

Fonctions discriminantes

1

- N classes

- $g_1(\mathbf{x}), g_2(\mathbf{x}), \dots, g_N(\mathbf{x}) : \mathbb{R}^d \mapsto \mathbb{R}$

- fonction de décision:

$$f(\mathbf{x}) = C_i \text{ si } g_i(\mathbf{x}) \geq g_j(\mathbf{x}) \text{ pour tout } j = 1, \dots, N$$

- régions de décision: R_1, R_2, \dots, R_N

$$\mathbf{x} \in R_i \text{ si } f(\mathbf{x}) = C_i$$

- frontière de décision: $F \subset \mathbb{R}^d$

$$F = \{\mathbf{x} | \exists i, j : i \neq j, g_i(\mathbf{x}) = g_j(\mathbf{x})\}$$

Fonctions discriminantes

3

- La densité normale (gaussienne)

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}} \exp \left[-\frac{1}{2} (\mathbf{x} - \mu)' \Sigma^{-1} (\mathbf{x} - \mu) \right]$$

- μ : vecteur de moyenne

- Σ : matrice de covariance (positive semidefinite)

- $|\Sigma|$: déterminante (≥ 0)

- Σ^{-1} : inverse (existe)

Fonctions discriminantes

2

- Deux classes $\{-1, 1\}$

- $g(\mathbf{x}) = g_1(\mathbf{x}) - g_{-1}(\mathbf{x})$

- fonction de décision: $f(\mathbf{x}) = \text{signe}(g(\mathbf{x}))$

- régions de décision: $R_1 = \{\mathbf{x} : g(\mathbf{x}) \geq 0\}, R_{-1} = \{\mathbf{x} : g(\mathbf{x}) < 0\}$

- frontière de décision: $F = \{\mathbf{x} | g(\mathbf{x}) = 0\}$

Fonctions discriminantes

4

- Fonctions discriminantes

- $g_i(\mathbf{x}) = \ln p_i(\mathbf{x}) = \ln p(\mathbf{x} | Y = C_i) + \ln P(Y = C_i)$

- $g_i(\mathbf{x}) = -\frac{1}{2} (\mathbf{x} - \mu_i)' \Sigma_i^{-1} (\mathbf{x} - \mu_i) - \frac{d}{2} \ln 2\pi - \frac{1}{2} \ln |\Sigma_i| + \ln P(Y = C_i)$

- Cas 1: $\Sigma_i = \sigma^2 \mathbf{I}$ (matrices de covariance sphériques)

- $|\Sigma_i| = \sigma^{2d}$

- $\Sigma^{-1} = \frac{1}{\sigma^2} \mathbf{I}$

$$\begin{aligned} g_i(\mathbf{x}) &= -\frac{\|\mathbf{x} - \mu_i\|^2}{2\sigma^2} + \ln P(Y = C_i) + \text{const.} \\ &= -\frac{1}{2\sigma^2} (\mathbf{x}' \mathbf{x} - 2\mu_i' \mathbf{x} + \mu_i' \mu_i) + \ln P(Y = C_i) + \text{const.} \end{aligned}$$

Fonctions discriminantes

5

• Cas 1: $\Sigma_i = \sigma^2 \mathbf{I}$

• fonctions discriminantes **linéaires**: $g'_i(\mathbf{x}) = \mathbf{w}'_i \mathbf{x} + w_{i0}$

• vecteur de **poids**: $\mathbf{w}_i = \frac{1}{\sigma^2} \mu_i$

• **seuil** ou **biais**: $w_{i0} = -\frac{1}{2\sigma^2} \mu_i' \mu_i + \ln P(Y = C_i)$

• frontière de décision **linéaire** entre R_i et R_j
 $(\mu_i - \mu_j)'(\mathbf{x} - \mathbf{x}_0) = 0$

• $\mathbf{x}_0 = \frac{1}{2}(\mu_i + \mu_j) + \left[\frac{\sigma^2}{\|\mu_i - \mu_j\|^2} \ln \frac{P(Y = C_i)}{P(Y = C_j)} \right] (\mu_j - \mu_i)$

• $P(Y = C_i) = P(Y = C_j)$: classifieur de **plus proche moyenne**

Fonctions discriminantes

7

• Cas 2: $\Sigma_i = \Sigma$

• **frontière de décision** entre R_i et R_j :

$$\Sigma^{-1}(\mu_i - \mu_j)'(\mathbf{x} - \mathbf{x}_0) = 0$$

• $\mathbf{x}_0 = \frac{1}{2}(\mu_i + \mu_j) + \left[\frac{1}{(\mu_i - \mu_j)' \Sigma^{-1} (\mu_i - \mu_j)} \ln \frac{P(Y = C_i)}{P(Y = C_j)} \right] (\mu_j - \mu_i)$

• $P(Y = C_i) = P(Y = C_j)$: classifieur de plus proche moyenne selon la **distance de Mahalanobis**:

$$d(\mathbf{x}, \mu) = \sqrt{(\mathbf{x} - \mu)' \Sigma^{-1} (\mathbf{x} - \mu)}$$

Fonctions discriminantes

6

• Cas 2: $\Sigma_i = \Sigma$ (matrices de covariance identiques)

• $g_i(\mathbf{x}) = -\frac{1}{2}(\mathbf{x} - \mu_i)' \Sigma^{-1} (\mathbf{x} - \mu_i) + \ln P(Y = C_i) + const.$

• $g'_i(\mathbf{x}) = \mathbf{w}'_i \mathbf{x} + w_{i0}$

• vecteur de **poids**: $\mathbf{w}_i = \Sigma^{-1} \mu_i$

• **seuil** ou **biais**: $w_{i0} = -\frac{1}{2} \mu_i' \Sigma^{-1} \mu_i + \ln P(Y = C_i)$

Fonctions discriminantes

8

• Cas 3: $\Sigma_i = \text{arbitraire}$

• $g_i(\mathbf{x}) = -\frac{1}{2}(\mathbf{x} - \mu_i)' \Sigma_i^{-1} (\mathbf{x} - \mu_i) - \frac{1}{2} \ln |\Sigma_i| + \ln P(Y = C_i) + const.$

• fonctions discriminantes **quadratiques**: $g_i = \mathbf{x}' \mathbf{W}_i \mathbf{x} + \mathbf{w}'_i \mathbf{x} + w_{i0}$

• vecteur de **poids**: $\mathbf{w}_i = \Sigma_i^{-1} \mu_i$

• $\mathbf{W}_i = -\frac{1}{2} \Sigma_i^{-1}$

• **seuil** ou **biais**: $w_{i0} = -\frac{1}{2} \mu_i' \Sigma_i^{-1} \mu_i - \frac{1}{2} \ln |\Sigma_i| + \ln P(Y = C_i)$

• **frontière de décision** entre R_i et R_j : **hyperquadriques**