Fisher Linear Discriminant Analysis

Rahul Mehndiratta 2017A7PS1479H
BVS Ruthvik 2019A7PS0017H
R Vedang. 2019A7PS0150H

Introduction:

Fischer's Linear Discriminant Analysis is a method used in statistics, pattern recognition, and machine learning to find a linear combination of features that characterizes or separates two or more classes of objects or events. The resulting combination may be used as a linear classifier, or, more commonly, for dimensionality reduction before later classification. Fischer's LDA tries to maximize the difference between the means of the two classes and minimize the within-class covariances in their transformed direction. With this objective, we get the best possible direction to which the data points need to be transformed. After this, a point is decided in this direction which separates the two classes in the best way. This point can be the intersection point of the Gaussian curves for the two classes. When a new point comes, it is then tested against this point to classify it into one of the two classes. To deal with classification problems with 2 or more classes, most Machine Learning (ML) algorithms work the same way. Usually, they apply some kind of transformation to the input data with the effect of reducing the original input dimensions to a new (smaller) one. The goal is to project the data to a new space. Then, once projected, they try to classify the data points by finding a linear separation.

Problem Definition and Implementation:

We've been asked to implement Fischer's Linear Discriminant. This algorithm is basically a dimensional reduction algorithm that enables the user to separate points in an N-dimensional space by finding the projection of these points on a line in (N-1)-dimensions and then finding an appropriate threshold value which is a part of the required hyperplane that separates the N-Dimensional plots according to the given constraints. We've been given one dataset - FLD_dataset.csv. We have opted to use the entire dataset to find the appropriate plane of separation.

Unit Vector Of the Discriminant plane(Normal to discriminant Plane):

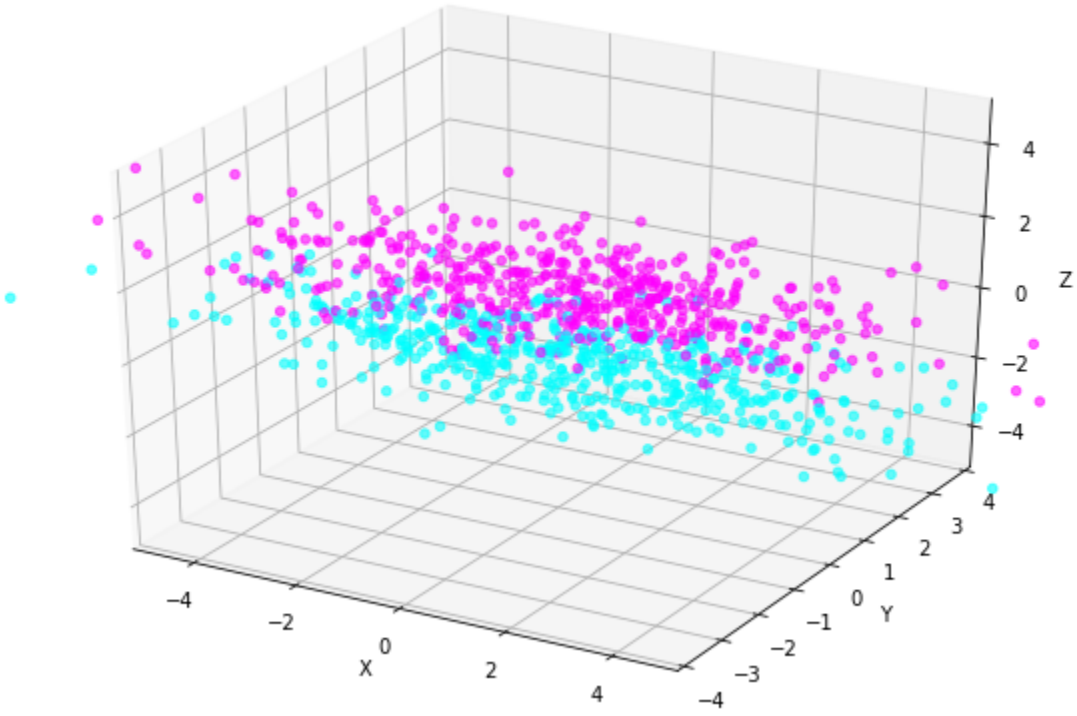
W Unit Vector= [-0.00655686 -0.01823739 0.99981218]

