MAS 622J: Pattern Recognition and Analysis

Problem Set 3
Date: Lecture 7   **Due:** Lecture 10,11

1. This is a MATLAB® problem.

   Download the datafile from the course webpage. The datafile consists of training and testing data for two different classes. The data given are 3-dimensional.

   - Reduce the dimensionality of the data to 1 (You should only use the training data for this),
     - Using Principal Component Analysis
     - Using Fisher Linear Discriminant

     Plot the histograms of the resulting data for each class and for each case.

   - Compute the maximum-likelihood mean and variance of the dimension-reduced data for both the classes and for each case.

   - Assuming that the classes are equally likely($P(\text{class0}) = P(\text{class1}) = \frac{1}{2}$) and that the likelihoods of observation given a class are Gaussian, write a MATLAB® script to classify the testing data using the estimated means and variances using 0-1 cost criterion. Report the recognition results for both the cases (PCA and Fisher Linear Discriminant).

2. Let x be a Gaussian distribution

$$p(x|\theta) \sim \mathcal{N}(\theta, \sigma^2)$$

   (a) What is the noninformative prior for the mean parameter θ? Is it proper?

   (b) What is the relation between the noninformative prior and a Gaussian prior $p(\theta) \sim \mathcal{N}(\mu, \alpha^2)$

   (c) Based on the noninformative prior for θ, compute the posterior of the mean $p(\theta|x)$. Is the posterior a proper distribution?

   (Hint: you might use the result derived in the problem 2 of homework set 1.)

3. From Duda, Hart and Stork: Chapter 3 - Problem 4, 14, and 38