STAT 51200

Applied Regression Analysis

Homework Assignments #08

By – Aditya Gaitonde

Q2)

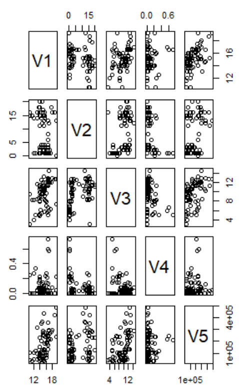
7.7) a)

Code –

```
data <- read.table("C:/Users/Aditya Gaitonde/Downloads/CH06PR18.txt", header = FALSE)
attach(data)
cor(data)
plot(data)
```

Output –

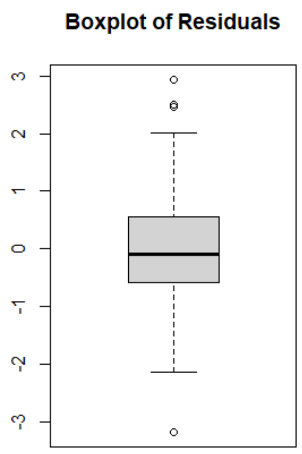
	V1	V2	V3	V4	V5
V1	1.00000000	-0.2502846	0.4137872	0.06652647	0.53526237
V2	-0.25028456	1.0000000	0.3888264	-0.25266347	0.28858350
V3	0.41378716	0.3888264	1.0000000	-0.37976174	0.44069713
V4	0.06652647	-0.2526635	-0.3797617	1.00000000	0.08061073
V5	0.53526237	0.2885835	0.4406971	0.08061073	1.00000000



Code –

```
model <- lm(V1 ~ V2+V3+V4+V5, data=data)
residuals <- model$residuals
boxplot(residuals, main="Boxplot of Residuals")
```

Plot –



Code –

```
y.hat <- fitted(model)
model$coefficients
anova(model)
SSTO = sum( anova(model)[,2] )
MSE = anova(model)[5,3]
SSR = sum( anova(model)[1:4,2] ) #SSR(X1, X2, X3, X4)
MSR = SSR / 4 #MSR(X1, X2, X3, X4) = SSR / df
SSE = anova(model)[5,2] #SSE(X1, X2, X3, X4)
MSE = anova(model)[5,3] #MSE(X1, X2, X3, X4)
model <- lm(V1 ~ V5, data=data)
anova(model)
SSR_X4 = anova(model)[1,2]
```

```
model <- lm(V1 ~ V2+V5, data=data)
SSR_X1_X4 = sum(anova(model)[1:2,2])
model <- lm(V1 ~ V2+V3+V5, data=data)
SSR_X1_X2_X4 = sum(anova(model)[1:3,2])
SSR_X1_X4 - SSR_X4
SSR_X1_X2_X4 - SSR_X1_X4
SSR - SSR_X1_X2_X4
```

The ANOVA table –

Sum of Variation	SS	df	MS
Regression	SSR(X_1, X_2, X_3, X_4) 138.326	4	34.58
X_4	SSR(X_4) 40.5033	1	40.50
$X_1 \mid X_4$	SSR$(X_1 \mid X_4)$ 42.2746	1	42
$X_2 \mid X_1, X_4$	SSR$(X_2 \mid X_1, X_4)$ 27.8575	1	27.8575
$X_3 \mid X_1, X_2, X_4$	SSR$(X_3 \mid X_1, X_2, X_4)$ 0.4195	1	0.4195
Error	SSE(X_1, X_2, X_3, X_4) 98.2306	76	1.29
Total	SSTO 236.55	79	

b) Code –

```
model <- lm(V1 ~ V4, data=data)
SSR_X3 = anova(model)[1,2]
n = length(data$V1)
f.stat = ((SSR - SSR_X1_X2_X4) * (n - 4))/SSE
f.value = qf(0.99,1,n-4)
```

To test if X_3 can be dropped from regression model

We test if $\beta_3 = 0$

$H_0: \beta_3 = 0$

$H_a: \beta_3 \neq 0$

$$F^* = \frac{\frac{SSR(X_3|X_1,X_2,X_4)}{1}}{\frac{SSE(X_1,X_2,X_3,X_4)}{n-4}} = \frac{\frac{0.42}{1}}{\frac{98.2306}{76}} = 0.3249$$

$F(0.99,1,76) = 6.9806$

$If\ F^* \leq F(1-\alpha,1,n-4)\ Conclude\ H_0$

$If\ F^* > F(1-\alpha,1,n-4)\ Conclude\ H_a$

$Since\ F^*(0.3249) \leq F(6.9806)\ Conclude\ H_0$

$P\ value = 0.5704$

7.8)

