



PRÉPASIMPLE

prepasimple.fr

HEC / ESSEC

Mathématiques

2025

Corrigé de mathématiques approfondies

Mathématiques approfondies — Prépa ECG

Annales de concours · Corrigé

Corrigé

On conserve les notations de l'énoncé. Les espaces sont réels et euclidiens. Les propriétés de l'adjoint sont démontrées dans la partie I; elles ne sont pas supposées connues. On écrit tA pour la transposée de A .

Partie I : autour de l'adjoint

1. On utilise les coordonnées dans la base orthonormale $\mathcal{B}_E$. Posons

$$a_0 = \sum_{i=1}^p \ell(e_i) e_i.$$

Pour

$$x = \sum_{i=1}^p x_i e_i,$$

la linéarité de ℓ et l'orthonormalité donnent

$$\ell(x) = \sum_{i=1}^p x_i \ell(e_i) = \langle a_0, x \rangle_E.$$

Si a_0 et b_0 conviennent, alors $\langle a_0 - b_0, x \rangle_E = 0$ pour tout $x \in E$. En choisissant $x = a_0 - b_0$, on obtient $\|a_0 - b_0\|^2 = 0$, donc $a_0 = b_0$. L'existence et l'unicité sont établies.

2. Pour $y \in F$ fixé, l'application $\ell_y : x \mapsto \langle u(x), y \rangle_F$ est linéaire, par linéarité de u et bilinéarité du produit scalaire. La question 1 s'applique et fournit un unique z_y tel que

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

En particulier,

$$z_y = \sum_{i=1}^p \langle u(e_i), y \rangle_F e_i.$$

3. Soient $y, y' \in F$ et $a, b \in \mathbb{R}$. Pour tout $x \in E$,

$$\langle u(x), ay + by' \rangle_F = a \langle z_y, x \rangle_E + b \langle z_{y'}, x \rangle_E = \langle az_y + bz_{y'}, x \rangle_E.$$

Par l'unicité établie en 2, $z_{ay+by'} = az_y + bz_{y'}$. Ainsi u^* est linéaire. La symétrie du produit scalaire donne l'identité (1).

4. Le coefficient de ligne i et colonne j de la matrice de u^* vaut

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

Sa matrice est donc tA . Le rang d'une matrice est égal à celui de sa transposée, d'où $\text{rg}(u^*) = \text{rg}(u)$.

La matrice de $(u^*)^*$ est ${}^t({}^tA) = A$. Deux applications linéaires de mêmes espaces de départ et d'arrivée ayant la même matrice dans les mêmes bases sont égales. Donc $(u^*)^* = u$.

5. Méthode 1 — Inclusion et dimensions

Si $z = u^*(y)$ et $x \in \text{Ker}(u)$, alors $\langle x, z \rangle_E = \langle u(x), y \rangle_F = 0$. Ainsi $\text{Im}(u^*) \subset \text{Ker}(u)^\perp$.

D'autre part,

$$\dim \text{Im}(u^*) = \text{rg}(u) = p - \dim \text{Ker}(u) = \dim \text{Ker}(u)^\perp.$$

L'inclusion entre ces sous-espaces de même dimension est donc une égalité.

Méthode 2 — Caractérisation de l'orthogonal

Un vecteur $x \in E$ appartient à $\text{Im}(u^*)^\perp$ si et seulement si $\langle x, u^*(y) \rangle_E = 0$ pour tout $y \in F$, soit $\langle u(x), y \rangle_F = 0$ pour tout y . En prenant $y = u(x)$, cette condition équivaut à $u(x) = 0$.

Ainsi $\text{Im}(u^*)^\perp = \text{Ker}(u)$. En dimension finie, le supplémentaire orthogonal du supplémentaire orthogonal d'un sous-espace est ce sous-espace. En prenant les orthogonaux, on obtient la même égalité. La première méthode convient lorsque les dimensions sont disponibles ; la seconde exploite directement l'identité définissant l'adjoint.

6. L'inclusion $\text{Ker}(u) \subset \text{Ker}(u^* \circ u)$ résulte de la composition. Réciproquement, si $u^*u(x) = 0$,

$$\|u(x)\|_F^2 = \langle u(x), u(x) \rangle_F = \langle x, u^*u(x) \rangle_E = 0.$$

Donc $u(x) = 0$, et les noyaux sont égaux. Le théorème du rang donne alors $\text{rg}(u^*u) = \text{rg}(u) = \text{rg}(u^*)$.

Or $\text{Im}(u^*u) \subset \text{Im}(u^*)$. L'égalité des dimensions entraîne $\text{Im}(u^*u) = \text{Im}(u^*)$.

Ces résultats restent vrais pour toute matrice réelle rectangulaire, indépendamment de la comparaison entre ses dimensions.

7. La question 6 assure que w prend ses valeurs dans $\text{Im}(u^*)$, et w est linéaire. Son noyau est

$$\text{Ker}(w) = \text{Im}(u^*) \cap \text{Ker}(u^*u) = \text{Ker}(u)^\perp \cap \text{Ker}(u) = \{0\}.$$

Un endomorphisme injectif d'un espace de dimension finie est bijectif ; w est un isomorphisme. Ceci inclut le cas de l'espace nul, lorsque $u = 0$.

