



PRÉPASIMPLE

prepasimple.fr

ESSEC

Mathématiques 1

1999

Corrigé de mathématiques approfondies

Mathématiques approfondies — Prépa ECG

Annales de concours · Corrigé

Préliminaire : trace d'une matrice et d'un endomorphisme

a.

On utilise la définition du produit matriciel. Pour chaque $k \in \{1, \dots, p\}$, le coefficient diagonal de AB vaut

$$(AB)_{k,k} = \sum_{j=1}^p a_{k,j} b_{j,k}.$$

Ainsi :

$$\text{Tr}(AB) = \sum_{k=1}^p \sum_{j=1}^p a_{k,j} b_{j,k}.$$

Les deux sommes sont finies ; on peut permuter leur ordre. Comme les coefficients sont réels, on peut aussi permuter les deux facteurs de chaque produit. Il vient :

$$\text{Tr}(AB) = \sum_{j=1}^p \sum_{k=1}^p b_{j,k} a_{k,j} = \text{Tr}(BA).$$

La trace est également linéaire, directement par sa définition : pour toutes matrices A, B et tous réels s, t , $\text{Tr}(sA + tB) = s \text{Tr}(A) + t \text{Tr}(B)$. Cette propriété sera utilisée dans la partie II.

b.

Les matrices M' et M sont semblables : il existe donc une matrice inversible P telle que $M' = P^{-1}MP$. On applique a. aux matrices carrées $P^{-1}M$ et P :

$$\text{Tr}(M') = \text{Tr}((P^{-1}M)P) = \text{Tr}(P(P^{-1}M)) = \text{Tr}(M).$$

Les matrices d'un même endomorphisme dans deux bases sont semblables. Leur trace

est donc la même, ce qui justifie la définition de la trace d'un endomorphisme donnée dans le sujet.

Partie I : étude des éléments de l'ensemble $T(E)$

1. Sous-espace orthogonal à un vecteur non nul x de E .

(a)

L'orthogonalité cherchée équivaut à :

$$\langle v - \lambda(v)x, x \rangle = 0 \iff \langle v, x \rangle - \lambda(v)\|x\|^2 = 0.$$

Comme $x \neq 0$, on a $\|x\|^2 > 0$ et cette équation possède une unique solution :

$$\lambda(v) = \frac{\langle x, v \rangle}{\|x\|^2}.$$

Le vecteur $\lambda(v)x$ est donc la projection orthogonale de v sur la droite $\mathbb{R}x$.

(b)

Notons $X = x^\perp = \{w \in E \mid \langle w, x \rangle = 0\}$. C'est un sous-espace vectoriel, puisque l'application $w \mapsto \langle w, x \rangle$ est linéaire.

La formule (1) écrit chaque $v \in E$ comme la somme de $\lambda(v)x \in \mathbb{R}x$ et de $v - \lambda(v)x \in X$. Elle donne donc $E = \mathbb{R}x + X$.

Si $w \in \mathbb{R}x \cap X$, il existe $s \in \mathbb{R}$ tel que $w = sx$, et :

$$0 = \langle w, x \rangle = s\|x\|^2.$$

On a $s = 0$, donc $w = 0$. La somme est directe et les deux sous-espaces sont supplémentaires :

$$E = \mathbb{R}x \oplus x^\perp.$$

En particulier, $\dim x^\perp = p - 1$. Lorsque $p = 1$, ce sous-espace est $\{0\}$ et la décomposition reste valable.

2. Élément de $T(E)$ associé à un vecteur de E .

(a)

Pour $x \neq 0$, la bilinéarité du produit scalaire entraîne la linéarité de l'application $v \mapsto u_x(v) = \langle x, v \rangle x$.

Pour tous $v, w \in E$, on a :

$$\langle u_x(v), w \rangle = \langle x, v \rangle \langle x, w \rangle = \langle v, u_x(w) \rangle.$$

L'endomorphisme u_x est donc symétrique. Son image est contenue dans $\mathbb{R}x$; comme $u_x(x) = \|x\|^2 x \neq 0$, son image est exactement $\mathbb{R}x$ et son rang vaut 1.

Enfin, pour tout $v \in E$:

$$\langle u_x(v), v \rangle = \langle x, v \rangle^2 \geq 0.$$

Ces trois propriétés montrent que $u_x \in T(E)$.

Choisissons une base $(w_2, \dots, w_p)$ de $X = x^\perp$ et complétons-la par x . D'après I-1.(b), $\mathcal{B} = (x, w_2, \dots, w_p)$ est une base de E . Comme $u_x(x) = \|x\|^2 x$ et $u_x(w_j) = 0$, on obtient :

$$\text{Mat}_{\mathcal{B}}(u_x) = \text{diag}(\|x\|^2, 0, \dots, 0).$$

La matrice de $u_x \circ u_x$ est le carré de cette matrice. Par définition de la trace :

$$\text{Tr}(u_x) = \|x\|^2, \quad \text{Tr}(u_x \circ u_x) = \|x\|^4.$$

Pour $p = 1$, la base est simplement (x) et les matrices sont d'ordre un ; les mêmes formules s'appliquent.

