coursera

Discussion Forums

Get help and discuss course material with the community.

THIS WEEK'S FORUM

Week 6

Discuss and ask questions about Week 6.

19 threads · Last post 5 hours ago

Go to forum

Forums | **All Threads**

[Search   🔍]

← All Course Discussions

---

ex1 Tutorial and test case for gradientDescent() ⌄

Tom Mosher   Mentor   · 2 years ago · Edited

Here is a tutorial on implementing gradientDescent() and gradientDescentMulti().

I use the vectorized method, hopefully you're comfortable with vector math. Using this method means you don't have to fuss with array indices, and your solution will automatically work for any number of features or training examples.

What follows is a vectorized implementation of the gradient descent equation on the bottom of Page 5 in ex1.pdf.

Reminder that 'm' is the number of training examples (the rows of X), and 'n' is the number of features (the columns of X). 'n' is also the size of the theta vector (n x 1).

Perform all of these steps within the provided for-loop from 1 to the number of iterations. Note that the code template provides you this for-loop - you only have to complete the body of the for-loop. The steps below go immediately below where the script template says "======= YOUR CODE HERE ======".

1 - The hypothesis is a vector, formed by multiplying the X matrix and the theta vector. X has size (m x n), and theta is (n x 1), so the product is (m x 1). That's good, because it's the same size as 'y'. Call this hypothesis vector 'h'.

2 - The "errors vector" is the difference between the 'h' vector and the 'y' vector.

3 - The change in theta (the "gradient") is the sum of the product of X and the "errors vector", scaled by alpha and 1/m. Since X is (m x n), and the error vector is (m x 1), and the result you want is the same size as theta (which is (n x 1), you need to transpose X before you can multiply it by the error vector.

The vector multiplication automatically includes calculating the sum of the products.

When you're scaling by alpha and 1/m, be sure you use enough sets of parenthesis to get the factors correct.

4 - Subtract this "change in theta" from the original value of theta. A line of code like this will do it:

```
1    theta = theta - theta_change;
```

That's it. Since you're never indexing by m or n, this solution works identically for both gradientDescent() and gradientDescentMulti().

There is a test case below (or use this link):

https://www.coursera.org/learn/machine-learning/discussions/-m2ng_KQEeSUBCIAC9QURQ/replies/jCkbzfQsEeSkXCIAC4tJTg

============================

Note: Replies to this thread tend to get lost due to a glitch in the forum. Please use the link below to post new questions.

https://www.coursera.org/learn/machine-learning/discussions/uCXYyH6wEeWU3RJpSD4VQQ

The thread you are reading is closed to new comments.

============================

Keywords: ex1 tutorial gradientdescent gradientdescentmulti gradient

🖒 71 Upvote   ·   Follow 73   ·   Reply to Tom Mosher

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

Earliest   Top   Most Recent

RJ   Rajesh jaiswal · a year ago   ⌄

!! Submission failed: unexpected error: urlread: Peer certificate cannot be authenticated with given CA certificates

🖒 1 Upvote   ·   Hide 1 Reply

    Tom Mosher Mentor · a year ago   ⌄

    See my reply to your previous post.

    🖒 1 Upvote

jim · a year ago   ⌄

Very nice tutorial!

I tried both with the vectorized implementation and use for loops without vectors multiplication. Both works well.

🖒 1 Upvote   ·   Reply

A   Acas · a year ago   ⌄

hi Tom, I didn't get if it's ok to look at these tutorials before submitting or it's a violation of the Honor Code .. because it feels quite like cheating :)

🖒 0 Upvote   ·   Hide 5 Replies

    Tom Mosher Mentor · a year ago   ⌄

You do not have to use the tutorials, if you can write the code without any other assistance. If they make you uncomfortable, then please ignore them.

Using the tutorials does not violate the Honor Code. They exist for students who need a little help understanding how to implement the equations in vectorized form, using the Octave/MATLAB script language.

👍 2 Upvote

A   Acas · a year ago   ⌄

thank you Tom

👍 0 Upvote

A   Acas · a year ago   ⌄

what kept me going mad for two days was that in every slide the vectorialization the prediction is theta' * x; and we have learned that in matrix A*B is different from B*A.. now ...you know what the answer is for the computeCost.m and I cannot write down here the solution not to violate the Honor Code but i solved the exercise only after seeing this tutorial ... the solution was completely unexpected for me, suddenly I had to do something i tought it was wrong to do ...

I dunno if i made myself clear, if you want i can contact you in private to explain better without violating the honour code

👍 2 Upvote

Tom Mosher Mentor · a year ago   ⌄

I think your comment is clear.

"h = theta' * x" is true only when theta and x are both column vectors.

"h = X * theta" is true when X is the whole matrix of training examples.

This can be seen clearly by looking at the dimensions of each variable.

