✓ 1. A single parameter model has mean-square error (MSE) defined below . We have a half term in the front because, 1/1

$$\frac{1}{2N}\sum_{n=1}^{N}(y_n - \beta_0)^2$$

- ○ Scaling MSE by half makes gradient descent converge faster.
- ○ Presence of half makes it easy to do grid search.
- ◉ It does not matter whether half is there or not. ✓
- ○ None of the above

✓ 2. A researcher wants to perform a simple linear regression to find out if the socio-economic status of a teacher can predict whether they work at a primary or a secondary school. Why can't this be done? 1/1

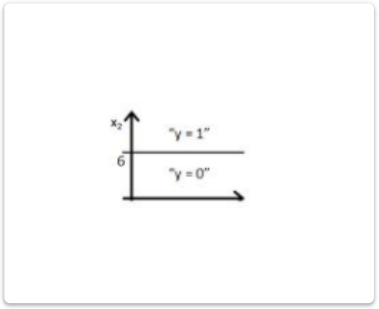
- ○ Because there are not enough variables for the analysis
- ◉ Because the outcome variable is nominal not continuous ✓
- ○ Can't Say
- ○ Because socio-economic status can not be used as a predictor variable

✓ 3. Hypothesis function of a logistic regression classifier is given below. Which of the following will represent the decision boundary of the classifier? 2/2

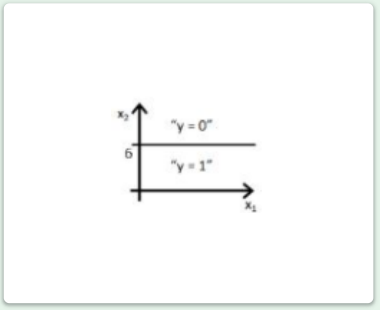
$$h_\theta(x) = g(\theta_0 + \theta_1 x_1 + \theta_2 x_2)$$

Where

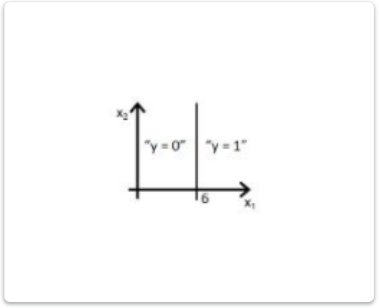
$$\theta_0 = 6, \theta_1 = 0, \theta_2 = -1.$$



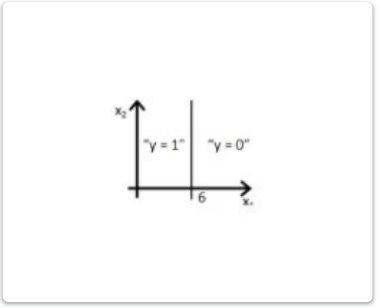
A



B ✓



C



D

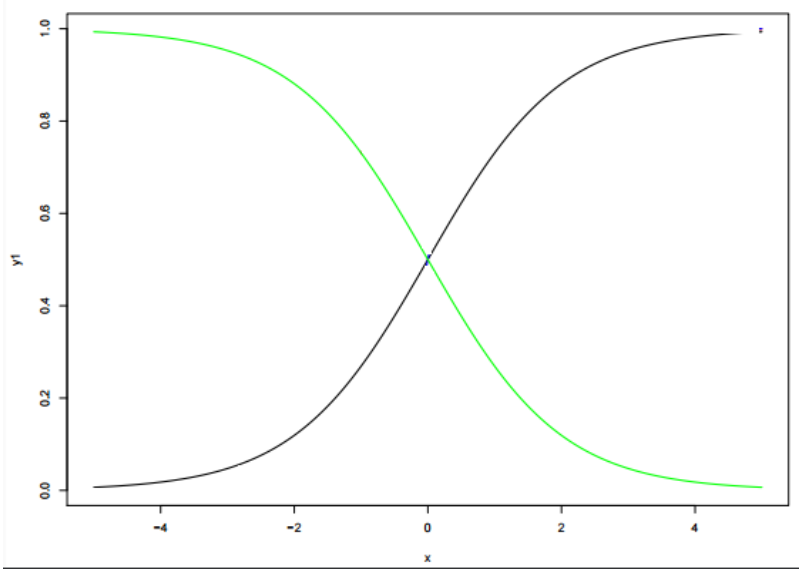
✓ 4. If a linear regression model has zero training set error, then: 2/2

- Test error = Train error
- Can't comment on test error ✓
- Test error is also always zero
- Test error is non-zero

