Homework1 solution

授课时间: 2019 年 7 月 5 日　授课教师: John E. Hopcroft　批改：孙奕灿，吴铭东

所有结论正确，过程合理的解答均可得到全部分数，结论错误，过程正确也可以得到相应的步骤分。

1　3.5

Manually find the left and right-singular vectors, the singular values, and the SVD decomposition of the folloing two matrices

(a):

$$\begin{pmatrix} 1 & 1 \\ 0 & 3 \\ 3 & 0 \end{pmatrix}$$

(b):

$$\begin{pmatrix} 0 & 2 \\ 2 & 0 \\ 1 & 3 \\ 3 & 1 \end{pmatrix}$$

Solution: （20 分）

(a):

$$U = (u_1, u_2) = \begin{pmatrix} \frac{\sqrt{22}}{11} & 0 \\ \frac{3\sqrt{22}}{22} & -\frac{\sqrt{2}}{2} \\ \frac{3\sqrt{22}}{22} & \frac{\sqrt{2}}{2} \end{pmatrix}$$

$$D = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} = \begin{pmatrix} \sqrt{11} & 0 \\ 0 & 3 \end{pmatrix}$$

$$V = (v_1, v_2) = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{pmatrix}$$

$$M = \sqrt{11} u_1 v_1^T + 3 u_2 v_2^T.$$

(b):

$$U=(u_1,u_2)=\begin{pmatrix}\frac{\sqrt{10}}{10} & -\frac{1}{2}\\ \frac{\sqrt{10}}{10} & \frac{1}{2}\\ \frac{\sqrt{10}}{5} & -\frac{1}{2}\\ \frac{\sqrt{10}}{5} & \frac{1}{2}\end{pmatrix}$$

$$D=\begin{pmatrix}\sigma_1 & 0\\ 0 & \sigma_2\end{pmatrix}=\begin{pmatrix}2\sqrt{5} & 0\\ 0 & 2\sqrt{2}\end{pmatrix}$$

$$V=(v_1,v_2)=\begin{pmatrix}\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2}\end{pmatrix}$$

$$M=2\sqrt{5}u_1v_1^T+2\sqrt{2}u_2v_2^T$$

2 3.12

Let $\sum_i \sigma_i u_i v_i^T$ be the singular value decomposition of a rank r matrix A, Let $A_k=\sum_{i=1}^{k}\sigma_i u_i v_i^T$ be a rank k approximation to A for some $k<r$,Express the following quantities in terms of the singular values $\{\sigma_i, 1\le i\le r\}$

1. $||A_k||_F^2$
2. $||A_k||_2^2$
3. $||A-A_k||_F^2$
4. $||A-A_k||_2^2$

Solution:（20 分，每小题 5 分）

2-norm 对应的是最大奇异值，F-norm 是逐元素平方和开根号.

1. $\sum_{i=1}^k \sigma_i^2 \quad (||A_k||_F^2=\mathrm{tr}(A_k^TA_k)=\mathrm{tr}(\sum_{1\le i,j\le k}\sigma_i\sigma_j v_i u_i^T u_j v_j^T)=\sum_{i=1}^k \sigma_i^2)$
2. σ_1^2
3. $\sum_{i=k+1}^r \sigma_i^2$
4. σ_{k+1}^2.

3 3.18

1. For $n=5,10,...,25$ create random graphs by generating random vectors $x=(x_1,x_2,\ldots,x_n)$, and $y=(y_1,y_2,\ldots,y_n)$. Create edges $(x_i,y_i)-(x_{i+1},y_{i+1})$ for $i=1:n$ and an edge $(x_n,y_n)-(x_1,y_1)$.

2. For each graph create a new graph by selecting the midpoint of each edge for the coordinates of the vertices and add edges between vertices corresponding to the midpoints of two adjacent edges of the original graph. What happens when you iterate this process? It is best to draw the graphs.

3. Repeat the above step but normalize the vectors x and y to have unit length after each iteration. What happens?

4. One could implement the process by matrix multiplication where $x(t)$ and $y(t)$ are the vectors at the t iteration. What is the matrix A such that $x(t+1)=Ax(t)$.

5. What is the first singular vector of A and the first two singular values of A. Does this explain what happens and how long the process takes to converge?

6. If A is invertible what happens when you run the process backwards.

Solution: (40 分, 每小题 8 分)

2. 现象：这些点组成的图形渐渐地变成了椭圆，然后这个椭圆越来越小，最终收敛到了一个点。

3. 现象：和 2 基本相同。(特别指出：如果我们每次迭代以后把 x 数组平均化到均值为 0，最后会变成一个椭圆.)

4.

$$\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & ... & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & ... & 0 & 0 \\ & & ... & & & \\ 0 & 0 & 0 & ... & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & ... & 0 & \frac{1}{2} \end{pmatrix}$$

5. 第一个奇异值为 1。第二个奇异值为 $\cos(\frac{\pi}{n}) < 1$。

这是因为基本循环矩阵 $B = F_n D F_n^T$，其中 F_n 为傅里叶矩阵，D 为 $\omega_n^0 \sim \omega_n^{n-1}$ 排成的对角矩阵。由

$$AA^T = \frac{1}{4}(I+B)(I+B^T) = \frac{1}{4}(2I + B + B^{n-1})$$

知 AA^T 的特征值就是

$$f(x) = \frac{1}{4}(2 + x + x^{n-1}), \forall x = \omega_n^k$$

即

$$\frac{1}{2}\left(1 + \cos\left(\frac{2\pi k}{n}\right)\right) = \cos^2\left(\frac{\pi k}{n}\right)$$

6.n 是偶数时 A 不可逆，n 是奇数时是可逆的，逆矩阵 B 为：$(y-x) \bmod p$ 是奇数时为-1，否则为 1。行列式为 2^{n-1}。乘这个矩阵相当于将平滑操作给倒着做。如果最初所有点坐标都相同，则逆操作不会改变它们，否则会使点之间距离扩大并使点的坐标趋向于 $+\infty$。

4 3.32

1. Consider the pairwise distance matrix for twenty US cities given below. Use the algorithm of Exercise 3.31 to place the cities on a map of the US. The algorithm is called classical multidimensional scaling, cmdscale, in Matlab. Alternatively use the pairwise distance matrix of 12 Chinese cities to place the cities on a map of China. Note: Any rotation or a mirror image of the map will have the same pairwise distance.

2. Suppose you had airline distances for 50 cities around the world. Could you use these distances to construct a 3-dimensional world model?

Solution: (20 分，第一小题 15 分，第二小题 5 分)

1. 可以用 Matlab 中的 cmdscale 直接做，也可以参考 3.31 中的做法.

2. 可以。如果是给出球面距离（非欧氏距离），内积矩阵可能出现负的特征值，不能用经典多维标度算法来做，需要用优化算法来求解。需要设立一个偏离度度量，比如

$$\text{Stress}\left(\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2, \cdots, \tilde{\mathbf{x}}_n\right) = \left(\sum_{i<j}\left(\|\tilde{\mathbf{x}}_i - \tilde{\mathbf{x}}_j\| - d_{ij}\right)^2 / \sum_{i<j} d_{ij}^2\right)^{1/2}$$

其中 $\tilde{\mathbf{x}}_1$ 为球坐标下的数据点。