PS9

William Townsend
Department of Economics, The University of Oklahoma, Norman

April 12, 2022

1 Problem 7

The dimension of my train data for $housing_train$ is 404 x 14 (404 observations and 14 variables). We have 59 more X variables because we must deduct the two Y outcome columns. The dimension of the new data is 404 x 73.

2 Problem 8

The optimal lambda is 0.00139. The in-sample RMSE in my data seems to be 0.170. The out-sample RMSE is 0.0632.

3 Problem 9

The optimal lambda is 0.0373. The outsample RMSE is 0.173.

I honestly tried on questions 8 and 9 but I know I'm not hitting the mark.

4 Problem 10

No, you would not be able to estimate a simple linear regression model on a data set that had more columns than rows. I would say that this model has a high bias-variance trade off based on its trends.