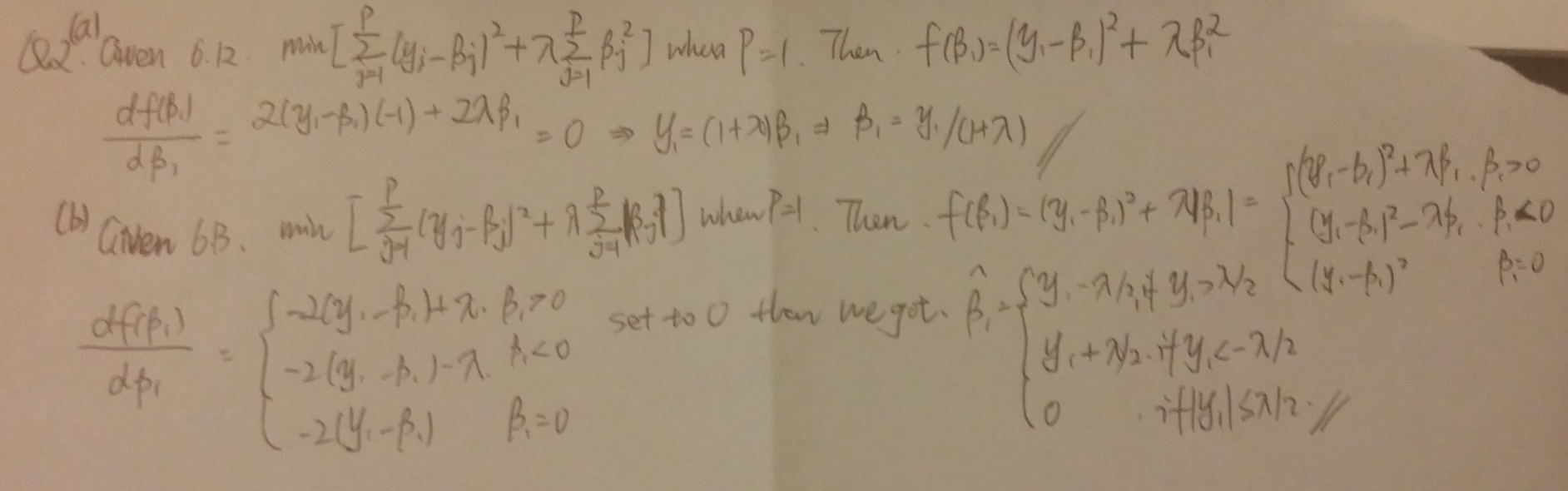
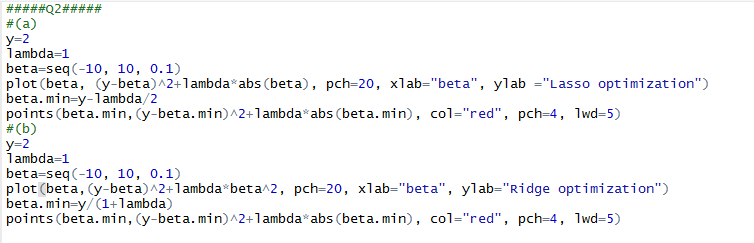
**Q2. We will now explore (6.12) and (6.13) further.**

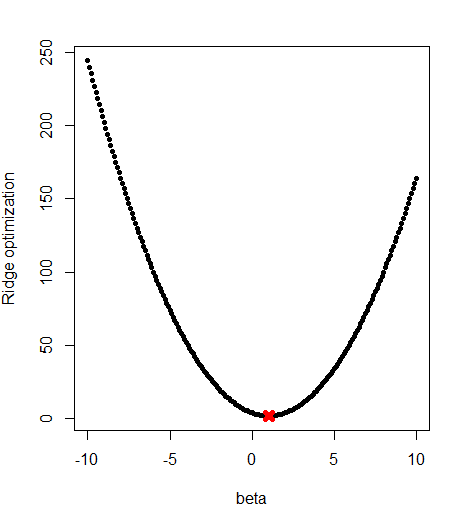
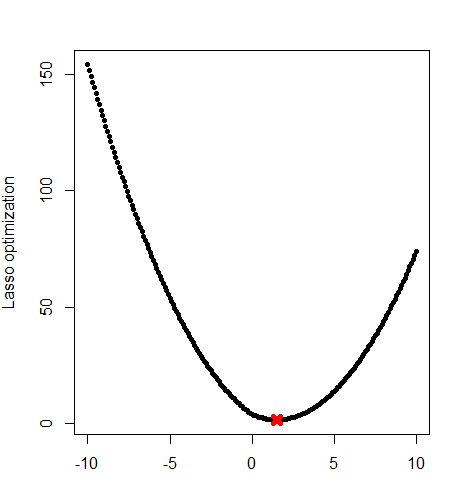
1. **Consider (6.12) with p = 1. For some choice of y1 and λ > 0, plot (6.12) as a function of β1. Your plot should confirm that (6.12) is solved by (6.14).**
2. **Consider (6.13) with p = 1. For some choice of y1 and λ > 0, plot (6.13) as a function of β1. Your plot should confirm that (6.13) is solved by (6.15).**

