

MATHÉMATIQUES APPROFONDIES 1

(conçu par l'ESSEC)

n et p sont deux entiers naturels, $n \geq 2$ et $p \geq 1$.

Pour tout entier naturel non nul k :

- on identifie chaque élément de \mathbb{R}^k à la matrice colonne de ses coordonnées dans la base canonique.
- on identifie chaque endomorphisme de \mathbb{R}^k à sa matrice dans la base canonique.
- on rappelle que si X et Y sont deux éléments de \mathbb{R}^k , le produit scalaire canonique de X et Y vaut tXY où tX désigne la transposée de X .

On utilise les notations suivantes :

- \mathcal{S}_n désigne l'ensemble des matrices carrées d'ordre n symétriques réelles, et \mathcal{H}_n le sous-ensemble de \mathcal{S}_n formé des matrices dont tous les coefficients diagonaux sont nuls.
- U_n est le vecteur colonne de \mathbb{R}^n dont toutes les composantes valent 1 et $H_n = I_n - \frac{1}{n}U_n {}^tU_n$ où I_n désigne la matrice identité d'ordre n .
- **Dans tout le problème**, si $A \in \mathcal{H}_n$ et est d'élément générique $a_{i,j}$, c'est à dire $A = (a_{i,j})_{1 \leq i,j \leq n}$, on notera :

$$\forall i \in [1, n], \quad a_i = \frac{1}{n} \sum_{k=1}^n a_{i,k} = \frac{1}{n} \sum_{k=1}^n a_{k,i} \quad \text{et} \quad a = \frac{1}{n^2} \sum_{1 \leq i,j \leq n} a_{i,j} = \frac{1}{n} \sum_{i=1}^n a_i.$$

- On dit qu'une matrice $A \in \mathcal{H}_n$ d'élément générique $a_{i,j}$ est **une matrice euclidienne, de type (n, p)** , s'il existe n vecteurs X_1, \dots, X_n de \mathbb{R}^p tels que :

$$\forall (i, j) \in [1, n]^2, \quad a_{i,j} = \|X_i - X_j\|^2$$

la norme étant la norme euclidienne canonique sur \mathbb{R}^p .

- On dit alors que les vecteurs X_1, \dots, X_n sont associés à A et réciproquement.
- On notera $\mathcal{E}_{n,p}$ l'ensemble des matrices euclidiennes de type (n, p) et \mathcal{E}_n la réunion des ensembles $\mathcal{E}_{n,p}$ pour $p \in \mathbb{N}^*$.

On admet que $\dim(\mathcal{S}_n) = \frac{n(n+1)}{2}$ et que si l'on note \mathcal{D}_n le sous-espace vectoriel de \mathcal{S}_n des matrices diagonales, $\dim(\mathcal{D}_n) = n$.

L'objet du problème est d'établir une caractérisation des matrices euclidiennes qui sont utilisées en analyse de données.

Partie I - Une application linéaire de \mathcal{H}_n dans \mathcal{S}_n

1. a) Montrer que \mathcal{H}_n est un sous-espace vectoriel de \mathcal{S}_n et que $\mathcal{S}_n = \mathcal{H}_n \oplus \mathcal{D}_n$.
- b) En déduire la dimension de \mathcal{H}_n .
2. a) Soit $X \in \mathbb{R}^n$. Montrer que $\frac{1}{n}U_n^t U_n X$ est la projection orthogonale de X sur $\text{Vect}(U_n)$, le sous-espace vectoriel de \mathbb{R}^n engendré par U_n .
- b) En déduire que H_n est le projecteur orthogonal sur $(\text{Vect}(U_n))^\perp$.
- c) H_n est-elle inversible ? Quel est le rang de H_n ?

On définit l'application Φ de \mathcal{H}_n dans \mathcal{S}_n par : $\forall A \in \mathcal{H}_n, \Phi(A) = -\frac{1}{2}H_n A H_n$.

3. Vérifier que Φ est bien définie et qu'elle est linéaire.
4. Écrire une fonction Python `Phi(A)` qui retourne $\Phi(A)$ lorsque le tableau numpy `A` représente la matrice A .

On supposera avoir importé `numpy` avec l'instruction `import numpy as np`.

Que renvoie `Phi(np.array([[0, 2], [2, 0]]))` ?

