MATH 427- Homework-5-Spring 2023

1 Exercise-Using R

This question should be answered using the Credit data set in Canvas.

(a) Use the *regsubsets()* function to perform best subset selection in order to choose the best model that predict *Balance* using other variables (in the data) as predictors. What is the best 4 variables model? Display R outputs to justify your answer.

(b) In part (a), what is the best model obtained according to each C_p, BIC and adjusted R^2. Show some plots to provide evidence for your answer, and report the coefficients of the best model obtained from each criterion.

(c) Repeat part (a) using forward step-wise selection. What is the best 4 variables model? Display R outputs to justify your answer.

(d) How does your answer in part (a) compare to the results in part (c)?

(e) In part (c), what is the best model obtained according to each C_p, BIC and adjusted R^2. Show some plots to provide evidence for your answer, and report the coefficients of the best model obtained from each criterion.

(f) Split your data set into a training set containing 200 observations and a test set containing 200 observations. Fit a linear regression of the best 4 variables model found in part (a) using only the observations corresponding to the training set.

(g) Compute the estimated test MSE for the linear regression fit in part (f).

(h) Repeat (f) and (g) using LOOCV and using K-FOLD (Use 10 folds). Which of the 3 methods gives the smallest test MSE?

***We want to select among all possible models using validation set approach
or Cross-validation.***

(i) Split the data into training set (200-obs) and a test set. Use the *regsubsets()* function
to perform best subset selection in order to choose the best model that predict *Balance* using
other variables (in the data) as predictors. Compute the test set MSE associated with the
best model of each size. Which size model has the smallest test MSE?

(j) Repeat (i) using K-FOLD Cross-validation.

(k) How do your results in (i) and (j) compare to the results in (b)?