

# Ridge Regression



6/9 points earned (66%)

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.

[Review Related Lesson](#)



1 / 1  
points

1.

**Which of the following is NOT a valid measure of overfitting?**



1 / 1  
points

2.

**In ridge regression, choosing a large penalty strength  $\lambda$  tends to lead to a model with (choose all that apply):**



1 / 1  
points

3.

**Which of the following plots best characterize the trend of bias, variance, and generalization error (all plotted over  $\lambda$ )?**



0 / 1  
points

4.

In ridge regression using unnormalized features, if you double the value of a given feature (i.e., a specific column of the feature matrix), what happens to the estimated coefficients for every other feature? They:

---



1 / 1  
points

5.

If we only have a small number of observations, K-fold cross validation provides a better estimate of the generalization error than the validation set method.

---



1 / 1  
points

6.

10-fold cross validation is more computationally intensive than leave-one-out (LOO) cross validation.

---



0 / 1  
points

7.

Assume you have a training dataset consisting of  $N$  observations and  $D$  features. You use the closed-form solution to fit a multiple linear regression model using ridge regression. To choose the penalty strength  $\lambda$ , you run leave-one-out (LOO) cross validation searching over  $L$  values of  $\lambda$ . Let  $\text{Cost}(N,D)$  be the computational cost of running ridge regression with  $N$  data points and  $D$  features. Assume the prediction cost is negligible compared to the computational cost of training the model. Which of the following represents the computational cost of your LOO cross validation procedure?

---



1 / 1  
points

8.

Assume you have a training dataset consisting of 1 million observations. Suppose running the closed-form solution to fit a multiple linear regression model using ridge regression on this data takes 1 second. Suppose you want to choose the penalty strength  $\lambda$  by searching over 100 possible values. How long will it take to run leave-one-out (LOO) cross-validation for this selection task?

---



0 / 1  
points

9.

Assume you have a training dataset consisting of 1 million observations. Suppose running the closed-form solution to fit a multiple linear regression model using ridge regression on this data takes 1 second. Suppose you want to choose the penalty strength  $\lambda$  by searching over 100 possible values. If you only want to spend about 1 hour to select  $\lambda$ , what value of  $k$  should you use for  $k$ -fold cross-validation?

---

