PS9

Adam Matthies

April 8, 2025

1 Modeling Questions

1. The optimal value of lambda for LASSO was .001 and the in-sample RMSE was 0.413. The out-of-sample RMSE was 0.390.

2. The optimal value of lambda for Ridge was 0.00464 and the out-of-sample RMSE was 0.167.

The ridge regression seemed a lot more accurate because all of the independent variables needed to be kept to give accurate predictions of home values. My lasso shrunk about every variable to 0 due to the penalty of .001 so I am convinced my optimization may have been done incorrectly. However, for this dataset, I do believe a ridge regression makes more sense to use–which can be seen in the out-of-sample RMSE.

No, you cannot have more observations than predictors in a simple linear regression. The RMSE from the ridge was quite low, so it shows a good tradeoff between bias and variance. The lasso RMSE was higher, so it is either showing an issue of overfitting or underfitting the data.