5. a) Quels sont les coefficients de la matrice H_n ?
- b) Soit $A = (a_{i,j})$ appartenant à \mathcal{H}_n et $a'_{i,j}$ l'élément générique de la matrice AH_n . Établir pour tout $(i, j) \in \llbracket 1, n \rrbracket^2$, l'égalité : $a'_{i,j} = a_{i,j} - a_i$.
- c) En déduire que $b_{i,j}$, l'élément générique de $\Phi(A)$, vérifie :

$$\forall (i, j) \in \llbracket 1, n \rrbracket^2, b_{i,j} = \frac{1}{2}(a_i + a_j - a_{i,j} - a)$$

- d) *Un exemple.* Calculer $\Phi(A)$ si $A = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 2 & 3 & 0 \end{pmatrix}$.

6. Soit $A \in \ker(\Phi)$ d'élément générique $a_{i,j}$.
 - a) Montrer que pour tout $i \in \llbracket 1, n \rrbracket, a_i = \frac{a}{2}$.
 - b) En déduire que $a = 0$ puis que A est la matrice nulle.
 - c) Que peut-on en déduire pour Φ ?
7. On considère l'ensemble \mathcal{K}_n des matrices M de \mathcal{S}_n telle que $MU_n = 0$.
 - a) Montrer que $\text{Im}(\Phi) \subset \mathcal{K}_n$.
 - b) Montrer que \mathcal{K}_n et \mathcal{D}_n sont en somme directe.
 - c) En déduire que $\dim(\mathcal{K}_n) = \frac{n(n-1)}{2}$ et que $\text{Im}(\Phi) = \mathcal{K}_n$.

Partie II - Propriétés des matrices euclidiennes

Dans cette partie, $A = (a_{i,j})$ appartient à $\mathcal{E}_{n,p}$ et est associée aux vecteurs X_1, \dots, X_n appartenant à \mathbb{R}^p .

On démontre dans cette partie que $\Phi(A)$ ne possède que des valeurs propres positives parmi lesquelles se trouve 0.

8. Un exemple lorsque $n = 3$.

A est la matrice associée aux vecteurs : $X_1 = {}^t(1 \ 0 \ 1)$, $X_2 = {}^t(1 \ 1 \ 0)$ et $X_3 = {}^t(1 \ 1 \ 1)$. Déterminer A , $\Phi(A)$, puis les valeurs propres de $\Phi(A)$.

9. Un autre exemple. Dans cette question A a tous ses coefficients non diagonaux qui valent 1.

a) En utilisant une base orthonormée de \mathbb{R}^n , montrer que A appartient à $\mathcal{E}_{n,n}$.

b) Si $b_{i,j}$ désigne l'élément générique de $\Phi(A)$, montrer que pour tout $(i, j) \in \llbracket 1, n \rrbracket^2$, $b_{i,j} = \begin{cases} -\frac{1}{2n} & \text{si } i \neq j \\ \frac{n-1}{2n} & \text{si } i = j \end{cases}$.

c) En déduire que $\Phi(A) = \frac{1}{2}H_n$ puis que $\text{rg}(\Phi(A)) = n - 1$ et $\text{Sp}(\Phi(A)) = \{0, \frac{1}{2}\}$.

10. Montrer que pour tout entier $k \geq p$, $A \in \mathcal{E}_{n,k}$.

11. a) Montrer que pour tout vecteur $Y \in \mathbb{R}^p$, les vecteurs $X_1 - Y, \dots, X_n - Y$ sont aussi associés à A .

b) En déduire que l'on peut choisir les X_i de sorte qu'ils soient associés à A et vérifient : $\sum_{i=1}^n X_i = 0$ (1).

