Question 1

Average classification accuracy over at least 10 runs =

Mean Accuracy = 73.020

Mean Stdev = 01.406

To select appropriate principal components or for **principal component analysis** we do Singular Value Decomposition on a matrix and then generates an eigen value matrix. To select the principal components, we take the first few eigen values and try to retain as much information as possible and this can be deduced by the fact using variance. So, we select those pairs which retains maximum variance.

Question 2

Using FLD we want to do 2 things-

- A large variance among the dataset classes.
- A small variance within each of the dataset classes.

We find m1, m2 of each distribution, within class scatter is the cumulative sum of covariance of each distribution. And between class scatter is the difference between the mean of the individual distribution.

Optimal line direction is one which increases the between class variance and decreases the in class variance.

Question 3

Average classification accuracy over at least 10 runs =

Mean Accuracy 77.656

Mean Stdev 1.700

Question 4

Mean Accuracy 100