

EXAMEN D'ANALYSE DE DONNÉES (ADD)

Niveau : L3 EA Session : Normale Année : 24/2025 Durée : 01h 30 min

Exercice 01 (6 pts) :

La Matrice $\mathbf{A}$ suivante, est-elle diagonalisable ?

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 2 \\ 1 & 2 & -1 \\ -1 & 1 & 4 \end{pmatrix}$$

Exercice 02 (14 pts) :

On considère le tableau suivant :

$j \backslash i$	1	2	3	4	5	6
1	1	1	2	0	3	1
2	0	1	2	1	1	3
3	1	2	4	1	4	4

Où la j ème ligne désigne la variable $\mathbf{X}_j$, et la i ème colonne désigne l'individu $\mathbf{e}_i$. On effectue l'ACP de $\mathbf{X}$ (matrice associée au tableau) sur matrice variance-covariance ($\mathbf{V}$), chaque individu étant muni du poids $\mathbf{p} = \frac{1}{6}$.

1. —Donner la matrice $\mathbf{X}$ associée au tableau, puis calculer le centre de gravité $\mathbf{g}$ et la matrice $\mathbf{Y}$ associée au tableau centré.

2. —Soit $\mathbf{V}$ la matrice variance-covariance du tableau : $\mathbf{V} = \frac{1}{18} \begin{pmatrix} 16 & 1 & 17 \\ 1 & 16 & 17 \\ 17 & 17 & 34 \end{pmatrix}$

Calculer l'inertie totale ($\mathbf{I}_g$) et les valeurs propres (λ_1, λ_2 , et λ_3). Déterminer $\mathbf{u}_1$ le vecteur propre de $\mathbf{V}$ associé à λ_1 de norme 1 (normé) qui dirige le premier axe factoriel ($\mathbf{F}_1$), et $\mathbf{u}_2$ le vecteur propre de la même matrice associé à λ_2 de norme 1 qui dirige le deuxième axe factoriel ($\mathbf{F}_2$)

3. —Quelle est la contribution absolue de l'axe $\mathbf{F}_1$ à l'inertie du nuage ? Quelle est la contribution absolue de l'axe $\mathbf{F}_2$ à l'inertie du nuage ? Représenter graphiquement l'ensemble des 6 individus dans le plan ($\mathbf{F}_1, \mathbf{F}_2$).

Corrigé :

Exercice 01 :

Définition : on dit que A est diagonalisable si elle possède n vecteurs propres $u_1, u_2, \dots, u_n$ linéairement indépendants.
Nous avons :

$$A = \begin{pmatrix} 1 & 2 & 2 \\ 1 & 2 & -1 \\ -1 & 1 & 4 \end{pmatrix} \quad \text{et :} \quad (A - \lambda.I_3) = \begin{pmatrix} 1-\lambda & 2 & 2 \\ 1 & 2-\lambda & -1 \\ -1 & 1 & 4-\lambda \end{pmatrix}$$

$$\det(A - \lambda.I_3) = (1-\lambda) \begin{vmatrix} 2-\lambda & -1 \\ 1 & 4-\lambda \end{vmatrix} - 2 \begin{vmatrix} 1 & -1 \\ -1 & 4-\lambda \end{vmatrix} + 2 \begin{vmatrix} 1 & 2-\lambda \\ -1 & 1 \end{vmatrix}$$

$$\det(A - \lambda.I_3) = (1-\lambda)(2-\lambda)(4-\lambda) + (1-\lambda) = (1-\lambda)(3-\lambda)^2$$

L'équation $(1-\lambda)(3-\lambda)^2$ a pour solutions les trois racines : $\lambda_1 = 1, \lambda_2 = \lambda_3 = 3$

- Les vecteurs propres associés à la valeur propre $\lambda_1 = 1$ sont ceux qui vérifient :

$$\begin{pmatrix} 0 & 2 & 2 \\ 1 & 1 & -1 \\ -1 & 1 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0 \iff \begin{cases} 2y + 2z = 0 \\ x + y - z = 0 \\ -x + y + 3z = 0 \end{cases} \iff y = -z \quad \text{et} \quad x = 2z \quad \text{soit :} \quad \mathbf{u}_1 = (2, -1, 1)$$

- Les vecteurs propres associés aux valeurs propres $\lambda_2 = \lambda_3 = 3$ sont ceux qui vérifient :

$$\begin{pmatrix} -2 & 2 & 2 \\ 1 & -1 & -1 \\ -1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0 \iff \begin{cases} -2x + 2y + 2z = 0 \\ x - y - z = 0 \\ -x + y + z = 0 \end{cases} \iff x = y + z \quad \text{soient :} \quad \mathbf{u}_2(1, 1, 0) \quad \text{et} \quad \mathbf{u}_3 = (1, 0, 1)$$

