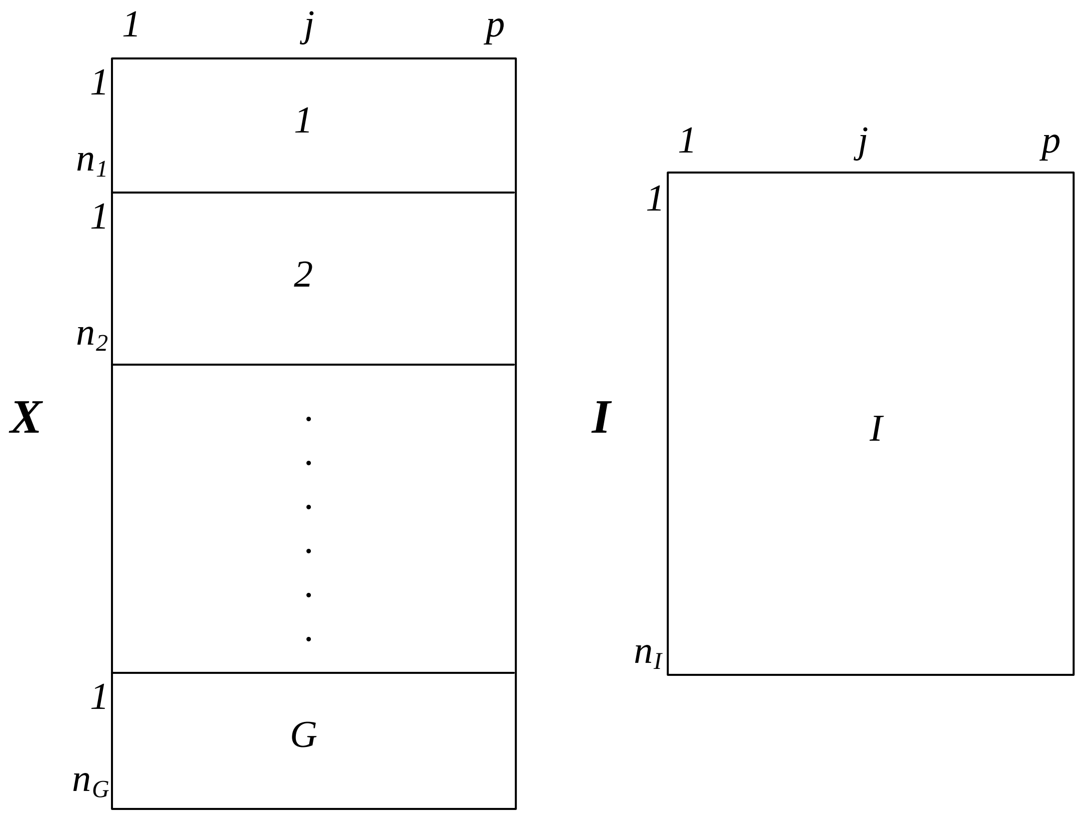
ANÁLISE DISCRIMINANTE

OBJECTIVO

1. Encontrar uma combinação linear das variáveis de partida que permita distinguir os grupos pré-definidos.
2. Classificar as amostras anónimas nos grupos pré-definidos.

