CONCOURS HEC ESSEC 2025
MATHÉMATIQUES APPROFONDIES 1

Pierre-Paul TACHER

Partie I. Autour de l'adjoint

1. Si on pose

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{p-1} \\ x_p \end{bmatrix}$$

les coordonnées de x dans la base orthogonale $\mathscr{B}_E$,

$$\begin{aligned} l(x) &= l(\sum_{i=1}^{p} x_i e_i) \\ &= \sum_{i=1}^{p} x_i l(e_i) \\ &= \langle a_0, x \rangle \end{aligned}$$

en posant

$$a_0 = \begin{bmatrix} l(e_1) \\ l(e_2) \\ \vdots \\ l(e_{p-1}) \\ l(e_p) \end{bmatrix}$$

ce qui démontre l'existence de a_0 tel que

$$\begin{aligned} l : E &\to \mathbb{R} \\ x &\mapsto \langle a_0, x \rangle \end{aligned}$$

Pour l'unicité, si on suppose

$$\forall x \in E, \quad \langle a_0, x \rangle = \langle u, x \rangle$$

il suffit d'appliquer pour i $\in [\![1,p]\!]$ à $x = e_i$ les vecteurs de la base orthogonale pour voir que les coordonnées de u et a_0 sont identiques et donc $u = a_0$.

2. Si on fixe $y \in F$, l'application

$$\begin{aligned} l_y : E &\to \mathbb{R} \\ x &\mapsto \langle u(x), y\rangle \end{aligned}$$

est linéaire.

D'après ce qu'on vient de voir, il existe donc un unique $z_y \in E$ (dépendant de y) tel que

$$\forall x \in E, \quad \begin{aligned} l_y(x) &= \langle u(x), y\rangle \\ &= \langle z_y, x\rangle \end{aligned}$$

3. Soit $\lambda \in \mathbb{R}$, $(y_1, y_2) \in F^2$.

$$\forall x \in E, \quad \begin{aligned} \langle z_{\lambda y_1 + y_2}, x\rangle &= \langle u(x), \lambda y_1 + y_2\rangle \\ &= \lambda\langle u(x), y_1\rangle + \langle u(x), y_2\rangle \\ &= \lambda\langle z_{y_1}, x\rangle + \langle z_{y_2}, x\rangle \\ &= \langle \lambda z_{y_1} + z_{y_2}, x\rangle \end{aligned}$$

Il suffit d'appliquer cette égalité aux vecteurs $x = e_i$ de la base orthogonale de E pour arriver à la conclusion

$$z_{\lambda y_1 + y_2} = \lambda z_{y_1} + z_{y_2}$$

i.e. l'application

$$\begin{aligned} u^* : F &\to E \\ y &\mapsto z_y \end{aligned}$$

est linéaire.

4. Notons

$$\tilde{A} = \mathrm{Mat}_{\mathscr{B}_E, \mathscr{B}_F}(u^*)$$

$$\forall (i,j) \in [\![1,n]\!] \times [\![1,p]\!], \quad \begin{aligned} \tilde{a}_{ij} &= \langle e_i, u^*(f_j)\rangle \\ &= \langle u(e_i), f_j\rangle \\ &= A_{ji} \\ &= {}^tA_{ij} \end{aligned}$$

donc on a bien $\mathrm{Mat}_{\mathscr{B}_E, \mathscr{B}_F}(u^*) = {}^tA$.

Il s'ensuit que

$$\begin{aligned} \mathrm{rang}(u^*) &= \mathrm{rang}({}^tA) \\ &= \mathrm{rang}(A) \\ &= \mathrm{rang}(u) \end{aligned}$$

puis

$$\begin{aligned}\mathrm{Mat}_{\mathscr{B}_E,\mathscr{B}_F}((u^*)^*) &= {}^t\mathrm{Mat}_{\mathscr{B}_E,\mathscr{B}_F}(u^*)\\ &= {}^t({}^tA)\\ &= A\\ &= \mathrm{Mat}_{\mathscr{B}_E,\mathscr{B}_F}(u)\end{aligned}$$

ainsi $(u^*)^* = u$.

5. Soit $x \in \ker u$ et $y' = u^*(y) \in \operatorname{Im} u^*$.

$$\begin{aligned}\langle x, y'\rangle &= \langle x, u^*(y)\rangle\\ &= \langle u(x), y\rangle\\ &= \langle 0, y\rangle\\ &= 0\end{aligned}$$

ce qui montre $\operatorname{Im} u^* \subset \ker u^\perp$.

Puis d'après le théorème du rang,

$$\begin{aligned}\dim \operatorname{Im} u^* &= \operatorname{rang}(u^*)\\ &= \operatorname{rang}(u)\\ &= \dim E - \dim \ker u\\ &= \dim \ker u^\perp\end{aligned}$$

On peut ainsi conclure $\operatorname{Im} u^* = \ker u^\perp$.

6. Soit $x \in E$.

$$\begin{aligned}\|u(x)\|^2 &= \langle u(x), u(x)\rangle\\ &= \langle x, u^* \circ u(x)\rangle\end{aligned}$$

cela montre $\ker u^* \circ u \subset \ker u$. L'inclusion inverse étant triviale, on a

$$\ker u^* \circ u = \ker u$$

On sait que $\operatorname{Im} u^* \circ u \subset \operatorname{Im} u^*$, et

$$\begin{aligned}\dim \operatorname{Im} u^* \circ u &= \dim E - \dim \ker u^* \circ u\\ &= \dim E - \dim \ker u\\ &= \dim \ker u^\perp\\ &= \dim \operatorname{Im} u^*\end{aligned}$$

donc finalement $\operatorname{Im} u^* \circ u = \operatorname{Im} u$.

