**Q1.** (40 points) PCA and LDA In dataset dataset 1.csv, columns correspond to variables and there are two variables named V1 and V2.

(1) Plot V2 vs V1. Do you see a clear separation of the raw data?

Ans. No, we do not see a clear separation of the raw data.

(2) Apply your own PCA function to this dataset without scaling the two variables. Project the raw data onto your first principal component axis, i.e. the PC1 axis. Do you still see a clear separation of the data in PC1, i.e. in projections of your raw data on the PC1 axis?

Ans.

(3) Add the PC1 axis to the plot you obtained in (1).

Ans. Added

(4) Apply your own LDA function to this dataset and obtain W. The class information of each data point is in the label column.

Ans. W = [[-0.71381103 1. ]

[ 0.70033836 0. ]]

(5) Project your raw data onto W. Do you see a clear separation of the data in the projection onto W?

Ans. Yes, we see a clear separation of the data.

(6) Add the W axis to your plot. At this point, your plot should contain the raw data points, the PC1 axis you obtain from the PCA analysis, and the W axis you obtain from the LDA analysis.

Ans. Plot contains 3 axes.

(7) Compute the variance of the projections onto PC1 and PC2 axes. What is the relationship between these two variances and the eigenvalues of the covariance matrix you use for computing PC1 and PC2 axes?

Ans. Variance explained by eigenvalue 1: 40.63%

Variance explained by eigenvalue 2 : 44.40%

(8) Compute the variance of the projections onto the W axis.

Ans. Variance explained by eigenvalue 1: 100.00%

Variance explained by eigenvalue 2: 0.00%

(9) What message can you get from the above PCA and LDA analyses?

Ans : LDA looks for PC2 so that we can preserve most of the discriminating information or separating information between the classes. LDA considers distance between classes as well as their overlapping.