<u>Machine Learning Assignment-1</u>

Documentation

1. Fisher's Linear Discriminant

Fisher's Linear Discriminant Analysis has applications in both Classification as well as Dimensionality reduction.

This algorithm seeks to obtain a vector **w** such that **w** maximises the separation between the points belonging to the two target classes (for example: positive and negative classes) after their projection onto the vector **w**.

In order to determine the vector **w**, we try to amalgamate the following optimisation problems simultaneously:

a) Maximisation of the distance between the means of the projected points of each class.
b) Minimise the sum of the variance of the projected points of each class

Implementation

For each class in the dataset, the d-dimensional mean vectors were calculated, and the difference between the two was evaluated. The deviation of the dataset from the mean was found out for each class and summed up to obtain the within-class scatter matrix.

Theory asserts that the vector **w** is proportional to the difference of means multiplied by the inverse of the within-class scatter matrix, both of which were obtained earlier. **w** was hence computed.

From the d-dimensions, the dataset was projected onto one-dimension by taking the projection of each data point onto vector **w**. A normal distribution was fitted to each class and the curve was plotted. The intersection of these curves was evaluated (by solving a quadratic equation). This intersection point serves as the threshold point.

Results

w = [-1.41766460e-04 1.49463712e-04 -9.99999979e-01]
Discriminating Line :
 0.3952724123143726 from projection of origin on **w**(denoted by the green in the following Figure)
Separating plane :

