Homework1_MSA8150

Anutida Sangkla

Problem 1

part (a) The RSS formulation

From

$$y = \beta_1 x_1 + \beta_2 x_2$$

We will get the RSS formulation for this problem as:

$$RSS = \sum_{i=1}^{n} (y_i - \beta_1 x_{1,i} - \beta_2 x_{2,i})^2 \quad (1)$$

part (b)

$$a\beta_1 + b\beta_2 = c \quad (1)$$

$$d\beta_1 + e\beta 2 = f \quad (2)$$

From(1), we will get

$$\beta_1 = \frac{c - b\beta_2}{a} \quad (3)$$

Substitute (3) into (2)

$$d(\frac{c - b\beta_2}{a}) + e\beta 2 = f$$

$$\frac{cd - bd\beta_2}{a} + \frac{ae\beta_2}{a} = f$$

Then, solve for

$$\beta_2 = \frac{af - cd}{ae - bd}$$

From(1), we will get

$$\beta_2 = \frac{c - a\beta_1}{b} \quad (4)$$

Substitute (4) into (2)

$$d\beta_1 + e(\frac{c - a\beta_1}{b}) = f$$

$$\frac{bd\beta_1}{b} + \frac{ce - ae\beta_1}{b} = f$$

Then, solve for

$$\beta_1 = \frac{ce - bf}{ae - bd}$$

part (c) Minimizing RSS

From (1), we can minimize RSS by using partial derivative

$$\frac{\partial RSS}{\beta_1}(\beta_1,\ \beta_2) = \sum_{i\,=\,1}^{n} -2x_{1,i}(y_i - \beta_1 x_{1,i} - \beta_2 x_{2,i}) = 0$$

$$\sum_{i\,=\,1}^{n} (\,y_i x_{1,i} - \beta_1 x_{1,i}^2 - \beta_2 x_{2,i} x_{1,i}) = 0$$

$$(\sum_{i\,=\,1}^{n} y_i\, x_{1,i} - \beta_1 \sum_{i\,=\,1}^{n} x_{1,i}^2 - \beta_2 \sum_{i\,=\,1}^{n} x_{2,i}\, x_{1,i}) = 0$$

Therefore, from part (b), we will get

$$\sum_{i\,=\,1}^{n} y_i\, x_{1,i} \quad (c)$$

$$\sum_{i\,=\,1}^{n} x_{1,i}^2 \quad (a)$$

$$\sum_{i\,=\,1}^{n} x_{1,i}\, x_{2,i} \quad (b)$$

$$\frac{\partial RSS}{\beta_2}(\beta_1,\ \beta_2) = \sum_{i\,=\,1}^{n} -2x_2(y_i - \beta_1 x_{1,i} - \beta_2 x_{2,i}) = 0$$

$$\sum_{i\,=\,1}^{n} (-y_i x_{2,i} - \beta_1 x_{2,i} x_{1,i} - \beta_2 x_{2,i}^2) = 0$$

$$\sum_{i\,=\,1}^{n} -y_i x_{2,i} + \beta_1 \sum_{i\,=\,1}^{n} x_{2,i}\, x_{1,i} + \beta_2 \sum_{i\,=\,1}^{n} x_{2,i}^2 = 0$$

$$\beta_1 \sum_{i=1}^{n} x_{2,i}\, x_{1,i} + \beta_2 \sum_{i=1}^{n} x_{2,i}^2 = \sum_{i=1}^{n} y_i\, x_{2,i}$$

From part (b)

$$\sum_{i=1}^{n} y_i\, x_{2,i} \quad (f)$$

$$\sum_{i=1}^{n} x_{2,i}\, x_{1,i} \quad (d)$$

$$\sum_{i=1}^{n} x_{2,i}^2 \quad (e)$$

From part (b), results

$$\hat{\beta}_1 = \frac{ce - bf}{ae - bd} = \frac{(\sum_{i=1}^{n} y_i\, x_{1,i})(\sum_{i=1}^{n} x_{2,i}^2) - (\sum_{i=1}^{n} y_i\, x_{2,i})(\sum_{i=1}^{n} x_{1,i}\, x_{2,i})}{(\sum_{i=1}^{n} x_{1,i}^2)(\sum_{i=1}^{n} x_{2,i}^2) - (\sum_{i=1}^{n} x_{2,i}\, x_{1,i})^2}$$