**Ans:**



By following codes, we can verify the formulas give us the right β1 values to optimize the method of Lasso and Ridge, respectively.





**Q3. In this exercise, we will generate simulated data, and will then use this data to perform best subset selection.**

1. **Use the *rnorm()* function to generate a predictor X of length n = 100,**

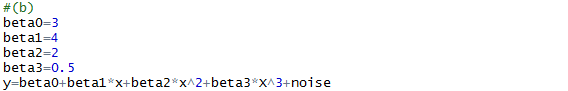
**as well as a noise vector of length n = 100.**



1. **Generate a response vector Y of length n = 100 according to the model**

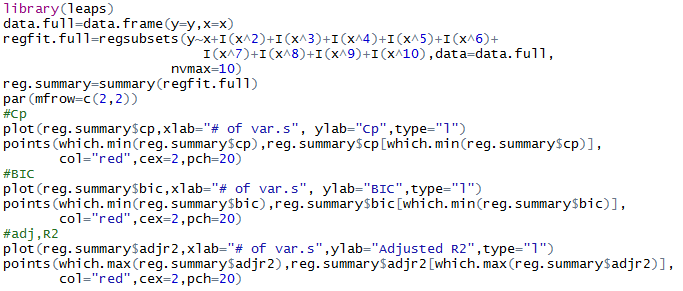
**Y = β0 + β1X + β2X2 + β3X3 + ε**

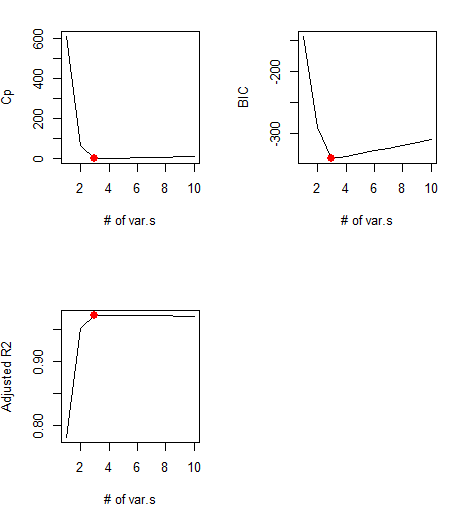
**where β0, β1, β2, and β3 are constants of your choice.**



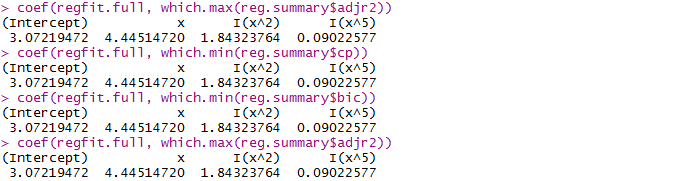
**Ans:** we set beta0 = 3, beta1 = 4, beta2 = 2, beta3 = 0.5.

1. **Use the *regsubsets()* function to perform best subset selection in order to choose the best model containing the predictors X, X2, . . . , X10. What is the best model obtained according to C*p*, BIC, and adjusted R2? Show some plots to provide evidence for your answer, and report the coefficients of the best model obtained. Note you will need to use the *data.frame()* function to create a single data set containing both X and Y .**





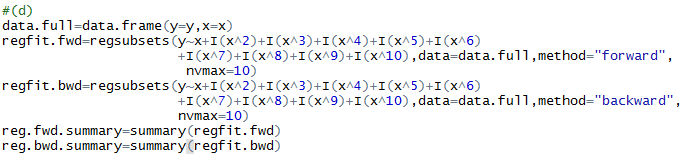
**Ans:** All the three index indicate that 3 variables (x, x2, x5) are the best fitting model here. The chosen best variables and their coefficients are shown as follows.