(b)

On a $u_x(x) = \|x\|^2 x$; le nombre $\|x\|^2 > 0$ est une valeur propre. Pour tout $w \in x^\perp$, $u_x(w) = 0$.

Pour déterminer tous les espaces propres, écrivons $v = sx + w$ avec $w \in x^\perp$. Alors $u_x(v) = s\|x\|^2 x$. L'équation $u_x(v) = \|x\|^2 v$ entraîne $\|x\|^2 w = 0$, donc $w = 0$. Ainsi :

$$E_{\|x\|^2}(u_x) = \mathbb{R}x.$$

D'autre part, $u_x(v) = 0$ équivaut à $\langle x, v \rangle = 0$, donc $\text{Ker } u_x = x^\perp$.

Si $p \geq 2$, cet espace est de dimension $p - 1 \geq 1$; 0 est une valeur propre et :

$$E_0(u_x) = x^\perp.$$

La matrice diagonale de (a) montre qu'il n'existe pas d'autre valeur propre. Si $p = 1$, le noyau est réduit à $\{0\}$: 0 n'est pas une valeur propre, et la seule valeur propre est $\|x\|^2$.

(c)

Méthode 1 — Employer la base adaptée et la formule (1), comme demandé.

Notons $r = \|x\|^2 > 0$. La formule (1), appliquée à $f(x)$, donne :

$$f(x) = \frac{\langle x, f(x) \rangle}{r} x + w, \quad w \in x^\perp.$$

Ainsi $(f \circ u_x)(x) = rf(x) = \langle x, f(x) \rangle x + rw$. Le premier élément diagonal de la matrice de $f \circ u_x$ dans la base $\mathcal{B} = (x, w_2, \dots, w_p)$ vaut donc $\langle x, f(x) \rangle$.

Pour $j \geq 2$, $u_x(w_j) = 0$, donc $(f \circ u_x)(w_j) = 0$. Toutes les colonnes d'indice au moins deux sont nulles, et les éléments diagonaux correspondants valent 0. Par conséquent :

$$\text{Tr}(f \circ u_x) = \langle x, f(x) \rangle.$$

La symétrie de f n'est pas nécessaire à cette identité.

Méthode 2 — Calculer dans une base orthonormale.

Choisissons une base orthonormale et notons $X = {}^t(x_1, \dots, x_p)$ la colonne des coordonnées de x , et $C = (c_{i,j})$ la matrice de f . Pour une colonne V , les coordonnées de $u_x(v)$ sont $X({}^tXV)$. La matrice de u_x est donc $X{}^tX$.

La matrice de $f \circ u_x$ est $CX{}^tX$. Son coefficient diagonal d'indice k vaut

$$\sum_{j=1}^p c_{k,j} x_j x_k.$$

Ainsi :

$$\text{Tr}(f \circ u_x) = \sum_{k=1}^p \sum_{j=1}^p x_k c_{k,j} x_j = {}^tX C X = \langle x, f(x) \rangle.$$

Cette seconde méthode fournit la formule directement dans des coordonnées usuelles ;

la première explique pourquoi une seule entrée diagonale contribue à la trace dans la base adaptée.

3. Vecteurs de E associés à un élément de $T(E)$.

On fixe $u \in T(E)$ non nul et $x \neq 0$ appartenant à $\text{Im } u$.

(a)

Comme u est non nul et de rang au plus un, son rang est exactement un. Ainsi $\text{Im } u = \mathbb{R}x$. Le vecteur $u(x)$ appartient à cette droite ; il existe donc un réel μ tel que $u(x) = \mu x$. Comme $x \neq 0$, c'est un vecteur propre de u associé à μ .

La positivité au sens de la définition de $T(E)$ donne :

$$0 \leq \langle u(x), x \rangle = \mu \|x\|^2.$$

En divisant par $\|x\|^2 > 0$, on obtient $\mu \geq 0$. À ce stade, « positive » signifie donc positive ou nulle ; la stricte positivité sera établie en (c).

(b)

On montre d'abord que u s'annule sur $x^\perp$. Pour $w \in x^\perp$, on a $u(w) \in \mathbb{R}x$, donc $u(w) = ax$ pour un réel a . La symétrie de u et l'égalité $u(x) = \mu x$ donnent :

$$a\|x\|^2 = \langle u(w), x \rangle = \langle w, u(x) \rangle = \mu \langle w, x \rangle = 0.$$

Ainsi $a = 0$ et $u(w) = 0$.

Pour $v \in E$, la formule (1) fournit $v = \lambda(v)x + w$ avec $w \in x^\perp$ et $\lambda(v) = \langle x, v \rangle / \|x\|^2$. Par linéarité :

$$u(v) = \lambda(v)u(x) + u(w) = \frac{\mu}{\|x\|^2} \langle x, v \rangle x.$$

Cette formule est valable pour tout $v \in E$.