$$\hat{\beta}_2 = \frac{af - cd}{ae - bd} = \frac{(\sum_{i=1}^{n} y_i\, x_{2,i})(\sum_{i=1}^{n} x_{1,i}^2) - (\sum_{i=1}^{n} y_i\, x_{1,i})(\sum_{i=1}^{n} x_{2,i}\, x_{1,i})}{(\sum_{i=1}^{n} x_{1,i}^2)(\sum_{i=1}^{n} x_{2,i}^2) - (\sum_{i=1}^{n} x_{2,i}\, x_{1,i})^2}$$

part (d)

Note that

$$\hat{\beta}_1 = \frac{(\sum_{i=1}^{n} y_i\, x_{1,i})(\sum_{i=1}^{n} x_{2,i}^2) - (\sum_{i=1}^{n} y_i\, x_{2,i})(\sum_{i=1}^{n} x_{1,i}\, x_{2,i})}{(\sum_{i=1}^{n} x_{1,i}^2)(\sum_{i=1}^{n} x_{2,i}^2) - (\sum_{i=1}^{n} x_{2,i}\, x_{1,i})^2}$$

Therefore,

$$E(\hat{\beta}_1) = E(\frac{(\sum_{i=1}^{n} y_i\, x_{1,i})(\sum_{i=1}^{n} x_{2,i}^2) - (\sum_{i=1}^{n} y_i\, x_{2,i})(\sum_{i=1}^{n} x_{1,i}\, x_{2,i})}{(\sum_{i=1}^{n} x_{1,i}^2)(\sum_{i=1}^{n} x_{2,i}^2) - (\sum_{i=1}^{n} x_{2,i}\, x_{1,i})^2})$$

Since

$$y = \beta_x^* + \epsilon$$

Then,

$$E(y_i) = \beta_x^*$$

Assume that

$$a = \sum_{i=1}^{n} x_{2,i}^2$$

$$b = \sum_{i=1}^{n} x_{1,i}\, x_{2,i}$$

$$c = (\sum_{i=1}^{n} x_{1,i}^2)(\sum_{i=1}^{n} x_{2,i}^2) - (\sum_{i=1}^{n} x_{2,i}\, x_{1,i})^2$$

$$E(\hat{\beta}_1) = \beta_1^*(\sum_{i=1}^{n} x_{1,i}^2)\frac{a}{c} - \beta_1^*(\sum_{i=1}^{n} x_{1,i}\, x_{2,i})\frac{b}{c}$$

$$\beta_1^*(\frac{a}{c}(\sum_{i=1}^{n} x_{1,i}^2) - \frac{b}{c}(\sum_{i=1}^{n} x_{1,i}\, x_{2,i}))$$

Substitute (a), (b), and (c)

$$\beta_1^*(\frac{\sum_{i=1}^{n} x_{2,i}^2)(\sum_{i=1}^{n} x_{1,i}^2) - (\sum_{i=1}^{n} x_{1,i}\, x_{2,i})(\sum_{i=1}^{n} x_{1,i}\, x_{2,i})}{(\sum_{i=1}^{n} x_{1,i}^2)(\sum_{i=1}^{n} x_{2,i}^2) - (\sum_{i=1}^{n} x_{2,i}\, x_{1,i})^2})$$

Therefore,

$$E(\hat{\beta}_1) = \beta^*$$

part (e)

$$\widetilde{RSS} = (\beta_1 - \beta_2)^2 + \sum_{i=1}^{n} (y_i - \beta_1 x_{1,i} - \beta_2 x_{2,i})$$

$$\frac{\partial \widetilde{RSS}}{\beta_1}(\beta_1,\ \beta_2) = \sum_{i=1}^{n} [-2(\beta_1 - \beta_2) - 2x_{1,i)}(y_i - \beta_1 x_{1,i} - \beta_2 x_{2,i})] = 0$$

