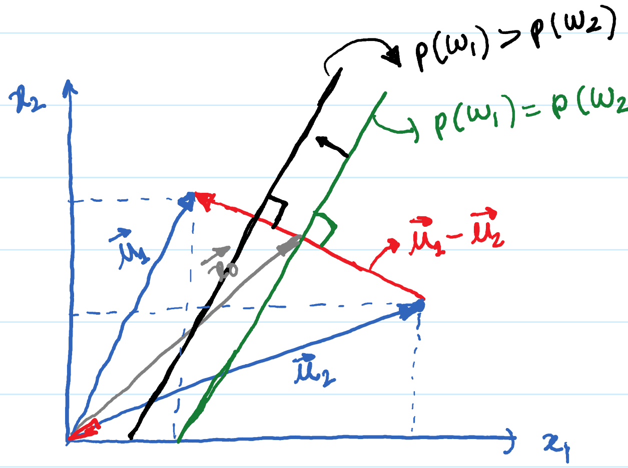
Bayesian Classification Part-10

03 March 2022 09:54

Decision Hyperplanes for Diagonal Covariance matrix:

$$g_{ij}(\vec{x}) = (\vec{\mu}_i - \vec{\mu}_j)^T \left[\vec{x} - \left(\frac{1}{2}(\vec{\mu}_i + \vec{\mu}_j) - \sigma^2 \ln\left(\frac{P(\omega_i)}{P(\omega_j)}\right) \frac{(\vec{\mu}_i - \vec{\mu}_j)}{(\vec{\mu}_i - \vec{\mu}_j)^T(\vec{\mu}_i - \vec{\mu}_j)} \right) \right]$$



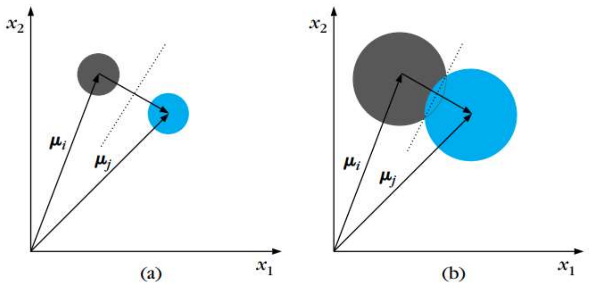


FIGURE 2.11

Decision line (a) for compact and (b) for noncompact classes. When classes are compact around their mean values, the location of the hyperplane is rather insensitive to the values of $P(\omega_1)$ and $P(\omega_2)$. This is not the case for noncompact classes, where a small movement of the hyperplane to the right or to the left may be more critical.

Decision Hyperplanes for Nondiagonal Covariance matrix:

$$g_i(\vec{x}) = \vec{\mu}_i^T \Sigma^{-1} \vec{x} + \ln(P(\omega_i)) - \frac{1}{2}\vec{\mu}_i^T \Sigma^{-1} \vec{\mu}_i$$

Hyperplane:

$$g_{ij}(\vec{x}) = g_i(\vec{x}) - g_j(\vec{x}) = 0$$

$$\Rightarrow (\vec{\mu}_i - \vec{\mu}_j)^T \Sigma^{-1} \vec{x} - \frac{1}{2}(\vec{\mu}_i - \vec{\mu}_j)^T \Sigma^{-1} (\vec{\mu}_i + \vec{\mu}_j) + \ln\left(\frac{P(\omega_1)}{P(\omega_2)}\right) = 0$$

$$\Rightarrow (\vec{\mu}_i - \vec{\mu}_j)^T \Sigma^{-1} \left(\vec{x} - \left[\frac{1}{2}(\vec{\mu}_i + \vec{\mu}_j) - \ln\left(\frac{P(\omega_1)}{P(\omega_2)}\right) \frac{(\vec{\mu}_i - \vec{\mu}_j)}{(\vec{\mu}_i - \vec{\mu}_j)^T \Sigma^{-1} (\vec{\mu}_i - \vec{\mu}_j)} \right] \right) = 0$$

$$\Rightarrow \Big(\underbrace{\Sigma^{-1}(\vec{\mu}_i-\vec{\mu}_j)}_{W}\Big)^T\Big(\vec{x} - \underbrace{\Big[\tfrac{1}{2}(\vec{\mu}_i+\vec{\mu}_j) - \ln\Big(\frac{P(W_1)}{P(W_2)}\Big)\frac{\vec{\mu}_i-\vec{\mu}_j}{\|\vec{\mu}_i-\vec{\mu}_j\|^2_{\Sigma^{-1}}}\Big]}_{\vec{x}_0}\Big) = 0,$$

where $\|\vec{y}\|_{\Sigma^{-1}} = (\vec{y})^T\Sigma^{-1}\vec{y} = \Sigma^{-1}$ Norm of $\vec{y}$.

