



Discussion Forums

Get help and discuss course material with the community.

THIS WEEK'S FORUM

Week 8

Discuss and ask questions about Week 8.

1907 threads · Last post 6 hours ago

[Go to forum](#)

Forums

All Threads

Search



← All Course Discussions

MA

Validation curve question



Mark Amott Assignment: Regularized Linear Regression and Bias/Variance · 3 years ago · Edited by moderator

I've passed the assignment with 80 marks but cannot get my validationCurve.m to work.

I won't post my added code here, just pseudocode. Hopefully my error will be apparent to someone.

```
for i = 1:length(lambda_vec) %already provided
```

```
lambda = lambda_vec(i); % already provided
```

{Mentor edit: pseudo-code-that-was-really-code-except-for-punctuation removed to protect the Honor Code}

end loop



My graph identifies that the optimum lambda is around 0.5, not 3 as the assignment suggests. I am training the model with the entire X and y and using the ith value of lambda on each occasion.

Given I've passed all other parts of the assignment I can only assume my cost function operates correctly for any order polynomial and my polynomial feature generation function is working correctly.

Ideas?

↑ 11 Upvotes

Reply

Follow this discussion

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

Earliest

Top

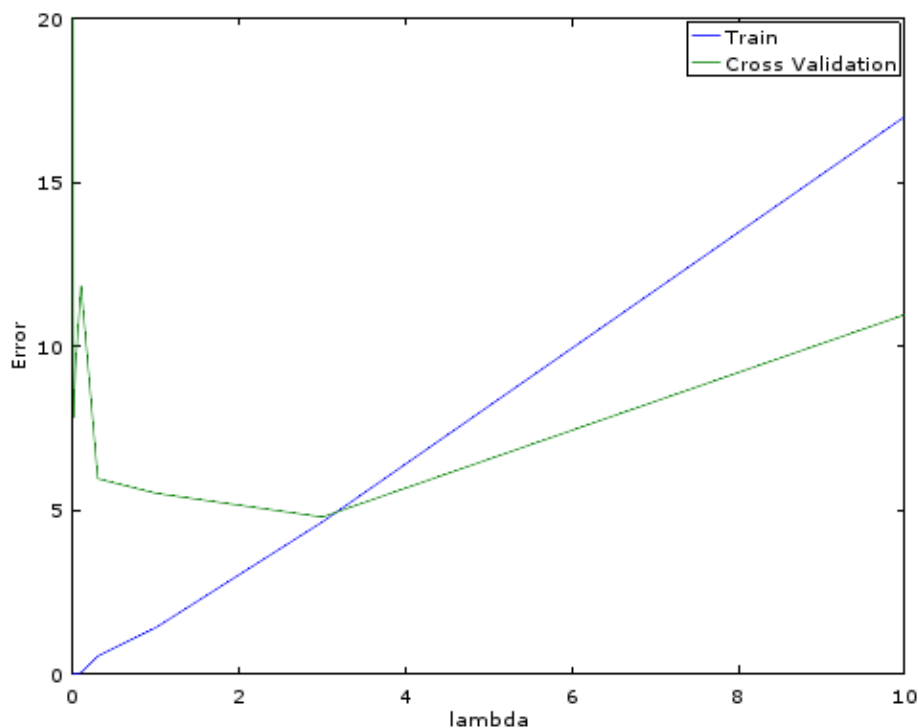
Most Recent



Charles Snyder · 3 years ago



In the validation curve exercise, I produce a graph that shows an optimum lambda value of around 3, but I get no credit when submitting. Looks like it should be ok to me, but I must be missing something. Has anyone else made this mistake and have any advice to offer?





Tom Mosher · Mentor · 3 years ago · Edited

The submit grader doesn't care what is in your plot. It uses a different test case, with a different data set.

- train using the whole training set and a value for lambda
- measure Jtrain and Jcv using the entire training and validation sets, with lambda set to 0

Here are my validation curve data results running ex5.m

1	lambda	Train Error	Validation Error
2	0.000000	0.096879	7.402780
3	0.001000	0.159985	17.399495
4	0.003000	0.176925	17.347999
5	0.010000	0.221821	17.033886
6	0.030000	0.281850	12.829462
7	0.100000	0.459318	7.587013
8	0.300000	0.921760	4.636833
9	1.000000	2.076188	4.260626
10	3.000000	4.901351	3.822907
11	10.000000	16.092213	9.945509

38 Upvotes



Charles Snyder · 3 years ago

Your data and training examples have saved me so much debugging time. Thank you.

If anyone else gets the error that I was having, here was my issue: I wasn't using the WHOLE data set as instructed. I had mindlessly copied my theta formula from the learning curve, so theta was only training on a set of size i.

