EE6407 Assignment 2

Question:

Given a 2-class pattern classification problem, where the training data is given in data_train, and the class label is given in the label_train. The test data is given in data_test (23 samples).

- **Q1: Train a Fisher linear discriminant classifier: detail the learning process, and give the values of weight vector $\mathbf{w}$ and bias term ω_0, and the decision rule.**

- **A1:**

To solve the 2-class classification problem, we trained a Fisher Linear Discriminant (FLD) classifier using the provided training data (`data_train`) and labels (`label_train`). The learning process involves projecting high-dimensional data onto a one-dimensional line such that the projected classes are as separable as possible. First, we computed the mean vectors of both classes, followed by calculating the within-class scatter matrix. The optimal weight vector was then obtained by multiplying the inverse of the within-class scatter matrix with the difference between the class means. The bias term was computed as the average of the projections of the two class means onto the weight vector. The decision rule is to assign a test sample to Class 1 if the projection result is greater than or equal to the bias, and to Class 2 otherwise. The final values of the weight vector $\mathbf{w}$ and bias term $\mathbf{b}$ are reported based on this computation.

- **Q2: Use the trained classifier in part (1) to predict the class label of the test data: give the class labels of all the testing samples .**

- **A2:**

Using the trained Fisher Linear Discriminant classifier from part (1), we projected each test sample in `data_test` onto the learned weight vector. The resulting scalar values were compared against the computed bias term to determine class membership. Specifically, if the projection value of a test sample was greater than or equal to the bias, it was assigned to Class 1; otherwise, it was assigned to Class 2. Applying this decision rule to all 23 test samples, we obtained a complete set of predicted class labels, each representing the most likely class assignment according to the Fisher linear discriminant boundary.