ACP simple . Exemple avec solution

Exercice

Objectif: Faire une **ACP** simple sur les données résumées dans un tableau $X_{8,4}$ croisant 8 individus de même poids avec 4 variables quantitatives homogènes de variances voisines.

$$\text{Soit } X_{n,p} = \begin{bmatrix} 10 & 8 & 6 & 6 \\ 5 & 5 & 3 & 1 \\ 4 & 6 & 0 & 4 \\ 7 & 11 & 5 & 7 \\ 5 & 7 & 7 & 11 \\ 0 & 4 & 4 & 6 \\ 3 & 1 & 5 & 5 \\ 6 & 6 & 10 & 8 \end{bmatrix}$$

(i) Justifier l'application de l'ACP simple et déterminer l'individu moyen g

(ii) Déduire la matrice centrée X^*

(iii) Déterminer la matrice Variance covariance V et calculer sa trace

(iv) Soient les vecteurs propres $U_\alpha(\alpha = 1,\ldots,4)$ de V.

Calculer les valeurs propres λ_1 et λ_2.

Indication: On vous donne $\lambda_3 = 4$ et $\lambda_4 = 1$

$$\text{vecteurs propres:} U_1 = \left\{\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}\right\} \leftrightarrow \lambda_1; U_2 = \left\{\begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \end{bmatrix}\right\} \leftrightarrow \lambda_2, U_3 = \left\{\begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \end{bmatrix}\right\} \leftrightarrow \lambda_3 = 4,$$

Quel est le vecteur propre U_4 correspondant à $\lambda_4 = 1$?

(v)Comparer $\sum_{i=1}^{4} \lambda_i$ avec $tr(V)$. Déterminer l'inertie globale et Compléter le tableau.

Valeurs propres λ_i	λ_1	λ_2	$\lambda_3 = 4$	$\lambda_4 = 1$
Taux d'inertie: τ_i(%)				
Taux cumulés (%)				

(vi) Calculer les F_α.

(vii) Faite la représentation sur le plan principal et donner son inertie.

(viii) A l'aide des relations entre les deux ajustements retrouver les vecteurs V_α et les G_α.

(ix) Donner la formule de reconstitution.

Solution

$$\text{Soit } X_{n,p} = \begin{bmatrix} 10 & 8 & 6 & 6 \\ 5 & 5 & 3 & 1 \\ 4 & 6 & 0 & 4 \\ 7 & 11 & 5 & 7 \\ 5 & 7 & 7 & 11 \\ 0 & 4 & 4 & 6 \\ 3 & 1 & 5 & 5 \\ 6 & 6 & 10 & 8 \end{bmatrix}$$

Faire une ACP simple(ACP de covariance)sur les données résumées dans $X_{8,4}$ en supposant que les 8 individus sont de même (masse m_i)(poids p_i)

Nuage des individus

Ajustement par un sous espace de $\mathbb{R}^4$

(i) Justifier l'application de l'ACP simple et déterminer L'individu moyen g

Justification de l'application de l'ACP simple: Type de croisement: individus×variables quantitaves et nature des variables: homogènes de variances voisines.

L'individu moyen $g = X^T D_{m_i} 1_n$ $\quad g = \frac{1}{8}\begin{bmatrix} 10 & 8 & 6 & 6 \\ 5 & 5 & 3 & 1 \\ 4 & 6 & 0 & 4 \\ 7 & 11 & 5 & 7 \\ 5 & 7 & 7 & 11 \\ 0 & 4 & 4 & 6 \\ 3 & 1 & 5 & 5 \\ 6 & 6 & 10 & 8 \end{bmatrix}^T \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 6 \\ 5 \\ 6 \end{bmatrix}$

(ii) On se ramène à l'origine à l'aide de la matrice centrée

$$X^*_{8,4} = X - 1_n \times G^T = \begin{bmatrix} 10 & 8 & 6 & 6 \\ 5 & 5 & 3 & 1 \\ 4 & 6 & 0 & 4 \\ 7 & 11 & 5 & 7 \\ 5 & 7 & 7 & 11 \\ 0 & 4 & 4 & 6 \\ 3 & 1 & 5 & 5 \\ 6 & 6 & 10 & 8 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 5 \\ 6 \\ 5 \\ 6 \end{bmatrix}^T = \begin{bmatrix} 5 & 2 & 1 & 0 \\ 0 & -1 & -2 & -5 \\ -1 & 0 & -5 & -2 \\ 2 & 5 & 0 & 1 \\ 0 & 1 & 2 & 5 \\ -5 & -2 & -1 & 0 \\ -2 & -5 & 0 & -1 \\ 1 & 0 & 5 & 2 \end{bmatrix}$$

