# **Assignment #5**

#### **Problem 1**

In this question, we will predict the number of applications received (Apps) using the other variables in the College data set (ISLR package).

```
> library(ISLR)
> dim(College)
[1] 777 18
 names (College)
[1] "Private"
[9] "Outstate"
                   "Apps"
                                                "Enroll"
                                                              "Top10perc"
                                                                             "Top25perc"
                                                                                           "F.Undergrad" "P.Undergrad"
                                  "Accept"
                    "Room.Board"
                                                "Personal"
                                                                             "Terminal"
                                                                                           "S.F.Ratio"
                                                                                                        "perc.alumni"
[17] "Expend"
                   "Grad.Rate"
 summary(College)
Private
                                                Enroll
                                                                              Top25perc
                                                             Top10perc
                Apps
                               Accept
                                                                                             F. Undergrad
No :212
                                                  : 35
           Min.
                           Min.
                                      72
                                           Min.
                                                           Min.
                                                                 : 1.00
                                                                            Min.
                                                                                            Min.
           1st Qu.:
 Yes:565
                           1st Qu.:
                                     604
                                            1st Qu.: 242
                                                           1st Qu.:15.00
                                                                            1st Qu.: 41.0
                                                                                                      992
           Median : 1558
                           Median : 1110
                                           Median: 434
                                                           Median :23.00
                                                                           Median: 54.0
                                                                                            Median: 1707
                 : 3002
                                                  . 780
                                  : 2019
                                                                                  : 55.8
           Mean
                           Mean
                                           Mean
                                                           Mean : 27.56
                                                                            Mean
                                                                                            Mean
           3rd Qu.: 3624
                           3rd Qu.: 2424
                                            3rd Qu.: 902
                                                           3rd Qu.:35.00
                                                                            3rd Qu.: 69.0
                                                                                            3rd Qu.: 4005
                  :48094
           Max.
                           Max.
                                  :26330
                                           Max.
                                                           Max.
                                                                  :96.00
                                                                            Max.
                                                                                   :100.0
                                                                                            Max. :31643
                                                      Books
1. : 96.0
                                                                                       PhD
  P.Undergrad
                      Outstate
                                     Room.Board
                                                                       Personal
                   Min. : 2340
1st Qu.: 7320
                   Min.
                                           :1780
Min.
             1.0
                                                   Min.
                                                                    Min. : 250
1st Qu.: 850
                                                                    Min.
                                                                                    Min
                                                                                             8 00
                                                                                                     Min.
                                                                                                     Min. : 24.0
1st Qu.: 71.0
                                   Min.
            95.0
                                   1st Qu.:3597
                                                   1st Qu.: 470.0
                                                                                    1st Qu.: 62.00
1st Qu.:
Median :
           353.0
                   Median: 9990
                                   Median:4200
                                                   Median : 500.0
                                                                    Median:1200
                                                                                    Median : 75.00
                                                                                                     Median: 82.0
           855.3
                   Mean
                         :10441
                                   Mean
                                          :4358
                                                   Mean
                                                            549.4
                                                                    Mean
                                                                           :1341
                                                                                    Mean
                                                                                             72.66
                                                                                                     Mean
                                                                                                              79.7
 3rd Qu.:
           967.0
                   3rd Qu.:12925
                                   3rd Qu.:5050
                                                   3rd Qu.: 600.0
                                                                    3rd Qu.:1700
                                                                                    3rd Qu.: 85.00
                                                                                                     3rd Qu.: 92.0
        :21836.0
                                                          :2340.0
Max.
                  Max.
                          :21700
                                   Max.
                                          :8124
                                                  Max.
                                                                    Max.
                                                                           :6800
                                                                                    Max.
                                                                                           :103.00
                                                                                                     Max.
                                                                                                            :100.0
  S.F.Ratio
                                                    Grad.Rate
                  perc.alumni
                                     Expend
                                                        : 10.00
 Min.
                       : 0.00
                                        : 3186
                 Min.
                                 Min.
                                                  Min.
 1st Qu.:11.50
                 1st Qu.:13.00
                                 1st Qu.: 6751
                                                  1st Qu.: 53.00
Median :13.60
                 Median :21.00
                                 Median : 8377
                                                  Median : 65.00
                                        : 9660
                       :22.74
Mean
       :14.09
                 Mean
                                 Mean
                                                  Mean : 65.46
 3rd Qu.:16.50
                 3rd Qu.:31.00
                                 3rd Qu.:10830
                                                  3rd Qu.: 78.00
                        :64.00
        :39.80
                 Max.
                                 Max.
                                         :56233
                                                  Max.
Max.
```

