



PRÉPASIMPLE

prepasimple.fr

HEC

Mathématiques 1

2003

Corrigé de mathématiques approfondies

Mathématiques approfondies — Prépa ECG

Annales de concours · Corrigé

Nuages de points et approximation d'un nuage

Partie I : Étude d'un exemple

1.a) Somme des normes des projections

Le vecteur $d = u_1 + mu_2$ a pour coordonnées $(1, m)$ et vérifie $\|d\|^2 = 1 + m^2 > 0$. La formule de projection sur une droite donne :

$$v'_j = \frac{\langle v_j, d \rangle}{\|d\|^2} d, \quad \|v'_j\|^2 = \frac{\langle v_j, d \rangle^2}{1 + m^2}.$$

Les trois produits scalaires valent respectivement $1 + 2m$, $-3 - m$ et $2 - m$. Leur somme de carrés est :

$$(1 + 2m)^2 + (-3 - m)^2 + (2 - m)^2 = 14 + 6m + 6m^2.$$

La quantité demandée est donc :

$$S(m) = \frac{6m^2 + 6m + 14}{1 + m^2}.$$

1.b) Maximum et valeur de m_0

Méthode 1 — Dérivation et variations

La fonction S est dérivable sur $\mathbb{R}$, car son dénominateur est toujours strictement positif.

La règle de dérivation d'un quotient donne :

$$S'(m) = \frac{(12m + 6)(1 + m^2) - 2m(6m^2 + 6m + 14)}{(1 + m^2)^2} = \frac{-2(3m - 1)(m + 3)}{(1 + m^2)^2}.$$

Le signe de S' est négatif sur $] -\infty, -3[$, positif sur $] -3, 1/3[$ et négatif sur $]1/3, +\infty[$. De plus, $\lim_{m \rightarrow \pm\infty} S(m) = 6$, $S(-3) = 5$ et $S(1/3) = 15$.

Ces variations montrent que le maximum global est 15, atteint uniquement pour $m = 1/3$. Ainsi $m_0 = 1/3$ et $\lambda_1 = 15$.

Méthode 2 — Majoration exacte par un carré

À partir de l'expression de 1.a :

$$15 - S(m) = \frac{15(1 + m^2) - (6m^2 + 6m + 14)}{1 + m^2} = \frac{(3m - 1)^2}{1 + m^2} \geq 0.$$

Donc $S(m) \leq 15$ pour tout réel m . L'égalité équivaut à $3m - 1 = 0$, soit $m = 1/3$. Cela établit directement le même maximum et son unicité.

La dérivation donne aussi toutes les variations ; la seconde méthode fournit une preuve courte de la majoration globale.

2.a) Valeur propre maximale

La multiplication matricielle donne :

$$X^t X = \begin{pmatrix} 1^2 + (-3)^2 + 2^2 & 1 \cdot 2 + (-3)(-1) + 2(-1) \\ 1 \cdot 2 + (-3)(-1) + 2(-1) & 2^2 + (-1)^2 + (-1)^2 \end{pmatrix} = \begin{pmatrix} 14 & 3 \\ 3 & 6 \end{pmatrix}.$$

Le vecteur $u_1 + m_0 u_2$ a pour coordonnées $(1, 1/3)$ et est non nul. Or :

$$\begin{pmatrix} 14 & 3 \\ 3 & 6 \end{pmatrix} \begin{pmatrix} 1 \\ 1/3 \end{pmatrix} = \begin{pmatrix} 15 \\ 5 \end{pmatrix} = 15 \begin{pmatrix} 1 \\ 1/3 \end{pmatrix}.$$

Ainsi $\lambda_1 = 15$ est bien une valeur propre, associée à $u_1 + (1/3)u_2$.

2.b) Autre valeur propre

Méthode 1 — Construction d'un second vecteur propre

Le vecteur $w = (1, -3)$ est non nul et orthogonal à $(1, 1/3)$. On vérifie :

$$\begin{pmatrix} 14 & 3 \\ 3 & 6 \end{pmatrix} \begin{pmatrix} 1 \\ -3 \end{pmatrix} = \begin{pmatrix} 5 \\ -15 \end{pmatrix} = 5 \begin{pmatrix} 1 \\ -3 \end{pmatrix}.$$

Les deux vecteurs propres sont orthogonaux et non nuls, donc forment une base de E_2 . Dans cette base, l'endomorphisme est représenté par $\text{diag}(15, 5)$. Ses seules valeurs propres sont donc 15 et 5, et la seconde est strictement inférieure à λ_1 .

Méthode 2 — Théorème spectral et trace

La matrice $X^t X$ est réelle symétrique. Le théorème spectral lui assure une base orthonormale de vecteurs propres. L'une des valeurs propres est déjà 15. Si μ est l'autre valeur sur la diagonale dans cette base, l'invariance de la trace par similitude donne :

$$15 + \mu = \text{Tr}(X^t X) = 14 + 6 = 20.$$

Ainsi $\mu = 5 < 15$. Cette méthode est utile lorsque la trace est connue et qu'une seule valeur propre reste à déterminer ; la première donne en plus un vecteur propre explicite.

Partie II : Les axes principaux d'inertie d'un nuage

1.a) Diagonalisation et positivité des valeurs propres

La matrice $V = X^t X$ est réelle symétrique, car ${}^t V = X^t X = V$. Le théorème spectral réel assure sa diagonalisation dans une base orthonormale de E_p .

