Pattern Classification Homework #7

1. Given A=(0,1) B=(0,2) C=(1,1) D=(1,2)

   E=(-1,0) F=(-1,1) G=(-2,0) H=(-2,1)

   Group #1 A, B, C, D vs. E, F, G, H

   Group #2 A, B, C, D, F vs. E. G. H

   Group #3 B, C, D vs. A, E, F, G, H

   (a) Compute between cluster and within cluster scatter matrices for each grouping

   (b) Compute sum-of-square-error criterion function J_e for each grouping

   (c) Which grouping gives the minimum variance (lowest J_e) clustering?

2. K-means Clustering

Given the following data x_1=(1,0), x_2=(1,1), x_3=(0,1), x_4=(0,0), x_5=(3,1), x_6=(4,1), x_7=(4,0), x_8=(3,0), x_1 through x_9, try k-means clustering with k=2 and k=3 and Euclidian distance. Discuss about the results. Use the silhouette plot for clustering evaluation.