



Multiple Regression



4/9 points earned (44%)

[Review Related Lesson \(/learn/ml-regression/home/week/2\)](/learn/ml-regression/home/week/2)

You haven't passed yet. You need at least 80% to pass.
Review the material and try again! You have 3 attempts every 8 hours.



1. Which of the following is **NOT** a **linear** regression model. *Hint: remember that a linear regression model is always linear in the parameters, but may use non-linear features.*

1 / 1
points



2. Your estimated model for predicting house prices has a large positive weight on 'square feet living'. This implies that if we remove the feature 'square feet living' and refit the model, the new predictive performance will be **worse** than before.

0 / 1
points



3. *Complete the following:* Your estimated model for predicting house prices has a positive weight on 'square feet living'. You then add 'lot size' to the model and re-estimate the feature weights. The new weight on 'square feet living' [] be positive.

0 / 1
points



4. If you double the value of a given feature (i.e. a specific column of the feature matrix), what happens to the least-squares estimated coefficients for every **other** feature? (assume you have no other feature that depends on the doubled feature i.e. no interaction terms).

0 / 1
points



5. Gradient descent/ascent is...

1 / 1
points



6. Gradient descent/ascent allows us to...

1 / 1
points



7. Which of the following statements about step-size in gradient descent is/are **TRUE** (select all that apply)

0 / 1
points



8. Let's analyze how many computations are required to fit a multiple linear regression model *using the closed-form solution* based on a data set with 50 observations and 10 features. In the videos, we said that computing the inverse of the 10×10 matrix $(H^T)H$ was on the order of D^3 operations. Let's focus on forming this matrix **prior** to inversion. How many multiplications are required to form the matrix $(H^T)H$?

1 / 1
points

Please enter a number below.



9. More generally, if you have D features and N observations what is the total complexity of computing $((H^T)H)^{-1}$?

0 / 1
points