7.

$$\begin{aligned}\ker w &= \ker u^* \circ u \cap \operatorname{Im} u^*\\ &= \ker u \cap \operatorname{Im} u^* && \text{(cf. question 6)}\\ &= (\operatorname{Im} u^*)^\perp \cap \operatorname{Im} u^* && \text{(cf. question 5)}\\ &= \{0\}\end{aligned}$$

et comme on est en dimension finie cela équivaut à w est un isomorphisme de $\operatorname{Im} u^*$.

Si on choisit une base adaptée à la décomposition $E = \ker u \oplus \operatorname{Im} u^*$, la matrice de $u^* \circ u$ est diagonale par blocs:

$$\widetilde{M} = \begin{array}{cc} \left[\begin{array}{cc} \boldsymbol{W} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{array}\right] & \begin{array}{l} \} \text{ taille } \operatorname{rang}(u^*) \\ \} \text{ taille } p - \operatorname{rang}(u^*) \end{array} \\ \underbrace{\quad\quad}_{\operatorname{rang}(u^*)} \underbrace{\quad\quad}_{p - \operatorname{rang}(u^*)} & \end{array}$$

avec la matrice $W \in \mathcal{M}_{\operatorname{rang}(u^*)}(\mathbb{R})$ inversible. La matrice $M = {}^tAA$ est donc semblable à cette matrice, i.e. il existe un matrice orthogonale P telle que

$$M = P\widetilde{M}\,{}^tP \tag{1}$$

8.

8a. D'après la question 5 en inversant le rôle de u et u^*, on a

$$F = \operatorname{Im} u \oplus \ker u^* \tag{2}$$

par unicité du supplémentaire orthogonal, on a

$$\ker u^* = \ker \pi$$

Soit $X \in \operatorname{Im} u$. $QX = X$, donc

$${}^tAQX = {}^tAX$$

Soit $X \in \ker u^*$. $QX = 0$, donc

$$\begin{aligned} {}^tAQX &= 0 \\ &= {}^tAX \end{aligned}$$

En vue de l'égalité 2, on peut conclure

$${}^tAQ = {}^tA$$

8b.

$$\begin{aligned} \operatorname{Tr} Q &= \operatorname{rang}(\pi) && \text{(}pi\text{ est un projecteur)} \\ &= \dim \operatorname{Im} u \\ &= \operatorname{rang}(u) \end{aligned}$$

9. D'après l'équation 1, M est inversible si et seulement si $\widetilde{M}$ est inversible. Cela équivaut à

$$\begin{aligned} & \dim \operatorname{Im} u^* = p \\ \Leftrightarrow \ & \operatorname{rang}(u^*) = p \\ \Leftrightarrow \ & \operatorname{rang}(u) = p \qquad \text{(cf. question 4)} \\ \Leftrightarrow \ & \operatorname{rang}(A) = p \end{aligned}$$

10.

10a. $A(M)^{-1}\,{}^tA$ est la matrice d'un projecteur orthogonal $\tilde{\pi}$ car

$$\begin{aligned} (A(M)^{-1}\,{}^tA)^2 &= A(M)^{-1}\,\underbrace{{}^tAA}_{=M}(M)^{-1}\,{}^tA \\ &= A(M)^{-1}\,{}^tA \end{aligned}$$

et

$$\begin{aligned} {}^t(A(M)^{-1}\,{}^tA) &= A\,{}^t(M)^{-1}\,{}^tA \\ &= A({}^tM)^{-1}\,{}^tA \\ &= A(M)^{-1}\,{}^tA \end{aligned}$$

Or

$$\tilde{\pi} = u \circ (u^* \circ u)^{-1} \circ u^*$$

donc $\operatorname{Im} \tilde{\pi} \subset \operatorname{Im} u$. et

$$\begin{aligned} \dim \operatorname{Im} \tilde{\pi} &= \operatorname{Tr}(A(M)^{-1}\,{}^tA) \\ &= \operatorname{Tr}(\underbrace{{}^tAA}_{=M}(M)^{-1}) \\ &= \operatorname{Tr}(I_p) \\ &= p \\ &= \operatorname{rang}(u) \\ &= \dim \operatorname{Im} u \end{aligned}$$

donc finalement $\tilde{\pi} = \pi$ i.e.

$$Q = A(M)^{-1}\,{}^tA$$

10b.

```python
import numpy as np
import matplotlib.pyplot as plt
import numpy.random as rd
import numpy.linalg as al

def Calcule_Q(a):
  if np.shape(a)[1] != al.matrix_rank(a):
    print("La matrice A n'est pas de rang p.")
    return
```

```
m = np.transpose(a) @ a
minv = al.inv(m)

return a @ minv @ np.transpose(a)
```

11.

$$\begin{aligned}{}^tXMX &= {}^tX\,{}^tAAX \\ &= {}^t(AX)AX \\ &= \langle AX, AX\rangle \\ &= \|AX\|^2 \geqslant 0\end{aligned}$$

Partie II. Minimisation d'une fonction quadratique

12.

$$\begin{aligned}2(J_0(X+H)-J_0(X)) &= \|A(X+H)-Y\|^2 - \|AX-Y\|^2 \\ &= \langle A(X+H)-Y, A(X+H)-Y\rangle - \|AX-Y\|^2 \\ &= \langle AX-Y+AH, AX-Y+AH\rangle - \|AX-Y\|^2 \\ &= 2\langle AH, AX-Y\rangle + \langle AH, AH\rangle \\ &= 2\langle H, {}^tA(AX-Y)\rangle + \langle AH, AH\rangle \\ &= 2\langle H, {}^tAAX - {}^tAY\rangle + {}^tH\,{}^tAAH \\ 2(J_0(X+H)-J_0(X)) &= 2\langle H, D(X)\rangle + {}^tHMH \qquad (3)\end{aligned}$$

Voir l'encadré plus bas pour faire le lien avec le gradient de J_0.

