Université Mohamed Khider, Biskra 2019/2020
Faculté des Sc. Exactes et Sc. de la Nature et la Vie
Master 1: **M1.4 (Modèle linéaire)**

Série N°2 : ACP

Exercice 1 *On considère la matrice des données:*

$$\mathbf{X} = \begin{pmatrix} -1 & 0 & 1 \\ 0 & -1 & 1 \end{pmatrix}.$$

1. Calculer le produit matriciel $\mathbf{X}^t\mathbf{X}$ *et s'assurer que c'est une matrice carré et symétrique.*
2. Calculer les valeurs propres λ_i *de* $\mathbf{X}^t\mathbf{X}$ *et ces vecteurs propres* $\mathbf{u}_i$ *associés. Donner la matrice diagonale* Λ *semblable à* $\mathbf{X}^t\mathbf{X}$ *et la matrice de passage* $\mathbf{P}$.
3. Vérifier que trace$(\mathbf{X}^t\mathbf{X}) = \sum_i \lambda_i$.

Exercice 2 *Soit la matrice des données suivante*

$$\mathbf{X}^* = \begin{pmatrix} 4 & 5 \\ 6 & 7 \\ 8 & 0 \end{pmatrix}.$$

1. Centrer et réduit (normer) les deux vecteurs colonnes, X_1^* *et* X_2^*, *de* $\mathbf{X}^*$.
2. Déterminer la matrice de variances-covariances $\mathbf{V}$ *et la matrice des corrélations* $\mathbf{R}$.
3. Diagonaliser $\mathbf{V}$. *On note* λ_i *ses valeurs propres.*
4. Déterminer les vecteurs propres $\mathbf{u}_i$ *associés à ces valeurs propres.*
5. Dans le contexte de l'analyse en composantes principales, déterminer les axes principaux du nuage de points défini par la matrice $\mathbf{X}$.

Exercice 3 *On s'intéresse à l'ACP sur le nuage de points défini par la matrice*

$$\mathbf{X}^* = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

Donc il s'agit de 4 lignes-individus et 4 colonnes-variables.
1. Donner les moyennes et les variances des quatre variables puis déterminer la matrice de variances-covariances $\mathbf{V}$ *associée à* $\mathbf{X}^*$.
2. Donner les valeurs propres et les vecteurs propres de $\mathbf{V}$.
3. Donner les coordonnées des lignes sur le deuxième axe principal de l'ACP de $\mathbf{X}^*$.
4. Donner les coordonnées des colonnes sur le deuxième axe principal de l'ACP de $\mathbf{X}^*$.

Exercice 4 *Une étude gastronomique a conduit à apprécier le service, la qualité et le prix de quatre restaurants. Pour cela, un expert à note ces restaurants avec des notes allant de* -3 *à* 3. *Les résultats sont les suivants:*

Restaurant	***Service***	***Qualité***	***Prix***
$\mathbf{R}_1$	-2	$+3$	-1
$\mathbf{R}_2$	-1	$+1$	0
$\mathbf{R}_3$	$+2$	-1	-1
$\mathbf{R}_4$	$+1$	-3	2

La matrice de variances-covariances est

$$\mathbf{V} = \begin{pmatrix} 5/2 & -3 & 1/2 \\ -3 & 5 & -2 \\ 1/2 & -2 & 3/2 \end{pmatrix},$$

et celle de corrélations (aux erreurs arrondies près) est

$$\mathbf{R} = \begin{pmatrix} 1 & -0.85 & 0.26 \\ -0.85 & 1 & -0,73 \\ 0.26 & -0.73 & 1 \end{pmatrix}.$$

***1**. Etude de valeurs propres:*
i) Vérifier que $\mathbf{V}$ *admet une valeur propre* $\lambda_3 = 0$.
ii) On donne $\lambda_1 = 30.5/4$. *Déduire la valeur de* λ_2.
iii) Calculer les pourcentages d'inerties. Quelle est la dimension à retenir?
***2**. Les vecteurs propres associés à* λ_1 *et* λ_2, *aux erreurs arrondies près, sont*

$$\mathbf{u}_1^* = \begin{pmatrix} 0.5 \\ -0.8 \\ 0.3 \end{pmatrix} \text{ et } \mathbf{u}_2^* = \begin{pmatrix} 0.65 \\ 0.11 \\ -0.75 \end{pmatrix}.$$

i) Déterminer les composantes principales qui correspondent aux axes principaux associés à $\mathbf{u}_1^*$ *et* $\mathbf{u}_2^*$ *respectivement.*
ii) Représenter les individus dans le plan principal $(1, 2)$.
***3**. Représentation:*
i) Déterminer les corrélations entre les variables originelles et les composantes principales.
ii) Représenter les variables sur le cercle des corrélations dans le plan factoriel $(1, 2)$.
iii) Interpréter les résultats.

Références:

- Mohamed El Merouani: https://elmerouani.jimdofree.com/analyse-des-donn%C3%A9es/
- Jean François Durand: "Eléments de Calcul matriciel et d'analyse factorielle Données", polycopie de l'Université de Montpellier II, Licence MASS, Maitrise MASS, Maitrise d'ingénierie Mathématique, DEA de Biostatistique, Novembre 2002.
- Le site Web: foad.refer.org/IMG/pdf/