

Engr 421 – HW #6 Report

Firstly, I defined three functions for updating the centroids, memberships, and plotting the current state. These functions are taken from the lab.

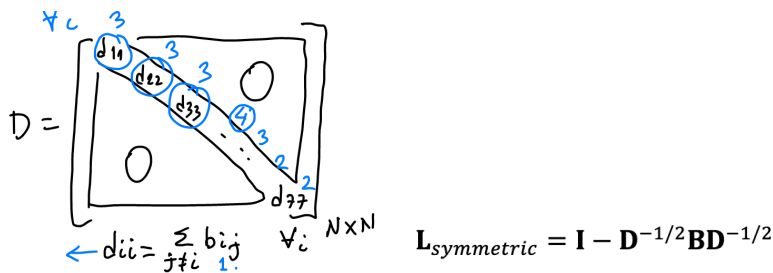
Euclidean Distance

$$d(x_i, x_j) = \|x_i - x_j\|_2 = \sqrt{x_i^T x_i - 2x_i^T x_j + x_j^T x_j}$$

$$= \sqrt{\sum_{d=1}^D (x_{i,d} - x_{j,d})^2}$$

I constructed the B matrix by using the Euclidean distance formula above. Then, I plot the connectivity matrix to show connected data points.

I calculated D and L matrices using the following formulas.



$D =$

$\leftarrow d_{ii} = \sum_{j=1}^N b_{ij}$

$L_{\text{symmetric}} = I - D^{-1/2} B D^{-1/2}$

I applied following steps where $R = 5$ smallest eigenvectors and $K = 5$. Using these eigenvectors, I constructed Z matrix. I initialized the centroids as the 85, 129, 167, 187, and 270th rows of Z matrix.

- Algorithm
- STEP ① Find the eigenvectors of normalized L matrix.
- STEP ② Pick R largest eigenvectors.
- STEP ③ Construct Z matrix as follows:
- $$Z = [v_1 \ v_2 \ \dots \ v_R]_{N \times R}$$
- STEP ④ Run k-means clustering algorithm on Z matrix to find K clusters.
- Parameters
- δ : threshold
- R : # of eigenvectors to be included
- K : # of clusters.

Lastly, I run k-means clustering algorithm on Z matrix and plot the clustering result. Formulas above are taken from the lecture notes.