23 Upvotes

陈

陈秋蕾 · 3 years ago

Can anyone tell me why we need to set lambda = 0 when measuring Jtrain and Jcv?

16 Upvotes



Tom Mosher · Mentor · 3 years ago

When you measure Jtrain and Jcv, you want the true cost without any additional penalties. Regularization is already included in theta - you don't need to include it twice.



↑ 78 Upvotes

coursera



Yuluis@gmail.com · 2 years ago



Thank you Tom, superhero!

↑ 10 Upvotes



Tom Mosher · Mentor · 3 years ago · Edited



For `validationCurve()`, you always use the entire training set, and the entire validation set. The only item you are varying is the value of `lambda` when you compute `theta` on the training set.

Also, do not use regularization when measuring the training error and the validation error.

=====

keywords: tutorial validationcurve

↑ 56 Upvotes

Hide 8 Replies



Zunran Guo · 3 years ago



Thank you so much Tom!

I got it correct now :]

↑ 0 Upvotes



Tom Mosher · Mentor · 3 years ago



Nice work.

↑ 3 Upvotes



Koo Zhengqun · 3 years ago



Hi Tom, may I ask why do we use the entire training and validation sets?

↑ 1 Upvote



Tom Mosher · Mentor · 3 years ago





Using the whole set of training examples gives you the best measurement of cost. If you had the entire dataset, you could then perhaps use a randomly-selected subset, just to speed up the process. That's not an issue in this exercise.

Coursera



In learningCurve, we only used a subset of the training examples specifically because we were demonstrating how cost varies with the size of the training set. The lesson learned there is that larger data sets give better results.

↑ 14 Upvotes

GK Gautam Karmakar · 3 years ago



Hi Tom, for validation curve instruction in script says that use lambda (not zero out after theta calculation) as follows (line #3): Please let me know if it is okay to use non regularized cost for this too.

Instructions: Fill in this function to return training errors in

% error_train and the validation errors in error_val. The

% vector lambda_vec contains the different lambda parameters

% to use for each calculation of the errors, i.e,

% error_train(i), and error_val(i) should give

% you the errors obtained after training with

% lambda = lambda_vec(i)

↑ 0 Upvotes



Tom Mosher · Mentor · 3 years ago



I don't understand your question.

Jtrain and Jcv are always computed without regularization.

↑ 1 Upvote

AS Anu Sabu · 3 years ago



Thank you so much Tom. I was stuck on this curve. I was using the same training set as in the previous problem during linear regularization. I rectified it, after seeing this post. I forgot that in the validation curve, we use the entire training set, and the entire validation set. Thank you for the explanation :)



↑ 1 Upvote

coursera



Farkhod Makhmudkhujaev · 2 years ago



Thank you Tom!



↑ 0 Upvotes

Zunran Guo · 3 years ago



I followed the instruction but somehow couldn't get the correct answers for training error and validation error.

I used the suggested for-loop and used each lambda to calculate the train_error and val_error in each case, using

`linearRegCostFunction(X(1:i,:), y(1:i,:), theta, lambda)` and

`linearRegCostFunction(Xval, yval, theta, lambda)`

like we did in the previous section.

But the answers I got are off a lot. I don't know why :[

lambda Train Error Validation Error

0.000000 0.000000 160.721900

0.001000 0.000196 143.551067

0.003000 0.066071 11.108577

0.010000 0.169094 7.847414

0.030000 0.397854 9.365090

0.100000 1.134154 12.162487

0.300000 3.415910 6.916314

1.000000 8.256224 8.119588

3.000000 18.514901 10.738650

10.000000 36.606671 20.296707

↑ 1 Upvote

💬 Reply



John Hui · 3 years ago





I understand that not using regularisation to compute for training error and validation gets the same answer as your reference solution.



But it does not seem to make sense. The whole point of validation curve using different values of lambda is precisely to evaluate the optimal value of lambda. We are not comparing the right thing if the training and validation errors are calculated with $\lambda = 0$.

I can understand that for learning curve you set lambda to zero when calculating errors. It does not make sense to me that for validation curve, you also do that. What's the math behind this?

↑ 1 Upvote

Hide 11 Replies



Tom Mosher · Mentor · 3 years ago



When training the system, you do use regularization in computing the theta values that minimize the training cost.

When you measure how well that system works, you don't use regularization with the cost function. Regularization is already baked-in to the set of theta values, and you want to measure how well they fit the data. Including the regularization terms would simply add additional penalty values based purely on the numerical values of theta - and that's not what you want in measuring the system performance.

↑ 19 Upvotes



John Hui · 3 years ago



OK. I understand that now. When you use the model, regularisation is no longer a consideration. Regularisation is useful for training purposes only.