On suppose dans la suite de cette partie que cette relation (1) est vérifiée.

c) Exprimer $a_{i,j}$ en fonction de $\|X_i\|^2$, $\|X_j\|^2$ et tX_iX_j . En déduire que :

$$a_i = \|X_i\|^2 + \frac{1}{n} \sum_{j=1}^n \|X_j\|^2, \text{ puis que } a = \frac{2}{n} \sum_{j=1}^n \|X_j\|^2$$

d) On note $b_{i,j}$ l'élément générique de $\Phi(A)$. En utilisant la question 5.c), établir l'égalité :

$$\forall (i, j) \in \llbracket 1, n \rrbracket^2, \quad b_{i,j} = {}^tX_iX_j.$$

puis que, si l'on note X la matrice de $\mathcal{M}_{p,n}(\mathbb{R})$ dont les colonnes sont X_1, \dots, X_n , on a $\Phi(A) = {}^tXX$.

e) Montrer que $\text{rg}(\Phi(A)) = \text{rg}(X)$ et en déduire que $\text{rg}(\Phi(A)) \leq p$.

12. Montrer que les valeurs propres de $\Phi(A)$ sont positives ou nulles et que 0 est une de ces valeurs propres.

Partie III - Caractérisation de Schönberg-Young-Householder des matrices euclidiennes

On considère dans les questions 13 et 14 cette partie $A \in \mathcal{H}_n$.

On note alors $G = \Phi(A)$, $p = \text{rg}(G)$ et on suppose que $\text{Sp}(G) \subset \mathbb{R}^+$ et $p \neq 0$.

On démontre que sous ces hypothèses A est euclidienne. On établit ensuite une caractérisation des matrices euclidiennes.

13. a) Montrer qu'il existe des réels strictement positifs $\alpha_1, \dots, \alpha_p$ et une matrice orthogonale P tels que $G = P\Delta^2({}^tP)$ où Δ est la matrice diagonale dont les p premiers éléments diagonaux sont $\alpha_1, \dots, \alpha_p$ et les autres coefficients sont nuls.

b) Montrer que les $n - p$ dernières lignes de $\Delta {}^tP$ sont nulles.

c) On considère X la matrice de $\mathcal{M}_{p,n}(\mathbb{R})$ dont les lignes sont les p premières lignes de $\Delta {}^tP$. Montrer que $G = {}^tXX$.

14. On note X_j la j -ème colonne de X .

a) Montrer que l'élément générique de G est tX_iX_j .

b) En utilisant que $GU_n = 0$, en déduire que $\sum_{j=1}^n X_j = 0$.

c) En conclure que A est la matrice euclidienne associée aux points X_1, \dots, X_n et que $A \in \mathcal{E}_{n,p}$.

15. a) En utilisant les résultats de la partie 2 et ceux de cette partie, démontrer que l'on a, pour toute $A \in \mathcal{H}_n$, l'équivalence suivante :

$$A \in \mathcal{E}_n \text{ si et seulement si } \text{Sp}(\Phi(A)) \subset \mathbb{R}^+$$

et que, si $A \in \mathcal{E}_n$ n'est pas la matrice nulle, $p = \text{rg}(\Phi(A))$ est la valeur minimale de p pour laquelle on a $A \in \mathcal{E}_{n,p}$.

- b) En déduire que $\mathcal{E}_n = \mathcal{E}_{n,n-1}$.

c) *Application.* En utilisant la matrice de la question 9 et le résultat précédent, montrer que dans \mathbb{R}^{n-1} , il existe n vecteurs X_1, \dots, X_n tels que $\|X_i - X_j\| = 1$ pour tout $i \neq j$ mais pas $n+1$ ou plus.

16. On importe la bibliothèque `linalg` de `numpy` en exécutant, `import numpy.linalg as al` et on saisit une matrice `A` qui représente une matrice de \mathcal{H}_n .

Donner une expression Python comportant `Phi(A)` et la fonction `al.eig` qui vaut `True` lorsque `A` représente une matrice euclidienne et `False` sinon.

17. Soit $M \in \mathcal{S}_n$. Montrer que $\text{Sp}(M) \subset \mathbb{R}^+$ si et seulement si pour tout élément X de \mathbb{R}^n , ${}^tXMX \geq 0$.

En déduire que, pour toute $A \in \mathcal{H}_n$:

$$A \in \mathcal{E}_n \text{ si et seulement si pour tout vecteur } X \in \mathbb{R}^n, {}^tX\Phi(A)X \geq 0.$$

18. Dans cette question, on suppose comme dans la question 5.d) que $A = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 2 & 3 & 0 \end{pmatrix}$ et on pose $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$.

a) Montrer que ${}^tX\Phi(A)X = \frac{1}{3}(x^2 + 2y^2 + 3z^2 - 2xz - 4yz)$.

b) En déduire que A est euclidienne.