$S = \{u_1, u_2, u_3\} = \{(2, -1, 1), (1, 1, 0), (1, 0, 1)\}$, S est un ensemble maximal de vecteurs propres linéairement indépendants. Et comme S est une base de $\mathbb{R}^3$, la matrice A est diagonalisable. En utilisant la base S , A est représentée par la matrice diagonale $D = \text{diag}(1, 3, 3)$. Soit P la matrice de colonnes $u_1; u_2; u_3$ alors :

$$D = P^{-1}AP = \frac{1}{2} \begin{pmatrix} 1 & -1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 2 \\ 1 & 2 & -1 \\ -1 & 1 & 4 \end{pmatrix} \begin{pmatrix} 2 & 1 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

Exercice 02 :

1. La matrice X associée au tableau est :

$$X = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 2 \\ 2 & 2 & 4 \\ 0 & 1 & 1 \\ 3 & 1 & 4 \\ 1 & 3 & 4 \end{pmatrix} \quad ; \text{ matrice de 6 lignes (individus) et de 3 colonnes (variables)}$$

2. Le centre de gravité g est égal par définition à : $g = {}^t X D \mathbf{1}_n$ où $\mathbf{1}_n$ désigne le vecteur colonne de $\mathbb{R}^n$ dont toutes les coordonnées sont égales à 1 et D la matrice des poids $\frac{1}{n} \mathbb{I}_n$ ($\mathbb{I}_n$ représente la matrice identité avec $n = 6$ dans ce cas) :

$$g = \begin{pmatrix} 1 & 1 & 2 & 0 & 3 & 1 \\ 0 & 1 & 2 & 1 & 1 & 3 \\ 1 & 2 & 4 & 1 & 4 & 4 \end{pmatrix} \begin{pmatrix} \frac{1}{6} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{6} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{6} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{6} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{6} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{6} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \bar{X}_1 = \frac{4}{3} \\ \bar{X}_2 = \frac{4}{3} \\ \bar{X}_3 = \frac{8}{3} \end{pmatrix}$$

3. La matrice $\mathbf{Y}$ associée au tableau centré est définie par : $\mathbf{Y} = \mathbf{X} - \mathbf{1}_n \cdot {}^t \mathbf{g}$ où $y_{ij} = x_{ij} - \bar{X}_j$

$$\mathbf{Y} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 2 \\ 2 & 2 & 4 \\ 0 & 1 & 1 \\ 3 & 1 & 4 \\ 1 & 3 & 4 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} \frac{4}{3} & \frac{4}{3} & \frac{8}{3} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 2 \\ 2 & 2 & 4 \\ 0 & 1 & 1 \\ 3 & 1 & 4 \\ 1 & 3 & 4 \end{pmatrix} - \begin{pmatrix} \frac{4}{3} & \frac{4}{3} & \frac{8}{3} \\ \frac{4}{3} & \frac{4}{3} & \frac{8}{3} \\ \frac{4}{3} & \frac{4}{3} & \frac{8}{3} \\ \frac{4}{3} & \frac{4}{3} & \frac{8}{3} \\ \frac{4}{3} & \frac{4}{3} & \frac{8}{3} \\ \frac{4}{3} & \frac{4}{3} & \frac{8}{3} \end{pmatrix} = \begin{pmatrix} -\frac{1}{3} & -\frac{4}{3} & -\frac{5}{3} \\ -\frac{1}{3} & -\frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & \frac{4}{3} \\ -\frac{4}{3} & -\frac{1}{3} & -\frac{5}{3} \\ \frac{5}{3} & -\frac{1}{3} & \frac{4}{3} \\ -\frac{1}{3} & \frac{5}{3} & \frac{4}{3} \end{pmatrix}$$

