Chap 9

- For controlled experiments model reduction is not very important. P 347

- For exploratory observational studies, model reduction is important.

Criteria for model selection **p353**

R-Sq criterion

Data : Surgical room data Chap 9

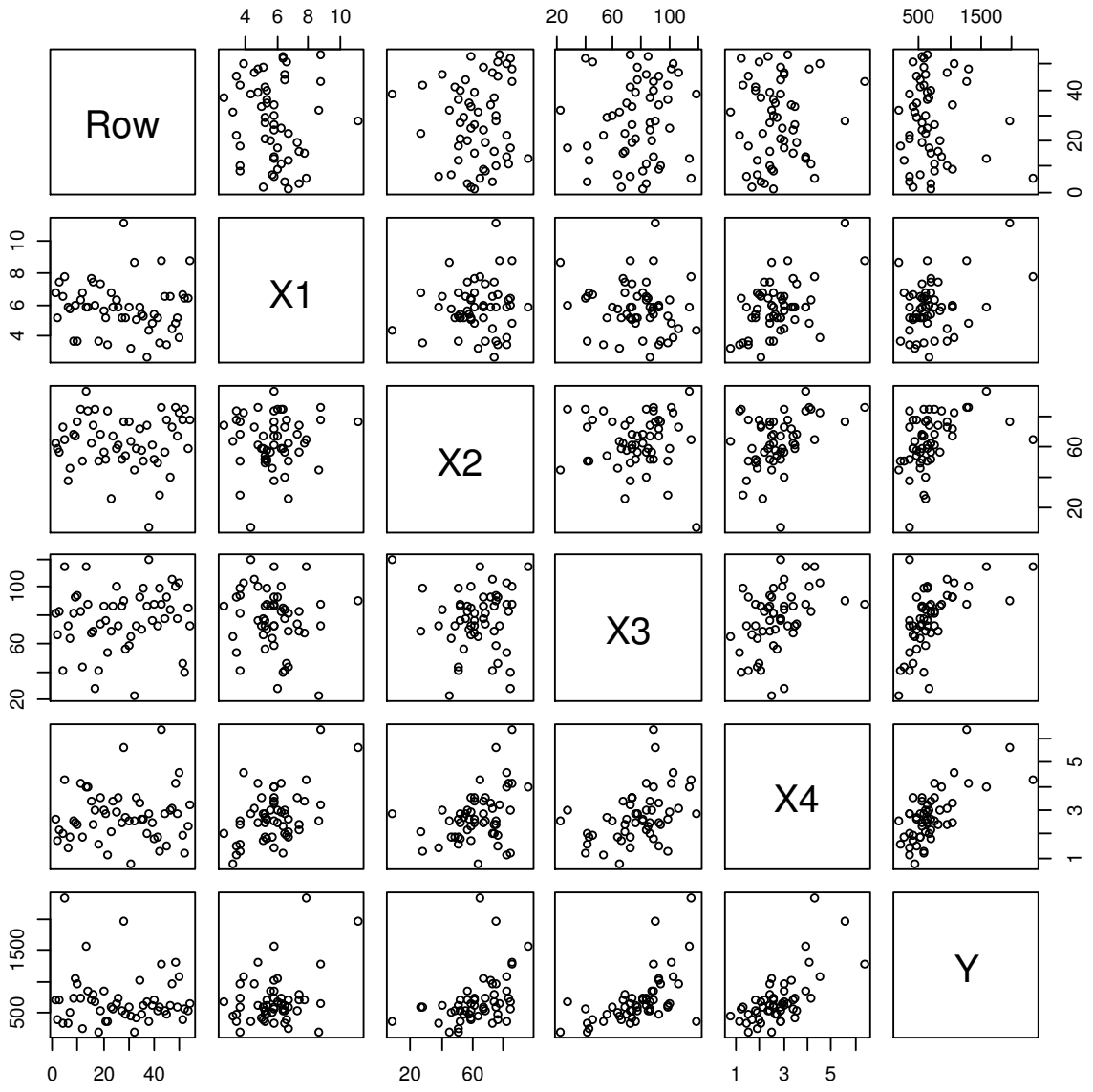
```
#Model selection
surgical=read.table("C:/Users/Mihinda/Desktop/surgical.txt",
header=1) #the data file
surgical
plot(surgical)
cor(surgical)
fit <- lm(Y ~ X1+X2+X3+X4, data=surgical)
summary(fit)
anova(fit)
plot(fitted(fit),residuals(fit))

# Try a log (i.e. ln) transformation on Y
surgical$lnY=log(surgical$Y)
fit2 <- lm(lnY ~ X1+X2+X3+X4, data=surgical)
summary(fit2)
```

```
anova(fit2)
plot(fitted(fit2),residuals(fit2))

#Variable selection
library(leaps)
X <- model.matrix(fit2)
X
X <- model.matrix(fit2)[,-1]
X

# R_sq method
r2.leaps <- leaps(X, surgical$lnY, nbest=3, method='r2')
r2.leaps
plot(r2.leaps$size, r2.leaps$r2)


# Adj R_sq method
adjr2.leaps <- leaps(X, surgical$lnY, nbest=3, method='adjr2')
adjr2.leaps
plot(adjr2.leaps$size, adjr2.leaps$adjr2)

# Cp method
Cp.leaps <- leaps(X, surgical$lnY, nbest=3, method='Cp')
Cp.leaps
```

