

## Homework 4

### Collaborators:

Name: Shen Zijin

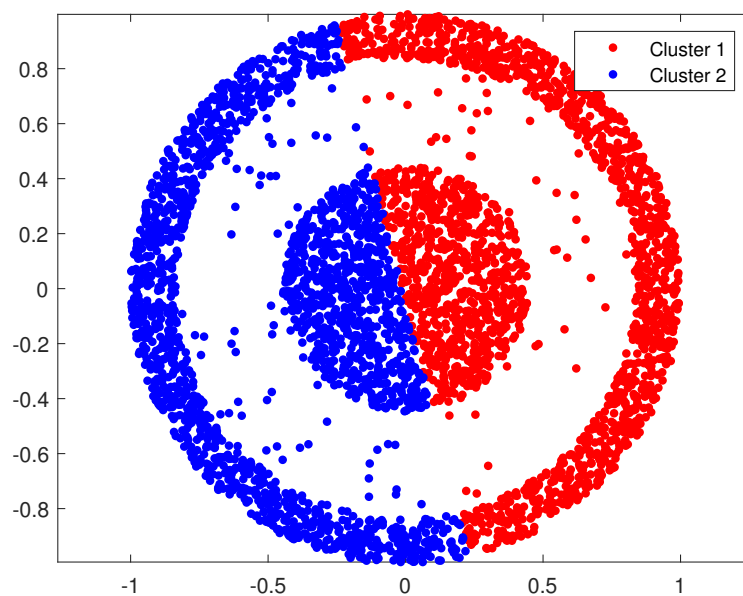
Student ID: 3160104734

### Problem 4-1. Spectral Clustering

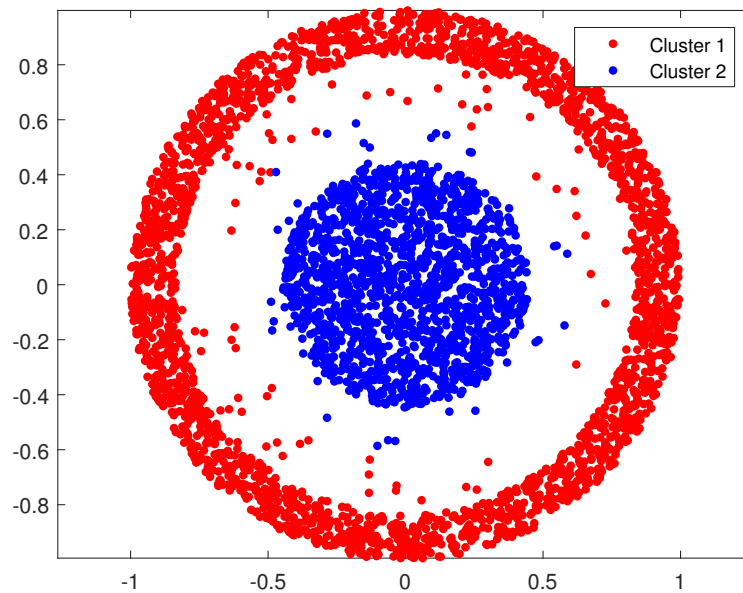
In this problem, we will try a dimensionality reduction based clustering algorithm – Spectral Clustering.

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

**Answer:**



**Figure 1:** Raw kmeans.



**Figure 2:** Spectral Clustering.

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

**Answer:** After examining the efficiency of both, I decided to solve the generalized eigenvalues directly without processing the Laplacian matrix. We set WeightMode to 'Binary', and  $k$  to 8, then we find that  
 if we choose raw Kmeans, accuracy is 0.52133, NMI is 0.39358;  
 if we choose Spectral Clustering, accuracy is 0.78104, NMI is 0.43346.

**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:**



33072

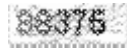
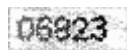
31372

31372

16468

16468

88375

A small, low-resolution image of a license plate with the number 88375. The plate is white with black text and is slightly tilted.A small, low-resolution image of a license plate with the number 06823. The plate is white with black text and is slightly tilted.A small, low-resolution image of a license plate with the number 06823. The plate is white with black text and is slightly tilted.

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

**Answer:**

i.



**Figure 3:** Eigenfaces

ii.

1. Testing Error rate with dimension 8 is 24.5%
2. Testing Error rate with dimension 16 is 20.0%
3. Testing Error rate with dimension 32 is 18.0%
4. Testing Error rate with dimension 64 is 15.0%
5. Testing Error rate with dimension 128 is 15.0%

iii.

Answer: Yes. To some extent, it prevents the loss of information. But the loss is not important because it has no significant performance impact.



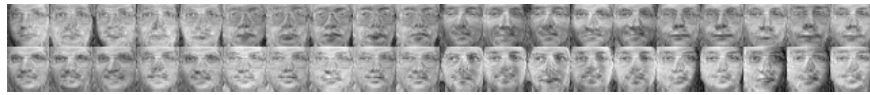


32





64



128