(a) Perform best subset selection to the data. What is the best model obtained according to Cp, BIC and adjusted R 2? Show some plots to provide evidence for your answer, and report the coefficients of the best model.

```
> library(leaps)
> fit.max = regsubsets(Apps~..College)
  fit.max = regsubsets(Apps~.,College)
fit.max = regsubsets(Apps~.,data=College,nvmax=17)
  fit.summary = summary(fit.max)
> fit.summary
Subset selection object
Call: regsubsets.formula(Apps ~ ., data = College, nvmax = 17)
17 Variables
               (and intercept)
             Forced in Forced out
PrivateYes
                 FALSE
                              FALSE
                 FALSE
Accept
                              FALSE
Enroll
                  FALSE
                              FALSE
Top10perc
                  FALSE
                              FALSE
Top25perc
                  FALSE
                              FALSE
F. Undergrad
                  FALSE
                              FALSE
P. Undergrad
                  FALSE
                              FALSE
Outstate
                  FALSE
                              FALSE
Room.Board
                  FALSE
                              FALSE
Books
                  FALSE
                              FALSE
Personal
                  FALSE
                              FALSE
PhD
                  FALSE
                              FALSE
Terminal
                  FALSE
                              FALSE
S.F.Ratio
                  FALSE
                              FALSE
perc.alumni
Expend
                  FALSE
                              FALSE
Grad.Rate
                  FALSE
                              FALSE
```

```
1 subsets of each size up to 17
Selection Algorithm: exhaustive
           PrivateYes Accept Enroll Top10perc Top25perc F.Undergrad P.Undergrad Outstate Room.Board Books Personal
                                                     0 = 0
                                                                                             0.0
   (1)
          0.0
3
  (1)
                                                                 0.0
                                                                               .....
           . .
                                .....
                                                                                                         .....
                                                                                                                              .....
   (1)
           .....
                        11 ½ 11
                                пұп
                                         11 ½ 11
                                                     . .
                                                                 0.0
                                                                               0.00
                                                                                              11 ½ 11
                                                                                                         . .
                                                                                                                       . .
                                                                                                                             0.0
   (1)
                                                                 (1)
           .....
                        11 ½ 11
                                11 18 11
                                         11 1/2 11
                                                     .....
                                                                               . . . . . . . . . . .
                                                                                              11 1/2 11
                                                                                                         11 1/2 11
                                                                                                                       0.00
           .....
                        11 1/2 11
                                n y n
                                         11 1/2 11
                                                     11 14 11
                                                                               0.00
                                                                                             11 1/2 11
                                                                                                         11 1/2 11
                                                                               0.00
           n w n
                                                     0.0
                                                                                                         n y n
                                                                                                                      0.0
                        0.80
                                 0.36.0
                                          11 1/2 11
                                                                                             11 1/2 11
                                                     11 18 11
                                                                                                         11 ½ 11
          11 % 11
                        11 1/2 11
                                11 1/2 11
                                          11 1/2 11
                                                                                             11 1/2 11
    (1) "*"
                        11 ½ 11
                                          11 12 11
                                                     11 ½ 11
                                                                                              11 ½ 11
                                                                                                         11 ½ 11
                                 11 1/2 11
10
                                                                 11 g 11
                                                                               0.0
    (1) "*"
                         11 1/2 11
                                          11 ½ 11
                                                     11 1/2 11
                                                                                              11 ½ 11
                                                                                                         11 ½ 11
11
                                                                 "*"
                                                                               11 18 11
12
                                         11 14 11
                                                     " * "
                                                                               11 1/2 11
                                                                                                         " * "
13
    (1) "*"
                        11 ½ 11
                                 11 1/2 11
                                         11 1/2 11
                                                     11 % 11
                                                                 11.44.11
                                                                               11 1/2 11
                                                                                              11 ½ 11
                                                                                                         11 ½ 11
                                                                                                                      .....
                                                                                                                             ......
14
                                                                               "*"
    (1)"*"
                                                                 11 ½ 11
11 ½ 11
                                                                                                         11 ½ 11
                        11 1/2 11
                                 n g n
                                         n g n
                                                     11 1/2 11
                                                                                              11 1/2 11
                                                                                                                      0.00
                                                                                                                             0.60
15
    (1)"*"
                                                                                                         11 1/2 11
16
                        0.40
                                0.50
                                         11 1/2 11
                                                     11 1/2 11
                                                                                              11 1/2 11
                                                                                                                      11 1/2 11
                                                                                                                              0.50
                                                                                                         " ½ "
    (1) "*"
                         11 ½ 11
                                 11 1/2 11
                                         11 1/2 11
                                                                 11 1/2 11
                                                                                              11 1/2 11
                                                                                                                       11 1/2 11
                                                                                                                              11 1/2 11
            PhD Terminal S.F.Ratio perc.alumni Expend Grad.Rate
1
   (1)
                        . . .
                                        .....
            . . . . .
                                                        . . .
    (1)
            . . . . .
                            .....
                                         . .
3
           . . . . .
                            .....
                                         .....
    (1)
           0 0 0 0
                            0.0
                                        0.00
                                                        11 yk 11
                                                                 ......
5
    (1)
    (1) """"
                            0.00
                                         0.0
6
            . . . . .
                            0.0
                                        0.0
                                                        11 % 11
    (1)
            n*n n n
                            0.0
                                        0.0
                                                        11 % 11
8
    (1)
    (1) "*" "
                            0.0
                                        0.00
                                                        11 18 11
9
     (1) "*" " "
                             .....
                                         .....
10
    (1) "*" " "
                            0.00
                                        0.0
                                                        11 ½ 11
11
     · ī ′ "*" " "
                             0.0
                                         0.0
                                                        11 1/2 11
12
     (1) "*" " "
                             11 ½ 11
                                         0.0
                                                        11 14 11
13
    (1) "*" "*"
                             11 1/2 11
                                         . .
                                                        11 1/2 11
14
15 (1) "*" "*"
                             11 ½ 11
                                         ....
                                                        11 1/2 11
                                                                 11 1/2 11
16 (1) "*" "*"
                                         .....
                                                        11 1/2 11
                                                                 n ķ n
    (1) "*" "*"
                                         11 18 11
                                                        11 1/11
17
> names(fit.summary)
[1] "which" "rsq"
                           "rss"
                                    "adjr2" "cp" "bic"
                                                                         "outmat" "obj"
```

## → Finding best models:

```
1.
```

```
> par(mfrow=c(2,2))
> plot(fit.summary$rss,xlab="No. of predictors",ylab="RSS",type="l")
> |
```

#### 2.

```
> plot(fit.summary$adjr2,xlab="No. of predictors",ylab="Adjusted R^2",type="1")
> which.max(fit.summary$adjr2)
[1] 13
> |
```

This shows that max adjusted R<sup>2</sup> is for model with 13 predictors. Coefficient estimates:

```
> points(13,fit.summary$adjr2[13], col="green",cex=2,pch=20)
> coef(fit.max,13)
  (Intercept)
                PrivateYes
                                  Accept
                                                Enroll
                                                           Top10perc
                                                                        Top25perc
                                                                                    F.Undergrad
                                                                                                  P. Undergrad
-440.74148270 -484.77261885
                             1.58542302
                                          -0.87824288 50.41461998 -14.63667155
                                                                                     0.05762769
                                                                                                   0.04642270
    Outstate
                Room, Board
                                    PhD
                                            S.F.Ratio
                                                             Expend
                                                                        Grad.Rate
                0.14696204 -10.91804823 15.15475056
  -0.08823311
                                                         0.07786425
                                                                       8.58578735
```

3.

```
> plot(fit.summary$cp,xlab="No. of predictors",ylab="Cp",type='l')
> which.min(fit.summary$cp)
[1] 12
> |
```

This shows minimum Cp corresponds to model with 12 predictors. Coefficient estimates:

```
> points(12,fit.summary$cp[12],col="green",cex=2,pch=20)
> coef(fit.max,12)
               PrivateYes
                                             Enrol1
                                                       Top10perc
                                                                    Top25perc
  (Intercept)
                                                                               F. Undergrad
                                                                                            P. Undergrad
                                Accept
-157.28685883 -511.78760196
                            1.58691470
                                        -0.88265385
                                                     50.41131660
                                                                 -14.74735373
                                                                                0.05945481
                                                                                             0.04593068
                                  PhD
                                                       Grad.Rate
    Outstate
               Room.Board
                                             Expend
  -0.09017643
               0.14776586 -10.70502848
                                         0.07246655
                                                      8.63961002
4.
> plot(fit.summary$bic,xlab="No. of predictors",ylab="BIC",type='1')
> which.min(fit.summary$bic)
 [1] 10
> |
```

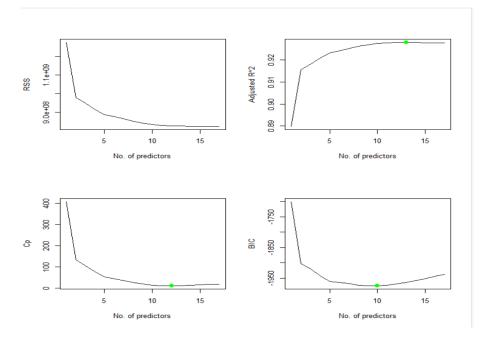
This shows minimum BIC corresponds to model with 12 predictors. Coefficient estimates:

```
> points(10,fit.summary$bic[10],col="green",cex=2,pch=20)
> coef(fit.max,10)
                PrivateYes
                                                           Top10perc
                                                                         Top25perc
  (Intercept)
                                   Accept
                                                Enroll
                                                                                        Outstate
                                                                                                    Room, Board
-100.51668243 -575.07060789
                              1.58421887
                                            -0.56220848
                                                                                                    0.16373674
                                                         49.13908916 -13.86531103
                                                                                     -0.09466457
                                Grad.Rate
                    Expend
                 0.07273776
 -10.01608705
                              7.33268904
> |
```

### Summary and Inference with plots:

Model Selection with:	Criteria	Inference
Adjusted R^2	Select model with maximum	Model with 13 predictors
	value of R^2	has maximum R^2
Ср	Select model with smallest	Model with 12 predictors
	value of Cp	has minimum Cp
BIC	Select model with smallest	Model with 10 predictors
	value of BIC	has the least BIC

## Plots with coefficient (pointed in green):



(b) Repeat (a) using forward stepwise selection and backwards stepwise selection. How does your answer compare to the results in (a)?

