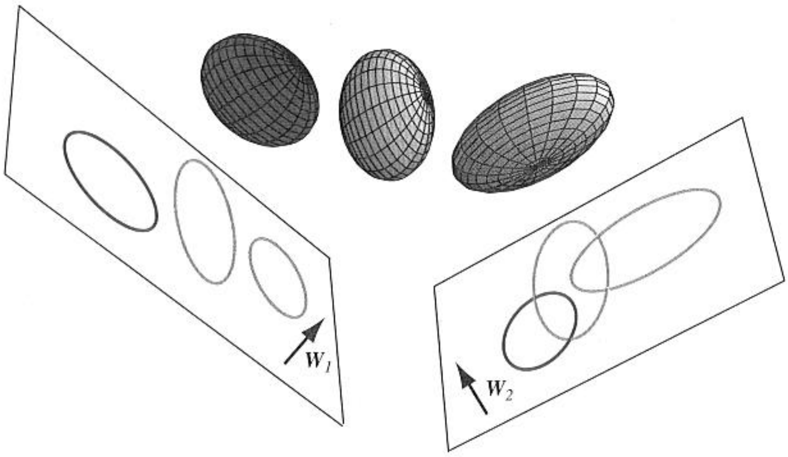

FIGURE 3.6. Three three-dimensional distributions are projected onto two-dimensional subspaces, described by a normal vectors $\mathbf{W}_1$ and $\mathbf{W}_2$. Informally, multiple discriminant methods seek the optimum such subspace, that is, the one with the greatest separation of the projected distributions for a given total within-scatter matrix, here as associated with $\mathbf{W}_1$.

These equations show how the within-class and between-class scatter matrices are transformed by the projection to the lower dimensional space (Fig. 3.6). What we seek is a transformation matrix $\mathbf{W}$ that in some sense maximizes the ratio of the between-class scatter to the within-class scatter. A simple scalar measure of scatter is the determinant of the scatter matrix. The determinant is the product of the eigenvalues, and hence is the product of the "variances" in the principal directions, thereby measuring the square of the hyperellipsoidal scattering volume (see also Eq. 46 in Chapter 2). Using this measure, we obtain the criterion function

$$J(\mathbf{W}) = \frac{|\tilde{\mathbf{S}}_B|}{|\tilde{\mathbf{S}}_W|} = \frac{|\mathbf{W}^t\mathbf{S}_B\mathbf{W}|}{|\mathbf{W}^t\mathbf{S}_W\mathbf{W}|}. \qquad (125)$$

The problem of finding a rectangular matrix $\mathbf{W}$ that maximizes $J(\cdot)$ is tricky, though fortunately it turns out that the solution is relatively simple. The columns of an optimal $\mathbf{W}$ are the generalized eigenvectors that correspond to the largest eigenvalues in

$$\mathbf{S}_B\mathbf{w}_i = \lambda_i\mathbf{S}_W\mathbf{w}_i. \qquad (126)$$

A few observations about this solution are in order. First, if $\mathbf{S}_W$ is nonsingular, this can be converted to a conventional eigenvalue problem as before. However, this is actually undesirable, since it requires an unnecessary computation of the inverse of $\mathbf{S}_W$. Instead, one can find the eigenvalues as the roots of the characteristic polynomial

$$|\mathbf{S}_B - \lambda_i\mathbf{S}_W| = 0 \qquad (127)$$

and then solve

$$(\mathbf{S}_B - \lambda_i\mathbf{S}_W)\mathbf{w}_i = 0 \qquad (128)$$