En coordonnées, pour tout $C \in \text{Im}({}^tA)$, l'équation $MX = C$ admet une unique solution dans $\text{Im}({}^tA) = \text{Ker}(A)^\perp$.

En particulier, pour tout $Y \in \mathbb{R}^n$, l'équation $MX = {}^tAY$ possède une unique solution orthogonale à $\text{Ker}(A)$. Les solutions dans l'espace entier peuvent, en revanche, être multiples.

8. (a) Puisque π est la projection orthogonale sur $\text{Im}(u)$, pour tout $y \in F$,
 $y - \pi(y) \in \text{Im}(u)^\perp$.

Pour tout $x \in E$,

$$\langle x, u^*(y - \pi(y)) \rangle_E = \langle u(x), y - \pi(y) \rangle_F = 0.$$

Donc $u^*(y - \pi(y)) = 0$ et $u^*\pi = u^*$. En matrices, ${}^tAQ = {}^tA$.

8. (b) Posons $r = \text{rg}(u)$. Une base de $\text{Im}(u)$ complétée par une base de son orthogonal donne une base de F dans laquelle la matrice de π est diagonale, avec r coefficients égaux à 1 et $n - r$ coefficients égaux à 0.

La trace est invariante par changement de base. Ainsi $\text{Tr}(Q) = r = \text{rg}(u)$, sans recours à un résultat supplémentaire sur la trace.

9. Méthode 1 — Noyau et rang

La question 6 donne $\text{Ker}(M) = \text{Ker}(A)$. La matrice carrée M est inversible si et seulement si son noyau est nul ; par le théorème du rang, $\text{Ker}(A) = \{0\}$ équivaut à $\text{rg}(A) = p$. D'où l'équivalence demandée.

Méthode 2 — Identité de normes

Pour tout X , ${}^tXMX = \|AX\|^2$. Si A est de rang p , $MX = 0$ impose $\|AX\|^2 = 0$, puis $AX = 0$ et $X = 0$. Ainsi M est inversible.

Réciproquement, si M est inversible et $AX = 0$, alors $MX = {}^tAAX = 0$, donc $X = 0$. Le rang de A est p . La première méthode réutilise les résultats abstraits ; la seconde donne un argument matriciel autonome.

10. (a) Méthode 1 — Caractérisation de la projection

La question 9 permet de définir $R = AM^{-1}A$. Pour tout Y , $RY \in \text{Im}(A)$, et

$${}^tA(Y - RY) = {}^tAY - MM^{-1}{}^tAY = 0.$$

Ainsi $Y - RY \in \text{Ker}({}^tA) = \text{Im}(A)^\perp$; cette dernière égalité résulte de la question 5 appliquée à l'adjoint de u . La décomposition orthogonale de Y impose $RY = QY$. Donc $Q = AM^{-1}A$.

Méthode 2 — Projecteur symétrique

La matrice M est symétrique; son inverse est symétrique car ${}^t(M^{-1}) = ({}^tM)^{-1} = M^{-1}$. Ainsi ${}^tR = R$. De plus,

$$R^2 = AM^{-1}({}^tAA)M^{-1}A = AM^{-1}A = R.$$

Enfin, $\text{Im}(R) \subset \text{Im}(A)$ et $RA = A$, donc $\text{Im}(R) = \text{Im}(A)$. Un projecteur symétrique est le projecteur orthogonal sur son image : $R = Q$. Cette seconde méthode permet de reconnaître une matrice de projection par ses propriétés algébriques.

10. (b) On teste le rang avant de calculer l'inverse.

Python

```
import numpy as np
import numpy.linalg as al

def Calcule_Q(A):
    n, p = np.shape(A)
    if al.matrix_rank(A) == p:
        At = np.transpose(A)
        M = np.dot(At, A)
        return np.dot(np.dot(A, al.inv(M)), At)
    else:
        return "Erreur : A n'est pas de rang p."
```

Les produits sont matriciels. Dans la branche où l'inverse est calculé, son existence est justifiée par la question 9. `matrix_rank` évalue numériquement le rang ; pour des matrices très proches d'une matrice de rang inférieur, ce test dépend de la précision des calculs flottants. Le code a été exécuté sur des matrices de rang plein et sur une matrice de rang inférieur.

11. En associant les produits,

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

L'égalité a lieu exactement lorsque $X \in \text{Ker}(A)$.

Partie II : minimisation d'une fonction quadratique

12. On développe le carré de la norme :

$$\|A(X + H) - Y\|^2 = \|AX - Y\|^2 + 2\langle AX - Y, AH \rangle + \|AH\|^2.$$

Or $\langle AX - Y, AH \rangle = \langle {}^tA(AX - Y), H \rangle = \langle D(X), H \rangle$ et $\|AH\|^2 = {}^tHMH$. En divisant par 2, on obtient exactement (2).

