**Modified Ex3**

from google.colab import drive

drive.mount('/content/drive')

import numpy as np

import matplotlib.pyplot as plt

import pandas as pd

from numpy import linalg as lg

%matplotlib inline

# Load dataset

df = pd.read\_csv('/content/drive/MyDrive/iris1.csv', header=None)

df = df.drop(df.index[0])  # Remove header if present in data

df = df.drop(df.columns[0], axis=1)  # Drop first column (e.g., index column)

X = df.values.astype(float)

num\_features = X.shape[1]

# Split data

X1 = X[0:20, :]

X2 = X[20:, :]

# Compute means

m1 = X1.mean(axis=0)

m2 = X2.mean(axis=0)

mean\_all = X.mean(axis=0)

# Compute Within-Class Scatter Matrix

S1 = (X1 - m1).T @ (X1 - m1)

S2 = (X2 - m2).T @ (X2 - m2)

S\_W = S1 + S2

# Compute Between-Class Scatter Matrix

S\_B = 20 \* np.outer((m1 - mean\_all), (m1 - mean\_all)) + \

      20 \* np.outer((m2 - mean\_all), (m2 - mean\_all))

e\_val, e\_vec = np.linalg.eig(np.linalg.inv(S\_W) @ S\_B)

# FIX GOES HERE

eig\_pairs = [(np.abs(e\_val[i]), e\_vec[:, i]) for i in range(len(e\_val))]

eig\_pairs = sorted(eig\_pairs, key=lambda x: x[0], reverse=True)

# Project to 1D

W = eig\_pairs[0][1].reshape(num\_features, 1)

lda\_project = X @ W

# Plot manual LDA

plt.figure()

plt.title("Manual LDA")

plt.plot(lda\_project[0:20], np.zeros(20), 'bo', label='NSCLC')

plt.plot(lda\_project[20:], np.zeros(20), 'ro', label='SCLC')

plt.legend()

plt.show()

# ===========================

# SKLEARN LDA Implementation

# ===========================

from sklearn.discriminant\_analysis import LinearDiscriminantAnalysis as LDA

# Labels: 1 for NSCLC, 2 for SCLC

y = np.array([1]\*20 + [2]\*20)

# Fit and transform

sklearn\_lda = LDA(n\_components=1)

X\_lda\_sklearn = sklearn\_lda.fit\_transform(X, y)

X\_lda\_sklearn = -X\_lda\_sklearn  # Optional: flip direction for consistency

# Plot sklearn LDA

plt.figure()

plt.title("Sklearn LDA")

plt.plot(X\_lda\_sklearn[0:20], np.zeros(20), 'bo', label='NSCLC')

plt.plot(X\_lda\_sklearn[20:], np.zeros(20), 'ro', label='SCLC')

plt.legend()

plt.show()

Dataset: Iris1.csv