Week 9

SUBFORUMS

All

Assignment: Anomaly Detection and Recommender Systems

Week 9

ex8 tutorial for cofiCostFunc()

Tom Mosher Mentor Week 9 · 3 years ago · Edited

Vectorized tutorial for cost and gradients with regularization

Definitions:

R: a matrix of observations (binary values). Dimensions are (movies x users)

Y: a matrix of movie ratings: Dimensions are (movies x users)

X: a matrix of movie features (0 to 5): Dimensions are (movies x features)

Theta: a matrix of feature weights: Dimensions are (users x features)

- Compute the predicted movie ratings for all users using the product of X and Theta. A transposition may be needed.

Dimensions of the result should be (movies x users).

- Compute the movie rating error by subtracting Y from the predicted ratings.

- Compute the "error_factor" my multiplying the movie rating error by the R matrix. The error factor will be 0 for movies that users have not rated. Use the type of multiplication by R (vector or element-wise) so the size of the error factor matrix remains unchanged (movies x users).

(**Note**: there is a quirk in the submit grader's test case that requires you to use the R matrix to ignore movies that have had no ratings).

Calculate the cost:

- Using the formula on Page 9 of ex8.pdf, compute the unregularized cost as a scaled sum of the squares of all of the terms in error_factor. The result should be a scalar.

- Test your code using ex8_cofi.m and the additional test cases. You should get a passing grade for this portion from the submit script.

Calculate the gradients (ref: the formulas on Page 10 of ex8,pdf):

- The X gradient is the product of the error factor and the Theta matrix. The sum is computed automatically by the vector multiplication. Dimensions are (movies x features)

- The Theta gradient is the product of the error factor and the X matrix. A transposition may be needed. The sum is computed automatically by the vector multiplication. Dimensions are (users x features)

- Test your code, then submit this portion.

Calculate the regularized cost:

- Using the formula on the top of Page 13 of ex8.pdf, compute the regularization term as the scaled sum of the squares of all terms in Theta and X. The result should be a scalar. Note that for Recommender Systems there are no bias terms, so regularization should include all columns of X and Theta.

- Add the regularized and un-regularized cost terms.

- Test your code, then submit this portion.

Calculate the gradient regularization terms (ref: the formulas in the middle of Page 13 of ex8.pdf)

- The X gradient regularization is the X matrix scaled by lambda.

- The Theta gradient regularization is the Theta matrix scaled by lambda.

- Add the regularization terms to their unregularized values.

- Test your code, then submit this portion.

Done.

==========

keywords: tutorial coficostfunc

55 Upvotes Reply Follow this discussion

This thread is closed. You cannot add any more responses.

Earliest **Top** **Most Recent**

JM Jon Machtynger · 2 years ago

Excellent step by step approach to getting this done. Thanks!

12 Upvotes Reply

Nabheet Sandhu · 2 years ago

Hi, I realize that other people may have found this a lot easier than me - but I was really lost on why the grader was not grading my assignment. I wasted at least 30mins while scratching my head until I found this comment by Tom Mosher https://www.coursera.org/threads/92NKXCLBEeWM2iIAC0KUpw/comments/Hp1gOUPlEeWSTg7F_5iDhQ:

"You must include the Y matrix when you multiply by R, because the submit grader's test case has some Y values that are non-zero even though the R matrix says it has not been rated. It's a defect in the submit grader's test case."

Posting it here so other people don't have to dig for it. It clearly explained why my solution was not being graded. And grading was successful after I fixed my issue. I guess posting the fix might violate the rules of engagement here. :-)

And maybe, we could add it to the original post.

16 Upvotes Reply

M Melissa · 2 years ago

Thank you so much, Tom!

2 Upvotes Reply

Haniball · 3 years ago

Thanks Tom! I'm a bit late but just finished the last assignment ex8 and wanted to say thanks again. I was stuck once a[illegible]rt but just reading a bit on this part in your tutorial made a question on the forum unneccessary and saved time :-)

2 Upvotes Hide 1 Reply

Tom Mosher Mentor · 3 years ago

Good news.

