

BAYES' SINIFLANDIRICI



Kayıp Fonksiyonu

Koşullu ortalama kayıp veya koşullu ortalama risk $r_k(x)$ şu şekilde tanımlanabilir:

$$r_k(x) = \sum_{i=1}^M L_{ik} p(\omega_i | x)$$

Bu, x 'in ω_k sınıfında olarak (aslında başka bir sınıfta (ω_i $i=1,2,\dots,M$ ve $i \neq k$) olmasına rağmen) yanlış sınıflandırılmasının beklenen veya ortalama değeridir.

Karar verme kuralı:

- Yanlış sınıflandırmalardan kaynaklanan beklenen kayıpları hesapla.,
- Eğer $r_k(x) \leq r_i(x) \quad \forall i, i \neq k$ ise $x \in \omega_k$ olarak karar ver.

$$d_k(x) = -r_k(x)$$

$$L = \begin{bmatrix} 0 & \dots & 1 \\ \vdots & & \vdots \\ 1 & \dots & 0 \end{bmatrix} \quad L_{ik} = 1 - \delta(k-i)$$
$$\delta(k-i) = \begin{cases} 1 & \text{eğer } k = i \text{ ise} \\ 0 & \text{aksi halde} \end{cases}$$

Simetrik kayıp matrisi

Kayıp Fonksiyonu

Negatif Kayıp matrisi:

$$L_{\text{neg}} = \begin{bmatrix} -h_1 & & & 0 \\ & -h_2 & & \cdot \\ & & -h_3 & \cdot \\ & & & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & -h_N \end{bmatrix} \quad L_{ik} = \begin{cases} -h_i & i = k \\ 0 & i \neq k \end{cases}$$

Matristeki h_i 'nin aynı olma zorunluluğu yok

$$L = \begin{bmatrix} 0 & L_{21} \\ L_{12} & 0 \end{bmatrix} \quad L_{ik}, x' \text{in } \omega_i \text{'de olması gerektiğini ama } \omega_k \text{'da olarak (yanlış) sınıflandırıldığı anlamına gelir.}$$

Bayes' Ayırma Fonksiyonu

$$p(\omega_i | x) = \frac{p(x | \omega_i) p(\omega_i)}{p(x)}$$

$$r_k(x) = \frac{1}{p(x)} \sum_{i=1}^M L_{ik} p(x | \omega_i) p(\omega_i)$$

$p(x)$ bütün $r_j(x)$, $j=1, \dots, M$ ler için aynı olduğundan kaldırılabilir:

$$\min_k r_k(x) = \min_k \sum_{i=1}^M L_{ik} p(x | \omega_i) p(\omega_i)$$

İki sınıflı problem için

$$\begin{vmatrix} r_1 \\ r_2 \end{vmatrix} = \begin{vmatrix} L_{11} & L_{21} \\ L_{12} & L_{22} \end{vmatrix} \begin{vmatrix} p(x | \omega_1) p(\omega_1) \\ p(x | \omega_2) p(\omega_2) \end{vmatrix}$$

$$x \in \omega_1 \text{ eğer } \frac{p(x | \omega_1)}{p(x | \omega_2)} > \frac{(L_{21} - L_{22})p(\omega_2)}{(L_{12} - L_{11})p(\omega_1)}$$

$$L_{11}p(x | \omega_1)p(\omega_1) + L_{21}p(x | \omega_2)p(\omega_2) < L_{12}p(x | \omega_1)p(\omega_1) + L_{22}p(x | \omega_2)p(\omega_2)$$

Çok sınıflı problem için

$$x \in \omega_1, \text{ eğer } l_{12}(x) > \theta_{12}$$

$$l_{ki} = \frac{p(x | \omega_k)}{p(x | \omega_i)} \quad \theta_{ki} = \frac{(L_{ik} - L_{ii})p(\omega_i)}{(L_{ki} - L_{kk})p(\omega_k)}$$

Eğer her i için $l_{ki} > \theta_{ki}$ ise $x \in \omega_k$ dır.

