

DeClaris, N., Ijaola, O. O., Newcomb, R. W.  
 University of Maryland  
 College Park, Maryland 20742

An algorithmic scheme is presented for generating quadratic surfaces suitable for computer implementation and utilization for disease classification purposes. That is, given N subjects with  $N\alpha$ -real physiological parameters (real numbers) characterizing a known and confirmed disease, we present a computational scheme to cluster the known subjects inside an n-dimensional quadratic surface having minimal volume. The theoretical basis for this algorithmic approach can be found in the Proceedings of the 14th Allerton Conference on Circuit and System Theory, Univ. of Illinois, Sept. 1976 in our article "Novel Uses of Quadratic Surfaces for Medical Diagnosis".

**Step 1:** With the given data of N subjects with  $N\alpha$ -real parameters we form a data matrix called A(I, J) where  $1 \leq I \leq N$ , and  $1 \leq J \leq K(N\alpha)$ . We then look for the maximum positive parameters in every one of the columns of A(I, J). We then store all the K maximum positive parameters obtained above. We shall call these obtained maximum positive parameters the normalization (normalizing) factors. That is,  $NF(1), NF(2), \dots, NF(K)$ .

**Step 2:** We divide each member (or element) of every column of A[I, J] by the corresponding maximum positive parameters obtained. These are the normalization factors which we use to perform the scaling of the given data matrix A[I, J]. This scaling directly gives us a modified data matrix which is dimensionless.

**Step 3:** We determine the center of gravity of the system, and compute the shifted center of gravity which is the difference between the modified or normalized data matrix and the center of gravity of our multidimensional system:

$$\begin{aligned} [\phi_{N,1}] &= [C_{\alpha}^1(1) - \alpha_{N,1}^1; \phi_{N,2} = C_{\alpha}^2(2) - \\ &- \alpha_{N,2}^1; \dots; \phi_{N,K} = C_{\alpha}^k(k) - \alpha_{N,K}^1] \end{aligned}$$

**Step 4:** We look at each of the rows of the  $\phi$ -matrix and then determine the Norm or measure of each of the new vectors for the whole N subjects.

$$\rho_N = \|\phi_N\|^{1/2} = (\phi_{N,1}^2 + \phi_{N,2}^2 + \phi_{N,3}^2 + \dots + \phi_{N,K}^2)^{1/2}$$

**Step 5:** We now re-order the N subjects according to the ascending order of magnitudes of their norms as follows: Of all the N values of these rhos (norms or measures) obtained in Step 4, we sort our the longest K rhos where  $K \ll N$ . We then rearrange the K-largest norms obtained such that the biggest of them all is in ascending reorder of magnitude with the remaining (K-1) norms, forming a  $K \times K$  matrix

$$\{\phi^{\max(i)}\} \text{ where } 1 \leq i \leq K.$$

**Step 6:** We now determine, by using the Gramian if the elements of  $\phi^{\max}$  are linearly independent, in which case we proceed. Otherwise select another set. If none stop, classification not yet possible.

**Step 7:** We now create new coordinate vectors from the  $\phi^{\max}$  as follows:  $X_a^{\phi} = \|\phi^{\max(1)}\|^{-1} \cdot \phi^{\max(1)}$   
 Let  $H_b = [\phi^{\max(2)} \dots \phi^{\max(2)} \cdot X_a^{\phi} > X_a^{\phi}]$  and then define:  $X_b^{\phi} = \|H_b\|^{-1} H_b$ .

Let  $H_c = [\phi^{\max(3)} \dots \phi^{\max(3)} \cdot X_b^{\phi} > X_b^{\phi} \dots \phi^{\max(3)} \cdot X_a^{\phi} > X_a^{\phi}]$ .  
 Then define  $X_c^{\phi} = \|H_c\|^{-1} \cdot H_c$ .

We continue this process until we obtain the final  $X_t^{\phi}$  where  $t = K(N\alpha)$  the number of variable parameters for the N subjects. So  $H_t^{\phi} = [\phi^{\max(K)} \dots \phi^{\max(K)} \cdot X_s^{\phi} > X_s^{\phi} \dots \phi^{\max(K)} \cdot X_a^{\phi} > X_a^{\phi}]$

with  $X_t^{\phi} = \|H_t^{\phi}\|^{-1} \cdot H_t^{\phi}$ .

**Step 8:** We now find  $d_i$ 's through a transformation  $V^{\phi} \rightarrow V^X$ , which gives  $\langle X, \phi^P \rangle = d_1 \vec{X}_1 + d_2 \vec{X}_2 + d_3 \vec{X}_3 + \dots + d_K \vec{X}_K$  where  $\phi^P = d_i$ .

**Step 9:** For each of the N subjects we now form an ellipsoid by "shrinking" a sphere in K-dimensional space to satisfy the boundary condition:

$$\frac{d_1^2}{(\rho^{\max(1)})^2} + \dots + \frac{d_2^2}{(\rho^{\max(1)} - P_K \Delta)^2} \leq 1$$

where a convenient value for the shrinking factor is  $\Delta = \rho^{\max(1)} \times 0.01$  and the  $P_i$ 's, the 'shrinking' coefficients, are determined by a special computer program labeled SHRINK.

Thus, for a patient whose re-ordered number is 1, will have his (or her) ellipsoidal boundary condition satisfying the following equation:

$$\frac{d_{11}^2}{(\rho^{\max(1)})^2} + \dots + \frac{d_{1,K}^2}{(\rho^{\max(1)} - P_K \Delta)^2} = 1$$

where  $P_1 = 0$ , and  $P_1, P_2, \dots, P_K$  are called the 'shrinking' coefficients. The new coordinates  $X_i$  and the ellipsoid axis:  $\delta_i = \rho^{\max(1)} - P_i \Delta$  where  $i=1, 2, \dots, K$  suffice to classify and identify the disease. The work was supported in part by U.S. Public Grant No. RR0-7042-11.