0 Upvotes

DI daniele iovinelli · 3 years ago

I tom ! How can I calculate epsilon in the selectThreshold.m?? There is a formula or what??

thank you

daniele

0 Upvotes Hide 7 Replies

Tom Mosher Mentor · 3 years ago · Edited

See line 14 in the selectThreshold.m script. The value 'epsilon' is defined for you. You add your code below that, in the "YOUR CODE HERE" section.

0 Upvotes

DI daniele iovinelli · 3 years ago

Ok! Thanks I have put the code with formulas in the provided pdf document , but when I launch the program it says this:

Program paused. Press enter to continue.

warning: division by zero

Best epsilon found using cross-validation: ***8.990853e-02***

Best F1 on Cross Validation Set: 0.057143

(you should see a value epsilon of about 8.99e-05).

What could be wrong in your opinion?

Thanks Tom!

0 Upvotes

Tom Mosher Mentor · 3 years ago

Why are these posts in the thread for the cofiCostFunc() tutorial?

0 Upvotes

DI

daniele iovinelli · 3 years ago

Sorry! But I didn't see a "thread" tutorial. Where can I post my questions about it?

Thanks

0 Upvotes

DI

daniele iovinelli · 3 years ago

However for this part I have done , I don't need help anymore, it is all OK!

0 Upvotes

DI

daniele iovinelli · 3 years ago

Thank you anyway

0 Upvotes

Tom Mosher Mentor · 3 years ago

If you've got a question on a new topic, start a new thread in the area for that lesson.

1 Upvote

daniele iovinelli · 3 years ago

DI

how can I calculate epsilon in selectThreshold?

thanks

0 Upvotes Reply

JL John Le · 3 years ago

Tom, you are really a lifesaver! Thank you!

0 Upvotes Reply

Anton Piatygin · 3 years ago

Just wanted to say " Thank you!". Amazing help along the whole course. However, I feel that sometimes I simply follow your instructions, it works, but I don't fully put my thinking into it. At the same time, it would take me multiple time more time if I were to do it without your help.

Thank you again!

0 Upvotes Hide 1 Reply

Tom Mosher Mentor · 3 years ago

I think the tutorials work best as a learning tool after students have made an effort to solve the problem for themselves. Often the issue is lack of familiarity with the programming language or in how to write vectorized code, rather than a lack of understanding of the ML techniques.

12 Upvotes

YS Yan Stein · 3 years ago

Great tutorial! Thanks Tom!

0 Upvotes Reply

AP Antoine Tuan PHAM · 3 years ago

Nice tutorial

0 Upvotes Reply

Ian Ferreira · 3 years ago

Tom you rock!

4 Upvotes Reply

AB Alejandro Barredo · 3 years ago · Edited

Hello Tom,

I have done all the code and it is running properly, the analytical gradient and the numerical one are equals, and the predictions at the end of the script are also correct. But I am still not getting the grade when submiting. Could you give me a test case?

Thank you in advance.

0 Upvotes Hide 1 Reply

AB Alejandro Barredo · 3 years ago

Said nothing.

I had a line commented from a test done before. All was correct.

0 Upvotes

Nuno Henriques · 3 years ago

Thanks, Tom! **You just cleared** the somehow confusing explanation in the ex8.pdf, I was trapped in half way between some vectorized and some single for-loop. I went for **full**, **clean** and **efficient vectorized solution** and it's done.

Cheers :-)

1 Upvote Hide 2 Replies

Tom Mosher Mentor · 3 years ago

Nice work!

0 Upvotes

LB Laurent Borderie · 3 years ago

I second that. I had a hard time figuring out how to use the references to

idx = find(R(i, :)==1) to etc in page 12, in a vectorized implementation, since I had the examples of multiple ratings from the courses and quizzes.

You made this clearer, then I just had to go for pure vectors and matrices. Thanks!

1 Upvote

AG Azadeh Gholami · 3 years ago

Hello Tom, I couldnt download the ex8 yet. The link looks incorrect. What should I do? If I cant pass this assignment I cant pass the course and I am very worry about that. :((

0 Upvotes Hide 1 Reply

Tom Mosher Mentor · 3 years ago · Edited

I am able to download the assignment from this link:

https://s3.amazonaws.com/spark-public/ml/exercises/on-demand/machine-learning-ex8.zip

0 Upvotes

GN Guy Nicholson · 3 years ago

Thanks Tom.

1 Upvote Reply