Homework 4

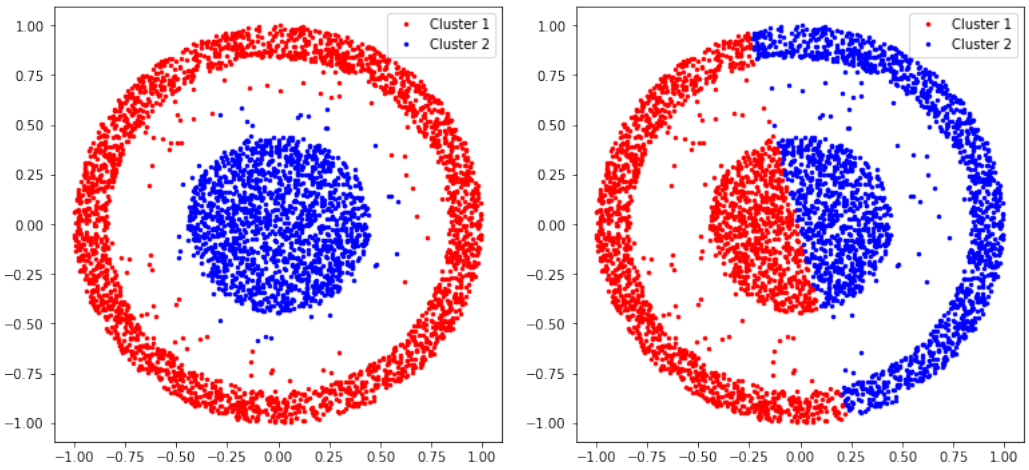
Collaborators:
Name: Wang Yanwei
Student ID: 11821049

Problem 4-1. Spectral Clustering

In this problem, we will try a dimensionality reduction based clustering algorithm – Spectral Clustering.

(a) We will first experiment Spectral Clustering on synthesis data

Answer: Abviously, Spectral Cluster can classiify the data correctly, but Kmeans not.



(a) spectral cluster (b) kmeans

Figure 1: spectral cluster VS kmeans

(b) Now let us try Spectral Clustering on real-world data.

Answer:

Spectral Clustering: Accuracy=0.5782410917361638, mutual information=0.6684473568298703

kmeans: Accuracy=0.5134950720242608, mutual information=0.36631839357492624

Problem 4-2. Principal Component Analysis Let us deepen our understanding of PCA by the following problems.

(a) Your task is to implement *hack_pca.m* to recover the rotated CAPTCHA image using PCA.

Answer:

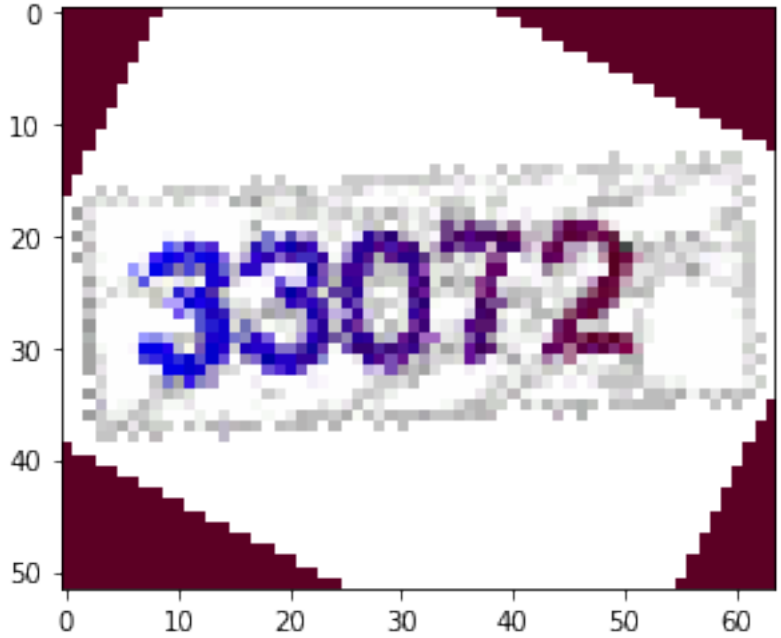


Figure 2: rotated image

(b) Now let us apply PCA to a face image dataset.

Answer:

Figure 3: eigen face

When low dimensional number is 8, accuracy is 0.745
When low dimensional number is 16, accuracy is 0.825
When low dimensional number is 32, accuracy is 0.845
When low dimensional number is 64, accuracy is 0.865
When low dimensional number is 128, accuracy is 0.865

Figure 4: 128 dim