1. Collaborative Filtering, Kernels, Linear Regression

In this question, we will use the alternating projections algorithm for low-rank matrix factorization, which aims to minimize

$$J(U,V) = \underbrace{\frac{1}{2}\sum_{(a,i)\in D}\left(Y_{ai} - [UV^T]_{ai}\right)^2}_{\text{Squared Error}} + \underbrace{\frac{\lambda}{2}\sum_{a=1}^{n}\sum_{j=1}^{k}U_{aj}^2 + \frac{\lambda}{2}\sum_{i=1}^{m}\sum_{j=1}^{k}V_{ij}^2}_{\text{Regularization}}.$$

In the following, we will call the first term the squared error term, and the two terms with λ the regularization terms.

Let Y be defined as

$$Y = \begin{bmatrix} 5 & ? & 7 \\ ? & 2 & ? \\ 4 & ? & ? \\ ? & 3 & 6 \end{bmatrix}$$

D is defined as the set of indices (a,i), where $Y_{a,i}$ is not missing. In this problem, we let $k=\lambda=1$. Additionally, U and V are initialized as $U^{(0)} = [6,0,3,6]^T$, and $V^{(0)} = [4,2,1]^T$.

1. (a)

1.0/1 point (graded)

Compute X, the matrix of predicted rankings UV^T given the initial values for U and V.

[[24, 12, 6],[0, 0, 0],[12, ✔ **Answer:** [[24, 12, 6], [0, 0, 0], [12, 6, 3], [24, 12, 6]]

Solution:

- the predicted rankings should be the matrix produce between U and V^T.

$$X = UV^T = \begin{bmatrix} 24 & 12 & 6 \\ 0 & 0 & 0 \\ 12 & 6 & 3 \\ 24 & 12 & 6 \end{bmatrix}$$

Submit You have used 1 of 3 attempts

Answers are displayed within the problem

1. (b)

2/2 points (graded)

Compute the squared error term, and the regularization terms in for the current estimate X.

Enter the squared error term (including the factor $1/2$):

255.5

✔ **Answer:** 255.5

Enter the regularization term (the sum of all the regularization terms):

51

✔ **Answer:** 51

Solution:

$$\begin{aligned} J_{\text{square}} &= \sum_{i,j \in D} (Y_{ij} - X_{ij})^2/2 \\ &= \frac{1}{2}\left((5-24)^2 + (7-6)^2 + (2-0)^2 + (4-12)^2 + (3-12)^2 + (6-6)^2\right) = 255.5 \\ J_{\text{reg}} &= \frac{\lambda}{2}\|U\|_F^2 + \frac{\lambda}{2}\|V\|_F^2 \\ &= \frac{\lambda}{2}\sum_{a=1}^{n}(U_a)^2 + \frac{\lambda}{2}\sum_{i=1}^{m}(V_i)^2 = 51 \end{aligned}$$

Submit You have used 1 of 3 attempts

Answers are displayed within the problem

1. (c)

1.0/1 point (graded)

Suppose V is kept fixed. Run one step of the algorithm to find the new estimate $U^{(1)}$.

Enter the $U^{(1)}$ as a list of numbers, $[U_1^{(1)}, U_2^{(1)}, U_3^{(1)}, U_4^{(1)}]$:

[3/2,4/5,16/17,2]

✔ **Answer:** [3/2, 4/5, 16/17, 2]

Solution:

With V fixed as $[4,2,1]^T$, we can represent prediction X as:

$$X = UV^T = \begin{bmatrix} 4U_1 & 2U_1 & 1U_1 \\ 4U_2 & 2U_2 & 1U_2 \\ 4U_3 & 2U_3 & 1U_3 \\ 4U_4 & 2U_4 & 1U_4 \end{bmatrix}$$

Let D be the set of index of observation, the estimate $U^{(1)}$ should be:

$$\begin{aligned} U^{(1)} &= \arg\min_U \; J(U) \\ &= \arg\min_U \sum_{(a,i)\in D} (Y_{ai} - (UV)_{ai})^2/2 + \sum_{a=1}^{4}\frac{\lambda}{2}\|U_a\|^2 \\ &= \arg\min_U \; [(5-4U_1)^2 + (7-U_1)^2 + (2-2U_2)^2 + (4-4U_3)^2 + (3-2U_4)^2 + (6-U_4)^2]/2 + \sum_{a=1}^{4}\frac{1}{2}U_a^2 \end{aligned}$$

To minimize this loss, we take the gradient with respect to U and equate it to zero.

$$0 = \nabla J(U) = \begin{pmatrix} -4(5-4U_1) - (7-U_1) + U_1 \\ -2(2-2U_2) + U_2 \\ -4(4-4U_3) + U_3 \\ -2(3-2U_4) - (6-U_4) + U_4 \end{pmatrix} = \begin{pmatrix} -27 + 18U_1 \\ -4 + 5U_2 \\ -16 + 17U_3 \\ -12 + 6U_4 \end{pmatrix}$$

Hence,

$$U_1^{(1)} = \frac{3}{2}$$
$$U_2^{(1)} = \frac{4}{5}$$
$$U_3^{(1)} = \frac{16}{17}$$
$$U_4^{(1)} = 2$$

Submit You have used 3 of 3 attempts

Answers are displayed within the problem

Discussion

Show Discussion

Topic: Unit 2 Nonlinear Classification, Linear regression, Collaborative Filtering (2 weeks):Homework 3 / 1. Collaborative Filtering, Kernels, Linear Regression