HW 5 AMSC 808N

Konstantinos Pantazis

November 18, 2020

1 Solution of exercise 1

Let Y be the $n \times m$ matrix which corresponds to the Laplacian eigenmap to $\mathbb{R}^m$ for all data points x_i. This means that each row of Y corresponds to the image of the Laplacian eigemap of each datum. The Lagrangian function is given by

$$L(Y, \lambda, \mu) = \text{trace}(Y^T LY) - \lambda(Y^T QY - I) - \mu Y^T Q \mathbb{1}_{n\times 1}$$

The derivative in terms of Y is

$$\begin{aligned} \frac{dL}{dY} &= (L + L^T)Y - \lambda(Q + Q^T)Y - \mu := 0 \\ \Rightarrow LY &= \lambda QY \text{ and } \mu = 0 \end{aligned}$$

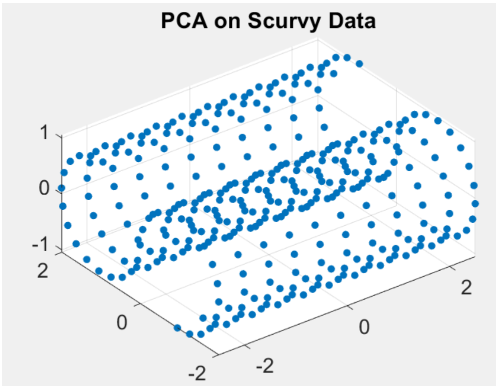
which is the generalized eigenvalue problem. Also,

$$\frac{dL^2}{d^2Y} > 0$$

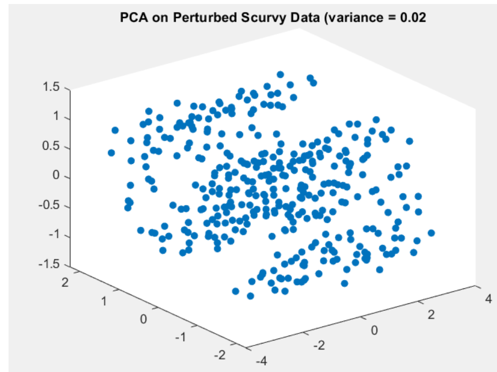
Hence, the truncated matrix of right eigenvectors of P is the optimal solution for this optimization problem.

2 Solution of exercise 2

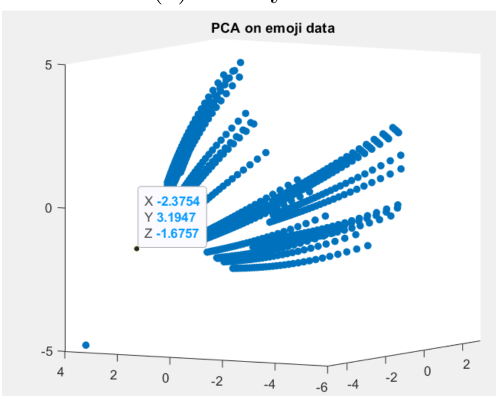
I. **PCA:**



(a) Scurvy data



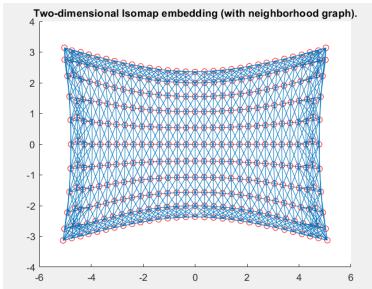
(b) Perturbed scurvy data



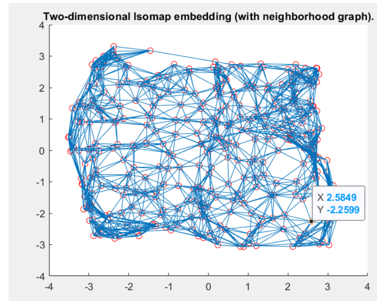
(c) Emoji data

Figure 1: PCA

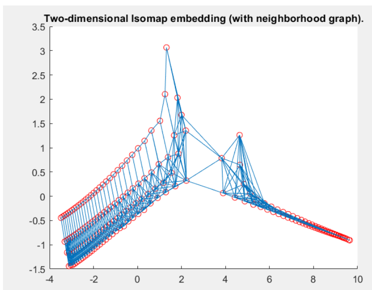
II. **ISOMAP:** The retreived embedding dimension by ISOMAP algorithm is $d = 2$.



