PS9

Christin Bivens

April 3rd, 2018

5. 404 rows ,16 columns

6. in sample rmse = 0.2072434 Out of sample rmse = 0.1917285 lamdba = 0.01466217

I ran this code about 5 times and my answers could vary up to .05, which seems like a high amount of variation between just a few iterations.

7. lambda = 0.006271684 in sample rmse = 0.1961141 oos rmse = 0.2043624

8. lambda = 0.07564617 in samples rmse = 0.2007728 oos rmse = 0.2056368

9. I would not use a simple OLS regression because it could overfit my data. OLS does not provide a means of regularization for the data. This means my model could fit too closely to the sample data and therefore would not be a good predictor for out of sample.

RMSE from all the methods sticks fairly close to .20. Being under .5 this means that the variance of the model is relatively low and therefore more regularized. A high variance is usually associated with overfitted data where noise is taken too much into account. In this situation we trade a lower variance through regularization for a higher bias on our model.