

PS9_{Quick}

Samuel Quick

April 2019

1 Question 5

There are observations with 450 variables

2 Question 6

Optimal Lambda was found to be $\lambda = 0.0435$ — The in sample RMSE was $= 0.2041295$ — The out of sample RMSE $= 0.1842767$

3 Question 7

The new optimal value of $\lambda = 0.0976$ — In sample $= 0.1645316$ — Out of sample $= 0.1424507$ —

4 Question 8

The new optimal value of $\lambda = 0.0137$ — In sample $= 0.1672074$ — Out of sample $= 0.1461131$ — Of all options, I would choose the LASSO model for this example.

5 Question 9

You would not be able to use a simple linear regression because the data is clearly not uniformly linear at any point - e.i there are too many variables that need to be categorized. Considering the above analysis, I would say low variance and high bias.