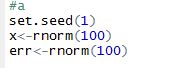
**HW 4 Chapter 6**

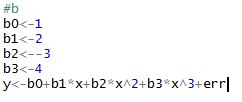
Luyao Zhang (NetID: lzhang94)

**Ex 8**

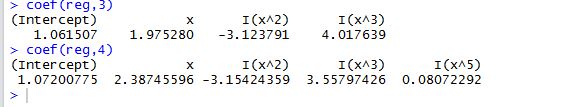
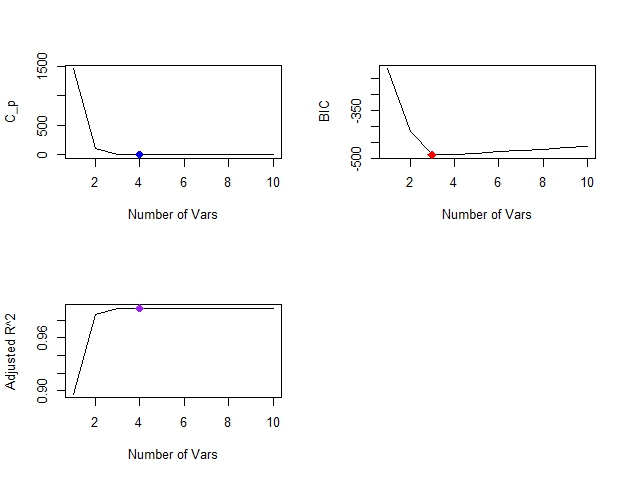
1. The code for generating simulated data is as below:



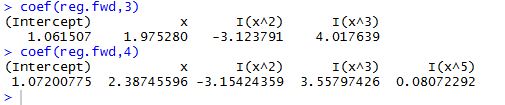
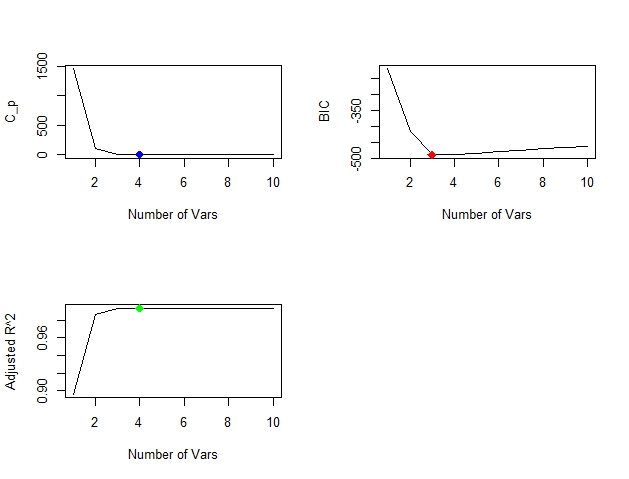
1. The code for generating the response vector Y with coefficients of my choice is as below. In my regression model, .



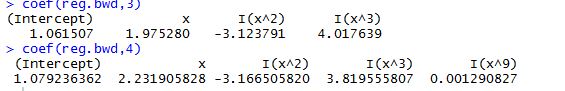
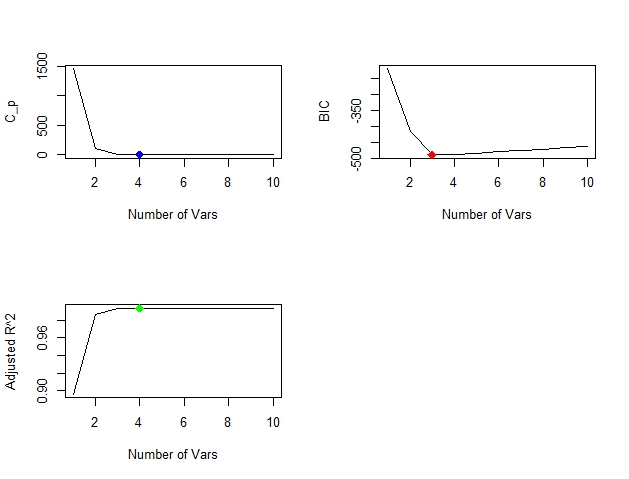
1. The 3 plots are as below.  and adjusted R^2 choose the 4-variable model, with predictor being x, x^2, x^3, and x^5. BIC chooses a 3-variable model, which is more accurate, with predictors being x, x^2, and x^3.



