Name and Student ID:

Pattern Recognition and Analysis BBL514E, Nov 3, 2010, Midterm Exam.

1	2	3	4	5	Total

Name:
Number:
Signature:
Duration: 120 minutes.
Write your answers neatly in the space provided for them.
Write your name on each sheet.
Books and notes are closed. Good Luck!

QUESTIONS

QUESTION1) [20 points, 4 points each] What is (use at most three sentences per question, you can use drawings, formulas, etc. also):

a) reject region

b) confusion matrix of a classifier

d) Bayes rule

e) Naïve Bayes classification

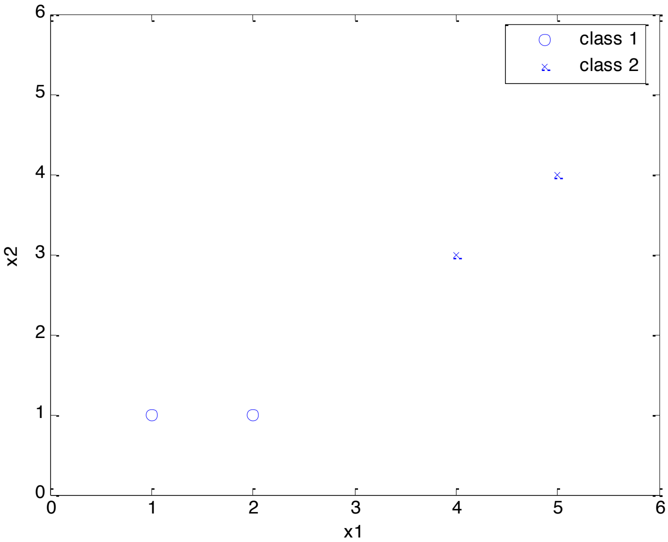
f) Cross validation

Name and Student ID:

QUESTION2) [30 points]

Consider the labeled data points given as follows:

$$X = \left\{ \left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}, 1 \right), \left(\begin{bmatrix} 2 \\ 1 \end{bmatrix}, 1 \right), \left(\begin{bmatrix} 4 \\ 3 \end{bmatrix}, 2 \right), \left(\begin{bmatrix} 5 \\ 4 \end{bmatrix}, 2 \right) \right\}$$



Assuming that inputs are normally distributed with class covariance matrices as follows:

$$S_1 = S_2 = \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}$$

4a) [20 pts] Compute the discriminant functions for both classes, $g_1(\underline{x})$ and $g_2(\underline{x})$.

4b) [5 pts] Compute and draw the discriminant function that separates the two classes.

Hint1: If $\underline{\mathbf{x}} \sim \mathrm{N}_d(\underline{\boldsymbol{\mu}}, \Sigma)$, then the pdf for x is given by:

$$p(\underline{\mathbf{x}}) = \frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}} \exp\left[-\frac{1}{2} (\underline{\mathbf{x}} - \underline{\boldsymbol{\mu}})^T \Sigma^{-1} (\underline{\mathbf{x}} - \underline{\boldsymbol{\mu}}) \right]$$

Hint2: Use the log likelihood for the discriminant function.

Hint3: $(1.5)^2 = 2.25$, $(2.5)^2 = 6.25$, $(3.5)^2 = 12.25$, $(4.5)^2 = 20.25$

Hint4: $\det\left(\begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}\right) = ad$

Name and Student ID:

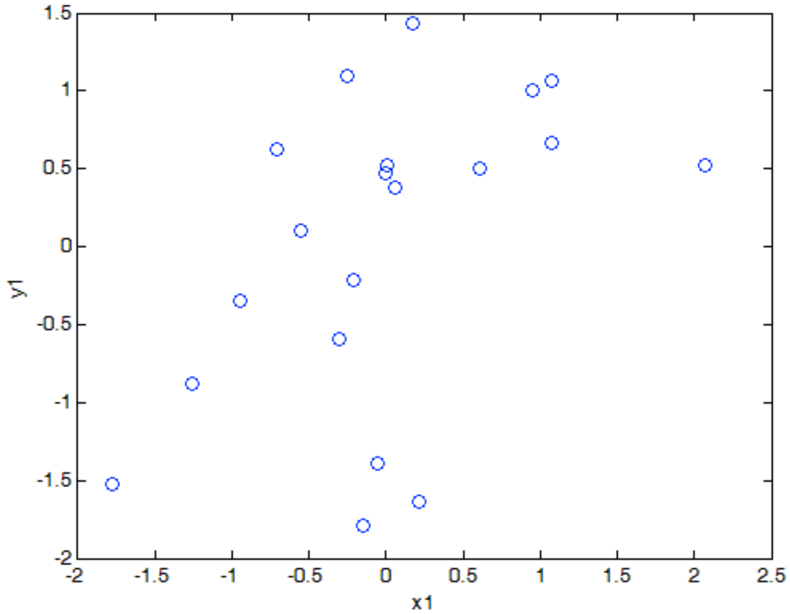
Name and Student ID:

QUESTION 3. [20 points]

Assume that you observe two different coins being tossed as follows:
Coin1={H,H,H,H,T,T,H,H,H,H,T,T,H,H,H,H,T}
Coin1={H,H,T,T,T,T,H,H,T,T ,T,T,H,H,T,T,T}
Assume that coin tosses are i.i.d. random variables. Each coin will be tossed one more time and for each correct guess you will be given 10000TLs. What would you guess for Coin1 and Coin2's next toss and why?

Name and Student ID:

QUESTION 4. [20 points]



Data points shown in the figure above have been normalized to have zero mean $\mu = [0\ 0]^T$
The covariance matrix $\sum$, eigen vectors D of the covariance matrix are:

$$\Sigma = \begin{bmatrix} 0.7697 & 0.4255 \\ 0.4255 & 0.9905 \end{bmatrix} \quad D = \begin{bmatrix} -0.7909 & 0.6119 \\ 0.6119 & 0.7909 \end{bmatrix},$$

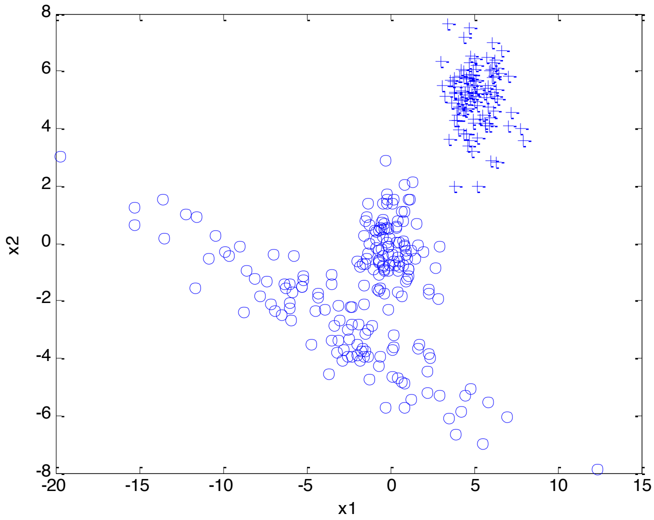
and the corresponding eigen values are $\lambda_1 = 0.4405$ and $\lambda_2 = 1.3197$.

Given this information, reduce the dimensionality of the following vector to one dimension using PCA:

$$x = [-1\ 1\ ]^T$$

Name and Student ID:

QUESTION 5. [10 points]
Given the following labeled data with two classes (+, o), discuss which dimensionality reduction technique you would prefer, why and how it would reduce the dimensionality of this data set.



Name and Student ID:

Extra sheet

Name and Student ID:

Extra sheet