13. Méthode 1 — Accroissements dans une direction

Si $D(X) = 0$, (2) et la question 11 donnent $J_0(X + H) - J_0(X) = \frac{1}{2}\|AH\|^2 \geq 0$ pour tout H . Donc X réalise un minimum global.

Réciproquement, supposons X minimiseur. Pour tout H fixé, la fonction $g_H(t) = J_0(X + tH)$ est un polynôme en t et possède un minimum en 0. Sa dérivée en 0 est nulle. Par (2), $g'_H(0) = \langle D(X), H \rangle$. En prenant $H = D(X)$, on obtient $\|D(X)\|^2 = 0$, donc $D(X) = 0$.

Méthode 2 — Projection orthogonale

Pour tout X , les vecteurs $AX - QY \in \text{Im}(A)$ et $QY - Y \in \text{Im}(A)^\perp$ sont orthogonaux. Par

Pythagore,

$$J_0(X) = \frac{1}{2} \|AX - QY\|^2 + \frac{1}{2} \|QY - Y\|^2.$$

Comme $QY \in \text{Im}(A)$, la valeur minimale est atteinte exactement lorsque $AX = QY$. Si cette égalité est satisfaite, 8(a) donne ${}^tA(AX - Y) = 0$. Réciproquement, si cette dernière égalité est satisfaite, $AX - Y$ est orthogonal à $\text{Im}(A)$; l'unicité de la décomposition orthogonale impose $AX = QY$. Ainsi la condition équivaut à $D(X) = 0$.

La première méthode est réutilisable pour les conditions d'optimalité; la seconde explique géométriquement les moindres carrés.

14. Par la question 7, $MX = {}^tAY$ admet une unique solution X_0 dans $\text{Im}({}^tA) = \text{Ker}(A)^\perp$.

Cette équation équivaut à $D(X) = 0$. Donc $\mathcal{S}_0 \cap \text{Ker}(A)^\perp = \{X_0\}$. En particulier, un minimum global existe même si A n'est pas de rang p .

15. (a) Puisque $X \in \mathcal{S}_0$, ${}^tA(AX - Y) = 0$. Pour tout $V \in \mathbb{R}^p$,

$$\langle AV, AX - Y \rangle = {}^tV {}^tA(AX - Y) = 0.$$

Donc $Y - AX \in \text{Im}(A)^\perp$, tandis que $AX \in \text{Im}(A)$. Par caractérisation de la projection orthogonale, $AX = QY$.

15. (b) Les vecteurs X et X_0 vérifient la même équation normale, donc $M(X - X_0) = 0$. Comme $\text{Ker}(M) = \text{Ker}(A)$, $X - X_0 \in \text{Ker}(A)$.

Le vecteur X_0 est orthogonal à $\text{Ker}(A)$; ainsi

$$\|X\|^2 = \|X_0 + (X - X_0)\|^2 = \|X_0\|^2 + \|X - X_0\|^2.$$

Si $X \neq X_0$, le dernier terme est strictement positif, d'où $\|X\| > \|X_0\|$. En fait, $\mathcal{S}_0 = X_0 + \text{Ker}(A)$: tout vecteur de cette forme satisfait l'équation normale. X_0 est l'unique minimiseur de norme minimale.

15. (c) Si $\text{rg}(A) = p$, la question 9 donne l'inversibilité de M . L'équation $MX = {}^tAY$ a alors une unique solution,

$$X = M^{-1} {}^t A Y.$$

$$\text{Ainsi } \mathcal{S}_0 = \{M^{-1} {}^t A Y\}.$$

16. (a) On utilise $Y = AU_0 + Z$ et la formule (4) :

$$X = U_0 + M^{-1} {}^t A Z,$$

$$\text{puis } A(X - U_0) = AM^{-1} {}^t A Z = QZ. \text{ Donc } T = \|QZ\|^2.$$

La matrice de la projection orthogonale satisfait ${}^t Q = Q$ et $Q^2 = Q$. Ainsi

$$T = {}^t Z {}^t Q Q Z = {}^t Z Q Z.$$

16. (b) On simule des normales indépendantes d'écart-type σ , et non de variance σ^2 .

Python

```
import numpy as np

import numpy.random as rd
import numpy.linalg as al

def simuleT(A, sigma):
    n, p = np.shape(A)
    At = np.transpose(A)
    Q = np.dot(np.dot(A, al.inv(np.dot(At, A))), At)
    Z = rd.normal(0.0, sigma, n)
    QZ = np.dot(Q, Z)
    return np.dot(QZ, QZ)
```

Le vecteur Z est ici monodimensionnel, ce qui représente les mêmes coordonnées qu'un vecteur colonne mathématique. $\text{np.dot}(QZ, QZ)$ est son carré de norme. L'hypothèse de rang plein justifie l'inverse. Le code a été exécuté.

16. (c) On effectue une moyenne de réalisations indépendantes. L'existence de l'espérance et de la variance sera établie en (e)–(i), ce qui permettra d'appliquer la

loi faible des grands nombres.

