Algorithm Foundations of Data Science and Engineering
Welcome Tutorial :-)
Tutorial 6

GAO Ming

DaSE @ ECNU

1 Apr., 2019

Tutorial 6

1. Let A be

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

   a. Find the sigular values of matrix A;

$$A^T A = \begin{pmatrix} 5 & 5 \\ 5 & 5 \end{pmatrix}$$

$$det(A^T A - \lambda I) = \begin{vmatrix} 5-\lambda & 5 \\ 5 & 5-\lambda \end{vmatrix} = \lambda(\lambda - 10) = 0$$

$$\lambda_1 = 10, \lambda_2 = 0$$

so, we have the sigular values: $\sqrt{10}, 0$

Tutorial 6 Cont'd

1. b. Find the SVD of matrix A.
The eigenvector of A^TA corresponding to eigenvalues $\lambda_1 = 10, \lambda_2 = 0$ are: $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})^T$, $(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})^T$.

$$AA^T = \begin{pmatrix} 2 & 4 & 0 \\ 4 & 8 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

The eigenvector of AA^T corresponding to eigenvalues $\lambda_1 = 10, \lambda_2 = 0$ are: $(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}, 0)^T$, $(-\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}, 0)^T$, $(0, 0, 1)^T$.
So, the SVD of matrix A is:

$$A = \begin{pmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} & 0 \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{10} & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

Tutorial 6 Cont'd

2. Given the following 3-dimensional data

data	1	2	3
x	1	-1	4
y	2	1	3
z	1	3	-1

   a. Compute the co-variance matrix of the data;

$$C = \begin{pmatrix} \frac{19}{3} & \frac{5}{2} & -5 \\ \frac{5}{2} & 1 & -2 \\ -5 & -2 & 4 \end{pmatrix}$$

Tutorial 6 Cont'd

2. b. Find the principle components of the data.

$$det(C - \lambda I) = \begin{vmatrix} \frac{19}{3} & \frac{5}{2} & -5 \\ \frac{5}{2} & 1 & -2 \\ -5 & -2 & 4 \end{vmatrix} = 0$$

$$\lambda_1 = 11.296, \lambda_2 = 0.037, \lambda_3 = 0$$

The eigenvector of C corresponding to eigenvalues are:
$x_1 = [-0.748, -0.297, 0.594]^T$,
$x_2 = [0.664, -0.334, 0.689]^T, x_3 = [0, -0.894, -0.447]^T$.
Take the eigenvector x_1 corresponding to the eigenvalue λ_1 and project the data into this space, then, we have the principle components of the data, that is:
$z = [-0.748, 2.233, -4.477]^T$

Tutorial 6 Cont'd

3. Given $f(x) = x^3 - 6x^2 + 11x - 6$. Using Newton method, find all roots of the equation $f(x) = 0$.
   因为牛顿法得到的最优解是局部的，所以初值选择不同，结果可能也不同。
   1）选择初值$x_0 = 0$
   迭代得到的x分别为：
   $0.545455, 0.848953, 0.974674, 0.999092, 0.999999, 1.000000$，收敛。
   所以，得到一个根为1.
   2)选择初值$x_0 = 2$
   因为$f(2) = 0$，所以得到一个根为2.
   3）选择初值$x_0 = 4$
   迭代得到的x分别为：
   $3.454545, 3.151046, 3.025326, 3.000908, 3.000002, 3.000001$，收敛。
   所以，得到一个根为3.

4. Let J be the reconstruction error in matrix factorization,

$$J = \frac{1}{2} \sum_{(u,i) \in \mathscr{K}} (r_{ui} - \mathbf{p}_u^T \mathbf{q}_i)^2$$

where r_{ui} is the known rating of user u for item i, $\widehat{u}_{ui} = \mathbf{p}_u^T \mathbf{q}_i$ is the predicted rating given by user u for item i, and the set of all user-item pair (u, i), which are observed in R, be denoted by $\mathscr{K}$, i.e., $\mathscr{K} = \{(u, i) | r_{ui} \text{ is observed}\}$.

   a. Show J is a convex function of parameters P and Q; 反证法：若J是关于P和Q的凸函数，则最优解存在且唯一
   将最终迭代得到的P中的第一、第二列互换，Q中的第一、第二行互换，损失函数J的值不变，这与最优解P和Q存在且唯一的假设矛盾。所以，J不是关于P和Q的凸函数。

Tutorial 6 Cont'd

4. b. Compute $\frac{\partial J}{\partial p_{uj}}$ and $\frac{\partial J}{\partial q_{ji}}$.

$$\frac{\partial J}{\partial p_{uj}} = -\sum_i (r_{ui} - \sum_j p_{uj} q_{ji}) q_{ji}$$
$$\frac{\partial J}{\partial q_{ji}} = -\sum_u (r_{ui} - \sum_j p_{uj} q_{ji}) p_{uj}$$

c. When we employ Gradient descent method to learning the parameters, give the update rules for parameters p_{uj} and q_{ji}.

$$p_{uj}^{(t+1)} \leftarrow p_{uj}^{(t)} + \alpha e_{ui}^{(t)} q_{ji}^{(t)}$$
$$q_{ji}^{(t+1)} \leftarrow q_{ji}^{(t)} + \alpha e_{ui}^{(t)} p_{uj}^{(t)}$$

其中，$e_{ui}^{(t)} = r_{ui} - \sum_j p_{uj} q_{ji}$.