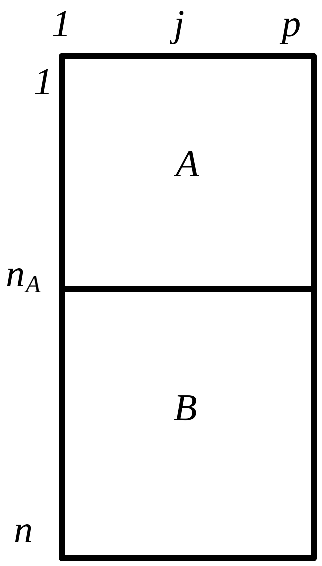


$$n = \sum_{k=1}^{G} n_k$$

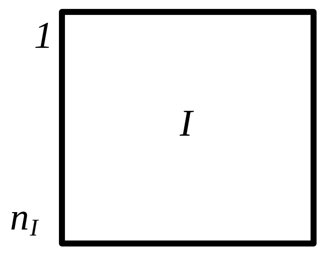
ANÁLISE DISCRIMINANTE

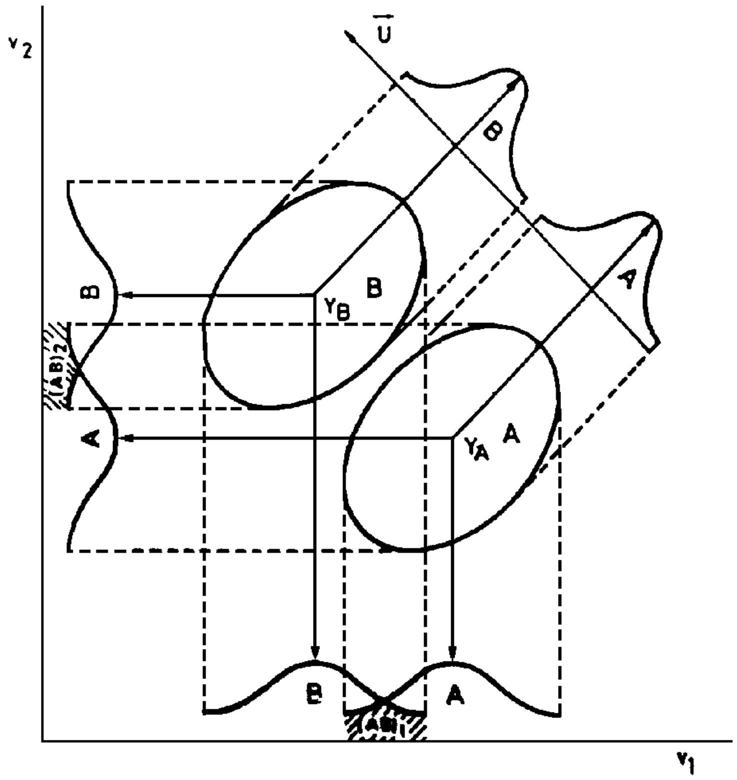
OBJECTIVO

Encontrar uma função discriminante que permita distinguir grupos de amostras, conhecidos *a priori* (*training set*). Esta função pode depois ser usada para classificar amostras anónimas.



$$n = n_A + n_B$$





ANÁLISE DISCRIMINANTE

Matriz de inércia total (**T**)

$$t_{jj'} = \frac{1}{n} \sum_{i=1}^{n} (x_{ij} - m_j)\,(x_{ij'} - m_{j'})$$

Matriz de inércia inter-classes (**B**)

$$b_{jj'} = \frac{n_A}{n}\,(m_{Aj} - m_j)\,(m_{Aj'} - m_{j'}) \;+\; \frac{n_B}{n}\,(m_{Bj} - m_j)\,(m_{Bj'} - m_{j'})$$

Matriz de inércia intra-classes (**W**)

$$w_{jj'} = \frac{1}{n} \left\{ \sum_{i=1}^{n_A} (x_{ij} - m_{Aj})\,(x_{ij'} - m_{Aj'}) + \sum_{i=1}^{n_B} (x_{ij} - m_{Bj})\,(x_{ij'} - m_{Bj'}) \right\}$$

$$\boxed{\mathbf{T} = \mathbf{B} + \mathbf{W}}$$

ANÁLISE DISCRIMINANTE

Matriz dos centros de gravidade (**M**) = matriz das médias

$$\mathbf{M} = \begin{bmatrix} m_{11} & \cdots & m_{1p} \\ \vdots & \ddots & \vdots \\ m_{G1} & \cdots & m_{Gp} \end{bmatrix}$$

$$m_{kj} = \frac{1}{n_k} \sum_{i=1}^{n_k} x_{ij} \qquad\qquad m_j = \frac{1}{n} \sum_{i=1}^{n} x_{ij}$$

Matriz de inércia total (**T**) = matriz variância - covariância

$$\mathbf{T} = \begin{bmatrix} s_{11}^2 & \cdots & s_{1p}^2 \\ \vdots & \ddots & \vdots \\ s_{p1}^2 & \cdots & s_{pp}^2 \end{bmatrix}$$

$$t_{jj'} = s_{jj'}^2 = \frac{1}{n} \sum_{i=1}^{n} (x_{ij} - m_j)\,(x_{ij'} - m_{j'})$$