1. (iii) La matrice à diagonaliser est V: Matrice variance covariance

$$V_{p,p} = X^{*T} D_m X^* = \frac{1}{8} X^{*T} X^* = \frac{1}{8}\begin{bmatrix} 5 & 2 & 1 & 0 \\ 0 & -1 & -2 & -5 \\ -1 & 0 & -5 & -2 \\ 2 & 5 & 0 & 1 \\ 0 & 1 & 2 & 5 \\ -5 & -2 & -1 & 0 \\ -2 & -5 & 0 & -1 \\ 1 & 0 & 5 & 2 \end{bmatrix}^T \begin{bmatrix} 5 & 2 & 1 & 0 \\ 0 & -1 & -2 & -5 \\ -1 & 0 & -5 & -2 \\ 2 & 5 & 0 & 1 \\ 0 & 1 & 2 & 5 \\ -5 & -2 & -1 & 0 \\ -2 & -5 & 0 & -1 \\ 1 & 0 & 5 & 2 \end{bmatrix} =$$

$$\frac{1}{8}\begin{bmatrix} 60 & 40 & 20 & 8 \\ 40 & 60 & 8 & 20 \\ 20 & 8 & 60 & 40 \\ 8 & 20 & 40 & 60 \end{bmatrix} = \begin{bmatrix} \frac{15}{2} & 5 & \frac{5}{2} & 1 \\ 5 & \frac{15}{2} & 1 & \frac{5}{2} \\ \frac{5}{2} & 1 & \frac{15}{2} & 5 \\ 1 & \frac{5}{2} & 5 & \frac{15}{2} \end{bmatrix}$$

$tr(V) = \sum_{i=1}^{4} v_{ii} = \frac{15}{2} + \frac{15}{2} + \frac{15}{2} + \frac{15}{2} = \mathbf{30}$

(iv)Recherche du spectre(valeurs et vecteurs propres correspondants) de la matrice Variance covariance V

Calculer les valeurs propres λ_1 et λ_2 on vous donne $\lambda_3 = 4$ et $\lambda_4 = 1$

vecteurs propres: $U_1 = \left\{\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}\right\} \leftrightarrow \boldsymbol{\lambda_1 = 16}, \mathbf{U_2} = \left\{\begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \end{bmatrix}\right\} \leftrightarrow \boldsymbol{\lambda_2 = 9}, U_3$

$$= \left\{ \begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \end{bmatrix} \right\} \leftrightarrow \lambda_3 = 4, \text{ et } \lambda_4 = 1$$

calculer le vecteur propre U_4 correspondant à $\lambda_4 = 1$

Le vecteur propre U_4 correspondant à $\lambda_4 = 1$ est solution de $VU_4 = U_4$

$$\begin{bmatrix} \frac{15}{2} & 5 & \frac{5}{2} & 1 \\ 5 & \frac{15}{2} & 1 & \frac{5}{2} \\ \frac{5}{2} & 1 & \frac{15}{2} & 5 \\ 1 & \frac{5}{2} & 5 & \frac{15}{2} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix} \Leftrightarrow \begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}. \ U_4 = \left\{ \begin{bmatrix} \mathbf{1} \\ \mathbf{-1} \\ \mathbf{-1} \\ \mathbf{1} \end{bmatrix} \right\} \leftrightarrow 1$$

Conclusion:

$$\begin{bmatrix} \frac{15}{2} & 5 & \frac{5}{2} & 1 \\ 5 & \frac{15}{2} & 1 & \frac{5}{2} \\ \frac{5}{2} & 1 & \frac{15}{2} & 5 \\ 1 & \frac{5}{2} & 5 & \frac{15}{2} \end{bmatrix}$$ Valeurs propres par ordre décroissant: $\boldsymbol{\lambda_1 = 16}, \boldsymbol{\lambda_2 = 9}, \lambda_3 = 4, \lambda_4 = 1$ sont

données

(v)

$$V = \begin{bmatrix} \frac{15}{2} & 5 & \frac{5}{2} & 1 \\ 5 & \frac{15}{2} & 1 & \frac{5}{2} \\ \frac{5}{2} & 1 & \frac{15}{2} & 5 \\ 1 & \frac{5}{2} & 5 & \frac{15}{2} \end{bmatrix} \text{ a pour trace: } 30 = \sum_{i=1}^{4} \lambda_i = 16 + 9 + 4 + 1 = 30$$

L'inertie globale=30 d'aprés tr(VM)= I_G dans ce cas $M=I_4$

compléter le tableau:

Valeurs propresλ_i	$\lambda_1 = 16$	$\lambda_2 = 9$	$\lambda_3 = 4$	$\lambda_4 = 1$
Taux d'inertie:$\tau_i = \frac{\lambda_i}{\sum \lambda_i}$(%)	53	30	14	3
Taux cumulés (%)	53	83	97	100

(vi)Calculer les F_α

Les composantes principales et les facteurs : La α-ème

$$F_\alpha = X^*_{8,4} U'_{\lambda_\alpha} \text{ où } \quad U'_{\lambda_i} = \frac{U_{\lambda_i}}{\| U_{\lambda_i} \|} \text{ le vecteur normalisé de la i-ème valeur propre ordonnée}$$

La première Composante principale:On a $U'_1 = \frac{U_1}{\|U_1\|} = \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ d'où $F_1 = \frac{1}{2}\begin{bmatrix} 5 & 2 & 1 & 0 \\ 0 & -1 & -2 & -5 \\ -1 & 0 & -5 & -2 \\ 2 & 5 & 0 & 1 \\ 0 & 1 & 2 & 5 \\ -5 & -2 & -1 & 0 \\ -2 & -5 & 0 & -1 \\ 1 & 0 & 5 & 2 \end{bmatrix}$

$$\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ -4 \\ -4 \\ 4 \\ 4 \\ -4 \\ -4 \\ 4 \end{bmatrix}$$

$$F_2 = \frac{1}{2}\begin{bmatrix} 5 & 2 & 1 & 0 \\ 0 & -1 & -2 & -5 \\ -1 & 0 & -5 & -2 \\ 2 & 5 & 0 & 1 \\ 0 & 1 & 2 & 5 \\ -5 & -2 & -1 & 0 \\ -2 & -5 & 0 & -1 \\ 1 & 0 & 5 & 2 \end{bmatrix}\begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -3 \\ -3 \\ -3 \\ -3 \\ 3 \\ 3 \\ 3 \\ 3 \end{bmatrix} =, \quad F_3 = \frac{1}{2}\begin{bmatrix} 5 & 2 & 1 & 0 \\ 0 & -1 & -2 & -5 \\ -1 & 0 & -5 & -2 \\ 2 & 5 & 0 & 1 \\ 0 & 1 & 2 & 5 \\ -5 & -2 & -1 & 0 \\ -2 & -5 & 0 & -1 \\ 1 & 0 & 5 & 2 \end{bmatrix}\begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \end{bmatrix} =$$

$$\begin{bmatrix} -2 \\ -2 \\ 2 \\ 2 \\ 2 \\ 2 \\ -2 \\ -2 \end{bmatrix} F_4 = \frac{1}{2}\begin{bmatrix} 5 & 2 & 1 & 0 \\ 0 & -1 & -2 & -5 \\ -1 & 0 & -5 & -2 \\ 2 & 5 & 0 & 1 \\ 0 & 1 & 2 & 5 \\ -5 & -2 & -1 & 0 \\ -2 & -5 & 0 & -1 \\ 1 & 0 & 5 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \\ 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}$$

On peut vérifier

$$F_\alpha{}' N F_\alpha = \frac{1}{8} F' F = \lambda_\alpha$$

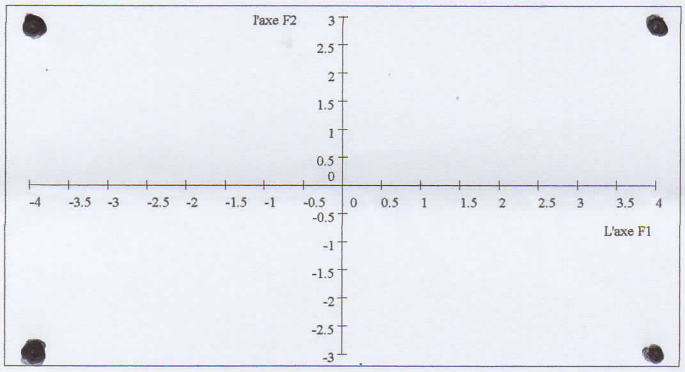
pour $\alpha = 2$ par exemple

$$\frac{1}{8}\begin{bmatrix} -3 \\ -3 \\ -3 \\ -3 \\ 3 \\ 3 \\ 3 \\ 3 \end{bmatrix}^T \begin{bmatrix} -3 \\ -3 \\ -3 \\ -3 \\ 3 \\ 3 \\ 3 \\ 3 \end{bmatrix} = 9$$

