\left(\frac{1}{\lambda_1+\mu}, \dots, \frac{1}{\lambda_p+\mu}\right){}^tV$

$({}^tJJ + \mu I)^{-1}\,{}^tJ = V\,\mathrm{Diag}\left(\frac{1}{\lambda_1+\mu}, \dots, \frac{1}{\lambda_p+\mu}\right){}^tV\,{}^tJ = V\,\mathrm{Diag}\left(\frac{1}{\lambda_1+\mu}, \dots, \frac{1}{\lambda_p+\mu}\right){}^t(JV)$

$J = US\,{}^tV$; $JV = US\,{}^tVV = US$; ${}^t(JV) = {}^tS\,{}^tU$

Ainsi $({}^tJJ + \mu I)^{-1}\,{}^tJ = V\,\mathrm{Diag}\left(\frac{1}{\lambda_1+\mu}, \dots, \frac{1}{\lambda_p+\mu}\right){}^tS\,{}^tU$

Montrons alors que $\mathrm{Diag}\left(\frac{1}{\lambda_1+\mu}, \dots, \frac{1}{\lambda_p+\mu}\right){}^tS = R$ et nous aurons le résultat.

Posons $\mathrm{Diag}\left(\frac{1}{\lambda_1+\mu}, \dots, \frac{1}{\lambda_p+\mu}\right) = (s_{ij})_{\substack{1\leq i\leq p\\ 1\leq j\leq p}}$ et $\mathrm{Diag}\left(\frac{1}{\lambda_1+\mu}, \dots, \frac{1}{\lambda_p+\mu}\right){}^tS = (w_{ij})_{\substack{1\leq i\leq p\\ 1\leq j\leq n}}$

$\forall (i,j) \in [\![1,p]\!] \times [\![1,n]\!]$, $w_{ij} = \sum_{k=1}^{p} s_{ik}\,\sigma_{jk} = s_{ii}\,\sigma_{ji} = \frac{1}{\lambda_i+\mu}\,\sigma_{ji} = \begin{cases} \frac{\sqrt{\lambda_i}}{\lambda_i+\mu} & \text{si } i=j \\ 0 & \text{sinon} \end{cases}$

donc $\forall (i,j) \in [\![1,p]\!] \times [\![1,n]\!]$, $w_{ij} = r_{ij}$.

Ceci achève de montrer que $\underline{\underline{({}^tJJ + \mu I)^{-1} \times {}^tJ = VR\,{}^tU}}$.

c)

(1) Posons $U = (u_{ij})_{\substack{1\leq i\leq n\\ 1\leq j\leq n}}$, $V = (v_{ij})_{\substack{1\leq i\leq p\\ 1\leq j\leq p}}$ et $Q = VR\,{}^tU = (q_{ij})_{\substack{1\leq i\leq p\\ 1\leq j\leq n}}$.

Soit $(i,j) \in [\![1,p]\!] \times [\![1,n]\!]$.

$$q_{ij} \underset{\substack{\uparrow \\ \text{calcul simple}}}{=} \sum_{k=1}^{p} v_{ik} \sum_{\ell=1}^{n} r_{k\ell}\,u_{j\ell} = \sum_{k=1}^{p} v_{ik}\,\frac{\sqrt{\lambda_k}}{\lambda_k+\mu}\,u_{jk}.$$

$$q_{ij} = \sum_{k=1}^{p} \frac{\sqrt{\lambda_k}}{\lambda_k + \mu} v_{ik} u_{jk}.$$

Posons $Q' = \sum_{\ell=1}^{q} \frac{\sqrt{\lambda_\ell}}{\lambda_\ell+\mu} V_\ell\, {}^tU_\ell = (q'_{ij})_{\substack{1\le i\le p\\ 1\le j\le n}}$.

On a encore $Q' = \sum_{\ell=1}^{p} \frac{\sqrt{\lambda_\ell}}{\lambda_\ell+\mu} V_\ell\, {}^tU_\ell$ car $\forall \ell \in [\![q+1, p]\!], \lambda_\ell = 0$.

