Manifold Learning and Sparse Representation: Assignment 5

ZHANG Yuan, 1601111332

Exercise 76.

Proof. Becasue Frobenius norm is unitary invariant, we have

$$\begin{aligned} L = \|\mathbf{X} - \mathbf{A}\|_F^2 =& \|\mathbf{U^T X U} - \mathbf{\Lambda}\|_F^2 \\ =& \|\tilde{\mathbf{X}} - \mathbf{\Lambda}\|_F^2 \\ =& \sum_i (\tilde{x}_{ii} - \lambda_i)^2 + 2\sum_{i<j} x_{ij}^2 \\ \geq& \sum_i (\tilde{x}_{ii} - \lambda_i)^2, \end{aligned}$$

where $\mathbf{A} = \mathbf{U\Lambda U^T}$ is the eigen-decomposition of $\mathbf{A}$ and $\tilde{\mathbf{X}} \triangleq \mathbf{U^T X U}$.

Since $dim\{\mathbf{a}|\mathbf{a^T X a} > 0\} = rank(\mathbf{X}) \leq r$, the number of non-zero diagonal elements (i.e. $\{x_i i\}$) of $\mathbf{X}$ is less or equal than r. Besides, the fact that X is p.s.d. implies all the non-zero diagonal elements should be positive. Thus, the solution to $\min .L$ is $\tilde{\mathbf{X}} = diag(\max(\lambda_1, 0), \max(\lambda_r, 0), 0, .., 0)$ so that $\mathbf{X} = \mathbf{U_r} \max(\mathbf{\Lambda_r}, \mathbf{0})\mathbf{U_r^T}$. □

Exercise 77.

The visualization of CMDS with $k = 2$ is shown as in Fig. 1. *Square Euclidean distance* that is widely used in many clustering scenarios is used in my settings.

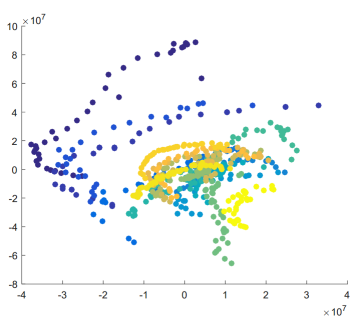

Figure 1: 2-D embedding visualization w/ square Euclidean distance.

Exercise 93.

The first 12 faces of each person are used for training and another 5 faces for testing. The average recognition rate is 0.8103.