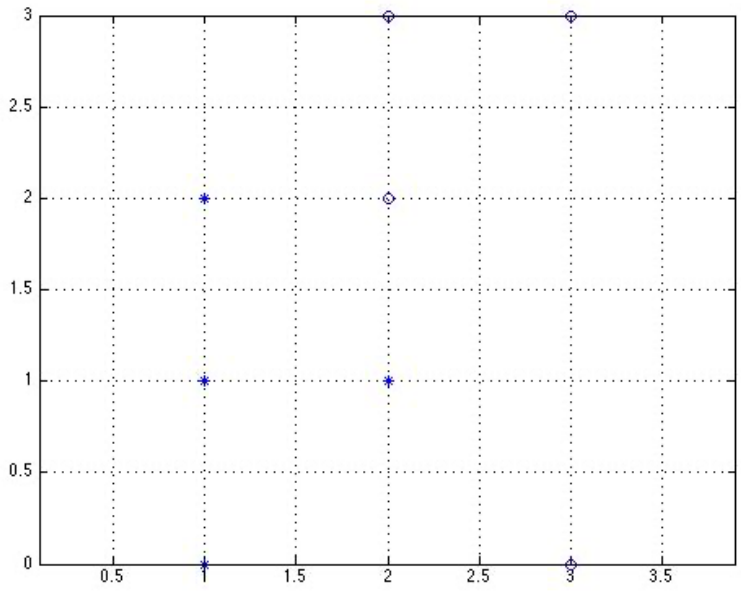
Fall 2021

COSC 522 – Homework #3 (Due 10/07)

Problem 1 (50 pts): A 2-class training set is given in the following table. Each class contains 4 samples. Each sample indicates a coordinate in a 2-d space. All the following questions are based on this data set. Use pencil and paper and only resort to calculator when performing matrix calculations.

Class 1 Samples	Class 2 Samples
(1, 1)	(2, 3)
(1, 2)	(2, 2)
(2, 1)	(3, 3)
(1, 0)	(3, 0)



a. If we need to reduce the dimension of the dataset from 2 to 1, find the projection direction using FLD. Show detailed steps, including class means, within-class scatter matrices, between class scatter matrix, etc.

b. Sketch the isotropic contour of the two scatter matrices on the same figure as shown in the previous page.

c. Sketch the projection vector on the same figure.

d. Project the samples on *w* and provide the projected result. Mark the projected samples on the vector in the same figure.

e. Based on the **projected** samples, classify a test sample $x=[2\ 1.4]^T$ using two different classifiers (show details)

    i) the minimum distance classifier, i.e., case 1 of discriminant function

    ii) the kNN classifier with k=1.

Problem 2 (50 pts). Redo Problem 1 but using PCA for dimensionality reduction.

a. If we need to reduce the dimension of the dataset from 2 to 1, find the projection direction using PCA. Show detailed steps, including how to calculate the covariance matrix of the entire dataset, perform eigenvalue decomposition on the covariance matrix.

b. If only 1 dimension is kept, calculate how much error you would introduce.

c. Sketch the isotropic contour of the covariance matrix on the same figure as shown in the previous page. Please use a different color.

d. Sketch the projection vector ($\mathbf{e}_1$) on the same figure.

e. Project the samples on $\mathbf{e}_1$ and provide the projected result. Mark the projected samples on the vector in the same figure.

f. Based on the **projected** samples, classify a test sample $x=[2\ 1.4]^T$ using two different classifiers (show details)

    i) the minimum distance classifier, i.e., case 1 of discriminant function

    ii) the kNN classifier with k=1.