

Analyse des données

Série 2

Analyse en Composantes Principales (Présentation et Interprétation)

Exercice 1

Considérons la matrice des données suivante ;

$I \backslash J$	$J1$	$J2$	$J3$
$I1$	0	2	3
$I2$	0	0	3
$I3$	4	2	1
$I4$	4	0	1

- 1) Déterminer la matrice des variances-covariances et celle des corrélations.
- 2) Déterminer le meilleur sous espace ajustant le nuage des points individus tout en justifiant le choix du nombre d'axes à retenir.
- 3) Déterminer les composantes principales choisies tout en précisant la variance de chaque composante principale.
- 4) Etudier les corrélations entre les variables initiales et les composantes principales.
- 5) En déduire les coordonnées des variables dans le plan principal (nouveau plan).
- 6) Représenter graphiquement ces corrélations dans le nouveau plan et donner une interprétation détaillée des résultats obtenus.

Exercice 2

Considérons une matrice de données X de type $(96, 8)$ dont la matrice des corrélations est donnée par :

$$R = \begin{pmatrix} 1 & 0.56 & 0.28 & 0.20 & 0.43 & 0.29 & 0.31 & 0.43 \\ 0.56 & 1 & 0.34 & 0.32 & 0.48 & 0.32 & 0.47 & 0.49 \\ 0.28 & 0.34 & 1 & 0.64 & 0.72 & 0.52 & 0.70 & 0.45 \\ 0.20 & 0.32 & 0.64 & 1 & 0.67 & 0.46 & 0.58 & 0.44 \\ 0.43 & 0.48 & 0.72 & 0.67 & 1 & 0.56 & 0.68 & 0.53 \\ 0.29 & 0.32 & 0.52 & 0.46 & 0.56 & 1 & 0.64 & 0.42 \\ 0.31 & 0.47 & 0.70 & 0.58 & 0.68 & 0.64 & 1 & 0.51 \\ 0.43 & 0.49 & 0.45 & 0.44 & 0.53 & 0.42 & 0.51 & 1 \end{pmatrix}.$$

- 1) Commenter les corrélations calculées entre les différentes variables.
- 2) La résolution de l'équation $\det(R - \lambda Id) = 0$ nous retourne les valeurs propres suivantes :

1.1366, 0.3758, 0.2287, 0.4580, 0.583, 0.5302, 4.4242, 0.2582.

Analyse des données

Déterminer les taux d'inertie portés par tous les axes principaux.

- 3) En déduire le meilleur sous espace principal de dimension 2 ainsi que le taux expliqué de l'inertie totale. Expliquer sans faire de calculs.
- 4) Donner les expressions des coordonnées des individus dans le plan principal ainsi que celles des variables.
- 5) La réalisation de l'ACP sur le nuage des variables nous retourne les résultats suivants :

<div>Axes Variables</div>	1 ^{er} Axe	2 ^{ème} Axe
X^1	0.55	0.68
X^2	0.65	0.55
X^3	0.81	- 0.32
X^4	0.75	- 0.35
X^5	0.87	- 0.1
X^6	0.72	- 0.2
X^7	0.84	- 0.18
X^8	0.71	0.25

Que représentent les valeurs données par ce tableau ? Expliquer.

- 6) Visualiser graphiquement les variables initiales sur le plan factoriel c'est-à-dire le cercle de corrélation.
- 7) Interpréter les résultats obtenus. Plus précisément les relations entre les variables et les axes principaux.
- 8) Expliquer comment pouvons-nous exploiter l'analyse de cette représentation graphique des variables pour l'interprétation de la représentation des individus dans le plan choisi par l'ACP.