

L'énoncé supposait, pour toutes les questions en langage Python, les bibliothèques usuelles déjà importées sous leurs raccourcis habituels.

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

Partie I : autour de l'adjoint

1. soit ℓ une application linéaire de E dans \mathbb{R} . Les formules dans la base orthonormée $\mathcal{B}_E = (e_1, e_2, \dots, e_p)$ de E et la linéarité de ℓ permettent d'écrire :

$$\forall x \in E, \quad \ell(x) = \ell\left(\sum_{k=1}^p \langle x, e_k \rangle_E e_k\right) = \sum_{k=1}^p \langle x, e_k \rangle_E \ell(e_k) = \langle x, \sum_{k=1}^p \ell(e_k) e_k \rangle_E = \langle a_0, x \rangle_E$$

avec $a_0 = \sum_{k=1}^p \ell(e_k) e_k$, qui est bien un vecteur de E comme combinaison linéaire (les $(\ell(e_k))_{1 \leq k \leq p}$ sont des réels) des vecteurs de \mathcal{B}_E .

On a montré l'existence, démontrons maintenant l'unicité de ce vecteur a_0 en supposant qu'il existe un deuxième vecteur b_0 de E tel que :

$$\forall x \in E, \quad \ell(x) = \langle b_0, x \rangle_E = \langle a_0, x \rangle_E.$$

On a alors : $\forall x \in E, \quad \langle a_0, x \rangle_E - \langle b_0, x \rangle_E = 0 \iff \forall x \in E, \quad \langle a_0 - b_0, x \rangle_E = 0.$

Le vecteur $a_0 - b_0$ est donc orthogonal à tout vecteur x de E : il appartient à $E^\perp = \{0_E\}$, donc $a_0 - b_0 = 0_E \iff a_0 = b_0.$

On a donc bien prouvé l'existence et l'unicité d'un vecteur a_0 de E tel que :

$$\forall x \in E, \quad \ell(x) = \langle a_0, x \rangle_E.$$

Remarque : question difficile sans indication ; la porte d'entrée au résultat était bien de faire apparaître le produit scalaire via la formule pour la décomposition d'un vecteur en base orthonormée.

2. Soit $y \in F$. L'application $\ell : E \rightarrow \mathbb{R}$ est clairement à valeurs réelles et linéaire, $x \mapsto \langle u(x), y \rangle_F$ par linéarité de $u \in \mathcal{L}(E, F)$ et par bilinéarité du produit scalaire.

D'après la question précédente, il existe donc un unique vecteur $z_y \in E$ tel que :

$$\forall x \in E, \quad \ell(x) = \langle z_y, x \rangle_E \iff \forall x \in E, \quad \langle u(x), y \rangle_F = \langle z_y, x \rangle_E.$$

3. L'application $u^* : F \rightarrow E$ est bien définie de F dans E , montrons qu'il s'agit d'une application linéaire.

$$y \mapsto z_y$$

Soient y et y' deux éléments de E , et soient $z_y = u^*(y)$, $z_{y'} = u^*(y')$; soit $\lambda \in \mathbb{R}$.

Pour tout x de E :

$$\begin{aligned} \langle u(x), \lambda \cdot y + y' \rangle_F &= \lambda \cdot \langle u(x), y \rangle_F + \langle u(x), y' \rangle_F \\ &= \lambda \cdot \langle x, z_y \rangle_E + \langle x, z_{y'} \rangle_E \quad \text{par définition de } z_y \text{ et } z_{y'} \\ &= \langle x, \lambda \cdot z_y + z_{y'} \rangle_E \quad \text{par bilinéarité du produit scalaire.} \end{aligned}$$

Par définition de $u^*(\lambda \cdot y + y')$, unique vecteur de E tel que $\langle u(x), \lambda \cdot y + y' \rangle_F = \langle x, u^*(\lambda \cdot y + y') \rangle_E$ on a donc :

$$u^*(\lambda \cdot y + y') = \lambda \cdot z_y + z_{y'} = \lambda \cdot u^*(y) + u^*(y'),$$

ce qui prouve bien, que u^* est une application linéaire (de F dans E).

L'application u^* s'appelle l'**application adjointe** de u , et on a de fait l'identité :

$$\forall x \in E, \forall y \in F, \quad \langle u(x), y \rangle_F = \langle x, u^*(y) \rangle_E. \quad (1)$$

4. Les coefficients de la matrice $A^* = \text{Mat}_{\mathcal{B}_E, \mathcal{B}_F}(u^*)$ sont les coordonnées des images $u^*(f_1), \dots, u^*(f_n)$ des vecteurs de la base \mathcal{B}_F , dans la base \mathcal{B}_E .

Toujours d'après les formules en base orthonormée (ici encore dans E), pour tout $(i, j) \in \llbracket 1; p \rrbracket \times \llbracket 1; n \rrbracket$:

$$A_{i,j}^* = \langle e_i, u^*(f_j) \rangle_E = \langle u(e_i), f_j \rangle_F = A_{j,i}$$

puisque cette fois, la matrice A de u contient colonne par colonne, les coordonnées des images $u(e_i)$ ($1 \leq i \leq p$) dans la base orthonormée \mathcal{B}_F .

Ces relations entre les coefficients de A^* et A montrent bien que :

$$\text{Mat}_{\mathcal{B}_E, \mathcal{B}_F}(u^*) = {}^t A.$$

Or une matrice et sa transposée ont le même rang (théorème du cours), donc il est en de même des applications linéaires que ces deux matrices représentent, c'est-à-dire :

$$\text{rg}(u^*) = \text{rg}(u).$$

L'adjoint de l'adjoint $(u^*)^*$ est une application linéaire de E dans F dont la matrice représentative de la base \mathcal{B}_E dans la base \mathcal{B}_F , est ${}^t A^* = {}^t({}^t A) = A$.

Par unicité de la représentation matricielle d'une application linéaire dans des bases données, on a donc :

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

5. Soit $y \in \text{Im}(u^*)$: alors y est un vecteur de E pour lequel il existe $z \in F$ tel que $y = u^*(z)$.

Mais alors, pour tout vecteur x de $\text{Ker}(u)$:

$$\langle y, x \rangle_E = \langle u^*(z), x \rangle_E = \langle z, u(x) \rangle_F = \langle z, 0_F \rangle_F = 0,$$

ce qui prouve que $y \in \text{Im}(u^*)$ appartient toujours à $\text{Ker}(u)^\perp$: on a donc prouvé ici l'inclusion

$$\text{Im}(u^*) \subset \text{Ker}(u)^\perp.$$

Or d'après 5., $\dim \text{Im}(u^*) = \text{rg}(u^*) = \text{rg}(u) = \dim E - \dim \text{Ker}(u)$ d'après le théorème du rang, et puisqu'on est en dimension finie, alors $\text{Ker}(u)$ et son orthogonal sont supplémentaires dans E , ce qui prouve que : $\dim \text{Im}(u^*) = \dim E - \dim \text{Ker}(u) = \dim \text{Ker}(u)^\perp$.

Une inclusion et l'égalité des dimensions donnent bien :

$$\text{Im}(u^*) = \text{Ker}(u)^\perp.$$

6. La linéarité de u^* donne directement :

$$x \in \text{Ker}(u) \iff u(x) = 0_E \implies u^*(u(x)) = u^*(0_E) \implies u^* \circ u(x) = 0_E,$$

donc $\text{Ker}(u) \subset \text{Ker}(u^* \circ u)$.

Réciproquement, soit $x \in \text{Ker}(u^* \circ u)$, vérifiant donc $u^*(u(x)) = 0_E$.

Pour pouvoir utiliser la relation (1), on calcule alors :

$$\langle x, u^*(u(x)) \rangle_E = 0 \implies \langle u(x), u(x) \rangle_F = 0 \implies \|u(x)\|_F^2 = 0 \implies u(x) = 0_F \implies x \in \text{Ker}(u),$$

donc $\text{Ker}(u^* \circ u) \subset \text{Ker}(u)$, d'où l'égalité $\text{Ker}(u^* \circ u) = \text{Ker}(u)$ par double inclusion.

