

<u>Unit 2 Nonlinear Classification</u>, <u>Linear regression, Collaborative</u>

<u>Course</u> > <u>Filtering (2 weeks)</u>

6. Alternating Minimization

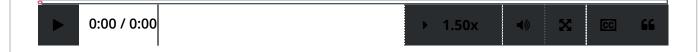
> Lecture 7. Recommender Systems >

### **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** 

# 6. Alternating Minimization Alternating Minimization





Video

Download video file

**Transcripts** 

<u>Download SubRip (.srt) file</u> Download Text (.txt) file

## **Alternating Minimization Concept Question**

1/1 point (graded)

As in the video above, we now want to find  $\boldsymbol{U}$  and  $\boldsymbol{V}$  that minimize our new objective

$$J = \sum_{(a,i) \in D} rac{\left(Y_{ai} - \left[UV^T
ight]_{ai}
ight)^2}{2} + rac{\lambda}{2} \Biggl(\sum_{a,k} U_{ak}^2 + \sum_{i,k} V_{ik}^2 \Biggr) \,.$$

To simplify the problem, we fix U and solve for V, then fix V to be the result from the previous step and solve for U, and repeat this alternate process until we find the solution.

Consider the case k=1. The matrices U and V reduce to vectors u and v, such that  $u_a=U_{a1}$  and  $v_i=V_{i1}$ .

When v is fixed, minimizing J becomes equivalent to minimizing ...

$$iggl( rac{\left( Y_{ai} - u_a v_i 
ight)^2}{2} + rac{\lambda}{2} \sum_a \left( u_a 
ight)^2$$

$$igotimes_{(a,i)\in D} rac{\left(Y_{ai}-u_av_i
ight)^2}{2} + rac{\lambda}{2} \sum_a \left(u_a
ight)^2$$

$$igcup_{(a,i)\in D}rac{\left(Y_{ai}-u_av_i
ight)^2}{2}$$

$$igcup_{(a,i)\in D} rac{\left(Y_{ai}-u_av_i
ight)^2}{2} + rac{\lambda}{2} \sum_i \left(v_i
ight)^2$$



#### **Solution:**

Regarding terms containing only  ${\cal V}$  as constants, minimizing  ${\cal J}$  is equivalent to minimizing

$$\sum_{(a,i)\in D}rac{\left(Y_{ai}-u_av_i
ight)^2}{2}+rac{\lambda}{2}\sum_a{(u_a)^2}.$$

Submit

You have used 1 of 3 attempts

Answers are displayed within the problem

# Fixing V and Finding U

2.0/2 points (graded)

Now, assume we have 2 users, 3 movies, and a 2 by 3 matrix Y given by

$$Y=egin{bmatrix}1&8&?\2&?&5\end{bmatrix}$$

Our goal is to find U and V such that  $X=UV^T$  closely approximates the observed ratings in Y.

Assume we start by fixing V to initial values of  $\begin{bmatrix} 4,2,1 \end{bmatrix}^T$ . Find the optimal  $2\times 1$  vector U in this case. (Express your answer in terms of  $\lambda$ ).

First element of U is:

20/(lambda+20)

✓ Answer: 20/(20+lambda)

The second element of U is:

13/(lambda + 17)

**✓ Answer:** 13/(17+lambda)

**STANDARD NOTATION** 

#### **Solution:**

To compute the first element ( $u_1$ ), compute the objective (ignore missing elements from Y), derive and compare to zero to find the minimum:

$$rac{\partial}{\partial u_1}[rac{(1-4u_1)^2}{2}+rac{(8-2u_1)^2}{2}+rac{\lambda}{2}u_1^2]=(\lambda+20)\,u_1-20=0.$$

Submit

You have used 1 of 3 attempts

**1** Answers are displayed within the problem

# Discussion

**Topic:** Unit 2 Nonlinear Classification, Linear regression, Collaborative Filtering (2 weeks):Lecture 7. Recommender Systems / 6. Alternating Minimization

**Hide Discussion** 

#### Add a Post

Show all posts by recent activity		
<b>∀</b>	[staff]Wrong rate for the second U  Dear TA, Could you please check the lambda coefficient for the second user in the answer? I t	6
2	Hint: Code to compute derivative with Python  Pseudo code to compute a derivative wrt X of a function in python import sympy as sym X = s	3
2	<u>Thank you</u> <u>Hi, Thank you for the very clear explanation. The instructor has an exceptional capability to e</u>	6
2	(Staff) I think that there is a mistake in one formula	4
<b>Y</b>	How/Why Lambda was set equal to 1 in the video?  Maybe a trivial question but I can't remember when or how the value of lambda was set equ	2
<b>\( \right\)</b>	Initialisation of V  How exactly was V initialised? The prof just took a V without any explanation.	5
<b>\( \right\)</b>	When will the lecture slides be available?  Dear Staffs, I have been waiting for the lecture slides for several days. It was promised to be	2
Ą	norm of multiplication of two vectors should norm of multiplication of two vectors, U and V, be product of norm of U and norm of	3
2	How to take such a derivative?  Is the derivative in the video at 9:00 correct? Can someone point me to a description on how	2
?	Regularization parameter mechanics – I don't get it.  The Squared error term is pretty clear. The more we deviate from the values picked by actual	3
?	HTML error  Ljust saved a reply for the second exercise, but now I cannot see the exercise. I can only read	2

Learn About Verified Certificates

© All Rights Reserved