TD 2 : Analyse en Composantes Principales - 2 Variables - Professeur : Harchli Fidae

Exercice 1 (Application du cours) :

Deux systèmes d'exploitation sont notés de 0 à 20 exprimant une appréciation du produit élevée :

INDIVIDUS	WIN	MAC
A	10	2
B	20	4
C	16	6
D	15	8
E	5	18

1. **Donner les éléments de cette ACP**
2. **Présenter les données sous forme de matrice que vous nommez M**
3. **Procéder au centrage de la matrice**
4. **Calculer la matrice de variance-covariance A**
5. **Calculer les valeurs et les vecteurs propres**
6. **Quel est le taux d'inertie du premier axe factoriel**
7. **Quel est le taux d'inertie du premier plan factoriel**
8. **Calculer les coordonnées F des individus dans le plan factoriel**
9. **Présenter ses coordonnées graphiquement**
10. **Interpréter les résultats obtenus**

Exercice 2 (Python) :

Considérons les notes (sur 40) obtenues par 5 étudiants (poids uniformes) dans 2 matières :

- $\boldsymbol{X_1}$: Analyse des Données
- $\boldsymbol{X_2}$: Statistique Descriptive

INDIVIDUS	$\boldsymbol{X_1}$	$\boldsymbol{X_2}$
1	18.9	21.9
2	19.3	22.8
3	20.1	25.8
4	21.4	27.9
5	22.4	30.3

On propose de traiter ces données par **l'ACP non normée**. Pour ce, répondez aux questions suivantes :

1. **Calculer la matrice A de variance-covariance**
2. **En admettant que** $A = \begin{pmatrix} 1.71 & 4.05 \\ 4.05 & 9.77 \end{pmatrix}$ **calculer ses valeurs propres ainsi que ses vecteurs propres normés**
3. **Écrire les expressions des deux composantes principales $\boldsymbol{F^1}$ et $\boldsymbol{F^2}$ en fonction de $\boldsymbol{Z_1}$ et $\boldsymbol{Z_2}$** (où Z_1 et Z_2 les variables centrées respectivement de X_1 et X_2) **et calculer la part de variance expliquée par chacune d'elles** (seulement pour cette question, on présente le résultat à la troisième décimale).
4. **Démontrer que les corrélations des variables $\boldsymbol{Z_1}$ et $\boldsymbol{Z_2}$ avec $\boldsymbol{F^1}$ sont respectivement ≈ 0.993 et ≈ 0.9998**
5. **En déduire une signification de l'axe $\boldsymbol{F^1}$**
6. **Calculer les coordonnées des étudiants dans le plan principal**
7. **Déterminer l'étudiant moyen ? Justifier**
8. **Comment ont peut interpréter les étudiants 1 et 2 ? Justifier**
9. **Même question pour les étudiants 4 et 5**

Exercice 3 :

Une étude gastronomique a conduit à apprécier le service, la qualité et le prix de quatre restaurants. Pour cela, un expert a noté ces restaurants avec des notes allant de -3 à 3. Les résultats sont les suivants :

Restaurant	**Service**	**Qualité**	**Prix**
1	-2	3	-1
2	-1	1	0
3	2	-1	-1
4	1	-3	2

La matrice des variances-covariances est :

$$\mathbf{X} = \begin{pmatrix} \frac{5}{2} & -3 & \frac{1}{2} \\ -3 & 5 & -2 \\ \frac{1}{2} & -2 & \frac{3}{2} \end{pmatrix}$$

Et celle des corrélations (aux erreurs d'arrondi près) est : La matrice des corrélations (aux erreurs d'arrondi près) est :

$$\Gamma = \begin{pmatrix} 1 & -0.85 & 0.26 \\ -0.85 & 1 & -0.73 \\ 0.26 & -0.73 & 1 \end{pmatrix}.$$

1. Étude des valeurs propres :
   (a) Vérifier que V admet une valeur propre $\lambda_3 = 0$.
   (b) On donne $\lambda_1 = \frac{30.5}{4}$. En déduire λ_2.
   (c) Calculer les pourcentages d'inertie. Quelle est la dimension à retenir ?
2. (a) On donne, aux erreurs d'arrondi près, $v_1 = \begin{pmatrix} 0.5 \\ -0.8 \\ 0.3 \end{pmatrix}$ et $v_2 = \begin{pmatrix} 0.65 \\ 0.11 \\ -0.75 \end{pmatrix}$. Calculer les composantes principales.
   (b) Représenter les individus dans le plan principal $(1, 2)$.
3. (a) Déterminer les corrélations entre les variables et les composantes.
   (b) Représenter les variables sur le cercle des corrélations dans le plan factoriel $(1, 2)$.
   (c) Interpréter les résultats.

‘

Exercice 4 (ACP à p variables) :

Considérons la matrice suivante :

$$\mathbf{X} = \begin{bmatrix} 2 & 0 & 0 & 8 \\ 4 & 0 & 0 & 8 \\ 6 & 0 & 0 & 8 \\ 8 & 8 & 8 & 8 \end{bmatrix}$$

1. **Centrer et réduire les données.**
2. **Calculer la matrice de covariance.**
3. **Déterminer les valeurs propres et les vecteurs propres.**
4. **Projeter les données sur les deux premières composantes principales.**
5. **Analyser la perte d'information si on conserve uniquement les deux premières composantes.**
6. **Appliquer l'ACP sur un jeu de données issu de la pratique (par exemple, le jeu `Iris` de la bibliothèque `scikit-learn`). Discuter des résultats obtenus.**