On montre ensuite une seule inclusion :

pour tout $y \in \text{Im}(u^* \circ u)$, il existe $x \in E$ tel que $y = u^* \circ u(x) \iff y = u^*(u(x))$, donc $y \in \text{Im}(u^*)$ et $\text{Im}(u^* \circ u) \subset \text{Im}(u^*)$.

Mais d'après le théorème du rang (R) pour les applications linéaires $u^* \circ u \in \mathcal{L}(E)$ et $u^* \in \mathcal{L}(F, E)$:

$$\dim \text{Im}(u^* \circ u) \stackrel{(R)}{=} \dim E - \dim \text{Ker}(u^* \circ u) = \dim E - \dim \text{Ker}(u) \stackrel{(R)}{=} \text{rg}(u) \stackrel{4.}{=} \text{rg}(u^*) = \dim \text{Im}(u^*),$$

donc l'inclusion démontrée et l'égalité des dimensions donnent bien :

$$\text{Im}(u^* \circ u) = \text{Im}(u^*).$$

7. Soit l'application $w : \text{Im}(u^*) \rightarrow \text{Im}(u^*)$ définie par :

$$\forall x \in \text{Im}(u^*), \quad w(x) = u^* \circ u(x).$$

La linéarité de u et u^* entraîne celle de w par composition, donc w est un endomorphisme de $\text{Im}(u^*)$; comme ce dernier est un espace vectoriel de dimension finie, il suffit de prouver que w est injective pour que ce soit un automorphisme de $\text{Im}(u^*)$ (isomorphisme de $\text{Im}(u^*)$ dans lui-même).

Soit $x \in \text{Ker}(w)$:

alors x est un vecteur de $\text{Im}(u^*) \stackrel{5.}{=} \text{Ker}(u)^\perp$ qui vérifie $w(x) = 0_E \iff u^* \circ u(x) = 0_E$,

donc $x \in \text{Ker}(u^* \circ u) \stackrel{6.}{=} \text{Ker}(u)$.

Ainsi, $x \in \text{Ker}(u)^\perp \cap \text{Ker}(u) = \{0_E\}$ puisqu'un sous-espace vectoriel et son orthogonal sont toujours en somme directe, et donc $x = 0_E$, ce qui prouve que $\text{Ker}(w) \subset \{0_E\}$.

Comme l'inclusion réciproque est toujours vraie du fait que $\text{Ker}(w)$ est un sous-espace vectoriel de $\text{Im}(u^*) \subset E$, alors $\text{Ker}(w)$ est bien réduit au vecteur nul (de E), et w est injective, donc bijective.

Matriciellement cela signifie que pour tout $Y \in \text{Im}({}^tA)$ de la forme $Y = {}^tAZ$ avec $Z \in \mathcal{M}_{n,1}(\mathbb{R})$, l'équation $Y = {}^tAAX$ d'inconnue $X \in \mathcal{M}_{p,1}(\mathbb{R})$ admet une unique solution.

Cela revient aussi à dire que pour tout $Z \in \mathcal{M}_{n,1}(\mathbb{R})$, l'équation ${}^tAZ = {}^tAAX$ d'inconnue $X \in \mathcal{M}_{p,1}(\mathbb{R})$ admet une unique solution.

8. Soit π le projecteur orthogonal de F sur $\text{Im}(u)$, et Q sa matrice dans la base \mathcal{B}_F .

(a) Pour tout x de E , puisque $u(x) \in \text{Im}(u)$, alors ce vecteur est invariant par π :

$$\forall x \in E, \quad \pi \circ u(x) = u(x).$$

Matriciellement dans les bases \mathcal{B}_E et \mathcal{B}_F , cela se traduit par l'égalité matricielle :

$$QA = A.$$

Or Q est la matrice représentative d'un *projecteur orthogonal* dans la *base orthonormée* \mathcal{B}_F : on sait d'après le cours, que Q est alors symétrique, et en appliquant la transposition aux deux membres de l'égalité matricielle précédente, on obtient :

$${}^tA {}^tQ = {}^tA \iff {}^tAQ = {}^tA.$$

- (b) Dans une base orthonormée de F obtenue en concaténant une base orthonormée de $\text{Im}(u) = \text{Im}(\pi)$ avec une base orthonormée de son orthogonal $\text{Im}(u)^\perp = \text{Ker}(\pi)$, on sait la matrice de π est la matrice diagonale $D = \text{Diag}(\underbrace{1, 1, \dots, 1}_{\dim \text{Im}(u)}, 0, \dots, 0)$.

En effet, tout élément x de $\text{Im}(u)$ est invariant par π , donc chaque élément de la base orthonormée de $\text{Im}(u)$ choisie ici est égal à son image par π .

On sait d'après le cours, que puisque les matrices Q et D sont semblables, alors elles ont la même trace : comme $\text{Tr}(D)$ est clairement égale à $\dim \text{Im}(u) = \text{rg}(u)$, on a également :

$$\text{Tr}(Q) = \text{Tr}(D) = \text{rg}(u).$$

9. La matrice $M = {}^tAA$, qui est carrée d'ordre p , est inversible si et seulement si son rang est maximal, c'est-à-dire égal à p .

L'égalité $\text{Im}(u^* \circ u) = \text{Im}(u^*)$ obtenue en 6. assure que $\text{rg}(u^* \circ u) = \text{rg}(u^*)$, qui est aussi égal à $\text{rg}(u)$ d'après 4., donc :

$$\text{rg}(u^* \circ u) = \text{rg}(u),$$

et ainsi, puisque A représente matriciellement u et $M = {}^tAA$, représente $u^* \circ u$:

$$M \text{ est inversible } \iff \text{rg}(M) = p \iff \text{rg}(u^* \circ u) = p \iff \text{rg}(u) = p \iff \text{rg}(A) = p.$$

10. On suppose que le rang de A est égal à p .

- (a) Soit $x \in F$; alors $y = \pi(x)$, projection orthogonale de x sur $\text{Im}(u)$, est caractérisé par :

$$y = \pi(x) \in \text{Im}(u) \quad (*) \quad \text{et} \quad x - y \in \text{Im}(u)^\perp \quad (**)$$

ce qui se traduit matriciellement dans les bases orthonormées \mathcal{B}_E et \mathcal{B}_F , avec $X = [x]_{\mathcal{B}_F}$ et $Y = [y]_{\mathcal{B}_E}$:

$$(*) \exists W \in \mathcal{M}_{p,1}(\mathbb{R}); Y = AW \quad \text{et} \quad (***) \forall Z \in \mathcal{M}_{p,1}(\mathbb{R}), X - Y \perp AZ \iff \forall Z \in \mathcal{M}_{p,1}(\mathbb{R}), {}^t(AZ)(X - Y) = 0.$$

La dernière égalité s'écrit aussi :

$$\forall Z \in \mathcal{M}_{p,1}(\mathbb{R}), {}^tZ {}^tA(X - Y) = 0 \iff \forall Z \in \mathcal{M}_{p,1}(\mathbb{R}), {}^tAX - {}^tAY \perp Z \iff {}^tAX - {}^tAY = 0$$

puisque le seul vecteur orthogonal à tout vecteur, est le vecteur nul.

