ML:Errata: Week 1

From Coursera

Contents

Introduction

- Supervised Learning: 1:25: Describing the curve as quadratic is confusing since the independent variable is price, but the plot's X-axis represents area.
- Unsupervised Learning: 5:01: 'surprisingly interesting useful theories of how galaxies are formed'.
- Unsupervised Learning: 6:56 - the mouse does not point to the correct audio sample being played on the slide. Each subsequent audio sample has the mouse pointing to the previous sample.
- Unsupervised Learning: 12:50 - the slide shows first option "Given email labelled as span/not spam, learn a spam filter" as one of the answers as well. Whereas, in the audio Professor puts it in Supervised Learning category.
- Unsupervised Learning: 12:50 - the slide shows fourth option 'Given a dataset of patients diagnosed.....having diabetes of not.' as one of the answers as well. Whereas, in the audio Professor puts it in Supervised Learning category.

Cost Function

- In the transcript, at 0:26: 'the parameters of the maldo' should be changed to 'the parameters of the model'.
- In the transcript, at 2:24: 'the parana state of zero and' should be changed to 'the parameters theta zero and'.
- In the transcript, at 2:50: 'the value we predict on equals x' should be changed to 'the value we predict on input x'.
- In the transcript, at 3:10: 'values for the perimeters so that' should be changed to 'values for the parameters so that'.
- In the transcript, at 3:21: 'so then regression' should be changed to 'so linear regression'.
- In the transcript, at 7:09 and 7:18: 'cause function' should be changed to 'cost function'.
- In the transcript, at 7:11: 'squares of the areas' should be changed to 'squares of the errors'.

Linear Regression With One Variable

- A general note about the graphs that Prof Ng sketches when discussing the cost function. The vertical axis can be labeled either 'y' or 'h(x)' interchangeably. 'y' is the true value of the training example, and is indicated with a marker. 'h(x)' is the hypothesis, and is typically drawn as a curve. The scale of the vertical axis is the same, so both can be plotted on the same axis.
- On the cost function explanation page, the use of $\bar{x}$ could be misleading. Often that would represent the mean of all the x_i values. But here (as stated) it's the mean of the squared "deviations," $(h_\theta(x_i) - y_i)^2$.
- In the video "Cost Function - Intuition I", at about 5:30, Prof Ng says, "...the sum of squared values for the height of this line...". etc. Not quite. For each line segment in his example, the term is just the square of the height of the line, not the *sum* of the squares.
- In the video "Cost Function - Intuition I", at about 6:34, the value given for J(0.5) is incorrect. Should be 0.58, not 0.68.
- In the video "Cost Function - Intuition I", at about 6:52, the vertical axis name: "h(x)" is incorrect, it should be "y".
- Parameter Learning: Video "Gradient Descent for Linear Regression": At 6:15, the equation Prof Ng writes in blue "h(x) = -900 - 0.1x" is incorrect, it should use "+900".

Gradient Descent for Linear Regression

- At Timestamp 3:27 of this video lecture, the equation for θ1 is wrong, please refer to first line of Page 6 of ex1.pdf (Week 2 programming Assignment) for model equation (The last x is X superscript i, subscript j (Which is 1 in this case, as it is of θ1)). θ0 is correct as it will be multiplied by 1 anyways(value of X superscript i, subscript 0 is 1), as per the model equation.
- At 6:30 there is a clarification that states j=0 and j=1. This statement is just confusing while the code is actually clear.
- On the **Gradient Descent for Linear Regression** page (not the video):

The derivation of the derivative could confuse some students. Some subscripts are misplaced, etc. I *think* the equation is trying to show us how to do the derivative for θ_1. Here is another try! (I will use x_i instead of $x^{(i)}$. Also, sorry I can't see how to make display equations properly in this markdown!)

$$\frac{\partial}{\partial\theta_1} J(\theta_0, \theta_1) = \frac{\partial}{\partial\theta_1} \sum_{i=1}^{m} \frac{1}{2} \left(h(x_i) - y_i\right)^2$$

Move the derivative inside the sum and the $\frac{1}{2}$ outside the sum:

$$\frac{\partial}{\partial\theta_1} J(\theta_0, \theta_1) = \frac{1}{2} \sum_{i=1}^{m} \frac{\partial}{\partial\theta_1} \left(h(x_i) - y_i\right)^2$$

Use the chain rule to get:

$$\frac{\partial}{\partial\theta_1} J(\theta_0, \theta_1) = \frac{1}{2} \sum_{i=1}^{m} 2(h(x_i) - y_i) \frac{\partial}{\partial\theta_1} \left(h(x_i) - y_i\right)$$

Now, $h(x_i) = \theta_1 x_i + \theta_0$, so we substitute in to get

$\frac{\partial}{\partial \theta_1} J(\theta_0, \theta_1) = \frac{1}{2} \sum_{i=1}^{m} 2(h(x_i) - y_i) \frac{\partial}{\partial \theta_1} (\theta_1 x_i + \theta_0 - y_i)$

And *that* derivative $\frac{\partial}{\partial \theta_1} (\theta_1 x_i + \theta_0 - y_i) = x_i$. So

$\frac{\partial}{\partial \theta_1} J(\theta_0, \theta_1) = \sum_{i=1}^{m} (h(x_i) - y_i) x_i,$

which is what you see in the "repeat-until-convergence" block up above.

Linear Algebra Review

- Matrix-Matrix Multiplication: 7:14 to 7:33 - While exploring a matrix multiplication, Andrew solved the problem correctly below, but when he tried to rewrite the answer in the original problem, one of the numbers was written incorrectly. The correct result was (matrix 9 15) and (matrix 7 12), but when it was rewritten above it was written as (matrix 9 15) and (matrix 4 12). The 4 should have been a 7. (Thanks to John Kemp and others). This has been partially corrected in the video - third subresult matrix shows 7 but the sound is still 4 for both subresult and result matrices. Subtitle at 6:48 should be “two is seven and two”, and subtitle at 7:14 should be “seven twelve and you”.
- 3.4: Matrix-Matrix Multiplication: 8:12 - Andrew says that the matrix on the bottom left shows the housing prices, but those are the house sizes as written above
- 3.6: Transpose and Inverse: 9:23 - While demonstrating a transpose, an example was used to identify B(subscript 12) and A(subscript 21). The correct number 3 was circled in both cases above, but when it was written below, it was written as a 2. The 2 should have been a 3. (Thanks to John Kemp and others)

Addition and scalar multiplication video

- Spanish subtitles for this video are wrong. Seems that those subtitles are from another video.