Partie IV - Une condition nécessaire et une condition suffisante utilisant la trace

On établit dans cette partie une condition nécessaire et une condition suffisante, assez faciles à vérifier, pour qu'une matrice de \mathcal{H}_n soit euclidienne.

Pour toute matrice $M \in \mathcal{S}_n$ semblable à une matrice diagonale dont les valeurs propres non nulles sont $\lambda_1 \geq \dots \geq \lambda_r$ où $r \geq 1$, on pose :

$$m = \frac{\text{Tr}(M)}{r} = \frac{1}{r} \sum_{k=1}^r \lambda_k \text{ et } s = \sqrt{\frac{\text{Tr}(M^2)}{r} - m^2}$$

19. Soit $M \in \mathcal{S}_n$. On pose $\Lambda = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_r \end{pmatrix}$ et on rappelle que $U_r = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$.

a) Montrer que $m = \frac{1}{r} {}^tU_r \Lambda$ puis que $m^2 \leq \frac{\text{Tr}(M^2)}{r}$. En déduire que la définition de s est toujours possible.

b) Montrer que $s^2 = \frac{1}{r} {}^t\Lambda H_r \Lambda$, où $H_r = I_r - \frac{1}{r} U_r {}^tU_r$.

c) Soit $Y \in \mathbb{R}^r$, montrer que :

$$\left| {}^tY H_r \Lambda \right| \leq \sqrt{{}^tY H_r Y} \sqrt{{}^t\Lambda H_r \Lambda}$$

en déduire que :

$$-s\sqrt{r}\sqrt{{}^tY H_r Y} \leq {}^tY H_r \Lambda \leq s\sqrt{r}\sqrt{{}^tY H_r Y}$$

d) On suppose que Y est le r -ième vecteur de la base canonique de \mathbb{R}^r .

Montrer que ${}^tYH_rY = \frac{r-1}{r}$ et ${}^tYH_r\Lambda = \lambda_r - m$.

En déduire que :

$$\lambda_r \geq m - s\sqrt{r-1}$$

e) Montrer que : $r^2(m - \lambda_r)^2 \geq \sum_{k=1}^r (\lambda_k - \lambda_r)^2$ et que $\sum_{k=1}^r (\lambda_k - \lambda_r)^2 = rs^2 + r(m - \lambda_r)^2$.

En conclure que si $r \geq 2$:

$$\lambda_r \leq m - \frac{s}{\sqrt{r-1}}$$

20. On suppose, dans cette question, que $n \geq 3$ et l'on considère $A \in \mathcal{H}_n$.

On utilise les notations de la question précédente avec $M = -H_nAH_n = 2\Phi(A)$, en supposant que le rang de M , noté r , est supérieur ou égal à 1.

a) Montrer que $m = \frac{1}{nr} {}^tU_nAU_n$ et que $s^2 = \frac{1}{r}\text{Tr}(A^2) - \frac{2}{nr} {}^tU_nA^2U_n + (r-1)m^2$.

b) Montrer que si A est euclidienne alors $\frac{2}{n} {}^tU_nA^2U_n \geq \text{Tr}(A^2)$.

c) On rappelle que $r \leq n-1$. Montrer que si l'on a,

$$\frac{2}{n} {}^tU_nA^2U_n - \frac{n-3}{n^2(n-2)} ({}^tU_nAU_n)^2 \geq \text{Tr}(A^2)$$

alors A est euclidienne.

21. *Exemples*

a) On suppose dans cette question que $n = 3$ et $A \in \mathcal{H}_3$.

Montrer que $A \in \mathcal{E}_3$ si et seulement si $\frac{2}{3} {}^tU_3A^2U_3 \geq \text{Tr}(A^2)$. En déduire pour quelles valeurs de a , la

matrice $\begin{pmatrix} 0 & a & 1 \\ a & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$ est euclidienne.

b) Montrer que la matrice de la question 9 satisfait la condition suffisante de la question 20.c).

22. Écrire une fonction Python `CsEuclidienne(A)` qui renvoie `True` si A , représentant une matrice de symétrique à élément diagonaux tous nuls, vérifie la condition suffisante de la question 20.c) et `False` sinon.