Homework 5 **Hongyu Yan (516030910595)**
CS 259 Numerical Methods for Data Science ACM Class, Zhiyuan College, SJTU
Prof. **David Bindel** Due Date: June 27, 2018
TA. **Yurong You, Xinran Zhu** Submit Date: July 25, 2018

Problem 1

As A is symmetric, the difference between Λ and Σ is that the values in the diagonal of Σ are non-negative and in descending order, while those in Λ are not necessarily positive and unordered.

To compute the SVD, we can make it by change the sign of negative entries of Λ and sort them in descending order.

So the alogrithm can be described as:

1. Generate proper matrix L s.t. $L\Lambda$ is non-negative. In other words, if $\Lambda_{ii} \leq 0$, then $L_{ii} = -1$, otherwise $L_{ii} = 1$.

2. Generate proper elementary matrix T s.t. $T(L\Lambda)T$ is the matrix that the values in the diagonal are in descending order.

3. So that $\Sigma = TL\Lambda T$, and we have $A = Q\Lambda Q^T = Q(TL)^{-1}(TL\Lambda T)T^{-1}Q = Q(TL)^{-1}\Sigma T^{-1}Q$. So $U = Q(TL)^{-1}$, $V = (T^{-1}Q)^T$.

The time complexity of this algorithm is $O(n^2)$, because generating the sorted matrix needs $O(n^2)$.

By the way, I think it's also ok by sorting values in $O(nlogn)$ then generating the matrix T in $O(n)$.

Problem 2

Suppose there are n points $x_1, x_2, \cdots, x_n$, so $A \in \mathbb{R}^{n\times n}$. $X = \begin{bmatrix} x_1^T \\ x_2^T \\ \vdots \\ x_n^T \end{bmatrix}$, $X \in \mathbb{R}^{n\times d}$

The purpose is to prove that $B = XX^T$, and we have one solution for X as $X = \begin{bmatrix} B_{11} \\ B_{12} \end{bmatrix} R_{11}^{-1}$.

First, from the definition of B, we got:

$$\begin{aligned} B &= -\frac{1}{2}JAJ \\ &= -\frac{1}{2}(I - \frac{ee^T}{n})A(I - \frac{ee^T}{n}) \\ &= -\frac{1}{2}(A - A\frac{ee^T}{n} - \frac{ee^T}{n}A + \frac{ee^T}{n}A\frac{ee^T}{n}) \end{aligned}$$

Specifically, we have:

$$B_{ij} = -\frac{1}{2}(A_{ij} - \frac{1}{n}\sum_{k=1}^{n} A_{ik} - \frac{1}{n}\sum_{l=1}^{n} A_{lj} + \frac{1}{n^2}\sum_{k=1}^{n}\sum_{l=1}^{n} A_{lk})$$

Because $A_{ij} = \|x_i - x_j\|^2 = \|x_i\|^2 - 2 < x_i, x_j > + \|x_j\|^2$, substitute it into previous equation, we get:

$$B_{ij} = -\frac{1}{2}(A_{ij} - \|x_i\|^2 - \|x_j\|^2) = < x_i, x_j >$$

So $B = XX^T$ and B is positive semi-definite.

Second, we will prove that $X = \begin{bmatrix} B_{11} \\ B_{12} \end{bmatrix} R_{11}^{-1}$ is a solution.

Suppose there are d "landmark" points, then we can split X into $\begin{bmatrix} X_1^T \\ X_2^T \end{bmatrix}$, where X_1 contains the d "landmark" points, and X_2 contains the other $n - k$ points. So

$$B = XX^T = \begin{bmatrix} X_1^T \\ X_2^T \end{bmatrix} \begin{bmatrix} X_1 & X_2 \end{bmatrix} = \begin{bmatrix} X_1^T X_1 & X_1^T X_2 \\ X_2^T X_1 & X_2^T X_2 \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

$$B_{12} = B_{21}^T$$

Since $B_{11} = R_{11}^T R_{11}$, so

$$\begin{aligned}
B_{11} &= R_{11}^T R_{11} \\
B_{12} &= R_{11}^T (B_{21} R_{11}^{-1})^T \\
B_{21} &= B_{21} R_{11}^{-1} R_{11} \\
B_{22} &= X_2^T X_2 \\
&= X_2^T X_1 (X_1^T X_1)^{-1} X_1^T X_2 \\
&= X_2^T X_1 (B_{11})^{-1} X_1^T X_2 \\
&= B_{21} R_{11}^{-1} (R_{11}^T)^{-1} B_{21}^T \\
&= (B_{21} R_{11}^{-1})(B_{21} R_{11}^{-1})^T
\end{aligned}$$

so that:

$$B = \begin{bmatrix} R_{11}^T R_{11} & R_{11}^T (B_{21} R_{11}^{-1})^T \\ B_{21} R_{11}^{-1} R_{11} & (B_{21} R_{11}^{-1})(B_{21} R_{11}^{-1})^T \end{bmatrix} = \begin{bmatrix} B_{11} \\ B_{21} \end{bmatrix} R_{11}^{-1} (\begin{bmatrix} B_{11} \\ B_{21} \end{bmatrix} R_{11}^{-1})^T$$

Problem 3

$$\begin{aligned}
\|A - LMR\|_F^2 &= tr((A - LMR)^T (A - LMR)) \\
&= tr((A^T - R^T M^T L^T)(A - LMR)) \\
&= tr(A^T A - R^T M^T L^T A - A^T LMR + R^T M^T L^T LMR) \\
&= tr(A^T A) - 2tr(R^T M^T L^T A) + tr(R^T M^T L^T LMR)
\end{aligned}$$

We get the minimum when the derivative of $\|A - LMR\|_F^2$ equals 0. Because

$$\begin{aligned}
\frac{\partial tr(A^T A)}{\partial M} &= 0 \\
\frac{\partial tr(R^T M^T L^T A)}{\partial M} &= \frac{\partial tr(M^T L^T A R^T)}{\partial M} = L^T A R^T \\
\frac{\partial tr(R^T M^T L^T LMR)}{\partial M} &= 2L^T LMRR^T
\end{aligned}$$

so that

$$\frac{\partial \|A - LMR\|_F^2}{\partial M} = -2L^T A R^T + 2L^T L M R R^T = 0$$

which means

$$\begin{aligned} L^T A R^T &= L^T L M R R^T \\ M &= (LTL)^{-1} L^T A R^T (R R^T)^{-1} = L^\dagger A R^\dagger \end{aligned}$$