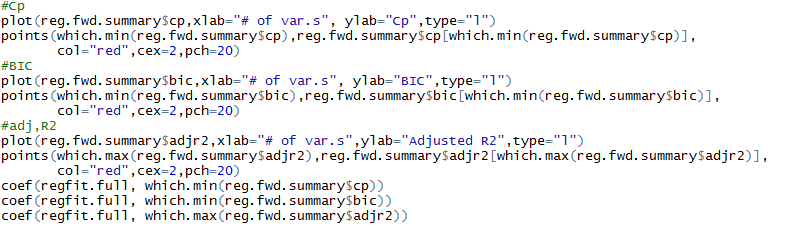
1. **Repeat (c), using forward stepwise selection and also using backwards stepwise selection. How does your answer compare to the results in (c)?**

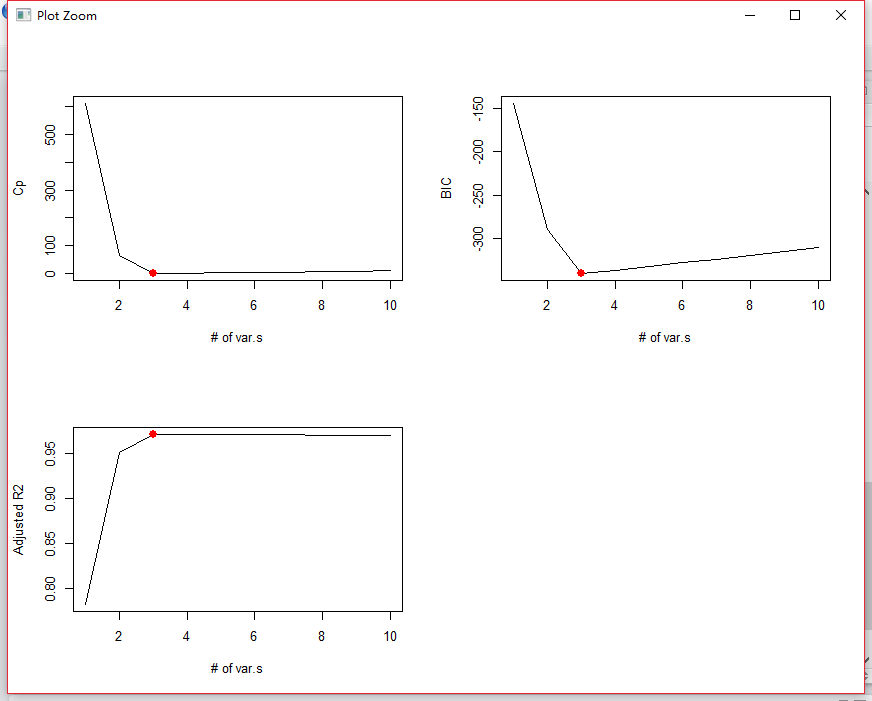
**Ans:**

**Codes for forward and backward selection.**

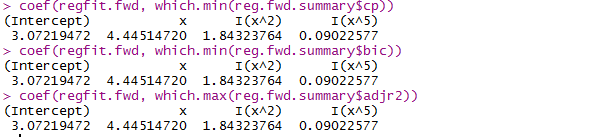


**Forward method summary:**

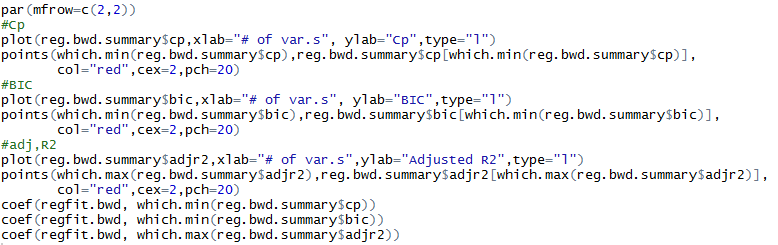


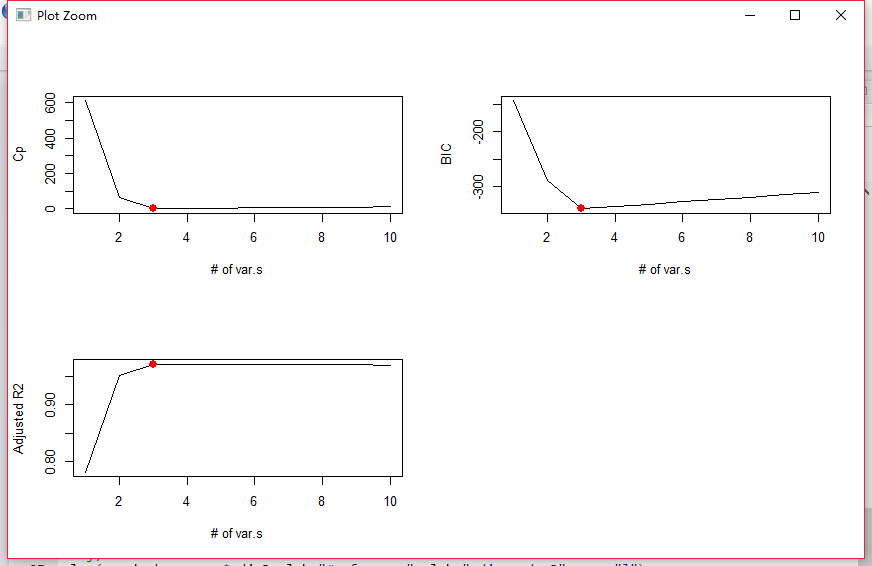


