Decision Boundaries For The Quadratic Discriminant Analysis

Saffet Gökçen Şen

May 18, 2020

Assume that a classification is to be made. Given the parameters $\boldsymbol{\theta}$ and the input vector $\mathbf{x}$, the class of the output variable y is to be determined. The probability density function of y being from the class c is modeled using a discrete generative model as follows:

$$p\left(y=c|\mathbf{x},\boldsymbol{\theta}\right)=\frac{p\left(\mathbf{x}|y=c,\boldsymbol{\theta}\right)p\left(y=c|\boldsymbol{\theta}\right)}{p\left(\mathbf{x}|\boldsymbol{\theta}\right)} \tag{1}$$

If a multivariate Gaussian distribution is used to model the class conditional density of the input vector, i.e. $p\left(\mathbf{x}|y=c,\boldsymbol{\theta}\right)$, then this is called the quadratic discriminant analysis:

$$p\left(\mathbf{x}|y=c,\boldsymbol{\theta}\right)=\frac{1}{\left(2\pi\right)^{D/2}\left|\boldsymbol{\Sigma}_{c}\right|^{1/2}}\exp\left[-\frac{1}{2}\left(\mathbf{x}-\boldsymbol{\mu}_{c}\right)^{T}\boldsymbol{\Sigma}_{c}^{-1}\left(\mathbf{x}-\boldsymbol{\mu}_{c}\right)\right] \tag{2}$$

The name "quadratic" comes from the fact that the decision boundaries separating classes are in general quadratic. Let the decision boundary between two classes be drawn to illustrate the quadratic boundary. First, the following equation is to be solved:

$$p\left(y=c_{1}|\mathbf{x},\boldsymbol{\theta}\right)=p\left(y=c_{2}|\mathbf{x},\boldsymbol{\theta}\right)\Rightarrow \tag{3}$$

$$\frac{p\left(\mathbf{x}|y=c_{1},\boldsymbol{\theta}\right)p\left(y=c_{1}|\boldsymbol{\theta}\right)}{p\left(\mathbf{x}|\boldsymbol{\theta}\right)}=\frac{p\left(\mathbf{x}|y=c_{2},\boldsymbol{\theta}\right)p\left(y=c_{2}|\boldsymbol{\theta}\right)}{p\left(\mathbf{x}|\boldsymbol{\theta}\right)}\Rightarrow \tag{4}$$

$$p\left(\mathbf{x}|y=c_{1},\boldsymbol{\theta}\right)p\left(y=c_{1}|\boldsymbol{\theta}\right)=p\left(\mathbf{x}|y=c_{2},\boldsymbol{\theta}\right)p\left(y=c_{2}|\boldsymbol{\theta}\right)\Rightarrow \tag{5}$$

$$\pi_{1}\frac{1}{\left(2\pi\right)^{D/2}\left|\boldsymbol{\Sigma}_{c_{1}}\right|^{1/2}}\exp\left[-\frac{1}{2}\left(\mathbf{x}-\boldsymbol{\mu}_{c_{1}}\right)^{T}\boldsymbol{\Sigma}_{c_{1}}^{-1}\left(\mathbf{x}-\boldsymbol{\mu}_{c_{1}}\right)\right]=$$

$$\pi_{2}\frac{1}{\left(2\pi\right)^{D/2}\left|\boldsymbol{\Sigma}_{c_{2}}\right|^{1/2}}\exp\left[-\frac{1}{2}\left(\mathbf{x}-\boldsymbol{\mu}_{c_{2}}\right)^{T}\boldsymbol{\Sigma}_{c_{2}}^{-1}\left(\mathbf{x}-\boldsymbol{\mu}_{c_{2}}\right)\right] \tag{6}$$

After taking the natural logarithm of both sides, the following simplified form is obtained:

$$0=$$

$$\frac{1}{2}\left[\left(\mathbf{x}-\boldsymbol{\mu}_{c_{1}}\right)^{T}\boldsymbol{\Sigma}_{c_{1}}^{-1}\left(\mathbf{x}-\boldsymbol{\mu}_{c_{1}}\right)-\left(\mathbf{x}-\boldsymbol{\mu}_{c_{2}}\right)^{T}\boldsymbol{\Sigma}_{c_{2}}^{-1}\left(\mathbf{x}-\boldsymbol{\mu}_{c_{2}}\right)\right]+\frac{1}{2}\log\left(\frac{\left|\boldsymbol{\Sigma}_{c_{1}}\right|}{\left|\boldsymbol{\Sigma}_{c_{2}}\right|}\right)-\log\left(\frac{\pi_{c_{1}}}{\pi_{c_{2}}}\right) \tag{7}$$

