



[Unit 2 Nonlinear Classification,](#)
[Linear regression, Collaborative](#)
[Course](#) > [Filtering \(2 weeks\)](#)
4. Collaborative Filtering: the Naive
Approach

> [Lecture 7. Recommender Systems](#) >

Audit Access Expires May 11, 2020

You lose all access to this course, including your progress, on May 11, 2020.
Upgrade by Mar 25, 2020 to get unlimited access to the course as long as it
exists on the site. **[Upgrade now](#)**

4. Collaborative Filtering: the Naive Approach

Collaborative Filtering: the Naive Approach



Video

[Download video file](#)

Transcripts

[Download SubRip \(.srt\) file](#)

[Download Text \(.txt\) file](#)

Compute the Derivative of the Regression Objective

2.0/2 points (graded)

Recall that each user a has a set of movies that (s)he has already rated. Let Y be a matrix with n row and m columns whose $(a, i)^{\text{th}}$ entry Y_{ai} is the rating by user a of movie i if this rating has already been given, and blank if not. Our goal is to come up with a matrix X that has no blank entries and whose $(a, i)^{\text{th}}$ entry X_{ai} is the prediction of the rating user a will give to movie i .

Let D be the set of all (a, i) 's for which a user rating Y_{ai} exists, i.e. $(a, i) \in D$ if and only if the rating of user a to movie i exists.

A naive approach to solve this problem would be to minimize the following objective:

$$J(X) = \sum_{a,i \in D} \frac{(Y_{ai} - X_{ai})^2}{2} + \frac{\lambda}{2} \sum_{(a,i)} X_{ai}^2$$

where the first term is the sum of the squared errors for entries with observed rating, and the second term is a regularization term roughly to prevent the predictions to become extremely large, and the parameter λ controls the balance between these two terms.

Compute the derivative $\frac{\partial J}{\partial X_{ai}}$ of the objective function $J(X)$. (Note that $J(X)$ can be viewed as a function of the variables X_{ai} .)

(Type $x_{\{ai\}}$ for matrix entries X_{ai} , $y_{\{ai\}}$ for matrix entries Y_{ai} and "lambda" for λ . Note that X and Y are capital letters to represent matrices.)

For (any fixed) $(a, i) \in D$,

$$\frac{\partial J}{\partial X_{ai}} = \boxed{-(Y_{\{ai\}} - X_{\{ai\}}) + \text{lambda} * X_{\{ai\}}}$$

✔ Answer: $X_{\{ai\}} - Y_{\{ai\}} + \text{lambda} * X_{\{ai\}}$

For (any fixed) $(a, i) \notin D$:

$$\frac{\partial J}{\partial X_{ai}} = \text{lambda} * X_{ai}$$

✓ Answer: $\text{lambda} * X_{ai}$

STANDARD NOTATION

Solution:

Derive the objective and remember to treat any entry in the matrix that is not the one that we are deriving by as a constant. Hence, the derivative of all components of the sum that are not (a, i) will be zero.

Submit

You have used 2 of 2 attempts

📘 Answers are displayed within the problem

Performance of the Naive Approach

2.0/2 points (graded)

Let us now check the quality of the solution when using this wrong approach. Recall the naive approach assumes independence between all entries of the matrix.

What value of the matrix X will minimize the loss

$$J(X) = \sum_{a,i \in D} \frac{(Y_{ai} - X_{ai})^2}{2} + \frac{\lambda}{2} \sum_{(a,i)} X_{ai}^2? \text{ That is, for each } (a, i), \text{ solve the}$$

following equation for X_{ai} :

$$\frac{\partial J}{\partial X_{ai}} = 0.$$

We will denote the argmin as \widehat{X} and its components as \widehat{X}_{ai} .

For $(a, i) \in D$:

$\hat{X}_{ai} =$ $Y_{\{ai\}}/(1+\lambda)$ For $(a, i) \notin D$: $\hat{X}_{ai} =$

0

STANDARD NOTATION

Submit

You have used 1 of 3 attempts

Discussion






[Hide Discussion](#)

Topic: Unit 2 Nonlinear Classification, Linear regression, Collaborative Filtering (2 weeks):Lecture 7. Recommender Systems / 4. Collaborative Filtering: the Naive Approach

[Add a Post](#)

Show all posts

by recent activity

- | | | |
|---|---|---|
|  | [STAFF] Bug:Answer marked correct but asks to keep submitting. 2nd problem I got the right answer but it wants me to keep submitting saying I have used 1/3 attempts. | 4 |
|  | Why can't I see "Show Answer" for the question "Performance of the Naive Approach"??
Why can't I see "Show Answer" for the question "Performance of the Naive Approach"?? | 4 |
|  | At 5min05s shouldn't it be norm(X) squared? | 2 |
|  | lamda lambda | 2 |
|  | Staff - For question #1, I entered the correct answers, but grader reported problems evaluating the input
For question #1, on second attempt, I entered the correct answers for both parts, and tried t... | 6 |

✓	<u>What would be wrong with $(Y_{\{a,i\}})/(1+\lambda)$. OR $Y_{\{a,i\}}/(1+\lambda)$</u>	3
	<u>Please, could someone help me with my formatting because I am trying to submit the answe...</u>	
?	<u>Could not format HTML for problem. Contact course staff in the discussion forum for assistance.</u>	4
	<u>On the first question with the gradients the grader would not check my answers, though I ha...</u>	
💬	<u>"Compute the Derivative of the Regression Objective": For $(a,i) \notin D$</u>	3
	<u>In the 1st question, for $(a,i) \notin D$, why partial derivative is zero? ... my result is $(1+\lambda) \cdot X_{\{ai\}}$... ca...</u>	
?	<u>Exercise 01 grader issue?</u>	5
	<u>is the grader for the first part of the Compute the Derivative of the Regression Objective exer...</u>	

Learn About Verified Certificates

© All Rights Reserved