Homework 6
MATH 123 - Spring 2023
Tufts University, Department of Mathematics
Due: March 14, 2023

Question 1

Let $L = D - W$ be the unnormalized graph Laplacian associated to a graph $\mathcal{G} = (V, W)$ on points $\{x_i\}_{i=1}^n$ with symmetric weight matrix W and diagonal degree matrix D. Let $\{C, \overline{C}\}$ be any partition of $\{x_i\}_{i=1}^n$, and let

$$f_i^C = \begin{cases} -\sqrt{\text{vol}(\overline{C})/\text{vol}(C)} & x_i \in C \\ \sqrt{\text{vol}(C)/\text{vol}(\overline{C})} & x_i \in \overline{C} \end{cases}$$

(a) Prove that $\langle Df^c, \mathbb{1} \rangle = 0$, where $\mathbb{1} = (1, 1, \ldots, 1)$ is the vector of all 1's.

(b) Prove that $(f^C)^T D f^C = \text{vol}(V)$.

(c) Prove that $(f^C)^T L f^C = \text{vol}(V)\text{Ncut}(C, \overline{C})$.

Question 2

Recall that one construction of the weight matrix for a graph on data $\{x_i\}_{i=1}^n$ is to use the Gaussian kernel $W_{ij} = \exp(-\|x_i - x_j\|_2^2/\sigma^2), i \neq j$ and $W_{ij} = 0, i = j$ for some choice of $\sigma > 0$.

(a) What happens to the resulting Laplacian matrix L as $\sigma \to 0^+$?

(b) What happens to the resulting Laplacian matrix L as $\sigma \to \infty$?

Question 3

Load the dataset "SalinasA_corrected.mat" and "Salinas-S-groundtruth" from `http://www.ehu.eus/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes#Salinas_scene`.

(a) Run spectral clustering on this data, using a sparse Laplacian with different numbers of nearest neighbors and $K = 6$ clusters. How do the results compare to the ground truth data?

(b) Plot the first 10 eigenvalues of the data for different choices of σ. What does the eigengap estimate as the number of clusters for these choices of σ?

(c) Compare the projections onto the first three principle components with the first three Laplacian eigenvectors by plotting both sets in different figures using 'scatter3'. How do the representations differ qualitatively?