Soit $k \in [\![1,p]\!]$.

$$V_\ell\, {}^tU_\ell = \begin{pmatrix} v_{1\ell} \\ v_{2\ell} \\ \vdots \\ v_{p\ell} \end{pmatrix} (u_{1\ell}\ u_{2\ell} \cdots u_{n\ell}) = (v_{i\ell} u_{j\ell})_{\substack{1\le i\le p\\ 1\le j\le n}}.$$

Alors $q'_{ij} = \sum_{\ell=1}^{p} \frac{\sqrt{\lambda_\ell}}{\lambda_\ell+\mu} v_{i\ell} u_{j\ell} = q_{ij}$! Ainsi $Q' = Q$.

Alors $VR\,{}^tU = \sum_{k=1}^{q} \frac{\sqrt{\lambda_k}}{\lambda_k+\mu} V_k\, {}^tU_k = \sum_{i=1}^{q} \frac{\sqrt{\lambda_i}}{\lambda_i+\mu} V_i\, {}^tU_i$.

donc $\underline{({}^tJJ + \mu I)^{-1}\, {}^tJ = \sum_{i=1}^{q} \frac{\sqrt{\lambda_i}}{\lambda_i+\mu} V_i\, {}^tU_i}$.

(V2) Reprenons les bases canoniques $(E_1, \dots, E_p)$ et $(E'_1, \dots, E'_n)$ de $M_{p,1}(\mathbb{R})$ et de $M_{n,1}(\mathbb{R})$. Soit $(i,j) \in [\![1,p]\!] \times [\![1,n]\!]$.

$E_i\, {}^tE'_j \in M_{p,n}(\mathbb{R})$ et un calcul simple montre que cette matrice a tous ses coefficients nuls sauf celui situé à l'intersection de la i^{e} ligne et de la j^{e} colonne qui vaut 1. Ainsi $(E_i\, {}^tE'_j)_{\substack{1\le i\le p\\ 1\le j\le n}}$ est la base canonique de $M_{p,n}(\mathbb{R})$.

donc $R = \sum_{i=1}^{p}\sum_{j=1}^{n} r_{ij} E_i\, {}^tE'_j = \sum_{i=1}^{q} r_{ii} E_i\, {}^tE'_i$.

donc $VR\,{}^tU = V\left(\sum_{i=1}^{q} r_{ii} E_i\, {}^tE'_i\right){}^tU = \sum_{i=1}^{q} r_{ii} VE_i\, {}^tE'_i\, {}^tU = \sum_{i=1}^{q} r_{ii} VE_i\, {}^t(UE'_i)$

or $\forall i \in [\![1,n]\!]$, $VE_i = V_i$ et $UE'_i = U_i$.

donc $VR\,{}^tU = \sum_{i=1}^{q} r_{ii} V_i\, {}^tU_i = \sum_{i=1}^{q} \frac{\sqrt{\lambda_i}}{\lambda_i+\mu} V_i\, {}^tU_i$. On retrouve ainsi le résultat.

Q6 a) $\forall h \in \mathbb{R}^p - \{0\}$, $\dfrac{F(x+h)-\Pi(h)}{\|h\|} = \dfrac{F(x+h)-L(h)-\frac{\mu}{2}\|h\|^2}{\|h\|} = \dfrac{F(x+h)-L(h)}{\|h\|} - \frac{\mu}{2}\|h\|$.

Or $\lim\limits_{h\to 0} \dfrac{F(x+h)-L(h)}{\|h\|} = 0$ (d'après II Q3 b)) et $\lim\limits_{h\to 0} \left(\frac{\mu}{2}\|h\|\right) = 0$

Ainsi $\lim\limits_{h\to 0} \dfrac{F(x+h)-\Pi(h)}{\|h\|} = 0$.

b) Posons $\forall h \in \mathbb{R}^p$, $N(h) = \|h\|^2$.