R output

```
> #Model selection
> surgical=read.table("C:/Users/Mihinda/Desktop/surgical.txt",
header=1) #the data file
> surgical
   Row   X1 X2  X3   X4    Y
1    1  6.7 62  81 2.59  695
2    2  5.1 59  66 1.70  403
3    3  7.4 57  83 2.16  710
4    4  6.5 73  41 2.01  349
5    5  7.8 65 115 4.30 2343
6    6  5.8 38  72 1.42  348
7    7  5.7 46  63 1.91  518
8    8  3.7 68  81 2.57  749
9    9  6.0 67  93 2.50 1056
10  10  3.7 76  94 2.40  968
11  11  6.3 84  83 4.13  745
12  12  6.7 51  43 1.86  257
13  13  5.8 96 114 3.95 1573
14  14  5.8 83  88 3.95  858
15  15  7.7 62  67 3.40  702
16  16  7.4 74  68 2.40  809
17  17  6.0 85  28 2.98  682
18  18  3.7 51  41 1.55  205
19  19  7.3 68  74 3.56  550
20  20  5.6 57  87 3.02  838
21  21  5.2 52  76 2.85  359
22  22  3.4 83  53 1.12  353
23  23  6.7 26  68 2.10  599
24  24  5.8 67  86 3.40  562
25  25  6.3 59 100 2.95  651
26  26  5.8 61  73 3.50  751
27  27  5.2 52  86 2.45  545
28  28 11.2 76  90 5.59 1965
29  29  5.2 54  56 2.71  477
30  30  5.8 76  59 2.58  600
31  31  3.2 64  65 0.74  443
32  32  8.7 45  23 2.52  181
33  33  5.0 59  73 3.50  411
34  34  5.8 72  93 3.30 1037
35  35  5.4 58  70 2.64  482
36  36  5.3 51  99 2.60  634
37  37  2.6 74  86 2.05  678
38  38  4.3  8 119 2.85  362
39  39  4.8 61  76 2.45  637
40  40  5.4 52  88 1.81  705
```

```
41  41  5.2 49  72 1.84  536
42  42  3.6 28  99 1.30  582
43  43  8.8 86  88 6.40 1270
44  44  6.5 56  77 2.85  538
45  45  3.4 77  93 1.48  482
46  46  6.5 40  84 3.00  611
47  47  4.5 73 106 3.05  960
48  48  4.8 86 101 4.10 1300
49  49  5.1 67  77 2.86  581
50  50  3.9 82 103 4.55 1078
51  51  6.6 77  46 1.95  405
52  52  6.4 85  40 1.21  579
53  53  6.4 59  85 2.33  550
54  54  8.8 78  72 3.20  651
> plot(surgical)
```

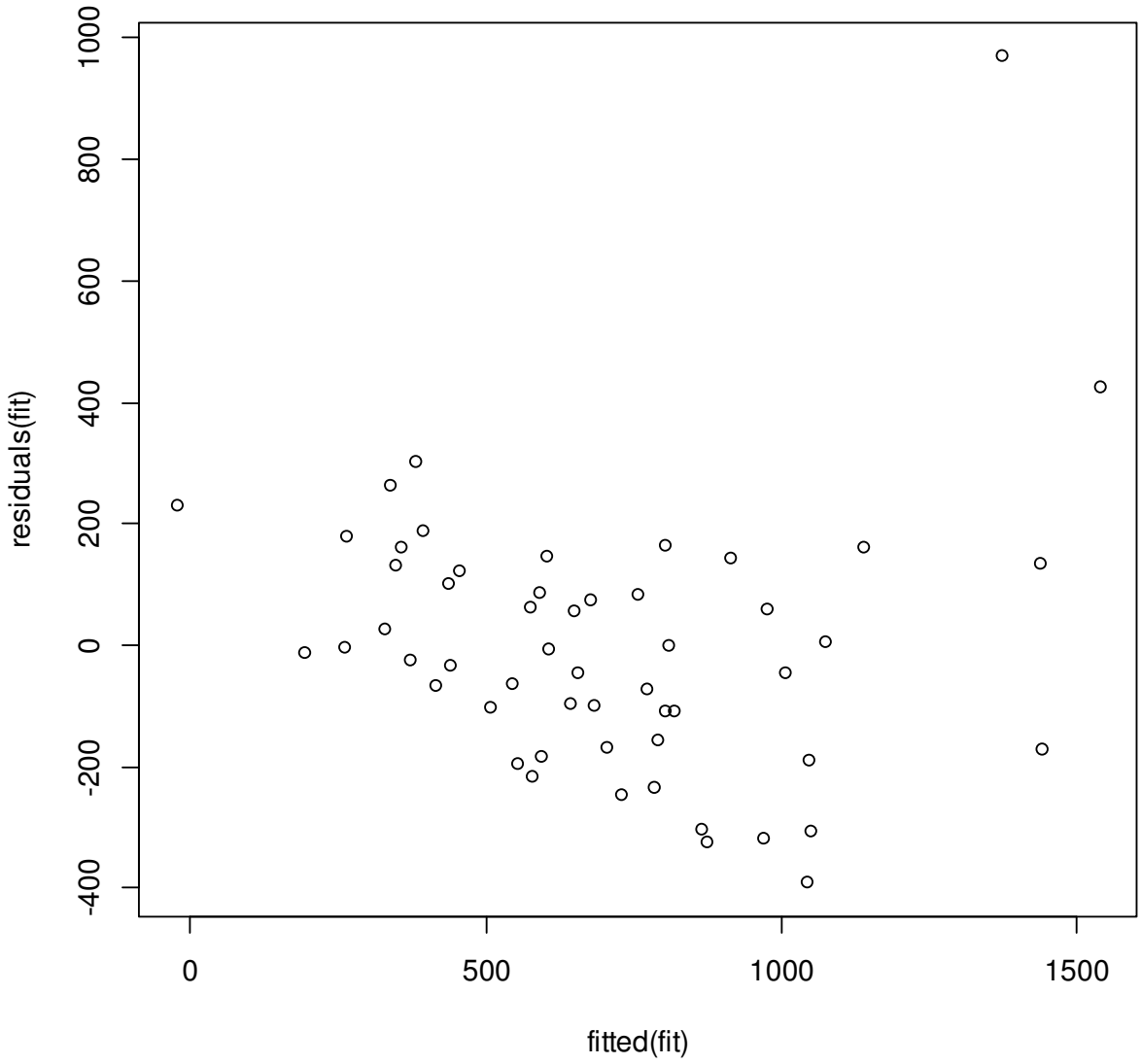



```
> cor(surgical)
```

```
            Row          X1          X2          X3          X4           Y
Row  1.00000000 -0.09812105  0.02813377  0.11488890 0.06163258 -0.07831648
X1  -0.09812105  1.00000000  0.09011973 -0.14963411 0.50241567  0.34654968
X2   0.02813377  0.09011973  1.00000000 -0.02360544 0.36902563  0.42048097
X3   0.11488890 -0.14963411 -0.02360544  1.00000000 0.41642451  0.57822600
X4   0.06163258  0.50241567  0.36902563  0.41642451 1.00000000  0.67419499
Y   -0.07831648  0.34654968  0.42048097  0.57822600 0.67419499  1.00000000
> fit <- lm(Y ~ X1+X2+X3+X4, data=surgical)
```

