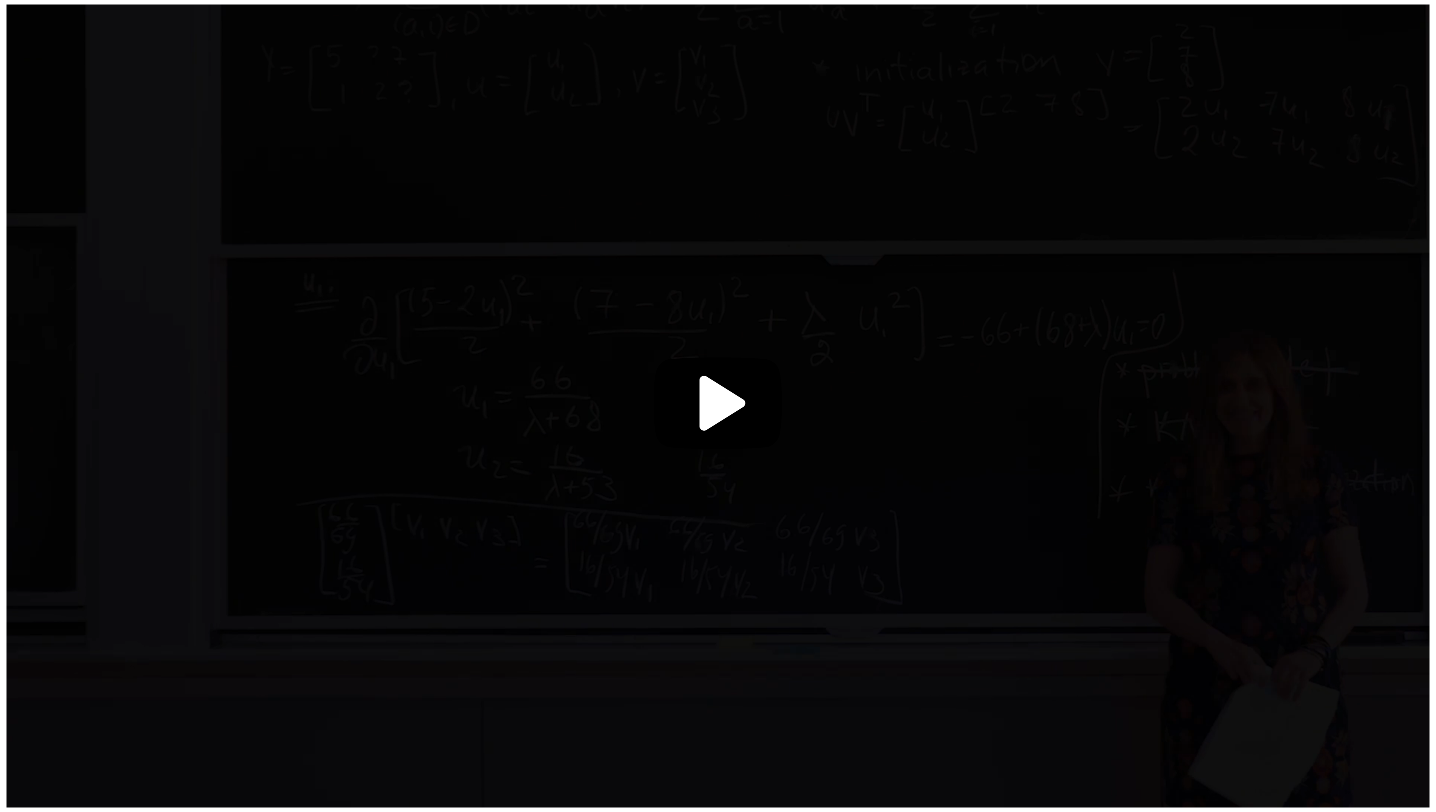
6. Alternating Minimization
Alternating Minimization



14:17 / 14:17 Speed 1.50x

But exactly the same idea will translate to rank 2, 3, and k.

And you will try it in your exercises.

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.

Video
Download video file

Transcripts
Download SubRip (.srt) file

Download Text (.txt) file

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

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

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.

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

- ( ) $\frac{\left(Y_{ai} - u^{(a)}v^{(i)}\right)^2}{2} + \frac{\lambda}{2}\left\|u^{(a)}\right\|^2$
- (•) $\sum_{(a,i)\in D} \frac{\left(Y_{ai} - u^{(a)}v^{(i)}\right)^2}{2} + \frac{\lambda}{2}\left\|u^{(a)}\right\|^2$ ✔
- ( ) $\sum_{(a,i)\in D} \frac{\left(Y_{ai} - u^{(a)}v^{(i)}\right)^2}{2}$
- ( ) $\sum_{(a,i)\in D} \frac{\left(Y_{ai} - u^{(a)}v^{(i)}\right)^2}{2} + \frac{\lambda}{2}\left\|v^{(i)}\right\|^2$

Solution:

Regarding terms related to only V as constants, J becomes equivalent to $\sum_{(a,i)\in D} \frac{\left(Y_{ai} - u^{(a)}v^{(i)}\right)^2}{2} + \frac{\lambda}{2}\left\|u^{(a)}\right\|^2$.

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

Discussion

Hide Discussion

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

Add a Post

Show all posts ▼ by recent activity ▼

[STAFF] How does the professor reduce the expression down to -66 + (68 + Lambda)u1?
Is there any additional material you can provide to see how she reduced this expression? It has been a while since I took Linear Algebra, Matrix computing and multivariate ca... 1

Grateful for blackboard lectures
Maybe it's just me, but I understand the subject better when taught at a blackboard by a human instead of a voice accompanied by premade slides with no focus pointer. 5

I hope there are more lectures by Prof Barzilay :)
I love the lectures by Prof Barzilay. She clearly brings out even the most complex concepts 1

Fixing V and Finding U
Hi, Any hint on the number of iterations? I have completed 2 iterations and getting a complex term with respect to lambda. Is there any easy step that I am missing? 3

[Staff] Alternating Minimization Concept Question
We work in the general k rank case, I can't see why the regularization part is not a sum of the noms of all the lines in U why is it just the (a) line ? 2

Understanding Alternating Minimization
Hi, In the example as show in the lecture we first randomly select V and use it to get U1 and U2. Now when we use this U1 and U2 to calculate V1, V2, V3 should we take terms... 1

Alternating Least Squares 2

How do we call Y and X ? (Terminologies in recommender system) 3

Learn About Verified Certificates