When the variance–covariance matrices are equal but the covariance is not zero, the geometry is a little more complicated. Because the contours are not circular, but rather elliptical, a first step is to transform the ellipses to circles by an appropriate shrinking and rotation. The matrix that will produce the appropriate transformation is $\Sigma^{-1/2}$. This is somewhat similar to what is done in a two-sample Student's t-test. One wants to compare the difference between two means, but because the absolute difference may change with changes of scale, one divides by the appropriate standard deviation to remove scale effects. Once in the scaled space, the ellipses become circles and the problem is like the one in Fig. 7.4.

The slope of the line that separates the means in the scaled space is defined by $(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^T \Sigma^{-1/2}$. In the unscaled space, this corresponds to a line with slope defined by $(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^T \Sigma^{-1/2} \Sigma^{-1/2}$, or $(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^T \Sigma^{-1}$. One possible line is illustrated in Fig. 7.12(a). The contours in this figure have equal variance covariance matrices and nonzero correlation. The line that is drawn there is a line with slope defined by $(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^T \Sigma^{-1}$. Note that the scaling on this new line is arbitrary. From a classification point of view, it is only the location of a point relative to the others that is important, not the numerical value of that point. A common scaling to use is in terms of standard deviation units. Thus $\boldsymbol{l}$ may be adjusted so that $\text{Var}(y)$