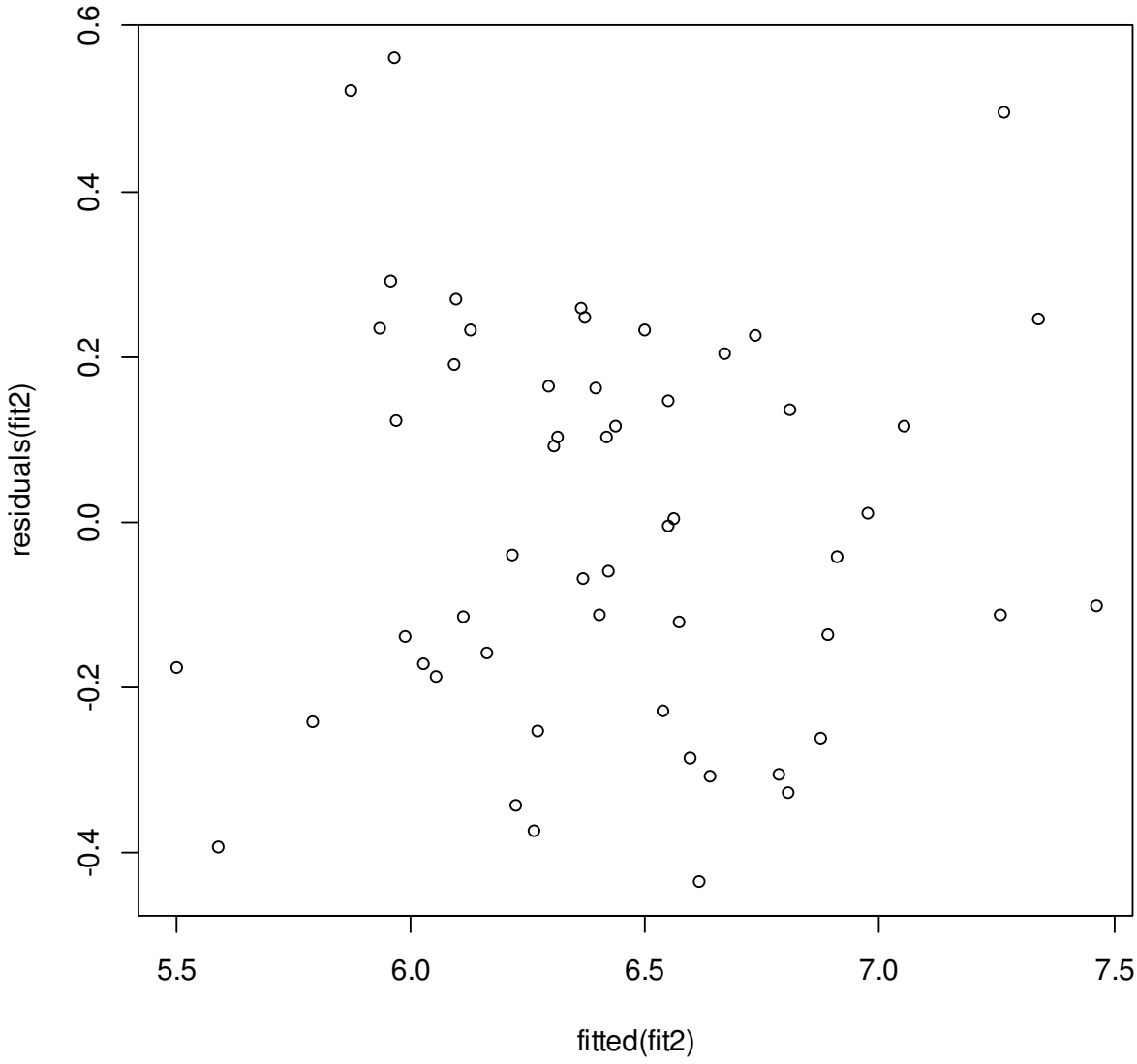
```
> summary(fit)

Call:
lm(formula = Y ~ X1 + X2 + X3 + X4, data = surgical)

Residuals:
    Min      1Q  Median      3Q     Max
-391.55 -144.81   -8.34  129.51  970.26

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -1279.242    243.808  -5.247 3.30e-06 ***
X1             82.988     26.402   3.143  0.00284 **
X2              8.346      2.120   3.937  0.00026 ***
X3             10.870      1.923   5.652 8.01e-07 ***
X4             49.346     47.126   1.047  0.30018
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 229.7 on 49 degrees of freedom
Multiple R-squared: 0.691,      Adjusted R-squared: 0.6658
F-statistic:  27.4 on 4 and 49 DF,  p-value: 5.704e-12

> anova(fit)
Analysis of Variance Table

Response: Y
          Df  Sum Sq Mean Sq F value    Pr(>F)
X1         1 1005152 1005152 19.0470 6.567e-05 ***
X2         1 1278496 1278496 24.2267 1.010e-05 ***
X3         1 3442172 3442172 65.2269 1.461e-10 ***
X4         1   57862   57862  1.0964    0.3002
Residuals 49 2585839   52772
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

plot(fitted(fit),residuals(fit))




```
> # Try a log (i.e. ln) transformation on Y
> surgical$lnY=log(surgical$Y)
> fit2 <- lm(lnY ~ X1+X2+X3+X4, data=surgical)
> summary(fit2)

Call:
lm(formula = lnY ~ X1 + X2 + X3 + X4, data = surgical)

Residuals:
     Min       1Q   Median       3Q      Max
```

```
-0.43514 -0.17436 -0.02156  0.18475  0.56054

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 3.851933   0.266263  14.467  < 2e-16 ***
X1          0.083739   0.028834   2.904  0.00551 **
X2          0.012671   0.002315   5.474 1.50e-06 ***
X3          0.015627   0.002100   7.440 1.38e-09 ***
X4          0.032056   0.051466   0.623  0.53627
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.2509 on 49 degrees of freedom
Multiple R-squared: 0.7591,     Adjusted R-squared: 0.7395
F-statistic: 38.61 on 4 and 49 DF,  p-value: 1.398e-14