Pour tout $z \in E_p$, la définition du produit scalaire canonique donne :

$${}^t z V z = {}^t z X^t X z = ({}^t X z)({}^t X z) = \|{}^t X z\|^2 \geq 0.$$

Si $z \neq 0$ est un vecteur propre associé à une valeur propre λ , alors ${}^t z V z = \lambda \|z\|^2$. Puisque $\|z\|^2 > 0$, on obtient $\lambda \geq 0$.

On peut ordonner les valeurs de la diagonale de façon décroissante et réordonner la base orthonormale correspondante. Il existe donc une base orthonormale $(e_1, \dots, e_p)$ telle que $V e_j = \lambda_j e_j$, avec $\lambda_1 \geq \dots \geq \lambda_p \geq 0$.

1.b) Noyau, rang et base de F

Si ${}^t X z = 0$, alors $V z = X({}^t X z) = 0$. Réciproquement, si $V z = 0$, l'identité de 1.a donne $\|{}^t X z\|^2 = {}^t z V z = 0$, donc ${}^t X z = 0$. Ainsi :

$$\text{Ker } \Phi_V = \text{Ker } \Phi_{{}^t X}.$$

Les deux applications ont pour espace de départ E_p . Le théorème du rang donne donc $\text{rg}(V) = \text{rg}({}^t X)$. Le rang d'une matrice est égal à celui de sa transposée, et l'image de Φ_X est le sous-espace engendré par les colonnes de X , c'est-à-dire F . Ainsi :

$$\text{rg}(V) = \text{rg}(X) = \dim F = r.$$

Dans la base propre de 1.a, le rang de la matrice diagonale est le nombre de coefficients diagonaux non nuls. Il y en a exactement r . Comme les valeurs propres sont positives ou nulles et ordonnées :

$$\lambda_1 \geq \dots \geq \lambda_r > 0, \quad \lambda_{r+1} = \dots = \lambda_p = 0.$$

Enfin, pour $j \leq r$, $e_j = \lambda_j^{-1} V e_j$ appartient à l'image de V , qui est incluse dans l'image de X , donc dans F . La famille $(e_1, \dots, e_r)$ est orthonormale, donc libre, et possède $r = \dim F$ éléments : c'est une base de F .

On en déduit aussi $F^\perp = \text{Vect}(e_{r+1}, \dots, e_p)$; ces vecteurs appartiennent au noyau de V .

2.a) Expression matricielle de l'inertie

Puisque v est unitaire, la formule de la projection orthogonale sur D_v est

$P_{D_v}(c_j) = \langle v, c_j \rangle v$. Ainsi :

$$\|P_{D_v}(c_j)\|^2 = \langle v, c_j \rangle^2.$$

Le vecteur tXv a pour composantes ces produits scalaires. Par conséquent :

$$I(v) = \sum_{j=1}^n \langle v, c_j \rangle^2 = \|{}^tXv\|^2 = {}^tVv.$$

L'hypothèse $\|v\| = 1$ est essentielle à cette dernière expression de I . Pour un vecteur quelconque non nul, on aurait $I(v) = {}^tVv/\|v\|^2$.

2.b) Inerties des vecteurs propres

Chaque e_j est unitaire. Par 2.a et l'équation propre :

$$I(e_j) = {}^t(e_j)V e_j = \lambda_j {}^t(e_j)e_j = \lambda_j.$$

Ainsi $I(e_j) = \lambda_j$ pour tout j , et cette inertie est nulle pour $j > r$.

2.c) Sous-espaces successifs et maxima

Pour F_1 , la formule demandée est celle de 1.b : $F_1 = \text{Vect}(e_1, \dots, e_r)$. Supposons que $F_{j-1} = \text{Vect}(e_{j-1}, \dots, e_r)$ pour un indice $j \geq 2$. Un vecteur de cet espace s'écrit

$$\sum_{k=j-1}^r \alpha_k e_k.$$

Il est orthogonal à e_{j-1} exactement lorsque $\alpha_{j-1} = 0$, grâce à l'orthonormalité. Donc :

$$F_j = F_{j-1} \cap D_{e_{j-1}}^\perp = \text{Vect}(e_j, \dots, e_r).$$

Par récurrence, la formule vaut pour tous les indices $1 \leq j \leq r$, y compris lorsque $r = 1$.

Pour un vecteur unitaire

$$v = \sum_{k=1}^p \alpha_k e_k,$$

on a

$$\sum_{k=1}^p \alpha_k^2 = 1$$

et, par 2.a :

$$I(v) = \sum_{k=1}^p \lambda_k \alpha_k^2 \leq \lambda_1.$$

La valeur λ_1 est atteinte en e_1 , qui appartient à F_1 . Ainsi les maxima sur la sphère de E_p et sur celle de F_1 valent tous les deux $I(e_1) = \lambda_1$.

Plus généralement, si $v \in F_j$ est unitaire, seuls les coefficients d'indices $j, \dots, r$ peuvent être non nuls. Alors :

$$I(v) = \sum_{k=j}^r \lambda_k \alpha_k^2 \leq \lambda_j \sum_{k=j}^r \alpha_k^2 = \lambda_j.$$

Cette borne est atteinte en $e_j \in F_j$. Donc $I(e_j) = \max\{I(v) : v \in F_j, \|v\| = 1\}$. L'existence de chacun de ces maxima est établie par un vecteur qui réalise la borne, sans recours à un théorème de topologie générale.