(a) Scurvy data

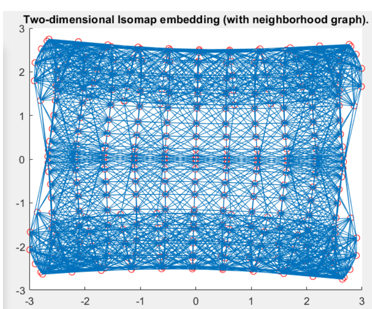


(b) Perturbed scurvy data

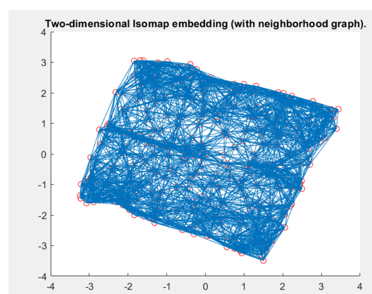


(c) Emoji data

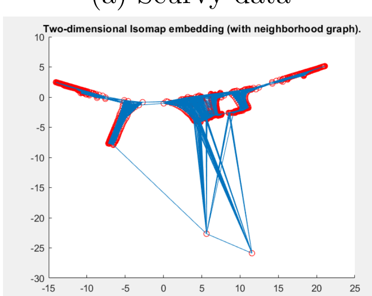
Figure 2: ISOMAP and kNN with $k = 10$ neighbors.



(a) Scurvy data



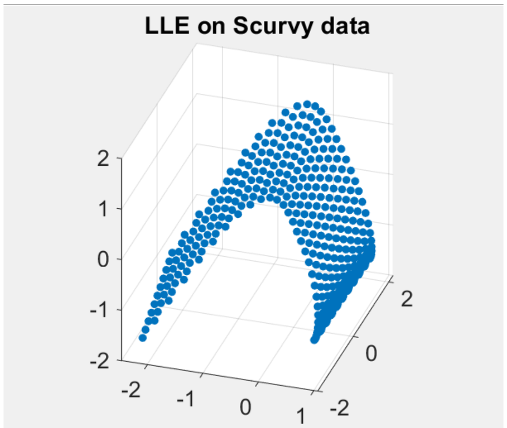
(b) Perturbed scurvy data



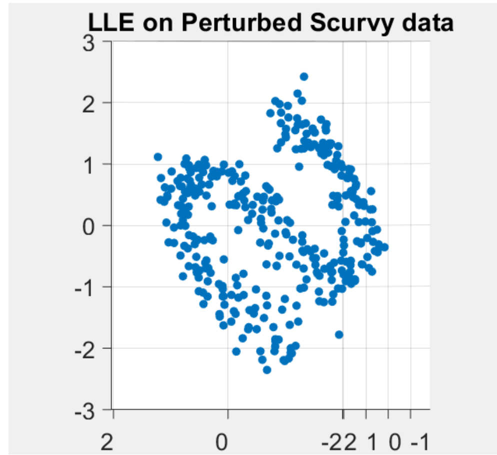
(c) Emoji data

Figure 3: ISOMAP and kNN with $k = 20$ neighbors.

III. **LLE:**

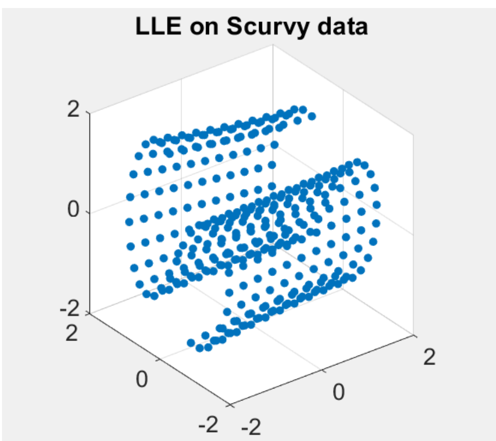


(a) Scurvy data

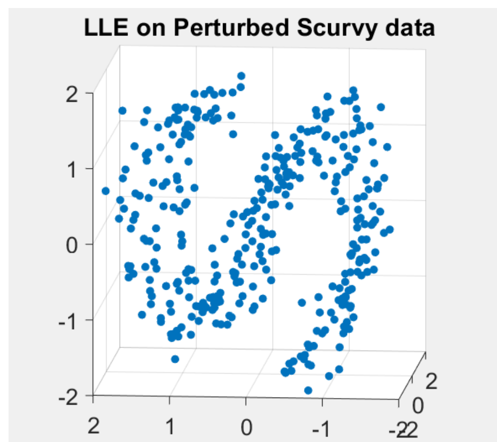


(b) Perturbed scurvy data

Figure 4: LLE and kNN with $k = 10$ neighbors.