> anova(fit2)
Analysis of Variance Table

Response: lnY
          Df Sum Sq Mean Sq  F value    Pr(>F)
X1         1 0.7770  0.7770  12.3443 0.0009618 ***
X2         1 2.5904  2.5904  41.1565 5.341e-08 ***
X3         1 6.3286  6.3286 100.5490 1.838e-13 ***
X4         1 0.0244  0.0244   0.3879 0.5362698
Residuals 49 3.0841  0.0629
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
> plot(fitted(fit2),residuals(fit2))
```


residuals(fit2)
0.6
0.4
0.2
0.0
-0.2
-0.4
5.5
6.0
6.5
7.0
7.5
fitted(fit2)

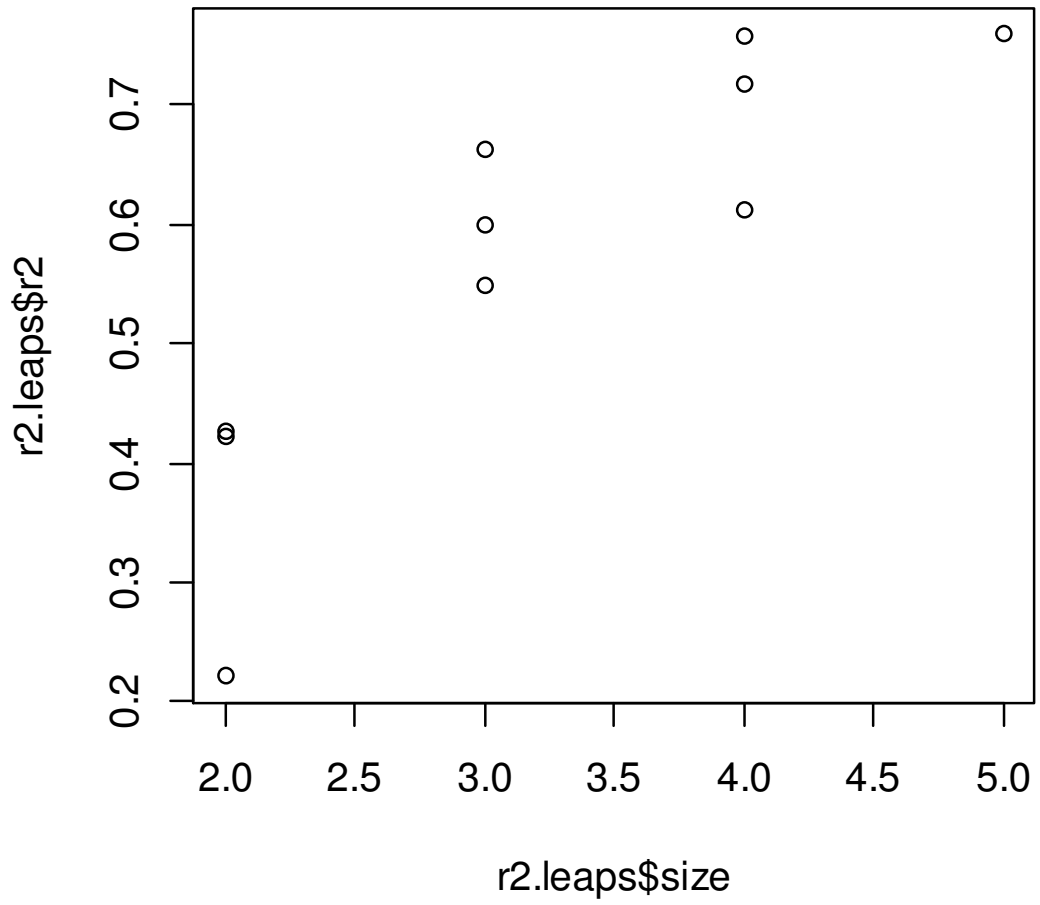
Variable selection methods

```
> #Variable selection
> library(leaps)
> X <- model.matrix(fit2)
> X
   (Intercept)   X1 X2  X3   X4
1            1  6.7 62  81 2.59
2            1  5.1 59  66 1.70
3            1  7.4 57  83 2.16
4            1  6.5 73  41 2.01
5            1  7.8 65 115 4.30
6            1  5.8 38  72 1.42
7            1  5.7 46  63 1.91
8            1  3.7 68  81 2.57
9            1  6.0 67  93 2.50
10           1  3.7 76  94 2.40
11           1  6.3 84  83 4.13
12           1  6.7 51  43 1.86
13           1  5.8 96 114 3.95
14           1  5.8 83  88 3.95
15           1  7.7 62  67 3.40
16           1  7.4 74  68 2.40
17           1  6.0 85  28 2.98
18           1  3.7 51  41 1.55
19           1  7.3 68  74 3.56
20           1  5.6 57  87 3.02
21           1  5.2 52  76 2.85
22           1  3.4 83  53 1.12
23           1  6.7 26  68 2.10
24           1  5.8 67  86 3.40
25           1  6.3 59 100 2.95
26           1  5.8 61  73 3.50
27           1  5.2 52  86 2.45
28           1 11.2 76  90 5.59
29           1  5.2 54  56 2.71
30           1  5.8 76  59 2.58
31           1  3.2 64  65 0.74
32           1  8.7 45  23 2.52
33           1  5.0 59  73 3.50
34           1  5.8 72  93 3.30
35           1  5.4 58  70 2.64
36           1  5.3 51  99 2.60
37           1  2.6 74  86 2.05
38           1  4.3  8 119 2.85
39           1  4.8 61  76 2.45
```

```
40           1  5.4 52  88 1.81
41           1  5.2 49  72 1.84
42           1  3.6 28  99 1.30
43           1  8.8 86  88 6.40
44           1  6.5 56  77 2.85
45           1  3.4 77  93 1.48
46           1  6.5 40  84 3.00
47           1  4.5 73 106 3.05
48           1  4.8 86 101 4.10
49           1  5.1 67  77 2.86
50           1  3.9 82 103 4.55
51           1  6.6 77  46 1.95
52           1  6.4 85  40 1.21
53           1  6.4 59  85 2.33
54           1  8.8 78  72 3.20
attr(,"assign")
[1] 0 1 2 3 4
> X <- model.matrix(fit2)[,-1]
> X
     X1 X2  X3   X4
1   6.7 62  81 2.59
2   5.1 59  66 1.70
3   7.4 57  83 2.16
4   6.5 73  41 2.01
5   7.8 65 115 4.30
6   5.8 38  72 1.42
7   5.7 46  63 1.91
8   3.7 68  81 2.57
9   6.0 67  93 2.50
10  3.7 76  94 2.40
11  6.3 84  83 4.13
12  6.7 51  43 1.86
13  5.8 96 114 3.95
14  5.8 83  88 3.95
15  7.7 62  67 3.40
16  7.4 74  68 2.40
17  6.0 85  28 2.98
18  3.7 51  41 1.55
19  7.3 68  74 3.56
20  5.6 57  87 3.02
21  5.2 52  76 2.85
22  3.4 83  53 1.12
23  6.7 26  68 2.10
24  5.8 67  86 3.40
25  6.3 59 100 2.95
26  5.8 61  73 3.50
27  5.2 52  86 2.45
```