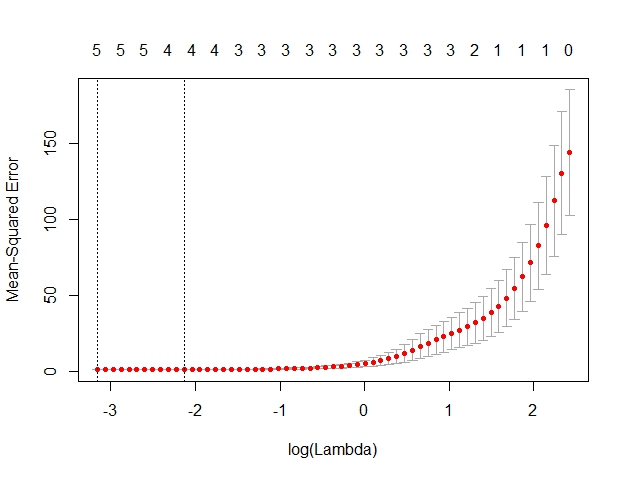
1. Results for forward selection are as below. Forward selection results are the same as what was obtained in (c), which is that  and adjusted R^2 support a 4-variable model, with predictors being x, x^2, x^3, and x^5. BIC supports a more accurate 3-varaiable model, with predictors being x, x^2, and x^3.



Results for backward selection are as below. Backward selection results are the same as what was obtained in (c), which is that  and adjusted R^2 support a 4-variable model, with predictors being x, x^2, x^3, and x^5. BIC supports a more accurate 3-varaiable model, with predictors being x, x^2, and x^3.



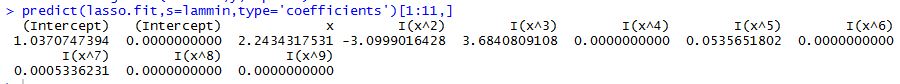
1. The plot of the cv error as a function of lambda is as below:



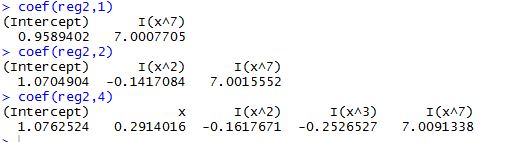
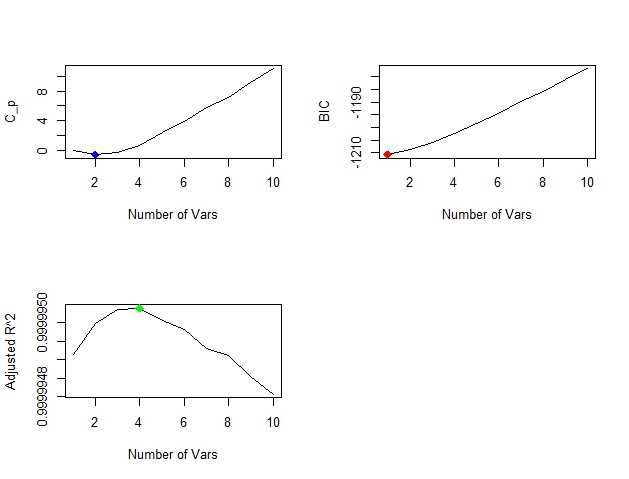
The optimal value of lambda, according to cross-validation, equals 0.0426.



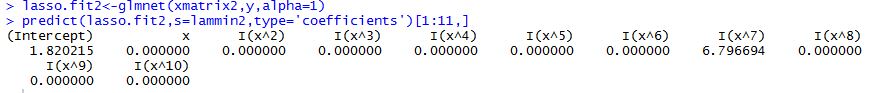
Predictor selection based on the lasso method returns x, x^2, x^3, x^5, and x^7 as appropriate predictors for the model.



1. Plots are as below. Turns out BIC supports a more accurate 1-variable model with the single predictor being x^7.

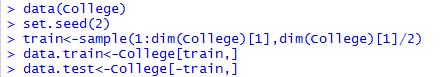


Using the lasso and cross-validation, we also obtained the 1-variable model, where x^7 is the only predictor, and the coefficient is also close to 7.