Python

```
import numpy as np

import numpy.random as rd

import numpy.linalg as al

def esperance(A, sigma):
    n, p = np.shape(A)
    At = np.transpose(A)
    Q = np.dot(np.dot(A, al.inv(np.dot(At, A))), At)
    N = 10000
    total = 0.0
    for r in range(N):
        Z = rd.normal(0.0, sigma, n)
        QZ = np.dot(Q, Z)
        total = total + np.dot(QZ, QZ)
    return total / N
```

La projection est calculée une seule fois. Le nombre $N = 10000$ est un choix numérique, non une donnée du sujet. Cette fonction fournit une approximation Monte-Carlo ; elle ne démontre aucune formule exacte. Le code a été exécuté.

Le programme source utilise le nom `esperanceT` alors que la question (c) demande `esperance`. La version transcrite utilise partout `esperance` pour rendre le programme exécutable.

- 16.** (d) Les trois matrices ont respectivement 2, 3 et 5 colonnes indépendantes. En effet, leurs premières lignes, en nombre égal au nombre de colonnes, forment une matrice triangulaire à diagonale égale à 1, donc inversible.

Pour $\sigma = 1$, les valeurs 1,99, 3,01, 4,97 suggèrent $E(T) = p$.

Le graphique pour A_1 , qui a $p = 2$, représente approximativement les points (1; 2,1), (2; 8,0), (3; 17,9), (4; 31,4), (5; 50,5), (6; 71,0). Ces valeurs sont lues par mesure sur le scan et ne sont pas les valeurs théoriques exactes.

Elles sont proches de $2\sigma^2$; la conjecture générale est donc $E(T) = p\sigma^2$. Les fluctuations observées viennent de la simulation. L'exemple A_1 utilise $n = 3$, contrairement à la convention générale $n \geq 4$; cette exception locale est explicitée dans le sujet, sans changer la matrice. Les résultats démontrés ici sont valables aussi pour cette dimension.

16. (e) L'hypothèse $\sigma > 0$ précise le domaine du paramètre de la loi normale non dégénérée. Posons $G = z_1/\sigma$; G suit $\mathcal{N}(0, 1)$, de densité

$$\varphi(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}.$$

Les intégrales des valeurs absolues des puissances jusqu'à l'ordre 4 convergent : pour $|t|$ assez grand, $|t|^r e^{-t^2/2} \leq e^{-t^2/4} \leq e^{-|t|/4}$, pour chaque $r \leq 4$. Les moments utilisés existent donc.

Par symétrie de la densité, $E(G^3) = 0$. La moyenne est nulle et la variance vaut 1, donc $E(G^2) = 1$.

Pour le quatrième moment, $\varphi'(t) = -t\varphi(t)$. Sur le segment $[-R, R]$, une intégration par parties donne

$$\int_{-R}^R t^4 \varphi(t) dt = -[t^3 \varphi(t)]_{-R}^R + 3 \int_{-R}^R t^2 \varphi(t) dt.$$

Le terme de bord tend vers 0 par croissance comparée, et les intégrales convergent. En passant à la limite, $E(G^4) = 3E(G^2) = 3$. Par transfert,

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

Les mêmes résultats valent pour toutes les composantes. Plus généralement, tout moment absolu entier d'une normale existe par la même comparaison. Les produits de moments utilisés ci-dessous sont donc intégrables ; l'indépendance permet de factoriser leurs espérances. Les sommes étant finies, aucun échange avec une limite n'est nécessaire.

16. (f) En développant la forme quadratique et en utilisant $Q_{i,j} = Q_{j,i}$,

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

Pour $i \neq j$, l'indépendance et le centrage donnent $E(z_i z_j) = E(z_i)E(z_j) = 0$. Ainsi $E(T_2) = 0$, et

$$E(T) = \sigma^2 \sum_{i=1}^n Q_{i,i} = \sigma^2 \operatorname{Tr}(Q) = p\sigma^2,$$

car $\operatorname{rg}(A) = p$ et la question 8(b) s'applique. Les deux conjectures sont donc démontrées.

16. (g) En développant,

$$E(T_1 T_2) = \sum_{a=1}^n \sum_{i < j} Q_{a,a} Q_{i,j} E(z_a^2 z_i z_j).$$

Si a est distinct de i, j , l'espérance se factorise en $E(z_a^2)E(z_i)E(z_j) = 0$.

Si $a = i$, elle vaut $E(z_i^3)E(z_j) = 0$. Si $a = j$, elle vaut $E(z_i)E(z_j^3) = 0$. Ces trois cas couvrent tous les indices puisque $i < j$. Tous les termes sont nuls, donc $E(T_1 T_2) = 0$.