```
28 11.2 76  90 5.59
29  5.2 54  56 2.71
30  5.8 76  59 2.58
31  3.2 64  65 0.74
32  8.7 45  23 2.52
33  5.0 59  73 3.50
34  5.8 72  93 3.30
35  5.4 58  70 2.64
36  5.3 51  99 2.60
37  2.6 74  86 2.05
38  4.3  8 119 2.85
39  4.8 61  76 2.45
40  5.4 52  88 1.81
41  5.2 49  72 1.84
42  3.6 28  99 1.30
43  8.8 86  88 6.40
44  6.5 56  77 2.85
45  3.4 77  93 1.48
46  6.5 40  84 3.00
47  4.5 73 106 3.05
48  4.8 86 101 4.10
49  5.1 67  77 2.86
50  3.9 82 103 4.55
51  6.6 77  46 1.95
52  6.4 85  40 1.21
53  6.4 59  85 2.33
54  8.8 78  72 3.20
>
> # R_sq method
> r2.leaps <- leaps(X, surgical$lnY, nbest=3, method='r2')
> r2.leaps
$which
      1     2     3     4
1 FALSE FALSE  TRUE FALSE
1 FALSE FALSE FALSE  TRUE
1 FALSE  TRUE FALSE FALSE
2 FALSE  TRUE  TRUE FALSE
2 FALSE FALSE  TRUE  TRUE
2  TRUE FALSE  TRUE FALSE
3  TRUE  TRUE  TRUE FALSE
3 FALSE  TRUE  TRUE  TRUE
3  TRUE FALSE  TRUE  TRUE
4  TRUE  TRUE  TRUE  TRUE

$label
[1] "(Intercept)" "1"           "2"           "3"           "4"
```

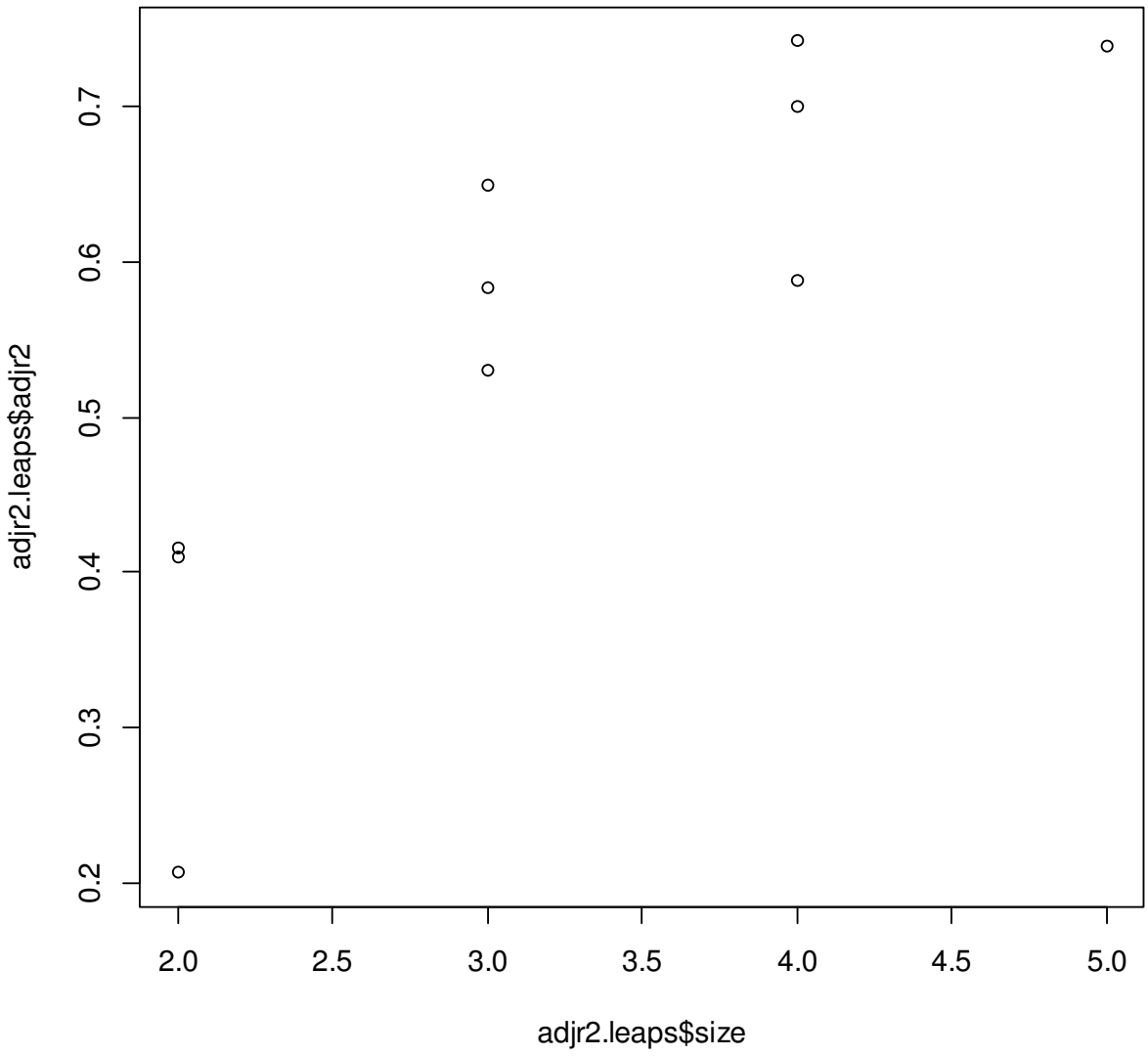
```
$size
 [1] 2 2 2 3 3 3 4 4 4 5

$r2
 [1] 0.4272545 0.4214649 0.2210432 0.6631757 0.5992084 0.5483980 0.7572331
 [8] 0.7176812 0.6118710 0.7591401

> plot(r2.leaps$size, r2.leaps$r2)
```



Adjusted R sq method

```
> # Adj R_sq method
> adjr2.leaps <- leaps(X, surgical$lnY, nbest=3, method='adjr2')
> adjr2.leaps
$which
      1     2     3     4
1 FALSE FALSE  TRUE FALSE
1 FALSE FALSE FALSE  TRUE
1 FALSE  TRUE FALSE FALSE
2 FALSE  TRUE  TRUE FALSE
2 FALSE FALSE  TRUE  TRUE
2  TRUE FALSE  TRUE FALSE
3  TRUE  TRUE  TRUE FALSE
3 FALSE  TRUE  TRUE  TRUE
3  TRUE FALSE  TRUE  TRUE
4  TRUE  TRUE  TRUE  TRUE

$label
[1] "(Intercept)" "1"           "2"           "3"           "4"