$H_0: \beta_2 = \beta_3 = 0$

$H_a: \beta_2\ and\ \beta_3 \neq 0$

$SSR(X_2,X_3 \mid X_1,X_4) = 28.277$

$SSE(X_1,X_2,X_3,X_4) = 98.2306$

$$F^* = \frac{\frac{SSR(X_2,X_3 \mid X_1,X_4)}{2}}{\frac{SSE(X_1,X_2,X_3,X_4)}{n-4}} = \frac{\frac{28.277}{2}}{\frac{98.2306}{76}} = 10.9288$$

$F(0.99, 0.2, 20) = 4.8958$

$If\ F^* \leq F(1-\alpha, 1, n-4)\ Conclude\ H_0$

$If\ F^* > F(1-\alpha, 1, n-4)\ Conclude\ H_a$

$$Since\ F^*(10.9288) > F(4.8958)\ Conclude\ H_a$$

7.10)

$H_0: \beta_1 = -1, \qquad \beta_2 = 0$

$H_a: Not\ both\ equalities\ hold$

$Full\ Model: Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \varepsilon_i$

$Reduced\ Model = Y_i = \beta_0 + \beta_3 X_{i3} + \varepsilon_i$

$SSE(F) = 4{,}248.84$

$df_F = 42$

$SSE(R) = 4{,}427.7$

$df_R = 44$

$$F^* = \frac{\frac{4{,}427.7 - 4{,}248.84}{2}}{\frac{4{,}248.84}{42}} = 0.8840$$

$F(0.99, 2, 76) = 4.89584$

$Since\ F^*(0.8840) \leq F(4.89584)\ Conclude\ H_0$

This implies that the reduced model is adequate.

7.27)

a)

Code –

```
model <- lm(V1 ~ V2+V5, data=data)
coefficients(model)
```

Output –

```
(Intercept)          V2          V5
 1.436128e+01 -1.144670e-01  1.044493e-05
```

Therefore, the regression equation is

$$Y = 1.436128\mathrm{e} + 01 + -1.144670\mathrm{e} - 01 \times X_1 + 1.044493\mathrm{e} - 05 \times X_4$$

b)

$$Y = 12.2 - 0.142 \times X_1 + 0.282 \times X_2 + 0.6193 \times X_3 + 0.000007924 \times X_4$$

Regression parameters obtained from 6.18c are smaller than that of the current model.

The current model is a better fit

c)

SSR(X4) = 67.7751

SSR(X4|X3) = 66.8582

Clearly, they are not equal.

SSR(X1) = 14.8185

SSR(X1|X3) = 13.7744

Clearly, they are not equal.

d)

X3 and X4 are weakly correlated and hence SSR(X4) and SSR(X4|X3) are close.

X1 and X3 are also weakly correlated but they are more correlated as compared to X3 and X4 and hence their difference for SSR is slightly more

r12 = .4670, r13 = .3228, r23 = .2538

8.3)

a) The $R^2 = 0.991$ is very high can imply that a few of the independent variables are highly correlated with Y.

b) ${R_a}^2$ is the adjusted R^2 and it decreases as we keep on adding more independent variables so it is better to use ${R_a}^2$

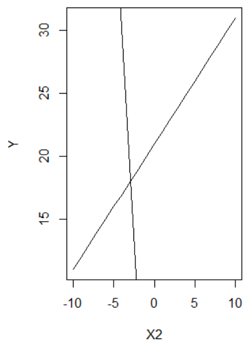
8.10)

a)

Code –

```
X1<-seq(-10, 10, by=1)

X2<-seq(-10, 10, by=1)

n<-length(X1)

X1 = 1

Y<- 14 + 7*X1 + 5*X2 - 4*(X1*X2)

plot(X2, Y, type='l', lty=1)

X1 = 4

Y<- 14 + -7*X1 + 5*X2 - 4*(X1*X2)

abline(lm(Y~X2))
```

Plot –



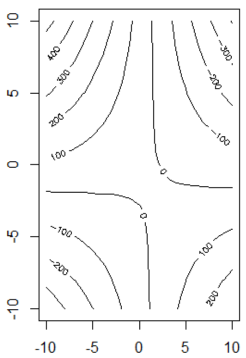
As the plot is not linear it is not additive in nature. The effect of X1 and X2 on Y are not additive as the lines are not parallel. The interaction effect of X1 and X2 is an interference.

