

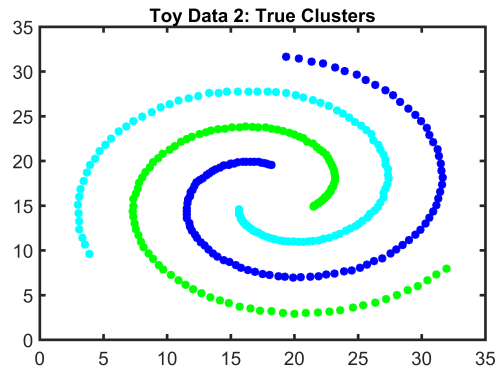
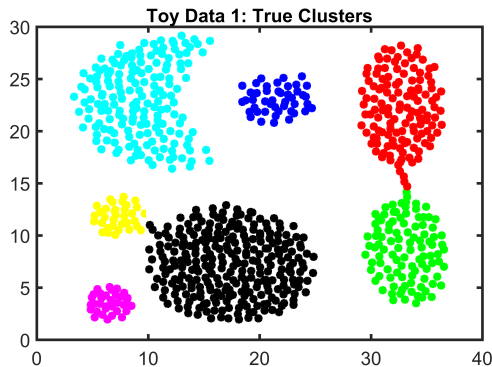
CmpE 482 - Homework 3

(Due: Dec. 18th, 2017 midnight - No extensions)

Part 1: Spectral Clustering & Toy Data Sets (90 pts)

In this exercise, we will implement two different approaches for spectral clustering, test the methods on toy data sets, and compare with k-means clustering.

- Method 1: Implement unnormalized spectral clustering as described on page 399 of the tutorial by von Luxburg [1].
- Method 2: Implement normalized spectral clustering as described on the second page of the paper by Ng et al. [2].
- You are given two toy data sets, i.e., `toy1.mat` and `toy2.mat`. Each toy data set is in the form of a samples by features matrix. Show that you can cluster the samples as in the following plots using spectral clustering. For `toy1`, use $K = 5, 6, 7$ as the number of clusters. For `toy2`, you can use $K = 3$.
- Compare your results with what you get using k-means clustering. You can use the `kmeans` function in MATLAB.



Part 2: Spectral Clustering & Real Data (10 pts)

You have many choices when doing spectral clustering such as how to form the similarity matrix and the Laplacian [1]. Use spectral clustering to divide the **features** in the real data set into clusters as follows:

- Load `realdata.mat`. It is a 292 samples by 1456 features matrix. Contrary to what we would usually be interested in, i.e., clustering samples, we will cluster features in this exercise.

- Use different approaches to see if spectral clustering can divide the features into several clusters. The ground truth is an open research question so for this part, I only want to see what you have tried and what you have observed.

Deliverables

Work in pairs and send the following by email to evrim.acarataman@gmail.com:

- MATLAB scripts for spectral clustering called `unnormalized_spec.m` and `normalized_spec.m`
- For each toy data set, a MATLAB script performing two different approaches for spectral clustering and plotting the clusters as well as k-means clustering & its plot.
- For the real data set, a MATLAB script showing what you have tried and a brief description.

References

- [1] U. von Luxburg. “A tutorial on spectral clustering,” *Statistics and Computing*, vol. 17, pp. 395-416, 2007.
- [2] A. Y. Ng, M. I. Jordan, and Y. Weiss, “On Spectral Clustering: Analysis and an algorithm,” *NIPS’01: Proceedings of the 14th International Conference on Neural Information Processing Systems*, pp. 849–856, 2001.