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 for computeCost

Tom Mosher Mentor · 2 years ago · Edited

This is a step-by-step tutorial for how to complete the computeCost() function portion of ex1. You will still have to do some thinking, because I'll describe the implementation, but you have to turn it into Octave script commands.

All the programming exercises in this course follow the same procedure; you are provided a starter code template for a function that you need to complete. You never have to start a new script file from scratch.

This is a vectorized implementation. You're only going to write a few simple lines of code.

With a text editor (NOT a word processor), open up the computeCost.m file. Scroll down until you find the "====== YOUR CODE HERE =====" section. Below this section is where you're going to add your lines of code. Just skip over the lines that start with the '%' sign - those are instructive comments.

We'll write these three lines of code by inspecting the equation on Page 5 of ex1.pdf

The first line of code will compute a vector 'h' containing all of the hypothesis values - one for each training example (i.e. for each row of X).

The hypothesis (also called the prediction) is simply the product of X and theta. So your first line of code is...

```
1   h = {multiply X and theta, in the proper order that the inner dimensions
        match}
```

Since X is size (m x n) and theta is size (n x 1), you arrange the order of operators so the result is size (m x 1).

The second line of code will compute the difference between the hypothesis and y - that's the error for each training example. Difference means subtract.

```
error = {the difference between h and y}
```

The third line of code will compute the square of each of those error terms (using element-wise exponentiation),

An example of using element-wise exponentiation - try this in your workspace command line so you see how it works.

```
v = [-2 3]
v_sqr = v.^2
```

So, now you should compute the squares of the error terms:

```
error_sqr = {use what you have learned}
```

Next, here's an example of how the sum function works (try this from your command line)

```
q = sum([1 2 3])
```

Now, we'll finish the last two steps all in one line of code. You need to compute the sum of the error_sqr vector, and scale the result (multiply) by 1/(2*m). That completed sum is the cost value J.

```
J = {multiply 1/(2*m) times the sum of the error_sqr vector}
```

That's it. If you run the ex1.m script, you should have the correct value for J. Then you should run one of the unit tests (available in the Forum).

Then you can run the submit script, and hopefully it will pass.

Note: Be sure that every line of code ends with a semicolon. That will suppress the output of any values to the workspace. Leaving out the semicolons will surely make the grader unhappy.

==========

This thread is closed to new comments. If you have a question, please start a new thread in the Week 2 discussion forum area.

==========

keywords: tutorial computeCost

85 Upvote · Follow 68 · Reply to Tom Mosher

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

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

JW

JINGYUAN WU · 8 months ago

hello.

could you tell me where i can find "Page 5 of ex1.pdf"

0 Upvote · Hide 1 Reply

Tom Mosher Mentor · 8 months ago

The exercise zip file "machine-learning-ex1.zip" contains the ex1.pdf file. Those are the instructions you should be using for the programming exercise.

0 Upvote

Sri Ravi Teja Kolipakula · 8 months ago

I am wrong in defining my theta I guess. can you tell me how to define theta? i allotted zero vector to vector theta.

1 Upvote · Hide 1 Reply

Tom Mosher Mentor · 8 months ago

Theta is defined for you when the ex1.m script calls your computeCost() function. You do not need to define theta.

1 Upvote

Rajesh Yerragunta · 9 months ago

Hello Tom,

I am able to get to the correct answer on cost function , but still not clearly understand the algorithm.

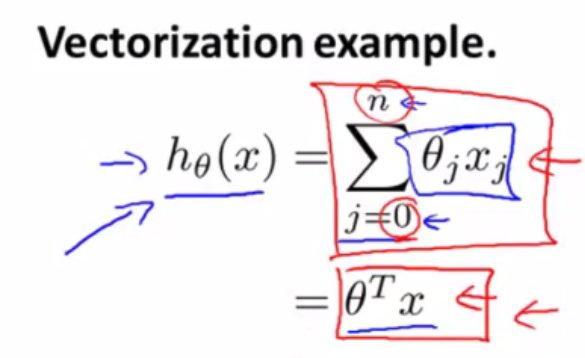
X is a matrix of m*(n+1) dimension. i.e. dimension is 97*2

theta is a (n+1) vector. i.e. 2*1

Based on the course example, prediction is calculated as (theta)' * X . And matrix dimensions does not support this multiplication.

Can you please clarify how the cost function calculation is different from the example in below image. ?

Rajesh



0 Upvote · Hide 1 Reply

Tom Mosher Mentor · 9 months ago

Read the page "Programming tips from Mentors" in the Week 2 materials.

