6. Alternating Minimization

Bookmark this page

Exercises due Oct 14, 2020 05:29 IST *Past Due*

Alternating Minimization

Start of transcript. Skip to the end.

OK.

So now we will take this objective

and rewrite it in terms of our u and v, which are just vectors.

So let's start.

And I will keep the formula above so that you

can see the connection.

So let me start first writing that you

0:00 / 0:00 1.25x

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

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

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 ...

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

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

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 points possible (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:

Answer: 20/(20+lambda)

The second element of U is:

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 0 of 3 attempts

Answers are displayed within the problem

Discussion

Hide Discussion

Topic: Unit 2 Nonlinear Classification, Linear regression, Collaborative Filtering (2

Add a Post

Show all posts | by recent activity

- How to understand X_{ai} — 2
  The professor wrote X_{ai} = U_aV_i, is it a dot product or just a scalar product? I thought X_{ai} represents something like {U_a, V_i}.
- Don't understand what happens in the video from 1:41 to 1:59 — 4
  It may be trivial or I may just be tired but I cannot find exactly how we go from "sum(a,i) of Xai" to "sum(a=1 to n) of ua^2 + sum(i=1 t...
- Splitting the regularisation term — 1
  Hi all, I think I'm missing a trick here - when X was split into U*V, how did we get from the regularisation term being lambda times th...
- Fixing V and finding U: it says express your terms lamda but lamda not permitted in answer — 4
  Invalid Input: \'lamda\' not permitted in answer as a variable
- [STAFF] Why are we using the development set to find the best initialization? — 2
  Since we are trying to find the global minimum of our objective function. Mine intuition is to compute the objective function and cho...
- initialization of vector V — 2
  what is the process of taking initial value of vector V. As the professor stated, wrong initialization may lead to very bad results. So is t...
- justifying the alternating minimization procedure — 2
  I understand the how the procedure of alternating minimization goes, but I don't get the intuition behind it. Are there any resources ...
- Not really important to class but... — 6
  Is the teacher not chuckling to herself at 11:27 and 11:33 when she is just repeating the number 69 a few times? At first I thought I m...
- [Staff] Why use an iterative method if we could use the Hessian Matrix, for example? — 3
  I think I am mixing concepts here, but wouldn't it be possible to find the global minimum using the Hessian Matrix of this problem? S...
- formula written on black board at time 01:24 is missing a divided by 2 — 1
  formula written on black board at time 01:24 is missing a divided by 2
- Is the UV factorization technique the same as SVD? — 3
  Is the UV factorization technique mentioned in the the lecture same as SVD (Singular value decomposition) ?
- [Staff] Derivative calculation — 2
  At 8:35 in video, in transcript where it says We will take the derivative of this expression with respect to u1. How is derivative of u1 c...
- Understanding the meaning of rank 2 matrix in the context of movie recommenders — 3
  Community TA

< Previous | Next >

edX

About