16. (h) Développons d'abord T_1^2 . Pour les indices égaux, le quatrième moment vaut $3\sigma^4$; pour les indices distincts, l'indépendance donne $E(z_i^2 z_j^2) = \sigma^4$. D'où

$$E(T_1^2) = 3\sigma^4 \sum_i Q_{i,i}^2 + 2\sigma^4 \sum_{i < j} Q_{i,i} Q_{j,j}.$$

Comme

$$\left(\sum_i Q_{i,i} \right)^2 = q + 2 \sum_{i < j} Q_{i,i} Q_{j,j} = p^2,$$

$$E(T_1^2) = \sigma^4(2q + p^2).$$

Dans le développement de T_2^2 , un terme correspond à $E(z_i z_j z_a z_b)$ avec $i < j$ et $a < b$. Si les paires sont distinctes, soit elles sont disjointes, soit elles ont un unique indice commun. Dans les deux cas, au moins un indice apparaît une seule fois ; l'indépendance et le centrage annulent l'espérance. Si les paires sont identiques, l'espérance vaut $E(z_i^2 z_j^2) = \sigma^4$. Donc

$$E(T_2^2) = \sigma^4 \sum_{i < j} Q_{i,j}^2.$$

Enfin, $Q^2 = Q$ et Q est symétrique. Le coefficient diagonal d'indice i de cette identité donne

$$\sum_j Q_{i,j}^2 = Q_{i,i}.$$

En sommant sur i ,

$$q + 2 \sum_{i < j} Q_{i,j}^2 = \text{Tr}(Q) = p.$$

Ainsi $E(T_2^2) = \frac{\sigma^4}{2}(p - q)$, comme demandé.

16. (i) Les questions précédentes donnent

$$E(T^2) = E(T_1^2) + 4E(T_1 T_2) + 4E(T_2^2) = \sigma^4(p^2 + 2p).$$

L'existence du second moment est établie. Par la formule de Huygens,

$$V(T) = E(T^2) - E(T)^2 = 2p\sigma^4.$$

Pour $\alpha > 0$, $p\sigma^2 > 0$, et

$$\{T > (1 + \alpha)p\sigma^2\} \subset \{|T - E(T)| > \alpha p\sigma^2\}.$$

L'inégalité de Bienaymé-Tchebychev entraîne

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

Si ce majorant dépasse 1, l'inégalité reste vraie. Aucune identification à une loi extérieure au référentiel n'a été nécessaire.

Partie III : minimisation d'une fonction non différentiable

17. Méthode 1 — Quantité conjuguée réelle

Pour $t \neq 0$,

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

Le dénominateur est strictement positif puisque $u \neq 0$. La continuité de la norme, qui résulte de $\|a\| - \|b\| \leq \|a - b\|$, permet de passer à la limite. Le résultat vaut $\langle u, v \rangle / \|u\|$.

Méthode 2 — Dérivée d'une fonction d'une variable

La fonction $h(t) = \|u + tv\|^2 = \|u\|^2 + 2t\langle u, v \rangle + t^2\|v\|^2$ est polynomiale et $h(0) > 0$. Elle reste strictement positive sur un voisinage de 0. La fonction $g(t) = \sqrt{h(t)}$ y est dérivable, avec

$$g'(0) = \frac{h'(0)}{2\sqrt{h(0)}} = \frac{\langle u, v \rangle}{\|u\|}.$$

La limite demandée est la définition de cette dérivée. La première preuve ne nécessite que des opérations sur les limites ; la seconde reconnaît une composition de fonctions dérivables. Ces deux arguments sont valables en toute dimension euclidienne, notamment dans $\mathbb{R}^m$ pour les questions suivantes.

18. Fixons $H \in \mathbb{R}^p$. La fonction $g_H(t) = J(X_0 + tH)$ possède un minimum en 0. Puisque

$BX_0 \neq 0$, la question 17, appliquée dans $\mathbb{R}^m$, montre qu'elle est dérivable en 0 et que

$$g'_H(0) = \langle D(X_0), H \rangle + \frac{\langle BX_0, BH \rangle}{\|BX_0\|} = \left\langle D(X_0) + {}^tB \frac{BX_0}{\|BX_0\|}, H \right\rangle.$$

Cette dérivée est nulle pour tout H . En prenant H égal au vecteur entre accolades dans le produit scalaire, on en déduit

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

Posons $V = BX_0 / \|BX_0\|$. Alors $\|V\| = 1$ et $\|BX_0\|V - BX_0 = 0$, donc $-D(X_0) \in \mathcal{N}(X_0)$.

- 19.** (a) Question adaptée au programme actuel : la dimension de H est rectifiée en p au lieu de n , pour que $X_0 + tH$ et BH soient définis.

Puisque $BX_0 = 0$, pour tout $t \in \mathbb{R}$,

$$J(X_0 + tH) - J(X_0) = t\langle D(X_0), H \rangle + \frac{t^2}{2}\|AH\|^2 + |t|\|BH\|.$$

Pour $t < 0$, le minimum impose que cette expression soit positive ou nulle. En divisant par $-t > 0$,

$$-\langle D(X_0), H \rangle + \frac{-t}{2}\|AH\|^2 + \|BH\| \geq 0.$$

En faisant tendre t vers 0 par valeurs négatives, on obtient $\langle D(X_0), H \rangle \leq \|BH\|$.

En remplaçant H par $-H$, on obtient aussi $-\langle D(X_0), H \rangle \leq \|BH\|$, donc en fait $|\langle D(X_0), H \rangle| \leq \|BH\|$.

