EXHIBIT 4
PART

(b) Show that the transfer of a sample $\hat{\mathbf{x}}$ from $\mathscr{X}_i$ to $\mathscr{X}_j$ causes J_T to change to

$$J_T^* = J_T - \left[\frac{n_j}{n_j+1}(\hat{\mathbf{x}} - \mathbf{m}_j)^t S_T^{-1}(\hat{\mathbf{x}} - \mathbf{m}_j) - \frac{n_i}{n_i - 1}(\hat{\mathbf{x}} - \mathbf{m}_i)^t S_T^{-1}(\hat{\mathbf{x}} - \mathbf{m}_i)\right].$$

(c) Suggest an iterative procedure for minimizing J_T.

17. Use the facts that $S_T = S_W + S_B$, $J_e = \text{tr } S_W$, and $\text{tr } S_B = \sum n_i \|\mathbf{m}_i - \mathbf{m}\|^2$ to derive the equations given in Section 6.9 for the change in J_e resulting from transferring a sample $\hat{\mathbf{x}}$ from cluster $\mathscr{X}_i$ to cluster $\mathscr{X}_j$.

18. Let cluster $\mathscr{X}_i$ contain n_i samples, and let d_{ij} be some measure of the distance between two clusters $\mathscr{X}_i$ and $\mathscr{X}_j$. In general, one might expect that if $\mathscr{X}_i$ and $\mathscr{X}_j$ are merged to form a new cluster $\mathscr{X}_k$, then the distance from $\mathscr{X}_k$ to some other cluster $\mathscr{X}_h$ is not simply related to d_{hi} and d_{hj}. However, consider the equation

$$d_{hk} = \alpha d_{hi} + \alpha_j d_{hj} + \beta d_{ij} + \gamma |d_{hi} - d_{hj}|.$$

Show that the following choices for the coefficients α_i, α_j, β, and γ lead to the distance functions indicated. (For other cases, see Lance and Williams, 1967.)

(a) $d_{\text{min}}: \alpha_i = \alpha_j = 0.5,\ \beta = 0,\ \gamma = -0.5.$
(b) $d_{\text{max}}: \alpha_i = \alpha_j = 0.5,\ \beta = 0,\ \gamma = 0.5.$
(c) $d_{\text{avg}}: \alpha_i = \dfrac{n_i}{n_i + n_j},\ \alpha_j = \dfrac{n_j}{n_i + n_j},\ \beta = \gamma = 0.$
(d) $d^2_{\text{mean}}: \alpha_i = \dfrac{n_i}{n_i + n_j},\ \alpha_j = \dfrac{n_j}{n_i + n_j},\ \beta = -\alpha_i\alpha_j,\ \gamma = 0.$

19. Consider a hierarchical clustering procedure in which clusters are merged so as to produce the smallest increase in the sum-of-squared error at each step. If the ith cluster contains n_i samples with sample mean $\mathbf{m}_i$, show that the smallest increase results from merging the pair of clusters for which

$$\frac{n_i n_j}{n_i + n_j}\|\mathbf{m}_i - \mathbf{m}_j\|^2$$

is minimum.

20. Consider the representation of the points $\mathbf{x}_1 = (1\ 0)^t$, $\mathbf{x}_2 = (0\ 0)^t$ and $\mathbf{x}_3 = (0\ 1)^t$ by a one-dimensional configuration. To obtain a unique solution, assume that the image points satisfy $0 = y_1 < y_2 < y_3$.

(a) Show that the criterion function J_{ee} is minimized by the configuration with $y_2 = (1 + \sqrt{2})/3$ and $y_3 = 2y_2$.
(b) Show that the criterion function J_{ff} is minimized by the configuration with $y_2 = (2 + \sqrt{2})/4$ and $y_3 = 2y_2$.