$size
 [1] 2 2 2 3 3 3 4 4 4 5

$adjr2
 [1] 0.4162401 0.4103392 0.2060632 0.6499669 0.5834910 0.5306881 0.7426671
 [8] 0.7007420 0.5885833 0.7394781
```

```
> plot(adjr2.leaps$size, adjr2.leaps$adjr2)
```



Mallows C_p criterion p357

$$C_p = \frac{SSE_p}{MSE_P} - (n - 2p)$$

P-1 is the number of potential *X* variables.

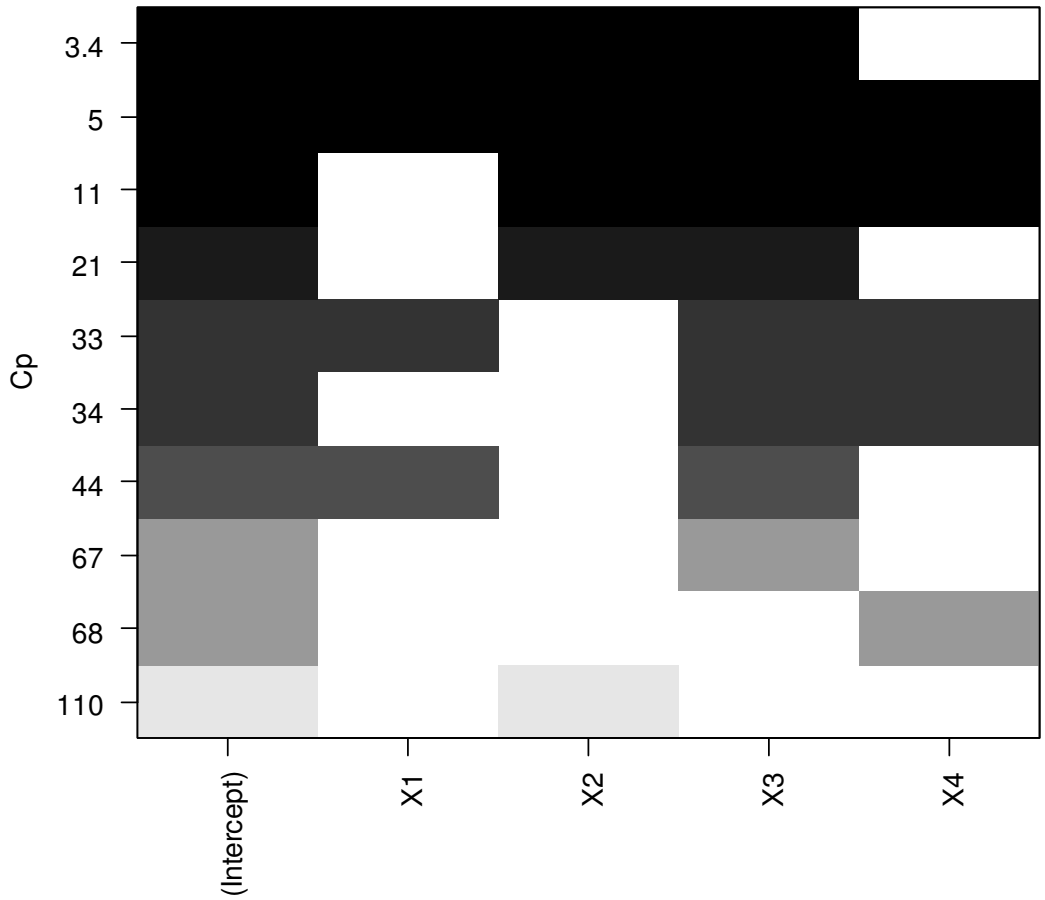
p-1 is the *X* variables in a subset.

When *p* = *P*, $C_p = P$

The basic idea is to find the smallest value of p such that $C_p \approx p$.

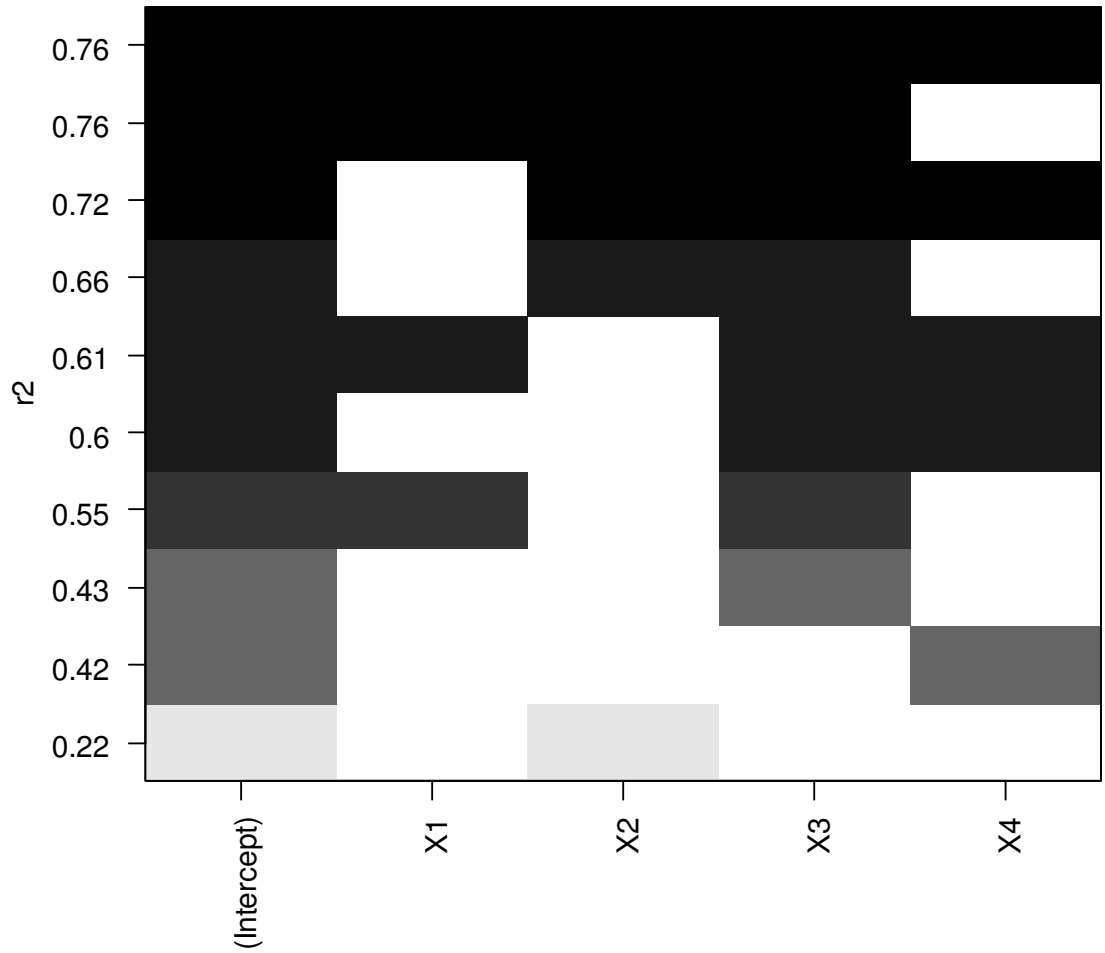
```
> # Cp method
> Cp.leaps <- leaps(X, surgical$lnY, method='Cp')
> Cp.leaps
$which
      1     2     3     4
1 FALSE FALSE  TRUE FALSE
1 FALSE FALSE FALSE  TRUE
1 FALSE  TRUE FALSE FALSE
1  TRUE FALSE FALSE FALSE
2 FALSE  TRUE  TRUE FALSE
2 FALSE FALSE  TRUE  TRUE
2  TRUE FALSE  TRUE FALSE
2 FALSE  TRUE FALSE  TRUE
2  TRUE FALSE FALSE  TRUE
2  TRUE  TRUE FALSE FALSE
3  TRUE  TRUE  TRUE FALSE
3 FALSE  TRUE  TRUE  TRUE
3  TRUE FALSE  TRUE  TRUE
3  TRUE  TRUE FALSE  TRUE
4  TRUE  TRUE  TRUE  TRUE