13. Si $D(X) = 0$ alors

$$\forall H \in \mathcal{M}_{p,1}(\mathbb{R}), \quad J_0(X+H) - J_0(X) = \frac{1}{2}\,{}^tHMH \geqslant 0 \qquad \text{(cf. question 11)}$$

ce qui montre que X est un minimum global.

Réciproquement, supposons $D(X) \neq 0$ et appliquons l'équation 3 à $H_\lambda = -\lambda D(X), \quad \lambda > 0$. On obtient

$$\begin{aligned}J_0(X+H_\lambda) - J_0(X) &= -\lambda \underbrace{\langle D(X), D(X)\rangle}_{>0} + \frac{1}{2}\lambda^2\,{}^tD(X)MD(X) \\ &\underset{\lambda\to 0}{=} -\lambda\langle D(X), D(X)\rangle + \mathrm{o}(\lambda)\end{aligned}$$

Cette expression est du signe de $-\lambda\langle D(X), D(X)\rangle$ au voisinage de 0, c'est-à-dire si on choisit $\lambda > 0$ suffisamment petit, mais alors

$$J_0(X+H_\lambda) < J_0(X)$$

contredisant le fait que $J_0(X)$ est un minimum global.

14. Notons $y \in F$ le vecteur dont la représentation matricielle est Y. D'après la question 7, $u^* \circ u$ est un isomorphisme de $\operatorname{Im} u^*$, il existe donc un unique $x_0 \in \operatorname{Im} u^* = \ker u^\perp$ tel que $u^* \circ u(x_0) = u^*(y)$. Ce qui s'exprime matriciellement par exactement ce que l'on veut:

$$\exists X_0 \in \mathbb{R}^p, \quad \mathbb{S}_0 \cap \ker A^\perp = \{X_0\}$$

15.

On doit faire ici le lien avec le gradient de la fonction à optimiser:

$$\begin{aligned}
J_0(X) &= \frac{1}{2}\|AX - Y\|^2 \\
&= \frac{1}{2}\langle AX - Y, AX - Y\rangle \\
&= \frac{1}{2}\,{}^tAX - Y(AX - Y) \\
&= \frac{1}{2}\,{}^tX\,{}^tAAX - {}^tYAX + \frac{1}{2}\,{}^tYY
\end{aligned}$$

Si on note, pour $f : \mathbb{R}^p \to \mathbb{R}$,

$$\begin{aligned}
\nabla_X f &= \frac{\partial f}{\partial X} \\
&= \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \vdots \\ \frac{\partial f}{\partial x_{p-1}} \\ \frac{\partial f}{\partial x_p} \end{bmatrix}
\end{aligned}$$

Il est utile de connaitre les formules

$$\frac{\partial\,{}^tXMX}{\partial X} = 2MX$$

si M est une matrice symétrique, et

$$\frac{\partial BX}{\partial X} = {}^tB$$

(Faire l'analogie avec la dérivation par rapport à une variable réelle.)
Ainsi

$$\begin{aligned}
\nabla_X J_0(X) &= \frac{\partial J_0}{\partial X}(X) \\
&= {}^tAAX - {}^tAY \\
&= D(X)
\end{aligned}$$

15a. D'après la question 5, $\mathcal{M}_{n,1}(\mathbb{R}) = \ker {}^tA \oplus \operatorname{Im} A$. ou encore

$$F = \ker u^* \oplus \operatorname{Im} u$$

On sait alors que $u^*_{|\operatorname{Im} u}$ est un isomorphisme de $\operatorname{Im} u$ vers $\operatorname{Im} u^*$.

$$\begin{aligned}
&D(X) = 0 \\
\Leftrightarrow\, &{}^tAAX = {}^tAY
\end{aligned}$$

$$\begin{aligned}
&\Leftrightarrow {}^tAAX = {}^tAQY && \text{(cf. question 8a)}\\
&\Leftrightarrow \quad AX = QY && \text{(au vu de la remarque précédente car } AX, QY \in \operatorname{Im} A)
\end{aligned}$$

15b. D'après 15a,

$$\begin{aligned}
A(X - X_0) &= AX - AX_0\\
&= QY - QY\\
&= 0
\end{aligned}$$

i.e. $X - X_0 \in \ker A$.

On a ainsi

$$X = \underbrace{X_0}_{\in \ker A^\perp} + \underbrace{X - X_0}_{\in \ker A}$$

Le théorème de Pythagore nous donne

$$\|X\|^2 = \|X_0\|^2 + \underbrace{\|X - X_0\|^2}_{>0} > \|X_0\|^2$$

si $X \neq X_0$.

15c. Si $\operatorname{rang}(A) = p$, d'après 9 M est inversible,

$$\begin{aligned}
&X \in \mathbb{S}_0\\
\Leftrightarrow &D(X) = 0\\
\Leftrightarrow &\ MX = {}^tAY\\
\Leftrightarrow &\quad X = M^{-1}\,{}^tAY
\end{aligned}$$

16.

