Analyse Factorielle Discriminante

Y qualitative : variable à expliquer

$x^1, \dots, x^p$: p variables quantitative explicative

vecteur des valeurs prises par la variable x^j

$$X = \begin{array}{c|c} & x^1 \quad \dots \quad x^p \\ p_1 & [\quad\quad\quad] \\ & [\quad\quad\quad] \\ x_n & \end{array} \quad \begin{array}{c} Y \\ \begin{bmatrix} 1 \\ 2 \\ \vdots \end{bmatrix} \end{array}$$

classe

q classes

↳ vecteur des valeurs prises par l'individus i pour les p var expli.

- p_i : masse de l'individus i avec $\sum_i p_i = 1$
- D_p : $Diag(p_i)$
- Π_x = nuage des x_i
- g = centre de gravité de $\Pi_x = \sum_{i \in I} p_i x_i$
- $V = X' D_p X$
- I_K : l'ensemble des individus ayant adopté la modalité K

  ↳ classe K

- $\boxed{n_K = |I_K|}$, $\boxed{m_K = \sum_{i \in I_K} p_i}$; $\boxed{g_K = \frac{1}{m_K} \sum_{i \in I_K} p_i x_i}$

  ↓ centre de gravité du nuage des indiv de I_K

- V_K : matrice de variance associée

$$G = \begin{array}{c} g_1 \\ \vdots \\ g_q \end{array} \begin{bmatrix} \quad\quad\quad \end{bmatrix} (q, p)$$

↓ tableau centre

centre de gravité

$\rightarrow \sum_{k=1}^{q} m_K g_K = 0$ ⟹ Matrice centrée

* Matrice de variance de G : B

$$B = G' D_m G$$

matrice de var interclasses

D_m : diag (m_K) → matrice des poids des centres de gravité

G : centrés $\begin{pmatrix} g_1 - g \\ g_2 - g \end{pmatrix}$

* Matrice variance intraclasses :

$$\boxed{W = \sum_{K=1}^{q} m_K V_K}$$

Relation fondamentale :

$$\boxed{V = B + W}$$

V → matrice variance totale

* Problème :

score ← $Z = b_1 x^1 + b_2 x^2 + \cdots + b_p x^p$ tq

C1/ les classes sont séparés de façon optimale
=> variance de Z entre les classes (interclasses) est max

C2/ Variance intra minimale

$Var(Z) = cov(Z, Z) = b' V b \qquad b' = (b_1, \ldots, b_p)$

$V = B + W$

$$Var(Z) = \underbrace{b' B b}_{\text{variance interclasses de } Z} + \underbrace{b' W b}_{\text{variance intraclasses}} = \sum_K m_K \left(\frac{1}{m_K} \sum_{i \in I_K} P_i (Z_i - \bar{Z}^K)^2 \right)$$

=> on cherche à maximiser $\boxed{\eta^2 = \dfrac{b' B b}{b' V b}}$

$$\begin{cases} V^{-1}Bb = \lambda b \Rightarrow b \text{ est un vecteur propre de } V^{-1}B \\ b'Vb = 1 \end{cases}$$

$$\boxed{\lambda = \frac{b'Bb}{b'Vb}}$$: $\lambda \in [0,1]$ solution du problème obtenue

$b'Vb$ ↓ $var(Z)$

* λ_1 : la plus grande valeur propre
* $Z_{(1)}$: score solution
* $b_{(1)}$ le vecteur propre (normé pour V) associé à λ_1

⟶ On cherchera ensuite un second score $Z_{(2)} = \sum_j b_{(2)j} x^j$ centré de la m^ manière non corrélé à $Z_{(1)}$.

⋮

$$Z_{(\alpha)} = \sum_j b_{(\alpha)} x^j$$

* r : rang de $V^{-1}B = rg(g^1, \dots, g^q)$

$$r \leq \min(p, q-1)$$

* $b_{(1)}, \dots, b_{(r)}$: facteurs discriminants
* $\lambda_1, \dots, \lambda_r$: pouvoirs discriminants

Axes factoriels discriminants I :

$\boxed{u^\alpha = Vb_{(\alpha)}}$ ⟶ vecteurs associés factoriels du nuage M_G

$BV^{-1}u^\alpha = \lambda_\alpha u^\alpha$ ⟶ analyse en composante principale sur les nuages des points.

- L'objectif de l'analyse discriminante est de produire un nouvel espace de représentation qui permet de distinguer le mieux les q classes.

Pourquoi l'ACP sur la matrice G : car on travaille sur des groupes / classes

- La démarche consiste à produire une suite de variables discriminantes Z_k non-corrélés deux à deux,

tel que : les indiv du m̂ grp projetés sur ces axes soient le plus proches possibles les un des autres Var intra $b'Wb$

- les individus de groupes différents soient le plus éloignés possibles. Var inter $b'Bb$ >>

$$\boxed{W = m_A V_A + m_B V_B}$$

- Vecteur axial factoriel discriminant

$$u = \frac{g_A - g_B}{\|g_A - g_B\|_{V^{-1}}}$$

facteur discriminant b est donné par :

$$\boxed{b = V^{-1} u} \neq \frac{1}{\lambda b}$$

- Valeur propre discriminante

$$\lambda = m_A m_B \|g_A - g_B\|^2_{V^{-1}}$$

$$\lambda = tr(BV^{-1})$$

- AFD descriptive → déterminer des nouvelles variables qui sont des combinaisons linéaires des anciennes

$$\boxed{z(\alpha_i) = \sum_j b_{\ell\alpha_j} x^j}$$

↳ qui séparent le mieux possible les classes

→ pour cela on effectue l'ACP sur le tableau G des centres de gravités des classes

(ici g^1 et g^2 c.d.g des classes I_1 et I_2)

avec la métrique V^{-1}

→ les nouvelles variables z_k appelées scores, sont les coord de tous les indiv x_i projetés après centrage sur les axes de cette AC

→ il n'y a que 2 centres de g_i

↳ un seul axe factoriel

↓

axe factoriel discriminant

1

2

3