

Machine Learning with Python-From Linear Models to Deep Learning

Discussion Course **Progress Resources** Dates

A Course / Unit 2. Nonlinear Classification, Linear regression, ... / Lecture 7. Red

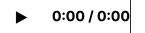
< Previous

6. Alternating Minimization

 \square Bookmark this page

Exercises due Mar 8, 2023 08:59 -03 Completed

Alternating Minimization



▶ 1.0x

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

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

In order to break a big optimization problem into smaller pieces that we know how to so the best V for that U. But a subtle and important point is that even if V^* is best for U, the for V^* might not be the original U! It's like how we might be some lonely person's best fare not our best friend. In light of this, we repeat, like this: we fix U and solve for V, therefrom the previous step and solve for U, and repeat this alternate process until we find the example of iterative entirization, where we grandly take steps in a good direction, but

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



Submit

You have used 1 of 3 attempts

Fixing V and Finding U

0/2 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 rating encourage U, V to be 'simple' in the sense that their entries are small; we quantify the sencouragement by a parameter λ , so that our optimal solution (U, V) depends on λ . (**BEC** encouragement resolves cases that are strictly ambiguous (for example, in

$$\begin{bmatrix} ? & ? & 16 \\ 2 & 3 & ? \end{bmatrix}$$

that 16 could be a product of 4×4 or a product $(0.00004) \times (400000)$; we regard the former unless special domain knowledge suggests otherwise). Moreover, this encouragement and V entries toward zero, even in non-ambiguous cases. It's like a TA tells you they entries toward zero, even in non-ambiguous cases. It's like a TA tells you they entries toward zero, even in non-ambiguous cases. It's like a TA tells you they entries the special product of the data perfective that the tell of the

Assume we start by fixing V to initial values of $[4, 2, 1]^T$. Find the optimal 2×1 vector U your answer in terms of λ).

The first element of *U* is:

weeks):Lecture 7.

Previous

Δ

иехт

Show all posts

- How can we model temporal effect/dynamics in collaborative filtering along with what was taged. How can we model temporal effect/dynamics in collaborative filtering along with what was taught? for e.g. r
- ? Is there a missed "/2" in loss function in video at 1:21?
 Just noticed that while writing the 1st part of loss function (at 1:21) Prof. Barzilay missed the dividing by 2 in
- why not explaining the next steps in calculating the v's?

 Its a pitty that the next steps are not explained. How do we continue now for calculating the v's knowing the
- ? I don't understand how we got 66/ (68 + lambda) from -66 + 68 + lambda*u_1 = 0
 what is wrong in the train of thought that if -66 + 68 + lambda*u_1 = 0 then lambda*u_1 = -2 hence U_1 = -2/
- How does X_ai^2 = U_a^2 + V_i^2?
 If X=UV how does X_ai^2 = U_a^2 + V_i^2? Shouldn't X_ai^2 = (U_a*V_i)^2?
- Matrix rank
 Hello, I have a question, maybe someone can help. In the video the teacher assumes that the range is equal
- Sidenote placed ambiguously
 I understand that this is my own fault, but sidenote is definitely placed in a way to confuse two matrices: the
- Question about partial derivatives and critical points
 I have a bit of a naive question about the reason behind equating the partial derivative of the objective funct
- © Clarity on the last question Error Message Invalid Input: \'lamda\', \'u_1\' not permitted in an So, I followed the lecture and calculated the appropriate derivatives. My derivatives are in the term of lamda
- Meaning of the last part of the sidenote
 Is the last part of the sidenote talking about how not to choose initial values for one of the vectors or is it's s
- Need help on last question
 Hi, I don't understand how the professor got the solution at 9:05 in the video. Can anybody explain it to me?
- ? What are typical real world values for lambda and k?
 I am wondering what are typical real world values of lambda and k for an online store recommendation engin
- Reviewing what the first part of the objective function represents
 Hi, I've successfully answered the final question by imitating the professor's work. My question, as a coder w

Legal

Terms of Service & Honor Code

Privacy Policy

Accessibility Policy

Trademark Policy

Sitemap

Cookie Policy

Do Not Sell My Personal Information

Connect

Blog

Contact Us

Help Center

Security

Media Kit















© 2023 edX LLC. All rights reserved.

深圳市恒宇博科技有限公司 粤ICP备17044299号-2