PS9 Garcia

Vladimir Garcia

March 4th 2018

1. Answer the questions posed in the preceding four questions. Explain why you would not be able to estimate a simple linear regression model on the housing.train dataframe. Using the RMSE values of the tuned models in the previous three questions, comment on where your model stands in terms of the bias-variance tradeoff.

   - From the LASSO model, the optimal value of lambda is 0.0362 and alpha is 1. The in-sample RMSE is 0.2019874 and the out-of-sample error is 0.141141.

   - From the Ridge Regression Model, the optimal value of lambda is 0.135 and alpha is 0, the in-sample RMSE is 0.16706614 and the out-of-sample RMSE is 0.1470856.

   - From the elastic net model, the optimal lambda is 0.0113 and alpha is 0.129. The in-sample RMSE is 0.1807778, and the out-of-sample is RMSE 0.1539898. The more general elastic net is more similar to the Ridge Regression model.

   - We cannot run a simple linear model because this will result in underfitting because the function and feature the model would use to predict would be considered too simplistic, and this, will have higher bias.

   Within the bias-variance trade-off our model is regularized. For each model, the prediction accuracy measured in in-sample RMSE is close to ou-of-sample RMSE. In this case the model trades off really well between out-of-sample overfitting with high variance, and in-sample underfitting with high variance.