(vii) coord(i,1,2,3,4 axe)=$[F_1, F_2, F_3, F_4] = \begin{bmatrix} 4 & -3 & -2 & 1 \\ -4 & -3 & -2 & -1 \\ -4 & -3 & 2 & 1 \\ 4 & -3 & 2 & -1 \\ 4 & 3 & 2 & 1 \\ -4 & 3 & 2 & -1 \\ -4 & 3 & -2 & 1 \\ 4 & 3 & -2 & -1 \end{bmatrix}$

(viii)Représentation sur le plan principal: Plan engendré par le 1ièr et le 2ièm axe:

coord(i,1,2)=$\begin{bmatrix} 4 & -3 \\ -4 & -3 \\ -4 & -3 \\ 4 & -3 \\ 4 & 3 \\ -4 & 3 \\ -4 & 3 \\ 4 & 3 \end{bmatrix}$



Représentation sur le plan principal

L'inertie du plan principal vaut 83% ($\frac{\lambda_1+\lambda_2}{\sum \lambda_i}(\%) = \frac{16+9}{30}(\%)$)

(viii) On a d'aprés les relation entre les deux ajustements.

$$G_\alpha(j) = \sqrt{\lambda_\alpha}\, u^{*}_\alpha$$

$$G_1 = \sqrt{16}\,\tfrac{1}{2}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 2 \\ 2 \end{bmatrix};\ G_2 = \sqrt{9}\,\tfrac{1}{2}\begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{3}{2} \\ -\frac{3}{2} \\ \frac{3}{2} \\ \frac{3}{2} \end{bmatrix};\ G_3 = \sqrt{4}\,\tfrac{1}{2}\begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \end{bmatrix};$$

$$G_4 = \sqrt{1}\,\tfrac{1}{2}\begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$$

$$V_\alpha = \frac{1}{\sqrt{\lambda_\alpha}} F_\alpha$$

$$v_1 = \frac{1}{4}\begin{bmatrix}4\\-4\\-4\\4\\4\\-4\\-4\\4\end{bmatrix} = \begin{bmatrix}1\\-1\\-1\\1\\1\\-1\\-1\\1\end{bmatrix} \text{ et ainsi } v_2 = \frac{1}{3}\begin{bmatrix}-3\\-3\\-3\\-3\\3\\3\\3\\3\end{bmatrix} = \begin{bmatrix}-1\\-1\\-1\\-1\\1\\1\\1\\1\end{bmatrix},\ v_3 = \frac{1}{2}\begin{bmatrix}-2\\-2\\2\\2\\2\\2\\-2\\-2\end{bmatrix} = \begin{bmatrix}-1\\-1\\1\\1\\1\\1\\-1\\-1\end{bmatrix} \text{ et}$$

$$v_4 = \begin{bmatrix}1\\-1\\1\\-1\\1\\-1\\1\\-1\end{bmatrix}$$

(ix) **Formule de reconstitution:**

$$X^* = \sum_{\alpha} \sqrt{\lambda_\alpha}\, v_\alpha u_\alpha'$$

$$X^*=2\begin{bmatrix}1\\-1\\-1\\1\\1\\-1\\-1\\1\end{bmatrix}\begin{bmatrix}1\\1\\1\\1\end{bmatrix}^T + \frac{3}{2}\begin{bmatrix}-1\\-1\\-1\\-1\\1\\1\\1\\1\end{bmatrix}\begin{bmatrix}-1\\-1\\1\\1\end{bmatrix}^T + \begin{bmatrix}-1\\-1\\1\\1\\1\\1\\-1\\-1\end{bmatrix}\begin{bmatrix}-1\\1\\-1\\1\end{bmatrix}^T + \frac{1}{2}\begin{bmatrix}1\\-1\\1\\-1\\1\\-1\\1\\-1\end{bmatrix}\begin{bmatrix}1\\-1\\-1\\1\end{bmatrix}^T =$$

$$\begin{bmatrix}5 & 2 & 1 & 0\\0 & -1 & -2 & -5\\-1 & 0 & -5 & -2\\2 & 5 & 0 & 1\\0 & 1 & 2 & 5\\-5 & -2 & -1 & 0\\-2 & -5 & 0 & -1\\1 & 0 & 5 & 2\end{bmatrix} \text{ C'est identique avec } X^* = \begin{bmatrix}5 & 2 & 1 & 0\\0 & -1 & -2 & -5\\-1 & 0 & -5 & -2\\2 & 5 & 0 & 1\\0 & 1 & 2 & 5\\-5 & -2 & -1 & 0\\-2 & -5 & 0 & -1\\1 & 0 & 5 & 2\end{bmatrix}$$