## **Forward Stepwise selection:**

```
> fit.fwd=regsubsets(Apps~.,data=College,nvmax=17,method="forward")
Subset selection object
Call: regsubsets.formula(Apps ~ ., data = College, nvmax = 17, method = "forward")
17 Variables (and intercept)
             Forced in Forced out
PrivateYes
                 FALSE
                             FALSE
Accept
Enroll
Top10perc
                 FALSE
                             FALSE
                 FALSE
                             FALSE
                 FALSE
                             FALSE
Top25perc
F.Undergrad
                             FALSE
                 EALSE
                 FALSE
                             FALSE
                 FALSE
P. Undergrad
                             FALSE
Outstate
                 FALSE
                             FALSE
Room.Board
                 FALSE
                             FALSE
Books
                 FALSE
                             FALSE
Personal
                 FALSE
                             FALSE
PhD
                 FALSE
                             FALSE
Terminal
                 FALSE
                             FALSE
S.F.Ratio
                 FALSE
                             FALSE
perc.alumni
                 FALSE
                             FALSE
Expend
                 FALSE
                             FALSE
Grad.Rate
                 FALSE
                             FALSE
1 subsets of each size up to 17
```

```
Selection Algorithm: forward
            PrivateYes Accept Enroll Top10perc Top25perc F.Undergrad P.Undergrad Outstate Room.Board Books Personal
                          n k n
                                            11 1/2 11
                                                        \mathbf{u} = \mathbf{u}
                                                                    0.0
                                                                                   0 = 0
                                                                                                  0 = 0
                                                                                                             0 \leq 0
2
                          11 1/2 11
                                            11 ½ 11
3
    (1)
                          11 % 11
                                            11 1/2 11
4
    (1)
                                                                                                             0 = 0
                          11 1/2 11
5
    (1)
            .....
                          n k n
                                   11 % 11
                                           11 1/2 11
                                                        0.0
                                                                    0.0
                                                                                   0.00
                                                                                                             11 % 11
                                                                                                                           .....
6
    (1)
            0.0
                          " ½ "
                                           11 ½ 11
                                                        11 1/2 11
                                                                     0.0
                                                                                   0.0
                                                                                                             11 14 11
                                                                                                                           . .
    (1)
            11 ½ II
                          11 1/2 11
                                   11 ½ 11
                                           11 1/2 11
                                                        11 11 11
                                                                    .....
                                                                                   .....
                                                                                                  11 1/2 11
                                                                                                             11 ½ 11
                                                                                                                           .....
8
      1
         )
    (1)
            11 1/2 11
                          11 % 11
                                   11 % 11
                                           11 14 11
                                                        11 18 11
                                                                    0.00
                                                                                   0.00
                                                                                                  11 18 11
                                                                                                             11 % 11
                                                                                                                           0.00
9
     (1) "*"
                          11 14 11
                                   11 14 11
                                           11 ½ 11
                                                        11 12 11
                                                                    .....
                                                                                   0.0
                                                                                                  11 1/2 11
                                                                                                             11 14 11
                                                                                                                           .....
10
                                                                                   0.0
                                                                                                                           .....
     (1) "*"
                          11 1/2 11
                                   11 1/2 11
                                           n y n
                                                        0.80
                                                                    11 % 11
                                                                                                  n g n
                                                                                                             11 1/2 11
11
            11 % 11
                                           11 1/2 11
                                                       11 12 11
                                                                                   11 12 11
                                                                                                  11 1/2 11
                          11 % 11
                                   11 ½ 11
                                                                    11 % 11
                                                                                                             11 14 11
            11 1/4 11
                          11 1/2 11
                                   11 % 11
                                           11 1/2 11
                                                        11 18 11
                                                                    11 14 11
                                                                                   11 18 11
                                                                                                  11 1/2 11
                                                                                                             11 1/4 11
13
     (1) "*"
                          11 14 11
                                   11 1/2 11
                                            11 ½ 11
                                                        0.80
                                                                    11 1/2 11
                                                                                   11 14 11
                                                                                                  11 ½ 11
                                                                                                              11 1/2 11
14
     (1) "*"
                                   " * "
                                           " ½ "
                                                                                                                           . .
                          11 % 11
15
     (1) "*"
                                   n <sub>th</sub> n
                                                                                                                           11 1/2 11
16
     (1) "*"
                          11 ½ 11
                                                                    11 ½ II
                                                                                                  " ½ "
                                                                                                             11 ½ 11
                                                                                                                           "*"
                PhD Terminal S.F.Ratio perc.alumni Expend Grad.Rate
     (1)
                                                     .....
                0 0 0 0
                                     \mathbf{n} = \mathbf{n}
                                                                         .....
                                                                                    .....
2
     (1)
                0 \quad 0 \quad 0 \quad 0
                                     \mathbf{n} = \mathbf{n}
                                                     \mathbf{n} = \mathbf{n}
                                                                                    . .
                                                                         11 1/2 11
     (1)
                0 \quad 0 \quad 0 \quad 0
                                     0.0
                                                     .....
                                                                        11 1/2 11
                                                                                    ......
4
     (1)
                0 0 0 0
                                    0.0
                                                     11 1/2 11
                                                                                    5
     (1)
                0 0 0 0
                                    0.00
                                                     .....
                                                                                    .....
6
     (1)
                                                                         11 1/2 11
                . . . . . .
                                     0 \leq 0
                                                     ....
                                                                         11 18 11
     (1
                0 \quad 0 \quad 0 \quad 0
                                     .....
                                                     .....
                                                                         0.3 \pm 0
                                                                                    ......
8
        1
           )
                11 m m m
                                    .....
                                                    m_{\frac{1}{N}}m
                                                                                    . .
9
     (1)
      (1)"*"""
                                     ......
                                                     .........
                                                                         n g n
                                                                                    m \le m
10
      (1)"*"""
                                     \mathbf{u}=\mathbf{u}
                                                     0.00
                                                                         11 1/2 11
                                                                                    11 1/2 11
11
      (1)"*"""
                                     0.00
                                                     .....
                                                                         m \times m
                                                                                    11 ½ 11
12
      (1)"*"""
                                                     ....
                                     11 14 11
                                                                         m_{\frac{1}{N}}m
                                                                                    пұп
13
      (1) "*" "*"
                                                     ......
                                     11 14 11
                                                                         11 1/2 11
                                                                                    11 ½ 11
14
      (1) "*" "*"
                                     n y n
                                                     0.00
                                                                         n g n
                                                                                    пуп
15
      (1) "*" "*"
                                                     \mathbf{n} = \mathbf{n}
                                     11 \pm 11
                                                                         m \times m
                                                                                    11 1/2 11
16
       (1) "*" "*"
                                                     11 18 11
                                                                                    11 1/2 11
17
Coefficient estimates of 7 – predictor model:
> coef(fit.max,7)
                                                                         Top25perc
  (Intercept)
                         Accept
                                          Fnroll
                                                        Top10perc
                                                                                           Outstate
                                                                                                          Room, Board
                                                                                                                                Expend
 -466.5685474
                     1.5988303
                                     -0.5159191
                                                      49.0236357
                                                                      -14.3277749
                                                                                         -0.1199474
                                                                                                           0.1533666
                                                                                                                             0.0699095
> coef(fit.fwd,7)
  (Intercept)
                         Accept
                                          Enroll
                                                       Top10perc
                                                                        Top25perc
                                                                                           Outstate
                                                                                                          Room.Board
                                                                                                                                Expend
 -466.5685474
                     1.5988303
                                     -0.5159191
                                                      49.0236357 -14.3277749
                                                                                         -0.1199474
                                                                                                           0.1533666
                                                                                                                            0.0699095
> fit.fwd.summary = summary(fit.fwd)
 > names(fit.fwd.summary)
                                                                                          "outmat" "obj"
 [1] "which" "rsq"
                                  "rss"
                                                "adjr2" "cp"
                                                                            "bic"
```

