MATH 567: Mathematical Techniques in Data Science
Lab 12

Dominique Guillot

Departments of Mathematical Sciences
University of Delaware

May 10, 2017

Exercise 1: Graph Clustering

Challenge: You are given a graph containing 3 communities. Follow the steps below to discover them.

1. Use the `load` command to load the file `graph.RData` (on Sakai). The variable `Ap` is the adjacency matrix of a graph.
2. Use the `image` command to draw the adjacency matrix of the graph.
3. Use the following code to draw the graph

```
library(igraph)
G = graph_from_adjacency_matrix(Ap, "undirected")
plot(G)
# Use tkplot(G) if you want an interactive graph
```

4. Construct the (unnormalized) graph Laplacian $L = D - A_p$.
5. Compute the eigenvalues and eigenvectors of L.
6. Order the eigenvalues/eigenvectors:

```
I = order(e$values)
eval = e$values[I]
evec = e$vectors[,I]
```

Exercise 1 (cont.)

7. Plot the 2nd and 3rd eigenvectors: `plot(evec[,2], evec[,3])`.
8. Use the k-means algorithm to find 3 clusters in `evec[,2:3]`.
9. Permute the vertices of the graph according to the k-means clustering:

```
perm = order(clus$cluster)
Arec = Ap[perm, perm]
```

10. Draw the adjacency matrix of `Arec`.

Exercise 2: Spectral Clustering of Images

1. Load the file "scat.jpg" (on Sakai) using the `readJPEG` command from the `jpeg` package.

```
cat = readJPEG('scat.jpg')
```

2. Convert the image to a list of pixel colors

```
p1 = dim(cat)[1]
p2 = dim(cat)[2]
p = p1*p2
catflat = matrix(cat, nrow=p)
```

3. Use spectral clustering to find 4 clusters in the image:

```
sc = specc(catflat, centers=4)
```

4. Reshape the cluster assignment, and plot it as an image:

```
Ccat = matrix(sc, nrow=p1, ncol=p2)
image(Ccat)
```