This is also discussed in the "Tips from mentors" document in the Week 1 materials.

👍 4 Upvote

A   Acas · a year ago   ⌄

thanks again, Tom

👍 0 Upvote

RJ   Rajesh jaiswal · a year ago   ⌄

what is theta_change? how to calculate it

👍 0 Upvote   ·   Hide 2 Replies

RJ   Rajesh jaiswal · a year ago   ⌄

@Tom Mosher

👍 0 Upvote

Tom Mosher Mentor · a year ago   ⌄

It's what you compute in step 3 of the tutorial.

👍 0 Upvote

RJ **Rajesh jaiswal** · a year ago ⌄

what is theta change ? how to find

👍 0 Upvote · Reply

A **Athresh** · a year ago · Edited ⌄

Hi Tom,

This is the result I'm getting for the test case you provided. However, I'm not getting any credit for the gradient descent() function.

```
1   >> [theta J_hist] = gradientDescent([1 5; 1 2; 1 4; 1 5],[1 6 4 2]',[0 0]'
        ,0.01,1000);
2   >> theta
3
4   theta =
5
6       5.2148
7      -0.5733
8
9   >> J_hist(1)
10
11  ans =
12
13      5.9794
14
15  >> J_hist(1000)
16
17  ans =
18
19      0.8543
```

Thank you for your help.

👍 0 Upvote · Hide 3 Replies

Tom Mosher Mentor · a year ago ⌄

That is very curious. The results look correct to me.

👍 0 Upvote

A **Athresh** · a year ago ⌄

Do you have any idea as to what could be causing this? I've been stuck at this for days now.

👍 0 Upvote

Tom Mosher Mentor · a year ago ⌄

Are you still stuck on this issue?

👍 0 Upvote

BL **Brian Lewis** · a year ago ⌄

So I am not sure exactly where my problem is.. I hope someone here can help me. My theta values for the gradient descent functions are -3.878138 1.19126, are they correct? My cost function passed the online grader but my gradient descent will not. First, i find hypothesis by multiplying X by theta, then the error by subtracting y from the hypothesis. Next, the variable theta_change is equal to transpose X multiplied by error which is then multiple by learning rate times 1 over m. Finally theta is equal to theta minus theta_change. I did have the 'sum' operator in the theta_change variable, however I read earlier in this thread that that is already calculated due to the multiplication of matrices, and that did seem to help my debugging process significantly. However, I am still unsure if I am calculating the correct theta for global minimum of cost. Any help, or reasons why I am not passing the online grader would be greatly appreciated.

Thank you!

👍 0 Upvote   ·   Hide 4 Replies

BL   Brian Lewis · a year ago   ⌄

I forgot to ask, in the implementation of the gradient descent should
theta 1 and theta 2 be referenced some how? I just use the theta
vector, but I have seen some reference to a 'theta(1)' and 'theta(2)'.

👍 0 Upvote

Tom Mosher Mentor · a year ago   ⌄

My values:

```
1   Running Gradient Descent ...
2   ans =  32.073
3   Theta found by gradient descent: -3.630291 1.166362
4   For population = 35,000, we predict a profit of 4519.767868
5   For population = 70,000, we predict a profit of 45342.450129
```

If you're using the vectorized method given in the Tutorial, then it's
just "theta", all one vector.

Check if you're using enough sets of parenthesis. Your method
sounds correct, but the details are important. I use (h - y) for the
error.

👍 0 Upvote

Tom Mosher Mentor · a year ago   ⌄

Keep in mind also that the submit grader doesn't use this test case at
all - it has a different test case, with a different size of data set. Your
code should work with any data set.

There are additional test cases you can use also - find them in the
General Discussion forum area.

👍 0 Upvote

BL   Brian Lewis · a year ago   ⌄

I figured out my problem. I was multiplying the learning rate alpha
by 1/m instead of 1/2m. I got the same results that you posted, and I
submitted it to the grader and was successful.

Thank you for your help Tom.

👍 0 Upvote

A   Athresh · a year ago   ⌄

Hi Tom,

My program passes your test case but I'm not getting any credit for the
gradient descent function. Could you please tell me what I might be doing
wrong? Thanks!

👍 0 Upvote   ·   Hide 3 Replies

Tom Mosher Mentor · a year ago   ⌄

Please post your results from running the test case.

👍 0 Upvote

A   Athresh · a year ago   ⌄

Sorry for the late reply.

[theta J_hist] = gradientDescent([1 5; 1 2; 1 4; 1 5],[1 6 4 2]',[0
0]',0.01,1000);

>> theta

theta =

5.2148

-0.5733

>> J_hist(1)

ans =

5.9794

>> J_hist(1000)

ans =

0.8543

👍 0 Upvote

Tom Mosher Mentor · a year ago · Edited

That's the correct result for that test case.