All the three index indicate that 3 variables (x, x2, x5) are the best fitting model here. The chosen best variables and their coefficients are shown as follows.

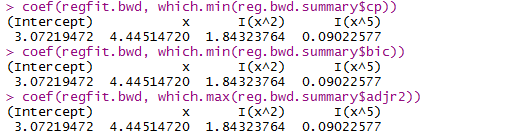


**Backward method summary:**



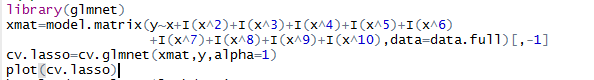


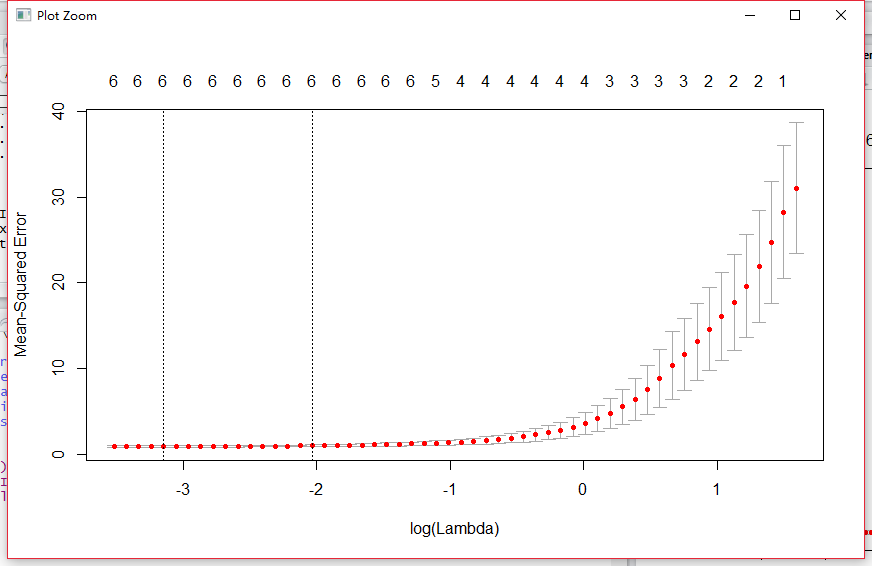
All the three index indicate that 3 variables (x, x2, x5) are the best fitting model here. The chosen best variables and their coefficients are shown as follows.



In these variable selection methods, all three methods select same variables (x, x2, x5).