1 Upvote

RH Rohith Hy · 9 months ago

Hello Tom,

I need some clarity about computeCost, I just wanted to clarify if the value of theta supplied to computeCost is a zero vector or not. Please help me out.

0 Upvote · Hide 3 Replies

Tom Mosher Mentor · 9 months ago

Theta is a vector of real values. It is not a vector of zeros.

0 Upvote

RH

Rohith Hy · 9 months ago

But in the ex1.m file, isnt theta initialized to a zero vector before being passed to the computeCost function?

0 Upvote

Tom Mosher Mentor · 9 months ago

That is only the initial value. Gradient descent modifies theta as it progresses.

0 Upvote

gloria · 9 months ago · Edited by moderator

Hi, mentor. I entered the codes as you wrote on this page, then I ran 'ex1', there is an error 'error: 'multiply' undefined near line 16 column 6', which referred to ''h = {multiply X and theta, in the proper order that the inner dimensions match}''

(Mentor edit: Removing image that contains code script due to honor code violation}

0 Upvote · Hide 2 Replies

Tom Mosher Mentor · 9 months ago

The tutorial describes the code you need to write. You need to translate it into the correct math operators and variable names.

0 Upvote

Tom Mosher Mentor · 9 months ago

Be sure you watch all of the Week 1 and Week 2 video lectures, including the "Octave Tutorial" videos.

0 Upvote

Jeff Davis · 9 months ago

Thanks - very helpful!

0 Upvote · Reply

CO

Christian Pérez Ortiz · 9 months ago

Thanks for showing the difference between the one training example and the vector example (h= theta' * x vs. h=X*theta). I knew I had to do the arrangement in order to get the answer, but I was so confused by why the formula taught was the one training exapmle case and not the vector one.

2 Upvote · Reply

AR

Abhilash R · 10 months ago

Thank you so much, Tom. The tutorial was of super help.

0 Upvote · Reply

Guilherme Campos Duarte · 10 months ago

Hi Tom, Where can I download this file you mentioned "computeCost.m" and "ex1.pdf"?

0 Upvote · Hide 2 Replies

Tom Mosher Mentor · 10 months ago

Be sure you watch all of the Week 1 and Week 2 video lectures before you try the programming exercise.

The programming assignment zip files are available in the 'Assignments" menu.

0 Upvote

Guilherme Campos Duarte · 10 months ago

Oh, I'm sorry, I haven't started week 2 just yet.

Thanks a lot.

1 Upvote

SC

Samuel Camps · 10 months ago

But I don't believe it, I'm using vectors. It works with the ex1 data, it works with the test case data, so logic tells me there it should work with all data sets. I'm multiplying in the correct order. If that data set should matter, unless it gets loaded in in different way.

0 Upvote · Hide 1 Reply

Tom Mosher Mentor · 10 months ago

There's nothing unusual about the submit grader's test case. It has the same orientation as the exercise script, and the additional test cases. It does have a different number of training examples, but that's to be expected - your code should solve the general case.

You can inspect the test case that the submit.m script uses. It's right there in plain text.

0 Upvote

SC

Samuel Camps · 10 months ago

I programmed the costfunction using vectors, nothing hard coded. Wrote it in one line amd also wrote it in steps like in this tutorual. In both cases it gave me the correct asnwer for J and the correct answer for the testcases. However, if I submit I always get the error "error using * Inner matrix dimensions must agree"

But they do! The order is correct and running testdata gives the right answer.....

0 Upvote · Hide 1 Reply

Tom Mosher Mentor · 10 months ago

Your code does not work correctly on the test case that the submit grader uses. Believe the error message - you are not handling the sizes of the data set correctly.

0 Upvote

Richie Thomas · 10 months ago

Hi Tom, great walk-through. My question concerns theta vs. theta' (theta-transpose) when calculating the hypothesis. Page 5 of ex1.pdf says:

```
1  h(x) = theta' * x
```

But in your tutorial, we're just meant to use theta, not theta-transpose. I would expect theta to have the same number of columns in both the PDF and the example above, so why does the pdf say we should use theta-transpose?

0 Upvote · Hide 1 Reply

Tom Mosher Mentor · 10 months ago · Edited

The PDF is written presuming you're going to compute one hypothesis value at a time, and put it inside a for-loop over all of the training examples. That's fine if you want to use the slow and inefficient (and hard to debug) iterative method.

The tutorial gives the vectorized method.

Take your choice. Either method works.

0 Upvote

CP

Cristian Popa · a year ago

Installed the patches and everything worked great, thank you! Octave 4.0.0@Win10 x64

0 Upvote · Hide 1 Reply

Tom Mosher Mentor · a year ago

Good news. Thanks for the report.