Please post a screen capture of your workspace after you run the submit script. I would like to see any error messages or warnings that might appear.

👍 0 Upvote

Iago Maciel de Souza · a year ago

Hey Tom!

My gradient descent function was not working properly because I was using a sum to calculate the theta_change and had some syntax errors. Now I've already corrected these mistakes but still not getting the right solution. I got theta(1) = 87.58 and theta(2) = 979.93... When I take out the semi-solon of the cost function, I can see that it is increasing (which should not happen)... Any clues? I cannot paste this bit of my code here, can I?

Thank you very much for all the help!

👍 0 Upvote · Hide 7 Replies

Tom Mosher Mentor · a year ago

Sorry, you can't post your code on the Forum. The course Honor Code says so.

In implementing the formulas, perhaps you have a sign error, or an addition instead of a subtraction.

👍 0 Upvote

Iago Maciel de Souza · a year ago

Okay... I'm getting the theta_change values negative, I think there is the error as the theta = theta - theta_change and thus it will increase the theta value as - * - = +. This is weird as I am doing exactly as you corrected a classmate earlier...

```
1    theta_change = alpha*(1/m)*X'*e
```

Another problem is that even though i'm getting wrong values for the regression, when I plot the regression it does not shows the line correctly through the data samples. The data samples become several points along the x = 0 line..
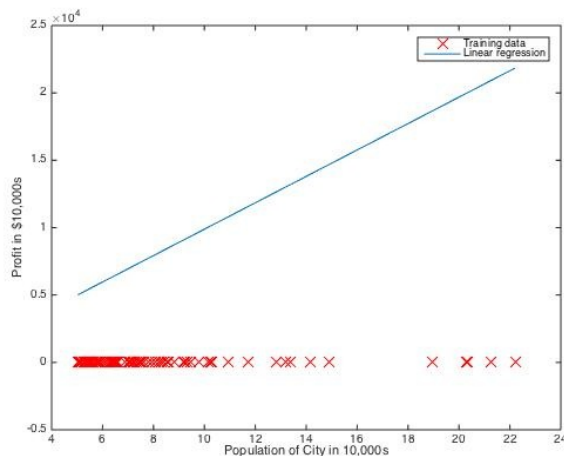
I've tried as on the ex1.m script and adding again the plotData(x,y) in the new figure (using hold on and off).. This is really weird.

👍 0 Upvote

Tom Mosher Mentor · a year ago · Edited   ⌄

Buggy code often acts really weird, until all the bugs are gone.

The problem with your plot is that the 'y' values are incorrect. Those should have been plotted at line 47 from the original data, then "hold on" at line 73 keeps that data intact while the hypothesis values are plotted (at line 74).

Messing around with "hold on" and "hold off" can upset this sequence. So can re-assigning the 'y' values.

👍 0 Upvote

Iago Maciel de Souza · a year ago   ⌄

my code is exactly as you said above. Plotting the original data at 47 and the linear regression at line 74...

👍 0 Upvote

Iago Maciel de Souza · a year ago · Edited   ⌄

Well, I've managed to correct the gd function and now I've got these results:

theta found by gradient descent: -3.630291 1.166362

For population = 35,000, we predict a profit of 4519.767868

For population = 70,000, we predict a profit of 45342.450129.

I saw on a older thread that these are the correct values but the grader is not accepting my code. I don't get any "submit error" messages, just it does not count my gd bit. Any thoughts in how can I debug it?

👍 0 Upvote

Iago Maciel de Souza · a year ago ⌄

Managed to debug it! I was hard-coding num_iters. Cheers tom

👍 0 Upvote

Tom Mosher Mentor · a year ago ⌄

Nice catch.

👍 0 Upvote

Shixiao Cui · a year ago ⌄

Very useful！

👍 0 Upvote · Reply

Will Brocklebank · a year ago ⌄

I am banging my head against the wall with this: after my gradientdecent runs and also on submission I am getting this:

```
1   !! Submission failed: unexpected error: A(I) = X: X must have the same size
        as I!! Please try again later.
```

Anyone able to help? I'm running out of time!

👍 0 Upvote · Hide 5 Replies

Tom Mosher Mentor · a year ago ⌄

The submit grader uses a different test case than the one in the exercise script. The test case has a different size of data set. Your code needs to work with any size of data set.

👍 0 Upvote

Will Brocklebank · a year ago ⌄

I get that. I am not specifying any particular data set in the gradientDescent script (although 'ex1data1.txt' is loaded in the plotData module which runs before computeCost and gradientDescent).

And that's the thing: I don't just get this on submission but also after my theta_change value is output I then get

```
1   gradientDescent: A(I) = X: X must have the same size as I
```

But I have modified nothing else in gradientDescent.m after the last line of code to make theta = theta - theta_change!