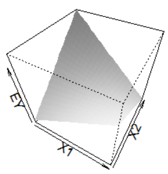
b)

Code –

```
X1<-seq(-10, 10, by=1)
X2<-seq(-10, 10, by=1)
n<-length(X1)
EY<-matrix(rep(0, n^2), n, n)
j<-1
while(j<=n){
  k<-1
  while(k<=n){
    EY[j,k]<- 14 + -7*X1[j] + 5*X2[k] - 4*(X1[j]*X2[k])
    k<-k+1
  }
  j<-j+1
}
contour(X1, X2, EY)
persp(X1, X2, EY, phi=45, theta=30, shade=0.3, border=NA)
```

Plot –





Additive or non-interacting predictor variables lead to parallel contour curves.

The plot has curved contour lines and hence the interaction effect is not additive.

From the plot we can say that it is a interference interaction effect.

8.12)

This could be due to the Linearly dependent columns. This can occur due to one or more reason the most probable one being that person might have misused the indicator variables including another indicator column, which resulted in the creation of linearly dependent columns and the result will be the creation of a Singular matrix where the determinant of the matrix is zero.

8.14)

T test statistic is used to calculate whether there is effect on learning time for a certain task or not

T = b2 coefficient / its standard error

= 22.3/3.8 = 5.868

Decision rule : If p value < 0.05, then there is regression model is effective

Comment : There is strong effect of learning time for certain task since

p value = p [T > 5.868] = 0.000 > 0.05

8.16)

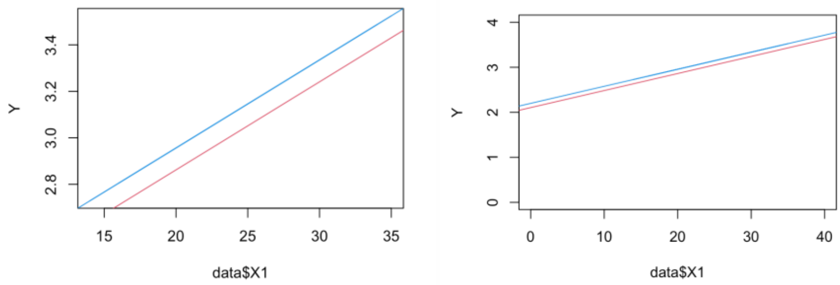
a)

Code –

```
data <- read.table("C:/Users/Aditya
Gaitonde/Downloads/CH01PR19.txt", header = FALSE,
                    col.names =c('Y','X1'))
data['X2'] <- read.table("C:/Users/Aditya
Gaitonde/Downloads/CH08PR16.txt", header = FALSE)
model <- lm(Y ~ ., data=data)
X2 = 0
Y<- 2.19841929 + (0.03789396)*data$X1 + (-0.09430392)*X2
plot(data$X1, Y, type='l', lty=1, col=4)
abline(lm(Y~data$X1), col=4)
X2 = 1
Y<- 2.19841929 + (0.03789396)*data$X1 + (-0.09430392)*X2
abline(lm(Y~data$X1), col=2)
X2 = 0
Y<- 2.19841929 + (0.03789396)*data$X1 + (-0.09430392)*X2
plot(data$X1, Y, type='l', lty=1, col=4, xlim=c(0,40), ylim=c(0,4))
```

```
abline(lm(Y~data$X1), col=4)
X2 = 1
Y<- 2.19841929 + (0.03789396)*data$X1 + (-0.09430392)*X2
abline(lm(Y~data$X1), col=2)
```

Plot –



The red line represents the case when X2 = 1 and blue represents when X2 = 0

$$E\{Y\} = \beta_0 + \beta_1 X_1 + \beta_2 X_2$$

Where, X_2 is an indicator variable.

Case 1: When $X_2 = 0$

$$E\{Y\} = \beta_0 + \beta_1 X_1.$$

This is a straight line with Y intercept β_0 and slope β_1

Case 2: When $X_2 = 1$

$$E\{Y\} = \beta_0 + \beta_2 + \beta_1 X_1$$

This is a straight line with the same slope but with Y intercept $(\beta_0 + \beta_2)$

b)

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2$$

$$Y = 2.1984 + (0.0378)X_1 - (0.094)X_2$$

c)

$$H_0: \beta_2 = 0$$

$$H_a: \beta_2 \neq 0$$

$$s\{b_2\} = 0.11997$$

$$t^* = \frac{-0.09430}{0.11997} = -0.786$$

$$t(0.995, 117) = 2.6185$$

$$If\ |t^*| \leq t(0.995, 117)\ Conclude\ H_0$$

$$If\ |t^*| > t(0.995, 117)\ Conclude\ H_a$$

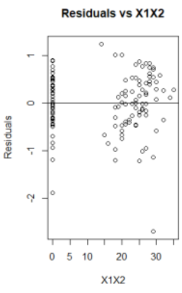
$$Since\ |t^*|(0.786) \leq t(2.6185)\ Conclude\ H_0$$

d)

