Minimum Risk Bayes Decision Theoretic Classifier

Bayes Theorem: $p(y|x) = p(x|y).p(y)$ → posterior α likelihood $\times$ prior

$p(x|y)$ is multivariate Gaussian

$$p(x|y) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}}\exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right)$$

where μ and Σ are μ_{MLE} and Σ_{MLE}, respectively.

Training

Compute the values of μ_{MLE} and Σ_{MLE}.

$$\mu_{MLE_i}{}^T = \frac{1}{m}\sum_{j=1}^{m} x_{train\_i}{}^{(j)}$$

- μ_{MLE_i} is the mean vector of the i^{th} class training set.
- m is the number of samples in the i^{th} class training set.
- $x_{train\_i}{}^{(j)}$ is the j^{th} sample of the i^{th} class training set.

$$\Sigma_{MLE_i} = \frac{1}{m}\left(X_{train\_i} - M_{train\_i}\right)^T\left(X_{train\_i} - M_{train\_i}\right)$$

- X_{train_i} is the feature matrix of the i^{th} class training set.
- M_{train_i} is the mean matrix of the i^{th} class training set

Classification

Compute the discriminant function for each class to find the final class label.

$$g_i(x_{test}) = -\frac{1}{2}\left(x_{test} - \mu_{MLE_i}\right)^T \Sigma_{MLE_i}{}^{-1}\left(x_{test} - \mu_{MLE_i}\right) - \frac{d}{2}\ln(2\pi) - \frac{1}{2}\ln\left|\Sigma_{MLE_i}\right| + \ln\left(p(w_i)\right)$$

The term $\frac{d}{2}\ln(2\pi)$ can be dropped.

- $p(w_i)$ is the i^{th} class prior probability

- Class label = $argmax(g_1, g_2, \ldots)$

EXAMPLE

2D dataset with 2 classes and equal prior probabilities.

petal_length	petal_width	class
1.5	0.2	1
3.0	1.1	2
1.3	0.2	1
4.9	1.5	2
1.4	0.3	1
5.0	1.7	2
3.7	1.0	2
1.9	0.2	1

Training

Class 1 training data: $\begin{bmatrix} 1.5 & 0.2 \\ 1.3 & 0.2 \\ 1.4 & 0.3 \\ 1.9 & 0.2 \end{bmatrix}$

Class 1 mean vector: ${\mu_{MLE1}}^T = [1.525 \quad 0.225]$

Class 1 mean matrix: $\begin{bmatrix} 1.525 & 0.225 \\ 1.525 & 0.225 \\ 1.525 & 0.225 \\ 1.525 & 0.225 \end{bmatrix}$

Class 1 covariance:

$$\Sigma_{MLE1} = \frac{1}{4} \begin{bmatrix} 1.5 - 1.525 & 0.2 - 0.225 \\ 1.3 - 1.525 & 0.2 - 0.225 \\ 1.4 - 1.525 & 0.3 - 0.225 \\ 1.9 - 1.525 & 0.2 - 0.225 \end{bmatrix}^T \begin{bmatrix} 1.5 - 1.525 & 0.2 - 0.225 \\ 1.3 - 1.525 & 0.2 - 0.225 \\ 1.4 - 1.525 & 0.3 - 0.225 \\ 1.9 - 1.525 & 0.2 - 0.225 \end{bmatrix} = \begin{bmatrix} 0.0692 & -0.0042 \\ -0.0042 & 0.0025 \end{bmatrix}$$

Class 2 training data: $\begin{bmatrix} 3.0 & 1.1 \\ 4.9 & 1.5 \\ 5.0 & 1.7 \\ 3.7 & 1.0 \end{bmatrix}$

Class 2 mean vector: ${\mu_{MLE2}}^T = [4.15 \quad 1.325]$

Class 2 covariance: $\Sigma_{MLE2} = \begin{bmatrix} 0.9367 & 0.2850 \\ 0.2850 & 0.1091 \end{bmatrix}$

Classification

Class 1 test data: $x_{test}^T = [1.5 \quad 0.2]$

$$g_1 = -\frac{1}{2}\begin{bmatrix} 1.5 - 1.525 \\ 0.2 - 0.225 \end{bmatrix}^T \begin{bmatrix} 0.0692 & -0.0042 \\ -0.0042 & 0.0025 \end{bmatrix}^{-1} \begin{bmatrix} 1.5 - 1.525 \\ 0.2 - 0.225 \end{bmatrix} - \frac{1}{2}\ln\left(\begin{vmatrix} 0.0692 & -0.0042 \\ -0.0042 & 0.0025 \end{vmatrix}\right) + \ln\left(\frac{1}{2}\right) = 3.5304$$

$g_2 = -4.7728$

Predicted class label: $\hat{y} = argmax\ (g_1, g_2) = 1$

Note:

$A. \mathrm{A}^{-1} = \mathrm{I}$

$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$