0 Upvote

Guilherme Carlos · a year ago

Thanks a lot Tom! This part helped me a lot!!

"Since X is size (m x n) and theta is size (n x 1), you arrange the order of operators so the result is size (m x 1)."

0 Upvote · Reply

JL

Jesse Linson · a year ago

Hi Tom,

Thank you for helping out with this. I'm receiving the following error message for the second line of code where I write error = {h - y}: Undefined operator '-' for input arguments of type 'cell'. Do you know what I'm doing wrong?

Thanks,

Jesse

0 Upvote · Hide 11 Replies

Tom Mosher Mentor · a year ago

Don't use curly-braces. Curly-braces are used for creating a cell array.

Use square brackets [ ] if you want to create a vector or matrix, but in this case you don't need them. h and y are already vectors, so you can subtract them directly. No other notation is needed.

1 Upvote

Tom Mosher Mentor · a year ago

Since you're getting started in the course, I recommend you read this list of tips:

Greetings,

Welcome to Machine Learning! Please read this list of tips for new ML students.

https://www.coursera.org/learn/machine-learning/discussions/v2YppY8FEeWIeBJxvl1elQ

1 Upvote

JL

Jesse Linson · a year ago · Edited

Hi Tom,

Thanks for the quick response and the list of tips! Unfortunately, that led me to the same error. To be clear, I wrote error = h - y. I tried brackets and parentheses as well. The line before gives me a result of h = [97x1 double].

0 Upvote

Tom Mosher Mentor · a year ago

What program are you using to edit your script files?

Please enter these commands and post back the results:

```
1  whos y
2  whos h
```

0 Upvote

JL

Jesse Linson · a year ago

I'm editing the scripts directly in MATLAB. This is what I received when I entered those commands:

Name Size Bytes Class Attributes

y 97x1 776 double

0 Upvote

Tom Mosher Mentor · a year ago

That seems to be the correct size. So that's good.

I seem to recall a similar issue from another student recently, but I can't find it in the Forums at the moment. It was very unusual.

Try entering these commands at your console, just as an experiment. Please post back your results:

```
1  v = [1 ; 2 ; 3]
2  w = ones(3,1)
3  v - w
```

0 Upvote

JL

Jesse Linson · a year ago

Here are my results, they look good to me.

v = [1 ; 2 ; 3]

w = ones(3,1)

v - w

v =

1

2

3

w =

1

1

1

ans =

0

1

2

0 Upvote

Tom Mosher Mentor · a year ago · Edited

Since it works fine in the console, I'd have to guess that somehow your editor is inserting a non-standard character for the '-' minus sign in the script file.

I have no idea how to fix that.

0 Upvote

JL

Jesse Linson · a year ago · Edited

Hi Tom,

Good news! I was able to advance onward with the problem after I removed the curly braces from my definition of h. When I run the code now, however, my J value does not appear in the results. I get the following:

Theta found by gradient descent: 0.000000 0.000000

For population = 35,000, we predict a profit of 0.000000

For population = 70,000, we predict a profit of 0.000000

Then, when I hit enter again the code plots a graph of different sized parabolas.

Since the code passes, is this correct?

0 Upvote

Tom Mosher Mentor · a year ago

The computeCost() function just computes the cost value.

To get the theta value, you have to work on the gradientDescent() function. There is a separate tutorial for that.

0 Upvote

Jesse Linson · a year ago

JL Thanks Tom! I greatly appreciate your help. I will get on it now.

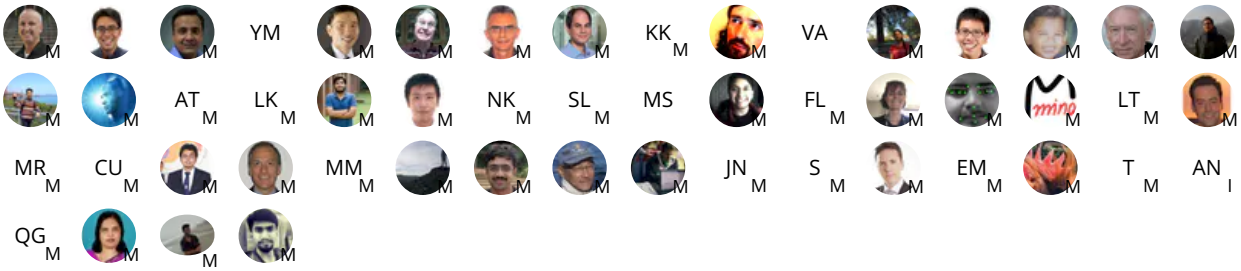
0 Upvote

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



Learn more about becoming a Mentor

Forum guidelines