Code –

```
model <- lm(Y ~ ., data=data)

residuals <- residuals(model)

plot(data$X1*data$X2, residuals, main="Residuals vs X1X2",
xlab="X1X2", ylab="Residuals")

abline(0,0)
```

Plot –



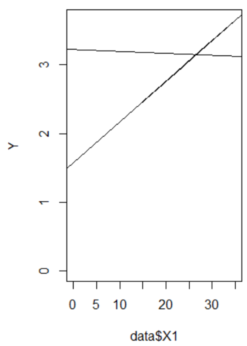
The residuals are identical along line (0,0) and maintain a constant variance with the presence of two outliers. Hence, we can include this interaction term in our model

8.20)

Code –

```
model <- lm(Y ~ X1+X2+X1*X2, data=data)
coefficients(model)
mse = anova(model)[4,3]
mean.x1x2 <- mean(data$X1*data$X2)
b <- (sum((data$X1*data$X2-mean.x1x2)^2))
s.square.b3 <- mse/b
s.b3 <- sqrt(s.square.b3)
t.stat = 0.06224465/s.b3
t.value = qt(0.975,77) # n-4
2*pt(t.stat, 77, lower.tail=FALSE)
X2 = 1
Y<- 3.22 + - 0.0027*data$X1 - 1.649*X2 + 0.0622*(data$X1*X2)
plot(data$X1, Y, type='l', lty=1, ylim=c(0,max(Y)),
xlim=c(0,max(data$X1)))
abline(lm(Y~data$X1))
X2 = 0
Y<- 3.22 + - 0.0027*data$X1 - 1.649*X2 + 0.0622*(data$X1*X2)
abline(lm(Y~data$X1))
```

Plot –



a)

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2$$

$$Y = 3.226 - 0.0027X_1 - 1.649X_2 + 0.0622X_1X_2$$

b)

$$H_0: \beta_3 = 0$$

$$H_a: \beta_3 \neq 0$$

$$s\{b_3\} = 0.02649$$

$$t^* = \frac{0.06224}{0.02649} = 2.3496$$

$$t(0.975, 116) = 1.9806$$

$$If\ |t^*| \leq t(0.975, 116)\ Conclude\ H_0$$

$$If\ |t^*| > t(0.975, 116)\ Conclude\ H_a$$

$$Since\ |t^*|(2.3496) > t(1.9806)\ Conclude\ H_a$$

Therefore, the term X_1X_2 cannot be dropped

$$p\ value = 4.084 \times 10^{-22}$$

which is less than 0.5 therefore, conclude H_a

Here we have one quantitative and one qualitative variable. Thus, the non-parallel response functions do not mean non additive.

According to the plot, the lines intersect within the scope of the model it is a disordinal interaction.

Q3)

Y = sale price (x $1,000)

x_1 = square footage (x 100)

x_2 = number of bedrooms

x_3 = number of bathrooms

x_4 = total number of rooms

x_5 = age of the home

x_6 = car garage (yes=1, no=1)

x_7 = good view (yes=1, no=0),

H0 : no difference in two models (full and reduced) so variable can be excluded

Ha : full model is significantly better, so variable must be included

p-value < f.stat => reject H0 - significantly lower SSE

Code –

```
data <- read.table("C:/Users/Aditya Gaitonde/Downloads/Homes1.txt", header = FALSE)
 head(data)
```

Output –

```
    V1   V2 V3 V4 V5 V6 V7 V8
1 50.6  8.0  2  1  5  5  0  0
2 51.5  9.5  2  1  5  8  0  0
3 53.3  9.1  3  1  6  2  0  0
4 65.9  9.5  3  1  6  6  0  0
5 67.4 12.0  3  2  7  5  0  0
6 68.9 10.0  3  1  6 11  0  0
```

Code –

```
model_full <- lm(V1 ~ ., data=data)
summary(model_full)
```

Output –

```
Residuals:
    Min       1Q   Median       3Q      Max
-16.5406  -4.4641  -0.5364   3.6054  30.3260

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -2.183e+01  1.014e+01  -2.153 0.037078 *
V2           2.933e+00  3.335e-01   8.794  4.5e-11 ***
V3           2.401e+00  2.902e+00   0.827 0.412686
V4          -8.747e+00  3.820e+00  -2.290 0.027139 *
V5           9.449e+00  2.568e+00   3.680 0.000659 ***
V6          -1.357e-04  3.169e-01   0.000 0.999660
V7           1.772e+00  3.040e+00   0.583 0.562972
V8           2.555e+00  4.544e+00   0.562 0.576941
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 9.092 on 42 degrees of freedom
Multiple R-squared:  0.8989,    Adjusted R-squared:  0.8821
F-statistic: 53.37 on 7 and 42 DF,  p-value: < 2.2e-16
```

