1) (30/20) Mahalanobis distance vs. Euclidean distance.

   a. (5/5) Write the equations to calculate these two distances. (Note: ONLY the equation.)

   b. (5/5) Explain intuitively (in no more than three sentences) the differences between the two distances.

   c. (20/10) Use the following example to understand the differences these two distances make in classification. Here, the minimum distance classifier (i.e., Case I) is used.

      i. Plot the above data set on the same figure.

      ii. Given a test sample \(x = [0.85 \ 1.15]^T\), calculate the Euclidean distance to the two classes. Based on the distances, which class should \(x\) belong to?

      iii. Use the same test sample, calculate the Mahalanobis distance to the two classes. Based on this pair of distances, which class should \(x\) belong to?
iv. Plot the test sample \( x \) on the same figure as the data set. Just by observing the plot, which decision do you think makes more sense?

2) (70/70) Dimensionality reduction with a two-feature two-class data set. (Use the same set as shown in 1)

a. (10/10) Preprocessing steps (Need to show step-by-step details):
   i. Calculate the mean of each class \((m_1, m_2)\) manually.
   ii. Calculate the covariance matrix of each class \((\Sigma_1, \Sigma_2)\) manually.

b. (10/10) Using Fisher’s Linear Discriminant (FLD) to find the projection vector \((w)\) which optimally (in the Fisher sense) separates the projections of these two classes.

c. (10/10) Is the vector derived from FLD along the same direction as \((m_1-m_2)\)? Plot \(m_1-m_2\) and \(w\) on the same figure.

d. (10/10) Using principal component analysis (PCA) to reduce the dimension to 1 and plot the principal component on the same figure. (You can use existing functions to find the eigenvectors and eigenvalues.)

ever. (10/10) Comment on the differences between FLD and PCA and \((m_1-m_2)\). Make up a scenario where FLD will be aligned, perpendicular to \((m_1-m_2)\), if possible at all.

f. (10/10) Project the test sample \( x \) onto \( w \) derived from FLD and determine its label

g. (10/10) Project the test sample \( x \) onto the principal axis from PCA and determine its label

3) (+10/10) Using maximum likelihood method to derive the equation for mean and variance assuming the pdf (or likelihood) is modeled by 1-D Gaussian.