3. Appartenance à F d'un vecteur maximisant l'inertie

Méthode 1 — Décomposition dans la base propre

Écrivons

$$w = \sum_{k=1}^p \alpha_k e_k,$$

avec

$$\sum_k \alpha_k^2 = 1.$$

Le maximum vaut $\lambda_1 > 0$ par 1.b et 2.c. L'égalité $I(w) = \lambda_1$ donne :

$$0 = \lambda_1 - I(w) = \sum_{k=1}^p (\lambda_1 - \lambda_k) \alpha_k^2.$$

Tous les termes sont positifs ou nuls. Pour $k > r$, $\lambda_k = 0$, donc ces termes valent $\lambda_1 \alpha_k^2$ et ne peuvent être nuls que si $\alpha_k = 0$. Ainsi $w \in \text{Vect}(e_1, \dots, e_r) = F$.

Méthode 2 — Projection sur F

Décomposons orthogonalement $w = a + b$, avec $a \in F$ et $b \in F^\perp$. Comme chaque $c_j \in F$, $\langle w, c_j \rangle = \langle a, c_j \rangle$ et :

$$I(w) = \sum_{j=1}^n \langle a, c_j \rangle^2.$$

Si $a = 0$, cette quantité est nulle, ce qui contredit $I(w) = \lambda_1 > 0$. Donc $a \neq 0$, et le vecteur $a/\|a\|$ est unitaire. Par 2.c :

$$I(w) = \|a\|^2 I(a/\|a\|) \leq \|a\|^2 \lambda_1.$$

Comme $I(w) = \lambda_1 > 0$, on en déduit $\|a\|^2 \geq 1$. Mais Pythagore donne $1 = \|a\|^2 + \|b\|^2$. Ainsi $b = 0$ et $w = a \in F$.

La première méthode décrit les coefficients des maximisateurs ; la seconde explique géométriquement pourquoi une composante orthogonale au nuage ne peut pas améliorer l'inertie.

4.a) Bases orthonormales

Les espaces G_j sont emboîtés et inclus dans F . Si $i < j$, $G_j \subset G_{i+1} \subset D_{\varepsilon_i}^\perp$. Comme $\varepsilon_j \in G_j$, les vecteurs ε_i et ε_j sont orthogonaux.

Les r vecteurs sont unitaires par hypothèse et appartiennent à F . Ils forment donc une famille orthonormale de $r = \dim F$ éléments, c'est-à-dire une base orthonormale de F .

Par 1.b, $(e_{r+1}, \dots, e_p)$ est une base orthonormale de $F^\perp$. La réunion des deux familles est donc une base orthonormale de E_p :

$$(\varepsilon_1, \dots, \varepsilon_r, e_{r+1}, \dots, e_p).$$

On dispose en outre de l'expression, issue de la définition récursive :

$$G_i = F \cap \bigcap_{k=1}^{i-1} D_{\varepsilon_k}^\perp = \text{Vect}(\varepsilon_i, \dots, \varepsilon_r).$$

Pour $i = 1$, l'intersection sans contrainte est simplement F .

4.b) Expression de J

Le vecteur tXv a pour composantes $\langle c_j, v \rangle$, et de même pour w . Par définition du produit scalaire canonique :

$$J(v, w) = \langle {}^tXv, {}^tXw \rangle = {}^tVX^tXw = {}^tVw = \langle v, \Phi_V(w) \rangle.$$

En particulier, J est bilinéaire et symétrique, et $J(v, v) = I(v)$ lorsque v est unitaire.

4.c) Rotation dans un plan

L'orthonormalité de v_1, v_2 et l'identité $\cos^2 t + \sin^2 t = 1$ montrent que $v(t) = (\cos t)v_1 + (\sin t)v_2$ est unitaire pour tout t . Par 2.a et la bilinéarité de J :

$$\varphi(t) = I(v_1) \cos^2 t + I(v_2) \sin^2 t + 2J(v_1, v_2) \cos t \sin t.$$

Les fonctions sinus et cosinus sont bornées ; cette formule donne par exemple $|\varphi(t)| \leq I(v_1) + I(v_2) + 2|J(v_1, v_2)|$. La fonction est donc majorée sur $\mathbb{R}$.

Elle est continue et 2π -périodique. Sur le segment $[0, 2\pi]$, le théorème des extrema pour une fonction continue assure qu'elle atteint un maximum. La périodicité montre que ce maximum est aussi un maximum sur tout $\mathbb{R}$.

Supposons maintenant qu'il soit atteint en 0. La fonction φ est dérivable sur l'ouvert $\mathbb{R}$; la condition nécessaire d'extremum donne $\varphi'(0) = 0$. En dérivant l'expression précédente :

$$\varphi'(t) = 2(I(v_2) - I(v_1)) \sin t \cos t + 2J(v_1, v_2)(\cos^2 t - \sin^2 t).$$

Donc $\varphi'(0) = 2J(v_1, v_2)$ et $J(v_1, v_2) = 0$.

La donnée de deux vecteurs unitaires orthogonaux de F implique ici $r \geq 2$. Si $r = 1$, ce choix n'est pas possible ; la question conditionnelle ne s'applique pas et la question suivante n'a aucun couple d'indices distincts à examiner.

4.d) Axes propres et ordre des valeurs propres

Soient $i < j \leq r$. D'après 4.a, ε_i et ε_j sont orthonormaux et appartiennent tous deux à G_i . Pour tout réel t , le vecteur $(\cos t)\varepsilon_i + (\sin t)\varepsilon_j$ est donc unitaire dans G_i .

Le système (S) implique que l'inertie de ce vecteur est inférieure ou égale à $I(\varepsilon_i)$, valeur atteinte en $t = 0$. En appliquant 4.c dans ce plan, on obtient $J(\varepsilon_i, \varepsilon_j) = 0$. Par symétrie de J , cela vaut aussi pour $j < i$. Lorsque $r = 1$, l'affirmation est vraie puisqu'il n'existe pas de couple d'indices distincts.

