How to Choose Alpha and Lambda

For now, we will focus on trial and error. We touched on using cross validation in Section 5.2 and we will be using it here to select the best values for alpha and lambda.

1. Split the data into training and test sets (80%/20% or otherwise).

2. Split the training data into *n* folds (for *n*-fold cross validation).

3. Select a list of values you want to test, e.g., lambda= (0.01, 0.1, 0.5, 1) and alpha = (0, 0.25, 0.5, 0.75, 1). Note that if only performing ridge or lasso regression, you only need to worry about lambda.

4. For each pair of values (or single value for lasso or ridge):
   A. Train the model on (*n*–1) folds of the training data using those values.
   B. Calculate the prediction error on the remaining fold of training data.
   C. Repeat *n* times for each fold in the training data (cross validation).
   D. Calculate the average error rate across all iterations in the cross validation (cross validation error).

5. Take the pair of values that gives the lowest cross validation error.

6. Train a model on all of the training data using the optimal hyperparameters.

7. Test the model on the test data.