16a.

$$\begin{aligned}
T &= \|QZ\|^2\\
&= {}^t(QZ)QZ\\
&= {}^tZ\,{}^tQQZ\\
&= {}^tZQQZ && \text{(Q est un projecteur orthogonal)}\\
&= {}^tZQZ && \text{(Q est un projecteur)}
\end{aligned}$$

16b.

```
def simuleT(a,sigma):
  n = np.shape(a)[0]
  z = np.random.normal(0, sigma, size=(n, 1))

  return np.transpose(z) @ Calcule_Q(a) @ z
```

16c.

```
def esperance(a,sigma):
  sum=0
  for i in range(1000):
    sum += simuleT(a,sigma)
  return sum/1000
```

16d. Pour $\sigma = 1$, on peut conjecturer que $\mathbb{E}(T) = \text{rang}(A) = p$.

16e. Le prgramme représente $\mathbb{E}(T)$ en fonction de σ, pour $\sigma \in [\![1, 6]\!]$. On peut penser que les points sont sur une parabole, et comme de plus comme pour $\sigma = 1$, on peut conjecturer que $\mathbb{E}(T) = \text{rang}(A) = p$ et pour $\sigma = 0$, $\mathbb{E}(T) = 0$, une conjecture raisonnable est:

$$\mathbb{E}(T) = p\sigma^2$$

16f. Déjà

$$\begin{aligned}\mathbb{E}[z_1^2] &= \sigma_{z_1}^2 + (\mathbb{E}[z_1])^2 \\ &= \sigma^2 + (\mathbb{E}[z_1])^2 \\ &= \sigma^2\end{aligned}$$

Puis,

$$\begin{aligned}\mathbb{E}[z_1^3] &= \int_0^{+\infty} x^3 e^{-\frac{x^2}{2\sigma^2}}\,\mathrm{d}x \\ &= \left[-\sigma^2 x^2 e^{-\frac{x^2}{\sigma^2}}\right]_0^{+\infty} + 2\sigma^2 \int_0^{+\infty} x e^{-\frac{x^2}{\sigma^2}}\,\mathrm{d}x\end{aligned}$$

Par puissances comparées

$$\lim_{x \to +\infty} x^2 e^{-\frac{x^2}{\sigma^2}} = 0$$

donc

$$\begin{aligned}\mathbb{E}[z_1^3] &= 2\sigma^2 \int_0^{+\infty} x e^{-\frac{x^2}{\sigma^2}}\,\mathrm{d}x \\ &= 2\sigma^2 \mathbb{E}[z_1] \\ &= 0\end{aligned}$$

De la même manière avec une intégration par parties,

$$\begin{aligned}\mathbb{E}[z_1^4] &= \int_0^{+\infty} x^4 e^{-\frac{x^2}{\sigma^2}}\,\mathrm{d}x \\ &= \left[-\sigma^2 x^3 e^{-\frac{x^3}{\sigma^2}}\right]_0^{+\infty} + 3\sigma^2 \int_0^{+\infty} x^2 e^{-\frac{x^2}{\sigma^2}}\,\mathrm{d}x \\ &= 3\sigma^2 \int_0^{+\infty} x^2 e^{-\frac{x^2}{\sigma^2}}\,\mathrm{d}x \\ &= 3\sigma^2 \mathbb{E}[z_1^2] \\ &= 3\sigma^4\end{aligned}$$

16g.

$$\begin{aligned}
T &= {}^tZQZ \\
&= \sum_{(i,j)\in[\![1,n]\!]^2} z_i Q_{i,j} z_j \\
&= \sum_{i=1}^n Q_{i,i} z_i^2 + \sum_{\substack{(i,j)\in[\![1,n]\!]^2 \\ i<j}} z_i Q_{i,j} z_j + \sum_{\substack{(i,j)\in[\![1,n]\!]^2 \\ j<i}} z_j \underbrace{Q_{i,j}}_{=Q_{j,i}} z_i \\
&= \sum_{i=1}^n Q_{i,i} z_i^2 + 2 \sum_{\substack{(i,j)\in[\![1,n]\!]^2 \\ i<j}} z_i Q_{i,j} z_j \\
T &= \sum_{i=1}^n Q_{i,i} z_i^2 + 2 \sum_{i=1}^n \sum_{j=i+1}^n z_i Q_{i,j} z_j
\end{aligned}$$

En utilisant la linéarité de l'espérance et l'indépendance des z_i, on obtient

$$\begin{aligned}
\mathbb{E}[T] &= \sum_{i=1}^n Q_{i,i} \mathbb{E}[z_i^2] + 2 \sum_{i=1}^n \sum_{j=i+1}^n Q_{i,j} \mathbb{E}[z_i z_j] \\
&= \sum_{i=1}^n Q_{i,i} \sigma^2 + 2 \sum_{i=1}^n \sum_{j=i+1}^n Q_{i,j} \mathbb{E}[z_i]\mathbb{E}[z_j] \\
&= \sum_{i=1}^n Q_{i,i} \sigma^2 \\
&= \operatorname{Tr} Q \sigma^2 \\
&= \dim \operatorname{Im} A \sigma^2 \\
&= \operatorname{rang}(A) \sigma^2 \\
&= p\sigma^2
\end{aligned}$$

On vérifie la conjecture faite en 16e.