Posons $\mu_i = I(\varepsilon_i)$ pour $1 \leq i \leq r$. Dans la base orthonormale de 4.a, les coefficients de la matrice de Φ_V sont les produits scalaires des vecteurs de base avec les images des vecteurs de base. Le bloc associé aux ε_i est diagonal par le résultat précédent, de diagonale $\mu_1, \dots, \mu_r$.

Pour $j > r$, $Ve_j = 0$ par 1.b. Les coefficients des colonnes correspondantes sont donc nuls ; ceux des lignes correspondantes le sont aussi, puisque V est symétrique. La matrice est ainsi :

$$\text{diag}(\mu_1, \dots, \mu_r, 0, \dots, 0).$$

Cela montre déjà que chaque ε_i est un vecteur propre de valeur propre μ_i . Les nombres μ_i sont strictement positifs : pour un vecteur unitaire de F , l'expression dans la base $(e_1, \dots, e_r)$ donne une inertie au moins égale à $\lambda_r > 0$.

De plus, pour $i < j$, $\varepsilon_j \in G_i$. Le caractère maximal de ε_i dans G_i donne $\mu_i \geq \mu_j$. La liste $\mu_1, \dots, \mu_r$ est donc ordonnée de façon décroissante.

Les deux bases propres, celle des e_j et celle de 4.a, diagonalisent le même endomorphisme. Pour chaque valeur propre, le nombre de vecteurs de l'une ou l'autre base associés à cette valeur est la dimension de son espace propre : il est identique dans les deux bases. Les listes ordonnées de valeurs propres, avec ces répétitions, sont donc les mêmes. Ainsi $\mu_i = \lambda_i$ pour chaque $i \leq r$, et :

$$V\varepsilon_i = \lambda_i\varepsilon_i.$$

Cette conclusion n'exige pas que les valeurs propres positives soient toutes distinctes. En cas de valeurs répétées, les axes ne sont pas nécessairement uniques.

5. Covariances et matrice V

Question adaptée au programme actuel : la convention de covariance empirique, avec pondération uniforme et coefficient $1/n$, est précisée dans le sujet.

Les moyennes des deux séries x_k et x_ℓ sont nulles par hypothèse. La covariance empirique vaut donc :

$$\text{Cov}(x_k, x_\ell) = \frac{1}{n} \sum_{j=1}^n (x_{kj} - 0)(x_{\ell j} - 0) = \frac{1}{n} \sum_{j=1}^n x_{kj}x_{\ell j}.$$

Le coefficient d'indice (k, ℓ) de $V = X^tX$ est exactement

$$\sum_{j=1}^n x_{kj}x_{\ell j}.$$

Si C désigne la matrice des covariances :

$$C = \frac{1}{n}V, \quad V = nC.$$

Il s'agit de la covariance des séries sur une population de n individus uniformément pondérés, et non d'un estimateur dont le coefficient serait $1/(n-1)$.

Partie III : Une décomposition de la matrice X

Dans cette partie, les vecteurs e_i sont unitaires. Pour tout $z \in E_p$, la projection sur D_{e_i} s'écrit $P_{D_{e_i}}(z) = \langle z, e_i \rangle e_i$. Sa matrice est donc $\Pi_i = e_i {}^t(e_i)$. En particulier, Π_i est symétrique.

1. Somme des projections

Tout vecteur $z \in E_p$ possède la décomposition dans la base orthonormale $(e_1, \dots, e_p)$:

$$z = \sum_{i=1}^p \langle z, e_i \rangle e_i = \sum_{i=1}^p \Pi_i z.$$

L'égalité vaut pour tous les vecteurs, donc

$$\sum_{i=1}^p \Pi_i = I_p.$$

2. Produit de deux projections distinctes

Pour $i \neq j$, les vecteurs e_i, e_j sont orthogonaux. En utilisant la formule matricielle de la projection :

$$\Pi_i \Pi_j = e_i ({}^t(e_i) e_j) {}^t(e_j) = 0.$$

Le produit est la matrice nulle d'ordre p . Géométriquement, un vecteur de D_{e_j} a une projection nulle sur D_{e_i} . Pour la suite, la même formule donne aussi $\Pi_i^2 = \Pi_i$.

3. Décomposition du nuage

Pour $i > r$, $e_i \in F^\perp$ d'après II.1.b. Chaque colonne c_j de X appartient à F , donc sa projection sur D_{e_i} est nulle. Toutes les colonnes de $\Pi_i X$ sont donc nulles : $\Pi_i X = 0$.

En multipliant l'identité de la question 1 par X :

$$X = \left(\sum_{i=1}^p \Pi_i \right) X = \sum_{i=1}^r \Pi_i X.$$

4.a) Image de X_s

Pour tout $z \in \mathcal{M}_{n,1}(\mathbb{R})$, on a :

$$X_s z = \sum_{i=1}^s \Pi_i X z.$$

Chaque terme $\Pi_i X z$ appartient à D_{e_i} . Leur somme appartient donc à $\text{Vect}(e_1, \dots, e_s)$, ce qui établit l'inclusion demandée.