Ainsi : ${}^tAY = {}^tAX \xrightarrow{(*)} \underbrace{{}^tAA}_{=M} W = {}^tAX \implies W = (M)^{-1} {}^tAX \implies AW = A(M)^{-1} {}^tAX$, ce qui

s'écrit aussi :

$$Y = A(M)^{-1} {}^tAX.$$

Cette relation, vraie pour tout $x \in F$ avec $y = \pi(x)$, caractérise de façon unique la matrice représentative de π , qui est la matrice Q :

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

- (b) La fonction Python ci-dessous prend en entrée une matrice A de type `array` de taille $n \times p$ (format calculé au début de l'exécution), teste si A est de rang p et renvoie la matrice $Q = A(M)^{-1} {}^tA$ dans ce cas, un message d'erreur sinon.

```
def Calcule_Q(A):
    n, p = np.shape(A)
    if al.matrix_rank(A) == p:
        M = np.dot(np.transpose(A), A)
        return np.dot(A, np.dot(al.inv(M), np.transpose(A)))
    else:
        return "rang incompatible"
```

11. Pour toute matrice colonne $X \in \mathcal{M}_{p,1}(\mathbb{R})$:

$${}^tXMX = {}^tX {}^tAAX = {}^t(AX)AX = \|AX\|^2 \geq 0.$$

Partie II : minimisation d'une fonction quadratique

Dans cette partie, l'énoncé s'intéressait à la minimisation de la fonction J_0 , définie sur $\mathcal{M}_{p,1}(\mathbb{R})$ par :

$$J_0(X) = \frac{1}{2} \|AX - Y\|^2, \quad \text{pour tout } X \in \mathcal{M}_{p,1}(\mathbb{R}).$$

On rappelle que l'énoncé définit dans le préambule : $\forall X \in \mathcal{M}_{p,1}(\mathbb{R}), D(X) = MX - {}^tAY$.

12. Soient X et H deux éléments de $\mathcal{M}_{p,1}(\mathbb{R})$:

$$\begin{aligned} J_0(X+H) - J_0(X) &= \frac{1}{2} \|AH + (AX - Y)\|^2 - \frac{1}{2} \|AX - Y\|^2 \\ &= \frac{1}{2} (\|AH\|^2 + \|AX - Y\|^2 + 2\langle AH, AX - Y \rangle) - \frac{1}{2} \|AX - Y\|^2 \\ &= \frac{1}{2} {}^t(AH)AH + {}^t(AH)(AX - Y) = \frac{1}{2} {}^tH {}^tAAH + {}^tH {}^tA(AX - Y) \\ &= \frac{1}{2} {}^tHMH + {}^tH(MX - {}^tAY) = {}^tHD(X) + \frac{1}{2} {}^tHMH \\ &= \langle D(X), H \rangle + \frac{1}{2} {}^tHMH. \quad (2) \end{aligned}$$

13. Soit $X \in \mathcal{M}_{p,1}(\mathbb{R})$.

Si $D(X) = 0$, alors pour tout $H \in \mathcal{M}_{p,1}(\mathbb{R})$, la relation (2) s'écrit : $J_0(X+H) - J_0(X) = \frac{1}{2} {}^tHMH \geq 0$ d'après 11., donc :

$$\forall H \in \mathcal{M}_{p,1}(\mathbb{R}), \quad J_0(X+H) \geq J_0(X),$$

qui s'écrit encore, puisque pour tout $X' \in \mathcal{M}_{p,1}(\mathbb{R})$, $X' = X + H$ avec $H = X' - X$:

$$\forall X' \in \mathcal{M}_{p,1}(\mathbb{R}), \quad J_0(X') \geq J_0(X),$$

ce qui signifie bien que la fonction J_0 possède un minimum global en X .

Réciproquement, si J_0 possède un minimum global en X : alors pour tout $H \in \mathcal{M}_{p,1}(\mathbb{R})$, la fonction $\varphi_H : t \in \mathbb{R} \mapsto J_0(X + t \cdot H)$ admet un minimum global en 0.

Or :

$$\begin{aligned} \forall t \in \mathbb{R}, \varphi_H(t) &= J_0(X + t \cdot H) = J_0(X) + \langle D(X), t \cdot H \rangle + \frac{1}{2} {}^t(t \cdot H)M(t \cdot H) \\ &= J_0(X) + t \langle D(X), H \rangle + \frac{t^2}{2} {}^tHMH. \end{aligned}$$

La fonction φ_H est donc un trinôme du second degré, dérivable sur \mathbb{R} , et qui vérifie donc la condition nécessaire d'extremum en 0 :

$$\varphi_H'(0) = 0 \iff \langle D(X), H \rangle = 0.$$

Or cette dernière égalité doit être vraie pour tout vecteur H de $\mathcal{M}_{p,1}(\mathbb{R})$: on en déduit que $D(X)$ est orthogonal à tout vecteur de l'espace vectoriel auquel il appartient, c'est donc le vecteur nul.

Il n'aura pas échappé aux fins connaisseurs du cours sur les fonctions de plusieurs variables, qu'on a repris ici l'esprit de la preuve de la condition nécessaire d'extremum (local ou global) pour une fonction de plusieurs variables de classe \mathcal{C}^1 .

L'énoncé note dans la suite \mathbb{S}_0 l'ensemble :

$$\mathbb{S}_0 = \{M \in \mathcal{M}_{p,1}(\mathbb{R}) \mid D(X) = 0\}.$$

14. En reprenant l'isomorphisme w de la question 7. : puisque $u^*(y)$ appartient à $\text{Im}(u^*)$, alors il existe $x_0 \in \text{Im}(u^*) \stackrel{5.}{=} \text{Ker}(u)^\perp$ unique tel que $u^* \circ u(x_0) = u^*(y)$, ce qui se traduit matriciellement :

$$\exists! X_0 \in \text{Ker}(A)^\perp ; {}^t A A X_0 = {}^t A Y \iff M X_0 - {}^t A Y = 0 \iff D(X_0) = 0.$$

On a donc bien démontré qu'il existe $X_0 \in \text{Ker}(A)^\perp$ tel que

$$\mathbb{S}_0 \cap \text{Ker}(A)^\perp = \{X_0\}.$$

15. Soit $X \in \mathbb{S}_0$.

- (a) Le vecteur $\pi(y)$, représenté matriciellement par QY , est caractérisé par le fait que c'est le seul vecteur qui vérifie : $y - \pi(y) \in \text{Im}(u)^\perp$.

On cherche donc à montrer que $y - u(x)$, représenté matriciellement par $Y - AX$, appartient à $\text{Im}(A)^\perp$, donc que pour tout $Z \in \mathcal{M}_{p,1}(\mathbb{R})$, $Y - AX \perp AZ$.

Or, pour tout $Z \in \mathcal{M}_{p,1}(\mathbb{R})$, en reprenant matriciellement les propriétés de l'adjoint étudiées dans la partie 1 :

$$\langle AZ, Y - AX \rangle = \langle Z, {}^t A(Y - AX) \rangle = \langle Z, \underbrace{{}^t A Y - {}^t A A X}_{=-D(X)=0} \rangle = 0$$

ce qu'on pouvait aussi écrire, uniquement par voie matricielle :

$$\langle AZ, Y - AX \rangle = {}^t A Z(Y - AX) = {}^t Z {}^t A(Y - AX) = -{}^t Z \underbrace{({}^t A A X - {}^t A Y)}_{=D(X)=0} = 0.$$

On a bien prouvé que $Y - AX \in \text{Im}(A)^\perp$, ce qui prouve que $u(x) = \pi(y)$, ou encore :

$$AX = QY. \quad (3)$$

- (b) Les vecteurs X et X_0 appartiennent tous les deux à \mathbb{S}_0 , donc ils vérifient :

$$MX = {}^t A Y \text{ et } M X_0 = {}^t A Y \implies MX = M X_0 \implies M(X - X_0) = 0,$$

donc $X - X_0 \in \text{Ker}(M)$ qui est égal à $\text{Ker}(A)$ d'après la question 6.

Mais alors, tout $X \in \mathcal{M}_{p,1}(\mathbb{R})$ s'écrit $X = (X - X_0) + X_0$ avec $X - X_0 \in \text{Ker}(A)$ et $X_0 \in \text{Ker}(A)^\perp$ (d'après 14.).

Le théorème de Pythagore donne alors : $\|X\|^2 = \|X - X_0\|^2 + \|X_0\|^2$,

et si $X \neq X_0$, alors $\|X - X_0\| > 0 \implies \|X\|^2 > \|X - 0\|^2 \implies \|X\| > \|X_0\|$ par stricte croissance de la racine carrée sur \mathbb{R}^+ , et positivité de la norme.