- 19.** (b) Pour $H \in \text{Ker}(B)$, le majorant précédent est nul. Donc $\langle D(X_0), H \rangle = 0$ pour tout $H \in \text{Ker}(B)$. Ainsi $D(X_0) \in \text{Ker}(B)^\perp$.

- 19.** (c) La question 5 s'applique à l'application de matrice B et donne $\text{Ker}(B)^\perp = \text{Im}({}^tB)$. Il existe donc $W \in \mathbb{R}^m$ tel que $D(X_0) = {}^tBW$.

19. (d) La question 6, appliquée à B , donne $\text{Im}({}^tBB) = \text{Im}({}^tB)$. Le vecteur $D(X_0)$ appartient à cette image, donc il existe $V \in \mathbb{R}^p$ tel que $D(X_0) = {}^tBBV$.

L'application de ces résultats à B ne nécessite pas que $m \geq p$: leurs démonstrations ne font intervenir aucune telle hypothèse.

19. (e) Appliquons (a) à $H = V$:

$$\|BV\|^2 = \langle {}^tBBV, V \rangle = \langle D(X_0), V \rangle \leq \|BV\|.$$

Si $\|BV\| = 0$, l'inégalité $\|BV\| \leq 1$ est satisfaite. Sinon, on divise par $\|BV\| > 0$ et on obtient la même conclusion.

Posons $U = -BV \in \mathbb{R}^m$. Alors $\|U\| \leq 1$, ${}^tBU = -D(X_0)$ et, comme $BX_0 = 0$, $\|BX_0\|U - BX_0 = 0$. Ainsi $-D(X_0) \in \mathcal{N}(X_0)$.

Il faut distinguer le vecteur $V \in \mathbb{R}^p$ construit en (d) du vecteur $U \in \mathbb{R}^m$ utilisé dans la définition de $\mathcal{N}(X_0)$.

20. (a) En développant séparément les deux carrés,

$$\|u + v\|^2 = \|u\|^2 + 2\langle u, v \rangle + \|v\|^2,$$

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

La soustraction donne exactement l'identité demandée. Les divisions sont licites car $u \neq 0$.

20. (b) Méthode 1 — Exploitation de l'identité précédente

Cauchy-Schwarz donne $\langle u, v \rangle^2 \leq \|u\|^2 \|v\|^2$. Le membre de droite de (a) est donc positif ou nul. Posons $c = \|u\| + \langle u, v \rangle / \|u\|$.

Si $c \leq 0$, on a directement $\|u + v\| \geq 0 \geq c$. Si $c > 0$, (a) donne $\|u + v\|^2 \geq c^2$, d'où $\|u + v\| \geq c$ puisque les deux membres sont positifs. Dans les deux cas,

$$\|u + v\| - \|u\| \geq \frac{\langle u, v \rangle}{\|u\|}.$$

Méthode 2 — Produit scalaire avec un vecteur unitaire

Par Cauchy-Schwarz,

$$\left\langle \frac{u}{\|u\|}, u + v \right\rangle \leq \|u + v\|.$$

Le membre de gauche vaut $\|u\| + \langle u, v \rangle / \|u\|$, ce qui donne immédiatement le résultat. La première méthode suit la démarche imposée par «en déduire» et contrôle le signe avant de prendre les racines ; la seconde montre que l'inégalité exprime la domination d'une projection scalaire par une norme.

- 21.** (a) Par définition de $\mathcal{N}(X_0)$, il existe $V \in \mathbb{R}^m$ tel que $W = {}^tBV$, $\|V\| \leq 1$ et $\|BX_0\|V = BX_0$.

Si $BX_0 \neq 0$, alors $V = BX_0 / \|BX_0\|$. L'inégalité (7), en dimension m , avec $u = BX_0$ et $v = B(X - X_0)$, donne

$$\|BX\| - \|BX_0\| \geq \langle V, B(X - X_0) \rangle = \langle W, X - X_0 \rangle.$$

Si $BX_0 = 0$, Cauchy-Schwarz et $\|V\| \leq 1$ donnent

$$\langle W, X - X_0 \rangle = \langle V, BX \rangle \leq \|V\| \|BX\| \leq \|BX\|.$$

Le résultat est donc vrai dans tous les cas.

- 21.** (b) Appliquons (2) avec $H = X - X_0$ et ajoutons le résultat de (a) :

$$J(X) - J(X_0) \geq \langle D(X_0), X - X_0 \rangle + \frac{1}{2} \|A(X - X_0)\|^2 + \langle W, X - X_0 \rangle.$$

En supprimant le terme de carré de norme positif ou nul, on obtient

$$J(X) - J(X_0) \geq \langle D(X_0) + W, X - X_0 \rangle.$$

- 22.** Si X_0 réalise un minimum global, la question 18 donne $-D(X_0) \in \mathcal{N}(X_0)$ lorsque $BX_0 \neq 0$, et la question 19 donne cette inclusion lorsque $BX_0 = 0$.

Réciproquement, si $-D(X_0) \in \mathcal{N}(X_0)$, on choisit $W = -D(X_0)$ dans (8). Pour tout X , on obtient $J(X) - J(X_0) \geq 0$. Ainsi X_0 réalise un minimum global.

