DIMENSIONALITY REDUCTION USING PRINCIPAL COMPONENT ANALYSIS

Linear Algebra and Random Processes (CS6015)
Assignment 4

1 Problem Statement

Given N (number of samples in each class), D-dimensional data points for 2 classes, reduce the dimensionality of the data using PCA and classify the reduced features using Bayes classification (done in assignment 3).

2 Input

- N = 7k, 70k (for each class)
- D = 10, 50

Input contains data from three different distributions for each of the two combinations of N, D given above.
Look into the table attached, to see the allotment of input data.

3 Output

- The top 2 principal components for the given data i.e, reduce the given data to 2-dimensional data.
- Scatter plot of the reduced 2-dimensional data (in different colors indicating that they are data points from two different classes).
- Classification accuracy using training and validation set.

4 References

- Richard O. Duda, Peter E. Hart, David G. Stork. Pattern Classification. Second Edition. Section 3.8.1.

- Christopher M. Bishop. Pattern Recognition and Machine Learning. Section 12.1