Since $\mathbf{x}$, $\boldsymbol{\mu}_{c_1}$ and $\boldsymbol{\mu}_{c_2}$ are column vectors, $\left(\mathbf{x} - \boldsymbol{\mu}_{c_1}\right)^T = \mathbf{x}^T - \boldsymbol{\mu}_{c_1}^T$ and $\left(\mathbf{x} - \boldsymbol{\mu}_{c_2}\right)^T = \mathbf{x}^T - \boldsymbol{\mu}_{c_2}^T$. After these equalities are used in the above equation, the following final form is obtained:

$$0 =$$

$$\frac{1}{2}\mathbf{x}^T \left(\boldsymbol{\Sigma}_{c_1}^{-1} - \boldsymbol{\Sigma}_{c_2}^{-1}\right) \mathbf{x} - \mathbf{x}^T \left(\boldsymbol{\Sigma}_{c_1}^{-1}\boldsymbol{\mu}_{c_1} - \boldsymbol{\Sigma}_{c_2}^{-1}\boldsymbol{\mu}_{c_2}\right) + \frac{1}{2}\left(\boldsymbol{\mu}_{c_1}^T\boldsymbol{\Sigma}_{c_1}^{-1}\boldsymbol{\mu}_{c_1} - \boldsymbol{\mu}_{c_2}^T\boldsymbol{\Sigma}_{c_2}^{-1}\boldsymbol{\mu}_{c_2}\right) +$$

$$\frac{1}{2}\log\left(\frac{|\boldsymbol{\Sigma}_{c_1}|}{|\boldsymbol{\Sigma}_{c_2}|}\right) - \log\left(\frac{\pi_{c_1}}{\pi_{c_2}}\right) \tag{8}$$

This is a quadratic form. Hence, the decision boundary between the two classes is quadratic. Now, let this boundary be drawn for $D = 2$, i.e. for the case when the random vectors are of dimension 2. If $A \triangleq \boldsymbol{\Sigma}_{c_1}^{-1} - \boldsymbol{\Sigma}_{c_2}^{-1}$, $B \triangleq \boldsymbol{\Sigma}_{c_1}^{-1}\boldsymbol{\mu}_{c_1} - \boldsymbol{\Sigma}_{c_2}^{-1}\boldsymbol{\mu}_{c_2}$ and $\mathbf{x} = [x_1 \ x_2]^T$, then:

$$0 =$$

$$\frac{1}{2}a_{22}x_2^2 + \left[\frac{1}{2}(a_{12} + a_{21})x_1 - b_{21}\right]x_2 +$$

$$\left[\frac{1}{2}a_{11}x_1^2 - b_{11}x_1 + \frac{1}{2}\left(\boldsymbol{\mu}_{c_1}^T\boldsymbol{\Sigma}_{c_1}^{-1}\boldsymbol{\mu}_{c_1} - \boldsymbol{\mu}_{c_2}^T\boldsymbol{\Sigma}_{c_2}^{-1}\boldsymbol{\mu}_{c_2}\right) + \frac{1}{2}\log\left(\frac{|\boldsymbol{\Sigma}_{c_1}|}{|\boldsymbol{\Sigma}_{c_2}|}\right) - \log\left(\frac{\pi_{c_1}}{\pi_{c_2}}\right)\right] \tag{9}$$

gives element x_2 given element x_1 of $\mathbf{x}$.

The covariance matrices of the distributions are to be formed. They must be positive definite. Let the distribution's eigenvectors be v_1 and v_2. Let the corresponding eigenvalues be λ_1 and λ_2 respectively. Then:

$$\boldsymbol{\Sigma}_c = \begin{bmatrix} v_1 & v_2 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} v_1 & v_2 \end{bmatrix}^{-1} \tag{10}$$

References

Machine Learning A Probabilistic Perspective, Kevin P. Murphy