16h.

$$\begin{aligned}
\mathbb{E}[T_1 T_2] &= \mathbb{E}[\sum_{k=1}^n Q_{k,k} \sum_{i=1}^n \sum_{j=i+1}^n Q_{i,j} z_k z_i z_j] \\
&= \sum_{k=1}^n Q_{k,k} \sum_{i=1}^n \sum_{j=i+1}^n Q_{i,j} \mathbb{E}[z_k z_i z_j]
\end{aligned}$$

Si $k \neq i$ et $k \neq j$,

$$\begin{aligned}
\mathbb{E}[z_k z_i z_j] &= \mathbb{E}[z_k]\mathbb{E}[z_i]\mathbb{E}[z_j] \\
&= 0
\end{aligned}$$

Si $k = i$,

$$\begin{aligned}
\mathbb{E}[z_k z_i z_j] &= \mathbb{E}[z_i^2]\mathbb{E}[z_j] \\
&= 0
\end{aligned}$$

car z_i et z_j étant indépendante, z_i^2 et z_j le sont aussi. Le cas $k = j$ est similaire.

Au final,

$$\mathbb{E}[T_1T_2] = 0$$

16i.

$$\begin{aligned}
\mathbb{E}[T_1^2] &= \mathbb{E}[\sum_{k=1}^{n} Q_{k,k} \sum_{i=1}^{n} Q_{i,i} z_i^2 z_k^2] \\
&= \sum_{k=1}^{n} Q_{k,k} \sum_{i=1}^{n} Q_{i,i} \mathbb{E}[z_i^2 z_k^2] \\
&= \sum_{k=1}^{n} Q_{k,k}^2 \mathbb{E}[z_k^4] + \sum_{\substack{(i,k)\in[\![1,n]\!]^2 \\ i\neq k}} Q_{k,k}Q_{i,i}\mathbb{E}[z_k^2 z_i^2] \\
&= 3\sum_{k=1}^{n} Q_{k,k}^2 \sigma^4 + \sum_{\substack{(i,k)\in[\![1,n]\!]^2 \\ i\neq k}} Q_{k,k}Q_{i,i}\mathbb{E}[z_k^2]\mathbb{E}[z_i^2] \\
&= 3\sum_{k=1}^{n} Q_{k,k}^2 \sigma^4 + \sum_{\substack{(i,k)\in[\![1,n]\!]^2 \\ i\neq k}} Q_{k,k}Q_{i,i}\sigma^4 \\
&= \left(2\sum_{k=1}^{n} Q_{k,k}^2 + \sum_{k=1}^{n} Q_{k,k}^2 + \sum_{\substack{(i,k)\in[\![1,n]\!]^2 \\ i\neq k}} Q_{k,k}Q_{i,i} \right) \sigma^4 \\
&= \left(2q + \sum_{(i,k)\in[\![1,n]\!]^2} Q_{k,k}Q_{i,i} \right) \sigma^4 \\
&= \left(2q + (\sum_{i=1}^{n} Q_{i,i})^2 \right) \sigma^4 \\
&= \left(2q + (\mathrm{rang}(A))^2 \right) \sigma^4 \\
\mathbb{E}[T_1^2] &= (2q + p^2)\sigma^4
\end{aligned}$$

$$T_2^2 = \frac{1}{4} \left(\sum_{\substack{(i,j)\in[\![1,n]\!]^2 \\ i\neq j}} Q_{i,j} z_i z_j \right) \left(\sum_{\substack{(k,l)\in[\![1,n]\!]^2 \\ k\neq l}} Q_{k,l} z_k z_l \right)$$

On remarque que:

$$\mathbb{E}[z_i z_j z_k z_l] = \begin{cases} \sigma^4 & \text{si } (i,j) = (k,l) \text{ ou } (i,j) = (l,k) \\ 0 & \text{sinon} \end{cases}$$

d'où

$$\begin{aligned}
\mathbb{E}[T_2^2] &= \frac{1}{4}\left(2\sum_{\substack{(i,j)\in[\![1,n]\!]^2\\ i\neq j}} Q_{i,j}^2\right)\sigma^4\\
&= \frac{1}{2}\left(\sum_{(i,j)\in[\![1,n]\!]^2} Q_{i,j}Q_{j,i} - \sum_{k=1}^{n} Q_{k,k}^2\right)\sigma^4\\
&= \frac{1}{2}\left(\sum_{i=1}^{n}\underbrace{(\sum_{j=1}^{n} Q_{i,j}Q_{j,i})}_{=Q_{i,i}^2=Q_{i,i}} - q\right)\sigma^4\\
&= \frac{1}{2}\left(\sum_{i=1}^{n} Q_{i,i} - q\right)\sigma^4\\
&= \frac{1}{2}(\operatorname{Tr} Q - q)\sigma^4\\
&= \frac{1}{2}(p-q)\sigma^4
\end{aligned}$$

16j.

$$\begin{aligned}
\mathbb{V}[T] &= \mathbb{E}[T^2] - (\mathbb{E}[T])^2\\
&= \mathbb{E}[T_1^2] + 4\underbrace{\mathbb{E}[T_1T_2]}_{=0} + 4\mathbb{E}[T_2^2] - (\mathbb{E}[T])^2\\
&= (2q + p^2 + 2p - 2q - p^2)\sigma^4\\
&= 2p\sigma^4
\end{aligned}$$

L'inégalité de Bienaymé-Tchebychev nous donne alors:

$$\begin{aligned}
P(T > \underbrace{p\sigma^2}_{=\mathbb{E}[T]} + \alpha p\sigma^2) &\leqslant P(|T-\mathbb{E}[T]| \geqslant \alpha p\sigma^2) \qquad \text{car } (T > (1+\alpha)p\sigma^2) \subset (|T-\mathbb{E}[T]| \geqslant \alpha p\sigma^2)\\
&\leqslant \frac{2p\sigma^4}{\alpha^2p^2\sigma^4} = \frac{2}{p\alpha^2}
\end{aligned}$$

Partie III. Minimisation d'une fonction non différentiable

Voir l'encadré pour l'explication du "non différentiable".