↑ 1 Upvote

AN

Anup Nair · 3 years ago



I had the same question as John. Thanks for your very well crafted explanation Tom.

↑ 1 Upvote

BK

Boris Kozlov · 3 years ago



Thanks for the explanation, Tom. Much appreciated.

↑ 1 Upvote



Eloi Teixeira Pereira · 3 years ago



I had the same question. Thanks for the discussion.

↑ 1 Upvote



Ajit Kshirsagar · 2 years ago



Kudos Tom !!!

↑ 0 Upvotes



xiang zhou · 2 years ago



Thanks Tom! So actually lambda is a mean for training the ideal theta and after we got the ideal theta, we apply it to the new data set for prediction without lambda, is that right?

thanks!

erik

↑ 5 Upvotes



Tom Mosher Mentor · 2 years ago



Yes.

↑ 1 Upvote

SF

Stephen Filios · 2 years ago



I'm working on the validation as well, and running into a coding problem that is throwing me for a loop (haha...get it?) - anyhow, I'm getting this weird "A(I) index out of bounds error", citing the line of ex5.m that prints out my lambda/train error/validation error. It prints out only the first line of values with no problems, and then says value 2 out of bound 1 instead of proceeding through the loop. I use the normal for / end structure provided. Any clue what might be causing this?

↑ 0 Upvotes



Tom Mosher Mentor · 2 years ago



You've got an indexing problem with how you are storing the values.

I recommend you read the tutorial from the Resources menu.



↑ 0 Upvotes

coursera



Farkhod Makhmudkhujaev · 2 years ago



Thank you Tom! Your answer about regularization was so helpful!

↑ 1 Upvote

MA

Mark Amott · 3 years ago



Thanks Shreyas and Tom. Much appreciated. I will take a look at the regularisation aspect of the validation (and exclude it). Thanks both for pointing me in the right direction.

↑ 1 Upvote

Reply



Tom Mosher · Mentor · 3 years ago



Don't use regularization when computing the training error and the validation error.

↑ 4 Upvotes

Hide 9 Replies



Hang Dong · 3 years ago



Thank you. :)

↑ 1 Upvote

BW

Beck Weathers · 3 years ago



Why don't we use regularization when computing the training error? Is it to make the errors comparable across different lambdas?

↑ 1 Upvote

BW

Beck Weathers · 3 years ago



Think you answered it below and I was hasty in my reply. Thanks.

↑ 0 Upvotes

AE

Ahmed EL-Shaer · 3 years ago



Tom, your suggestions are dead on. Thanks a lot.

↑ 1 Upvote





GW Guohui Wang · 3 years ago



Thanks. I made the same mistake...

↑ 0 Upvotes



Bertrand Wilson · 3 years ago



Thanks Tom! My code works now.

↑ 0 Upvotes



Tom Mosher Mentor · 3 years ago



Good news.

↑ 0 Upvotes



李莹莹 · 2 years ago



Thanks Tom! I made the same mistake.

↑ 0 Upvotes



Anastasia Voronenko · 2 years ago · Edited



Did the same. Don't know, why, but that's how I read PDF instruction: regularize it. [<https://www.youtube.com/watch?v=fYL7fa3rxV0>]

Thanks to all mentors for our happy childhood!

↑ 0 Upvotes

SP Shreyas Padmanabhan · 3 years ago



Hey,

I just managed to rectify the same error. So if you go back to the part of the exercise where you are calculating the learning curve and pay close attention to the training set error equation. There is a small nuance regarding the regularization while calculating the error for the validation/learning curve. Once you identify that you ought to get the given graph.

↑ 5 Upvotes Reply

MA Mark Amott · 3 years ago



Here's my lambda, error_train and error_val outputs if it helps:

lambda Train Error Validation Error



0.000000 0.155038 18.829079

0.001000 0.184591 16.189554

0.003000 0.251276 18.492268

0.010000 0.385069 17.095362

0.030000 0.669275 13.050780

0.100000 1.443470 8.149386

0.300000 3.101591 5.882451

1.000000 7.268148 7.227460

3.000000 15.867688 10.089385

10.000000 33.372203 19.819789


0 Upvotes Reply


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


J

M


M


BC

M


M


M


M

M

M

M


M

M

M

M

M

M


AC


KK


M

DT


VA

M

M

M

M

M

CT


K

M

KS

M

LK

M

SL

S


MS

M

FL

M


PS

M


JN

M

MR

M

CU

M


T

M


YC

M

AT

M

LG

M

FL

M

JN

M

JD

S

T

YC

AT

M

LG

M

FL



[Learn more about becoming a Mentor](#)



[Forum guidelines](#)

