Master 1 Data-Science -INPHB

Analyse des données
Mory Ouattara

Exercice 1

On considère le tableau K suivant où a est un entier non nul :

I/J	j_1	j_2	j_3	j_4	j_5
i_1	a	a	a	0	0
i_2	a	a	0	a	0
i_3	0	a	0	a	a

On pose :
$I = \{i_1; i_2; i_3\}$ et $J = \{j_1; j_2; j_3; j_4; j_5\}$
On effectue l'analyse factorielle des corresponsdances (AFC) de K.
1-) Déterminer les centres de gravité des nuages $N(I)$ et $N(J)$
2-) Déterminer la matrice des profils colonnes F_1 ainsi que la matrice des profils lignes F_2 de K.
3-) Calculer le produit F_1F_2
4-) Quel est l'influence du réel a sur l'AFC de ce tableau.
5-) Quel est l'axe factoriel trivial, à quelle valeur propre est-il associé ?
6-) Quelle est l'inertie du nuage N(J).

7-) On pose

$$w_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \text{ et } w_2 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$$

Montrer que w_1 et w_2 sont des vecteurs propres de F_1F_2, en déduire les axes factoriels non triviaux u_1 et u_2 ainsi que les valeurs propres associés. On choisira u_1 de manière que la première coordonnée soit positive, de même pour u_2.
8-) Calculer $\varphi_\alpha(i)$ l'abscisse de la projection du profil de la ligne i sur le α avec la contrainte $\varphi_\alpha(i_1) \geq 0$
9-) Calculer à l'aide des formules de transition, ψ_α^j l'abscisse de la projection du profil de la colonne j sur le α ème axe factoriel.
10-) Représenter les deux nuages N(I) et N(J) simultanément dans le plan factoriel 1-2.
11-) Calculer la contribution de i_1 a chacun des axes factoriels non triviaux ainsi que la qualité de représentation de i_1 dans le plan factoriel 1-2 c'est-à-dire $COR1(i_1) + COR2(i_1)$.

Exercice 2

On considère le tableau de données, noté X, et défini par :

$$X = \begin{array}{|c|cccccc|} \hline & j_1 & j_2 & j_3 & j_4 & j_5 & j_6 \\ \hline i_1 & 1 & 1 & 2 & 0 & 3 & 1 \\ i_2 & 0 & 1 & 2 & 1 & 1 & 3 \\ i_3 & 1 & 2 & 4 & 1 & 4 & 4 \\ \hline \end{array}$$

où la i ème ligne désigne la variable x_i et la j ème colonne désigne l'individu x^j.
Par la suite, on considère les résultats de l'ACP sur matrice variance du tableau X.

1 - Calculer les coordonnées du centre de gravité g du nuage $\mathcal{M}$ constitué des vecteurs colonnes de X (munis du même poids $1/6$), et en déduire le tableau Y centré qui est associé à X. On présentera Y sous la forme $Y = \frac{1}{3}Y_1$ où Y_1 est une matrice à coefficients entiers.

2 - Soit V la matrice variance du tableau X. Compléter les valeurs manquantes dans l'expression de la matrice V ci-dessous :

$$V = \frac{1}{18}\begin{pmatrix} 16 & 1 & 17 \\ 1 & 16 & ? \\ 17 & ? & ? \end{pmatrix}.$$

3 - Expliquer pourquoi le nombre d'axes factoriels non triviaux est égal à 2.

4 - Calculer l'inertie totale du nuage étudié.

5 - Montrer que $\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$ est un vecteur directeur d'un axe factoriel non trivial.

6 - Calculer le pourcentage d'inertie expliquée par l'axe factoriel déterminé à la question 5. Cet axe est-il le premier ou le second axe factoriel ?

7 - Déterminer les coordonnées du premier vecteur axial factoriel, noté u^1 (on choisira sa première coordonnée de façon à ce qu'elle soit positive).

8 - Calculer la première composante principale de l'individu j_2, notée $\Psi_1^{j_2}$.

9 - Calculer la contribution de l'individu j_2 à l'inertie du premier axe, notée $CTR_1(j_2)$.

10 - Calculer la qualité de représentation de l'individu j_2 sur le premier axe, notée $COR_1(j_2)$.

11 - Calculer la contribution de la variable i_1 à l'inertie du premier axe, notée $CTR_1(i_1)$.