$label
[1] "(Intercept)" "1"           "2"           "3"           "4"

$size
 [1] 2 2 2 2 3 3 3 3 3 3 4 4 4 4 5

$Cp
 [1]  66.518070  67.695892 108.469236 141.093441  20.522784  33.536154
 [7]  43.872905  57.174572  67.961243 101.936972   3.387945  11.434313
[13]  32.960084  58.357601   5.000000
```

```
plot(leaps, scale="Cp")
```



Black indicates that a variable is included in the model, while white indicates that they are not.

```
> plot(leaps, scale="r2")
```



Step wise selection with AIC_p criterion p359

$$AIC_p = n\ln SSE_p - n\ln n + 2p$$

W look for models with smaller AIC_p

Note: SSE_p decreases as p increases, the second term is fixed and the third term (i.e. penalty) increases as p increases. The models with smaller SSE_p will do well by this criterion as long as the penalties ($2p$) is not too large.

```
> #Stepwise regression
> # Remove the unnecessary columns i.e. row and Y
> # from the data set to avoid them going to candidate set.
> surgical=surgical[-1]
> surgical
     X1 X2  X3   X4    Y      lnY
1   6.7 62  81 2.59  695 6.543912
2   5.1 59  66 1.70  403 5.998937
3   7.4 57  83 2.16  710 6.565265
4   6.5 73  41 2.01  349 5.855072
5   7.8 65 115 4.30 2343 7.759187
6   5.8 38  72 1.42  348 5.852202
7   5.7 46  63 1.91  518 6.249975
8   3.7 68  81 2.57  749 6.618739
9   6.0 67  93 2.50 1056 6.962243
10  3.7 76  94 2.40  968 6.875232
11  6.3 84  83 4.13  745 6.613384
12  6.7 51  43 1.86  257 5.549076
13  5.8 96 114 3.95 1573 7.360740
14  5.8 83  88 3.95  858 6.754604
15  7.7 62  67 3.40  702 6.553933
16  7.4 74  68 2.40  809 6.695799
17  6.0 85  28 2.98  682 6.525030
18  3.7 51  41 1.55  205 5.323010
19  7.3 68  74 3.56  550 6.309918
20  5.6 57  87 3.02  838 6.731018
21  5.2 52  76 2.85  359 5.883322
22  3.4 83  53 1.12  353 5.866468
23  6.7 26  68 2.10  599 6.395262
24  5.8 67  86 3.40  562 6.331502
25  6.3 59 100 2.95  651 6.478510
26  5.8 61  73 3.50  751 6.621406
27  5.2 52  86 2.45  545 6.300786
28 11.2 76  90 5.59 1965 7.583248
29  5.2 54  56 2.71  477 6.167516
30  5.8 76  59 2.58  600 6.396930
31  3.2 64  65 0.74  443 6.093570
32  8.7 45  23 2.52  181 5.198497
33  5.0 59  73 3.50  411 6.018593
34  5.8 72  93 3.30 1037 6.944087
35  5.4 58  70 2.64  482 6.177944
36  5.3 51  99 2.60  634 6.452049
37  2.6 74  86 2.05  678 6.519147
38  4.3  8 119 2.85  362 5.891644
39  4.8 61  76 2.45  637 6.456770
40  5.4 52  88 1.81  705 6.558198
41  5.2 49  72 1.84  536 6.284134
42  3.6 28  99 1.30  582 6.366470
43  8.8 86  88 6.40 1270 7.146772
44  6.5 56  77 2.85  538 6.287859
45  3.4 77  93 1.48  482 6.177944
46  6.5 40  84 3.00  611 6.415097
47  4.5 73 106 3.05  960 6.866933
48  4.8 86 101 4.10 1300 7.170120
49  5.1 67  77 2.86  581 6.364751
50  3.9 82 103 4.55 1078 6.982863
51  6.6 77  46 1.95  405 6.003887
52  6.4 85  40 1.21  579 6.361302
53  6.4 59  85 2.33  550 6.309918
54  8.8 78  72 3.20  651 6.478510
```