4.b) Images des vecteurs propres et rang

Comme $X^t X = V$ et $V e_j = \lambda_j e_j$:

$$X_s^t X e_j = \sum_{i=1}^s \Pi_i V e_j = \lambda_j \sum_{i=1}^s \Pi_i e_j.$$

L'orthonormalité donne $\Pi_i e_j = e_j$ si $i = j$, et 0 sinon. Ainsi :

$$X_s^t X e_j = \begin{cases} \lambda_j e_j & \text{si } 1 \leq j \leq s, \\ 0 & \text{si } s < j \leq p. \end{cases}$$

Pour $j \leq s \leq r$, $\lambda_j > 0$. Par conséquent :

$$e_j = X_s \left(\frac{1}{\lambda_j} {}^t X e_j \right) \in \text{Im } \Phi_{X_s}.$$

L'image de X_s contient donc $\text{Vect}(e_1, \dots, e_s)$, et l'inclusion inverse est celle de 4.a. Ainsi $\text{Im } \Phi_{X_s} = \text{Vect}(e_1, \dots, e_s)$ et $\text{rg}(X_s) = s$.

Partie IV : Une norme euclidienne de matrices carrées

1. Produit scalaire et norme

Pour $M = (m_{ij})$ et $N = (n_{ij})$ dans $\mathcal{M}_{p,n}(\mathbb{R})$, les coefficients diagonaux de $M^t N$ donnent :

$$\Theta(M, N) = \text{Tr}(M^t N) = \sum_{i=1}^p \sum_{j=1}^n m_{ij} n_{ij}.$$

Cette expression est bilinéaire en M, N , car chaque somme est finie, et symétrique, car $m_{ij} n_{ij} = n_{ij} m_{ij}$.

De plus,

$$\Theta(M, M) = \sum_{i,j} m_{ij}^2 \geq 0.$$

Cette somme est nulle si et seulement si chaque coefficient m_{ij} est nul, donc si et seulement si $M = 0$. La forme est ainsi définie positive : Θ est un produit scalaire réel.

Sa norme associée vérifie :

$$\|M\|^2 = \text{Tr}(M^t M) = \sum_{i=1}^p \sum_{j=1}^n m_{ij}^2.$$

En particulier, si les colonnes de M sont $m_1, \dots, m_n$, alors

$$\|M\|^2 = \sum_{j=1}^n \|m_j\|^2.$$

Le produit $M^t M$ est carré même lorsque M est rectangulaire.

2. Orthogonalité des composantes du nuage

Méthode 1 — Calcul par la trace

Les matrices Π_i et Π_j sont symétriques. Par définition de Θ et par l'identité de trace rappelée dans le sujet :

$$\Theta(\Pi_i X, \Pi_j X) = \text{Tr}(\Pi_i X^t X \Pi_j) = \text{Tr}(\Pi_j \Pi_i V).$$

Si $i \neq j$, III.2 donne $\Pi_j \Pi_i = 0$, donc cette quantité vaut zéro.

Si $i = j$, $\Pi_i^2 = \Pi_i$, donc $\Theta(\Pi_i X, \Pi_i X) = \text{Tr}(\Pi_i V)$. Puisque V est symétrique et $V e_i = \lambda_i e_i$, on a ${}^t(e_i) V = \lambda_i {}^t(e_i)$, d'où :

$$\Pi_i V = e_i {}^t(e_i) V = \lambda_i \Pi_i.$$

Enfin,

$$\text{Tr}(\Pi_i) = \sum_{k=1}^p (e_i)_k^2 = \|e_i\|^2 = 1.$$

Donc $\Theta(\Pi_i X, \Pi_i X) = \lambda_i$.

Méthode 2 — Calcul colonne par colonne

L'expression du produit scalaire en IV.1 donne :

$$\Theta(\Pi_i X, \Pi_j X) = \sum_{k=1}^n \langle \Pi_i c_k, \Pi_j c_k \rangle.$$

Si $i \neq j$, les deux vecteurs de chaque produit scalaire appartiennent à des droites orthogonales, donc tous les produits scalaires sont nuls.

Si $i = j$, la somme vaut

$$\sum_k \|P_{D_{e_i}}(c_k)\|^2 = I(e_i) = \lambda_i$$

par II.2.b.

Dans les deux méthodes :

$$\Theta(\Pi_i X, \Pi_j X) = \begin{cases} \lambda_i & \text{si } i = j, \\ 0 & \text{si } i \neq j. \end{cases}$$

La première méthode exploite les propriétés de la trace ; la seconde relie directement l'orthogonalité matricielle à celle des projections des points.

3. Erreur de l'approximation X_s

Par III.3 et la définition de X_s :

$$X - X_s = \sum_{i=s+1}^r \Pi_i X.$$

Les termes de cette somme sont deux à deux orthogonaux pour le produit scalaire Θ , d'après IV.2. En développant le carré de la norme, les termes croisés sont donc nuls :

$$\|X - X_s\|^2 = \sum_{i=s+1}^r \|\Pi_i X\|^2 = \sum_{i=s+1}^r \lambda_i.$$

Si $s = r$, la somme est vide et vaut zéro ; on retrouve $X_r = X$. Le même raisonnement appliqué à

$$X = \sum_{i=1}^r \Pi_i X$$

donne aussi :

$$\|X\|^2 = \sum_{i=1}^r \lambda_i, \quad \|X_s\|^2 = \sum_{i=1}^s \lambda_i.$$

Partie V : La meilleure approximation du nuage

Dans les questions où $1 \leq s \leq r - 1$, l'existence d'un tel entier suppose $r \geq 2$. La question 4.d couvre aussi le cas $r = 1$.

1. Base propre pour la matrice de l'erreur

Posons $B = (X - N)^t(X - N)$. Cette matrice réelle est symétrique, car la transposition du produit redonne B .