#### Finding the best models:

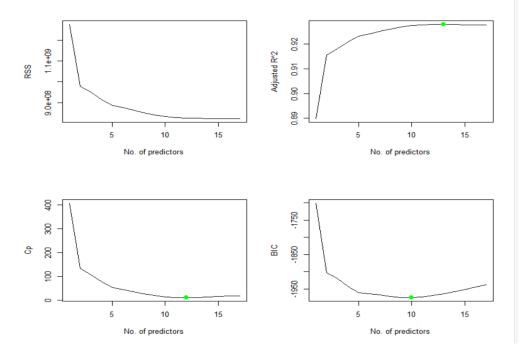
> |

```
> par(mfrow=c(2,2))
> plot(fit.fwd.summary$rss,xlab="No. of predictors",ylab="RSS",type="l")
> plot(fit.fwd.summary$adjr2,xlab="No. of predictors",ylab="Adjusted R^2",type="l")
> which.max(fit.fwd.summary$adjr2)
[1] 13
> points(13,fit.summary$adjr2[13], col="green",cex=2,pch=20)
> coef(fit.fwd,13)
(Intercept) Pr
                 PrivateYes
                                                                 Top10perc
                                                                                Top25perc
                                                                                             F.Undergrad
                                                                                                            P.Undergrad
-440.74148270 -484.77261885
                                                -0.87824288 50.41461998 -14.63667155
                                 1.58542302
                                                                                               0.05762769
                                                                                                              0.04642270
                                        PhD
                                                 S.F.Ratio
                                                                                Grad.Rate
     Outstate Room.Board
                                                                    Expend
  -0.08823311
                 0.14696204 -10.91804823
                                               15.15475056
                                                                0.07786425
                                                                               8.58578735
> plot(fit.fwd.summary$cp,xlab="No. of predictors",ylab="Cp",type='l')
> which.min(fit.fwd.summary$cp)
[1] 12
> points(12,fit.fwd.summary$cp[12],col="green",cex=2,pch=20)
> coef(fit.fwd,12)
  (Intercept)
                 PrivateYes
                                                                 Top10perc
                                                                                Top25perc
                                                                                             F.Undergrad
                                                                                                            P.Undergrad
-157.28685883 -511.78760196
                                1.58691470
                                               -0.88265385
                                                               50.41131660 -14.74735373
                                                                                               0.05945481
                                                                                                              0.04593068
                Room, Board
                                        PhD
                                                                 Grad.Rate
     Outstate
                                                    Expend
  -0.09017643
                 0.14776586 -10.70502848
                                                0.07246655
                                                                8.63961002
> plot(fit.fwd.summary$bic,xlab="No. of predictors",ylab="BIC",type='l')
> which.min(fit.summary$bic)
[1] 10
> points(10,fit.fwd.summary$bic[10],col="green",cex=2,pch=20)
> coef(fit.fwd,10)
  (Intercept)    PrivateYes
-100.51668243 -575.07060789
                                                     Enroll
                                                                 Top10perc
                                                                                Top25perc
                                                                                                 Outstate
                                                                                                              Room.Board
                                               -0.56220848
                                 1.58421887
                                                              49.13908916 -13.86531103
                                                                                              -0.09466457
                                                                                                              0.16373674
          PhD
                      Expend
                                  Grad.Rate
 -10.01608705
                  0.07273776
                                 7.33268904
```

## **Summary and Inference with plots:**

Model Selection with:	Criteria	Inference
Adjusted R^2	Select model with maximum	Model with 13 predictors
	value of R^2	has maximum R^2
Ср	Select model with smallest	Model with 12 predictors
	value of Cp	has minimum Cp
BIC	Select model with smallest	Model with 10 predictors
	value of BIC	has the least BIC

## Plots with coefficient (pointed in green):



## **Backward Stepwise selection:**

```
> fit.bwd=regsubsets(Apps~.,data=College,nvmax=17,method="backward")
> summary(fit.bwd)
Subset selection object
Call: regsubsets.formula(Apps ~ ., data = College, nvmax = 17, method = "backward")
17 Variables (and intercept)
             Forced in Forced out
PrivateYes
                 FALSE
                              FALSE
Accept
Enroll
Top10perc
                 FALSE
FALSE
                              FALSE
FALSE
                              FALSE
                 FALSE
Top25perc
                  FALSE
                              FALSE
F. Undergrad
                  FALSE
                              FALSE
P. Undergrad
                  FALSE
                              FALSE
Outstate
                  FALSE
                              FALSE
Room.Board
                  FALSE
                              FALSE
                 FALSE
FALSE
Books
                              FALSE
Personal
                              FALSE
PhD
                  FALSE
                              FALSE
Terminal
                  FALSE
                              FALSE
S.F.Ratio
                  FALSE
                              FALSE
perc.alumni
                  FALSE
                              FALSE
Expend
                  FALSE
                              FALSE
Grad.Rate
                 FALSE
                              FALSE
1 subsets of each size up to 17
Selection Algorithm: backward
```