- (c) On suppose que le rang de A est égal à p : alors d'après la question 9., M est inversible, et :

$$X \in \mathbb{S}_0 \iff MX = {}^t A Y \iff X = M^{-1} {}^t A Y,$$

ce qui prouve que \mathbb{S}_0 est un ensemble réduit au seul vecteur ci-dessus :

$$\mathbb{S}_0 = \{X\} \text{ avec } X = M^{-1} {}^t A Y. \quad (4)$$

16. Dans cette question, on suppose de nouveau que le rang de A est égal à p et que Y s'écrit sous la forme

$$Y = A U_0 + Z,$$

où U_0 est un vecteur constant fixé, et $Z = (z_1, \dots, z_n)$ est un vecteur aléatoire dont les composantes $(z_i)_{1 \leq i \leq n}$ sont des variables aléatoires indépendantes, identiquement distribuées telles que $z_i \hookrightarrow \mathcal{N}(0, \sigma^2)$.

L'énoncé définit X par la formule (4), c'est-à-dire : $X = M^{-1} {}^t A Y$, et on considère aussi la variable aléatoire $T = \|A(X - U_0)\|^2$.

(a) Avec la formule (3) :

$$T = \|AX - AU_0\|^2 = \|QY - AU_0\|^2 = \|Q(AU_0 + Z) - AU_0\|^2 = \|QZ + (QA - A)U_0\|^2.$$

Or, comme on l'a vu en 8.(a) : ${}^tAQ = {}^tA \iff {}^tQA = A \implies QA = A$ en transposant, et du fait que Q est symétrique comme matrice d'un projecteur orthogonal dans une base orthonormée.

On a alors $QA - A = 0$, et on obtient bien :

$$T = \|QZ\|^2 = {}^t(QZ)QZ = {}^tZ {}^tQQZ = {}^tZQ^2Z = {}^tZQZ$$

puisque $Q^2 = Q$, vu que Q représente matriciellement un projecteur.

(b) La fonction ci-dessous réutilise celle de la question 10.(b) : elle calcule la matrice Q à partir de la matrice A , simule un vecteur aléatoire z contenant n réalisations indépendantes de la loi $\mathcal{N}(0, \sigma^2)$, et rend ensuite le produit ${}^tZQZ = zQ {}^tz$, en prenant garde au fait que les vecteurs en Python sont considérés comme des matrices-lignes, et non des matrices-colonnes !

```
def simuleT(A, sigma):
    Q = Calcule_Q(A)
    z = rd.normal(0, sigma, n)
    return np.dot(z, np.dot(Q, np.transpose(z)))
```

(c) On calcule une valeur approchée de l'espérance de T en calculant la moyenne empirique d'un grand nombre de simulations de T :

```
def esperance(A, sigma):
    v = np.zeros(10000)
    for k in range(10000):
        v[k] = simuleT(A, sigma)
    return np.mean(v)
```

(d) Avec $\sigma = 1$, l'algorithme donne des valeurs approchées de l'espérance de T qui sont à chaque fois proches du nombre de colonnes de la matrice correspondante.

En remarquant la forme particulière de ces matrices A_k , à $p = k + 1$ colonnes et $n = k + 2$ lignes, les coefficients d'indices (i, j) avec $j > i$ sont nuls, ceux pour lesquels $i = j$ sont égaux à 1, ceux pour lesquels $i = j + 1$ sont égaux à 2, ceux pour lesquels $i = j + 2$ sont égaux à 3, etc..., il semble que $\mathbf{E}(T) = k + 1 = p$ pour la matrice A_k .

Le programme suivant fourni par l'énoncé, fixe ensuite la matrice $A_1 = \begin{pmatrix} 1 & 0 \\ 2 & 1 \\ 3 & 2 \end{pmatrix}$ et fait varier

l'écart-type σ en lui donnant les valeurs successives 1, 2, 3, 4, 5, 6.

Le graphique fait apparaître les valeurs suivantes (arrondies dans le tableau ci-dessous) pour l'espérance de T :

σ	1	2	3	4	5	6
$\mathbf{E}(T)$	2	8	18	32	50	72

Puisque $2 = 1 \times 2$, $8 = 2 \times 4$, $18 = 3 \times 6$, $32 = 4 \times 8$, $50 = 5 \times 10$ et $72 = 6 \times 12$, il semble qu'avec la matrice A_1 :

$$\mathbf{E}(T) = \sigma \times 2\sigma = 2\sigma^2.$$

(e) La variable aléatoire z_1 suit la loi $\mathcal{N}(0, \sigma^2)$: d'après le cours sur la loi normale, on sait alors que $Z_1 = \frac{z_1}{\sigma}$ suit la loi normale centrée réduite $\mathcal{N}(0, 1)$.

On sait alors que $\mathbf{E}(Z_1) = 0$, $\mathbf{E}(Z_1^2) = \mathbf{V}(Z_1) = 1$, et on cherche à calculer les deux moments suivants, qui valent respectivement $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^3 e^{-\frac{x^2}{2}} dx$ et $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^4 e^{-\frac{x^2}{2}} dx$.

Le moment d'ordre 3 est nul puisqu'on intègre une fonction impaire sur \mathbb{R} , et de même $\mathbf{E}(z_1^3) = \sigma^3 \mathbf{E}(Z_1^3) = 0$.

La fonction $x \mapsto x^4 e^{-\frac{x^2}{2}}$ est par contre paire sur \mathbb{R} , donc $\mathbf{E}(Z_1^4) = \frac{2}{\sqrt{2\pi}} \int_0^{+\infty} e^{-\frac{x^2}{2}} dx$.

Soit alors $A > 0$. Dans l'intégrale $\int_0^A x^4 e^{-\frac{x^2}{2}} dx$, on pose :

$$\begin{aligned} u(x) = x^3 &\longrightarrow u'(x) = 3x^2 \\ v'(x) = x e^{-\frac{x^2}{2}} &\longrightarrow v(x) = -e^{-\frac{x^2}{2}} \end{aligned}$$

Les fonctions u et v sont de classe \mathcal{C}^1 sur \mathbb{R}_+ , donc par intégration par parties :

$$\int_0^A x^4 e^{-\frac{x^2}{2}} dx = \left[-x^3 e^{-\frac{x^2}{2}} \right]_0^A + 3 \int_0^A x^2 e^{-\frac{x^2}{2}} dx = -A^3 e^{-\frac{A^2}{2}} + 3 \int_0^A x^2 e^{-\frac{x^2}{2}} dx.$$

La nouvelle intégrale converge et $\lim_{A \rightarrow +\infty} A^3 e^{-\frac{A^2}{2}} = 0$ par croissances comparées, donc lorsqu'on fait tendre A vers $+\infty$, on obtient :

$$\mathbf{E}(Z_1^4) = \frac{2}{\sqrt{2\pi}} \times 3 \int_0^{+\infty} x^2 e^{-\frac{x^2}{2}} dx = 3 \times \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^2 e^{-\frac{x^2}{2}} dx = 3\mathbf{E}(Z_1^2) = 3,$$

et ainsi, puisque $z_1^k = \sigma^k Z_1^k$ pour tout $k \in \mathbb{N}^*$:

$$\mathbf{E}(z_1^2) = \sigma^2, \quad \mathbf{E}(z_1^3) = 0, \quad \mathbf{E}(z_1^4) = 3\sigma^4.$$

- (f) La variable aléatoire $T = {}^t Z Q Z$ s'écrit, d'après les formules du produit matriciel et sachant que Q est une matrice carrée d'ordre n :

$$T = \sum_{i=1}^n z_i (QZ)_i = \sum_{i=1}^n z_i \left(\sum_{j=1}^n Q_{i,j} z_j \right) = \sum_{i=1}^n \sum_{j=1}^n Q_{i,j} z_i z_j.$$

