

PS9 Holt

richard.n.holt

March 2019

1 Problem Set Questions

5. The housing train data has observations of 450 different variables.

6. The optimal value of lamda is 0.00705. The in-sample RMSE is 0.1674909. The out-of-sample RMSE is 31600.48.

7. The new optimal value of lamda is 0.098. The in-sample RMSE is 0.1734887. The out-of-sample RMSE is also 31600.48.

8. The elastic optimal value of lamda is 0.00314. The optimal value of alpha is 0.829. The in-sample RMSE is 0.1658791. The out-of-sample RMSE is 31600.48. The Lasso regression should be used for this prediction test.

9. You would not be able to use a simple linear regression because the data takes a clear non-linear form. I would say that based on the RMSE values, my model has low variance, but high bias. The lowest bias belongs to the Lasso model, and the highest bias belongs to the ridge regression, with the elastic net model falling near the ridge regression.