(c)

Si $\mu = 0$, la formule de (b) entraînerait $u(v) = 0$ pour chaque $v \in E$, contrairement au choix de u non nul. Puisque $\mu \geq 0$, on a donc $\mu > 0$.

Définissons :

$$y = \frac{\sqrt{\mu}}{\|x\|} x.$$

Cette définition est licite car $\mu > 0$ et $x \neq 0$. Pour tout $v \in E$:

$$\langle y, v \rangle y = \frac{\mu}{\|x\|^2} \langle x, v \rangle x = u(v).$$

Ainsi $u = u_y$, et y est non nul. Tout élément non nul de $T(E)$ est donc de la forme u_y .

(d)

Pour $x = 0$, la même formule définit $u_0 = 0$, qui appartient à $T(E)$. L'application considérée est donc bien définie sur tout E .

Elle n'est pas injective : comme $p \geq 1$, il existe $x \neq 0$, et x et $-x$ sont deux vecteurs distincts tels que :

$$u_{-x}(v) = \langle -x, v \rangle (-x) = \langle x, v \rangle x = u_x(v).$$

Elle est surjective : l'endomorphisme nul est l'image de 0, et chaque élément non nul de $T(E)$ s'écrit u_y d'après (c).

Pour préciser cette non-injectivité, si $u_y = u_x \neq 0$, l'égalité de leurs images entraîne $y = ax$ pour un réel $a \neq 0$. Or $u_y = a^2 u_x$, donc $a^2 = 1$ et $y = x$ ou $y = -x$.

L'endomorphisme nul n'a que 0 comme antécédent, puisque $u_x(x) = \|x\|^2 x$ est non nul lorsque $x \neq 0$.

L'application est donc surjective, mais non injective ; chaque endomorphisme non nul possède exactement deux antécédents opposés.

Partie II : approximation des éléments de $S(E)$ par des éléments de $T(E)$

Rectification du titre : le PDF répète « Partie I » au début de cette partie. Il s'agit de la partie II, comme l'annonce le préambule. Seul cet intitulé est rectifié ; les questions et leurs numéros sont conservés.

1. Un produit scalaire sur $S(E)$.

(a)

Fixons une base orthonormale de E . Les matrices $A = (a_{i,j})$ et $B = (b_{i,j})$ de $f, g \in S(E)$ sont réelles symétriques. D'après le préliminaire :

$$[f, g] = \text{Tr}(AB) = \sum_{i=1}^p \sum_{j=1}^p a_{i,j} b_{j,i} = \sum_{i=1}^p \sum_{j=1}^p a_{i,j} b_{i,j}.$$

On vérifie les trois propriétés d'un produit scalaire réel.

La bilinéarité résulte de cette somme de produits de coefficients : la matrice de $sf + tg$ est $sA + tB$, et les coefficients se combinent linéairement. En particulier,

$[sf + tg, h] = s[f, h] + t[g, h]$, et de même dans la seconde variable.

La symétrie vient de $\text{Tr}(AB) = \text{Tr}(BA)$, ou directement de la dernière somme :

$[f, g] = [g, f]$.

Enfin :

$$[f, f] = \sum_{i=1}^p \sum_{j=1}^p a_{i,j}^2 \geq 0.$$

Cette somme finie de carrés est nulle si et seulement si tous les coefficients de A sont nuls, donc si et seulement si $f = 0$. On a établi la positivité et le caractère défini.

Ainsi $[\cdot, \cdot]$ est un produit scalaire sur $S(E)$ et la norme associée est :

$$N(f) = \sqrt{[f, f]} = \left(\sum_{i=1}^p \sum_{j=1}^p a_{i,j}^2 \right)^{1/2}.$$

Le produit $f \circ g$ n'est pas nécessairement symétrique, mais sa trace est bien définie pour tout endomorphisme ; cela ne gêne donc pas cette construction.

(b)

Puisque f et u_x appartiennent à $S(E)$, on peut développer le carré de la norme issue du produit scalaire précédent :

$$N^2(f - u_x) = [f - u_x, f - u_x] = [f, f] - 2[f, u_x] + [u_x, u_x].$$

Si $x \neq 0$, I-2.(a) et I-2.(c) donnent :

$$[f, u_x] = \text{Tr}(f \circ u_x) = \langle x, f(x) \rangle, \quad [u_x, u_x] = \text{Tr}(u_x \circ u_x) = \|x\|^4.$$

Si $x = 0$, on a $u_x = 0$, et ces deux identités restent vraies, leurs membres étant nuls. Pour tout $x \in E$, il vient donc :

$$N^2(f - u_x) = N^2(f) - 2\langle x, f(x) \rangle + \|x\|^4 = F(x).$$

Grâce à la surjectivité prouvée en I-3.(d), minimiser F sur E revient exactement à minimiser $N^2(f - u)$ pour $u \in T(E)$.

