Cross Validation Notes

- The book introduces Cross-Validation in the section of chapter 2 titled “Better Evaluation Using Cross-Validation”.
- sklearn’s brief introduction to cross validation
- sklearn.model_selection.cross_val_score
- Model evaluation, selection.cross_val_score’s scoring parameter
- Using [scoring="neg_mean_squared_error"] for [cross_val_score] means the scores that are returned are from the negated MSE. But, each of the k-fold fits will use the regressor’s internal fitting score to minimize. For example, Lasso will still minimize MSE + alpha*L1(theta), LinearRegression will minimize MSE and RidgeRegression will minimize MSE + alpha * L2(theta). However, [cross_val_score] won’t report their internal scores. Instead it just reports MSE.
- To do your own customized scoring function you can pass a callable object that calculates the score to report. Be careful though. You want to be able to directly compare the scores for the results of different models.