

Problem Set 9

Leah Pomerantz

April 13, 2022

1 Problem 7

Using the commands `dim(housing)` and `dim(housing_train)`, I found that they had the same number of columns (14), but that `housing_train` had fewer rows (404 instead of 506, as in the original data).

2 Problem 8

Optimal Lambda: Our optimal λ is 0.00139

In-sample RMSE: Our in-sample RMSE is 0.0682

Out-of-sample RMSE: Our out-of-sample RMSE is 0.188

3 Problem 9

Optimal Lambda: 0.0373 Out-of-sample RMSE: 0.185

4 Problem 10

You wouldn't be able to do simple linear regression with more columns than rows, it would cause too many problems with multicollinearity. Unlike the unbiased OLS, ridge and lasso models favor a lower variance in return for allowing some bias.