Problem Set 9

savannahjsimpson

April 9 2024

1 Creating the New Recipe

- The new dimensions are 404 X 75.
- There are now 61 more X variables than in the original housig data.

2 LASSO Model

- The optimal value of lambda is 0.002222996 .
- The in-sample RSME is 1.958302.
- The out of sample RSME is 1.950796.

3 Ridge Regression Model

- The optimal value of lambda is now 0.03727594 .
- The out of sample RSME is 1.949841.

4 Overview

- If the dataset were to have had more columns than rows, I would not have been able to run the regression. In my process, I experienced many errors regarding non-numeric columns, and having an uneven number of data would further complicate calculations.
- The ridge regression produced a higher value for lambda, meaning that it is generalizing the data more. The LASSO analysis is more specific, signalizing less bias overall. Given the similar RSME values, the variance for both models is relatively low. Looking deeper, I would utilize the RR model as it is more generalized and has a slightly lower RSME.