MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Media Arts and Sciences Section

Course 4.902: Pattern Recognition and Analysis — Assigned: 3/2/93

Problem Set No. 5 — Due: 3/9/93

QUIZ 1 ANNOUNCEMENT:

> Quiz 1 will be held during class time on Tuesday, March 16.
> Material from the following will be covered: Class lectures through March 4; Therrien Chapters 1-6; Problem Sets 1-5.

READING ASSIGNMENT: Therrien, Sections 6.3-7.3

This problem set consists of four problems. Remember that you are *not* allowed to look at old problem sets, but collaboration is allowed and encouraged.

Feel free to use Matlab or any software package of your choice on any of these problems.

Problem 1: (25 points)

Using a software package such as Matlab, plot the one-sigma contours and likelihood ratio boundaries for the following pairs of Gaussian distributions (assume equal priors):

a) Means $\mathbf{m}_1 = [0 \;\; 0]^{\mathrm{T}}$ and $\mathbf{m}_2 = [4 \;\; 0]^{\mathrm{T}}$, with covariances

$$\mathbf{K}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \mathbf{K}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

b) Means $\mathbf{m}_1 = [0 \;\; 0]^{\mathrm{T}}$ and $\mathbf{m}_2 = [4 \;\; 0]^{\mathrm{T}}$, with covariances

$$\mathbf{K}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \qquad \mathbf{K}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

c) Means $\mathbf{m}_1 = [0 \;\; 0]^{\mathrm{T}}$ and $\mathbf{m}_2 = [2\sqrt{2} \;\; 2\sqrt{2}]^{\mathrm{T}}$, with covariances

$$\mathbf{K}_1 = \begin{bmatrix} 3/2 & -1/2 \\ -1/2 & 3/2 \end{bmatrix} \qquad \mathbf{K}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

d) Means $\mathbf{m}_1 = [0 \;\; 0]^{\mathrm{T}}$ and $\mathbf{m}_2 = [4 \;\; 0]^{\mathrm{T}}$, with covariances

$$\mathbf{K}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \qquad \mathbf{K}_2 = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$$

Problem 2: (25 points)

There are many types of decision boundaries that can occur with the quadratic classifiers we've discussed. Construct an example set of *diagonal* covariance matrices, $\mathbf{K}_1$, $\mathbf{K}_2$, and of mean vectors, $\mathbf{m}_1$, $\mathbf{m}_2$ for each of the following forms of decision boundary. Your examples should be different from those given in Problem 1.

(a) Linear	(b) Hyperbolic
(c) Parabolic	(d) Elliptical

For each form, show:

i) A simplified equation for your decision boundary in terms of the variables y_1 and y_2.

ii) A plot in the y_1-y_2 plane which shows the one-sigma covariance ellipses and decision boundary you chose.

Assume that the two classes have equal priors, so that the threshold $T = 0$.

Problem 3: (25 points)

The quadratic classifiers can be extended to more than two classes. Suppose we have three equally likely classes, with the following covariances and means.

$$\mathbf{K}_1 = \mathbf{K}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1/4 \end{bmatrix} \qquad \mathbf{K}_3 = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}$$

$$\mathbf{m}_1 = [0 \quad -2]^{\mathrm{T}} \qquad \mathbf{m}_2 = [0 \quad 0]^{\mathrm{T}} \qquad \mathbf{m}_2 = [0 \quad 2]^{\mathrm{T}}$$

a) Plot the one-sigma covariance ellipses in the y_1-y_2 plane.

b) Determine simple expressions for the quadratic discriminant functions, $g_1(\mathbf{y})$, $g_2(\mathbf{y})$, and $g_3(\mathbf{y})$ in terms of y_1 and y_2.

c) Solve for the decision boundaries $g_1(\mathbf{y}) = g_2(\mathbf{y})$, $g_1(\mathbf{y}) = g_3(\mathbf{y})$, and $g_2(\mathbf{y}) = g_3(\mathbf{y})$, and plot them on the graph you made in (a).

Problem 4: (25 points)

This problem is concerned with the nonlinear feature extraction method described in Section 5.7.2.

a) Consider the following set of training samples:

$$\mathbf{x}^{(1)} = \begin{bmatrix} 0.3 \\ 1.1 \\ 5.5 \end{bmatrix} \quad \mathbf{x}^{(2)} = \begin{bmatrix} 4.2 \\ 1.4 \\ 3.3 \end{bmatrix} \quad \mathbf{x}^{(3)} = \begin{bmatrix} 5.3 \\ 0.1 \\ 2.5 \end{bmatrix} \quad \mathbf{x}^{(4)} = \begin{bmatrix} 5.8 \\ 0.5 \\ 1.8 \end{bmatrix}$$

$$\mathbf{x}^{(5)} = \begin{bmatrix} 3.4 \\ 1.9 \\ 4.9 \end{bmatrix} \quad \mathbf{x}^{(6)} = \begin{bmatrix} 4.6 \\ 2.2 \\ 3.7 \end{bmatrix} \quad \mathbf{x}^{(7)} = \begin{bmatrix} 1.3 \\ 0.6 \\ 6.9 \end{bmatrix}$$

Samples $\mathbf{x}^{(1)}$, $\mathbf{x}^{(3)}$, $\mathbf{x}^{(4)}$, and $\mathbf{x}^{(7)}$ belong to class ω_1; the remaining three belong to class ω_2. Using Euclidean distance for d_x, and Eq. (5.62) with $\mu = 1$ for d_y, determine the between-class and within-class distances d_x and d_y for this training set. Present your results graphically as in Figure 5.5 (a). You should have 9 within-class distances and 12 between-class distances, for a total of 21 points.

b) Now suppose that d_y is to be approximated by a polynomial in d_x, as in Eq. (5.63). Do Problem 5.7 in the text: show that the given expression for the polynomial coefficients is optimum.

c) Suppose we want to use a third-order polynomial ($p = 3$) to approximate d_y in terms of d_x, so that for any particular pair of points,

$$d_y \approx a_0 + a_1 d_x + a_2 d_x^2 + a_3 d_x^3.$$

Using the distance values you obtained in part (a) along with Eq. (5.61), write down the matrices $\mathbf{D}$, $\mathbf{q}$, and $\mathbf{H}$ as defined in Problem 5.7. Use Matlab to evaluate the solution,

$$\mathbf{a} = (\mathbf{D}^{\mathrm{T}}\mathbf{H}\mathbf{D})^{-1}\mathbf{D}^{\mathrm{T}}\mathbf{H}\mathbf{q},$$

and plot the resulting polynomial on the graph you made in part (a).