PS9$_T homasson$

Campbell Thomasson

April 9, 2024

1 PS9

housing - 14 X Variables

housing train - 14 X variables

housing train prepped - 75 X variables

housing train X - 74 X variables

housing train Y - 1 X Variable

There are no more additional X variables in the housing train data

For the LASSO model:

The optimal value of λ is 0.00139. The in-sample RMSE is 0.413. The out-of-sample RMSE is 0.390.

For the Ridge model:

The optimal value of λ is 0.0373.

No, you would not be able to to estimate a simple linear regression model on a dataset that has more columns than rows, as there would be more predictors than observations.

The LASSO model with the optimal λ value of 0.00139 has a slightly lower out-of-sample RMSE (0.390) compared to the in-sample RMSE (0.413). This indicates lower variance and potentially higher bias.

The Ridge model with the optimal λ value of 0.0373 also has slightly higher out-of-sample RMSE compared to the in-sample RMSE. However, the difference is smaller compared to the LASSO model, suggesting a better balance between bias and variance.