

Coursera Beta

- Abhishek Singh Sambyal
- Contact Us
- My Contributions
- Log Out

# ML:Errata: Week 3

From Coursera

## Contents

- 1 VI. Logistic Regression
  - 1.1 Decision Boundary
  - 1.2 Cost Function
  - 1.3 Simplified Cost Function and Gradient Descent
  - 1.4 Advanced Optimization
- 2 VII. Regularization
  - 2.1 The Problem of Overfitting
  - 2.2 Cost Function
  - 2.3 Regularized linear regression
  - 2.4 Regularized logistic regression
- 3 Quizzes
- 4 Programming Exercise Errata

## VI. Logistic Regression

PDF: Lecture6.pdf (<https://d396qusza40orc.cloudfront.net/ml/docs/slides/Lecture6.pdf>)

### Decision Boundary

At 1:56 in the transcript, it should read 'sigmoid function' instead of 'sec y function'.

### Cost Function

The section between 8:30 and 9:20 is then repeated from 9:20 to the quiz. The case for  $y = 0$  is explained twice. At 10:28 in the transcript, it should read 'nailed it' instead of 'melted'. At 11:21 in the transcript, it should read 'gradient' instead of 'grading'.

### Simplified Cost Function and Gradient Descent

These following mistakes also exist in the video:

- 6.5: On page 19 in the PDF, the leftmost square bracket seems to be slightly misplaced.
- 6.5: It seems that the factor  $1/m$  is accidentally omitted between pages 20 and 21 when the handwritten expression is converted to a typeset one (starting at 6:53 of the video) At 9:07, 9:13, and 9:25 in the transcript, it should read 'vectorized' instead of 'vector rise'.

## Advanced Optimization

In the video at 7:30, the notation for specifying MaxIter is incorrect. The value provided should be an integer, not a character string. So (...MaxIter, '100') is incorrect. It should be (...MaxIter, 100). This error only exists in the video - the exercise script files are correct.

## VII. Regularization

PDF: Lecture7.pdf (<https://d396qusza40orc.cloudfront.net/ml/docs/slides/Lecture7.pdf>)

### The Problem of Overfitting

At 2:07, a curve is drawn using predicting function  $\theta_0 + \theta_1 x + \theta_2 x^2$ , which is said as "just right". But when size of house is large enough, the prediction of this function will increase much faster than linear if  $\theta_2 > 0$ , or will decrease to  $-\infty$  if  $\theta_2 < 0$ , which neither could correspond to reality. Instead,  $\theta_0 + \theta_1 x + \theta_2 \sqrt{x}$  may be "just right".

At 2:28, a curve is drawn using a quartic (degree 4) polynomial predicting function  $\theta_0 + \theta_1 x + \theta_2 x^2 + \theta_3 x^3 + \theta_4 x^4$ ; however, the curve drawn is at least quintic (degree 5).

### Cost Function

In the video at 5:17, the sum of the regularization term should use 'j' instead of 'i', giving  $\sum_{j=1}^n \theta_j^2$  instead of  $\sum_{i=1}^n \theta_j^2$ .

### Regularized linear regression

In the video starting at 8:04, Prof Ng discusses the Normal Equation and invertibility. He states that X is non-invertible if  $m \leq n$ . The correct statement should be that X is non-invertible if  $m < n$ , and may be non-invertible if  $m = n$ .

### Regularized logistic regression

In the video at 3:52, the lecturer mistakenly said "gradient descent for regularized linear regression". Indeed, it should be "gradient descent for regularized logistic regression".

In the video at 5:21, the cost function is missing a pair of parentheses around the second log argument. It should be  $J(\theta) = \left[ -\frac{1}{m} \sum_{i=1}^m y^{(i)} \log(h_{\theta}(x^{(i)})) + (1 - y^{(i)}) \log(1 - h_{\theta}(x^{(i)})) \right] + \frac{\lambda}{2m} \sum_{j=1}^n \theta_j^2$

In the original videos for the course (ML-001 through ML-008), there were typos in the equation for regularized logistic regression in both the video lecture and the PDF lecture notes. In the slides for "Gradient descent" and "advanced optimization", there should be positive signs for the regularization term of the gradient. The formula on page 10 of 'ex2.pdf' is correct. These issues in the video were corrected for the 'on-demand' format of the course.

## Quizzes

- Typo "it's" in question «Because logistic regression outputs values  $0 \leq h_{\theta}(x) \leq 1$ , it's range [...]»

- $\frac{1}{m}$  factor missing in the definition of the gradient in question «For logistic regression, the gradient is given by  $\frac{\partial}{\partial \theta_j} J(\theta) = \sum_{i=1}^m [\dots]$ ». It should be  $\frac{\partial}{\partial \theta_j} J(\theta) = \frac{1}{m} \sum_{i=1}^m [\dots]$ . See also #Simplified Cost Function and Gradient Descent above.

## Programming Exercise Errata

- In ex2.pdf on page 5, Section 1.2.3, "gradient descent" should be "gradient descent".
- If you are using a linux-derived operating system, you may need to remove the attribute "MarkerFaceColor" from the plot() function call in plotData.m.
- In ex2.m at lines 10 through 13, the list of files the student needs to complete should include plotData.m
- In costFunction.m on line 12, the gradient vector is incorrectly initialized to an  $(n + 1) \times (n + 1)$  zero matrix via `grad = zeros(size(theta))`. The gradient vector should instead be initialized to an  $(n + 1) \times 1$  zero vector via `grad = zeros(size(theta),1)`.
- In costFunctionReg.m on line 12, the gradient vector is incorrectly initialized to an  $(n + 1) \times (n + 1)$  zero matrix via `grad = zeros(size(theta))`. The gradient vector should instead be initialized to an  $(n + 1) \times 1$  zero vector via `grad = zeros(size(theta),1)`.

Retrieved from "[https://share.coursera.org/wiki/index.php?title=ML:Errata:\\_Week\\_3&oldid=34425](https://share.coursera.org/wiki/index.php?title=ML:Errata:_Week_3&oldid=34425)"

Category: ML:Errata

- 
- This page was last modified on 3 November 2016, at 18:40.
  - This page has been accessed 8,829 times.
  - [Privacy policy](#)
  - [About Coursera](#)
  - [Disclaimers](#)