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

Examen (1h)

Exercice 1 ***(ACP, 16 pts)***
Considérons le tableau de données suivant :

	Math	**Info**	**Gestion**
Ahmed	0	4	0
Ibrahim	1	3	1
Youcef	2	2	3
Imane	3	1	1
Hicham	4	0	0

où les lignes représentent les individus (noms de quelques étudiants d'une formation) et les colonnes les variables (notes sur 4 *en mathématiques, informatique et gestion). Le tableau de données peut être représenté par la matrice de données brutes :*

$$X^* := \begin{pmatrix} 0 & 4 & 0 \\ 1 & 3 & 1 \\ 2 & 2 & 3 \\ 3 & 1 & 1 \\ 4 & 0 & 0 \end{pmatrix}.$$

1. Calculer la matrice X des données centrées.(2pts)
2. Vérifier que la matrice de variance-covariance vaut

$$V := \begin{pmatrix} 2 & -2 & 0 \\ -2 & 2 & 0 \\ 0 & 0 & 1.2 \end{pmatrix}. (2pts)$$

3. Vérifier que les vecteurs $u_1 := \left(-1/\sqrt{2}, 1/\sqrt{2}, 0\right)^t$ et $u_2 := (0,0,1)^t$, sont les vecteurs propres de V associés aux valeurs propres non-nulles. Trouvez-les (2pt).
4. Vérifier l'orthonormalité des deux vecteurs u_1 et u_2. (2pt)
5. Définir les deux axes principaux du nuage de point X. (1pt)
6. Déterminer l'inertie totale par rapport au premier plan principal.(1pt)
7. Déterminer l'inertie relative à chaque axe du premier plan principal.(1pt)
8. Quelles sont coordonnées des individus (de X) par rapport au plan défini par la base $\{u_1, u_2\}$. (2pt)
9. Déduire de la question 8, les composantes principales du premier plan principal.(1pt)
10. Peut-on représenter parfaitement le nuage des individus à deux dimensions ? Si oui, représentez les individus de X sur le plan défini par la base $\{u_1, u_2\}$. (2pt)

Exercice 2 ***(AFC, 4pts)***
Choisir les bonnes réponses qui correspondent aux 4 questions suivantes:
***1**. L'analyse factorielle des correspondances est basée sur:*
a) La matrice V_r b) la matrice M_rV_r c) la matrice V_rM_r
***2**. La dimension de la matrice V_rM_r égale à:*
a) $p \times q$ b) $p \times p$ c) $q \times q$
***3**. Le nombre de valeurs propres de la matrice V_cM_c est:*
a) p b) q c) pq.
***4**. Le degrés de liberté de la statistique de khi-deux qui correspond à la table de contingence égale à:*
a) $p+q$ b) pq c) $(p-1)(q-1)$
***5**. La p-value du test de khi-deux vaut 0.01. Peut-on effectuer l'AFC?*
***6**. Nous avons $p = 3$ et $q = 10$. Que choisissez vous?*
a) AFC des profils-lignes b) AFC des profils-colonnes
***7**. Nous avons $p = 3$ et $q = 10$. Le rang(V_rM_r) est:*
a) $= 30$ b) ≤ 18 c) ≤ 20.
***8**) L'écart à l'indépendance égale à:*
a) $n\chi^2_{obs}$ b) χ^2_{obs}/n c) l'inertie totale.

Corrigé type de l'Examen

Solution de l'exercice 1.

1) Calcul de la matrice X des données centrées. Le centre de gravité est

$$g = (10/5, 10/5, 5/5)^t = (2, 2, 1)^t.$$

La matrice des données centrée est

$$X := \begin{pmatrix} 0-2 & 4-2 & 0-1 \\ 1-2 & 3-2 & 1-1 \\ 2-2 & 2-2 & 3-1 \\ 3-2 & 1-2 & 1-1 \\ 4-2 & 0-2 & 0-1 \end{pmatrix} = \begin{pmatrix} -2 & 2 & -1 \\ -1 & 1 & 0 \\ 0 & 0 & 2 \\ 1 & -1 & 0 \\ 2 & -2 & -1 \end{pmatrix}.$$

La matrice de variance-covariance est:

$$V = \frac{1}{5} \begin{pmatrix} -2 & 2 & -1 \\ -1 & 1 & 0 \\ 0 & 0 & 2 \\ 1 & -1 & 0 \\ 2 & -2 & -1 \end{pmatrix}^T \begin{pmatrix} -2 & 2 & -1 \\ -1 & 1 & 0 \\ 0 & 0 & 2 \\ 1 & -1 & 0 \\ 2 & -2 & -1 \end{pmatrix} =: \begin{pmatrix} 2 & -2 & 0 \\ -2 & 2 & 0 \\ 0 & 0 & 1.2 \end{pmatrix}.$$