Le théorème spectral réel assure donc l'existence d'une base orthonormale de vecteurs propres. De plus, pour tout $z \in E_p$:

$${}^t_z B z = \|{}^t(X - N)z\|^2 \geq 0.$$

En appliquant cette identité à chaque vecteur propre non nul, on montre comme en

II.1.a que toutes les valeurs propres sont positives ou nulles. On peut les ordonner en réordonnant la base, de sorte que $\gamma_1 \geq \dots \geq \gamma_p \geq 0$ et $Ba_j = \gamma_j a_j$.

2.a) Intersection de sous-espaces

Notons $L_i = \text{Vect}(a_i, \dots, a_p)$. La famille est libre, donc $\dim L_i = p - i + 1$. La formule de dimension rappelée dans le sujet donne :

$$\dim(G \cap L_i) = \dim G + \dim L_i - \dim(G + L_i).$$

Or $G + L_i \subset E_p$, donc $\dim(G + L_i) \leq p$, et $\dim G \geq i$. Ainsi :

$$\dim(G \cap L_i) \geq i + (p - i + 1) - p = 1.$$

2.b) Choix d'un vecteur unitaire

Par 2.a, l'intersection $G \cap L_i$ contient un vecteur non nul z . Posons $u = z/\|z\|$; il est unitaire et appartient à cette intersection.

Puisque $u \in L_i$, il s'écrit

$$u = \sum_{j=i}^p \alpha_j a_j,$$

avec

$$\sum_{j=i}^p \alpha_j^2 = 1.$$

Par l'orthonormalité et les équations propres :

$$\|{}^t(X - N)u\|^2 = {}^t u B u = \sum_{j=i}^p \gamma_j \alpha_j^2 \leq \gamma_i \sum_{j=i}^p \alpha_j^2 = \gamma_i.$$

Le vecteur u satisfait donc la propriété demandée.

2.c) Comparaison des valeurs propres

Notons $L = \text{Vect}(e_1, \dots, e_{s+i})$. Puisque $i \leq r - s$, on a $s + i \leq r$, et $\dim L = s + i$.

L'application Φ_{tN} a pour espace de départ E_p . Le théorème du rang et l'égalité des rangs d'une matrice et de sa transposée donnent :

$$\dim \text{Ker } \Phi_{tN} = p - \text{rg}(^tN) = p - \text{rg}(N) \geq p - s.$$

Par la formule de dimension d'une intersection, puisque la somme des deux espaces est incluse dans E_p :

$$\dim H \geq (p - s) + (s + i) - p = i.$$

Appliquons 2.b avec $G = H$. Il existe un vecteur unitaire $u \in H$ tel que

$\|{}^t(X - N)u\|^2 \leq \gamma_i$. Comme $u \in \text{Ker } \Phi_{tN}$, on a ${}^tNu = 0$, donc ${}^t(X - N)u = {}^tXu$.

D'autre part, $u \in L$, donc

$$u = \sum_{j=1}^{s+i} \beta_j e_j$$

avec

$$\sum_{j=1}^{s+i} \beta_j^2 = 1.$$

Alors :

$$\|{}^tXu\|^2 = {}^tVu = \sum_{j=1}^{s+i} \lambda_j \beta_j^2 \geq \lambda_{s+i} \sum_{j=1}^{s+i} \beta_j^2 = \lambda_{s+i}.$$

En réunissant les deux bornes :

$$\lambda_{s+i} \leq \|X u\|^2 = \|(X - N)u\|^2 \leq \gamma_i.$$

Donc $\lambda_{s+i} \leq \gamma_i$.

3.a) Norme de l'erreur et trace

Par définition de la norme matricielle, $\|X - N\|^2 = \text{Tr}(B)$. Dans la base orthonormale de V.1, B est représentée par la matrice diagonale $\text{diag}(\gamma_1, \dots, \gamma_p)$.

L'invariance de la trace par similitude donne donc :

$$\|X - N\|^2 = \sum_{i=1}^p \gamma_i.$$

3.b) Borne inférieure de l'erreur

Les nombres γ_i sont positifs ou nuls par V.1. On peut donc omettre les termes d'indices supérieurs à $r - s$ dans une minoration. La question 2.c donne ensuite :

$$\|X - N\|^2 = \sum_{i=1}^p \gamma_i \geq \sum_{i=1}^{r-s} \gamma_i \geq \sum_{i=1}^{r-s} \lambda_{s+i} = \sum_{j=s+1}^r \lambda_j.$$

C'est la borne demandée, valable pour toute matrice N de rang inférieur ou égal à s .

3.c) Meilleure approximation

Par III.4.b, $\text{rg}(X_s) = s$: la matrice X_s appartient à la classe des matrices autorisées. Par IV.3, elle vérifie :

$$\|X - X_s\|^2 = \sum_{j=s+1}^r \lambda_j.$$

La question 3.b montre que toute matrice N de rang inférieur ou égal à s a une erreur au moins égale à cette valeur. Donc X_s atteint le minimum et réalise une meilleure approximation de X dans cette classe.

Comme les normes sont positives ou nulles et que la racine carrée est croissante, la même conclusion vaut pour la norme et pour son carré :

$$\min_{\text{rg}(N) \leq s} \|X - N\| = \left(\sum_{j=s+1}^r \lambda_j \right)^{1/2}.$$

La question établit l'existence d'une meilleure approximation, sans affirmer son unicité.

