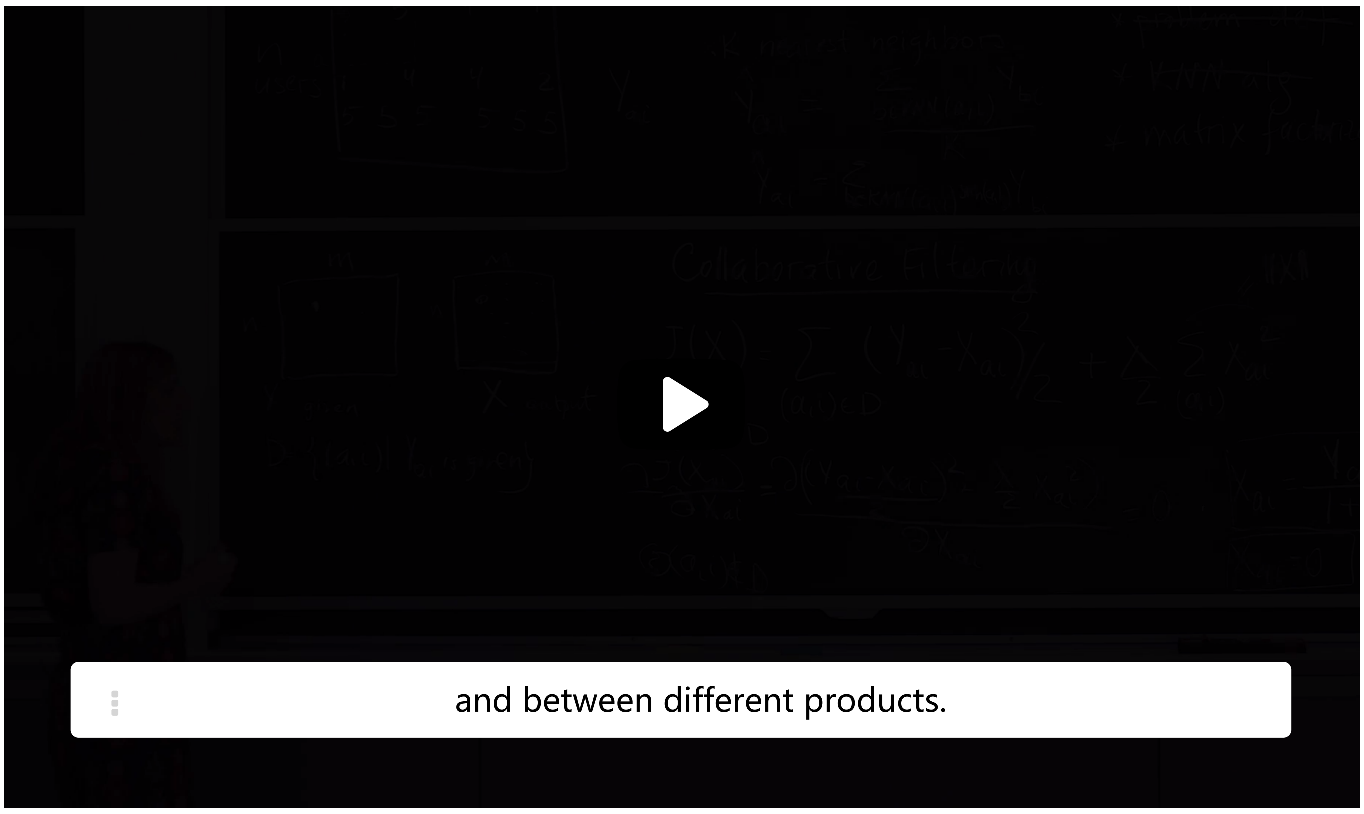
4. Collaborative Filtering: the Naive Approach

Collaborative Filtering: the Naive Approach



10:18 / 10:18 1.0x

making.
We're treating every choice independently.
And it is not surprising that since we're not modeling dependency in any way, we're really losing important connection, which was the first reason why we decided to look at this problem differently in order to find the connection-- the hidden connection between different users
and between different products.

End of transcript. Skip to the start.

Video
Download video file

Transcripts
Download SubRip (.srt) file
Download Text (.txt) file

Compute the Derivative of the Regression Objective

2/2 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 $\dfrac{\partial J}{\partial X_{ai}}$ 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.)

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

$\frac{\partial J}{\partial X_{ai}} =$ -Y_{ai}+(1+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 2 of 2 attempts

Answers are displayed within the problem

Performance of the Naive Approach

2/2 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)

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

$\widehat{X}_{ai} =$ 0

STANDARD NOTATION

Submit You have used 1 of 3 attempts

Correct (2/2 points)

Discussion

Show Discussion

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

© All Rights Reserved