On sait déjà que $F(x) \geq 0$ pour tout x et que $F(0) = N^2(f)$. Ainsi son infimum existe et appartient à $[0, N^2(f)]$. L'existence de vecteurs qui atteignent cet infimum sera démontrée ci-dessous.

2. Condition nécessaire de minimum pour F .

(a)

Fixons $x \in E$ et $y \in E$ unitaire. Posons, pour ce calcul :

$$r = \|x\|^2, \quad b = \langle x, y \rangle, \quad c = \langle x, f(x) \rangle, \quad d = \langle y, f(x) \rangle, \quad q = \langle y, f(y) \rangle.$$

La symétrie de f assure $\langle x, f(y) \rangle = \langle f(x), y \rangle = d$. On a donc :

$$\langle x + ty, f(x + ty) \rangle = c + 2dt + qt^2.$$

Comme $\|y\| = 1$, le carré de la norme vaut :

$$\|x + ty\|^2 = r + 2bt + t^2.$$

Son carré est :

$$(r + 2bt + t^2)^2 = r^2 + 4rbt + (2r + 4b^2)t^2 + 4bt^3 + t^4.$$

En substituant ces deux développements dans la définition de F , on obtient :

$$h(t) = F(x) + 4(rb - d)t + (2r + 4b^2 - 2q)t^2 + 4bt^3 + t^4.$$

En particulier, le coefficient de t^4 vaut 1, donc h est un polynôme de degré exactement 4 pour chaque choix de x et de y unitaire. Tous ses coefficients sont explicités par la formule précédente.

(b)

Si F présente un minimum en x , on a $F(x + ty) \geq F(x)$ pour tout réel t , quel que soit le vecteur unitaire y . Donc h présente un minimum en 0.

La fonction h est dérivable sur $\mathbb{R}$, car elle est polynomiale, et 0 est un point intérieur de son domaine. La condition nécessaire de minimum pour une fonction dérivable d'une variable réelle donne :

$$h'(0) = 0.$$

On peut aussi vérifier cette condition directement : pour $t > 0$, $(h(t) - h(0))/t \geq 0$, tandis que pour $t < 0$, ce quotient est inférieur ou égal à 0. Ses deux limites coïncident avec $h'(0)$, donc cette dérivée est à la fois positive ou nulle et négative ou nulle. Elle est nulle. Le raisonnement vaut pour chaque vecteur unitaire y .

(c)

Méthode 1 — Exploiter les dérivées des restrictions aux droites.

La formule de (a) donne $h'(0) = 4(rb - d)$. La condition de (b), valable pour chaque y unitaire, implique :

$$\langle y, \|x\|^2 x - f(x) \rangle = 0 \quad \text{pour tout } y \text{ unitaire.}$$

Posons $z = \|x\|^2 x - f(x)$. Si $z \neq 0$, le vecteur $y = z/\|z\|$ est unitaire et $\langle y, z \rangle = \|z\| > 0$, contradiction. Ainsi $z = 0$ et :

$$f(x) = \|x\|^2 x.$$

Méthode 2 — Calculer le gradient en coordonnées orthonormales.

Dans une base orthonormale, notons $A = (a_{i,j})$ la matrice symétrique de f et $X = {}^t(x_1, \dots, x_p)$ la colonne des coordonnées de x . La fonction F s'écrit :

$$F(X) = N^2(f) - 2 \sum_{i,j=1}^p a_{i,j} x_i x_j + \left(\sum_{j=1}^p x_j^2 \right)^2.$$

C'est un polynôme en les coordonnées, donc une fonction de classe C^2 , et même C^∞ , sur l'ouvert $\mathbb{R}^p$. La symétrie de A donne, pour chaque k :

$$\frac{\partial F}{\partial x_k}(X) = -4 \sum_{j=1}^p a_{k,j} x_j + 4 \left(\sum_{j=1}^p x_j^2 \right) x_k.$$

Ainsi :

$$\nabla F(X) = 4(\|X\|^2 X - AX).$$

En un minimum sur l'ouvert $\mathbb{R}^p$, le gradient s'annule. On retrouve donc $AX = \|X\|^2 X$, c'est-à-dire $f(x) = \|x\|^2 x$.

En dérivant une seconde fois, on obtient la hessienne, utile pour contrôler les extrema dans la suite :

$$\nabla^2 F(X) = 8X^t X + 4\|X\|^2 I_p - 4A.$$

La première méthode respecte directement la démarche du sujet et ne nécessite pas de coordonnées. La seconde relie cette condition à l'annulation du gradient d'une fonction de plusieurs variables. La seule annulation du gradient ne suffit pas à prouver un minimum.

(d)

Si F présente un minimum en x , la question (c) donne $f(x) = \|x\|^2 x$. Avec les notations de (a), cela entraîne $d = rb$; le terme de degré un dans $h(t) - h(0)$ est donc nul.