36166610	ii Aigoi i ciiii.	Dackwai	u								
	PrivateYes					F.Undergrad					
1 (1)	0.0	'' ½ ''		0.0		0.0		0.0		0.0	0.0
2 (1)	" "	*	0.0	11 % 11							
3 (1)		11 1K 11		11.7.11							
4 (1)	0.0		0.0	11 14 11	0.0	0.0	0.0	*		0.0	0.0
5 (1)	0.0	11 % II	11 1/4 11	11/2/11	0.0	0.0	0.0	11 14 11	0.0	0.0	0.0
6 (1)	0.0	" * "	11 1/2 11	11 1/2 11	0.0	0.0	0.0	''*''	" * "	0.0	0.0
7 (1)	"*"	"*"	11 14 11	11 ½ 11				*	" * "	0.00	0.0
8 (1)	п* п		11 ¼ II	11 14 11				11 14 11	11 ½ 11		0.0
9 (1)	11 ½ 11		11 14 11	11 14 11	11 ½ 11	0.0	0.0	*	11 ½ 11	0.0	0.0
10 (1	) "*"	" * "	11 1/4 11	11/2/11	" * "	0.0	0.0	11 14 11	" * "	0.0	0.0
11 (1	) "*"	" * "	11 14 11	11 14 11	"*"	11 × 11		11 14 11	" * "	0.0	0.0
12 (1	) "*"		11 × 11	11 % 11	11 ½ 11	"*"	11 ½ 11		" * "		
13 (1	) "*"	" * "	11 × 11	11 14 11	11 ½ 11	11 * 11	11 ½ 11	11 1/2 11	" * "		
14 (1	) "*"	11 % II	11 14 11	11 14 11	11 ½ 11	11 % II	" ½ "	11 1/2 11	11 ½ 11	0.0	0.0
15 (1	) "*"	" * "	11 14 11	11 14 11	11 ½ 11	11 % II	" ½ "	11 14 11	" * "	0.0	11 1/2 11
16 (1	) "*"	" * "	11 14 11	11 14 11	11 ½ 11	" * "	"*"	11 14 11	" * "	11 X 11	11 14 11
17 (1	) "*"		" * "	11 * 11	11 ½ 11	*	" * "	11 ½ 11	"*"	11 × 11	'' * ''

```
PhD Terminal S.F.Ratio perc.alumni Expend Grad.Rate
1
     (1)
               . . . . .
                                                   .. ..
2
     (1)
               0 0 0 0
                                   .....
                                                   .....
                                                                      пұп
                                                                                .. ..
3
       1
4
       1
               . . . . .
                                   .....
                                                   .....
                                                                     11 1/2 11
                                                                                **
                                                                                   ...
          )
               . . . . .
                                                   .. ..
                                                                                   ...
                                                                      11 12 11
5
     (1)
               . . . . .
                                   .....
                                                   .. ..
                                                                     11 ½ II
                                                                                . .
6
       1
          )
               . . . . .
                                   0.0
                                                   .....
                                                                                . .
                                                                     m \le m
7
       1
          )
     (
               11×11 11 11
                                   .....
                                                                      пұп
8
       1
          )
               n * n n n
                                   . .
                                                   .. ..
                                                                     0.3 \pm 0
                                                                                .. ..
9
     (1)
               11×11 11 11
                                   . .
                                                   . .
                                                                     m \ll m
                                                                                пуп
10
      (1)
               "*" " "
                                   . .
                                                   .. ..
                                                                                n g n
11
      (1)
               \mathbf{u}_{X}\mathbf{u}=\mathbf{u}_{-\mathbf{u}}
                                   .....
                                                   .....
                                                                     11 1/2 11
                                                                                11 1/2 11
      (1)
12
               "*" " "
                                   11 1/2 11
13
                                                                      m \le m
                                                                                0.80
      (1)
               11 11 11 11 11 11 11
                                   пұп
                                                   . .
                                                                                11 🕸 11
14
      (
         1
           )
              11411 11411
15
      (1)
                                   пуп
                                                   .....
                                                                     m \le m
                                                                                n g n
               11411 11411
                                   пұп
                                                   .....
                                                                      пұп
                                                                                пұп
16
      (1)
                                                                      11 14 11
               11 x 11 11 x 11
                                   пķп
                                                   11 ½ II
                                                                                11 1/2 11
17
      (1)
> |
```

### Coefficient estimates of 7 – predictor model:

```
> coef(fit.max,7)
                                    Enroll
 (Intercept)
                                                Top10perc
                                                               Top25perc
                                                                               Outstate
                                                                                           Room.Board
                                                                                                               Expend
                     Accept
                                                                             -0.1199474
-466.5685474
                  1.5988303
                                -0.5159191
                                              49.0236357
                                                            -14.3277749
                                                                                            0.1533666
                                                                                                           0.0699095
  coef(fit.bwd,7)
(Intercept) PrivateYes
-734.27312807 -383.62179137
                                                        Enroll
                                                                    Top10perc
                                                                                     Outstate
                                                                                                   Room.Board
                                        Accept
                                                                                                                        Expend
                                   1.57996313
                                                  -0.58191135
                                                                  33.88030869
                                                                                  -0.10169116
                                                                                                   0.15500420
                                                                                                                   0.07574393
> fit.bwd.summary = summary(fit.bwd)
> names(fit.bwd.summary)
[1] "which" "rsq" "rss"
                                   "adir2"
                                             "cp"
                                                        "bic"
                                                                  "outmat" "obi"
```

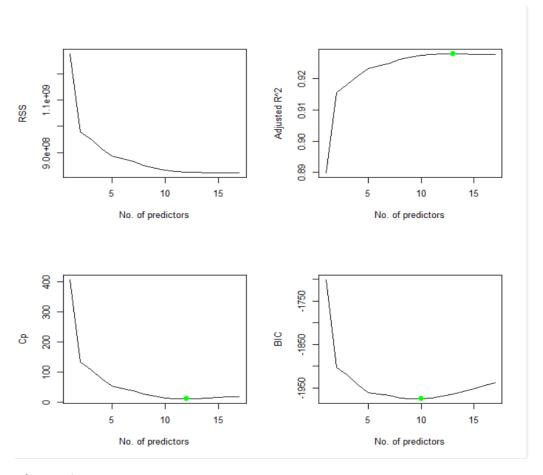
#### Finding the best models:

```
> par(mfrow=c(2,2))
> plot(fit.bwd.summary$rss,xlab="No. of predictors",ylab="RSS",type="l")
> plot(fit.bwd.summary$adjr2,xlab="No. of predictors",ylab="Adjusted R^2",type="1")
> which.max(fit.bwd.summary$adjr2)
[1] 13
> points(13,fit.summary$adjr2[13], col="green",cex=2,pch=20)
> coef(fit.bwd,13)
                 PrivateYes
                                                  Enrol1
                                                             Top10perc
                                                                           Top25perc
  (Intercept)
                                                                                        F. Undergrad
                                                                                                      P. Undergrad
                                    Accept
                                             -0.87824288
-440.74148270 -484.77261885
                               1.58542302
                                                           50.41461998
                                                                         -14.63667155
                                                                                         0.05762769
                                                                                                       0.04642270
                                      PhD
     Outstate
                 Room.Board
                                               S.F.Ratio
                                                                Expend
                                                                           Grad.Rate
                                                            0.07786425
                 0.14696204 -10.91804823
  -0.08823311
                                            15.15475056
                                                                          8.58578735
> plot(fit.bwd.summary$cp,xlab="No. of predictors",ylab="Cp",type='l')
> which.min(fit.bwd.summary$cp)
[1] 12
> points(12,fit.bwd.summary$cp[12],col="green",cex=2,pch=20)
> coef(fit.bwd,12)
                 PrivateVes
  (Intercept)
                                    Accept
                                                  Enroll
                                                             Top10perc
                                                                           Top25perc
                                                                                        F. Undergrad
                                                                                                      P. Undergrad
-157.28685883 -511.78760196
                                             -0.88265385
                                                           50.41131660
                                                                        -14.74735373
                               1.58691470
                                                                                         0.05945481
                                                                                                       0.04593068
                                      PhD
                                                             Grad.Rate
     Outstate
                 Room, Board
                                                  Expend
                 0.14776586 -10.70502848
                                              0.07246655
  -0.09017643
                                                            8.63961002
> plot(fit.bwd.summary$bic,xlab="No. of predictors",ylab="BIC",type='l')
  which.min(fit.summary$bic)
Γ17 10
> points(10,fit.bwd.summary$bic[10],col="green",cex=2,pch=20)
> coef(fit.bwd,10)
  (Intercept)
                 PrivateYes
                                    Accept
                                                             Top10perc
                                                                           Top25perc
                                                                                                       Room, Board
-100.51668243 -575.07060789
                               1.58421887
                                             -0.56220848
                                                           49.13908916 -13.86531103
                                                                                        -0.09466457
                                                                                                       0.16373674
         PhD
                     Expend
                                Grad.Rate
 -10.01608705
                 0.07273776
                               7 33268904
```

