TEMA 2.1 MAP

Ejercicio 1 (OPCIONAL)

Demuestre que el criterio MAP $\hat{\omega}_i = \max_{\omega_i}\left\{\Pr\left(\omega_i \,|\mathbf{x}\right)\right\}$ es equivalente a minimizar la probabilidad de error del clasificador (Páginas 17, 18).

Ejercicio 1 (OPCIONAL)

Obtenga la f.d.p del vector aleatorio $\mathbf{y} = \mathbf{U}^{\mathbf{T}}\mathbf{x} \in \mathbb{R}^k$, donde $\mathbf{x} \in \mathbb{R}^d$; $\mathbf{x} : N(0, \mathbf{\Sigma}_x)$; $\mathbf{\Sigma}_x = \mathbf{U\Lambda U}^{\mathbf{T}}$ y $\mathbf{U} = \left(\mathbf{v}_1, \mathbf{v}_2, .., \mathbf{v}_k\right)$; $\mathbf{U}^{\mathbf{T}}\mathbf{U} = \mathbf{I}_k$; $k < d$ $\mathbf{\Lambda} = diag(\lambda_1, \lambda_2, ..., \lambda_k)$ y comente los resultados.

Ejercicio 3

Demuestre que el hiperplano (Criterio MAP) $h_i(\mathbf{x}) = h_j(\mathbf{x}) \Rightarrow \mathbf{w}^{\mathbf{T}}\left(\mathbf{x} - \mathbf{x}_0\right) = 0$ que separa dos zonas de decisión correspondientes a dos clases con distribuciones $\mathbf{x}_i : N(\boldsymbol{\mu}_i, \sigma^2\mathbf{I})$ se determina mediante:

$$\mathbf{w} = \boldsymbol{\mu}_i - \boldsymbol{\mu}_j;$$

$$\mathbf{X}_0 = \tfrac{1}{2}\left(\boldsymbol{\mu}_i + \boldsymbol{\mu}_j\right) - \frac{\sigma^2}{\left\|\boldsymbol{\mu}_i - \boldsymbol{\mu}_j\right\|^2}\ln\left(\frac{\Pr(\omega_i)}{\Pr(\omega_j)}\right)\left(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j\right)$$

OPCIONAL: Realizar un programa en MATLAB para d=2, que dibuje las dos regiones de decisión, en función de los parámetros de entrada: $\mu_i, \sigma; \Pr\left(\omega_1\right); \Pr\left(\omega_2\right)$. Recuerde que $\Pr\left(\omega_1\right) + \Pr\left(\omega_2\right) = 1$.

Ejercicio 4

Demuestre que el hiperplano (Criterio MAP) $h_i(\mathbf{x}) = h_j(\mathbf{x}) \Rightarrow \mathbf{w}^{\mathbf{T}}\left(\mathbf{x} - \mathbf{x}_0\right) = 0$ que separa dos zonas de decisión correspondientes a dos clases con distribuciones $\mathbf{x}_i : N(\boldsymbol{\mu}_i, \mathbf{\Sigma})$ se determina mediante:

$$\mathbf{w} = \mathbf{\Sigma}^{-1}\left(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j\right);$$

$$\mathbf{X}_0 = \tfrac{1}{2}\left(\boldsymbol{\mu}_i + \boldsymbol{\mu}_j\right) - \frac{\ln\left(\Pr(\omega_i)\right) - \ln\left(\Pr(\omega_j)\right)}{\left(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j\right)^{\mathbf{T}}\mathbf{\Sigma}^{-1}\left(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j\right)}\left(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j\right)$$

OPCIONAL: Realizar un programa en MATLAB para d=2, que dibuje las dos regiones de decisión, en función de los parámetros de entrada: $\mu_i; \Sigma = \sigma^2\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}; \Pr\left(\omega_1\right); \Pr\left(\omega_2\right)$. Recuerde que $\Pr\left(\omega_1\right) + \Pr\left(\omega_2\right) = 1$.