Code –

```
anova(model_full)
```

Output –

```
Response: V1
          Df  Sum Sq Mean Sq  F value    Pr(>F)
V2         1 28650.4 28650.4 346.5691 < 2.2e-16 ***
V3         1   788.1   788.1   9.5336 0.0035675 **
V4         1     8.1     8.1   0.0980 0.7557843
V5         1  1374.7  1374.7  16.6295 0.0001981 ***
V6         1     2.0     2.0   0.0238 0.8780442
V7         1    35.5    35.5   0.4294 0.5158746
V8         1    26.1    26.1   0.3161 0.5769409
Residuals 42  3472.1    82.7
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Code –

```
model_red_1 <- lm(V1 ~ V3 + V4 + V5 + V6 + V7 + V8, data=data)
print(paste("Model Evaluation - Full vs Reduced (Excluding X", i," ))
cat("\n")
{
  value = anova(model_red_1, model_full)[2,6] < anova(model_red_1, model_full)[2,5]
  if(value == TRUE){
   print(paste("Reject H0 -> full model is significantly better -> X", 1, " should be included"))
```

```
} else{
 print(paste("Reject Ha -> reduced model is significantly better -> X", 1, " should be excluded"))  }}
```

Output –

"Reject H0 -> full model is significantly better -> X 1 should be included"

Code –

```
cat("\n")
model_red_2 <- lm(V1 ~ V2 + V4 + V5 + V6 + V7 + V8, data=data)
cat("\n")
 {
   value = anova(model_red_2, model_full)[2,6] < anova(model_red_2, model_full)[2,5]
   if(value == TRUE){
     print(paste("Reject H0 -> full model is significantly better -> X", 2, " should be included"))
   } else{
     print(paste("Reject Ha -> reduced model is significantly better -> X", 2, " should be excluded"))
   }
 }
```

Output –

"Reject H0 -> full model is significantly better -> X 2 should be included"

Code –

```
 model_red_3 <- lm(V1 ~ V2 + V3 + V5 + V6 + V7 + V8, data=data)
 cat("\n")
 {
   value = anova(model_red_3, model_full)[2,6] < anova(model_red_3,
model_full)[2,5]
   if(value == TRUE){
     print(paste("Reject H0 -> full model is significantly better -> X", 3, "
should be included"))
   } else{
     print(paste("Reject Ha -> reduced model is significantly better -> X", 3, "
should be excluded"))
   }
 }
```

Output –

```
"Reject H0 -> full model is significantly better -> X 3  should be included"
```

Code –

```
 model_red_4 <- lm(V1 ~ V2 + V3 + V4 + V6 + V7 + V8, data=data)
cat("\n")
 {
   value = anova(model_red_4, model_full)[2,6] < anova(model_red_4,
model_full)[2,5]
   if(value == TRUE){
     print(paste("Reject H0 -> full model is significantly better -> X", 4, "
should be included"))
   } else{
```

```
    print(paste("Reject Ha -> reduced model is significantly better -> X", 4, " should be excluded"))
  } }
```

Output –

"Reject H0 -> full model is significantly better -> X 4 should be included"

Code –

```
 model_red_5 <- lm(V1 ~ V2 + V3 + V4 + V5 + V7 + V8, data=data)
 cat("\n")
 {
   value = anova(model_red_5, model_full)[2,6] < anova(model_red_5, model_full)[2,5]
   if(value == TRUE){
     print(paste("Reject H0 -> full model is significantly better -> X", 5, " should be included"))
   } else{
     print(paste("Reject Ha -> reduced model is significantly better -> X", 5, " should be excluded"))
   }
 }
```

Output –

"Reject Ha -> reduced model is significantly better -> X 5 should be excluded"

Code –

```
model_red_6 <- lm(V1 ~ V2 + V3 + V4 + V5 + V6 + V8, data=data)
cat("\n")
```

```
 {
   value = anova(model_red_6, model_full)[2,6] < anova(model_red_6, model_full)[2,5]
   if(value == TRUE){
     print(paste("Reject H0 -> full model is significantly better -> X", 6, " should be included"))
   } else{
     print(paste("Reject Ha -> reduced model is significantly better -> X", 6, " should be excluded"))
   }
 }
```

Output –

"Reject Ha -> reduced model is significantly better -> X 6 should be excluded"

Code –