17.

$$\begin{aligned}
\|u+tv\| - \|u\| &= \frac{\|u+tv\|^2 - \|u\|^2}{\|u+tv\| + \|u\|}\\
&= \frac{\langle u+tv, u+tv\rangle - \|u\|^2}{\|u+tv\| + \|u\|}\\
&= \frac{\|u\|^2 + 2t\langle u,v\rangle + t^2\|v\|^2 - \|u\|^2}{\|u+tv\| + \|u\|}
\end{aligned}$$

En effet la fonction

$$\begin{aligned} J : \mathbb{R}^p &\to \mathbb{R} \\ X &\mapsto \frac{1}{2}\|AX - Y\|^2 + \|BX\| \end{aligned}$$

n'est pas différentiable sur $\mathbb{R}^p$ entier à cause du terme $\|BX\|$.
$\|BX\|$ est différentiable si et seulement si $BX \neq 0$, et dans ce cas

$$\frac{\partial\|BX\|}{\partial X} = \frac{{}^tBBX}{\|BX\|}$$

Du fait de la continuité de la norme $x \mapsto \|x\|$, cette expression admet un développement limité en 0:

$$\frac{2t\langle u, v\rangle + t^2\|v\|^2}{\|u + tv\| + \|u\|} \underset{t\to 0}{=} \frac{t\langle u, v\rangle}{\|u\|} + \mathrm{o}\,(t)$$

On a donc bien:

$$\lim_{t\to 0} \frac{\|u + tv\| - \|u\|}{t} = \frac{\langle u, v\rangle}{\|u\|}$$

18. En reprenant l'équation 3, appliquée à $H = -tX$, on obtient

$$\begin{aligned} J(X_0 - tX) - J(X_0) &= -t\langle D(X_0), X\rangle + \frac{1}{2}t^2 XMX + \underbrace{\|BX_0 - tBX\| - \|BX_0\|}_{= -t\frac{\langle BX_0, BX\rangle}{\|BX_0\|} + \mathrm{o}\,(t)} \\ &= -t\,{}^tXD(X_0) - t\,{}^tX\,{}^tB\frac{BX_0}{\|BX_0\|} + \mathrm{o}\,(t) \\ &= -t\,{}^tX(D(X_0) + {}^tB\frac{BX_0}{\|BX_0\|}) + \mathrm{o}\,(t) \end{aligned}$$

Si $D(X_0) + {}^tB\frac{BX_0}{\|BX_0\|} \neq 0$, alors $\mathrm{vect}(D(X_0) + {}^tB\frac{BX_0}{\|BX_0\|})^\perp \subsetneq \mathbb{R}^p$.

On peut alors trouver $X \in \mathbb{R}^p$ tel que $\langle X, D(X_0) + {}^tB\frac{BX_0}{\|BX_0\|}\rangle > 0$ (quitte à changer X en $-X$), mais alors pour $t > 0$ suffisamment petit, on a

$$J(X_0 - tX) - J(X_0) < 0$$

contredisant la minimalité de $J(X_0)$.

On conclut donc

$$D(X_0) + {}^tB\frac{BX_0}{\|BX_0\|} = 0$$

ou encore

$$-D(X_0) = {}^tBV$$

avec

$$V = \frac{BX_0}{\|BX_0\|} \in \mathcal{M}_{m,1}(\mathbb{R})$$

et donc

$$\begin{aligned} \|V\| &= 1 \\ \|BX_0\|V - BX_0 &= 0 \end{aligned}$$

i.e. $-D(X_0) \in \mathcal{N}(X_0)$.

19.

19a. On a, puisque $BX_0 = 0$, $\forall t > 0$,

$$J(X_0 - tH) - J(X_0) = -t(\langle D(X_0), H\rangle - \|BH\|) + \mathrm{o}\,(t)$$

Pour $t > 0$ suffisamment petit, $J(X_0 - tH) - J(X_0) \geqslant 0$ est donc du signe de $-t(\langle D(X_0), H\rangle - \|BH\|)$, ce qui implique

$$\langle D(X_0), H\rangle - \|BH\| \leqslant 0 \tag{4}$$

19b.

$$\forall X \in \ker B, \quad \langle D(X_0), X\rangle \leqslant \|BX\| = 0$$

et

$$\begin{aligned} -\langle D(X_0), X\rangle &= \langle D(X_0), -X\rangle \\ &\leqslant \|BX\| = 0 \end{aligned}$$

d'ou $\langle D(X_0), X\rangle = 0$.
On conclut $D(X_0) \in \ker B^{\perp}$.

19c. D'après la question 5 on a donc $D(X_0) \in \operatorname{Im} {}^tB$:

$$\exists W \in \mathcal{M}_{m,1}(\mathbb{R}), \quad D(X_0) = {}^tBW$$

19d. On a:

$$\mathcal{M}_{m,1}(\mathbb{R}) = \operatorname{Im} B \oplus \ker {}^tB$$

Posons

$$W = W_1 + W_2$$

avec $(W_1 = BV, W_2) \in \operatorname{Im} B \times \ker {}^tB$, avec $V \in \mathcal{M}_{p,1}(\mathbb{R})$.

$$\begin{aligned} D(X_0) &= {}^tBW \\ &= {}^tBW_1 + \underbrace{{}^tBW_2}_{=0} \\ &= {}^tBBV \end{aligned}$$

19e. On applique l'équation 4 à $H = V$ pour obtenir

$$\begin{aligned}\|BV\|^2 &= {}^tBVBV\\ &= {}^tV\,{}^tBBV\\ &\leqslant \|BV\|\end{aligned}$$

ce qui implique $\|BV\| \leqslant 1$.

