PS9 Yarberry

Megan N. Yarberry

March 31, 2020

1 Question 5

What is the dimension of your training data (housing.train)?

- 404 objects by 648 variables

2 Question 6 - LASSO Model

What is the optimal value of ?

- 7120.064

What is the in-sample RMSE?

- 0.1740835

What is the out-of-sample RMSE (i.e. the RMSE in the test data)?

- 0.1692024

3 Question 7 - Ridge Regression Model

What is the optimal value of ?

- 7120.064

What is the in-sample RMSE?

- 0.1545797

What is the out-of-sample RMSE (i.e. the RMSE in the test data)?

- 0.1548843

4 Question 8 - Elastic Net Model

What is the optimal value of ?

- 0.192089

What is the in-sample RMSE?

- 0.061

What is the out-of-sample RMSE (i.e. the RMSE in the test data)?

- 0.202

Does the optimal value of lead you to believe that you should use LASSO or ridge regression for this prediction task?

- It would be better to use the ridge regression model better

5 Question 9

Based on bias-variance tradeoff, estimate a simple linear regression model on the housing.train dataframe. You are more likely to go with the Ridge Regression Model since the in sample and out of sample are the closest.