7.5.1. Geometry of Fisher's Rule

Fisher's method is interesting from a geometrical viewpoint. We can gain some intuitive understanding of the formulas from Figs. 7.4–7.7. In Fig. 7.4 a simple case is displayed, because the variances are equal. In this case $\Sigma^{-1} = 1/\sigma^2 I$, where I is the identity matrix and we classify according to the rule defined by Eq. (7.26). If this is expanded, it becomes

$$\frac{\mu_{11} - \mu_{21}}{\sigma^2} x_{01} + \frac{\mu_{12} - \mu_{22}}{\sigma^2} x_{02} > \frac{1}{2}\left(\frac{\mu_{11}^2 - \mu_{21}^2}{\sigma^2} + \frac{\mu_{12}^2 - \mu_{22}^2}{\sigma^2}\right) \qquad (7.26')$$

or

$$(\mu_{11} - \mu_{21})x_{01} + (\mu_{12} - \mu_{22})x_{02} > (\mu_{11} - \mu_{21})\overline{\mu}_1 + (m_{12} - \mu_{22})\overline{\mu}_2 \qquad (7.26'')$$

where $\overline{\mu} = \frac{1}{2}(\mu_1 + \mu_2)$. Note that $\overline{\mu}$ is the midpoint of the line segment joining the means of the two classes, while the slope of the line segment is $(\mu_{12} - \mu_{22})/(\mu_{11} - \mu_{21})$. We see that in the simple case of Fig. 7.4 the discriminant function is the equation of any line parallel to the line connecting the means. If $y_0 = l^{\mathrm{T}}\mathbf{x}_0$, then y_0 can be thought of as a projection onto that line. The quantity y_0 is sometimes called a "discriminant score." By using y_0 for classification we have simplified the classification problem, as we have reduced it to a univariate problem. The object can be classified by comparison of y_0 with $\mu_y = l^{\mathrm{T}}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)/2$. In this way Fisher's method can be viewed as a distance method.