Bayes' Ayırma Fonksiyonu

Simetrik kayıp matrisi için:

$$\frac{p(x | \omega_k)}{p(x | \omega_i)} > \frac{p(\omega_i)}{p(\omega_k)}$$

Diğer bir deyişle şu koşulun sağlanması durumunda $x \in \omega_k$ diyebiliriz.

$$p(x | \omega_k)p(\omega_k) > p(x | \omega_i)p(\omega_i) \quad \forall i$$

Ayırma fonksiyonu:

$$d_k(x) = p(x | \omega_k)p(\omega_k)$$

$$d_k^*(x) = \log p(x | \omega_k) + \log p(\omega_k)$$

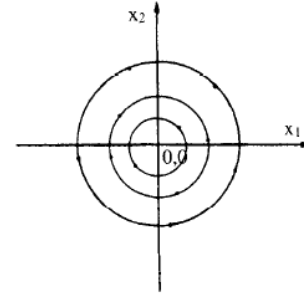
Normal Dağılım

$$p(x | \omega_k) = \frac{1}{(2\pi)^{n/2} |C_k|^{1/2}} \exp \left[-\frac{1}{2} (x - m_k)^T C_k^{-1} (x - m_k) \right]$$
$$= N(m_k, C_k) \quad k = 1, 2, \dots, M;$$

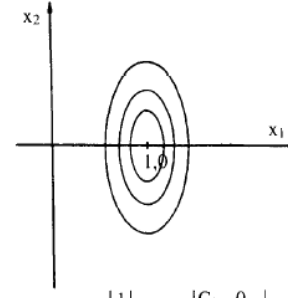
n =örüntü vektörünün boyutudur.

N , normal yoğunluk fonksiyonunun, m_k ortalama vektörünün ve C_k da k sınıfı için kovaryans matrisinin k sınıfına ait örüntülere göre beklenen değerlerini tanımlar.

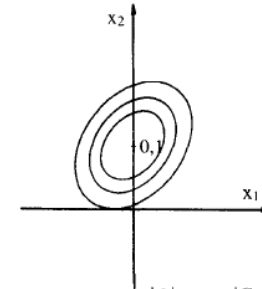
- $m_k = E_k[x]$ ve
- $C_k = E_k[(x - m_k)(x - m_k)^T]$



(a) $m = 0$
 $C = I$



(b) $m = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $C = \begin{bmatrix} C_{11} & 0 \\ 0 & C_{22} \end{bmatrix}$
 $C_{22} > C_{11}$



(c) $m = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $C = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}$

En Uygun Ayırma Fonksiyonu

$$d_k^*(x) = \log p(x | \omega_k) + \log p(\omega_k)$$

$$d_k^{**}(x) = -\frac{n}{2} \log(2\pi) - \frac{1}{2} \log |C_k| - \frac{1}{2} (x - m_k)^T C_k^{-1} (x - m_k) + \log p(\omega_k)$$

sağ taraftaki ilk terim bütün k'lar için aynı olduğundan elimine edilebilir:

$$d_k^{***}(x) = -\frac{1}{2} (x - m_k)^T C_k^{-1} (x - m_k) + \left[\log p(\omega_k) - \frac{1}{2} \log |C_k| \right]$$

$$d_k^{(iv)}(x) = -\frac{1}{2} r + f(k) \quad x \in \omega_k \text{ için}$$

$$r = (x - m_k)^T C_k^{-1} (x - m_k) \quad f(k) = \log p(\omega_k) - \frac{1}{2} \log |C_k|$$

Durum 1: ($C_i = C_j = C_k = C$) ise:

□ Genel eşitliği açarsak:

$$d_k(x) = -\frac{1}{2} x^T C^{-1} x - \frac{1}{2} m_k^T C^{-1} m_k + x^T C^{-1} m_k + \log p(\omega_k) - \frac{1}{2} \log |C|$$

ilk ve son terimler bütün sınıflar için aynıdır:

$$d_k(x) = x^T C^{-1} m_k + \left[\log p(\omega_k) - \frac{1}{2} m_k^T C^{-1} m_k \right] \quad k = 1, 2, \dots, M$$

B bu bir doğrusal ayırma fonksiyonudur

En Uygun Ayırma Fonksiyonu

Durum 2: $\sigma_k^2 = |C_k|$ ise:

$$d_k(x) = -\frac{1}{2} \frac{(x - m_k)^T (x - m_k)}{\sigma_k^2} + \left[\log p(\omega_k) - \frac{1}{2} \log \sigma_k^2 \right]$$

her bir özellik aynı varyansa sahip olduğunda (Durum 1 ile birleşimi):

$$d_k(x) = -\frac{1}{2} \frac{x^T x - 2x^T m_k + m_k^T m_k}{\sigma^2} + \log p(\omega_k) - \frac{1}{2} \log \sigma^2$$

$x^T x$ ve $\frac{1}{2} \log \sigma^2$ bütün k lar için aynı olduğundan:

$$d_k(x) = \frac{x^T m_k}{\sigma^2} + \left[\log p(\omega_k) - \frac{m_k^T m_k}{2\sigma^2} \right]$$