```
> surgical=surgical[-5]
> surgical
     X1 X2  X3   X4      lnY
1   6.7 62  81 2.59 6.543912
2   5.1 59  66 1.70 5.998937
3   7.4 57  83 2.16 6.565265
4   6.5 73  41 2.01 5.855072
5   7.8 65 115 4.30 7.759187
6   5.8 38  72 1.42 5.852202
7   5.7 46  63 1.91 6.249975
8   3.7 68  81 2.57 6.618739
9   6.0 67  93 2.50 6.962243
10  3.7 76  94 2.40 6.875232
11  6.3 84  83 4.13 6.613384
12  6.7 51  43 1.86 5.549076
13  5.8 96 114 3.95 7.360740
14  5.8 83  88 3.95 6.754604
15  7.7 62  67 3.40 6.553933
16  7.4 74  68 2.40 6.695799
17  6.0 85  28 2.98 6.525030
18  3.7 51  41 1.55 5.323010
19  7.3 68  74 3.56 6.309918
20  5.6 57  87 3.02 6.731018
21  5.2 52  76 2.85 5.883322
22  3.4 83  53 1.12 5.866468
23  6.7 26  68 2.10 6.395262
24  5.8 67  86 3.40 6.331502
25  6.3 59 100 2.95 6.478510
26  5.8 61  73 3.50 6.621406
27  5.2 52  86 2.45 6.300786
28 11.2 76  90 5.59 7.583248
29  5.2 54  56 2.71 6.167516
30  5.8 76  59 2.58 6.396930
31  3.2 64  65 0.74 6.093570
32  8.7 45  23 2.52 5.198497
33  5.0 59  73 3.50 6.018593
34  5.8 72  93 3.30 6.944087
35  5.4 58  70 2.64 6.177944
36  5.3 51  99 2.60 6.452049
37  2.6 74  86 2.05 6.519147
38  4.3  8 119 2.85 5.891644
39  4.8 61  76 2.45 6.456770
40  5.4 52  88 1.81 6.558198
41  5.2 49  72 1.84 6.284134
42  3.6 28  99 1.30 6.366470
43  8.8 86  88 6.40 7.146772
44  6.5 56  77 2.85 6.287859
45  3.4 77  93 1.48 6.177944
46  6.5 40  84 3.00 6.415097
47  4.5 73 106 3.05 6.866933
48  4.8 86 101 4.10 7.170120
49  5.1 67  77 2.86 6.364751
50  3.9 82 103 4.55 6.982863
51  6.6 77  46 1.95 6.003887
52  6.4 85  40 1.21 6.361302
53  6.4 59  85 2.33 6.309918
54  8.8 78  72 3.20 6.478510
>
```

```
> null=lm(lnY~1, data=surgical)
> null

Call:
lm(formula = lnY ~ 1, data = surgical)

Coefficients:
(Intercept)
      6.431

> full=lm(lnY~., data=surgical)
> full

Call:
lm(formula = lnY ~ ., data = surgical)

Coefficients:
(Intercept)           X1           X2           X3           X4
    3.85193      0.08374      0.01267      0.01563      0.03206
```

Foreward selection

```
> step(null, scope=list(lower=null, upper=full), direction="forward")
Start:  AIC=-75.72
lnY ~ 1

       Df Sum of Sq     RSS      AIC
+ X3    1    5.4708  7.3337 -103.811
+ X4    1    5.3967  7.4079 -103.268
+ X2    1    2.8303  9.9742  -87.205
+ X1    1    0.7770 12.0275  -77.096
<none>              12.8045  -75.716

Step:  AIC=-103.81
lnY ~ X3

       Df Sum of Sq    RSS     AIC
+ X2    1    3.0209 4.3129 -130.48
+ X4    1    2.2018 5.1319 -121.09
+ X1    1    1.5512 5.7825 -114.64
<none>              7.3337 -103.81

Step:  AIC=-130.48
lnY ~ X3 + X2
```

```
       Df Sum of Sq    RSS     AIC
+ X1    1   1.20436 3.1085 -146.16
+ X4    1   0.69792 3.6150 -138.01
<none>              4.3129 -130.48

Step:  AIC=-146.16
lnY ~ X3 + X2 + X1

       Df Sum of Sq    RSS     AIC
<none>              3.1085 -146.16
+ X4    1  0.024418 3.0841 -144.59

Call:
lm(formula = lnY ~ X3 + X2 + X1, data = surgical)

Coefficients:
(Intercept)           X3           X2           X1
    3.76644      0.01644      0.01334      0.09547


> fitx3 <- lm(lnY ~ X3, data=surgical)

> anova(fitx3)
Analysis of Variance Table

Response: lnY
          Df Sum Sq Mean Sq F value    Pr(>F)
X3         1 5.4708  5.4708  38.791 8.381e-08
***
Residuals 52 7.3337  0.1410
```

n=54
p=2

SSE =7.3337

AIC = n*ln(SSE)-n*ln(n)+2*p = -103.8112097

```
> fitx1x2x3 <- lm(lnY ~ X1+X2+X3, data=surgical)
> summary(fitx1x2x3)

Call:
lm(formula = lnY ~ X1 + X2 + X3, data = surgical)


Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 3.766441   0.226757  16.610  < 2e-16 ***
X1          0.095475   0.021692   4.401 5.66e-05 ***
X2          0.013344   0.002035   6.558 2.95e-08 ***
X3          0.016444   0.001630  10.089 1.19e-13 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.2493 on 50 degrees of freedom
Multiple R-squared: 0.7572,     Adjusted R-squared: 0.7427
F-statistic: 51.99 on 3 and 50 DF,  p-value: 2.137e-15
```

```
> step(full, scope=list(lower=null, upper=full), direction="backward")
Start:  AIC=-144.59
lnY ~ X1 + X2 + X3 + X4

       Df Sum of Sq    RSS     AIC
- X4    1    0.0244 3.1085 -146.16
<none>               3.0841 -144.59
- X1    1    0.5309 3.6150 -138.01
- X2    1    1.8857 4.9698 -120.82
- X3    1    3.4842 6.5683 -105.76

Step:  AIC=-146.16
lnY ~ X1 + X2 + X3

       Df Sum of Sq    RSS      AIC
<none>               3.1085 -146.161
- X1    1    1.2044 4.3129 -130.479
- X2    1    2.6740 5.7825 -114.644
- X3    1    6.3286 9.4371  -88.194

Call:
lm(formula = lnY ~ X1 + X2 + X3, data = surgical)

Coefficients:
(Intercept)           X1           X2           X3
    3.76644      0.09547      0.01334      0.01644
```

```
> step(full, scope=list(lower=null, upper=full), direction="both")
Start:  AIC=-144.59
lnY ~ X1 + X2 + X3 + X4

       Df Sum of Sq    RSS     AIC
- X4    1    0.0244 3.1085 -146.16
<none>               3.0841 -144.59
- X1    1    0.5309 3.6150 -138.01
- X2    1    1.8857 4.9698 -120.82
- X3    1    3.4842 6.5683 -105.76

Step:  AIC=-146.16
lnY ~ X1 + X2 + X3

       Df Sum of Sq    RSS      AIC
<none>               3.1085 -146.161
+ X4    1    0.0244 3.0841 -144.587
- X1    1    1.2044 4.3129 -130.479
- X2    1    2.6740 5.7825 -114.644
- X3    1    6.3286 9.4371  -88.194

Call:
lm(formula = lnY ~ X1 + X2 + X3, data = surgical)

Coefficients:
(Intercept)           X1           X2           X3
    3.76644      0.09547      0.01334      0.01644

>
```