In this homework, I implemented a spectral clustering algorithm in Python.

First imported libraries which will be needed in project. multivariate\_normal is to create multivariate normal Gaussians from specific mean vector and covariance matrix. numpy.linalg.eig is to create eigen values and eigen vectors.

Text

Description automatically generated

Figure 1: import libraries

**Part 1**

I read data from hw08\_data\_set which contains 300 data points generated randomly from five bivariate Gaussian densities.

**Part 2**

First I defined Euc\_Distance(point1, point2) function to calculate Euclidean distances.

I defined threshold delta parameter to the 1.25 as said in the pdf. Then, I created B matrix which is connectivity matrix by below formula.

Text

Description automatically generated with low confidence

Figure 2: Connectivity matrix formula

After defining B matrix, visualized connectivity matrix as follows:

Diagram

Description automatically generated

Figure 3: Connectivity matrix(B) visualization

**Part 3**

In this part, I calculated 𝐃 and 𝐋 matrices as described in the lecture notes.

Text

Description automatically generated

Text

Description automatically generated

Figure 4: D,L matrices formulas given in the lectures

. Calendar

Description automatically generated

Figure 5: D,L matrices outputs

**Part 4**

In this part I found eigenvectors and eigenvalues of normalized Laplacian matrix using linalg.eig function. Then I took 5 eigenvectors corresponding to the smallest eigenvalues to put them to the Z matrix .

Text

Description automatically generated with medium confidence

Figure 6: Z matrix representation given in the lectures

Table

Description automatically generated

Figure 7: Z matrix outputs

**Part 5**

In this part I run k-means clustering algorithm on 𝐙 matrix to find 𝐾 = 5 clusters.

First, I assigned 29, 143, 204, 271, and 277 rows of Z matrix as initial centroid. Then used similar codes with lab11. However, in this time, we work on the Z matrix instead of X matrix as whole. After iterations we lastly updated centroids according to the found memberships.

**Part 6**

Last I plotted the clustering results obtained by k-means.

Chart, scatter chart

Description automatically generated

Figure 8: Visualization of clustering obtained by k-means on Z matrix