4. Collaborative Filtering: the Naive Approach

Bookmark this page

Exercises due Oct 14, 2020 05:29 IST *Completed*

Collaborative Filtering: the Naive Approach

Start of transcript. Skip to the end.

So right now we will start talking
about collaborative filtering.
This is actually the main part of
today's lecture.
So let's write collaborative filtering.
So now I want to start thinking about
this problem
in the same terms as what we've done

▶ 0:00 / 0:00 ▸ 1.25x

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 theses 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) $(a,i) \in D$,

$\frac{\partial J}{\partial X_{ai}} =$ X_{ai}-Y_{ai}+lambda*X_{ai} ✔ **Answer:** X_{ai}-Y_{ai}+lambda*X_{ai}

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

$\frac{\partial J}{\partial X_{ai}} =$ lambda*X_{ai} **Answer:** 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 1 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 $J(X) = \sum_{a,i \in D} \frac{(Y_{ai} - X_{ai})^2}{2} + \frac{\lambda}{2} \sum_{(a,i)} X_{ai}^2$? That is, for each (a,i), 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 $(a,i) \in D$:

$\widehat{X}_{ai} =$ Y_{ai}/(1+lambda) **Answer:** Y_{ai}/(1+lambda)

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

$\widehat{X}_{ai} =$ 0 **Answer:** 0

STANDARD NOTATION

Solution:

Derive the objective (see previous question) and compare to zero to find the values at the minima. Using the results from the problem above, we have For $(a,i) \in D$:

$$\frac{\partial J}{\partial X_{ai}} = X_{ai} - Y_{ai} + \lambda X_{ai} = 0 \iff X_{ai} = \frac{Y_{ai}}{1+\lambda}$$

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

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

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):Lecture 7. Recommender Systems / 4. Collaborative Filtering: the Naive Approach

Add a Post

Show all posts — by recent activity

?	about finding minimum of the function Sorry for silly question, probably it should have poped up in my head during Unit 1 :) So, when we want to find a minimum: 1) we take d...	3
?	Why there is no summation in derivative result? Should derivative result sum all derivative forms together?	2
?	(staff)'Cannot parse formula' errors.. I get the errors both and I do not think I am wrong. Expression is clear.	4
	I cannot use lambda in any of the answers, why? the question says it can be used I get this error in any of the answers in this page Invalid Input: \'lamda\' not permitted in answer as a variable	3
?	Euclidean square norm	2
	Hint: if you receive 'Cannot parse formula' errors.. If you receive parse errors when submitting the formula, there is probably a mistake in the submitted formula, as in my case where I ov...	5
?	Why does "Performance of the Naive Approach" not show the correct answer? I got the second part of the answer correct and the first part wrong. How can I get the correct answer after I have used my 3/3 attempts?	8
?	First Question Is it possible that the answers for question one are inverted?. Perhaps i misunderstood the concepto of D. If (a,i) is on D.	4
?	(staff) Derivations details (Staff) How much do I have to simplify de derivation on the first question? I answered with the most simplificated form and i got it wron...	6
?	Could not format HTML for problem. Contact course staff in the discussion forum for assistance. Could not format HTML for problem. Contact course staff in the discussion forum for assistance.	5

Previous — Next