Course <u>Progress</u> <u>Dates</u> <u>Discussion</u> <u>Resources</u>

☆ Course / Unit 2 Nonlinear Classification, Linear regression, Collaborative Filtering (2 weeks) / Homework 2

Previous	ď	ď	ď	Z ~	ď	Next >

1. Collaborative Filtering, Kernels, Linear Regression

□ Bookmark this page

Homework due Oct 14, 2020 05:29 IST Past Due

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

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

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

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

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 point possible (graded)

Compute $X^{(0)}$, the matrix of predicted rankings UV^T given the initial values for $U^{(0)}$ and $V^{(0)}$.

(Enter your answer as a matrix, e.g., type **[[2,1],[1,0],[3,-1]]** for a 2×3 matrix $\begin{pmatrix} 2 & 1 \\ 1 & 0 \\ 3 & -1 \end{pmatrix}$. Note the square

brackets, and commas as separators.)

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

Solution:

ullet the predicted rankings should be the matrix product between U and V^T .

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

Submit

You have used 0 of 3 attempts

1 Answers are displayed within the problem

z points possible (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):

Answer: 255.5

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

Answer: 51

Solution:

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

Submit

You have used 0 of 3 attempts

1 Answers are displayed within the problem

1. (c)

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

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

Solution:

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

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

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

$$\begin{split} U^{(1)} &= \operatorname*{arg\,min}_{U} \quad J\left(U\right) \\ &= \operatorname*{arg\,min}_{U} \quad \sum_{(a,i) \in D} (Y_{ai} - (UV)_{ai})^{2}/2 + \sum_{a=1}^{4} \frac{\lambda}{2} \|U_{a}\|^{2} \\ &= \operatorname*{arg\,min}_{U} \quad \left[(5 - 4U_{1})^{2} + (7 - U_{1})^{2} + (2 - 2U_{2})^{2} + (4 - 4U_{3})^{2} + (3 - 2U_{4})^{2} + (6 - U_{4})^{2} \right]/2 + \sum_{a=1}^{4} \frac{1}{2} U \end{split}$$

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

$$0 =
abla J\left(U
ight) = egin{pmatrix} -4\left(5-4U_{1}
ight)-\left(7-U_{1}
ight)+U_{1} \ -2\left(2-2U_{2}
ight)+U_{2} \ -4\left(4-4U_{3}
ight)+U_{3} \ -2\left(3-2U_{4}
ight)-\left(6-U_{4}
ight)+U_{4} \end{pmatrix} = egin{pmatrix} -27+18U_{1} \ -4+5U_{2} \ -16+17U_{3} \ -12+6U_{4} \end{pmatrix}$$

Hence,

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

Submit

You have used 0 of 3 attempts

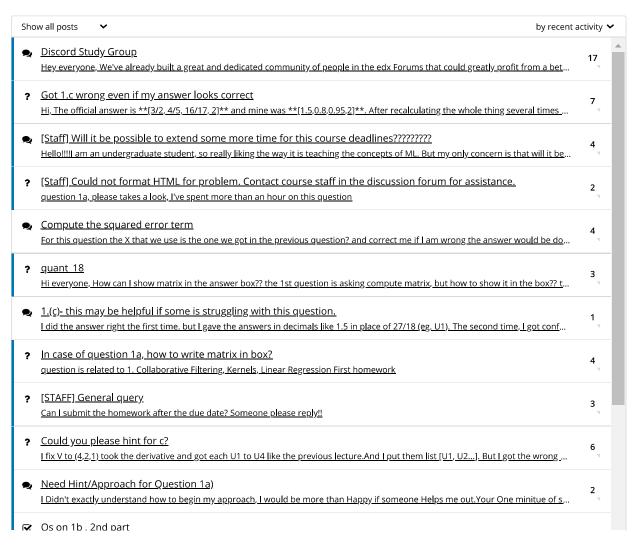
• Answers are displayed within the problem

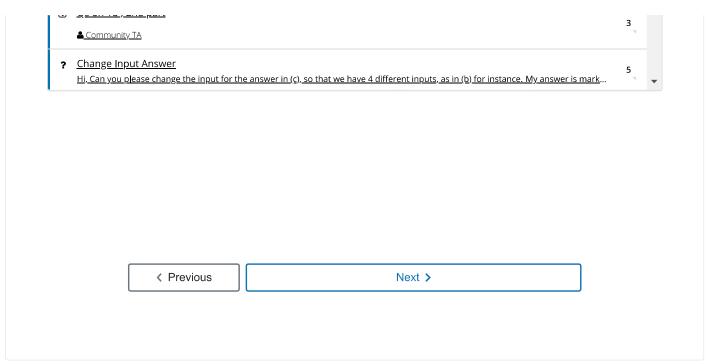
Discussion

Hide Discussion

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

Add a Post





© All Rights Reserved



edX

<u>About</u>

Affiliates

edX for Business

Open edX

Careers

News

Legal

Terms of Service & Honor Code

Privacy Policy

Accessibility Policy

Trademark Policy

<u>Sitemap</u>

Connect

<u>Blog</u>

Contact Us

Help Center

Media Kit

Donate