Vérifions que les vecteurs $u_1 := \left(-1/\sqrt{2}, 1/\sqrt{2}, 0\right)^t$ et $u_2 := (0,0,1)^t$, sont les vecteurs propres de V associés aux valeurs propres non-nulles. Nous devons montrer l'existence de deux constantes non nulles λ_1 et λ_2, appellées valeurs propres, telles que $Vu_i = \lambda_i u_i$ pour $i = 1, 2$. En effet,

$$Vu_1 = \begin{pmatrix} 2 & -2 & 0 \\ -2 & 2 & 0 \\ 0 & 0 & 1.2 \end{pmatrix} \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix} = \begin{pmatrix} -2\sqrt{2} \\ 2\sqrt{2} \\ 0 \end{pmatrix}.$$

Comme $Vu_1 = \lambda_1 u_1$, alors

$$\begin{pmatrix} -2\sqrt{2} \\ 2\sqrt{2} \\ 0 \end{pmatrix} = \lambda_1 \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix},$$

ce qui implique que $-2\sqrt{2} = -\lambda_1/\sqrt{2}$ et $2\sqrt{2} = \lambda_1/\sqrt{2}$. Les deux équations donnent une solution unique $\lambda_1 = 4$. En faisant la même procédure pour u_2, on trouve $\lambda_2 = 1.2$.

4. Vérifier l'orthonormalité des deux vecteurs u_1 et u_2 : on doit verifier que leur produit scalaire vaut zero et chacun de norme vaut 1. En effet

$$\langle u_1, u_2 \rangle = u_1^t u_2 = \left(-1/\sqrt{2}, 1/\sqrt{2}, 0\right) \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 0 + 0 + 0 = 0.$$

$$\|u_1\|^2 = u_1^t u_1 = \left(-1/\sqrt{2}, 1/\sqrt{2}, 0\right) \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix} = \frac{1}{2} + \frac{1}{2} + 0 = 1.$$

$$\|u_2\|^2 = u_2^t u_2 = (0,0,1) \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 0 + 0 + 1 = 1.$$

5. Les deux axes principaux du nuage de point X sont: $E_i = Vect\{u_i\}$, $i = 1, 2$.

6. L'inertie totale par rapport au premier plan principal: $I_T = \lambda_1 + \lambda_2 = 4 + 1.2 = 5.2$.

Remarque: certains étudiants écrivent $I_T = \lambda_1 + \lambda_2 + \lambda_2$. C'est faux, meme si le résultat est le même (coup de chance $\lambda_3 = 0!$).

7. Les inerties relatives à chaque axe du premier plan principal sont

$$I_1\left(E_1/E_1 \oplus E_2\right) = \frac{\lambda_1}{\lambda_1+\lambda_2} = \frac{4}{5.2} = 0.769\,23 \simeq 77\,\%$$

$$I_1\left(E_2/E_1 \oplus E_2\right) = \frac{\lambda_2}{\lambda_1+\lambda_2} = \frac{1.2}{5.2} = 0.230\,77 \simeq 23\,\%$$

8. Coordonnées des individus (de X) par rapport au plan défini par la base $\{u_1, u_2\}$: sont les lignes de la matrice

$$Y = XU = \begin{pmatrix} -2 & 2 & -1 \\ -1 & 1 & 0 \\ 0 & 0 & 2 \\ 1 & -1 & 0 \\ 2 & -2 & -1 \end{pmatrix} \begin{pmatrix} -1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2.82 & -1 \\ 1.41 & 0 \\ 0 & 2 \\ -1.41 & 0 \\ -2.82 & -1 \end{pmatrix}.$$

9. Les composantes principales du premier plan principal sont les collones de la matrice Y, c'est a dire

$$c_1 = \begin{pmatrix} 2.82 \\ 1.41 \\ 0 \\ -1.41 \\ -2.82 \end{pmatrix} \text{ et } c_2 = \begin{pmatrix} -1 \\ 0 \\ 2 \\ 0 \\ -1 \end{pmatrix}.$$

Peut-on représenter parfaitement le nuage des individus à deux dimensions: oui car l'enertie du troisieme axe principale vaut zero et par consequent l'enertie totale de l'espaces fromé par les trois axes principaux égale à l'inertie totale du paln formé par les deux premiers axe principaux. La représentation des individus de X sur le plan défini par la base $\{u_1, u_2\}$ est triviale.

Solution de l'exercice 2

Choisir les bonnes réponses qui correspondent aux 4 questions suivantes:

1. L'analyse factorielle des correspondances est basée sur:
c) la matrice $V_r M_r$
2. La dimension de la matrice $V_r M_r$ égale à:
c) $q \times q$
3. Le nombre de valeurs propres de la matrice $V_c M_c$ est:
a) p
4. Le degrés de liberté de la statistique de khi-deux qui correspond à la table de contingence égale à:
c) $(p-1)(q-1)$
5. La p-value du test de khi-deux vaut 0.01. Peut-on effectuer l'AFC? oui car $0.01 < 0.05$ donc les deux variables à étuider sont dépendantes.
6. Nous avons $p = 3$ et $q = 10$. Que choisissez vous?
AFC des profils-colonnes
7. Nous avons $p = 3$ et $q = 10$. Le rang$(V_r M_r)$ est:
b) ≤ 18
8) L'écart à l'indépendance égale à:
b) χ^2_{obs}/n c) l'inertie totale.