```
 model_red_7 <- lm(V1 ~ V2 + V3 + V4 + V5 + V6 + V7, data=data)
cat("\n")
 {
   value = anova(model_red_7, model_full)[2,6] < anova(model_red_7, model_full)[2,5]
   if(value == TRUE){
     print(paste("Reject H0 -> full model is significantly better -> X", 7, " should be included"))
   } else{
     print(paste("Reject Ha -> reduced model is significantly better -> X", 7, " should be excluded"))
   } }
```

Output –

"Reject Ha -> reduced model is significantly better -> X 7 should be excluded"

Conclusion -> X1, X2, X3, X4 should be included and X5, X6, X7 should be excluded from the model

Code –

```
new_model <- lm(V1 ~ V2+V3+V4+V5, data=data)
 summary(new_model) # R^2 - 0.8971
```

Output –

```
Residuals:
    Min      1Q  Median      3Q     Max
-16.085  -4.959  -1.420   3.864  30.434
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -23.1468     9.5802  -2.416 0.019808 *
V2            3.0084     0.2987  10.070 4.18e-13 ***
V3            2.0141     2.7827   0.724 0.472930
V4           -9.4913     3.5920  -2.642 0.011285 *
V5            9.9403     2.3764   4.183 0.000131 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 8.864 on 45 degrees of freedom
Multiple R-squared:  0.8971,    Adjusted R-squared:  0.8879
F-statistic: 98.07 on 4 and 45 DF,  p-value: < 2.2e-16
```

Code –

```
summary(model_full)
```

Output –

```
Residuals:
     Min       1Q   Median       3Q      Max
-16.5406  -4.4641  -0.5364   3.6054  30.3260

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -2.183e+01  1.014e+01  -2.153 0.037078 *
V2           2.933e+00  3.335e-01   8.794  4.5e-11 ***
V3           2.401e+00  2.902e+00   0.827 0.412686
V4          -8.747e+00  3.820e+00  -2.290 0.027139 *
V5           9.449e+00  2.568e+00   3.680 0.000659 ***
V6          -1.357e-04  3.169e-01   0.000 0.999660
V7           1.772e+00  3.040e+00   0.583 0.562972
V8           2.555e+00  4.544e+00   0.562 0.576941
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 9.092 on 42 degrees of freedom
Multiple R-squared:  0.8989,    Adjusted R-squared:  0.8821
F-statistic: 53.37 on 7 and 42 DF,  p-value: < 2.2e-16
```

Code –

anova(new_model, model_full)

Output –

Model 1: V1 ~ V2 + V3 + V4 + V5

Model 2: V1 ~ V2 + V3 + V4 + V5 + V6 + V7 + V8

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	45	3535.7				
2	42	3472.1	3	63.598	0.2564	0.8563

There isn't much difference in R^2 for both models so we have chosen the right variables on comparing anova tables we can see that p-value > f.stat => there is no significance difference between the new reduced model with variables selected and the full model

Therefore, we can exclude variables X1, X2, X3, X4.

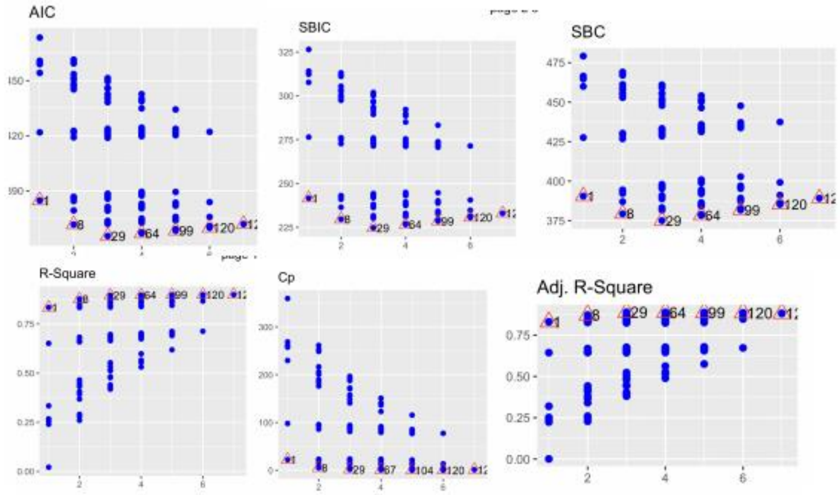
Code –

library('olsrr')

k <- ols_step_all_possible(model_full)

plot(k)

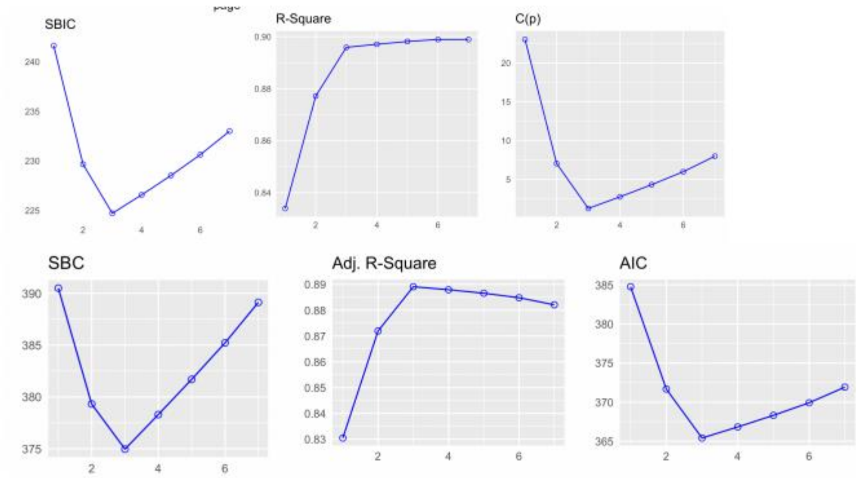
Output –



Code –