👍 0 Upvote

Enrique Maya Visuet · a year ago ⌄

Review your transpose or if you are including a sum(Gradient) this will affect your result.

👍 0 Upvote

Tom Mosher Mentor · a year ago · Edited ⌄

plotData() **should not** load any files from disc itself. It should plot the 'x' and 'y' data that is passed to it as parameters. ex1.m loads the data file - your function should not.

Read my previous post carefully. Your code is mis-handling the data that is used in the submit script - typically this is because something in your code only works correctly with the one data set (of a specific size) used in the exercise script.

Perhaps try comparing your code with the method described in the tutorial for this exercise. Tutorials are listed in the General Discussion forum area.

👍 0 Upvote

Will Brocklebank · a year ago   ⌄

Thanks very much for your help. I have removed the incorrect load line from plotData & after going back to the tutorial I had got confused and omitted the sum function because I had read it wasn't necessary... Phew!

👍 0 Upvote

IG   Ingolf Gehrhardt · a year ago   ⌄

Thanks a lot. This tutorial in combination of doing the first iteration on the paper, helped me to find my mistake. Great work!

👍 0 Upvote · Reply

Enrique Maya Visuet · a year ago   ⌄

Hello Tom,

I found this after running gradientDescent in ex1.m:

Plotting Data ...

Program paused. Press enter to continue.

Running Gradient Descent ...

ans = 32.073

Theta found by gradient descent: 0.720882 0.720882

warning: legend: ignoring extra labels

warning: called from

legend at line 458 column 9

ex1 at line 77 column 1

For population = 35,000, we predict a profit of 32439.671676

For population = 70,000, we predict a profit of 57670.527425

Program paused. Press enter to continue.

Visualizing J(theta_0, theta_1) ...

👍 4 Upvote · Reply

Yeasir Mohammad Akib · a year ago   ⌄

Hlw ,Tom after following the instructions I have got the following results :

```
1   [theta J_hist] = gradientDescent([1 5; 1 2; 1 4; 1 5],[1 6 4 2]',[0 0]',0
       .01,1000);
2   Error using error
3   Too many output arguments.
4
5   Error in gradientDescent (line 21)
6
```

👍 0 Upvote · Hide 6 Replies

Tom Mosher Mentor · a year ago · Edited   ⌄

Look at line 21 of your gradientDescent.m file, and see what the problem is.

👍 0 Upvote

Yeasir Mohammad Akib · a year ago   ⌄

According to the procedure theta_change is the multiplication of alpha ,reciprocal of m and sum of the error*X' .

then the following step is :

```
1   theta = theta - theta_change;
```

So what is the problem with the code ?

👍 0 Upvote

Tom Mosher Mentor · a year ago · Edited   ⌄

Please copy and paste your code from the surrounding lines of code here.

👍 0 Upvote

Tom Mosher Mentor · a year ago   ⌄

Other questions:

Have you created any new function definitions?

Are you using the gradientDescent.m script template that was provided?

👍 0 Upvote

Yeasir Mohammad Akib · a year ago · Edited   ⌄

My 21st code line is :

```
1   theta_change = (alpha)*(1/m)*(sum((error)*(X)'));
```

No, I don't create any new function definition .

👍 0 Upvote

Tom Mosher Mentor · a year ago · Edited   ⌄

Aha. That line of code has several syntax and logic errors.

Remove all the extra sets of parenthesis that enclose individual variable names.

(alpha) should be just alpha

(X) should be just X

Also remove the extra parenthesis that enclose the sum() function.

Then recognize that in a vector multiplication, the sum() function is not needed, because the vector multiplication includes the sum automatically.

The gradient computation would then be X' * error, because the size of X' is (n x m), and the error vector is (m x 1). The sum of the products is automatically computed, and this gives you a (n x 1) result. That's exactly what you want (it is the same size as theta).

(1/m) is good because the parenthesis surround an operator.

I believe the issue is that the sets of parenthesis around the variable names are being interpreted as an invalid function call.

👍 5 Upvote

Gwendolyn Sandy · a year ago ⌄

I keeping getting the error Gradient descents requires more inputs.

What does this mean?

👍 0 Upvote · Hide 1 Reply

　　Tom Mosher　Mentor　· a year ago ⌄

　　Please paste in a screen capture of exactly what you are seeing. Include the command you entered and the entire result.

　　👍 0 Upvote

‹ 1 2 3 4 ›

DESCRIPTION

Welcome to the course discussion forums! Ask questions, debate ideas, and find classmates who share your goals. Browse popular threads below or other forums in the sidebar.

MODERATORS

YM　　　　　　KK　　　VA

AT　LK　　　NK　SL　MS　　FL　　　LT　　　MR　CU

MM　　　　　　JN　S　　EM　　T　AN　QG

Learn more about becoming a Mentor

Forum guidelines　　　　　　　　　　　　　　　　　　　　　　　　›