# **Summary and Inference with plots:**

Model Selection with:	Criteria	Inference		
Adjusted R^2	Select model with maximum	Model with 13 predictors		
	value of R^2	has maximum R^2		
Ср	Select model with smallest	Model with 12 predictors		
	value of Cp	has minimum Cp		
BIC	Select model with smallest	Model with 10 predictors		
	value of BIC	has the least BIC		

## Plots with coefficient (pointed in green):



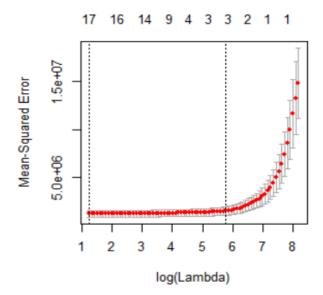
## **Observation:**

It is observed that both forward selection and the backward selection model give us the same results with respect to best model basis parameters such as adjusted R^2, BIC and Cp.

(c) Fit a lasso model on the data. Use cross-validation to select the optimal value of  $\lambda$ . Create plots of the cross-validation error as a function of  $\lambda$ . Report the resulting coefficient estimates.

- Lasso model is built using gelmet() function and setting the value of alpha to 1.
- Cross-validation is performed to obtain the best lambda to minimize the Cross-validation error.
- We predict by fitting a LASSO model to the College data with Apps as the response and all other 17 variables as predictors using the best lambda obtained.

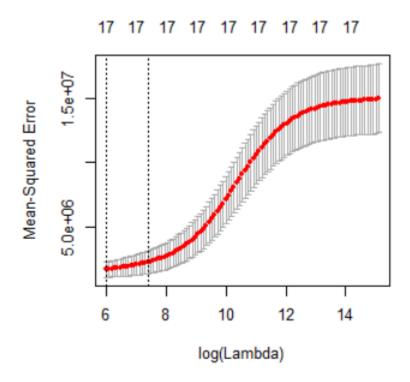
```
> set.seed(1)
> library(glmnet)
Loading required package: Matrix
Loading required package: foreach
foreach: simple, scalable parallel programming from Revolution Analytics
Use Revolution R for scalability, fault tolerance and more.
http://www.revolutionanalytics.com
Loaded glmnet 2.0-13
> x=model.matrix(Apps~.,College)[,-1]
> y=College$Apps
> cv.out=cv.glmnet(x,y,alpha=1)
> plot(cv.out)
> best_lambda=cv.out$lambda.min
> best_lambda
[1] 3.403063
> lasso.best_fit= glmnet(x, y, alpha = 1)
> predict(lasso.best_fit, s = best_lambda, type = "coefficients")
18 x 1 sparse Matrix of class "dgCMatrix"
(Intercept) -481.69122766
PrivateYes -489.47698922
Accept
              1.56285991
              -0.69952897
Enroll
             47.20524294
Top10perc
            -12.12210806
Top25perc
               0.03356097
F.Undergrad
P. Undergrad
               0.04415215
              -0.08184648
Outstate
Room.Board
               0.14813763
               0.01201765
Books
Personal
               0.02785918
              -8.24433269
PhD
              -3.21033519
Terminal
S.F.Ratio
              14.04536901
perc.alumni
              -0.13535398
               0.07662786
Expend
               8.06878113
Grad.Rate
```



(d) Fit a ridge regression model on the data. Use cross-validation to select the optimal value of  $\lambda$ . Create plots of the cross-validation error as a function of  $\lambda$ . Report the resulting coefficient estimates.

- Ridge regression model is built using glmet() function in R (and alpha=0).
- We perform cross-validation and obtain best value of lambda so that cross-validation error is minimized.
- Using the best lambda value, we predict by fitting a LASSO model Apps as response and all other 17 variables as predictors

```
> cv.out=cv.glmnet(x,y,alpha=0)
 plot(cv.out)
> best_lambda=cv.out$lambda.min
> best_lambda
[1] 400.4766
> ridge.best_fit= glmnet(x, y, alpha = 0)
> predict(ridge.best_fit, s = best_lambda, type = "coefficients")
18 x 1 sparse Matrix of class "dgCMatrix"
(Intercept) -1.514927e+03
PrivateYes -5.293325e+02
             9.780751e-01
Accept
             4.666917e-01
Enroll
Top10perc
             2.497314e+01
Top25perc
             1.056473e+00
F.Undergrad 7.662859e-02
P.Undergrad 2.445939e-02
            -2.136542e-02
Outstate
Room.Board 1.997980e-01
Books
             1.352799e-01
Personal
            -8.966624e-03
            -3.771159e+00
PhD
            -4.713593e+00
Terminal
S.F.Ratio
             1.282837e+01
perc.alumni -8.831661e+00
Expend
             7.527598e-02
Grad.Rate
             1.136663e+01
```