La condition est donc nécessaire et suffisante. La preuve ne suppose pas que J soit différentiable aux points où $BX_0 = 0$.

- 23.** (a) Pour $\lambda > 0$ et $1 \leq i \leq k$,

$$(\lambda + \alpha_i^2)^2 - 4\lambda\alpha_i^2 = (\lambda - \alpha_i^2)^2 \geq 0.$$

Les dénominateurs sont strictement positifs. Il en résulte

$$\frac{\alpha_i^2 y_i^2}{(\lambda + \alpha_i^2)^2} \leq \frac{y_i^2}{4\lambda}.$$

En sommant sur i puis en retranchant 1, on obtient

$$F(\lambda) \leq -1 + \frac{1}{4\lambda} \sum_{i=1}^k y_i^2.$$

- 23.** (b) Les réels α_i sont non nuls, donc F est continue sur $[0, +\infty[$ et dérivable sur $]0, +\infty[$. On a

$$F(0) = -1 + \rho(Y) > 0, \quad \lim_{\lambda \rightarrow +\infty} F(\lambda) = -1.$$

Par le théorème des valeurs intermédiaires, F s'annule au moins une fois dans $]0, +\infty[$. De plus,

$$F'(\lambda) = -2 \sum_{i=1}^k \frac{\alpha_i^2 y_i^2}{(\lambda + \alpha_i^2)^3}.$$

Comme $\rho(Y) > 1$, au moins un des y_i , $i \leq k$, est non nul. Tous les dénominateurs sont positifs, donc $F'(\lambda) < 0$. La fonction est strictement décroissante et la racine λ_0 est unique.

En appliquant (a) à $\lambda_0 > 0$,

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

En multipliant par $\lambda_0 > 0$, on obtient

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

23. (c) Ici $M = I_n$, $D(X) = X - Y$ et ${}^tB = B$. Puisque $X_0 = Y - BV$, on a $-D(X_0) = BV = {}^tBV$. Pour appliquer 22, il reste à vérifier $\|V\| \leq 1$ et $\|BX_0\|V = BX_0$.

Si $\rho(Y) \leq 1$, $\beta = 0$. Pour $i \leq k$, $v_i = y_i/\alpha_i$ et $x_{0,i} = y_i - \alpha_i v_i = 0$. Pour $i > k$, $v_i = 0$ et $x_{0,i} = y_i$. Donc $BX_0 = 0$ et $\|V\|^2 = \rho(Y) \leq 1$. Les deux conditions sont satisfaites, y compris lorsque $\rho(Y) = 1$.

Si $\rho(Y) > 1$, $\beta = \lambda_0 > 0$ et $F(\beta) = 0$. Ainsi

$$\|V\|^2 = \sum_{i=1}^k \frac{\alpha_i^2 y_i^2}{(\beta + \alpha_i^2)^2} = 1.$$

Pour $i \leq k$,

$$x_{0,i} = y_i - \frac{\alpha_i^2 y_i}{\beta + \alpha_i^2} = \frac{\beta y_i}{\beta + \alpha_i^2},$$

puis $(BX_0)_i = \beta \alpha_i y_i / (\beta + \alpha_i^2) = \beta v_i$. Pour $i > k$, les deux membres sont nuls. Donc $BX_0 = \beta V$ et $\|BX_0\| = \beta \|V\| = \beta$. D'où $\|BX_0\|V = BX_0$.

Dans les deux cas, $-D(X_0) \in \mathcal{N}(X_0)$; la question 22 montre que X_0 réalise un minimum global de J .

Ce minimum est même unique : l'inégalité plus précise établie avant (8), avec $A = I_n$ et

$W = -D(X_0)$, donne $J(X) - J(X_0) \geq \frac{1}{2}\|X - X_0\|^2$. Si $X \neq X_0$, le membre de droite est strictement positif.

23. (d)(i) On calcule les k termes de la somme, avec des indices Python allant de 0 à $k - 1$. Les entrées satisfont les hypothèses du sujet, notamment $\alpha[i] \neq 0$ et $\text{lda} \geq 0$.

Python

```
import numpy as np

def FoncSom(alpha, Y, lda):
    k = np.shape(alpha)[0]
    valeur = -1.0
    for i in range(k):
        a2 = alpha[i] ** 2
        valeur = valeur + a2 * Y[i] ** 2 / (lda + a2) ** 2
    return valeur
```

Les dénominateurs sont positifs. La fonction renvoie un scalaire représentant $F(\lambda)$. Code exécuté.

23. (d)(ii) Comme $F(0) = \rho(Y) - 1$, on peut décider directement si $\beta = 0$. Sinon, posons

$$L = \frac{1}{4} \sum_{i=1}^k y_i^2 > 0;$$

(a) donne $F(L) \leq 0$, et $F(0) > 0$. La racine est donc dans $[0, L]$.

