Worksheet on “Spectral Clustering”

PRML – CS5691 (Jul–Nov 2023)

October 10, 2023

1. Find the Laplacian matrix for the following graph, and use R/Python (R code provided here) to plot a heatmap of its eigen vectors. Does the eigen vectors corresponding to the smallest a eigen values reveal the two clusters in the graph?

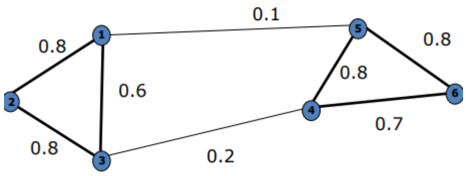

Figure 1:

2. You have a dataset of four data points in two-dimensional space:
Data Point 1: (1, 2)
Data Point 2: (2, 3)
Data Point 3: (3, 4)
Data Point 4: (4, 5)
Perform spectral clustering to group these data points into two clusters. The similarity matrix has been computed using a Gaussian kernel with a bandwidth parameter of $\sigma = 1$, resulting in the following similarity matrix:

```
S = [[1.0, 0.9, 0.6, 0.3],
     [0.9, 1.0, 0.8, 0.5],
     [0.6, 0.8, 1.0, 0.9],
     [0.3, 0.5, 0.9, 1.0]]
```

Figure 2:

Perform the following steps of spectral clustering:

i) Compute the (unnormalized) Laplacian matrix L.
ii) Compute the eigenvalues and eigenvectors of L.
iii) Select the eigenvector corresponding to the smallest eigenvalue, normalize the eigenvector so that it is unit-norm, and check if applying a threshold on it will reveal the two clusters?
iv) Select the eigenvector corresponding to the second smallest eigenvalue of L, normalize this eigenvector, and check if choosing a threshold of 0.7 will reveal two similar clusters in the dataset?

In the last two parts, provide the clusters obtained after performing spectral clustering and explain your reasoning for the cluster assignments based on the eigenvector and threshold.

3. Let G be a simple undirected graph over n nodes, with d_i denoting the degree of the ith node. If the eigen values of the graph Laplacian L of G are ordered from the smallest to the largest (e.g., second smallest eigenvalue is λ_2), then show that $\lambda_2 \leq \frac{n}{n-1}\bar{d} \leq \lambda_n$. Here, $\bar{d} = \frac{1}{n}\sum_i d_i$ is the average degree of a node.

4. Prove the following : Let G = (V, E) be a graph, and let i and j be vertices of degree one that are both connected to another vertex k. Then, the vector v given by

$$\boldsymbol{v}(u) = \begin{cases} 1 & u = i \\ -1 & u = j \\ 0 & otherwise, \end{cases}$$

Figure 3:

is an eigenvector of the Laplacian of G of eigenvalue 1.

5. Prove the following : The graph Sn has eigenvalue 0 with multiplicity 1, eigenvalue 1 with multiplicity n 2, and eigenvalue n with multiplicity 1.