Enfin

$$\|BX_0\|BV - BX_0 = 0 \qquad \text{(on rappelle que } BX_0 = 0)$$

donc $-D(X_0) \in \mathcal{N}(X_0)$.

20.

20a.

$$\begin{aligned}\|u+v\|^2 - (\|u\| + \frac{\langle u,v\rangle}{\|u\|})^2 &= \|u\|^2 + 2\langle u,v\rangle + \|v\|^2 - \|u\|^2 - 2\langle u,v\rangle - \frac{\langle u,v\rangle^2}{\|u\|^2}\\ &= \frac{\|u\|^2\|v\|^2 - \langle u,v\rangle^2}{\|u\|^2}\end{aligned}$$

20b. D'après l'inégalité de Cauchy-Schwartz,

$$\|u\|^2\|v\|^2 - \langle u,v\rangle^2 \geqslant 0$$

et donc

$$\begin{aligned}&\|u+v\|^2 - (\|u\| + \frac{\langle u,v\rangle}{\|u\|})^2 \geqslant 0\\ \Leftrightarrow \qquad & (0 \leqslant)\|u+v\|^2 \geqslant (\|u\| + \frac{\langle u,v\rangle}{\|u\|})^2\\ \Leftrightarrow \qquad & \|u+v\| \geqslant \left|\|u\| + \frac{\langle u,v\rangle}{\|u\|}\right|\\ & \geqslant \|u\| + \frac{\langle u,v\rangle}{\|u\|}\end{aligned} \tag{5}$$

21.

21a. On pose $W = {}^tBV \in \mathcal{N}(X_0)$. D'après l'inégalité 5,

$$\begin{aligned}\|BX_0\|(\|BX\| - \|BX_0\|) &= \|BX_0\|(\|BX_0 + B(X - X_0)\| - \|BX_0\|)\\ &\geqslant \langle BX_0, BX - BX_0\rangle\\ &= \|BX_0\|\langle V, BX - BX_0\rangle\\ &= \|BX_0\|\langle {}^tBV, X - X_0\rangle\\ &= \|BX_0\|\langle W, X - X_0\rangle\end{aligned}$$

ce qui donne dans le cas où $BX_0 \neq 0$,

$$\|BX\| - \|BX_0\| \geqslant \langle W, X - X_0\rangle$$

Dans le cas où $BX_0 = 0$,

$$\begin{aligned}\langle W, X - X_0\rangle &= \langle {}^tBV, X - X_0\rangle \\ &= \langle V, BX - BX_0\rangle \\ &= \langle V, BX\rangle \\ &\leqslant \|V\|\|BX\| && \text{(Inégalité de Cauchy-Schwartz)} \\ &\leqslant \|BX\| && (\|V\| \leqslant 1) \\ &= \|BX\| - \|BX_0\|\end{aligned}$$

21b. On applique l'égalité 3 à $H = X - X_0$:

$$\begin{aligned}J(X) - J(X_0) &= \langle D(X_0), X - X_0\rangle + \frac{1}{2}\underbrace{{}^t(X - X_0)M(X - X_0)}_{\geqslant 0 (\text{cf. } 11)} + \|BX\| - \|BX_0\| \\ &\geqslant \langle D(X_0), X - X_0\rangle + \|BX\| - \|BX_0\| \\ &\geqslant \langle D(X_0), X - X_0\rangle + \langle W, X - X_0\rangle \\ &= \langle D(X_0) + W, X - X_0\rangle\end{aligned} \tag{6}$$

22. Les questions 18 et 19 ont montré

$$X_0 \in \mathcal{M}_{p,1}(\mathbb{R}) \text{ réalise le minimum global de } J \Rightarrow -D(X_0) \in \mathcal{N}(X_0)$$

Supposons maintenant $-D(X_0) \in \mathcal{N}(X_0)$. D'après l'équation 6, appliquée à $W = -D(X_0)$,

$$\forall X \in \mathcal{M}_{p,1}(\mathbb{R}), \quad J(X) - J(X_0) \geqslant \langle 0, X - X_0\rangle = 0$$

ce qui montre que X_0 réalise le minimum global de J.

Ainsi

$$X_0 \in \mathcal{M}_{p,1}(\mathbb{R}) \text{ réalise le minimum global de } J \Leftrightarrow -D(X_0) \in \mathcal{N}(X_0)$$

23.

23a. Posons

$$\begin{aligned} f : \mathbb{R}^{+*} &\to \mathbb{R} \\ x &\mapsto \frac{x}{(\lambda + x)^2}\end{aligned}$$

f est définie et dérivable sur $\mathbb{R}^{+*}$, et

$$f'(x) = \frac{-x^2 + \lambda^2}{(\lambda + x)^4}$$

$$\begin{aligned}\forall x \in]0, \lambda[, &\quad f'(x) > 0 \\ \forall x \in]\lambda, +\infty[, &\quad f'(x) < 0\end{aligned}$$

ce aui montre que f admet un maximum atteint en λ, qui vaut

$$f(\lambda) = \frac{1}{4\lambda}$$

ce qui permet de conclure

$$\begin{aligned} \forall \lambda \in \,]0,+\infty[, \quad F(\lambda) &= -1 + \sum_{i=1}^{k} f(\alpha_i^2) y_i^2 \\ &\leqslant \quad -1 + \frac{1}{4\lambda} \sum_{i=1}^{k} y_i^2 \end{aligned} \tag{7}$$

23b.

$$F(0) = -1 + \rho(Y) > 0$$

F est strictement décroissante sur $[0,+\infty[$ et $\lim_{\lambda \to +\infty} F(\lambda) = -1$.