4. Par définition la matrice variance-covariance est une matrice carrée de dimension $(p \times p)$ où :

$$\mathbf{V} = {}^t \mathbf{Y} \mathbf{D} \mathbf{Y} = {}^t \mathbf{X} \mathbf{D} \mathbf{X} - \mathbf{g} \mathbf{g}^t \text{ ou } \mathbf{V} = \frac{1}{n} {}^t \mathbf{Y} \mathbf{Y} \text{ soit : } \mathbf{V} = \frac{1}{6} {}^t \mathbf{Y} \mathbf{Y}$$

$$\mathbf{V} = \frac{1}{6} \begin{pmatrix} -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} & -\frac{4}{3} & \frac{5}{3} & -\frac{1}{3} \\ -\frac{4}{3} & -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} & \frac{5}{3} \\ -\frac{5}{3} & -\frac{2}{3} & \frac{4}{3} & -\frac{5}{3} & \frac{4}{3} & \frac{4}{3} \end{pmatrix} \begin{pmatrix} -\frac{1}{3} & -\frac{4}{3} & -\frac{5}{3} \\ -\frac{1}{3} & -\frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & \frac{4}{3} \\ -\frac{4}{3} & -\frac{1}{3} & -\frac{5}{3} \\ \frac{5}{3} & -\frac{1}{3} & \frac{4}{3} \\ -\frac{1}{3} & \frac{5}{3} & \frac{4}{3} \end{pmatrix} = \begin{pmatrix} \frac{8}{9} & \frac{1}{18} & \frac{17}{18} \\ \frac{1}{18} & \frac{8}{9} & \frac{17}{18} \\ \frac{17}{18} & \frac{17}{18} & \frac{17}{9} \end{pmatrix} = \frac{1}{18} \begin{pmatrix} 16 & 1 & 17 \\ 1 & 16 & 17 \\ 17 & 17 & 34 \end{pmatrix}$$

5. L'inertie totale $\mathbf{I}_g$ du nuage de points associée à la matrice du tableau $\mathbf{X}$ est la moyenne pondérée des carrés des distances des points au centre de gravité $\mathbf{g}$:

$$\mathbf{I}_g = \sum_{i=1}^n p_i d^2(e_i, g) = \sum_{i=1}^n p_i \sum_{j=1}^p (x_{ij} - g)^2 = \sum_{j=1}^p \left[\frac{1}{n} \sum_{i=1}^n \left(x_{ij} - \frac{1}{n} \sum_{i=1}^n x_{ij} \right)^2 \right] = \sum_{j=1}^p \text{Var}(\mathbf{V}_j)$$

$$\mathbf{I}_g = \sum_{j=1}^3 \text{Var}(\mathbf{V}_j) = \text{Trace}(\mathbf{V}) = \frac{8}{9} + \frac{8}{9} + \frac{17}{9} = \frac{11}{3} \approx 3.6667$$

6. Le calcul des valeurs propres & des vecteurs propres de la matrice $\mathbf{V}$. Par définition, les valeurs propres de la matrice $\mathbf{V}$ sont les solutions en λ del l'équation $\det(\mathbf{V} - \lambda \mathbf{I}_n) = 0$. Et les vecteurs propres correspondants sont les éléments du noyau de $\mathbf{V} - \lambda \mathbf{I}_n$. Nous avons donc :

$$(\mathbf{V} - \lambda \mathbf{I}_3) = \frac{1}{18} \begin{pmatrix} 16 & 1 & 17 \\ 1 & 16 & 17 \\ 17 & 17 & 34 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \frac{1}{18} \begin{pmatrix} 16 - \lambda & 1 & 17 \\ 1 & 16 - \lambda & 17 \\ 17 & 17 & 34 - \lambda \end{pmatrix}$$

Et : $\det[\alpha(\mathbf{V} - \lambda \mathbf{I}_3)] = \alpha \cdot \det(\mathbf{V} - \lambda \mathbf{I}_3)$ alors :

$$\det(\mathbf{V} - \lambda \mathbf{I}_3) = (16 - \lambda) \begin{vmatrix} (16 - \lambda) & 17 \\ 17 & (34 - \lambda) \end{vmatrix} - \begin{vmatrix} 1 & 17 \\ 17 & (34 - \lambda) \end{vmatrix} + 17 \begin{vmatrix} 1 & (16 - \lambda) \\ 17 & 17 \end{vmatrix}$$

$$\det(\mathbf{V} - \lambda \mathbf{I}_3) = (16 - \lambda)^2 (34 - \lambda) - (34 - \lambda) - 2 \times 17^2 (15 - \lambda)$$

$$\det(\mathbf{V} - \lambda \mathbf{I}_3) = (34 - \lambda)(15 - \lambda)(17 - \lambda) - 2 \times 17^2 (15 - \lambda) = (51 - \lambda)(15 - \lambda)(- \lambda)$$

Cette dernière équation a pour solution : $\lambda_1 = 51$, $\lambda_2 = 15$ et $\lambda_3 = 0$. Et en multipliant chaque valeur propre par $\alpha = \frac{1}{18}$ nous obtenons :

$$\lambda_1 = \frac{1}{18} \times 51 = \frac{51}{18} = \frac{17}{6} \quad \text{et} : \quad \lambda_2 = \frac{1}{18} \times 15 = \frac{15}{18} = \frac{5}{6}$$

