

Due: 19 Nov 2020

1. **Laplacian eigenmap.** Show that the Laplacian eigenmap to \mathbb{R}^m is the solution to the following optimization problem:

$$\min \sum_{i,j} k_{ij} \|y_i - y_j\|_2^2 \quad \text{subject to} \quad Y^T Q Y = I, \quad Y^T Q \mathbf{1}_{m \times 1} = 0.$$

Here, the y_i are columns of $Y \in \mathbb{R}^{m \times m}$ and the rest of the notation is as in the lecture notes. Also show that $\sum_{i,j} k_{ij} \|y_i - y_j\|_2^2 = \text{tr} Y^T L Y$ where L is the graph Laplacian.

Solution. UNFINISHED

2. **Dimensional reduction comparison.** The goal of this problem is to practice and compare various methods for dimensional reduction. Use the following methods:

- PCA
- Isomap
- LLE
- t-SNE
- Diffusion maps.

Diffusion maps should be programmed from scratch, but the rest can be programmed using library functions. If you use library functions, specify their source, read their descriptions, and be ready to adjust their parameters. There are three datasets:

- an S-curve generated by a Matlab function
- an S-curve perturbed by Gaussian noise—try various intensities and push the methods' limits
- the “emoji” dataset generated by a Matlab function—picking a good ϵ for diffusion map might be tricky since the nearest neighbor distances are very nonuniform. You should be able to get a nice 2D surface in 3D with the right ϵ . Use either $\alpha = 0$ or $\alpha = 1$ (up to you).

Submit a report on the performance of these methods on each dataset.

Solution. I generated the datasets in Matlab and saved them to files, then read and processed them in Python.
UNFINISHED