En séparant les termes selon que $j = i$ ou non, et en tenant compte du fait que $Q_{i,j} = Q_{j,i}$ pour tout $(i, j) \in \llbracket 1; n \rrbracket^2$ puisque Q est symétrique, on obtient :

$$T = \sum_{i=1}^n Q_{i,i} z_i^2 + \sum_{1 \leq i < j \leq n} Q_{i,j} z_i z_j + \sum_{1 \leq j < i \leq n} Q_{i,j} z_i z_j = \sum_{i=1}^n Q_{i,i} z_i^2 + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^n Q_{i,j} z_i z_j = T_1 + 2T_2.$$

Les variables aléatoires $(z_i)_{1 \leq i \leq n}$ admettent chacune une variance et sont mutuellement indépendantes, donc avec aussi la linéarité de l'intégrale :

$$\mathbf{E}(T) = \sum_{i=1}^n Q_{i,i} \mathbf{E}(z_i^2) + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^n Q_{i,j} \mathbf{E}(z_i) \mathbf{E}(z_j) = \sigma^2 \sum_{i=1}^n Q_{i,i} = \text{Tr}(Q) \sigma^2 \stackrel{8.(b)}{=} \text{rg}(A) \sigma^2.$$

Lorsque $\sigma = 1$, les trois matrices A_1, A_2, A_3 fournies par l'énoncé sont de rangs respectifs 2, 3 et 4 (matrice échelonnées dont les colonnes sont clairement libres), donc la formule obtenue est cohérente avec les résultats des simulations.

C'est aussi le cas avec le programme qui utilise A_1 , matrice de rang 2 : on retrouve bien la formule $\mathbf{E}(T) = 2\sigma^2$ dans ce cas.

- (g) La variable aléatoire $T_1 T_2$ s'écrit : $\sum_{k=1}^n \sum_{i=1}^{n-1} \sum_{j=i+1}^n Q_{k,k} Q_{i,j} z_k^2 z_i z_j$ en développant le produit des deux (trois) sommes.

Or pour $k \in \llbracket 1; n \rrbracket$ et $1 \leq i < j \leq n$, il y a plusieurs cas possibles :

- Soit $k \neq i$ et $k \neq j$, auquel cas $\mathbf{E}(z_k^2 z_i z_j)$ existe et vaut, puisque z_k, z_i et z_j sont alors mutuellement indépendantes : $\mathbf{E}(z_k^2) \mathbf{E}(z_i) \mathbf{E}(z_j) = 0$ puisque les deux dernières espérances sont nulles.
- Soit $k = i$ et $i \neq j$, auquel cas $z_k^2 z_i z_j = z_k^3 z_j$ est le produit de deux variables indépendantes qui admettent une espérance, donc $\mathbf{E}(z_k^3 z_j) = \mathbf{E}(z_k^3) \mathbf{E}(z_j) = 0$ puisque les deux facteurs sont nuls.

Le cas où $k = j$ et $j \neq i$ est identique, et il est impossible d'avoir $k = i = j$ puisque $j > i$.

Bref, $T_1 T_2$ est une combinaison linéaire de variables aléatoires qui admettent toutes une espérance, mais cette espérance est systématiquement nulle !

On en déduit, par linéarité de l'espérance, que $\mathbf{E}(T_1 T_2)$ existe et vaut 0.

(h) notons $q = \sum_{k=1}^n Q_{k,k}^2$.

Avec $T_1^2 = \sum_{i=1}^n \sum_{j=1}^n Q_{i,i} Q_{j,j} z_i^2 z_j^2 = \sum_{i=1}^n Q_{i,i}^2 z_i^4 + \sum_{1 \leq i \neq j \leq n} Q_{i,i} Q_{j,j} z_i^2 z_j^2$, la linéarité de l'espérance donne

(les espérances existent au vu des résultats obtenus en (e)) :

$$\begin{aligned} \mathbf{E}(T_1^2) &= \sum_{i=1}^n Q_{i,i}^2 \mathbf{E}(z_i^4) + \sum_{1 \leq i \neq j \leq n} Q_{i,i} Q_{j,j} \mathbf{E}(z_i^2) \mathbf{E}(z_j^2) = 3\sigma^4 q + \sigma^4 \left(\left(\sum_{i=1}^n n Q_{i,i} \right)^2 - \sum_{i=1}^n Q_{i,i}^2 \right) \\ &= 3\sigma^4 q + \sigma^4 (p^2 - q) = \sigma^4 (2q + p^2). \end{aligned}$$

De la même façon :

$$\mathbf{E}(T_2^2) = \sum_{i=1}^{n-1} \sum_{j=i+1}^n \sum_{k=1}^{n-1} \sum_{\ell=k+1}^n Q_{i,j} Q_{k,\ell} \mathbf{E}(z_i z_j z_k z_\ell)$$

où dans cette somme multiple, j est forcément différent de i et ℓ différent de k .

Si les indices sont tous les 4 différents, ou bien si deux indices sont identiques mais différents des deux autres qui sont distincts entre eux, par indépendance on obtient des termes du type $\mathbf{E}(z_i) \mathbf{E}(z_j) \mathbf{E}(z_k) \mathbf{E}(z_\ell)$, ou $\mathbf{E}(z_i^2) \mathbf{E}(z_j) \mathbf{E}(z_k)$, qui sont tous nuls puisque $\mathbf{E}(z_i) = \mathbf{E}(z_i^3) = 0$. Il n'est pas possible d'avoir trois indices identiques et différents du quatrième.

Dans cette somme, ne subsistent donc que des termes où $i = k$ et $j = \ell$ avec $i < j$, soit :

$$\mathbf{E}(T_2^2) = \sum_{i=1}^{n-1} \sum_{j=i+1}^n Q_{i,j}^2 \mathbf{E}(z_i^2) \mathbf{E}(z_j^2) = \sigma^4 \times \frac{1}{2} \left(\sum_{i=1}^{n-1} \sum_{j=1}^n Q_{i,j}^2 - \sum_{i=1}^n Q_{i,i}^2 \right)$$

Or comme Q est symétrique :

$$\sum_{i=1}^n \sum_{j=1}^n Q_{i,j}^2 = \sum_{i=1}^n \left(\sum_{j=1}^n Q_{i,j} Q_{j,i} \right) = \sum_{i=1}^n (Q^2)_{i,i} \stackrel{Q^2=Q}{=} \sum_{i=1}^n Q_{i,i} = \text{Tr}(Q) = p,$$

donc :

$$\mathbf{E}(T_2^2) = \frac{\sigma^4}{2} (p - q).$$

- (i) De ce qui précède, on déduit que $T^2 = (T_1 + 2T_2)^2 = T_1^2 + 4T_1 T_2 + 4T_2^2$ admet une espérance, qui vaut :

$$\mathbf{E}(T^2) = \mathbf{E}(T_1^2) + 4\mathbf{E}(T_1 T_2) + 4\mathbf{E}(T_2^2) = \sigma^4 (2q + p^2) + 0 + 2\sigma^4 (p - q) = \sigma^4 (p^2 + 2p)$$

et T admet alors une variance donnée par la formule de Koenig-Huygens :

$$\mathbf{V}(T) = \mathbf{E}(T^2) - \mathbf{E}(T)^2 = \sigma^4 (p^2 + 2p) - (p\sigma^2)^2 = 2p\sigma^4.$$

Soit maintenant $\alpha > 0$: puisque $\mathbf{E}(T) = p\sigma^2$, alors

$$\begin{aligned} \mathbf{P}(T > (1 + \alpha)p\sigma^2) &= \mathbf{P}(T - \mathbf{E}(T) > \alpha p\sigma^2) \leq \mathbf{P}(T - \mathbf{E}(T) < -\alpha p\sigma^2) + \mathbf{P}(T - \mathbf{E}(T) > \alpha p\sigma^2) \\ &\leq \mathbf{P}(|T - \mathbf{E}(T)| > \alpha p\sigma^2). \end{aligned}$$

