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

Solution:

- the predicted rankings should be the matrix product 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 2 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):

51.0 ✔ **Answer:** 51

Solution:

$$\begin{aligned} J_{\text{square}} &= \sum_{i,j \in D} (Y_{ij} - X_{ij})^2 / 2 \\ &= \frac{1}{2}\left((5-24)^2 + (7-6)^2 + (2-0)^2 + (4-12)^2 + (3-12)^2 + (6-6)^2\right) = 255.5 \end{aligned}$$

$$J_{\text{reg}} = \frac{\lambda}{2}\|U\|_F^2 + \frac{\lambda}{2}\|V\|_F^2$$
$$= \frac{\lambda}{2}\sum_{a=1}^{n}(U_a)^2 + \frac{\lambda}{2}\sum_{i=1}^{m}(V_i)^2 = 51$$

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, $[U_1^{(1)}, U_2^{(1)}, U_3^{(1)}, U_4^{(1)}]$:

[1.5, 0.8, 0.94117647,2.0 ✔ **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:

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

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 1 of 3 attempts

Answers are displayed within the problem

Discussion

Hide Discussion

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

Add a Post

Show all posts ▼ by recent activity ▼

I can see regulation loss here, but what's the symptom of over fitting for Collaborative Filtering? 9

Community TA

[Staff] 1(a) Incomplete formulation of question 12

Staff my first answer was on the 3 percent difference I am learning and that is what matters to me, but my sloppy handwriting made me see a ? as a 7 (seven) and then my answer differs. On other co...	2
Sage way If you are tired of computing using pen and paper... sage: var('u1,u2,u3,u4,v1,v2,v3') (u1, u2, u3, u4, v1, v2, v3) sage: U = matrix([u1,u2,u3,u4]).tran... Community TA	23
Hints I'm very stuck on where to even begin on this problem. Can anyone give me a hint on where to start with 1a? It would be very much appreciated, ...	9
matrix produce typo On the answer	2
[Staff] Clarification 1 (b) For the regularization term in 1. (b) it would be helpful to make clear if we should include the half factor or not. I was doing something wrong an...	2
[STAFF] Problem 1. (a) Hi, Could you please take a look at my solution to the first problem ? I believe I got the correct figures but somehow my answer was marked 'Inco...	7
Why only one regularization term? Why do we only use one regularization for both the U and V matrices? Would it make sense to use unique regularization terms for each matrix?	3
[Hint] Deciding Which U0 V0 and Lambda to Choose how do we decide which U0,V0 and lambda to choose? We decide which rank to use on development set. I dont know what is development set b...	2
is it OK to write in (c) answer list elements as fractions a/b ? subj	4
Is there an elegant formula for the objective function using linear algebra notation?	5
Can someone from staff please check my answer for 1(c)? I checked my calculation thrice but always getting red light from the grader.	2