The decision hyperplane is no longer orthogonal to $(\vec{\mu}_i-\vec{\mu}_j)$, but orthogonal to $\Sigma^{-1}(\vec{\mu}_i-\vec{\mu}_j)$.

Minimum Distance Classifier:

Assume equi-probable classes with same covariance matrix.

$$\Rightarrow P(\omega_1) = P(\omega_2) = \cdots = P(\omega_M) = \frac{1}{M}$$

$$\Rightarrow g_i(\vec{x}) = -\frac{1}{2}(\vec{x}-\vec{\mu}_i)^T\Sigma^{-1}(\vec{x}-\vec{\mu}_i) + \underbrace{\cancel{\ln(P(\omega_i))}}_{\text{This can be dropped.}}$$

· Consider a diagonal covariance matrix:

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

$$\Rightarrow g_i(\vec{x}) = \frac{-1}{2\sigma^2}(\vec{x}-\vec{\mu}_i)^T(\vec{x}-\vec{\mu}_i) = \frac{-1}{2\sigma^2}(d_E)^2,$$

where d_E: Euclidean Distance $= \|\vec{x}-\vec{\mu}_i\|_2$

⇒ A given $\vec{x}$ is assigned to the class that has the minimum Euclidean distance.

· Consider a nondiagonal covariance matrix:

Mahalanobis distance $(d_m) \triangleq \left[(\vec{x}-\vec{\mu}_i)^T\Sigma^{-1}(\vec{x}-\vec{\mu}_i)\right]^{1/2}$

Σ can be diagonalized by a unitary transform

$$\Sigma = \Phi\Lambda\Phi^T, \text{ where } \Phi^T = \Phi^{-1},$$

$$\Phi = \text{Eigen values of } \Sigma$$

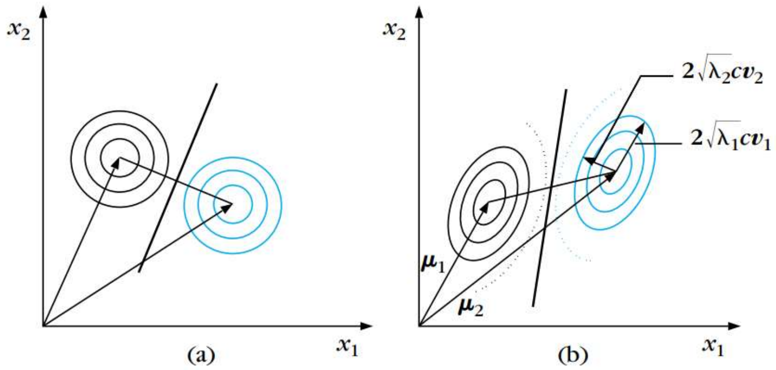


FIGURE 2.13

Curves of (a) equal Euclidean distance and (b) equal Mahalanobis distance from the mean points of each class. In the two-dimensional space, they are circles in the case of Euclidean distance and ellipses in the case of Mahalanobis distance. Observe that in the latter case the decision line is no longer orthogonal to the line segment joining the mean values. It turns according to the shape of the ellipses.

Example:

$M = 2, \; L = 2$

$$\Sigma_1 = \Sigma_2 = \begin{bmatrix} 1.1 & 0.3 \\ 0.3 & 1.9 \end{bmatrix}, \quad \mu_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \mu_2 = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$$

Let $\vec{x} = \begin{bmatrix} 1.0 \\ 2.2 \end{bmatrix}$. Classify $\vec{x}$ according to the Bayesian classifier

$$\Sigma^{-1} = \frac{1}{2}\begin{bmatrix} 1.9 & -0.3 \\ -0.3 & 1.1 \end{bmatrix} = \begin{bmatrix} 0.95 & -0.15 \\ -0.15 & 0.55 \end{bmatrix}$$

$$d_M^2 = (\vec{x} - \mu_i)^T \Sigma^{-1} (\vec{x} - \vec{\mu}_i)$$

d_M^2 for $i=1$ $= 2.952$

for $i=2$ $= 3.6672$

$\Rightarrow$ Assign $\vec{x}$ to ω_1.