On obtient :

$$\begin{aligned} F(x + ty) - F(x) &= (2r + 4b^2 - 2q)t^2 + 4bt^3 + t^4 \\ &= t^2(t^2 + 4bt + 4b^2 + 2(r - q)) \\ &= t^2((t + 2b)^2 + 2(r - q)). \end{aligned}$$

En remplaçant b , r et q par leurs définitions, c'est exactement :

$$F(x + ty) - F(x) = t^2 [(t + 2\langle y, x \rangle)^2 + 2(\|x\|^2 - \langle y, f(y) \rangle)].$$

Il est important pour la question suivante de remarquer que ce calcul utilise uniquement l'égalité $f(x) = \|x\|^2 x$. L'identité reste donc vraie dès que cette égalité est vérifiée, sans supposer préalablement que x réalise un minimum.

3. Condition nécessaire et suffisante de minimum pour F .

Méthode 1 — Utiliser l'identité des restrictions aux droites.

Supposons d'abord $F(x) = m(f)$. La condition (i), $f(x) = \|x\|^2 x$, a été démontrée en II-2.(c). Fixons un vecteur unitaire y et posons $b = \langle y, x \rangle$ et $r = \|x\|^2$. Pour tout $t \neq 0$, l'identité de II-2.(d) et la minimalité de x donnent :

$$(t + 2b)^2 + 2(r - \langle y, f(y) \rangle) \geq 0.$$

Si $b \neq 0$, on peut choisir $t = -2b \neq 0$ et obtenir directement $r - \langle y, f(y) \rangle \geq 0$. Si $b = 0$, l'inégalité est $t^2 + 2(r - \langle y, f(y) \rangle) \geq 0$ pour tout $t \neq 0$; en faisant tendre t vers 0, on obtient le même résultat. Ainsi la condition (ii) est nécessaire pour chaque y unitaire :

$$\langle y, f(y) \rangle \leq \|x\|^2.$$

Réciproquement, supposons (i) et (ii). L'identité de II-2.(d) est valable grâce à (i), et chacun des termes à l'intérieur du crochet y est positif ou nul grâce à (ii). Elle donne $F(x + ty) \geq F(x)$ pour tout $t \in \mathbb{R}$ et tout y unitaire.

Tout $z \neq x$ s'écrit $z = x + ty$ en prenant $t = \|z - x\| > 0$ et $y = (z - x)/\|z - x\|$. On a donc $F(z) \geq F(x)$ pour tous $z \in E$, l'égalité étant automatique pour $z = x$. Ainsi $F(x) = m(f)$.

Les deux conditions sont donc nécessaires et suffisantes, et leur réalisation implique bien que l'infimum est atteint.

Méthode 2 — Comparer globalement les valeurs de F .

Supposons que x réalise le minimum. La question II-2.(c) fournit (i). En posant $r = \|x\|^2$, on a alors $\langle x, f(x) \rangle = r^2$ et :

$$F(x) = N^2(f) - r^2.$$

Si (ii) était fausse, il existerait un vecteur unitaire y tel que $q = \langle y, f(y) \rangle > r \geq 0$. Définissons $z = \sqrt{q} y$. On a $\|z\|^2 = q$ et $\langle z, f(z) \rangle = q^2$, donc :

$$F(z) = N^2(f) - q^2 < N^2(f) - r^2 = F(x).$$

C'est impossible pour un minimum ; (ii) est donc nécessaire.

Réciproquement, si (i) et (ii) sont vérifiées, alors pour $z \neq 0$ la normalisation $y = z/\|z\|$ donne $\langle z, f(z) \rangle \leq r\|z\|^2$. Cette inégalité est aussi vraie pour $z = 0$. On obtient pour tout z :

$$F(z) - F(x) \geq \|z\|^4 - 2r\|z\|^2 + r^2 = (\|z\|^2 - r)^2 \geq 0.$$

Donc x réalise le minimum. La première méthode examine toutes les droites passant par x ; la seconde compare directement toutes les valeurs à l'aide d'une borne quadratique en $\|z\|^2$.

4. Étude du maximum de $\langle y, f(y) \rangle$ pour $\|y\| = 1$.

(a)

Le théorème spectral réel affirme que tout endomorphisme symétrique d'un espace euclidien réel de dimension finie possède une base orthonormale de vecteurs propres, avec des valeurs propres réelles.

Les hypothèses sont satisfaites : E est euclidien réel de dimension finie p , et $f \in S(E)$ est symétrique. Il existe donc une base orthonormale $(e_1, \dots, e_p)$ telle que :

$$f(e_j) = \lambda_j e_j \quad (1 \leq j \leq p).$$

En permutant les vecteurs de cette base, on peut classer les valeurs propres dans l'ordre :

$$\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_p.$$

Les valeurs peuvent être répétées ; on ne suppose pas qu'elles sont distinctes.

(b)

Dans la base orthonormale de (a), la matrice de f est diagonale de coefficients $\lambda_1, \dots, \lambda_p$. Celle de $f \circ f$ est donc diagonale de coefficients $\lambda_1^2, \dots, \lambda_p^2$.