✓ 5. Which of the following options is/are TRUE? 2/2

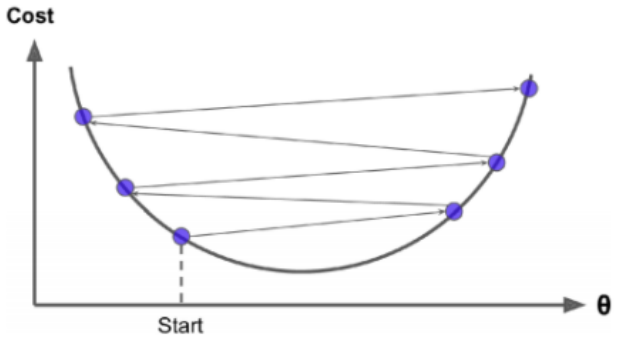
- [ ] Both Linear Regression and Logistic Regression error values have to be normally distributed
- [ ] Both Linear Regression and Logistic Regression error values have not to be normally distributed
- [x] Linear Regression errors values have to be normally distributed but in case of Logistic Regression it is not the case ✓
- [ ] Logistic Regression errors values have to be normally distributed but in case of Linear Regression it is not the case

✓ 6. Let the curves given below be of 2 logistic models, with different values of θ0 and θ1. Consider Y = θ0 + θ1*X where θ0 is the intercept and θ1 is the coefficient. Select the TRUE statement(s) 2/2



- [ ] θ1 for green is greater than black
- [ ] θ1 for both models is same
- [x] θ1 for green is lower than black ✓
- [ ] Can't say

✓ 7. Which case of learning rate would the following figure correspond to: 2/2



- (●) Learning rate is too large ✓
- ( ) Learning rate is too small
- ( ) Can't say
- ( ) Doesn't depend on the learning rate

✗ 8. Given a dataset with input x and output y. To test our linear model on this data, we split the data into training set and test set randomly. Now we increase the training set size gradually. As the training set size increases, what do you expect will happen with the mean training error? 0/2

- (●) Decrease ✗
- ( ) Remain constant
- ( ) Increase
- ( ) Can't Say

Correct answer

- (●) Can't Say

✓ 9. Gradient descent approach has computational complexity of order: 2/2

- ( ) O(n^3)
- ( ) O(2^n)
- ( ) O(n)
- (●) O(n^2) ✓

✗ 10. Select all the statements that are FALSE: 0/2

- [ ] A regression line can be used to predict one variable from another
- [x] A regression line is created by minimizing the difference between the line and the data points ✗
- [ ] A regression line can be calculated with data from a single variable
- [x] A regression line is created by maximizing the difference between the line and the data points ✓

Correct answer

- [x] A regression line is created by maximizing the difference between the line and the data points
- [x] A regression line can be calculated with data from a single variable

Assignment Questions 12 of 12 points

Make sure you have done the assignment before attempting this part of the quiz.

✓ 11. Computed cost (rounded off upto 3 decimal places) for A=0.69 and B = -0.42 is 2/2

- ( ) 13.717
- ( ) 73.322
- ( ) 10.388
- (•) 60.535 ✓

✓ 12. Initialize A = 0.69, B = 0.42 . Setting alpha = 0.001 what would be the value of B (rounded off upto 4 decimal places) computed by gradient descent after 2500 iterations : 2/2

- ( ) 0.7497
- ( ) 1.1879
- ( ) 0.8945
- (•) 0.8975 ✓

✓ 13. Initialize A = 1.0, B = -2.0 . Setting alpha = 0.01 what would be the value of A (rounded off upto 3 decimal places) computed by gradient descent after 2500 iterations : 2/2

- (●) -3.839 ✓
- ( ) -3.843
- ( ) -3.861
- ( ) -3.865

✓ 14. For population of 135000, what is estimated profit predicted by the model (in $) (For A=0.69,B=-0.42, Alpha = 0.01, Iterations=2500) 3/3

- ( ) 160364.89
- ( ) 16.04
- (●) 121900 ✓
- ( ) 12.19

✓ 15. For a population size of 80000, what is the predicted profit (in $)? (For A=2.0,B=-1.0, Alpha = 0.001, Iterations=1500) 3/3

- (●) 65600 ✓
- ( ) 5.58
- ( ) 6.56
- ( ) 55800

You've reached the end of the quiz! **0 of 0 points**

I have read all my answers and this is my final submission. *

- (●) Yes
- ( ) No