PS9$_T$*hatcher*

rachel.e.thatcher-1

April 2018

1 Question 5

What is the dimension of your training data (housing.train)?
6 rows, 450 columns

2 Question 6

What is the optimal value of ? .0054
What is the in-sample RMSE? .1566
What is the out-of-sample RMSE? . 1948

3 Question 7

What is the optimal value of now? .0886
What is the in-sample RMSE? .1465
What is the out-of-sample RMSE? .1844

4 Question 8

What are the optimal values of and after doing 6-fold cross validation?
lambda = .0371 alpha = .413

What is the in-sample RMSE? .1651

What is the out-of-sample RMSE? .2147

Does the optimal value of lead you to believe that you should use LASSO or ridge regression for this prediction task?
Ridge regression because it has alpha = 0 which is closer than LASSO which uses alpha = 1.

5 Question 9

Explain why you would not be able to estimate a simple linear regression model on the housing.train dataframe.

Because the simple linear regression is too simplistic and wouldn't be able to represent the data well.

Using the RMSE values of each of the tuned models in the previous three questions, comment on where your model stands in terms of the bias-variance tradeoff.

It seems to do worse, with a higher RMSE than the LASSO or Ridge Regression models.