A student submits an assignment 6 days late thanks to the grace of his professor

Luke Denton

April 20, 2021

1 Question 7

The prepped housing training data has 405 obs of 75 X Variables. The original housing data has 14 X variables. The prepped housing training data has 61 more X variables than the original housing data.

2 Question 8

The optimal value of lambda is 0.00356. The in sample RMSE is 0.220; the out of sample RMSE is 0.220.

3 Question 9

The optimal value of lambda is 0.0146; the out of sample RMSE is 0.218.

4 Question 10

You are unable to run a linear regression when k > n. There is not enough data to make distinctions about the effects of different values of all the variables. With fairly small RMSE in both our Lasso and Ridge models, it can be inferred that we have low variance within our predictions. This suggests we may have slightly more bias in our predictions. Although, it's important to remember our cross validation efforts are meant to bring us to the sweet spot between bias and variance, so we could be in a very good spot.