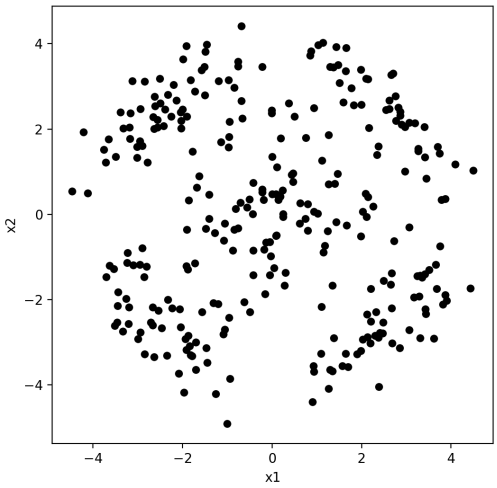
Ecre Çınar
69652

ENGR 421
Homework 8: **Spectral Clustering**

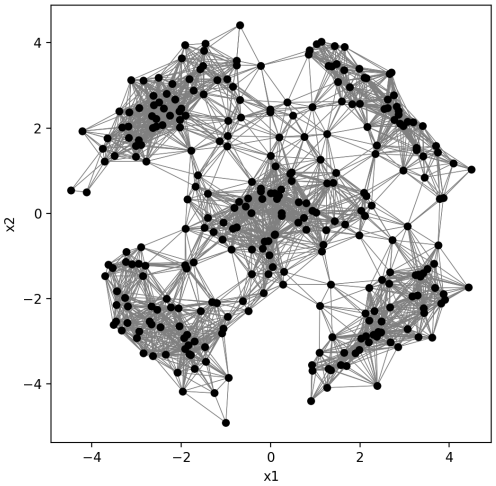


In this homework I first imported the necessary libraries which are numpy, matplotlib and scipy.spatial.

I then plotted the datapoints given in the csv file, using the means, deviations and class sizes they are generated from. The plotted data points are shown on the left.

After plotting the data I calculated the B matrix by calculating the Euclidian distances between points, using $\delta = 1.25$.

After calculating the B matrix, I used it to show the connectivities on a graph. For every entry == 1 for a given i, j pair I plotted a line between the points. The resulting connectivity graph looked like this:



After that I calculated the diagonal degree matrix, D, by summing up the diagonals of the B matrix.
To calculate the L symmetrical, I used the given formula:

$$\mathbf{L}_{symmetric} = \mathbf{I} - \mathbf{D}^{-1/2}\mathbf{B}\mathbf{D}^{-1/2}$$

I used numpy's ***identity*** function for I, and performed a matmul operation with D and B matrices.
After that I used the ***np.linalg.eig*** function to calculate the eigenvalues and corresponding eigenvectors of L symmetric. I then sorted these and took the first 5 smallest eigenvectors and constructed the Z matrix.

I initialized centroids with 29, 143, 204, 271, and 277th rows of the Z matrix, and then I used the same code from Lab11 to run K-means clustering algorithm. When the algorithm converged I plotted the datapoints showed in different colors for each class. The classes on the graph looked on the right.