# ISYE 8803 Homework 4 Problem 2

Nick DiNapoli, ndinapoli6@gatech.edu

July 10, 2022

## 1 Problem 2

In this problem, I analyze a political economy dataset. Specifically, I aim to predict the democracy index using a host of demographic features, namely, the logarithm of real GDP as well as see which features are most imperative for this prediction. The dataset used contains this index as well as each of the features for 195 countries from 1960 to 2000 (data points are taken every 5 years). I pool the data together and split it into training and testing sets (80%/20% split). Additionally, I standardize the data before analysis using several regression models. With that said, I make predictions and draw conclusions using four regularized regression models: ridge regression, lasso regression, adaptive lasso regression, and elastic net regression.

### 1.1 Ridge regression

First, I vary the lambda value from 1e-9 to 1e3 and perform 5-fold cross validation to get a sense of where the optimal lambda value lies. In order to complete this I determine the RSS for each model and look for the index of the lambda value that minimizes the RSS (called alpha in this library). A plot of this is shown in Figure 1. After ensuring this, I use 5-fold cross-validation and fit the model using the training dataset. After finding the optimal model (using MSE criterion) I analyze the optimal lambda, coefficients, and the MSE of the test set. These results are summarized in Table 1.

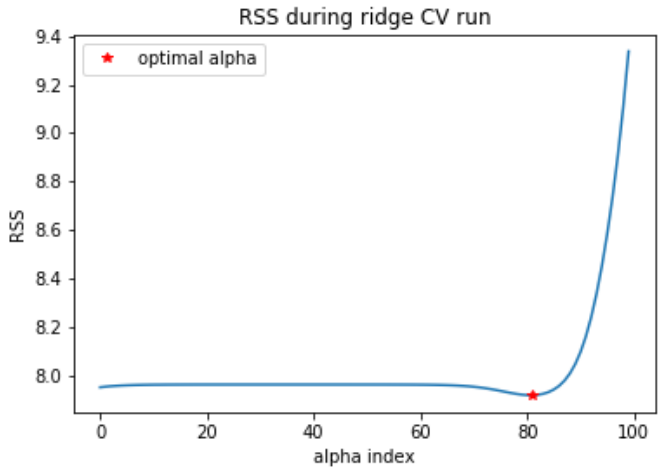


Figure 1: RSS as lambda is varied for ridge regression.

| | Values |
|---|---|
| lambda | 46.42 |
| coefficients | [0.1010, -0.0115, -0.0206, 0.0119, -0.0154, 0.0138, 0.0311, 0.0795, 0.0264, -0.0144] |
| MSE of test set | 0.05887 |

Table 1: Summary of optimal values for ridge regression.

## 1.2 Lasso regression

Next, I complete an identical task but use lasso regression. Every step is completed in a similar manner. Figure 2 shows my initial trial to search for the region with optimal lambda. Once again the optimal values are summarized in Table 2. Several coefficients tend to 0 as expected which is what I want to see.

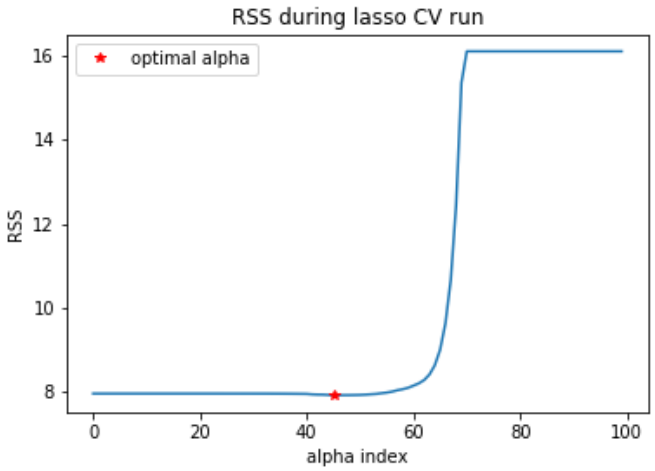


Figure 2: RSS as lambda is varied for lasso regression.

| | Values |
|---|---|
| lambda | 0.00811 |
| coefficients | [0.11945729, -0.00039932, -0.00344206, 0, 0, 0, 0.04488749, 0.08344771, 0, -0.00933468] |
| MSE of test set | 0.05912 |

Table 2: Summary of optimal values for lasso regression.

## 1.3 Adaptive lasso regression

The third model I analyzed was adaptive lasso regression using OLS. The results are shown in Table 3. Interestingly, the model picked out just one feature to keep and essentially forced all to 0 because of the large regularization parameter. The model also did not do a good job with prediction on the test set.

| | Values |
|---|---|
| lambda | 2567104 |
| coefficients | [0, 0, -3.2954e-10, 0, 0, 0, 0, 0, 0, 0] |
| MSE of test set | 0.55393 |

Table 3: Summary of optimal values for adaptive lasso regression.

### 1.4 Elastic net regression

Lastly, I complete elastic net regression. For this task I complete a parameter search not only on lambda but alpha as well (ratio of ridge to lasso regularization). Lambda is varied in the same way as the previous sections and alpha is varied from 0 to 1. Figure 3 shows a similar search for the rough region where lambda is optimal. As it can be seen, the plot follows lasso regression closely and this makes sense as the lambda parameter was found to be .99. Table 4 shows the summary of results. The results show to be very similar to lasso regression. Overall, the results for all the methods except the adaptive lasso show to be very similar. Basing decisions purely off MSE one might select ridge regression as the best for predicting democracy index but because the test set results are so similar, it is common place to select the most simple model in hopes that it could generalize well and lead to fast runtimes. Based on this approach, one should select lasso regression as it forces many of the coefficients to zero with a simple model (more so that elastic net) and provides competitive MSE on the test set.

| | Values |
|---|---|
| lambda | 0.00811 |
| alpha | .99 |
| coefficients | [0.1196917, -0.00043234, -0.00282838, 0, 0, 0, 0.04528941, 0.08350334, 0, -0.00941161] |
| MSE of test set | 0.05910 |

Table 4: Summary of optimal values for elastic net regression.