The Title

A. Author

December 4, 2016

1 Analysis

1.1 Modelling

In the Results section of the report, we determined the mean squared error values for each of our 3 models (one model with every variable from the data included and one where we preselected 10 variables). For the full Ridge, Lasso, and PCR the MSE's were 0.186, 0.195, and 0.271 respectively. For the shor Ridge, Lasso, and PCR the MSE's were 12.797, 17.114, and 1.446 respectively. Because our data was mean centered and standardized, all of these values exist on the same scale and can thus be prepared. Mean squared error values represent the average sum of the errors between actual and predicted response values. Therefore, the smaller the MSE value the smaller the error, and the better the predictive model. Given this, 0.186 is our smallest MSE value meaning Ridge is our best predictive model. These results are unsurprising to an extent, each of the MSE's for the shortened models were much higher, meaning that preselecting predictors negatively effects the modelling process.

In order to determine which variables most affected unemployment rates we look at the absolute values of the coefficients from our best model (ridge) and look at the top 5.

	Predictor	Coefficient
1	POVERTY_RATE	0.72
2	PCT_WHITE	-0.42
3	PCT_BA	-0.18
4	PCT_BLACK	-0.15
5	MARRIED	-0.13

Table 1: Information about top 5 Ridge Model Coefficients

1.2 Hypothesis Testing