Par la définition du produit scalaire et de la norme :

$$N^2(f) = [f, f] = \text{Tr}(f \circ f) = \sum_{j=1}^p \lambda_j^2.$$

Ainsi :

$$N(f) = \left(\sum_{j=1}^p \lambda_j^2 \right)^{1/2}.$$

(c)

Soit y unitaire. Écrivons

$$y = \sum_{j=1}^p y_j e_j.$$

Comme la base est orthonormale,

$$\sum_j y_j^2 = \|y\|^2 = 1.$$

Par diagonalisation :

$$\langle y, f(y) \rangle = \sum_{j=1}^p \lambda_j y_j^2 \leq \lambda_p \sum_{j=1}^p y_j^2 = \lambda_p.$$

Cette borne est atteinte par le vecteur unitaire e_p . Le supremum est donc un maximum et vaut λ_p .

Pour caractériser tous les cas d'égalité, on écrit :

$$\lambda_p - \langle y, f(y) \rangle = \sum_{j=1}^p (\lambda_p - \lambda_j) y_j^2.$$

Chaque terme est positif ou nul. La somme est nulle exactement lorsque $y_j = 0$ pour chaque indice tel que $\lambda_j < \lambda_p$. Autrement dit, l'égalité est vérifiée exactement par les vecteurs unitaires de l'espace propre :

$$E_{\lambda_p}(f) = \text{Ker}(f - \lambda_p \text{Id}_E).$$

Cela couvre aussi le cas où λ_p est multiple, ou celui où toutes les valeurs propres sont égales.

5. Conclusion et valeur de $m(f)$.

(a)

Méthode 1 — Comparer directement $F(x)$ à $F(0)$.

Supposons $\lambda_p \leq 0$. La décomposition dans une base orthonormale de vecteurs propres donne, pour tout

$$x = \sum_j x_j e_j :$$

$$\langle x, f(x) \rangle = \sum_{j=1}^p \lambda_j x_j^2 \leq 0,$$

car $\lambda_j \leq \lambda_p \leq 0$. Il s'ensuit :

$$F(x) = N^2(f) - 2\langle x, f(x) \rangle + \|x\|^4 \geq N^2(f) + \|x\|^4.$$

Or $F(0) = N^2(f)$. Si $x \neq 0$, le terme $\|x\|^4$ est strictement positif, donc $F(x) > F(0)$. Le minimum existe et est atteint uniquement en 0 :

$$m(f) = N^2(f) = \sum_{j=1}^p \lambda_j^2, \quad F(x) = m(f) \iff x = 0.$$

Méthode 2 — Appliquer les conditions de II-3.

Pour $x = 0$, la condition (i) est vérifiée puisque $f(0) = 0$. La condition (ii) l'est également, car pour chaque y unitaire, $\langle y, f(y) \rangle \leq \lambda_p \leq 0$. Donc 0 réalise le minimum.

Si un autre vecteur $x \neq 0$ réalisait le minimum, la condition (i) donnerait $f(x) = \|x\|^2 x$. Le nombre $\|x\|^2 > 0$ serait une valeur propre de f , alors que toutes ses valeurs propres sont inférieures ou égales à 0. C'est impossible. Le minimum est donc atteint uniquement en 0, avec la même valeur $N^2(f)$.

La première méthode fournit une minoration explicite et stricte hors de l'origine ; la seconde montre comment la condition de minimum exclut tout vecteur propre de valeur propre non positive.

Le calcul de la hessienne de II-2.(c) donne, à l'origine, $\nabla^2 F(0) = -4A$ dans une base orthonormale. Ses valeurs propres sont $-4\lambda_j \geq 0$. Si $\lambda_p < 0$, elles sont toutes strictement positives. Si $\lambda_p = 0$, au moins une est nulle : le seul critère hessien ne suffit alors pas à

conclure. La minoration $F(x) - F(0) \geq \|x\|^4$ prouve néanmoins que 0 reste l'unique minimum global, et un minimum strict.

(b)

On suppose maintenant $\lambda_p > 0$.

Méthode 1 — Employer les deux conditions nécessaires et suffisantes.

Si x réalise le minimum, la condition (ii) de II-3, combinée à II-4.(c), donne $\lambda_p \leq \|x\|^2$.

En particulier, $x \neq 0$. La condition (i) donne $f(x) = \|x\|^2 x$, donc $\|x\|^2$ est une valeur propre de f et satisfait $\|x\|^2 \leq \lambda_p$.

Les deux inégalités entraînent $\|x\|^2 = \lambda_p$ et $x \in E_{\lambda_p}(f)$.

Réciproquement, si $x \in E_{\lambda_p}(f)$ et $\|x\|^2 = \lambda_p$, alors $f(x) = \lambda_p x = \|x\|^2 x$. De plus, II-4.(c) donne $\langle y, f(y) \rangle \leq \lambda_p = \|x\|^2$ pour chaque y unitaire. Les deux conditions de II-3 sont satisfaites, donc x réalise le minimum.