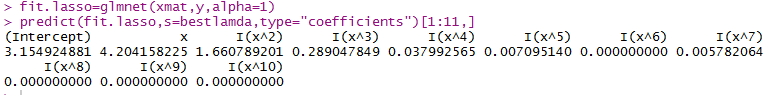
1. **Now fit a lasso model to the simulated data, again using *X, X*2*, . . . , X*10 as predictors. Use cross-validation to select the optimal value of *λ*. Create plots of the cross-validation error as a function of *λ*. Report the resulting coefficient estimates, and discuss the results obtained.**







**Ans:** Our model’s best lamda value is 0.0429 according to the cv value. By this value we can have the model with coefficients as follows. LASSO selects 6 variables (x,x2,x3,x4,x5,x7).



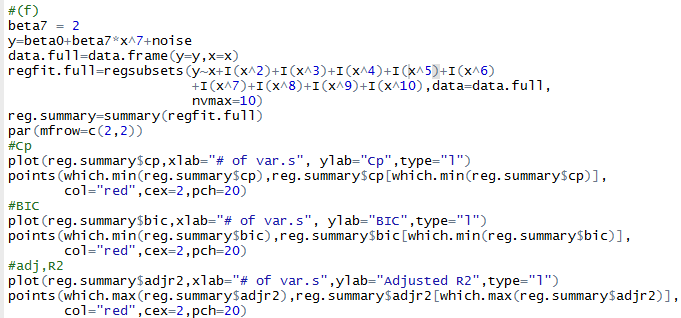
1. **Now generate a response vector *Y* according to the model**

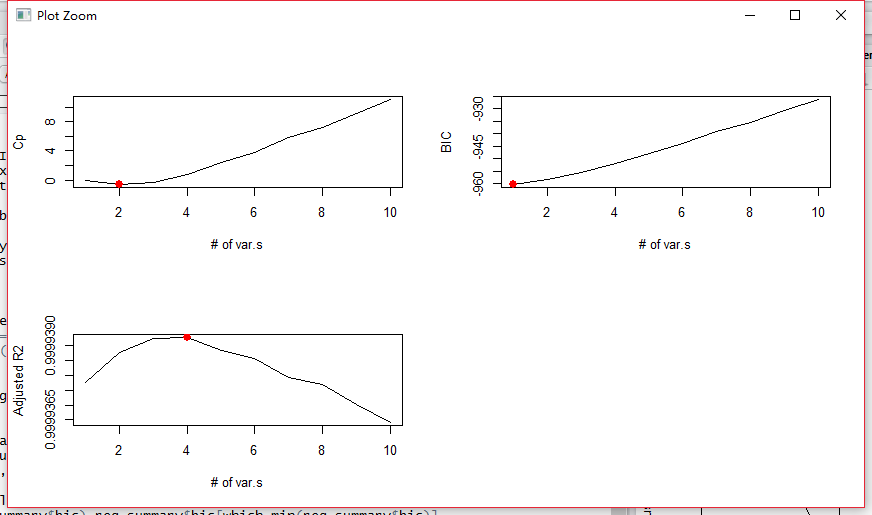
**Y *=* β0 *+* β7X7 *+ ε*,**

**and perform best subset selection and the lasso. Discuss the results obtained.**

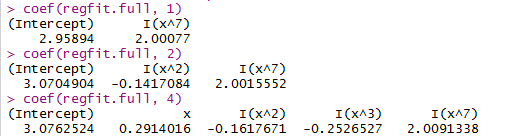
**Ans:**

**Best subset:**

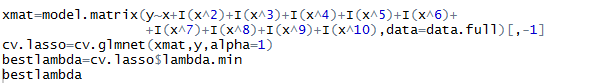


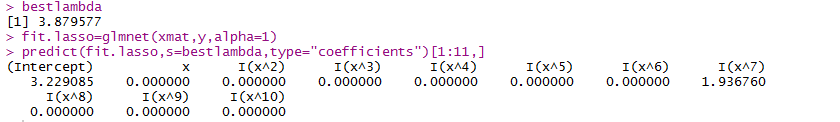


From best subsets method, three index do not give us the same number of variables. Cp implies 2 variables; BIC gives 1 variable, adjusted R2 uses 4 variables. The detail is shown below.



**LASSO:**





By the method of LASSO, we achieve the best model matching to our simulated question. In this case, LASSO performed better than best subset method.