Python

```
import numpy as np

def CalcBeta(alpha, Y, epsilon):
    if FoncSom(alpha, Y, 0.0) <= 0.0:
        return 0.0
```

```

k = np.shape(alpha)[0]
gauche = 0.0
droite = 0.0
for i in range(k):
    droite = droite + Y[i] ** 2 / 4.0
while droite - gauche > 2.0 * epsilon:
    milieu = (gauche + droite) / 2.0
    if FoncSom(alpha, Y, milieu) > 0.0:
        gauche = milieu
    else:
        droite = milieu
return (gauche + droite) / 2.0

```

L'invariant est $\lambda_0 \in [\text{gauche}, \text{droite}]$. Si F est positif au milieu, la stricte décroissance place la racine à droite ; sinon elle est à gauche ou au milieu. À chaque étape, la longueur est divisée par 2. Pour $\varepsilon > 0$, la boucle termine en arithmétique exacte. La valeur renvoyée est au plus à la moitié de la longueur finale de la racine, donc $|\beta_a - \beta| \leq \varepsilon$.

Comme toute dichotomie flottante, elle est à utiliser avec une tolérance compatible avec la précision machine. Code exécuté pour des tolérances usuelles, notamment 10^{-10} .

- 23.** (d)(iii) Question adaptée au programme actuel : une dichotomie donne une approximation de β , donc le tableau renvoyé approche X_0 ; il n'est généralement pas la solution exacte.

On utilise directement la formule de (c) et on conserve les composantes non pénalisées.

Python

```

import numpy as np

def CalcSolution(alpha, Y, epsilon):
    k = np.shape(alpha)[0]
    n = np.shape(Y)[0]
    beta = CalcBeta(alpha, Y, epsilon)
    X = np.zeros(n)

```

```

for i in range(k):
    X[i] = beta * Y[i] / (beta + alpha[i] ** 2)

for i in range(k, n):
    X[i] = Y[i]

return X

```

Lorsque $\rho(Y) \leq 1$, β vaut exactement 0 dans le modèle mathématique : les premières composantes sont nulles, les suivantes valent y_i .

Sinon, si $b = \beta_a$, la différence sur la composante $i \leq k$ vaut exactement

$$\frac{by_i}{b + \alpha_i^2} - \frac{\beta y_i}{\beta + \alpha_i^2} = \frac{\alpha_i^2 y_i (b - \beta)}{(b + \alpha_i^2)(\beta + \alpha_i^2)}.$$

Puisque $b, \beta \geq 0$ et $\alpha_i \neq 0$, sa valeur absolue est majorée par $|y_i| \varepsilon / \alpha_i^2$. Donc

$$\|X_a - X_0\| \leq \varepsilon \sqrt{\sum_{i=1}^k \frac{y_i^2}{\alpha_i^4}}.$$

La tolérance sur β n'est donc pas automatiquement la même tolérance sur le vecteur solution. Le code a été exécuté sur les cas $\rho(Y) < 1$, $\rho(Y) = 1$, $\rho(Y) > 1$, avec des coefficients α_i de signes variés.

Vérifications numériques des programmes

Les fonctions des questions 10(b), 16(b)–(c) et 23(d) ont été exécutées. Sur les trois matrices de 16(d), les rangs numériques obtenus sont 2, 3, 5 ; les matrices de projection calculées sont symétriques et idempotentes à la précision flottante. La branche de message d'erreur de 10(b) a aussi été contrôlée.

Les simulations de 16(c) sont cohérentes avec $p\sigma^2$ à leurs fluctuations Monte-Carlo près ; ces observations accompagnent les preuves exactes de (f)–(i).

Les cas de seuil et les deux branches de 23(d) ont été vérifiés, ainsi que le résidu $F(\beta_a)$ et la condition d'optimalité lorsque $\rho(Y) > 1$. Ces contrôles numériques ne remplacent pas les démonstrations.

Relevé des adaptations

- 16 : ajout explicite de $\sigma > 0$, domaine de l'écart-type d'une loi normale non dégénérée et condition permettant la division par $p\sigma^2$ en (i).
- 16(d) : exception locale $n = 3$ explicitée pour A_1 , dont les données sont conservées malgré la convention générale $n \geq 4$.
- 16(d) : le nom `esperanceT` du programme source est remplacé par `esperance`, nom imposé en 16(c), afin d'assurer la cohérence des appels.
- 16(d) : la figure est retranscrite en SVG dans l'énoncé à partir de l'original, avec les valeurs mesurées.
- 19(a) : $H \in \mathcal{M}_{n,1}(\mathbb{R})$ est remplacé par $H \in \mathcal{M}_{p,1}(\mathbb{R})$; la dimension originale rendait certains produits et sommes non définis lorsque $n \neq p$.
- 23(d)(iii) : «contenant la solution» est précisé en «approchant la solution», car β est calculé par dichotomie avec une tolérance positive.
- Formulaire I : les tableaux renvoyés par `linspace` et `zeros(n)` sont décrits comme monodimensionnels, plutôt que comme des matrices lignes, pour respecter leurs dimensions réelles dans NumPy.
- Formulaire III : le paramètre d'échelle de `rd.gamma(m, a, ...)` est explicité par la loi de aG , afin d'éviter toute ambiguïté avec un paramètre de taux. Le paramètre `d` de `normal` est l'écart-type.