Problem Set 9[1]

Chang Gao

Q8

The optimal value of λ is 0.0022229964825262, the in-sample RMSE is 0.138339826517746, the out-of-sample RMSE is 0.183976870925607.

Q9

The optimal value of λ is 0.0372759372031494, the in-sample RMSE is 0.140411729973063, the out-of-sample RMSE is 0.179788378367035.

Q10

We cannot estimate a simple linear regression model on a dataset that has more columns than rows. This is because the $X'X$ matrix becomes singular and non-invertible.

The Ridge model has slightly higher in-sample RMSE and lower out-of-sample RMSE, this means worse in-sample fit but better generalization to unseen data.

---

[1]I use paid version of Claude.ai to help with the problem sst