4.a) Inertie d'un sous-espace et norme matricielle

La colonne d'indice j de $\Pi_G X$ est $\Pi_G c_j = P_G(c_j)$. L'expression de la norme comme somme des normes au carré des colonnes, démontrée en IV.1, donne :

$$\|\Pi_G X\|^2 = \sum_{j=1}^n \|P_G(c_j)\|^2 = K(G).$$

Cette identité reste valable pour $G = \{0\}$, où les deux membres sont nuls.

4.b) Décomposition orthogonale de l'inertie

Pour chaque colonne c_j , la définition de la projection orthogonale donne $P_G(c_j) \in G$ et $c_j - P_G(c_j) \in G^\perp$. Par Pythagore :

$$\|c_j\|^2 = \|P_G(c_j)\|^2 + \|c_j - P_G(c_j)\|^2.$$

En sommant sur les colonnes et en utilisant IV.1 et 4.a :

$$\|X\|^2 = K(G) + \|X - \Pi_G X\|^2.$$

Ainsi $K(G) = \|X\|^2 - \|X - \Pi_G X\|^2$. En particulier, l'inertie totale $K(E_p)$ vaut $\|X\|^2$.

4.c) Maximum sous une contrainte de dimension

Supposons $1 \leq s \leq r - 1$ et $\dim G \leq s$. L'image de la matrice $\Pi_G X$ est incluse dans G , donc $\text{rg}(\Pi_G X) \leq \dim G \leq s$. La borne de V.3.b s'applique à $N = \Pi_G X$:

$$\|X - \Pi_G X\|^2 \geq \sum_{j=s+1}^r \lambda_j.$$

Par 4.b et IV.3 :

$$K(G) \leq \sum_{j=1}^r \lambda_j - \sum_{j=s+1}^r \lambda_j = \sum_{j=1}^s \lambda_j.$$

Pour $G_s = \text{Vect}(e_1, \dots, e_s)$, la famille est orthonormale et la projection sur G_s a pour matrice

$$\Pi_{G_s} = \sum_{j=1}^s \Pi_j.$$

On a donc $\Pi_{G_s} X = X_s$ et :

$$K(G_s) = \|X_s\|^2 = \sum_{j=1}^s \lambda_j.$$

L'espace G_s est de dimension s et atteint la borne. La valeur trouvée est donc le maximum de $K(G)$ pour $\dim G \leq s$.

4.d) Extension à tout $1 \leq s \leq p$

Méthode 1 — Distinction selon le rang du nuage

Si $s < r$, la conclusion est celle de 4.c.

Si $s \geq r$, toute projection orthogonale diminue la norme de chaque point, ou, de façon équivalente, 4.b donne $K(G) \leq \|X\|^2$. Or $G_s = \text{Vect}(e_1, \dots, e_s)$ contient $F = \text{Vect}(e_1, \dots, e_r)$. Tous les points du nuage sont donc déjà dans G_s , et leur projection ne les change pas. Ainsi :

$$K(G_s) = \|X\|^2 = \sum_{j=1}^r \lambda_j = \sum_{j=1}^s \lambda_j.$$

L'espace G_s a une dimension autorisée et réalise le maximum. Cette preuve couvre notamment $s = r$, $s = p$ et le cas $r = 1$.

Méthode 2 — Décomposition de l'inertie dans la base propre

Soit G de dimension $d \leq s$. Si $d = 0$, $K(G) = 0$ et la borne demandée est immédiate. Sinon, choisissons une base orthonormale $(g_1, \dots, g_d)$ de G , dont l'existence fait partie des propriétés des espaces euclidiens.

La formule de projection et Pythagore donnent, pour chaque point c_j :

$$\|P_G(c_j)\|^2 = \sum_{k=1}^d \langle c_j, g_k \rangle^2.$$

En échangeant deux sommes finies, puis en utilisant II.2.a pour chaque g_k :

$$K(G) = \sum_{k=1}^d I(g_k) = \sum_{k=1}^d {}^t(g_k) V g_k = \sum_{j=1}^p \lambda_j t_j,$$

avec

$$t_j = \sum_{k=1}^d \langle g_k, e_j \rangle^2 = \|P_G(e_j)\|^2.$$

Comme une projection orthogonale diminue la norme, $0 \leq t_j \leq \|e_j\|^2 = 1$.

Par décomposition de chaque vecteur unitaire g_k dans la base orthonormale des e_j :

$$\sum_{j=1}^p t_j = \sum_{k=1}^d \sum_{j=1}^p \langle g_k, e_j \rangle^2 = \sum_{k=1}^d \|g_k\|^2 = d \leq s.$$

Puisque les λ_j sont décroissants et positifs ou nuls :

$$\begin{aligned} \sum_{j=1}^s \lambda_j - K(G) &= \sum_{j=1}^s \lambda_j (1 - t_j) - \sum_{j=s+1}^p \lambda_j t_j \\ &\geq \lambda_s \left(\sum_{j=1}^s (1 - t_j) - \sum_{j=s+1}^p t_j \right) = \lambda_s (s - d) \geq 0. \end{aligned}$$

Lorsque $s = p$, la somme d'indices supérieurs à s est vide, et le raisonnement reste valide. On a donc

$$K(G) \leq \sum_{j=1}^s \lambda_j.$$

Pour G_s , la projection de e_j vaut e_j si $j \leq s$ et zéro sinon ; la borne est atteinte.

La première méthode réutilise directement la meilleure approximation obtenue précédemment ; la seconde explique le maximum d'inertie par le poids attribué à chaque direction propre.

