Homework 4
MATH 123 - Fall 2018
Tufts University, Department of Mathematics
Due: October 9, 2018

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

(a) Implement your own version of DBSCAN in MATLAB. You may find the pseudocode on wikipedia helpful. Do not use any pre-existing implementations.

(b) Run your algorithm on the dataset contained in DBSCAN_Data.mat for a range of ϵ and MinPts values to show that the algorithm can succeed or fail, depending on how these parameters are set.

3. Question 3

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.

4. Question 4

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?