Matriz de inércia inter-classes (**B**)

$$\mathbf{B} = \begin{bmatrix} b_{11} & \cdots & b_{1p} \\ \vdots & \ddots & \vdots \\ b_{p1} & \cdots & b_{pp} \end{bmatrix}$$

$$b_{jj'} = \sum_{k=1}^{G} \frac{n_k}{n} \, (m_{kj} - m_j)\,(m_{kj'} - m_{j'})$$

Matriz de inércia intra-classes (**W**)

$$\mathbf{W} = \begin{bmatrix} w_{11} & \cdots & w_{1p} \\ \vdots & \ddots & \vdots \\ w_{p1} & \cdots & w_{pp} \end{bmatrix}$$

$$w_{jj'} = \frac{1}{n} \sum_{k=1}^{G} \sum_{i=1}^{n_k} (x_{ij} - m_{kj})\,(x_{ij'} - m_{kj'})$$

$$\boxed{\mathbf{T} = \mathbf{B} + \mathbf{W}}$$

ANÁLISE DISCRIMINANTE

EIXOS DISCRIMINANTES

Os eixos discriminantes serão calculados de modo a maximizar o quociente entre a inércia inter-grupos e a inércia intra-grupos:

$$Max \ \frac{\hat{\mathbf{A}}}{\mathbf{W}}$$

A solução obtém-se extraindo os vectores e valores próprios da matriz $\mathbf{W}^{-1}\mathbf{B}$:

$$\mathbf{W}^{-1}\mathbf{B}\mathbf{u}_a - l \ \mathbf{u}_a$$

Notas:

- A matriz $\mathbf{W}^{-1}\mathbf{B}$ é não simétrica, logo os vectores próprios não são ortogonais.
- Os valores próprios $\lambda_\alpha \in [0, 1]$ medem o poder discriminante.

IMPORTÂNCIA DAS VARIÁVEIS

Coeficiente λ de Wilks

$\lambda \in [0, 1]$ dá o poder discriminatório do modelo depois da variável ter sido introduzida.

0 – discriminação perfeita

1 – nenhum poder discriminante

Coeficiente parcial λ de Wilks

$\lambda \in [0, 1]$ dá o poder discriminatório específico de cada variável

Tolerância

$Tol = 1 - r \in [0, 1]$ dá uma medida da redundância de cada uma das variáveis.

ANÁLISE DISCRIMINANTE

AGORITMO DA ANÁLISE DISCRIMINANTE

1. Encontrar os eixos discriminantes $\mathbf{u}_a$:

$$\mathbf{W}^{-1}\mathbf{B}\mathbf{u}_a - l\,\mathbf{u}_a$$

2. Seleccionar o número de eixos discriminantes:

$$100\;\frac{l_a}{\sum_{a}^{G-1} l_a}$$

3. Projectar as amostras nos eixos discriminantes:

$$\mathbf{Y}_a = u_{a0} + \sum_{j=1}^{p} u_{aj}\mathbf{X}_j$$

4. Encontrar, nos eixos discriminantes, as fronteiras que permitem minimizar a proporção de amostras do *training set* mal classificadas.

5. Validar os resultados.
   - Ponto fictício

6. Classificar as amostras anónimas.

ANÁLISE DISCRIMINANTE

CLASSIFICAÇÃO DAS AMOSTRAS ANÓNIMAS

- Coordenadas das amostras no eixos discriminantes
- Distância de Mahalanobis
- Máxima Verosimilhança
- Métodos bayesianos

CLASSIFICAÇÃO DAS AMOSTRAS ANÓNIMAS I_l

Critério da distância de Mahalanobis

- Calcular a distância de Mahalanobis para cada k:

$$D_k^2(I_l) = (I_l - M_k)' \, \mathbf{T}_k^{-1} \, (I_l - M_k)$$

- Atribuir I_l ao grupo k_0 mais próximo:

$$\min_k \left\{ D_k^2(I_l) \right\} \rightarrow k_0$$

Critério da máxima verosimilhança

- Calcular para cada k:

$$L(I_l|k) = f(I_l|k)$$

- Atribuir I_j ao grupo k_0 que maximiza o critério:

$$\max_k \{ L(I_l|k) \} \rightarrow k_0$$

ANÁLISE DISCRIMINANTE

CLASSIFICAÇÃO DAS AMOSTRAS ANÓNIMAS I_l

Métodos bayesianos

- Estabelecer a probabilidade a priori de cada k:

$$\text{p.e.:} \quad p(k) = \frac{1}{G}$$

- Calcular a probabilidade a posteriori para cada k:

$$p(k|I_l) = \frac{f(I_l|k)\,p(k)}{\sum_{k'=1}^{G} f(I_l|k')\,p(k')}$$

- Atribuir I_j ao grupo k_0 que maximiza o critério:

$$\max_k \{p(I_l|k)\} \rightarrow k_0$$