- (e) Now split the data set into a training set and a test set.
- i. Fit the best models obtained in the best subset selection (according to Cp, BIC or adjusted R 2 ) to the training set, and report the test error obtained.

```
> set.seed(1)
> train=sample(c(TRUE,FALSE), nrow(College), rep=TRUE)
> test=(!train)
> fit.best=regsubsets(Apps~.,data=College[train,],nvmax=17)
> summary(fit.best)
Subset selection object
Call: regsubsets.formula(Apps \sim ., data = College[train, ], nvmax = 17)
17 Variables (and intercept)
           Forced in Forced out
PrivateYes
               FALSE
                           FALSE
                           FALSE
               FALSE
Accept
                         FALSE
Enroll
              FALSE
Top10perc
Top25perc
              FALSE
                         FALSE
              FALSE
                         FALSE
F.Undergrad FALSE
                         FALSE
                         FALSE
P.Undergrad FALSE
                         FALSE
Outstate
               FALSE
           FALSE
FALSE
Room.Board
                           FALSE
Books
               FALSE
                           FALSE
             FALSE
FALSE
Personal
                           FALSE
PhD
                         FALSE
Terminal FALSE
S.F.Ratio FALSE
perc.alumni FALSE
Expend FALSE
                         FALSE
                         FALSE
                         FALSE
Expend
Grad.Rate
                         FALSE
              FALSE
                           FALSE
1 subsets of each size up to 17
```

```
Selection Algorithm: exhaustive
             PrivateYes Accept Enroll Top10perc Top25perc F.Undergrad P.Undergrad Outstate Room.Board
                                                                                          .....
                                                                                                                       . .
                                                            0.0
                                                                          0.0
                                                                                                          .....
             0.0
                            11 1/2 11
                                     0.0
                                               11 14 11
2
   (1)
                                                             0.00
                                                                                                         0.0
                            11 14 11
                                               11 1/2 11
3
   (1)
                                                                          .. ..
                            11 % 11
                                     0.00
                                                            0.0
                                                                                          0.00
                                                                                                          0.0
                                                                                                                       0.0
             11 1/2 11
                                               " * "
   (1)
                                                            ''* ''
             \alpha^{\frac{1}{N}}\alpha
                            11 14 11
                                               11 1/2 11
                                                                                         .....
                                                                                                          ........
5
   (1)
             пұп
                            11 14 11
                                               11 14 11
                                                             11 11 11
                                                                          11 ½ 11
                                                                                         0.00
                                                                                                          0.0
                                                                                                                       .....
6
   (1)
                                                                         ''*'
             11 1/2 11
                            11 14 11
                                     0.0
                                               11 14 11
                                                            11 14 11
                                                                                          0.00
                                                                                                          0.00
                                                                                                                      11411
    (1)
                                                            " * "
                                                                                          .....
                                                                                                          n g n
                            11 1/2 11
                                     11 1/2 11
                                                                                                                       11 1/2 11
8
   (1)
                                                                          "*"
                                     0.00
                                                             11 ½ 11
                                                                                          0.0
                                                                                                          0.0
             \Pi \not \cong \Pi
9
    (1)
10 (1) "*"
                            11 14 11
                                     0.0
                                               11 1/2 11
                                                            11 % 11
                                                                          11 ½ II
                                                                                                          0.0
                                                                                                                       11 1 11
                                                                          "*"
                                                            11 11 11
11 (1) "*"
                            11 1/2 11
                                     11 1/2 11
                                               11 1/2 11
                                                                                          .....
                                                                                                          11 1/2 11
   (1) "*"
                                                            " ½ "
                            " 1 "
                                     11 1/2 11
                                                                          11 ½ 11
                                                                                          .....
                                                                                                          "*"
12
13 (1) "*"
                            11 14 11
                                     11 1/2 11
                                               11 1/2 11
                                                            11 1/2 11
                                                                          11 14 11
                                                                                          0.00
                                                                                                          " * "
                                                                                                                       " ½ "
                                                            11 ½ 11
                                                                          " * "
14 (1) "*"
                            11 1/2 11
                                                                                          11 1/2 11
                                     11 % 11
                                               11 1/2 11
                                                                                                          11 1/2 11
                                                                                                                       11 1/2 11
15 (1) "*"
                           11 1/2 11
                                     11 18 11
                                              11 14 11
                                                            11 1/2 11
                                                                          11 1/2 11
                                                                                          11.76.11
                                                                                                          пжп
                                                                                                                       u \not \equiv u
16 (1) "*"
                            " ½ "
                                     11 1/2 11
                                               11 ½ II
                                                            " 18 "
                                                                          11 ½ 11
                                                                                          11 1/2 11
                                                                                                           11 1/2 11
                                                                                                                       11 ½ 11
    (1)"*"
                                                             11 11 11
                            11 1/2 11
                                     11 1/2 11
                                               11 18 11
                                                                          11 18 11
                                                                                          11 1/2 11
                                                                                                           11 1/2 11
                                                                                                                       11 % 11
```

```
Books Personal PhD Terminal S.F.Ratio perc.alumni Expend Grad.Rate
      1)
            .....
                    .....
                                0 0 0 0
                                                .. ..
                                                             ......
                                                                             .....
                                                                                      .. ..
      1)
            .....
                    .....
                                0 \quad 0 \quad 0 \quad 0
                                                .....
                                                             . .
                                                                             пұп
                                                                                      .....
    (1)
            \mathbf{u} = \mathbf{u}
                    \mathbf{u}=\mathbf{u}
                                0.000
                                                ......
                                                             ......
                                                                                      .....
4
    (1)
            .....
                    \mathbf{n} = \mathbf{n}
                                0 0 0 0
                                                0.00
                                                             0.00
                                                                             H & H
5
   (1)
            \mathbf{u} = \mathbf{u}
                    \mathbf{u}=\mathbf{u}
                                0.000
                                                0.0
                                                             0.00
                                                                             11 🔅 11
6
      1)
                                                0.0
            0.0
                    0.00
                                0.000
                                                             0.0
                                                                             11 18 11
    (1)
            .....
                    0.0
                                n*n n n
                                                0.0
                                                             0.00
                                                                             пķп
8
   (1)
                                n*n n n
            0.00
                    11 1/2 11
                                                ......
                                                             ......
                                                                             11 1/2 11
                                                                                      .....
9
   (1)
            0.0
                                *** " "
                                                0.0
                                                             0 - 0
                    11 1/2 11
                                                                             11 12 11
                                                                                      11 1/2 11
    (1)
10
         ) " "
                    11 ½ II
                                n*n n n
                                                0.0
                                                             0.00
                                                                             11 1/2 11
                                                                                      11 1/2 11
11
     (1
            0.0
                    пуп
                                11 m m m
                                                .....
                                                             ....
                                                                             n y n
                                                                                      пуп
12
     (1)
            .....
                                11 th 11 th
                                                .....
                                                             11 14 11
                    11 14 11
                                                                             11 1/2 11
                                                                                       11 1/2 11
13 (1)
                                \mathbf{u}_{\frac{1}{N}}\mathbf{u}=\mathbf{u}-\mathbf{u}
                                                .....
            11 1/2 11
    (1
                    11 % 11
                                                                             11 18 11
                                                                                      H \oplus H
14
          )
         .
) "*"
                                n * n n * n
                                                \mathbf{n} = \mathbf{n}
                    пұп
                                                              пұп
                                                                             пұп
                                                                                      пұп
15
     (1
    (1)"*"
                               n*n n n
                                                11 1/2 11
                                                             11 14 11
                    11 1/2 11
                                                                             11 14 11
                                                                                      пұп
16
    (1)"*"
                    11 1/2 11
                                11411 11411
                                                11 ½ 11
                                                              пұп
                                                                             0.8 \pm 0
                                                                                      11 1/2 11
17
> test.matrix=model.matrix(Apps~.,data=College[test,])
> val.errors=rep(NA,17)
> for(i in 1:17){
   coeff=coef(fit.best,id=i)
     pred=test.matrix[,names(coeff)]%*%coeff
    val.errors[i]=mean((College$Apps[test]-pred)^2)
> val.errors
 [1] 1714544 1542316 1510924 1512552 1492114 1635683 1645549 1610020 1626584 1616854 1555953 1520681
[13] 1526317 1519996 1522719 1520481 1520331
> which.min(val.errors)
[1] 5
> coef(fit.best,5)
                   PrivateYes
                                                       Top10perc
   (Intercept)
                                          Accept
                                                                        Top25perc
                                                                                             Expend
-114.81292106 -617.07313758
                                   1.28203523 50.65442163 -16.79622599
                                                                                        0.05489076
> fit.best=regsubsets(Apps~.,data=College,nvmax=17)
> coef(fit.best,5)
   (Intercept)
                         Accept
                                          Enroll
                                                       Top10perc
                                                                         Outstate
                                                                                             Expend
-478.93946837
                    1.60872093
                                    -0.55385758 33.47143032
                                                                      -0.10040567
                                                                                        0.08185744
> |
```

