PS9$_R$*einsch*

Emma Reinsch

April 2021

1 Questions from PS9 Output

1. (7 from PS9) After running the appropriate code, my training data now has 405 observations of 75 variables. In my original housing data, I had 14 X variables. The new training housing data has 74 X variables.

2. (8 from PS9) My in-sample RMSE is 0.220 (R-squared 0.768), and my out-of-sample RMSE is 0.219 (R-squared 0.770) The optimal value of lambda for the in-sample RMSE is 0.00222. The optimal value of lambda for the out-of-sample RMSE is 0.00139.

3. (9 from PS9) My out-of-sample RMSE is 0.219 (R-squared 0.776). The optimal value of lambda is now much smaller. It is 0.0000000001.

4. (10 from PS9) A simple linear regression allows us to observe the direct relationship between the independent variable (X) and the dependent variable (Y). One would not be able to estimate a simple linear regression if there are more columns than rows. This would make it a multiple linear regression. The bias-variance tradeoff refers to the tradeoff that takes place when we choose to lower the bias which typically increases variance or vice versa. In this case, we observe low variances across all three cases. This shows that there is a high amount of bias in the sample.