```
k <- ols_step_best_subset(model_full)
plot(k)
```

Output –



Code –

```
ols_step_best_subset(model_full)
```

Output –

```
      Best Subsets Regression
----------------------------------
Model Index    Predictors
----------------------------------
     1         V2
     2         V2 V5
     3         V2 V4 V5
     4         V2 V3 V4 V5
     5         V2 V3 V4 V5 V7
     6         V2 V3 V4 V5 V7 V8
     7         V2 V3 V4 V5 V6 V7 V8
```

Subsets Regression Summary

Model	R-Square	Adj. R-Square	Pred R-Square	C(p)	AIC	SBIC	SBC	MSEP	FPE	HSP	APC
1	0.8339	0.8304	0.8219	23.0304	384.7621	241.6135	390.4981	5944.6280	123.6441	2.5295	0.1799
2	0.8772	0.8719	0.8548	7.0480	371.6730	229.6552	379.3211	4491.6208	95.1760	1.9519	0.1385
3	0.8959	0.8891	0.8636	1.2672	365.4045	224.7071	374.9646	3891.6058	83.9781	1.7279	0.1222
4	0.8971	0.8879	0.8603	2.7693	366.8258	226.5654	378.2979	3934.2481	86.4277	1.7857	0.1258
5	0.8981	0.8866	0.8578	4.3297	368.3091	228.5241	381.6933	3984.3595	89.0740	1.8495	0.1296
6	0.8989	0.8848	0.8497	6.0000	369.9182	230.6184	385.2144	4047.4561	92.0506	1.9225	0.1340
7	0.8989	0.8821	0.8442	8.0000	371.9182	232.9994	389.1264	4146.1745	95.8956	2.0163	0.1396

AIC: Akaike Information Criteria

SBIC: Sawa's Bayesian Information Criteria

SBC: Schwarz Bayesian Criteria

MSEP: Estimated error of prediction, assuming multivariate normality

FPE: Final Prediction Error

HSP: Hocking's Sp

APC: Amemiya Prediction Criteria

The selection of the variables is correct as we have chosen one with large R squared, small MSE and smaller bias of Cp

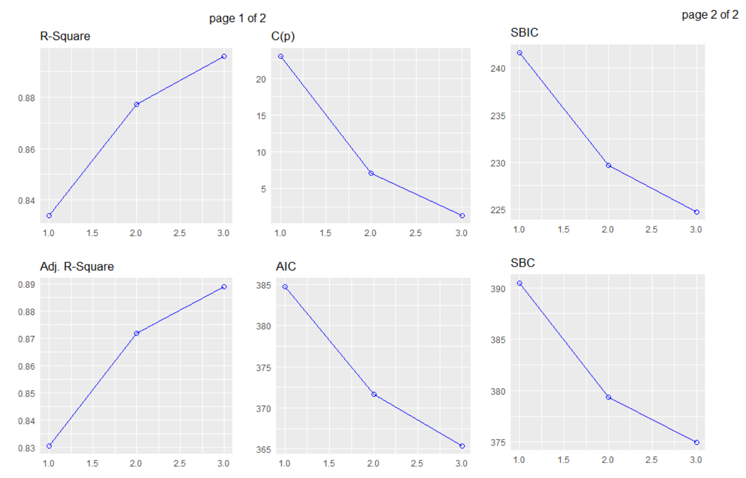
Best with Forward stepwise based on p

Code –

k <- ols_step_forward_p(model_full)

plot(k)

Plot –



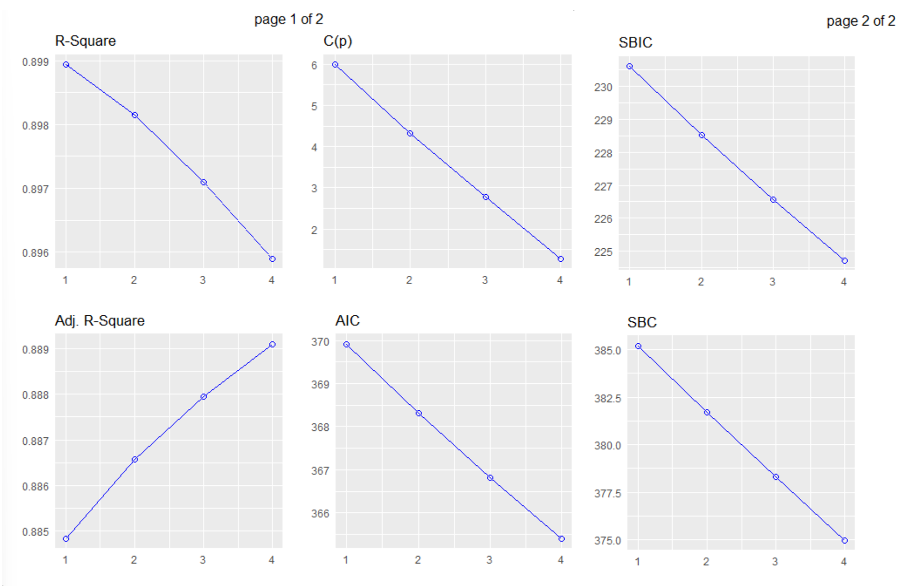
Best with Backward stepwise based on p

Code –

k <- ols_step_backward_p(model_full)

plot(k)

Plot –

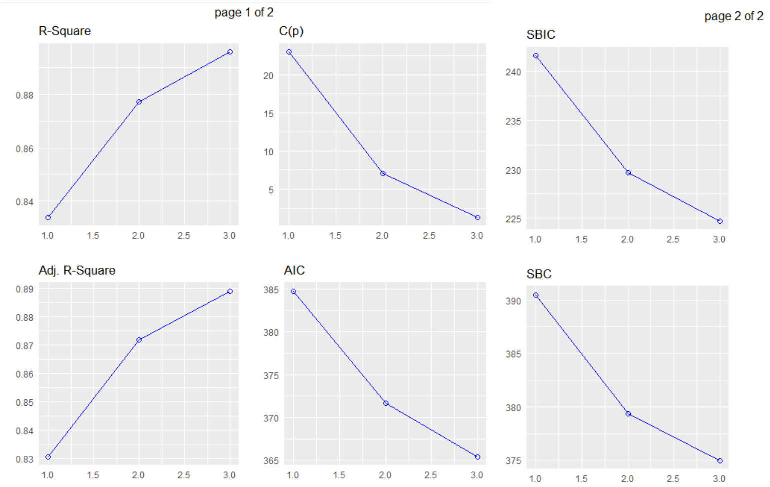


Best with Stepwise based on p

Code –