- We select the best model by checking the Validation (Test) Error rates.
- The Validation (Test) Error rates for various models are visible in the screenshots above.
- The test error rate is minimum for the model with 5 predictors combinations:
  - PrivatesYes
  - Accept
  - Top10perc
  - Top25perc
  - Expend
- We use this model and get the coefficients for the same. The coefficients of the best model are also visible in the screenshot above.

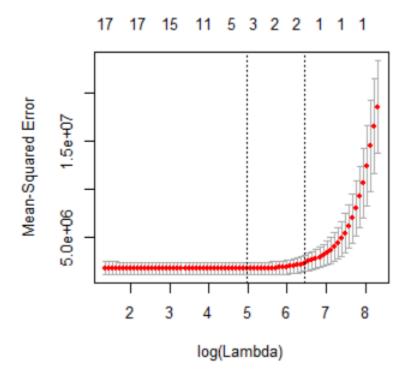
ii. Fit a lasso model to the training set, with  $\lambda$  chosen by cross validation. Report the test error obtained.

```
> set.seed(1)
> train=sample(1:nrow(x), nrow(x)/2)
> test=(-train)
> y.test=y[test]
> grid=10^seq(10,-2,length=100)
> lasso.mod=glmnet(x[train,],y[train],alpha=1,lambda=grid)
> cv.out=cv.glmnet(x[train,],y[train],alpha=1)
> plot(cv.out)
> bestlam=cv.out$lambda.min
> bestlam
[1] 144.2049
> lasso.pred=predict(lasso.mod,s=bestlam,newx=x[test,])
> mean((lasso.pred-y.test)^2)
[1] 1104498
> out=glmnet(x,y,alpha=1)
 {\tt lasso.coef=predict(out,type="coefficients",s=bestlam)[1:18,]}
> lasso.coef
  (Intercept)
                 PrivateYes
                                                  Enroll
                                                              Top10perc
                                                                            Top25perc
                                                                                        F. Undergrad
                                    Accept
               -59.24017671
                                              0.00000000
                                                                                          0.00000000
-708.20363844
                               1.38516192
                                                            22.63686463
                                                                           0.00000000
                                                              Personal
                                                   Books
                                                                                  PhD
                                                                                            Terminal
 P. Undergrad
                   Outstate
                               Room.Board
                                                                           0.00000000
                                              0.00000000
  0.00000000
                 0.00000000
                               0.00000000
                                                             0.00000000
                                                                                          0.00000000
                perc.alumni
                                               Grad.Rate
   S.F.Ratio
                                    Expend
                                              0.00000000
  0.00000000
                 0.00000000
                               0.03444127
> lasso.coef[lasso.coef!=0]
  (Intercept)
                 PrivateYes
                                    Accept
                                               Top10perc
                                                                 Expend
-708.20363844
               -59.24017671
                               1.38516192
                                             22.63686463
                                                             0.03444127
> |
```

Test MSE = 1104498 for best lambda (144.2049) The best model comprises 4 predictors –

### 1. PrivateYes 2. Accept 3. Top10perc 4. Expend

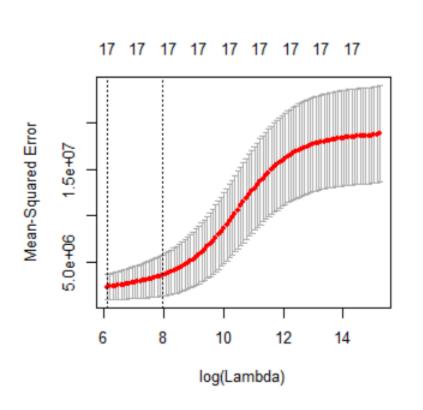
-- 3\----/



# iii. Fit a ridge regression model to the training set, with $\lambda$ chosen by cross validation. Report the test error obtained.

```
> ridge.mod=glmnet(x[train,],y[train],alpha=0,lambda=grid)
> set.seed(1)
> cv.out=cv.glmnet(x[train,],y[train],alpha=0)
> plot(cv.out)
> bestlam=cv.out$lambda.min
> bestlam
[1] 450.7435
> ridge.pred=predict(ridge.mod,s=bestlam,newx=x[test,])
> mean((ridge.pred-y.test)^2)
[1] 1036914
> out=glmnet(x,y,alpha=0)
      ridge.coef=predict(out,type="coefficients",s=bestlam)[1:18,]
> ridge.coef
                                                                                                                                                                                       Enroll
                                                                                                                                                                                                                                                                                     Top25perc
        (Intercept)
                                                               PrivateYes
                                                                                                                                                                                                                                Top10perc
                                                                                                                                                                                                                                                                                                                                 F. Undergrad
                                                                                                                                   Accept
-1.575123e + 03 -5.312141e + 02 9.443508e - 01 5.084882e - 01 2.395085e + 01 1.676068e + 00
                                                                                                                                                                                                                                                                                                                              8.195201e-02
       P. Undergrad
                                                                       Outstate
                                                                                                             Room.Board
                                                                                                                                                                                          Books
                                                                                                                                                                                                                                    Personal
                                                                                                                                                                                                                                                                                                           PhD
                                                                                                                                                                                                                                                                                                                                              Terminal
    2.519290 \\ e-02 \\ -1.804998 \\ e-02 \\ 2.008313 \\ e-01 \\ 1.443091 \\ e-01 \\ -9.669190 \\ e-03 \\ -3.428226 \\ e+00 \\ -4.551334 \\ e+00 \\ e-01 \\ -4.551334 \\ e+00 \\ -4.55134 \\ e+00 \\ -4
               S.F.Ratio perc.alumni
                                                                                                                                  Expend
                                                                                                                                                                           Grad.Rate
   1.260404e+01 -9.173762e+00 7.444322e-02 1.149378e+01
```

Test MSE = 1036914 for best lambda (450.7435)



# iv. Compare the test errors obtained in the above analysis (i-iii) and determine the optimal model.

Model	Test MSE	No. of predictors
Best subset selection	1492114	5
LASSO	1104498	4
Ridge regression	1036914	17

The optimal model is the Ridge regression model since it has the lowest Test MSE value.