Ces vecteurs existent : on peut prendre $x = \sqrt{\lambda_p} e_p$. Pour chacun d'eux,

$\langle x, f(x) \rangle = \lambda_p \|x\|^2 = \lambda_p^2$ et $\|x\|^4 = \lambda_p^2$. On obtient :

$$m(f) = N^2(f) - \lambda_p^2 = \sum_{j=1}^{p-1} \lambda_j^2.$$

L'ensemble des vecteurs qui atteignent le minimum est :

$$\{x \in E_{\lambda_p}(f) \mid \|x\| = \sqrt{\lambda_p}\} = \{\sqrt{\lambda_p} y \mid y \in E_{\lambda_p}(f), \|y\| = 1\}.$$

Méthode 2 — Établir une identité de minoration globale.

Pour tout x , II-4.(c), appliquée à $x/\|x\|$ si $x \neq 0$, donne $\langle x, f(x) \rangle \leq \lambda_p \|x\|^2$. L'inégalité est également vraie pour $x = 0$. On réécrit exactement $F(x)$ sous la forme :

$$F(x) = N^2(f) - \lambda_p^2 + (\|x\|^2 - \lambda_p)^2 + 2(\lambda_p \|x\|^2 - \langle x, f(x) \rangle).$$

Les deux derniers termes sont positifs ou nuls. Ainsi $F(x) \geq N^2(f) - \lambda_p^2$ pour chaque x . L'égalité exige simultanément $\|x\|^2 = \lambda_p$ et $\langle x, f(x) \rangle = \lambda_p \|x\|^2$.

La première égalité assure $x \neq 0$; la seconde, d'après le cas d'égalité de II-4.(c), équivaut à $x \in E_{\lambda_p}(f)$. Les vecteurs $\sqrt{\lambda_p} y$, avec y unitaire dans cet espace propre, satisfont les deux égalités. On retrouve donc le minimum et tous les minimiseurs donnés par la première méthode.

La première méthode exploite la caractérisation obtenue en II-3. La seconde donne une décomposition de l'écart au minimum en deux termes positifs ou nuls, ce qui prouve directement l'optimalité globale.

Pour contrôler le rôle de la hessienne, considérons un de ces minimiseurs x et sa colonne X dans une base orthonormale. La fonction F est de classe C^2 sur $\mathbb{R}^p$, son gradient s'annule en X , et :

$$\nabla^2 F(X) = 8X^t X + 4\lambda_p I_p - 4A.$$

Le vecteur unitaire $x/\sqrt{\lambda_p}$ est propre pour cette hessienne avec valeur propre $8\lambda_p > 0$. Comme f est symétrique et x est propre, l'espace $x^\perp$ est stable par f : pour $v \perp x$, $\langle f(v), x \rangle = \langle v, f(x) \rangle = 0$. Le théorème spectral appliqué à la restriction à cet espace fournit une base orthonormale de vecteurs propres $v \perp x$, avec valeurs propres $\mu \leq \lambda_p$. Sur chacun de ces vecteurs, la hessienne agit par :

$$\nabla^2 F(X)v = 4(\lambda_p - \mu)v.$$

Ses autres valeurs propres sont donc positives ou nulles. Elles sont toutes strictement positives si l'espace propre de λ_p est une droite. Si cet espace propre a dimension au moins deux, la hessienne possède une valeur propre nulle : elle ne suffit pas, à elle seule, à conclure. L'identité globale précédente prouve le minimum même dans ce cas et décrit l'ensemble entier des minimisateurs.

Pour $p = 1$, la somme

$$\sum_{j=1}^{p-1} \lambda_j^2$$

est vide et vaut 0. Les deux minimisateurs sont les vecteurs $\pm \sqrt{\lambda_1} e_1$. Ils définissent le même endomorphisme $u_x = f$, ce qui est cohérent avec une erreur minimale nulle.

6. Application à l'étude d'un exemple.

On travaille dans $E = \mathbb{R}^p$ avec le produit scalaire canonique. La matrice M est symétrique, tous ses coefficients sont strictement positifs, et la somme des coefficients de chacune de ses lignes vaut 1.

(a)

Notons $\mathbf{1} = {}^t(1, \dots, 1)$, colonne non nulle puisque $p \geq 1$. Pour chaque indice i :

$$(M\mathbf{1})_i = \sum_{j=1}^p m_{i,j} = 1.$$

Ainsi $M\mathbf{1} = \mathbf{1}$: 1 est une valeur propre de M et $\mathbf{1}$ est un vecteur propre associé.

(b)

La matrice M est réelle symétrique, donc ses valeurs propres sont réelles. Fixons une valeur propre λ et un vecteur propre $X = {}^t(x_1, \dots, x_p) \neq 0$. Choisissons un indice i tel que :

$$a = |x_i| = \max_{1 \leq j \leq p} |x_j| > 0.$$

La i -ième ligne de $MX = \lambda X$ donne

$$\lambda x_i = \sum_j m_{i,j} x_j.$$

Par l'inégalité triangulaire et la positivité des coefficients :

$$|\lambda|a = \left| \sum_{j=1}^p m_{i,j} x_j \right| \leq \sum_{j=1}^p m_{i,j} |x_j| \leq a \sum_{j=1}^p m_{i,j} = a.$$

Comme $a > 0$, on peut diviser par a et obtenir $|\lambda| \leq 1$.