Ek olarak bütün k lar için $p(\omega_k) = 1/k$ varsayımı da yapılırsa "log $p(\omega_k)$ " terimi $d_k(x)$ ifadesinden kaldırılabilir:

$$d_k(x) = x^T m_k - \frac{1}{2} |m_k|^2$$

Diğer bir ifadeyle minimum distance sınıflandırıcı elde edilir:

$$d_k(x) = -(x - m_k)^T (x - m_k) = -|x - m_k|^2$$

Örnek

Örüntü noktaları $(1,2)^T$, $(2,2)^T$, $(3,1)^T$, $(3,2)^T$ ve $(2,3)^T$ ün ω_1 sınıfına ait olduğu bilinsin. Diğer bir noktalar kümesi $(7,9)^T$, $(8,9)^T$, $(9,8)^T$, $(9,9)^T$ ve $(8,10)^T$ da ω_2 sınıfında olsun. Bu kümeleri birbirinden ayırt edecek Bayes' karar sınırını bulunuz..

$$m_1 = \frac{1}{N_1} \sum_{j=1}^{N_1} x_{1j} = \frac{1}{5} \begin{bmatrix} 11 \\ 10 \end{bmatrix} \quad m_2 = \frac{1}{N_2} \sum_{j=1}^{N_2} x_{2j} = \frac{1}{5} \begin{bmatrix} 41 \\ 45 \end{bmatrix}$$

$$C = E[(x - m)(x - m)^T] = E[xx^T] - mm^T$$

$$C_1 = \frac{1}{N_1} \sum_{j=1}^{N_1} x_{1j} x_{1j}^T - m_1 m_1^T$$

$$\begin{aligned} C_1 &= \frac{1}{5} \left[\begin{pmatrix} 1 \\ 2 \end{pmatrix} (1 \ 2) + \begin{pmatrix} 2 \\ 2 \end{pmatrix} (2 \ 2) + \begin{pmatrix} 3 \\ 1 \end{pmatrix} (3 \ 1) \right. \\ &\quad \left. + \begin{pmatrix} 3 \\ 2 \end{pmatrix} (3 \ 2) + \begin{pmatrix} 2 \\ 3 \end{pmatrix} (2 \ 3) \right] - \frac{1}{25} \begin{pmatrix} 11 \\ 10 \end{pmatrix} (11 \ 10) \\ &= \frac{1}{25} \begin{pmatrix} 14 & -5 \\ -5 & 10 \end{pmatrix} \end{aligned}$$

$$\begin{aligned} C_2 &= \frac{1}{N_2} \sum_{j=1}^{N_2} x_{2j} x_{2j}^T - m_2 m_2^T \\ &= \frac{1}{25} \begin{pmatrix} 14 & -5 \\ -5 & 10 \end{pmatrix} \end{aligned}$$

$$C_1 = C_2 = C = \frac{1}{25} \begin{pmatrix} 14 & -5 \\ -5 & 10 \end{pmatrix}$$

Örnek

$$C^{-1} = \frac{1}{|C|} \text{adj} C = \frac{1}{115} \begin{pmatrix} 10 & 5 \\ 5 & 14 \end{pmatrix}$$

$$C^{-1}m_1 = \frac{1}{115} \begin{pmatrix} 32 \\ 39 \end{pmatrix} \quad d_1(x) = x^T C^{-1} m_1 - \frac{1}{2} m_1^T C^{-1} m_1$$

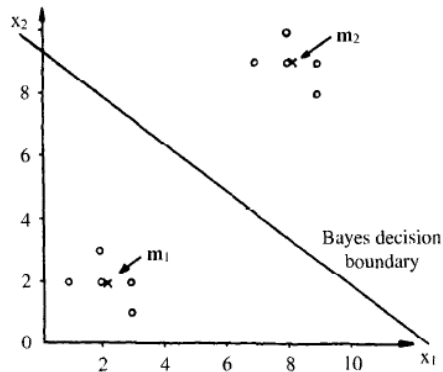
$$m_1^T C^{-1} m_1 = \frac{742}{5 \times 115} \quad = \frac{32}{115} x_1 + \frac{39}{115} x_2 - 0.65$$

$$C^{-1}m_2 = \frac{1}{115} \begin{pmatrix} 127 \\ 167 \end{pmatrix} \quad d_2(x) = x^T C^{-1} m_2 - \frac{1}{2} m_2^T C^{-1} m_2$$

$$m_2^T C^{-1} m_2 = 22 \quad = \frac{127}{115} x_1 + \frac{169}{115} x_2 - 11$$

$$d(x) = d_1(x) - d_2(x) = 0$$

$$d(x) = -0.826x_1 - 1.11x_2 + 10.35 = 0$$





Kaynak:

Sing T. Bow, “**Pattern Recognition and Image Preprocessing**”, CRC Press, 2002.