

## Homework 4

### Collaborators:

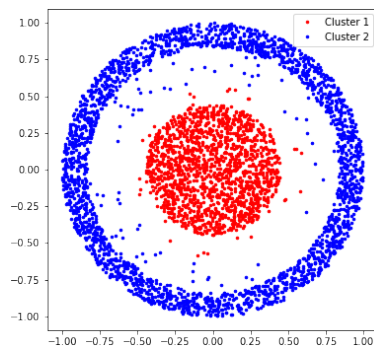
Name: Junlin Yin

Student ID: 3160104340

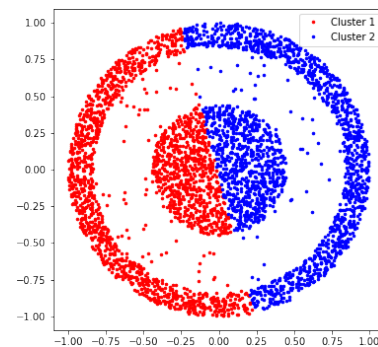
### Problem 4-1. Spectral Clustering

(a) First, experiment Spectral Clustering on synthesis data...

**Answer:** When  $K_{\text{in\_knn\_graph}}$  is 100 and threshold is 0.5, the Spectral Clustering system works well.



(a) Spectral Clustering



(b) KMeans

**Figure 1:** two results compare

(b) Try Spectral Clustering on real-world data...

**Answer:** When I use 5 nearest neighbor to construct adjacency matrix, I get the following results:

**Table 1:** results

	Spectral	KMeans
Accuracy	78.8%	59.4%
MutualInfo	0.626	0.357

### Problem 4-2. Principle Component Analysis

(a) Recover the rotated CAPTCHA image using PCA...

**Answer:** Main steps in my code:

1. turn the images into grayscale ones;
2. get the coordinates of these relatively white points (controlled by a threshold);
3. apply PCA to these white points and get the first Principle Component;
4. calculate the restoration angle using the first Principle Component, and construct the Rotating matrix;
5. rotate these white points and create the new image after restoration.

The results show that the PCA does work, but there is still some noise in the output image. It is created by the threshold which need to be set manually. What's more, some digits are hard to distinguish after the rotation, such as 5 and 6. In order to show the results in a clearer way, I use PIL.Image library. By doing this, I can directly rotate the original RGB image so the color information is preserved. Plus, the noise created by the threshold will also vanish.

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

**Answer:**

1. see the code part.
2. here are the results: (1 nearest neighbor)

**Table 2:** test error using different d

d	Test Error
8	26%
16	18.5%
32	14.5%
64	12%
128	12.5%
256	12%
512	12%
1024	12%

3. see the code part. We can know from the images and test errors that 64 dimensions are good enough to preserve the features of original inputs.