$$-0.000141766x + 0.0001494y + -0.99999z = 0.3952$$

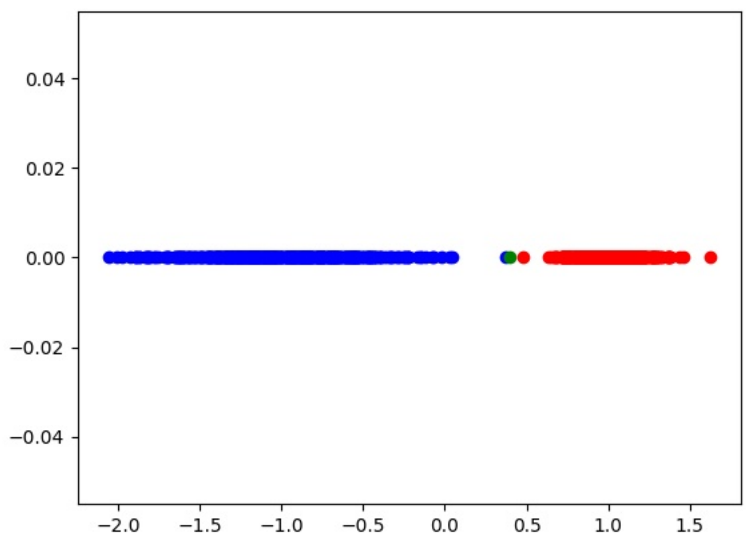

Figure 1

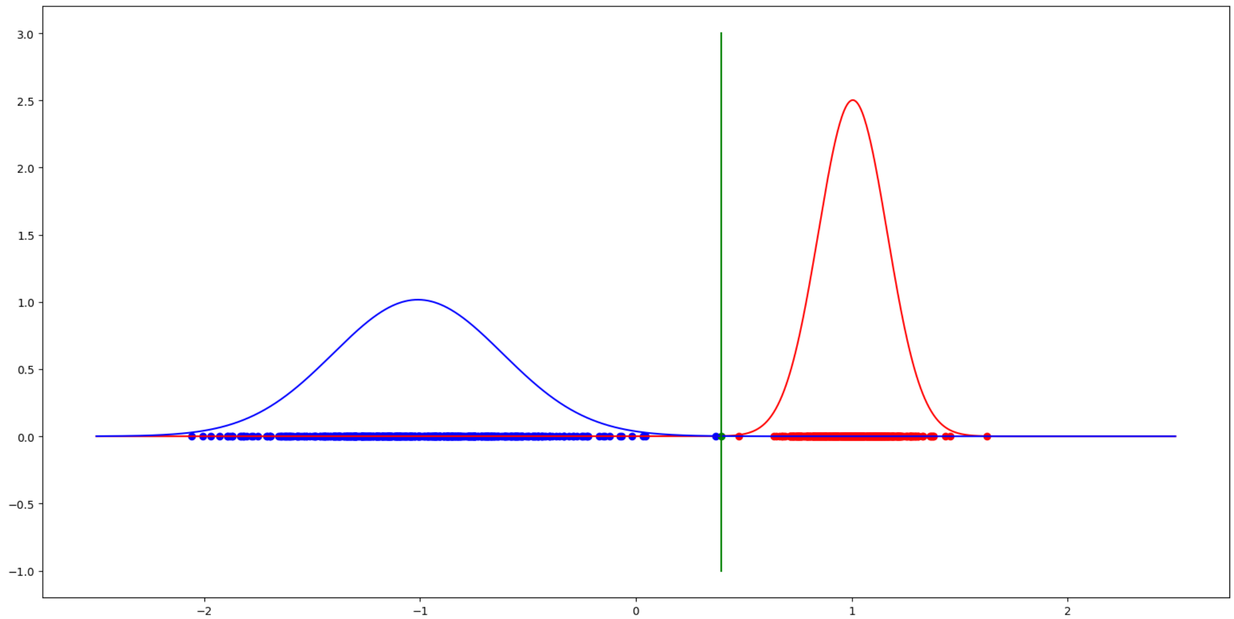

Figure 2

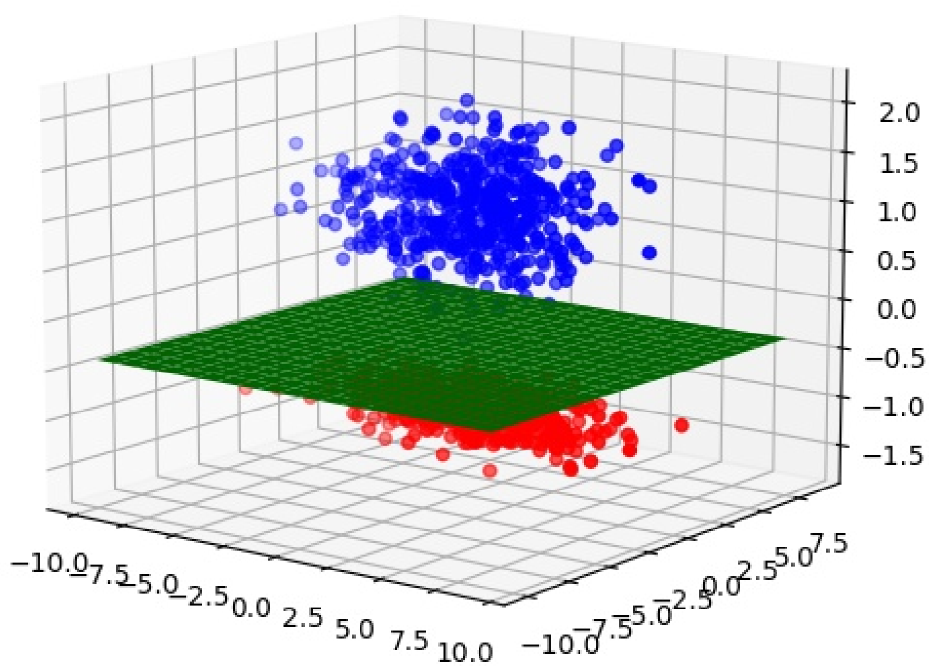

Figure 3

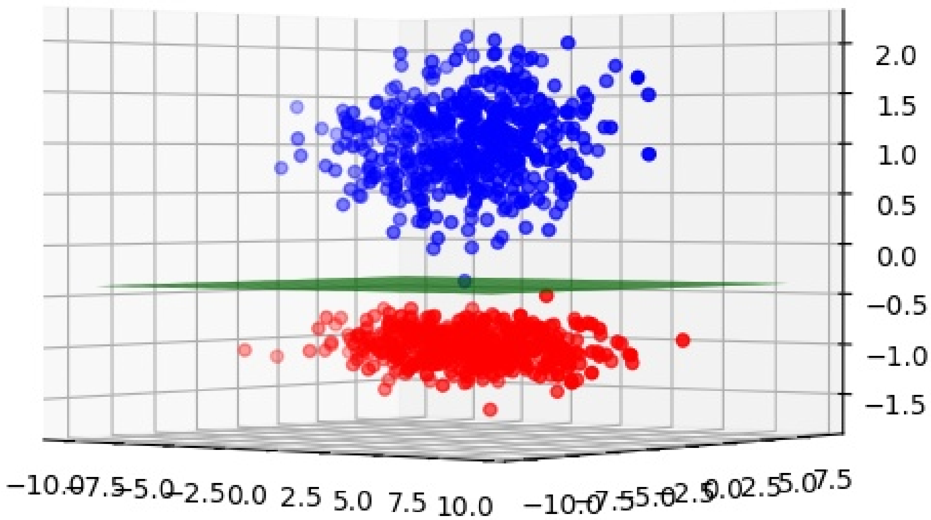

Figure 4

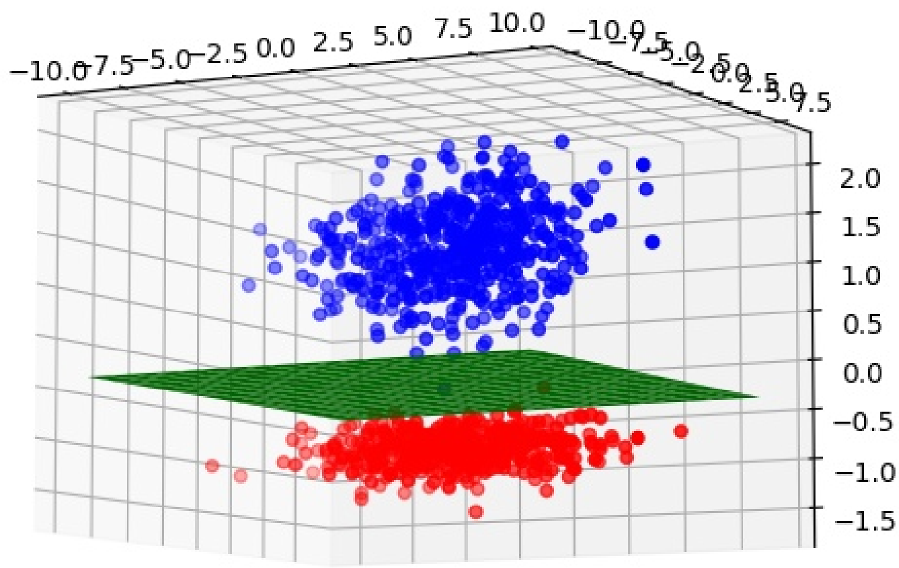

Figure 5