La variable aléatoire T admettant une variance, et puisque $\alpha p\sigma^2 > 0$, alors l'inégalité de Bienaymé-Tchebychev donne :

$$\mathbf{P}(|T - \mathbf{E}(T)| > \alpha p\sigma^2) \leq \frac{\mathbf{V}(T)}{\alpha^2 p^2 \sigma^4} = \frac{2p\sigma^4}{\alpha^2 p^2 \sigma^4} \implies \mathbf{P}(T > (1 + \alpha)p\sigma^2) \leq \frac{2}{p\alpha^2}.$$

Partie III : minimisation d'une fonction non différentiable

Dans cette partie, on suppose à nouveau que $Y \in \mathcal{M}_{n,1}(\mathbb{R})$ est un vecteur fixé (et non une variable aléatoire). On s'intéresse désormais à la fonction J définie sur $\mathcal{M}_{p,1}(\mathbb{R})$ par :

$$J(X) = \frac{1}{2} \|AX - Y\|^2 + \|BX\|, \quad \text{pour } X \in \mathcal{M}_{p,1}(\mathbb{R}). \quad (5)$$

Pour tout $X \in \mathcal{M}_{p,1}(\mathbb{R})$, on note $\mathcal{N}(X)$ l'ensemble

$$\mathcal{N}(X) = \{ {}^tBV \mid V \in \mathcal{M}_{m,1}(\mathbb{R}), \|V\| \leq 1 \text{ et } \|BX\|V - BX = 0 \}.$$

17. Soient u et v deux vecteurs tels que $u \neq 0$. Pour tout $t \neq 0$:

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

La continuité de la norme canonique sur \mathbb{R}^n assure que $\lim_{t \rightarrow 0} \|u + tv\| = \|u\|$, donc on a bien, par opérations sur les limites :

$$\lim_{t \rightarrow 0} \frac{\|u + tv\| - \|u\|}{t} = \frac{2\langle u, v \rangle}{2\|u\|} = \frac{\langle u, v \rangle}{\|u\|}. \quad (6)$$

18. Soit $X_0 \in \mathcal{M}_{p,1}(\mathbb{R})$ réalisant le minimum global de J .

Pour tout $H \in \mathcal{M}_{p,1}(\mathbb{R})$, la fonction $g_H : t \mapsto J(X_0 + tH)$ admet donc un minimum global en $t = 0$.

Or d'après les calculs précédents, et pour tout $t \neq 0$:

$$\begin{aligned} \frac{g(t) - g(0)}{t} &= \frac{\|(AX_0 - Y) + tAH\|^2 - \|AX_0 - Y\|^2}{2t} + \frac{\|BX_0 + tBH\| - \|BX_0\|}{t} \\ &= \frac{2t\langle AX_0 - Y, AH \rangle + t^2\|AH\|^2}{2t} + \frac{\|BX_0 + tBH\| - \|BX_0\|}{t} \\ &= \langle AX_0 - Y, AH \rangle + \frac{t}{2}\|AH\|^2 + \frac{\|BX_0 + tBH\| - \|BX_0\|}{t} \\ &\xrightarrow{t \rightarrow 0} \langle AX_0 - Y, AH \rangle + \frac{\langle BX_0, BH \rangle}{\|BH\|} \end{aligned}$$

Cette limite est finie, donc g est dérivable en 0 et $g'(0)$ est égal à cette limite ; et puisqu'en ce point elle admet un minimum global, alors nécessairement :

$$g'(0) = 0 \iff {}^tH {}^tA(AX_0 - Y) + \frac{{}^tH {}^tBBX_0}{\|BX_0\|} = 0 \iff {}^tH \left({}^tB \frac{BX_0}{\|BX_0\|} + \underbrace{({}^tAAX_0 - {}^tAY)}_{=D(X_0)} \right) = 0,$$

ce qui signifie :

$$\forall H \in \mathcal{M}_{p,1}(\mathbb{R}), \quad H \perp {}^tB \frac{BX_0}{\|BX_0\|} + D(X_0) \implies {}^tB \frac{BX_0}{\|BX_0\|} + D(X_0) = 0$$

puisque $\mathcal{M}_{p,1}(\mathbb{R})^\perp = \{0\}$, ou encore :

$$-D(X_0) = {}^tB \frac{BX_0}{\|BX_0\|}.$$

En posant $V = \frac{BX_0}{\|BX_0\|}$, alors $-D(X_0) = {}^tBV$ et $\|V\| = \frac{\|BX_0\|}{\|BX_0\|} = 1$,
et de plus $\|BX_0\|V - BX_0 = BX_0 - BX_0 = 0$ donc en effet, $-D(X_0) \in \mathcal{N}(X_0)$.

19. Soit $X_0 \in \mathcal{M}_{p,1}(\mathbb{R})$ réalisant le minimum global de J et vérifiant $BX_0 = 0$.

(a) Pour tout $H \in \mathcal{M}_{p,1}(\mathbb{R})$, $J(X_0 - H) \leq J(X_0)$ se réécrit :

$$\begin{aligned} \frac{1}{2} \|AX_0 - Y - AH\|^2 + \|BX_0 - BH\| &\geq \frac{1}{2} \|AX_0 - Y\|^2 + \underbrace{\|BX_0\|}_{=0} \\ \iff \frac{1}{2} \|AX_0 - Y\|^2 - \langle AX_0 - Y, AH \rangle + |-1| \cdot \|BH\| &\geq \frac{1}{2} \|AX_0 - Y\|^2 \end{aligned}$$

Sachant que $\langle AX_0 - Y, AH \rangle = {}^tH {}^tA(AX_0 - Y) = {}^tHD(X_0) = \langle D(X_0), H \rangle$, on a bien écrit :

$$\forall H \in \mathcal{M}_{p,1}(\mathbb{R}), \quad -\langle D(X_0), H \rangle + \|BH\| \geq 0 \iff \langle D(X_0), H \rangle \leq \|BH\|.$$

Remarque : on aura plutôt tendance à naturellement commencer par écrire $J(X_0 + H)$, et c'est au cours du calcul qu'on réalise que pour mieux gérer les signes, il est préférable de travailler avec $J(X_0 - H)$, ce qui ne change rien en réalité puisque $H \in \mathcal{M}_{p,1}(\mathbb{R})$ est pris quelconque.

(b) Tout $H \in \text{Ker}(B)$ vérifie $BX = 0$, et l'inégalité précédente se réécrit alors : $\langle D(X_0), H \rangle \leq 0$.
Mais $-H$ appartient encore à $\text{Ker}(B)$ puisque ce dernier est un espace vectoriel ; par conséquent, on a aussi $\langle D(X_0), -H \rangle \leq 0 \iff -\langle D(X_0), H \rangle \leq 0 \iff \langle D(X_0), H \rangle \geq 0$, donc par double inégalité :

$$\forall H \in \text{Ker}(B), \quad \langle D(X_0), H \rangle = 0, \quad \text{et donc } D(X_0) \in \text{Ker}(B)^\perp.$$

(c) En réécrivant le résultat de 5. sous forme matricielle : $\text{Ker}(B)^\perp = \text{Im}({}^tB)$, donc le fait que $D(X_0)$ appartienne à ce sous-espace assure l'existence d'un vecteur $W \in \mathcal{M}_{m,1}(\mathbb{R})$ tel que $D(X_0) = {}^tBW$.

(d) On applique ici le résultat de la question 7., qu'on avait exprimé matriciellement : il existe $V \in \mathcal{M}_{p,1}(\mathbb{R})$ (unique en fonction de W) tel que ${}^tBW = {}^tBBV$, donc tel que $D(X_0) = {}^tBBV$.

(e) Avec $H = V$, le résultat de 19.(a) se réécrit :

$$\langle {}^tBBV, V \rangle \leq \|BV\| \iff {}^tV {}^tBBV \leq \|BV\| \iff \|BV\|^2 \leq \|BV\|$$

Un réel positif est supérieur à son carré si et seulement s'il est inférieur ou égal à 1, donc $\|BV\| \leq 1$.