Ejercicio 5

Si las componentes del vector x, son de valores binarios para el caso binario de C=2 categorías y dimensión d, con componentes estadísticamente independientes entre sí.

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_d \end{pmatrix}, \qquad \begin{array}{l} p_i = \Pr\left(x_i = 1 \middle| \omega_1\right) = 1 - \Pr\left(x_i = 0 \middle| \omega_1\right) \\ q_i = \Pr\left(x_i = 1 \middle| \omega_2\right) = 1 - \Pr\left(x_i = 0 \middle| \omega_2\right) \end{array}, \qquad \begin{array}{l} \Pr\left(\mathbf{x} \middle| \omega_1\right) = \prod_{i=1}^{d} \left(p_i\right)^{x_i} \left(1 - p_i\right)^{1-x_i} \\ \Pr\left(\mathbf{x} \middle| \omega_2\right) = \prod_{i=1}^{d} \left(q_i\right)^{x_i} \left(1 - q_i\right)^{1-x_i} \end{array}$$

a) Demuestre que la función discriminate al aplicar MAP es de la forma

$$g(\mathbf{x}) = \mathbf{w}^{\mathbf{T}}\mathbf{x} + w$$

b) Identifique los valores del vector y del escalar en la función anterior.

Suponga a partir de este punto que $p_i = 1 - q_i = p; \quad i = 1, \ldots .d \quad \Pr\left(\omega_1\right) = \Pr\left(\omega_2\right) = \frac{1}{2}$

c) Demuestre que la decisión que minimiza el error de clasificación es $\sum_{i=1}^{d} x_i \underset{\omega_2}{\overset{\omega_1}{\gtrless}} \frac{d}{2}$.

Suponga para ello d impar.

d) Demuestre que la mínima probabilidad de error viene dada por $P_e\left(d, p\right) = \sum_{k=0}^{\frac{d-1}{2}} \begin{pmatrix} d \\ k \end{pmatrix} p^k \left(1 - p\right)^{d-k}$

Ejercicio 6

Si las componentes del vector x, son de valores ternarios para el caso binario de C=2 categorías y dimensión d, con componentes estadísticamente independientes entre sí.

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_d \end{pmatrix}, \qquad \begin{array}{l} a_{1i} = \Pr\left(x_i = 1 \middle| \omega_1\right); b_{1i} = \Pr\left(x_i = 0 \middle| \omega_1\right); c_{1i} = \Pr\left(x_i = -1 \middle| \omega_1\right); \\ a_{2i} = \Pr\left(x_i = 1 \middle| \omega_2\right); b_{2i} = \Pr\left(x_i = 0 \middle| \omega_2\right); c_{2i} = \Pr\left(x_i = -1 \middle| \omega_2\right); \end{array}$$

a) Calcule las funciones de probabilidad condicionadas: $\Pr\left(\mathbf{x} \middle| \omega_1\right); \Pr\left(\mathbf{x} \middle| \omega_2\right)$

b) Demuestre que la función discriminate al aplicar MAP es de la forma

$$g(\mathbf{x}) = \mathbf{x}^{\mathbf{T}}\mathbf{A}\mathbf{x} + \mathbf{w}^{\mathbf{T}}\mathbf{x} + w$$

Identifique los valores de la matriz, del vector y del escalar en la función anterior.

OPCIONAL: Realizar un programa en MATLAB para d=2, que dibuje las dos regiones de decisión, en función de los parámetros de entrada: $a_{1i}; b_{1i}; c_{1i}; a_{2i}; b_{2i}; c_{2i}; \Pr\left(\omega_1\right); \Pr\left(\omega_2\right)$. Recuerde que $a_{1i} + b_{1i} + c_{1i} = 1$, $a_{2i} + b_{2i} + c_{2i} = 1$, $\Pr\left(\omega_1\right) + \Pr\left(\omega_2\right) = 1$.