Supposons maintenant $|\lambda| = 1$. Les deux inégalités de la chaîne précédente sont alors des égalités. L'égalité dans la seconde impose :

$$\sum_{j=1}^p m_{i,j} (a - |x_j|) = 0.$$

Chaque terme est positif ou nul, et chaque $m_{i,j}$ est strictement positif. On en déduit $|x_j| = a$ pour tout j . En particulier, aucune composante de X n'est nulle.

L'égalité dans l'inégalité triangulaire implique que les réels $m_{i,j} x_j$ sont tous de même signe. Comme les coefficients $m_{i,j}$ sont strictement positifs, les x_j sont eux-mêmes tous de même signe. En combinant ce fait avec $|x_j| = a$, on obtient :

$$x_1 = x_2 = \dots = x_p = c, \quad c \neq 0.$$

Ainsi $X = c\mathbf{1}$. D'après (a), $MX = X$; l'équation propre donne donc $\lambda X = X$. Puisque $X \neq 0$, on a $\lambda = 1$. En particulier, -1 ne peut pas être une valeur propre.

Tout vecteur propre associé à 1 est un multiple de $\mathbf{1}$, et cette colonne est elle-même propre. Donc :

$$E_1(M) = \mathbb{R}\mathbf{1}, \quad \dim E_1(M) = 1.$$

Toutes les autres valeurs propres, s'il y en a, vérifient $|\lambda| < 1$. Les hypothèses de stricte positivité de tous les coefficients sont essentielles pour déduire l'égalité de toutes les composantes.

(c)

La valeur propre maximale est $\lambda_p = 1$: elle est présente d'après (a), et aucune valeur propre ne la dépasse d'après (b). Son espace propre est la droite $\mathbb{R}\mathbf{1}$.

II-5.(b) montre que les vecteurs qui minimisent F sont les vecteurs de cet espace propre de norme $\sqrt{\lambda_p} = 1$. Comme $\|\mathbf{1}\| = \sqrt{p}$, il y en a exactement deux :

$$x_+ = \frac{1}{\sqrt{p}}\mathbf{1}, \quad x_- = -\frac{1}{\sqrt{p}}\mathbf{1}.$$

Ces deux vecteurs opposés donnent le même endomorphisme $u_{x_+} = u_{x_-}$. Pour établir l'unicité parmi tous les éléments de $T(E)$, soit $u \in T(E)$ tel que $N^2(f - u) = m(f)$. Par I-3.(d), il existe $x \in E$ tel que $u = u_x$. II-1.(b) donne alors $F(x) = m(f)$, donc $x = x_+$ ou $x = x_-$. Ainsi u est nécessairement u_{x_+} . Cet endomorphisme réalise bien le minimum, ce qui prouve à la fois son existence et son unicité.

Dans la base canonique, sa matrice est :

$$P = x_+ x_+^T = \frac{1}{p} \mathbf{1}^t(\mathbf{1}) = \frac{1}{p} \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix}.$$

Pour $v = {}^t(v_1, \dots, v_p)$, elle agit par :

$$Pv = \frac{v_1 + \dots + v_p}{p} \mathbf{1}.$$

Elle remplace donc toutes les composantes par leur moyenne. Plus précisément, c'est la matrice de la projection orthogonale sur la droite $\mathbb{R}\mathbf{1}$ des vecteurs à composantes toutes égales. En effet, Pv appartient à cette droite et :

$$\langle v - Pv, \mathbf{1} \rangle = \sum_{j=1}^p v_j - p \frac{\sum_{j=1}^p v_j}{p} = 0.$$

Ce calcul vérifie directement le caractère orthogonal de la projection. La matrice P est symétrique, de rang un, et sa forme quadratique vaut

$$\langle Pv, v \rangle = \left(\sum_j v_j \right)^2 / p \geq 0,$$

conformément à son appartenance à $T(E)$.

Enfin, la valeur du minimum est :

$$m(f) = N^2(f) - 1 = \sum_{j=1}^{p-1} \lambda_j^2 = \sum_{i,j=1}^p m_{i,j}^2 - 1.$$

Pour $p = 1$, les hypothèses imposent $M = (1)$, les deux minimisateurs sont 1 et -1 , la matrice de l'unique meilleur approximant est $P = (1)$ et le minimum vaut 0. Toutes les conclusions restent donc valables dans ce cas limite.

Relevé des adaptations

- Partie II, intitulé précédant la question 1 : le PDF porte une seconde fois « Partie I ». Cet intitulé est rectifié en « Partie II », conformément à l'organisation annoncée dans le préambule. Aucun contenu mathématique ni numéro de question n'est modifié.