

# Decision Boundaries For The Linear 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:

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

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

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

A special case of the quadratic discriminant analysis is when the covariance matrix of the Gaussian distribution for the class conditional density of the input features is common to all of the classes. That is the mean,

$$\boldsymbol{\Sigma}_c = \boldsymbol{\Sigma} \quad (3)$$

for all of the classes  $c$ . Then,  $p(y = c|\mathbf{x}, \boldsymbol{\theta})$  can be calculated as follows:

$$p(y = c|\mathbf{x}, \boldsymbol{\theta}) = \frac{p(\mathbf{x}|y = c, \boldsymbol{\theta}) p(y = c|\boldsymbol{\theta})}{p(\mathbf{x}|\boldsymbol{\theta})} = \quad (4)$$

$$\frac{p(\mathbf{x}|y = c, \boldsymbol{\theta}) p(y = c|\boldsymbol{\theta})}{\sum_{s=1}^C p(\mathbf{x}, y = s|\boldsymbol{\theta})} = \frac{p(\mathbf{x}|y = c, \boldsymbol{\theta}) p(y = c|\boldsymbol{\theta})}{\sum_{s=1}^C p(\mathbf{x}|y = s, \boldsymbol{\theta}) p(y = s|\boldsymbol{\theta})} \Rightarrow \quad (5)$$

$$p(y = c|\mathbf{x}, \boldsymbol{\theta}) = \frac{\pi_c p(\mathbf{x}|y = c, \boldsymbol{\theta})}{\sum_{s=1}^C \pi_s p(\mathbf{x}|y = s, \boldsymbol{\theta})} \Rightarrow \quad (6)$$

$$p(y = c|\mathbf{x}, \boldsymbol{\theta}) = \frac{\pi_c \frac{1}{(2\pi)^{D/2} |\boldsymbol{\Sigma}|^{1/2}} \exp \left[ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu}_c)^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}_c) \right]}{\sum_{s=1}^C \pi_s \frac{1}{(2\pi)^{D/2} |\boldsymbol{\Sigma}|^{1/2}} \exp \left[ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu}_s)^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}_s) \right]} \quad (7)$$

Let the expression

$$-\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu}_c)^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}_c) \quad (8)$$

be simplified:

$$-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_c)^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}_c) = -\frac{1}{2}(\mathbf{x}^T - \boldsymbol{\mu}_c^T) (\boldsymbol{\Sigma}^{-1} \mathbf{x} - \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_c) = \quad (9)$$

$$-\frac{1}{2}(\mathbf{x}^T \boldsymbol{\Sigma}^{-1} \mathbf{x} - \mathbf{x}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_c - \boldsymbol{\mu}_c^T \boldsymbol{\Sigma}^{-1} \mathbf{x} + \boldsymbol{\mu}_c^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_c) = \quad (10)$$

$$-\frac{1}{2}(\mathbf{x}^T \boldsymbol{\Sigma}^{-1} \mathbf{x} - 2\mathbf{x}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_c + \boldsymbol{\mu}_c^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_c) \Rightarrow \quad (11)$$

$$-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_c)^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}_c) = -\frac{1}{2}\mathbf{x}^T \boldsymbol{\Sigma}^{-1} \mathbf{x} + \mathbf{x}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_c - \frac{1}{2}\boldsymbol{\mu}_c^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_c \quad (12)$$

The simplified expression is put into the class probability density conditioned on the input and parameters:

$$p(y = c | \mathbf{x}, \boldsymbol{\theta}) = \frac{\pi_c \frac{1}{(2\pi)^{D/2} |\boldsymbol{\Sigma}|^{1/2}} \exp \left[ -\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_c)^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}_c) \right]}{\sum_{s=1}^C \pi_s \frac{1}{(2\pi)^{D/2} |\boldsymbol{\Sigma}|^{1/2}} \exp \left[ -\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_s)^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}_s) \right]} = \quad (13)$$

$$\frac{\pi_c \frac{1}{(2\pi)^{D/2} |\boldsymbol{\Sigma}|^{1/2}} \exp \left[ -\frac{1}{2}\mathbf{x}^T \boldsymbol{\Sigma}^{-1} \mathbf{x} + \mathbf{x}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_c - \frac{1}{2}\boldsymbol{\mu}_c^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_c \right]}{\sum_{s=1}^C \pi_s \frac{1}{(2\pi)^{D/2} |\boldsymbol{\Sigma}|^{1/2}} \exp \left[ -\frac{1}{2}\mathbf{x}^T \boldsymbol{\Sigma}^{-1} \mathbf{x} + \mathbf{x}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_s - \frac{1}{2}\boldsymbol{\mu}_s^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_s \right]} = \quad (14)$$

$$\frac{\frac{1}{(2\pi)^{D/2} |\boldsymbol{\Sigma}|^{1/2}} \exp \left( -\frac{1}{2}\mathbf{x}^T \boldsymbol{\Sigma}^{-1} \mathbf{x} \right) \pi_c \exp \left[ \mathbf{x}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_c - \frac{1}{2}\boldsymbol{\mu}_c^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_c \right]}{\frac{1}{(2\pi)^{D/2} |\boldsymbol{\Sigma}|^{1/2}} \exp \left( -\frac{1}{2}\mathbf{x}^T \boldsymbol{\Sigma}^{-1} \mathbf{x} \right) \sum_{s=1}^C \pi_s \exp \left[ \mathbf{x}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_s - \frac{1}{2}\boldsymbol{\mu}_s^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_s \right]} = \quad (15)$$