F réalise donc une bijection de $[0,+\infty[$ dans $]-1,-1+\rho(Y)]$, ce qui prouve l'existence et l'unicité de $\lambda_0 \in \,]0,+\infty[$ tel que $F(\lambda_0) = 0$.

D'après l'inégalité 7, on a de plus

$$\lambda_0 \leqslant \frac{1}{4} \sum_{i=1}^{k} y_i^2$$

23c. D'après la question 22, il suffit de vérifier que $-D(X_0) \in \mathcal{N}(X_0)$. Or,

$$\begin{aligned} D(X_0) &= {}^t\!AAX_0 - {}^t\!AY \\ &= X_0 - Y \\ &= -\,{}^t\!BV \end{aligned}$$

On a

$$\|V\|^2 = \begin{cases} \rho(Y) & \text{si } \rho(Y) \leqslant 1 \\ F(\lambda_0) + 1 = 1 & \text{si } \rho(Y) < 1 \end{cases}$$

donc $\|V\| \leqslant 1$.

D'autre part,

$$B^2V = \begin{array}{ccc} \begin{array}{ccccccc} & & & k & & & \end{array} & & \\ \begin{bmatrix} \alpha_1^2 & & & & & & \\ & \alpha_2^2 & & & \mathbf{0} & & \\ & & \ddots & & & & \\ & & & \alpha_k^2 & & & \\ & \mathbf{0} & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{bmatrix} & \begin{array}{c} \\ \\ \\ k \\ \\ \\ \\ \end{array} & \begin{bmatrix} \frac{\alpha_1 y_1}{\beta+\alpha_1^2} \\ \frac{\alpha_2 y_2}{\beta+\alpha_2} \\ \vdots \\ \frac{\alpha_k y_k}{\beta+\alpha_k^2} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \end{array}$$

$$= \begin{bmatrix} \frac{\alpha_1^3 y_1}{\beta+\alpha_1^2} \\ \frac{\alpha_2^3 y_2}{\beta+\alpha_2} \\ \vdots \\ \frac{\alpha_k^3 y_k}{\beta+\alpha_k^2} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$BY = \begin{array}{c} \begin{array}{ccccc} & & k & & \end{array} \\ \begin{bmatrix} \alpha_1 & & & & \\ & \alpha_2 & & \mathbf{0} & \\ & & \ddots & & \\ & & & \alpha_k & \\ & \mathbf{0} & & 0 & \\ & & & & \ddots \\ & & & & 0 \end{bmatrix} \end{array} \begin{array}{c} \\ \\ \\ k \\ \\ \\ \\ \end{array} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} \alpha_1 y_1 \\ \alpha_2 y_2 \\ \vdots \\ \alpha_k y_k \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Donc on a:

$$BX_0 = BY - B^2 V$$

$$= \begin{bmatrix} \alpha_1 y_1 \\ \alpha_2 y_2 \\ \vdots \\ \alpha_k y_k \\ 0 \\ \vdots \\ 0 \end{bmatrix} - \begin{bmatrix} \frac{\alpha_1^3 y_1}{\beta+\alpha_1^2} \\ \frac{\alpha_2^3 y_2}{\beta+\alpha_2} \\ \vdots \\ \frac{\alpha_k^3 y_k}{\beta+\alpha_k^2} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$= \beta \begin{bmatrix} \frac{\alpha_1 y_1}{\beta+\alpha_1^2} \\ \frac{\alpha_2 y_2}{\beta+\alpha_2} \\ \vdots \\ \frac{\alpha_k y_k}{\beta+\alpha_k^2} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$= \beta V$$

(i) Si $\beta = 0$,

$$\begin{aligned} BX_0 &= 0 \\ &= \|BX_0\| V \end{aligned}$$

(ii) Si $\beta = \lambda_0$, on a vu que $\|V\| = 1$ et donc encore

$$BX_0 = \|BX_0\| V$$

Dnas tous les cas on a montré $-D(X_0) \in \mathcal{N}(X_0)$.

23d.

```
def FoncSom(alpha,y,lda):
  n = np.size(y)
  k=np.size(alpha)
  alpha = np.pad(alpha,(0,n-k))
  z=[k**2/(lda+k**2)**2 for k in alpha]
  return np.inner(z,np.square(y))-1


def CalcBeta(alpha,y,epsilon):
  a=0
  b=0.25*np.sum(np.square(y))

  while abs(b-a) >= epsilon:
    c = a + (b-a)/2
```

```
    if FoncSom(alpha,y,c)<0:
      b = c
    else:
      a = c

  return c


def CalcSolution(alpha,y,epsilon):
  n = np.size(y)
  k=np.size(alpha)
  alpha1 = np.pad(alpha,(0,n-k))
  rho = np.inner(np.square(y),np.pad([1/a**2 for a in alpha],(0,n-k)))
  beta=0

  if rho > 1:
    beta = CalcBeta(alpha1,y,epsilon)

  z=[k**2/(beta+k**2) for k in alpha1]

  return y-np.multiply(z,y)
```