HW 3

Elizabeth Hutton

September 12, 2018

1 5.7 Applied

a. Logistic Regression Model: $Direction \sim Lag1 + Lag2$

b. Logistic Regression Model excluding first observation: $Direction[-1] \sim Lag1[-1] + Lag2[-1]$

c. The model from part b. incorrectly predicts ”Up” for the first observation, using a cutoff of 0.5.

d./c. The LOOCV test error is 0.504.

2 5.6 Applied

a. Standard error estimates for the coefficients *balance* and *income* from the logistic regression model of $Default \sim income + balance$:

```
 (Intercept)       income      balance
4.347564e-01 4.985167e-06 2.273731e-04
```

b./c.

```
> #bootstrap function for logistic regression of default ~ income + balance
> #returns estimated coefficients for intercept, income, and balance
> boot.fn = function(data,index){
+   return (coef(glm(default ~ income + balance,
+           data = Default,family = binomial,subset=index)))
+ }
> #1000 bootstrap implementations
> #returns bootstrapped coefficients and standard errors
> library(boot)
> #boot(Default,boot.fn,1000)
```

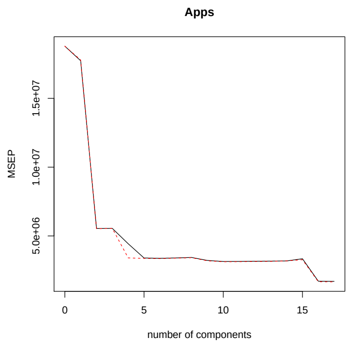
d. The bootstrap function did not produce significantly different estimates for the standard errors of coefficients, suggesting that the coefficients of the regression model are reliable.

3 6.9 Applied

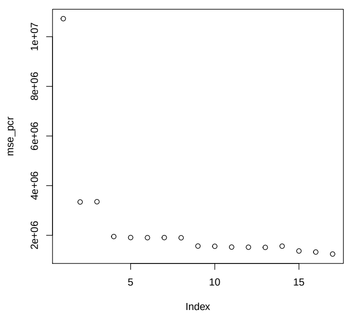
a./b. Test MSE for the linear model *Apps* $\sim$. is 1520331.5.

c. A 10-fold cross-validated ridge regression model *Apps* $\sim$. with $\lambda = 450.74$ has a test MSE of 1045387.1, an improvement over the linear model.

d. A 10-fold cross-validated lasso model with $\lambda = 32.55$ gives a test MSE of 1040223.2 and 3 out of the 17 predictors have coefficients equal to 0.

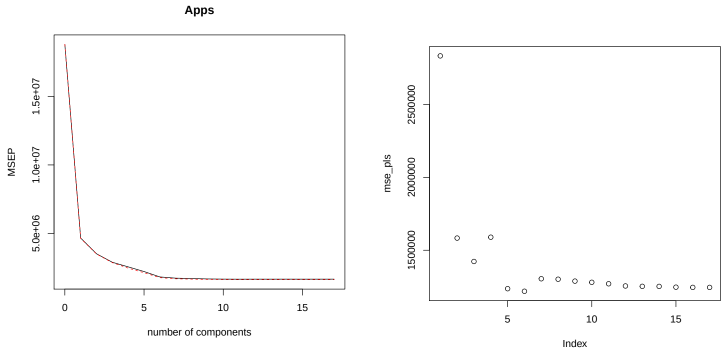


(a) PCR Results



(b) MSE vs. M components

e. Performing PCR on the training data for various M number of components yields the above results. From the plot of training MSEP vs. M, the lowest error is produced for M=16, but that is only 1 fewer than the total number of predictors. However, it appears that there is a steep drop in error that stays mostly steady from M=5 to M=15, indicating that as few as 5 components can capture most of the variation in the model. Plotting test MSE vs. the number of components also shows a slight dip in error at M=9 with a test MSE = 1566614.9.



(a) PLS Results

(b) MSE vs. M components

f. Both the plots of training MSEP and testing MSE vs. M number of components show a clear drop in error for M=6 components. Test MSE for 6 components is 1218544.8.

g. None of these methods seem to do a very good job of predicting the number of applications a college will receive, that is, they all have high test MSE's. There is not too big of a difference between the performance of these different models, although lasso seemed to perform best with the highest interpretability.

4 6.11 Applied

- Ridge regression: 35.07
- Lasso model: 38.94 with 7 coefficients
- PCR: 42.26 with $M = 5$ components
- PLS: 40.94 with $M = 9$ components

The above statistics present the test MSE for four 10-fold cross-validated models. Interestingly, ridge regression has the lowest error. From previous explorations of the dataset, I noticed that many of the variables were correlated with crime rate, so it is reasonable to believe that a model which includes all the predictors would perform well.