$\forall h \in \mathbb{R}^p$, $N(h) = \sum\limits_{i=1}^{p} h_i^2$. Ainsi N est de classe $\mathcal{C}^2$ sur $\mathbb{R}^p$ comme fonction polynôme.

$\forall i \in [\![1,p]\!]$, $\forall h = (h_1,\dots,h_p) \in \mathbb{R}^p$, $\dfrac{\partial N}{\partial h_i}(h) = 2h_i$. $\underline{\nabla N(h) = 2h}$ pour tout $h \in \mathbb{R}^p$.

$\forall (i,j) \in [\![1,p]\!]^2$, $\forall h \in (h_1,\dots,h_p) \in \mathbb{R}^p$, $\dfrac{\partial^2 N}{\partial h_j \partial h_i}(h) = \begin{cases} 2 \text{ si } i = j \\ 0 \text{ sinon} \end{cases}$.

Ainsi $\forall h \in \mathbb{R}^p$, $\nabla^2 N(h) = 2 I_p$.

L et N étant de classe $\mathcal{C}^2$ sur $\mathbb{R}^p$, Π est de classe $\mathcal{C}^2$ sur $\mathbb{R}^p$.

De plus $\nabla \Pi = \nabla L + \frac{\mu}{2} \nabla N$ et $\nabla^2 \Pi = \nabla^2 L + \frac{\mu}{2} \nabla^2 N$.

$\forall h \in \mathbb{R}^p$, $\nabla \Pi(h) = \nabla L(h) + \frac{\mu}{2} h = \nabla F(x) + G(x) h + \frac{\mu}{2} 2h$.

$\underline{\underline{\forall h \in \mathbb{R}^p, \ \nabla \Pi(h) = \nabla F(x) + (G(x) + \mu I) h.}}$

$\forall h \in \mathbb{R}^p$, $\nabla^2 \Pi(h) = \nabla^2 L + \frac{\mu}{2} \nabla^2 N = G(x) + \frac{\mu}{2} \times 2I = G(x) + \mu I$.

$\underline{\underline{\forall h \in \mathbb{R}^p, \ \nabla^2 \Pi(h) = G(x) + \mu I.}}$

c) Rappelons que $\nabla F(x) = {}^tJ(x) f(x)$ et que par hypothèse $\nabla F(x) \neq 0$.

Ainsi $J(x)$ n'est pas la matrice nulle. On peut donc appliquer ce qui précède à la matrice ${}^tJ(x)J(x) + \mu I$ donc à $G(x) + \mu I$.

$G(x) + \mu I$ est donc inversible.

Soit $\ell \in \mathbb{R}^p$. $\nabla\Pi(\ell)=0 \Leftrightarrow \nabla F(x) + (G(x)+\mu I)\ell = 0 \Leftrightarrow (G(x)+\mu I)\ell = -\nabla F(x)$

$\nabla\Pi(\ell)=0 \Leftrightarrow \ell = -(G(x)+\mu I)^{-1}\nabla F(x)$.

<u>Π admet un point critique et un seul : $\ell^* = -(G(x)+\mu I)^{-1}\nabla F(x)$.</u>

$$\ell^* = -({}^tJ(x)J(x)+\mu I)^{-1}\,{}^tJ(x)f(x) = -\sum_{i=1}^{q}\frac{\sqrt{\lambda_i(x)}}{\lambda_i(x)+\mu}V_i(x)\,{}^tU_i(x)\,f(x)$$ ▲

$$\ell^* = -\sum_{i=1}^{q}\frac{\sqrt{\lambda_i(x)}}{\lambda_i(x)+\mu}\langle U_i(x), f(x)\rangle\, V_i(x).$$

d) $\nabla^2\Pi(\ell) = G(x)+\mu I$ et les valeurs propres de $G(x)+\mu I$ sont strictement positives.

<u>Ainsi Π admet un minimum local en ℓ.</u>

▲ $\lambda_i(x)$, $U_i(x)$ et $V_i(x)$ sont définies à partir de ${}^tJ(x)J(x)$ comme λ_i, U_i et V_i dans Q1 et Q3 à partir de ${}^tJJ$.