Congratulations! You passed!

TO PASS 80% or higher **QUIZ • 10 MIN**

Keep Learning

GRADE
100%

Linear Regression with Multiple Variables

Linear Regression with Multiple Variables

LATEST SUBMISSION GRADE

100%

Submit your assignment

Try again

DUE Sep 30, 2:59 PM SGT **ATTEMPTS** 3 every 8 hours

1. Suppose *m*=4 students have taken some class, and the class had a midterm exam and a final exam. You have collected a dataset of their scores on the two exams, which is as follows: **1 / 1 point**

Receive grade

TO PASS 80% or higher

Grade
100%

View Feedback

We keep your highest score

midterm exam	(midterm exam)^2	final exam
89	7921	96
72	5184	74
94	8836	87
69	4761	78

You'd like to use polynomial regression to predict a student's final exam score from their midterm exam score. Concretely, suppose you want to fit a model of the form $h_\theta(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2$, where x_1 is the midterm score and x_2 is (midterm score)^2. Further, you plan to use both feature scaling (dividing by the "max-min", or range, of a feature) and mean normalization.

What is the normalized feature $x_2^{(2)}$? (Hint: midterm = 72, final = 74 is training example 2.) Please round off your answer to two decimal places and enter in the text box below.

Correct

2. You run gradient descent for 15 iterations with $\alpha = 0.3$ and compute $J(\theta)$ after each iteration. You find that the value of $J(\theta)$ decreases slowly and is still decreasing after 15 iterations. Based on this, which of the following conclusions seems most plausible? **1 / 1 point**

Correct

3. Suppose you have $m = 28$ training examples with $n = 4$ features (excluding the additional all-ones feature for the intercept term, which you should add). The normal equation is $\theta = (X^T X)^{-1} X^T y$. For the given values of m and n, what are the dimensions of θ, X, and y in this equation? **1 / 1 point**

Correct

4. Suppose you have a dataset with $m = 1000000$ examples and $n = 200000$ features for each example. You want to use multivariate linear regression to fit the parameters θ to our data. Should you prefer gradient descent or the normal equation? **1 / 1 point**

Correct

5. Which of the following are reasons for using feature scaling? **1 / 1 point**

Correct