Ainsi avec $V' = -BV \in \mathcal{M}_{m,1}(\mathbb{R})$: on a bien $-D(X_0) = {}^tBV'$, où $\|V'\| = \|-BV\| = \|BV\| \leq 1$

et $\|BX_0\|V' - BX_0 = 0 - 0 = 0$, donc $-D(X_0) \in \mathcal{N}(X_0)$.

20. Soient u et v deux vecteurs de \mathbb{R}^n avec $u \neq 0$.

(a) En développant l'expression, et avec les formules d'algèbre bilinéaire :

$$\|u+v\|^2 - \left(\|u\| + \frac{\langle u, v \rangle}{\|u\|} \right)^2 = \|u\|^2 + 2\langle u, v \rangle + \|v\|^2 - \|u\|^2 - 2\langle u, v \rangle - \frac{\langle u, v \rangle^2}{\|u\|^2} = \frac{\|u\|^2\|v\|^2 - \langle u, v \rangle^2}{\|u\|^2}.$$

(b) L'inégalité de Cauchy-Schwarz assure que : $|\langle u, v \rangle| \leq \|u\| \cdot \|v\| \implies \langle u, v \rangle^2 \leq \|u\|^2 \|v\|^2$ par stricte croissance de la fonction carrée sur \mathbb{R}^+ .

Mais alors, $\|u\|^2 \|v\|^2 - \langle u, v \rangle^2 \geq 0 \iff \|u+v\|^2 - \left(\|u\| + \frac{\langle u, v \rangle}{\|u\|}\right)^2 \geq 0$, soit :

$$\|u+v\|^2 \geq \left(\|u\| + \frac{\langle u, v \rangle}{\|u\|}\right)^2 \implies \|u+v\| \geq \left\| \|u\| + \frac{\langle u, v \rangle}{\|u\|} \right\| \geq \|u\| + \frac{\langle u, v \rangle}{\|u\|} \implies \|u+v\| - \|u\| \geq \frac{\langle u, v \rangle}{\|u\|}. \quad (7)$$

21. Soient maintenant $X_0, X \in \mathcal{M}_{p,1}(\mathbb{R})$ et $W \in \mathcal{N}(X_0)$.

a) On distingue ici, comme plus haut, deux cas selon que BX_0 est nul ou non.

• Si $BX_0 \neq 0$: alors, en écrivant $BX = BX_0 + B(X - X_0)$, le résultat précédent donne

$$\|BX\| - \|BX_0\| \geq \frac{\langle BX_0, B(X - X_0) \rangle}{\|BX_0\|} = \left\langle {}^t B \frac{BX_0}{\|BX_0\|}, X - X_0 \right\rangle = \langle W, X - X_0 \rangle$$

car lorsque $BX_0 \neq 0$, l'équation $\|BX_0\|V - BX_0 = 0$ d'inconnue V , admet une unique solution $V = \frac{BX_0}{\|BX_0\|}$ et $\mathcal{N}(X_0)$ est alors réduit au seul vecteur $W = {}^t B \frac{BX_0}{\|BX_0\|}$.

• Si $BX_0 = 0$, alors $\|BX\| - \|BX_0\| = \|BX\|$, et puisque $W \in \mathcal{N}(X_0)$, alors il existe $V \in \mathcal{M}_{m,1}(\mathbb{R})$ tel que $\|V\| = 1$ et $W = {}^t BV$, de sorte que :

$$\langle W, X - X_0 \rangle = \langle {}^t BV, X - X_0 \rangle = \langle V, \underbrace{BX - BX_0}_{=0} \rangle = \langle V, BX \rangle \leq \|V\| \cdot \|BX\| \leq \|BX\|$$

encore d'après l'inégalité de Cauchy-Schwarz, et du fait que $\|V\| \leq 1$.

On a donc bien démontré que dans tous les cas de figure, $\|BX\| - \|BX_0\| \geq \langle W, X - X_0 \rangle$.

b) On a alors :

$$\begin{aligned} J(X) - J(X_0) &= J(X_0 + X - X_0) - J(X_0) \\ &= \frac{1}{2} (\|AX_0 - Y + A(X - X_0)\|^2 - \|AX_0 - Y\|^2) + \|BX\| - \|BX_0\| \\ &= \langle D(X_0), X - X_0 \rangle + \|BX\| - \|BX_0\| \\ &\geq \langle D(X_0), X - X_0 \rangle + \langle W, X - X_0 \rangle = \langle D(X_0) + W, X - X_0 \rangle. \end{aligned} \quad (8)$$

22. En étudiant deux cas de figures complémentaires, les questions 18. et 19. ont montré que si J admet un minimum global en X_0 , alors $-D(X_0) \in \mathcal{N}(X_0)$.

Réciproquement, si $-D(X_0) \in \mathcal{N}(X_0)$, alors on peut réécrire (8) avec $W = -D(X_0)$, et on obtient :

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

ce qui prouve que X_0 réalise un minimum global de J sur $\mathcal{M}_{p,1}(\mathbb{R})$.

On a donc bien obtenu l'équivalence demandée par double implication.

23. On se place dans toute la suite, dans le cas où $n = p = m$, $A = I_n$ et B la matrice diagonale définie par :

$$B_{i,i} = \begin{cases} \alpha_i & \text{si } 1 \leq i \leq k, \\ 0 & \text{si } k+1 \leq i \leq n \end{cases}$$

où k est un entier vérifiant $1 \leq k \leq n$ et $\alpha_1, \dots, \alpha_k$ sont des nombres réels tous non nuls. On pose

$$\rho(Y) = \sum_{i=1}^k \frac{y_i^2}{\alpha_i^2}.$$

On définit la fonction F de $[0; +\infty[$ dans \mathbb{R} par

$$\forall \lambda \in [0; +\infty[, \quad F(\lambda) = -1 + \sum_{i=1}^k \frac{\alpha_i^2 y_i^2}{(\lambda + \alpha_i^2)^2}.$$

(a) Soit $\lambda \in]0; +\infty[$. Pour tout $i \in \llbracket 1; k \rrbracket$:

$$\frac{1}{4\lambda} - \frac{\alpha_i^2}{(\lambda + \alpha_i^2)^2} = \frac{(\lambda + \alpha_i^2)^2 - 4\lambda\alpha_i^2}{4\lambda(\lambda + \alpha_i^2)^2} = \frac{\lambda^2 + 2\lambda\alpha_i^2 + \alpha_i^4 - 4\lambda\alpha_i^2}{4\lambda(\lambda + \alpha_i^2)^2} = \frac{(\lambda - \alpha_i^2)^2}{4\lambda(\lambda + \alpha_i^2)^2} \geq 0,$$

donc puisque $y_i^2 \geq 0$:

$$\forall i \in \llbracket 1; k \rrbracket, \quad \frac{\alpha_i^2 y_i^2}{(\lambda + \alpha_i^2)^2} \leq \frac{y_i^2}{4\lambda} \implies -1 + \sum_{i=1}^k \frac{\alpha_i^2 y_i^2}{(\lambda + \alpha_i^2)^2} = F(\lambda) \leq -1 + \frac{1}{4\lambda} \sum_{i=1}^k y_i^2.$$

(b) On suppose dans cette question uniquement que $\rho(Y) > 1$.

La fonction F de la variable λ , est bien définie et continue sur $[0; +\infty[$ comme somme de fonctions qui le sont, avec $(\lambda + \alpha_i^2)^2 > 0$ pour tout $\lambda \geq 0$, vu que $\alpha_i^2 > 0$.

