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 0.0001325711. The out-of-sample RMSE (i.e. the RMSE in the test data) is 1.030134e+21. However, the out-of-sample RMSE seems to be very large and unlikely to be accurate.**

3 **Yes, it is possible to estimate a simple linear regression model on a dataset with more columns than rows, but it may lead to overfitting and unreliable results. Based on the RMSE values of the tuned models in questions 8 and 9, it appears that the ridge regression model with tuning by 6-fold cross-validation has lower RMSE and thus better performance, indicating a better balance between bias and variance.**