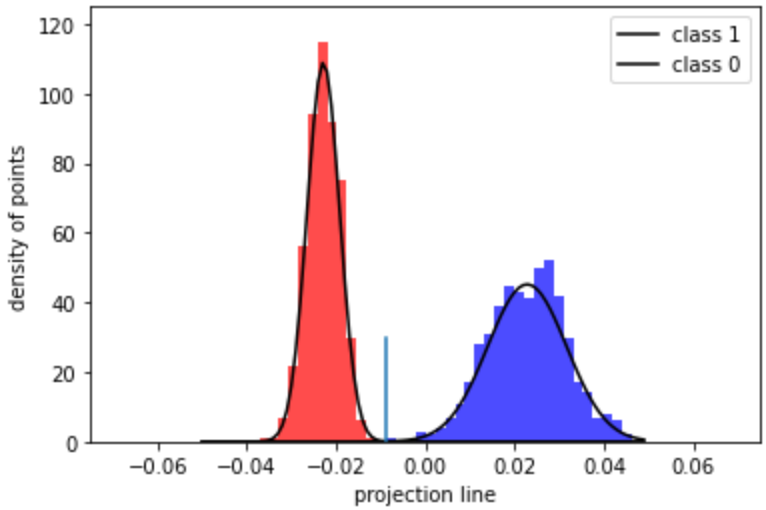
Discriminant Line Threshold:

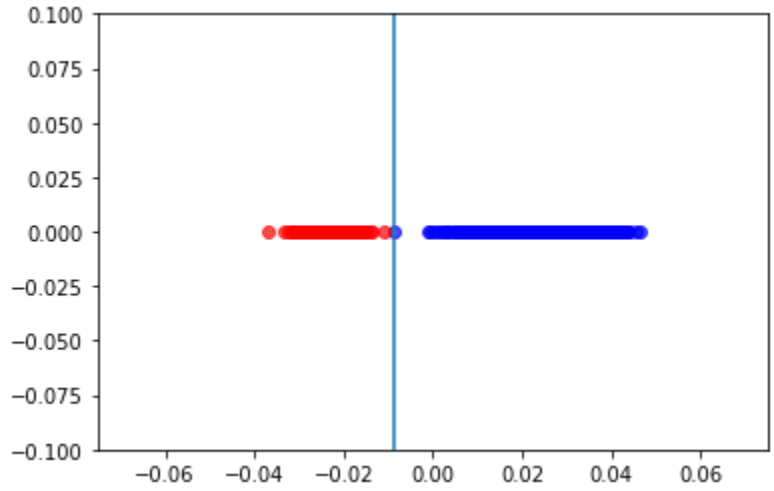
W=-0.008813863129467877

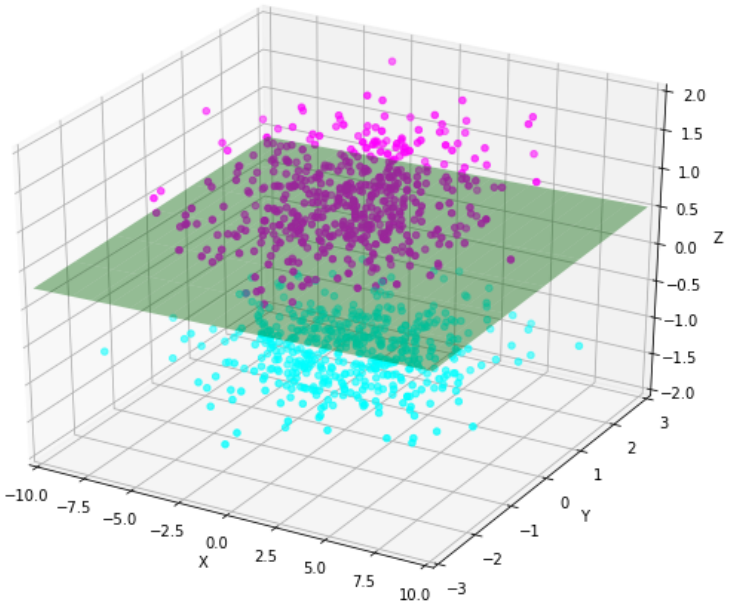
Plots:

1) Plotting the points in 3D




class 1
class 0
density of points
projection line
120
100
80
60
40
20
0
−0.06
−0.04
−0.02
0.00
0.02
0.04
0.06


0.100
0.075
0.050
0.025
0.000
−0.025
−0.050
−0.075
−0.100
−0.06
−0.04
−0.02
0.00
0.02
0.04
0.06


2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
Z
−10.0
−7.5
−5.0
−2.5
0.0
2.5
5.0
7.5
10.0
X
−3
−2
−1
0
1
2
3
Y