Assignment on LDA

1. Implement LDA (40 points)

Implement your own class to carry out two-class LDA. Use the iris dataset to test and compare your results with LDA in sklearn.

2. PCA and LDA (60 points)

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?

(2) Apply your own PCA class to this dataset without scaling the two variables. Project the raw data onto your first principal component axis, i.e. the PC1 axis.

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

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

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

(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.

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

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