PS9

Qinghua Feng

April 2023

1 **Question 8: The optimal value of is 0.0000324. The in-sample RMSE is 0.182, as shown in the first row of the first tibble. The out-of-sample RMSE is 0.0705, as shown in the first row of the second tibble.**

2 **Question 9: The optimal value of for the ridge regression model with tuning by 6-fold cross-validation is 0.0000000001. The out-of-sample RMSE (i.e. the RMSE in the test data) is 0.168. However, the out-of-sample RMSE seems to be very large and unlikely to be accurate.**

3 **It is possible to fit a simple linear regression model on a dataset that has more columns (features) than rows (observations). However, this will likely result in an overfit model due to the high dimensionality,both models appear to perform well in terms of the bias-variance tradeoff. However, the unusually low out-of-sample RMSE for the model in Question 8 should be carefully investigated to ensure the model's performance is accurate.**