La fonction F est également strictement décroissante sur $[0; +\infty[$: pour tous réels λ et μ de $[0; +\infty[$ tels que $0 \leq \lambda < \mu$,

$$\begin{aligned} \forall i \in \llbracket 1; k \rrbracket, \quad 0 < \lambda + \alpha_i^2 < \mu + \alpha_i^2 \stackrel{(*)}{\implies} 0 < (\lambda + \alpha_i^2)^2 < (\mu + \alpha_i^2)^2 \stackrel{(**)}{\implies} \frac{1}{(\lambda + \alpha_i^2)^2} > \frac{1}{(\mu + \alpha_i^2)^2} \\ \stackrel{(***)}{\implies} \sum_{i=1}^k \frac{\alpha_i^2 y_i^2}{(\lambda + \alpha_i^2)^2} > \sum_{i=1}^k \frac{\alpha_i^2 y_i^2}{(\mu + \alpha_i^2)^2} \implies F(\lambda) > F(\mu) \end{aligned}$$

(*) : par stricte croissance de la fonction carrée sur \mathbb{R}^+ .

(**) : par stricte décroissance de l'inverse sur \mathbb{R}_+^* . (***) : pour tout $i \in \llbracket 1; k \rrbracket$, $\alpha_i^2 y_i^2 \geq 0$ et le fait que $\rho(Y)$ assure que l'un au moins de ces termes est strictement positif (sinon on aurait $Y = 0$ et $\rho(Y) = 0$), donc l'inégalité stricte est conservée par passage à la somme.

On remarque aussi que $F(0) = -1 + \sum_{i=1}^k \frac{\alpha_i^2 y_i^2}{\alpha_i^4} = -1 + \rho(Y) > 0$ par hypothèse,

et $\lim_{\lambda \rightarrow +\infty} F(\lambda) = -1$ puisque $\lim_{\lambda \rightarrow +\infty} \frac{\alpha_i^2 y_i^2}{(\lambda + \alpha_i^2)^2} = 0$ pour tout $i \in \llbracket 1; k \rrbracket$.

La fonction F est ainsi continue, strictement décroissante sur $[0; +\infty[$: d'après le théorème éponyme, elle réalise donc une bijection de $[0; +\infty[$ dans son intervalle-image $] -1; \rho(Y) - 1[$ qui contient 0 puisque $\rho(Y) > 1$.

Il existe donc bien un unique réel $\lambda_0 \in [0; +\infty[$ tel que $F(\lambda_0) = 0$.

Puisque $F(0) \neq 0$, alors $\lambda_0 > 0$, et alors d'après (a) :

$$F(\lambda_0) = 0 \leq -1 + \frac{1}{4\lambda_0} \sum_{i=1}^k y_i^2 \iff 1 \leq \frac{1}{4\lambda_0} \sum_{i=1}^k y_i^2 \stackrel{\lambda_0 > 0}{\implies} \lambda_0 \leq \frac{1}{4} \sum_{i=1}^k y_i^2.$$

(c) On pose : $\beta = \begin{cases} 0 & \text{si } \rho(Y) \leq 1, \\ \lambda_0 & \text{si } \rho(Y) > 1. \end{cases}$ et $X_0 = Y - BV$, où $V \in \mathcal{M}_{n,1}(\mathbb{R})$ est défini par :

$$\forall i \in \llbracket 1; n \rrbracket, \quad v_i = \begin{cases} \frac{\alpha_i y_i}{\beta + \alpha_i^2} & \text{si } 1 \leq i \leq k, \\ 0 & \text{si } k+1 \leq i \leq n. \end{cases}$$

En se souvenant bien que dans cette question, $n = p = m$ et $A = I_n$, alors $-D(X_0) = -{}^t A A X_0 + {}^t A Y = Y - X_0 = B V = {}^t B V$ puisque B est diagonale, donc symétrique ici.

On vérifie alors que :

$$\|V\|^2 = \sum_{i=1}^n v_i^2 = \sum_{i=1}^k \frac{\alpha_i^2 y_i^2}{(\beta + \alpha_i^2)^2} = F(\beta) + 1 = \begin{cases} F(0) + 1 = \rho(Y) & \text{si } \rho(Y) \leq 1 \text{ car alors } \beta = 0 \\ F(\lambda_0) + 1 = 1 & \text{si } \rho(Y) > 1 \text{ car alors } \beta = \lambda_0 \end{cases}$$

Dans tous les cas, on a bien $\|V\| \leq 1$ (par positivité de la norme).

$$\text{Enfin : } B X_0 = B Y - B^2 V = \begin{pmatrix} \alpha_1 y_1 \\ \vdots \\ \alpha_k y_k \\ 0 \\ \vdots \\ 0 \end{pmatrix} - \begin{pmatrix} \frac{\alpha_1^3 y_1}{\beta + \alpha_1^2} \\ \vdots \\ \frac{\alpha_k^3 y_k}{\beta + \alpha_k^2} \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{\beta \alpha_1 y_1}{\beta + \alpha_1^2} \\ \vdots \\ \frac{\beta \alpha_k y_k}{\beta + \alpha_k^2} \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \beta V.$$

Ainsi, si $\rho(Y) \leq 1$ alors $\beta = 0$ donc $B X_0 = 0$ et $\|B X_0\|V - B X_0 = 0 - 0 = 0$ est vérifié.

sinon, $\rho(Y) > 1$ et alors $\beta = \lambda_0$, auquel cas on a vu que $\|V\| = 1$, et alors $\|B X_0\|V - B X_0 = \lambda_0 \|V\|V - \lambda_0 V = 0$.

On a donc vérifié ici que $-D(X_0)$ appartient à $\mathcal{N}(X_0)$ dans tous les cas, ce qui permet de conclure que X_0 réalise effectivement un minimum global de J .

(d) **Python.**

- i. Avec les définitions et les conventions de l'énoncé, la fonction ci-dessous prend en arguments les vecteurs alpha et Y, ainsi que le réel lda pour calculer $F(\lambda) = -1 + \sum_{i=1}^k \frac{\alpha_i^2 y_i^2}{(\lambda + \alpha_i^2)^2}$:

```
def FoncSom(alpha, Y, lda):
    S = 0
    k = np.size(alpha)
    for i in range(k):
        S = S + (alpha[i]*y[i])**2/(lda+alpha[i]**2)**2
    return -1+S
```

- ii. La fonction ci-dessous calcule une valeur approchée de l'unique solution de l'équation $F(\lambda) = 0$ sur $[0; +\infty[$ lorsque $\rho(Y) > 1$: on sait qu'elle est comprise entre 0 et $\frac{1}{4} \sum_{i=1}^k y_i^2$ qu'on calculera en premier lieu pour servir de borne de droite à l'intervalle de recherche initial.

```
def CalcBeta(alpha, Y, epsilon):
    k = np.size(alpha)
    S = 0
    for i in range(k):
        S = S + y[i]**2
    a = 0
    b = S/4
    while (b-a) > epsilon:
        c = (a+b)/2
        if F(c) * F(a) < 0:
            b = c
        else :
            a = c
    return (a+b)/2
```

iii. La fonction ci-dessous calcule donc, toujours dans le cas où $\rho(Y) > 1$, la solution

$$X_0 = Y - BV = \begin{pmatrix} y_1 \\ \vdots \\ y_k \\ y_{k+1} \\ \vdots \\ y_n \end{pmatrix} - \begin{pmatrix} \frac{\alpha_1^2 y_1}{\beta + \alpha_1^2} \\ \vdots \\ \frac{\alpha_k^2 y_k}{\beta + \alpha_k^2} \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} y_1 \left(1 - \frac{\alpha_1^2}{\beta + \alpha_1^2}\right) \\ \vdots \\ y_k \left(1 - \frac{\alpha_k^2}{\beta + \alpha_k^2}\right) \\ y_{k+1} \\ \vdots \\ y_n \end{pmatrix} \quad \text{où on donne à } \beta \text{ la valeur fournie}$$

par l'algorithme de dichotomie :

```
def CalcSolution(alpha, Y, epsilon):
    k = np.size(alpha)
    X = Y
    beta = CalcBeta(alpha, Y, epsilon)
    for i in range(k):
        X[i] = X[i]*(1-\alpha[i]**2/(beta+alpha[i]**2))
    return X
```

★★★ FIN DU SUJET ★★★