7. Nous déterminons juste après, u_1 le vecteur propre de $\mathbf{V}$ associé à $\lambda_1 = \frac{17}{6}$ de norme unitaire 1 (normé) qui dirige le premier axe factoriel ($\mathbf{F}_1$) et le vecteur propre u_2 de la même matrice associée à la valeur propre $\lambda_2 = \frac{5}{6}$ de norme 1 qui dirige le deuxième axe factoriel ($\mathbf{F}_2$) :

• pour $\lambda_1 = \frac{17}{6}$ nous avons $(\mathbf{V} - \frac{17}{6} \mathbf{I}_3)u_1 = 0$ et $u_1 = (x, y, z)$ c'est-à-dire :

$$\begin{pmatrix} -\frac{35}{18} & \frac{1}{18} & \frac{17}{18} \\ \frac{1}{18} & -\frac{35}{18} & \frac{17}{18} \\ \frac{17}{18} & \frac{17}{18} & -\frac{17}{18} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0 \iff \begin{cases} -\frac{35}{18}x + \frac{1}{18}y + \frac{17}{18}z = 0 \\ \frac{1}{18}x - \frac{35}{18}y + \frac{17}{18}z = 0 \iff 2x = 2y = z \quad \text{soit : } \mathbf{u}_1 = (\frac{1}{2}, \frac{1}{2}, 1) \\ \frac{17}{18}x + \frac{17}{18}y - \frac{35}{18}z = 0 \end{cases}$$

• Nous procédons au même calcul avec la deuxième valeur propre $\lambda_2 = \frac{5}{6}$ où $(\mathbf{V} - \frac{5}{6} \mathbf{I}_3)u_2 = 0$ et nous trouverons $\mathbf{u}_2 = (-1, 1, 0)$ le vecteur propre associé à la deuxième valeur propre λ_2 .

remarque : il faut que la norme de chaque vecteur qu'elle soit égale à 1. Sinon, on divise les composantes de chaque vecteur par la racine carrée de sa norme pour le rendre unitaire. Après la normalisation des deux vecteurs on obtient :

$$\mathbf{u}_1 = (0.4082, 0.4082, 0.8165)$$

$$\mathbf{u}_2 = (0.7071, -0.7071, 0)$$

8. Les contributions absolues des axes $\mathbf{F}_1$ et $\mathbf{F}_2$ à l'inertie du nuage :

$$cr(F_k/I_g) = \frac{\lambda_k}{\lambda_1 + \lambda_2 + \dots + \lambda_p} = \frac{\lambda_k}{Trace(V)}$$

$$cr(F_1/I_g) = \frac{\lambda_1}{Trace(V)} = \frac{17/6}{11/3} = 0.7727 \quad \text{et} \quad cr(F_2/I_g) = \frac{\lambda_2}{Trace(V)} = \frac{5/6}{11/3} = 0.2273$$

9. La représentation graphique :

Pour représenter les individus dans le nouveau plan, il suffit de calculer les coordonnées des individus dans les nouveaux axes. Pour obtenir y_{ik} , coordonnée de l'unité e_i sur l'axe F_k on projette orthogonalement le vecteur $\vec{ge}_i$ sur cet axe et on obtient :

$$y_{ik} = \langle \vec{ge}_i, \vec{u}_i \rangle = Y^t u$$

$$y_{6,3} = \begin{pmatrix} -\frac{1}{3} & -\frac{4}{3} & -\frac{5}{3} \\ -\frac{1}{3} & -\frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & \frac{4}{3} \\ -\frac{4}{3} & -\frac{1}{3} & -\frac{5}{3} \\ \frac{5}{3} & -\frac{1}{3} & \frac{4}{3} \\ -\frac{1}{3} & \frac{5}{3} & \frac{4}{3} \end{pmatrix} \begin{pmatrix} 0.4082 & 0.7071 \\ 0.4082 & -0.7071 \\ 0.8165 & 0.0000 \end{pmatrix} = \begin{pmatrix} -2.0412 & -0.7071 \\ -0.8165 & 0.0000 \\ 1.6330 & 0.0000 \\ -2.0412 & 0.7071 \\ 1.6330 & -1.4142 \\ 1.6330 & 1.4142 \end{pmatrix}$$

i	$\mathbf{F}_1$	$\mathbf{F}_2$
1	-2.04	-0.71
2	-0.82	0.00
3	1.63	0.00
4	-2.04	0.71
5	1.63	-1.41
6	1.63	1.41

graphe des individus