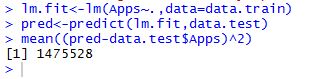


**Ex 9**

1. The code is as below:



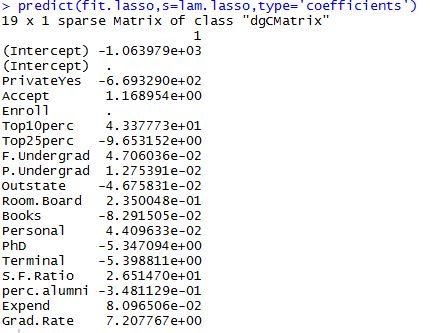
1. The mean square error obtained is 1475528.



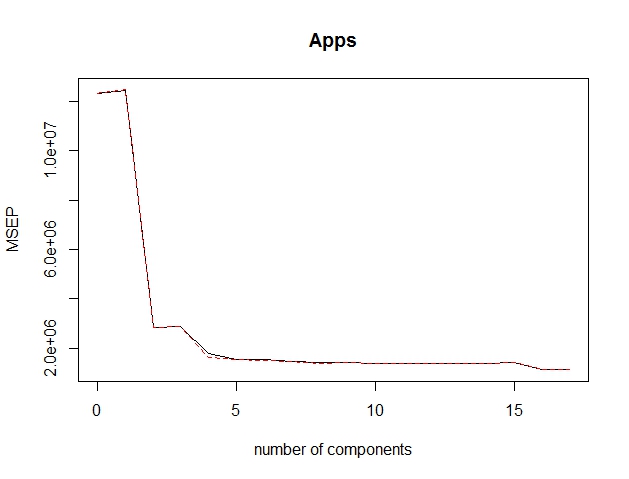
1. Test error for the ridge model is 1575990, which larger than the linear model.



1. The mean square error obtained using lasso is 1555355, which is still larger than the linear model. There are 17 non-zero coefficients estimates.



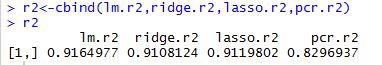
1. The plot is as below:



The mean square error is 3009399, which is larger than least square.

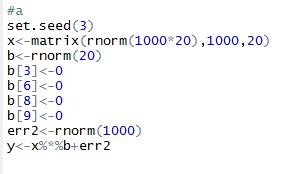


1. The accuracy of prediction using different approaches in terms of R^2 is shown below. Looks like pcr predicts the least accurately, while all the other 3 predict almost equally well.

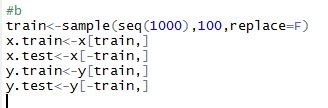


**Ex 10**

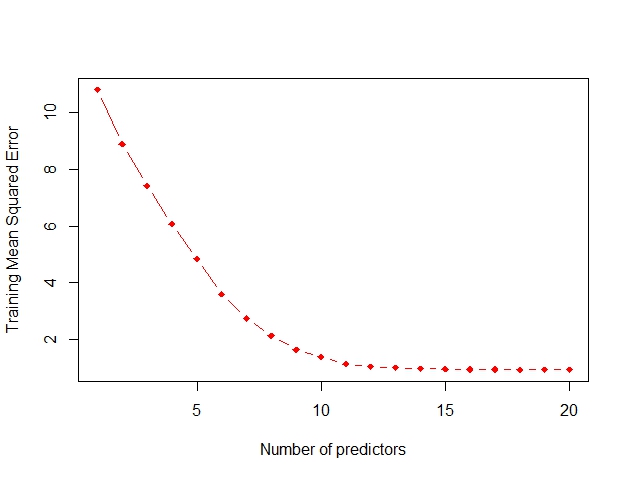
1. The data was generated as below:



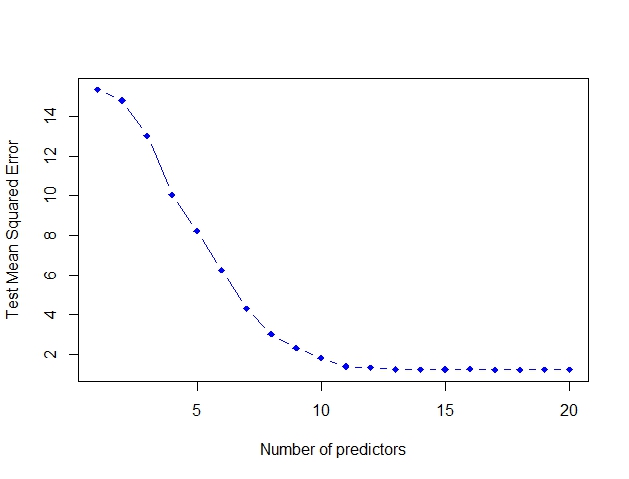
1. The code for splitting the data is as below:



1. The plot of training set MSE associated with the best model of each size is as below:



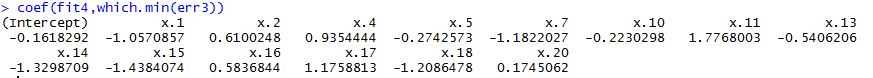
1. The plot of test MSE associated with the best model of each size is as below:



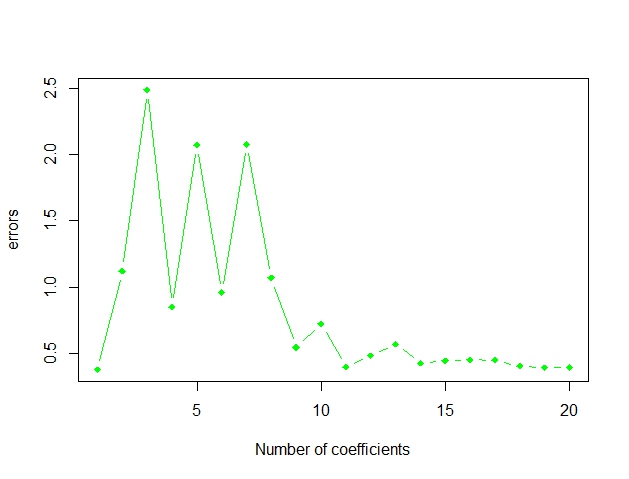
1. The 14-variable model has the smallest test mean squared error.



1. The coefficients are as below. The best model catches most of the 16 non-zero coefficients (all but x. 3, 6, 8, and 9), except for x. 12, and x.19.

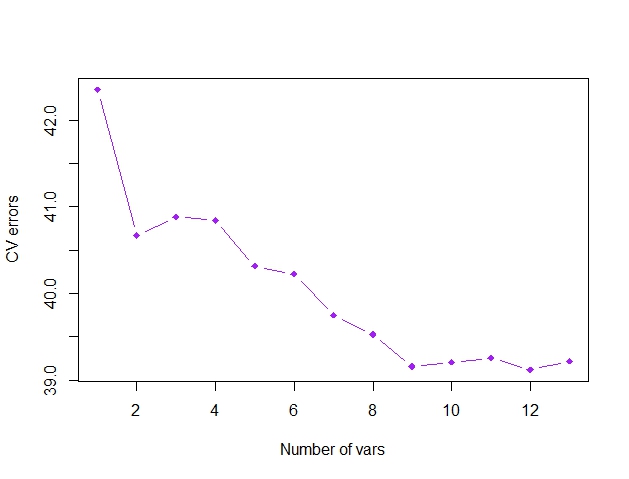


1. The plot is as below. The plot shows that a 1-variable model has the smallest error between true and estimated coefficient. However, according to (d), a 14-variable has the smallest test mean squared error. This discrepancy shows that being closer to the true coefficients does not guarantee a lower test mean squared error.

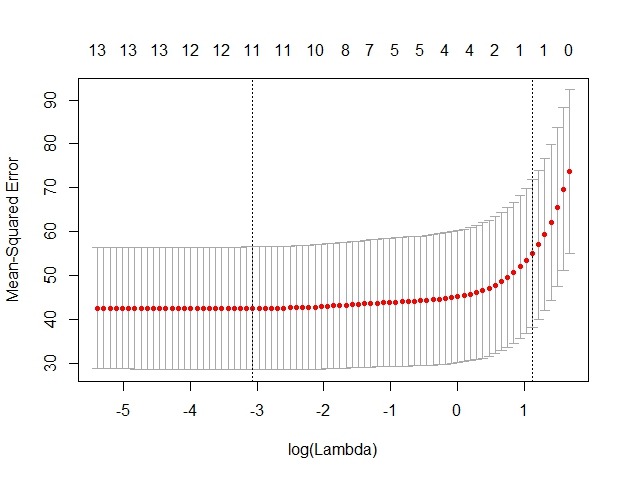
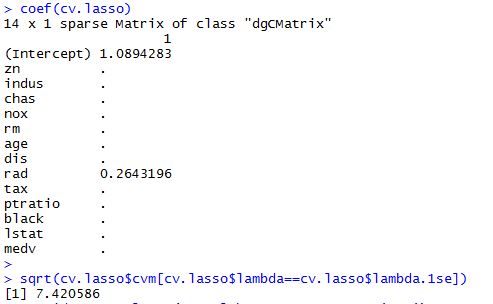