Partie VI : Non multa, sed multum

1. Part de l'inertie contenue dans le plan principal

Par IV.3 et V.4, l'inertie totale vaut la somme des valeurs propres, et celle du plan $\text{Vect}(e_1, e_2)$ vaut la somme des deux premières. Avec les valeurs approximatives données :

$$K(E_8) \simeq 70 + 20 + 5 + 3 + 2 = 100, \quad K(\text{Vect}(e_1, e_2)) \simeq 70 + 20 = 90.$$

Le quotient de ces inerties vaut donc environ $90/100 = 0,90$. Le plan principal contient approximativement 90 % de l'inertie totale du nuage. Les deux axes représentent respectivement environ 70 % et 20 % de cette inertie.

2. Signification du nuage dessiné

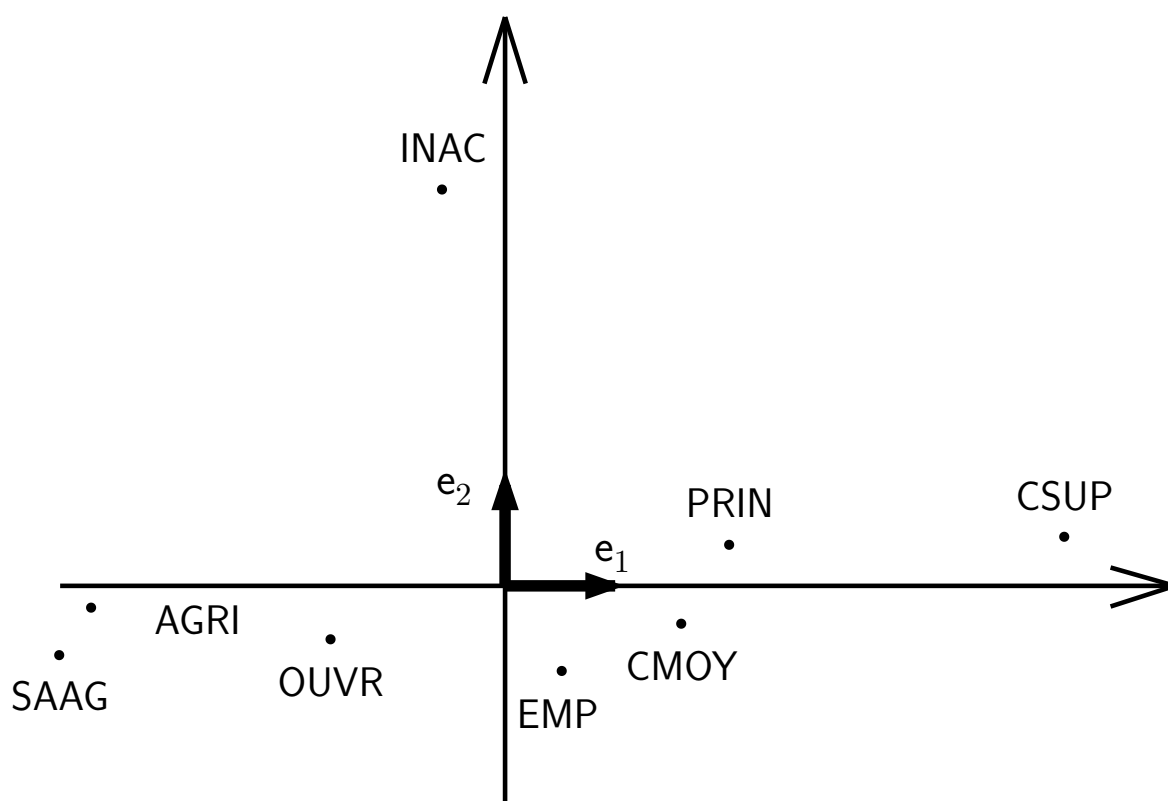
Chaque point du dessin est le projeté orthogonal $P_{\text{Vect}(e_1, e_2)}(c_j)$ d'une catégorie socioprofessionnelle. Dans la base orthonormale (e_1, e_2) du plan, ses deux coordonnées sont :

$$(\langle c_j, e_1 \rangle, \langle c_j, e_2 \rangle).$$

La matrice, dans la base canonique de E_8 , dont les colonnes sont ces huit projections est :

$$X_2 = (\Pi_1 + \Pi_2)X.$$

La figure représente donc le nuage obtenu en conservant les deux premiers axes principaux d'inertie. Par V.3.c, X_2 est une meilleure approximation du nuage initial par une matrice de rang inférieur ou égal à 2. Par V.4.c, ce plan est aussi celui qui conserve le plus d'inertie parmi les sous-espaces de dimension au plus 2.



Les noms portés par les points identifient les huit catégories. Le dessin utilise l'abréviation «OUVR» pour les ouvriers, tandis que le texte de l'énoncé utilise «OUV»; il s'agit dans les deux cas de la colonne c_7 .

L'erreur quadratique globale de projection vaut approximativement $5 + 3 + 2 = 10$, soit 10 % de l'inertie totale. Le dessin résume ainsi l'essentiel de la dispersion du nuage initial. Comme la figure source ne comporte pas de graduations numériques, aucune coordonnée numérique des points n'est déduite du schéma.

Relevé des adaptations

- II.5 : précision de la convention de covariance empirique, avec coefficient $1/n$ et pondération uniforme des individus, conformément au référentiel statistique.
- VI.1–VI.2 : la figure est retranscrite en SVG dans l'énoncé et le corrigé à partir de l'original, avec les positions des points et des axes mesurées en pixels. Les huit

points, les flèches, les vecteurs de base et les libellés sont conservés. La figure source n'étant pas graduée, aucune coordonnée numérique n'est inventée.