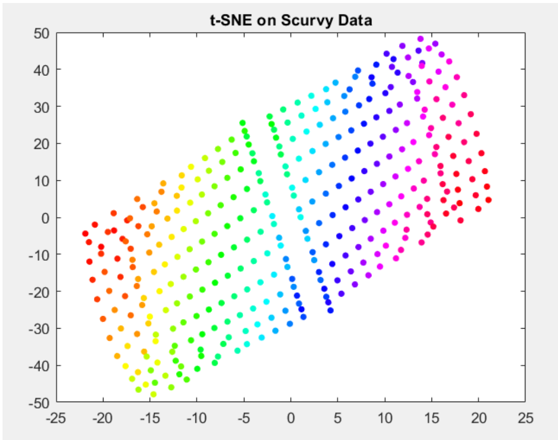
(a) Scurvy data



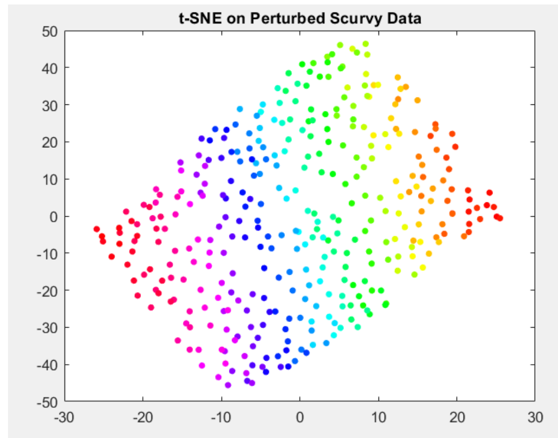
(b) Perturbed scurvy data

Figure 5: LLE and kNN with $k = 20$ neighbors.

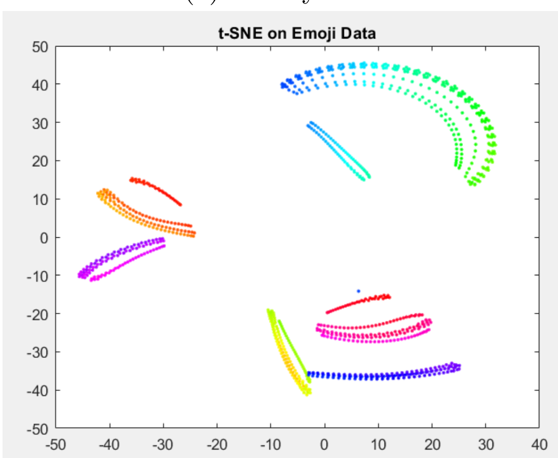
IV. **t-SNE:**



(a) Scurvy data



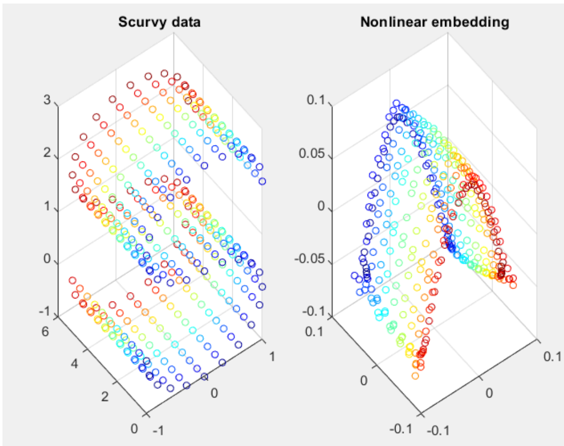
(b) Perturbed scurvy data



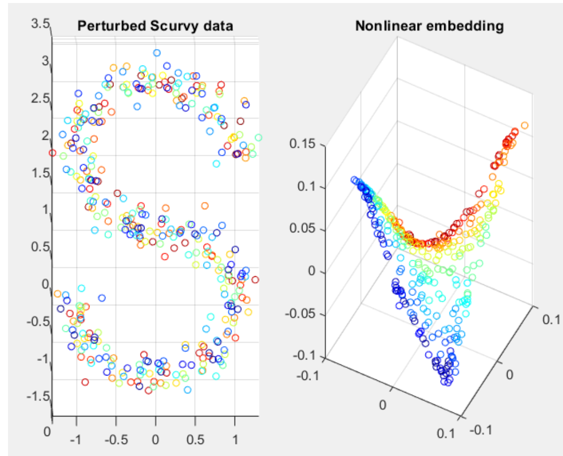
(c) Emoji data

Figure 6: tSNE

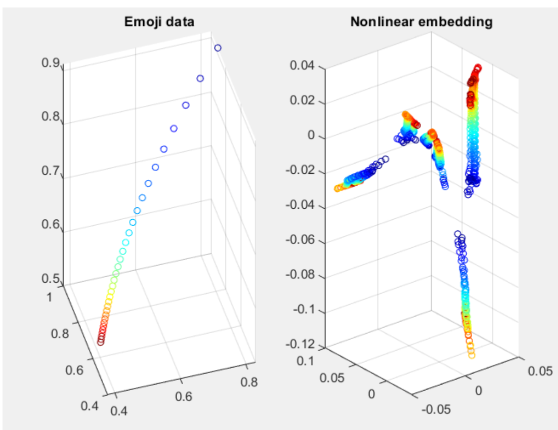
V. **Diffusion map:**



(a) Scurvy data



(b) Perturbed scurvy data



(c) Emoji data

Figure 7: Diffusion map with $\alpha = 0$, $\epsilon = 100$ and $t = 5$.