

Name: .....

Student ID#: .....

**Statistical Pattern Recognition (CE-725)**  
**Department of Computer Engineering**  
**Quiz #3 (Feature Extraction) - Spring 2012**

Consider 2D data points in 2 classes depicted below:

$$C_1 = \{(0,-1)^T, (1,0)^T, (2,1)^T\}$$

$$C_2 = \{(1,1)^T, (-1,1)^T, (-1,-1)^T, (-2,-1)^T\}$$

- a. (30 points)** Compute principal components for all data points.
- b. (15 points)** Compute the projected data points on the first principal component.
- c. (15 points)** Compute the coordinates of projected data points on the first principal component, in the original feature space.
- d. (30 points)** Find the best vector based on LDA and project data points on this vector.
- e. (10 points)** Find the border of two classes in the LDA space.

**Sol:**

**a.**

$$X = C_1 \cup C_2, \quad \mu_x = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad S_x = XX^T = \begin{bmatrix} 12 & 5 \\ 5 & 6 \end{bmatrix}$$

*Eigen value decomposition :*

$$\begin{vmatrix} 12-\lambda & 5 \\ 5 & 6-\lambda \end{vmatrix} = 0 \quad \Rightarrow (12-\lambda)(6-\lambda) - 25 = 0 \quad \Rightarrow \lambda \approx \begin{cases} 15 \\ 3 \end{cases}$$

*Corresponding eigen vectore to  $\lambda_1 = 15$ :*

$$S - \lambda I = \begin{bmatrix} -3 & 5 \\ 5 & -9 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = 0 \quad \Rightarrow v_1 \approx \begin{bmatrix} 5 \\ 3 \end{bmatrix} \Rightarrow \bar{v}_1 = \frac{v_1}{\|v_1\|} = \frac{1}{\sqrt{34}} \begin{bmatrix} 5 \\ 3 \end{bmatrix} = \begin{bmatrix} .9 \\ .5 \end{bmatrix}$$

**b.**

$$\bar{v}_1^T x_i = \{-.5, .9, 2.3, 1.4, -.4, -1.4, -2.3\}$$

**c.**

$$\bar{v}_1^T x_i \bar{v}_1 = \{(-.45, -.25)^T, (.81, .45)^T, (2, 1.1)^T, (1.2, .7)^T, (-.4, -.2)^T, (-1.2, -.7)^T, (-2, -1.1)^T\}$$

**d.**

$$\begin{aligned} \mu_1 &= (1, 0)^T, \mu_2 = (-.75, 0)^T & S_1 &= \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}, S_2 = \begin{bmatrix} 4.75 & 3 \\ 3 & 4 \end{bmatrix} \\ S_w &= S_1 + S_2 = \begin{bmatrix} 6.75 & 5 \\ 5 & 6 \end{bmatrix} & \Rightarrow S_w^{-1} &= \frac{1}{15.5} \begin{bmatrix} 6 & -5 \\ -5 & 6.75 \end{bmatrix} \approx \begin{bmatrix} .4 & -.33 \\ -.33 & .43 \end{bmatrix} \\ \Rightarrow v &= S_w^{-1}(\mu_1 - \mu_2) = \begin{bmatrix} .4 & -.33 \\ -.33 & .43 \end{bmatrix} \begin{bmatrix} 1.75 \\ 0 \end{bmatrix} = \begin{bmatrix} .7 \\ -.6 \end{bmatrix} & \Rightarrow \bar{v} &= \frac{v}{\|v\|} = \begin{bmatrix} .75 \\ -.65 \end{bmatrix} \end{aligned}$$

**In The Name of God, The Compassionate, The Merciful**

Projected C1 data points =  $v^T C_1 = \{.63, .77, .9\}$

Projected C2 data points =  $v^T C_2 = \{.13, -1.4, -.13, -.9\}$

**e.**

Best point for border is:  $(.63 + .13)/2 = .38$ , or  $v^T \frac{\mu_1 + \mu_2}{2} = 0.1$