Funciones discriminantes para la función de densidad normal

Funciones discriminantes para la función de densidad normal

Tags: Reconocimiento de Patrones, Clasificación, funciones discriminantes y superficies de decision

---

Funciones discriminantes para la función de densidad normal

A partir de la clasificación de tasa menor mínima

$$g_i(x) = Ln(P(x|w_i)) + Ln(P(w_i)) \qquad (2.3.18)$$

$$g_i(x) = Ln[P(x|w_i)] + Ln[P(w_i)] \qquad (2.3.7)$$

Si $P(x|w_i) \approx N(\mu_i, \sum_i)$

$$g_i(x) = Ln[\frac{1}{(2\pi)^{d/2}|\sum_i|^{-1/2}} e^{1/2[(x-\mu_i)^t \sum_i^{-1}(x-\mu)]}] + Ln(P(w_i)) \qquad (2.3.19)$$

$$g_i(x) = Ln[\frac{1}{(2\pi)^{d/2}|\sum_i|^{-1/2}}] + Ln[e^{1/2[(x-\mu_i)^t \sum_i^{-1}(x-\mu)]}] + Ln(P(w_i))$$

$$g_i(x) = \cancel{Ln(1)} - Ln[(2\pi)^{d/2}|\sum_i|^{1/2}] + (-1/2)[(x-\mu_i)^t \sum_i^{-} 1(x-\mu_i)] + Ln(P(w_i))$$

$$g_i(x) = (-1/2)[(x-\mu_i)^t \sum_i^{-} 1(x-\mu_i)] - Ln[(2\pi)^{d/2}] - Ln[|\sum_i|^{1/2}] + Ln(P(w_i))$$

(2.3.20)

$$g_i(x) = (-1/2)[(x-\mu_i)^t \sum_i^{-1}(x-\mu_i)] - (d/2)Ln(2\pi) - (1/2)Ln(|\sum_i|) + Ln(P(w_i)) \quad (2.3$$

Caso 1 : $\sum_i = \sigma^2 I$

Cuando las características son estadísticamente independientes y cuando cada característica tiene la misma varianza σ^2. En este caso, la matriz de covarianzas es:

$$\sum_i = \begin{bmatrix} \sigma^2 & \dots & 0 \\ 0 & \sigma^2 & \dots \\ 0 & \dots & \sigma^2 \end{bmatrix}$$

- Determinante

$$|\sum_i| = (\sigma^2)^d = \sigma^{2d} \text{ y } \Sigma_i^{-1} = \frac{1}{\sigma^2} I$$

Ya que $(d/2)Ln(2\pi)$ y el determinante $|\sum_i|$ no dependen de i, estos términos se pueden despreciar.

$$g_i(x) = (-1/2)[(x-\mu_i)^t \frac{1}{\sigma^2} I (x-\mu_i)] + Ln(P(w_i))$$

$$g_i(x) = -\frac{(x-\mu_i)^t (x-\mu_i)}{2\sigma^2} + Ln(P(w_i))$$

$$g_i(x) = -\frac{(x^t x - 2\mu_i^t + \mu_i^t \mu_i)}{s\sigma^2} + Ln(P(w_i))$$

Ya que $x^t x$ es la misma para toda i, este termino se ignora

$$g_i(x) = \frac{2\mu_i^t x - \mu_i^t \mu_i}{s\sigma^2} + Ln(P(w_i))$$

Caso 1 ($2.3.21$)

$$g_i(x) = \frac{\mu_i^t x}{\sigma^2} - \frac{\mu_i^t \mu_i}{2\sigma^2} + Ln(P(w_i) \qquad (2.3.21)$$

Caso 2: $\sum_i = \sum$

Matrices idénticas pero arbitrarias, en este caso $|\sum_i|$ y el termino $(d/2)Ln(2\pi)$ son independientes de i, por lo tanto, se ignoran. De esta manera:

Caso 2 $(2.3.22)$

$$g_i(x) = (-1/2)(x-\mu_i)^t \sum^{-1} (x-\mu_i) + Ln(P(w_i)) \qquad (2.3.22)$$

Si $P(w_i)$ es igual para todas las clases, entonces el termino $Ln(P(w_i))$ se ignora.

Caso 3: $\sum_i$ arbitraria

Solo el termino $(d/2)Ln(2\pi)$ puede omitirse

Caso 3 ($2.3.23$)

$$g_i(x) = (-1/2)(x-\mu_i)^t \sum_i^{-1}(x-\mu_i) - (1/2)Ln(|\sum_i|) + Ln(P(w_i)) \qquad (2.3.23)$$

Ejemplo:

Suponemos 2 categorías en 2 dimensiones con las siguientes distribuciones de probabilidad:

$$P(x|w_i) \approx N(\mu_i, \Sigma_i) \qquad (2.3.15)$$
$$P(x|w_1) \approx N(0, I) \rightarrow Lubina$$
$$P(x|w_2) \approx N(\binom{1}{1}, I)$$

con $p(w_i)$ = $1/2$, para $i = 1, 2$

- Clasificar el punto $x = \binom{0.3}{0.3}$ para una probabilidad mínima de error.

Matrices de covarianzas

$$\sum_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \sum_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \rightarrow \text{Matriz de identidad } I. \text{ Caso 2 , mismas matrices}$$

- Determinante

$$|\sum_1| = |\sum_2| = (1)(1) - (0)(0) = 1$$

- Inversa

$$\sum_1^{-1} = \sum_2^{-1} = I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$g_i(x) = (-1/2)(x-\mu_i)^t \sum^{-1}(x-\mu_i) + \cancel{Ln(P(w_i))} \rightarrow \text{Caso 2} \qquad (2.3.22)$$

Como $P(w_1) = P(w_2)$ se desprecia el termino $Ln(P(w_i))$

- Para $i = 1$

$$g_i(x) = (-1/2)(x - \mu_i)^t \sum^{-1}(x - \mu_i) \qquad (2.3.22)$$

$$g_1(x) = (-1/2)[\begin{pmatrix} 0.3 \\ 0.3 \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \end{pmatrix}]^t \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}(\begin{pmatrix} 0.3 \\ 0.3 \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \end{pmatrix})$$

$$g_1(x) = (-1/2)[0.3, 0.3] \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0.3 \\ 0.3 \end{pmatrix}$$

$$g_1(x) = (-1/2)[0.3, 0.3] \begin{pmatrix} 0.3 \\ 0.3 \end{pmatrix}$$

$$g_1(x) = -0.09$$

- Para $i = 2$

$$g_2(x) = (-1/2)[\begin{pmatrix} 0.3 \\ 0.3 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \end{pmatrix}]^t \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}(\begin{pmatrix} 0.3 \\ 0.3 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \end{pmatrix})$$

$$g_2(x) = (-1/2)[-0.7, -0.7] \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -0.7 \\ -0.7 \end{pmatrix}$$

$$g_2(x) = (-1/2)[-0.7, -0.7] \begin{pmatrix} -0.7 \\ -0.7 \end{pmatrix}$$

$$g_2(x) = -0.49$$

$$g_i(x) = (-1/2)(x - \mu_i)^t \sum_i^{-1}(x - \mu_i) - (1/2)Ln(|\sum_i|) + Ln(P(w_i)) \qquad (2.3.23)$$

References

Regiones de decision