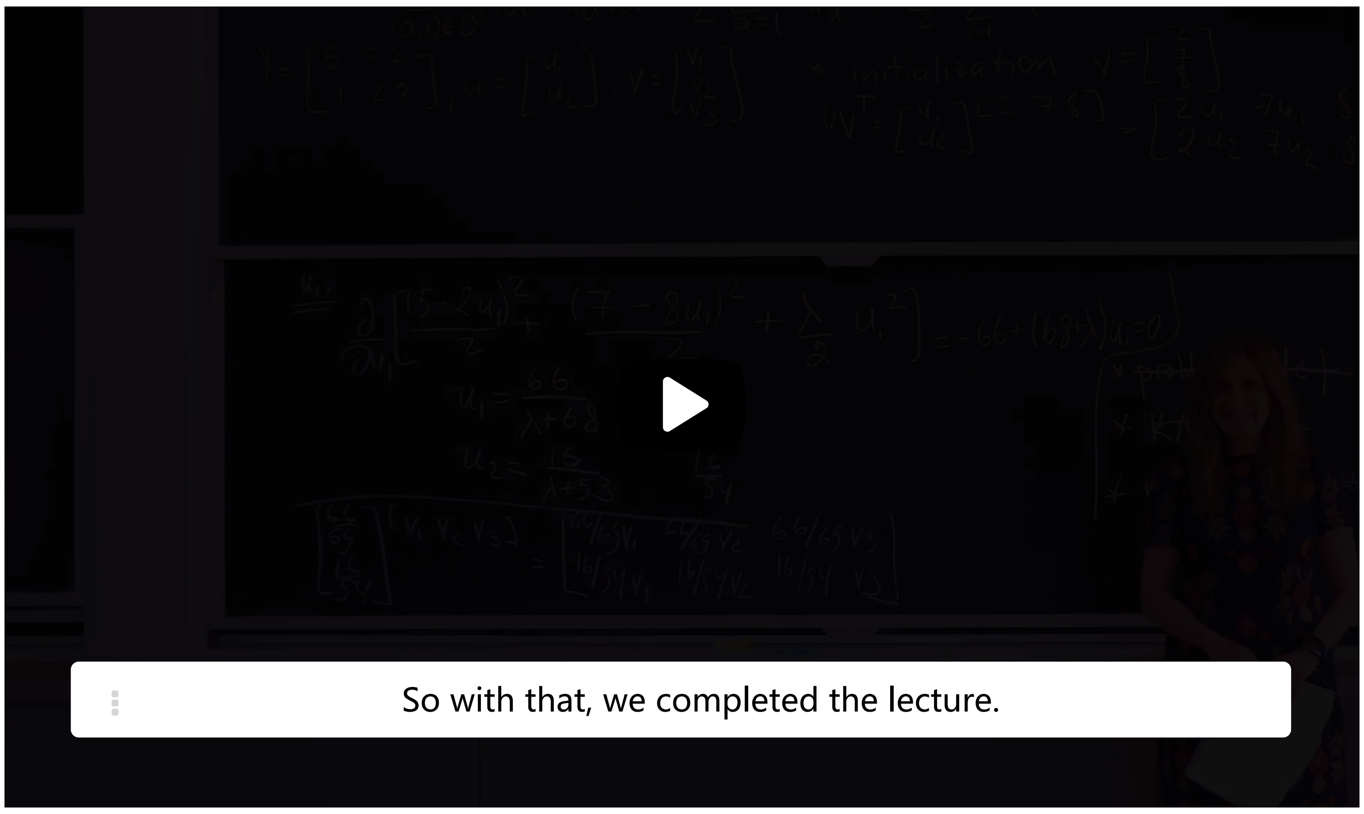
6. Alternating Minimization

Alternating Minimization



14:17 / 14:17 1.0x

But what I would like to tell you,
that this very simple algorithm actually
enables you, in a very interesting way,
to find connection between different users and products.
And given this very relatively simple machinery,
you can actually solve a very non-trivial problem
of product recommendation.
So with that, we completed the lecture.

End of transcript. Skip to the start.

Video
Download video file

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

xuetangX.com 学堂在线

Alternating Minimization Concept Question

1/1 point (graded)
As in the video above, we now want to find U and V that minimize our new objective

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

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 ...

- $\frac{(Y_{ai} - u_a v_i)^2}{2} + \frac{\lambda}{2}\sum_a (u_a)^2$
- $\sum_{(a,i)\in D} \frac{(Y_{ai} - u_a v_i)^2}{2} + \frac{\lambda}{2}\sum_a (u_a)^2$ ✔

$$\sum_{(a,i)\in D} \frac{(Y_{ai} - u_a v_i)^2}{2}$$

$$\sum_{(a,i)\in D} \frac{(Y_{ai} - u_a v_i)^2}{2} + \frac{\lambda}{2} \sum_i (v_i)^2$$

Solution:

Regarding terms containing only V as constants, minimizing J is equivalent to minimizing

$$\sum_{(a,i)\in D} \frac{(Y_{ai} - u_a v_i)^2}{2} + \frac{\lambda}{2} \sum_a (u_a)^2.$$

Submit You have used 1 of 3 attempts

Answers are displayed within the problem

Fixing V and Finding U

2/2 points (graded)

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

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

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 $[4, 2, 1]^T$. Find the optimal 2×1 vector U in this case. (Express your answer in terms of λ).

First element of U is:

20/(20+lambda) **Answer:** 20/(20+lambda)

The second element of U is:

13/(17+lambda) **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:

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

Submit You have used 2 of 3 attempts

Answers are displayed within the problem