**EX 11**

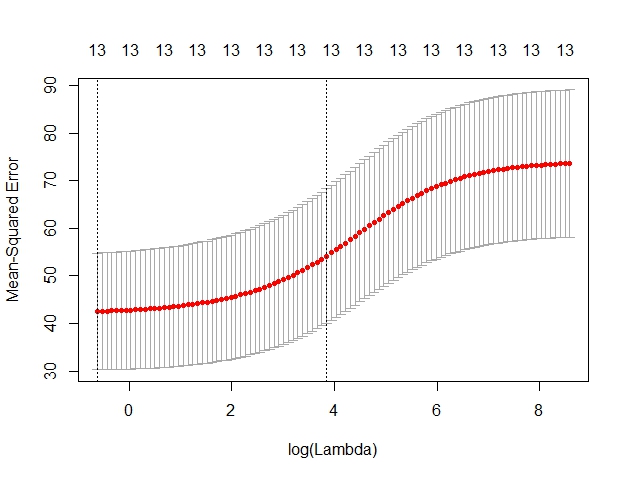
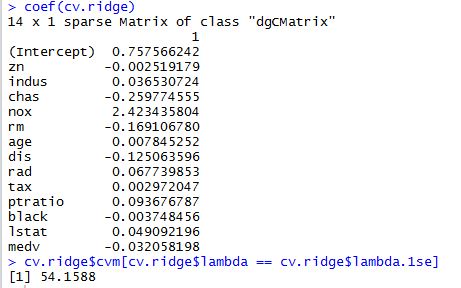
1. Cross-validation returns a 12-variable model. And the test mean squared error is 39.119.



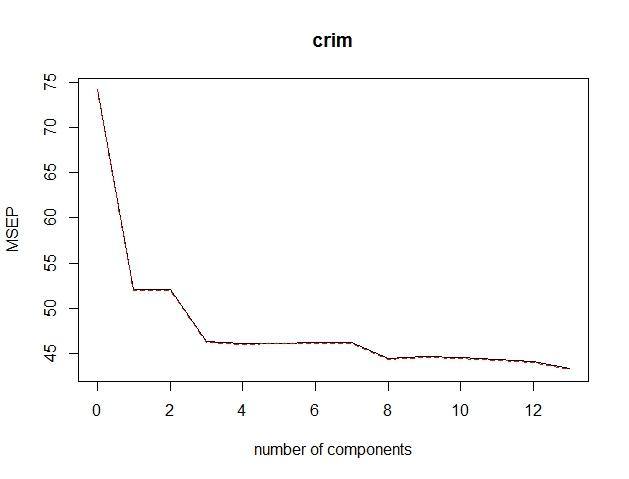
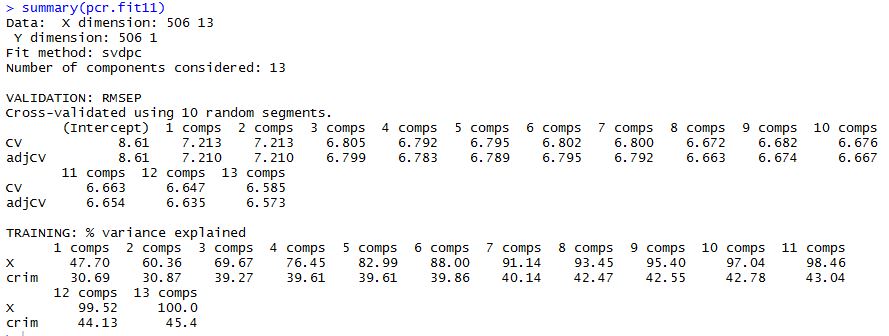
Under lasso, we obtained the plot as below. According to the plot, the selected model has a mean squared error equal to 55.06 (7.42^2).



Under ridge regression (as the plot below), the model selected has a mean squared as large as 54.16.



Under pcr, we obtained the summary of the model and the plot. The model selects a 14-variable model, which means there’s no dimension reduction.



1. Considering the results above, the best subset selection cross-validation yields the best model, which has effectively reduced dimensions and has the smallest mean squared error.
2. No. The best model includes only 12 predictors, instead of including all 13 of them.