Homework 5
MATH 123 - Spring 2023
Tufts University, Department of Mathematics
Due: March 7, 2023

1. QUESTION 1

Let $\{x_i\}_{i=1}^n \subset \mathbb{R}^D$ be a discrete set on unique points. Recall that the DBSCAN algorithm depends on two parameters: ϵ and MinPts.

(a) Describe the behavior of DBSCAN as $\epsilon \to +\infty$ and as $\epsilon \to 0^+$.

(b) Describe the behavior of DBSCAN as MinPts$\to +\infty$ and as MinPts$\to 0^+$

2. QUESTION 2

Let $L = D - W \in \mathbb{R}^{n \times n}$ be the graph Laplacian for data with associated symmetric weight matrix W with $W_{ij} \in [0,1]$ for all $i, j = 1, \ldots, n$.

(a) Show L is positive semidefinite.

(b) Show L is not positive definite by proving 0 is an eigenvalue of L.

3. QUESTION 3

Compute the graph Laplacian with $W_{ij} = \exp(-\|x_i - x_j\|_2^2/\sigma^2)$ on the image in Ncut_Data.mat using a range of σ. For each of these σ, use the second eigenvector (i.e. the eigenvector with second smallest eigenvalue) to segment the image by thresholding at 0. Discuss the results. Do they make sense? How do the results depend on σ? Note, this data is an image of a pepper, of size 31×100. To cluster it with spectral methods, one should treat it as a graph with 3100 nodes and an edge between each pixel.