$$\frac{\pi_c \exp \left[ \mathbf{x}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_c - \frac{1}{2}\boldsymbol{\mu}_c^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_c \right]}{\sum_{s=1}^C \pi_s \exp \left[ \mathbf{x}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_s - \frac{1}{2}\boldsymbol{\mu}_s^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_s \right]} \Rightarrow \quad (16)$$

$$p(y = c | \mathbf{x}, \boldsymbol{\theta}) = \frac{\exp \left[ \mathbf{x}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_c + \log \pi_c - \frac{1}{2}\boldsymbol{\mu}_c^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_c \right]}{\sum_{s=1}^C \exp \left[ \mathbf{x}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_s + \log \pi_s - \frac{1}{2}\boldsymbol{\mu}_s^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_s \right]} \quad (17)$$

The boundary separating two classes  $c_1$  and  $c_2$  is a linear function of  $\mathbf{x}$ . Hence, this analysis is called the linear discriminant analysis. The equation for this linear boundary can be derived as follows:

$$p(y = c_1 | \mathbf{x}, \boldsymbol{\theta}) = p(y = c_2 | \mathbf{x}, \boldsymbol{\theta}) \Rightarrow \quad (18)$$

$$\frac{\exp [\mathbf{x}^T \Sigma^{-1} \boldsymbol{\mu}_{c_1} + \log \pi_{c_1} - \frac{1}{2} \boldsymbol{\mu}_{c_1}^T \Sigma^{-1} \boldsymbol{\mu}_{c_1}]}{\sum_{s=1}^C \exp [\mathbf{x}^T \Sigma^{-1} \boldsymbol{\mu}_s + \log \pi_s - \frac{1}{2} \boldsymbol{\mu}_s^T \Sigma^{-1} \boldsymbol{\mu}_s]} = \quad (19)$$

$$\frac{\exp [\mathbf{x}^T \Sigma^{-1} \boldsymbol{\mu}_{c_2} + \log \pi_{c_2} - \frac{1}{2} \boldsymbol{\mu}_{c_2}^T \Sigma^{-1} \boldsymbol{\mu}_{c_2}]}{\sum_{s=1}^C \exp [\mathbf{x}^T \Sigma^{-1} \boldsymbol{\mu}_s + \log \pi_s - \frac{1}{2} \boldsymbol{\mu}_s^T \Sigma^{-1} \boldsymbol{\mu}_s]} \Rightarrow \quad (20)$$

$$\mathbf{x}^T \Sigma^{-1} \boldsymbol{\mu}_{c_1} + \log \pi_{c_1} - \frac{1}{2} \boldsymbol{\mu}_{c_1}^T \Sigma^{-1} \boldsymbol{\mu}_{c_1} = \mathbf{x}^T \Sigma^{-1} \boldsymbol{\mu}_{c_2} + \log \pi_{c_2} - \frac{1}{2} \boldsymbol{\mu}_{c_2}^T \Sigma^{-1} \boldsymbol{\mu}_{c_2} \Rightarrow \quad (21)$$

$$\mathbf{x}^T \Sigma^{-1} \boldsymbol{\mu}_{c_1} - \mathbf{x}^T \Sigma^{-1} \boldsymbol{\mu}_{c_2} = \frac{1}{2} \boldsymbol{\mu}_{c_1}^T \Sigma^{-1} \boldsymbol{\mu}_{c_1} - \frac{1}{2} \boldsymbol{\mu}_{c_2}^T \Sigma^{-1} \boldsymbol{\mu}_{c_2} + \log \pi_{c_2} - \log \pi_{c_1} \Rightarrow \quad (22)$$

$$\mathbf{x}^T (\Sigma^{-1} \boldsymbol{\mu}_{c_1} - \Sigma^{-1} \boldsymbol{\mu}_{c_2}) = \frac{1}{2} \boldsymbol{\mu}_{c_1}^T \Sigma^{-1} \boldsymbol{\mu}_{c_1} - \frac{1}{2} \boldsymbol{\mu}_{c_2}^T \Sigma^{-1} \boldsymbol{\mu}_{c_2} + \log \pi_{c_2} - \log \pi_{c_1} \quad (23)$$

Assume that there are only two features which means that  $\mathbf{x}$  vector is two dimensional. Then, the equation of the boundary is as follows:

$$\beta_1 x_1 + \beta_2 x_2 = \text{constant} \quad (24)$$

where

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \Sigma^{-1} \boldsymbol{\mu}_{c_1} - \Sigma^{-1} \boldsymbol{\mu}_{c_2}, \quad (25)$$

$$\text{constant} = \frac{1}{2} \boldsymbol{\mu}_{c_1}^T \Sigma^{-1} \boldsymbol{\mu}_{c_1} - \frac{1}{2} \boldsymbol{\mu}_{c_2}^T \Sigma^{-1} \boldsymbol{\mu}_{c_2} + \log \pi_{c_2} - \log \pi_{c_1} \quad (26)$$

The covariance matrix of the distribution is to be formed. It must be positive definite. Let the distribution's eigenvectors be  $v_1$  and  $v_2$ . Let the corresponding eigenvalues be  $\lambda_1$  and  $\lambda_2$  respectively. Then:

$$\Sigma_c = \begin{bmatrix} v_1 & v_2 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} v_1 & v_2 \end{bmatrix}^{-1} \quad (27)$$

## References

Machine Learning A Probabilistic Perspective, Kevin P. Murphy