```
k <- ols_step_both_p(model_full)
plot(k)
```

Plot –

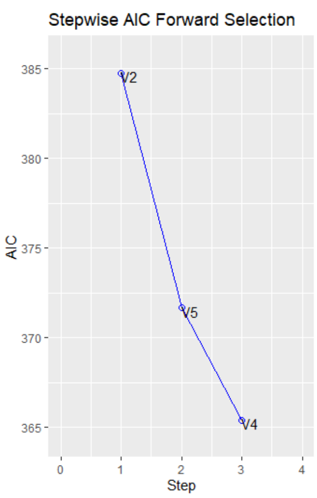


Best with Forward stepwise based on aic

Code –

```
k <- ols_step_forward_aic(model_full)
plot(k)
```

Plot –



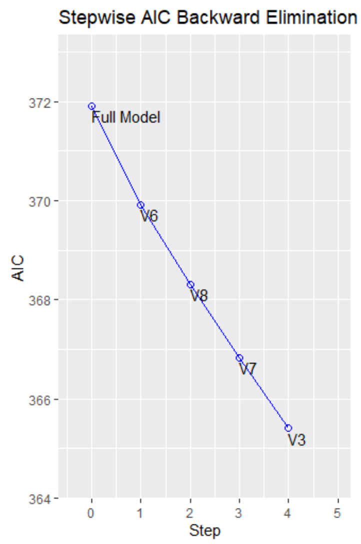
Best with Backward stepwise based on aic

Code –

k <- ols_step_backward_aic(model_full)

plot(k)

Plot –



Best with Stepwise based on aic

Code –

k <- ols_step_both_aic(model_full)

plot(k)

Plot –