

Valeur propre et vecteur propre

- A chaque axe sont associés une *valeur propre* et un *vecteur propre*
 - Valeur propre (λ) : *variance* exprimée par l'axe
 - Vecteur propre (u) : *direction* de l'axe
- La variance exprimée par un *plan* formé par deux axes principaux est la somme des variances (valeurs propres) exprimées par ces axes
- Il y a autant de λ et de u que de variables d'origine :
“redécoupage” et hiérarchisation de la variance

Analyse en Composantes Principales

Principales étapes du calcul

- Matrice de données de départ : centrage ou centrage et réduction des données
- Calcul de la matrice de covariance (= matrice de corrélation si données standardisées, c'est le cas général)
- Calcul des valeurs propres et vecteurs propres à partir de cette matrice
- Calcul des nouvelles coordonnées des objets dans le repère formé par les axes principaux : ce sont les composantes principales
- Projection des points dans un espace réduit, en général les deux premiers axes
- Calcul des coordonnées des descripteurs originaux dans le nouveau repère

Analyse en Composantes Principales

Interprétation dans l'espace réduit

- Position des objets
 - Proximité des points : préserve la distance Euclidienne
- Position des descripteurs
 - Flèches se terminant sur les points-descripteurs
 - Angle entre les flèches : covariance ou corrélation ($\cos(\alpha)$)
 - Interpréter les descripteurs ayant suffisamment d'influence sur le plan de projection : *dépassant le cercle des contributions équilibrées* (longueur des axes si contribution égale à chaque dimension) et *proches du cercle de rayon 1* (pour une matrice de corrélation)

Analyse Factorielle des Correspondances

Principales étapes du calcul

- Transformation des fréquences absolues en probabilités :
fréquences relatives centrées

f_{ij}		$\Sigma = f_{i.}$
$\Sigma = f_{.j}$		$\Sigma = f_{..}$

Fréquences absolues

p_{ij} $= f_{ij} / f_{..}$		$p_{i.}$ $= f_{i.} / f_{..}$
$p_{.j}$ $= f_{.j} / f_{..}$		$\Sigma = 1$

Fréquences relatives



Q

q_{ij}		

Fréquences relatives centrées

Avec $q_{ij} = (p_{ij} - p_{i.}p_{.j}) / \sqrt{p_{i.}p_{.j}}$ soit
 - centrage (- fréquence théorique)
 - pondération

Analyse Factorielle des Correspondances

- Calcul de la matrice de *covariance* à partir de la matrice **Q**
- Calcul des *valeurs propres* et *vecteurs propres* à partir de cette matrice de covariance
- Calcul des nouvelles coordonnées des objets dans le repère formé par les axes principaux
- Projection des points dans un espace réduit, en général les deux premiers axes
- On peut placer les lignes et les colonnes du tableau de contingence dans le même espace factoriel

Analyse Factorielle des Correspondances

Interprétation dans l'espace réduit

- Variance du plan liée à celles des valeurs propres
- Comme en ACP, on peut ne placer que les sites (objets) ou les espèces (descripteurs) sur le diagramme ; ou les deux (biplot)
- Proximité entre objets et descripteurs : ressemblance. Attention, cela est fonction de la variance exprimée par le plan considéré
- Plus les points sont proches du centre, plus ils sont proches des caractéristiques moyennes de l'échantillon

Ordination sur matrices de distances

MDS métrique (ou Analyse en coordonnées principales)

Permet de résumer une matrice de distances en un nombre de dimensions prédéfinies