$$\sum_{i=1}^{n} [(-y_i x_{1,i} + \beta_1(1 + x_{1,i}^2) + \beta_2(-1 + x_{1,i} x_{2,i})] = 0$$

$$(\sum_{i=1}^{n} - y_i x_{1,i} + \beta_1 \sum_{i=1}^{n} (1 + x_{1,i}^2) + \beta_2 \sum_{i=1}^{n} (-1 + x_{2,i} x_{1,i}) = 0$$

$$\beta_1 \sum_{i=1}^{n} (1 + x_{1,i}^2) + \beta_2 \sum_{i=1}^{n} (-1 + x_{2,i} x_{1,i}) = (\sum_{i=1}^{n} y_i\, x_{1,i})$$

Therefore, from part (b), we will get

$$\sum_{i=1}^{n} y_i\, x_{1,i} \quad (c)$$

$$(1+\sum_{i=1}^{n} 1+x_{1,i}^2) \quad (a)$$

$$(-1+\sum_{i=1}^{n} x_{2,i}\, x_{1,i}) \quad (b)$$

$$\frac{\partial \widetilde{RSS}}{\beta_2}(\beta_1,\ \beta_2)=\sum_{i=1}^{n}[-2(\beta_1-\beta_2)-2x_{2,i}(y_i-\beta_1 x_{1,i}-\beta_2 x_{2,i})]=0$$

$$\sum_{i=1}^{n}[(-y_i x_{2,i}+\beta_1(-1+x_{2,i}x_{1,i})+\beta_2(1+x_{2,i}^2)]=0$$

$$(\sum_{i=1}^{n}-y_i x_{2,i}+\beta_1\sum_{i=1}^{n}(-1+x_{2,i}x_{1,i})+\beta_2\sum_{i=1}^{n}(1+x_{2,i}^2)=0$$

$$\beta_1\sum_{i=1}^{n}(-1+x_{2,i}x_{1,i})+\beta_2\sum_{i=1}^{n}(1+x_{2,i}^2)=(\sum_{i=1}^{n} y_i\, x_{2,i})$$

From part (b)

$$(\sum_{i=1}^{n} y_i\, x_{2,i}) \quad (f)$$

$$(-1+\sum_{i=1}^{n} x_{2,i}\, x_{1,i}) \quad (d)$$

$$(1+\sum_{i=1}^{n} 1+x_{2,i}^2) \quad (e)$$

From part (b), results

$$\tilde{\beta}_1=\frac{ce-bf}{ae-bd}=\frac{(\sum_{i=1}^{n} y_i\, x_{1,i})(1+\sum_{i=1}^{n} x_{2,i}^2)-(\sum_{i=1}^{n} y_i\, x_{2,i})(-1+\sum_{i=1}^{n} x_{1,i}\, x_{2,i})}{(1+\sum_{i=1}^{n} x_{1,i}^2)(1+\sum_{i=1}^{n} x_{2,i}^2)-(-1+\sum_{i=1}^{n} x_{2,i}\, x_{1,i})^2}$$

$$\tilde{\beta}_2=\frac{af-cd}{ae-bd}=\frac{(\sum_{i=1}^{n} y_i\, x_{2,i})(1+\sum_{i=1}^{n} x_{1,i}^2)-(\sum_{i=1}^{n} y_i\, x_{1,i})(-1+\sum_{i=1}^{n} x_{2,i}\, x_{1,i})}{(1+\sum_{i=1}^{n} x_{1,i}^2)(1+\sum_{i=1}^{n} x_{2,i}^2)-(-1+\sum_{i=1}^{n} x_{2,i}\, x_{1,i})^2}$$

part (f)

```
x1 = c(89.09,84.24,98.77,95.44,90.98,97.39,89.27,88.51,97.06,84.45)
x2 = c(78.48,70.56,93.52,86.72,79.20,91.36,80,76.96,92.56,66.40)
y =  c(113.27,109.77,130.08,120.45,115.09,125.37,116.22,112.08,127.85,107.61)
n = 15; sumyx1 = sum(y*x1); sumx22 = sum(x2^2); sumyx2 = sum(y*x2); sumx1x2 =
sum(x1*x2); sumx12 = sum(x1^2); sumx2x1_2 = sum(x2*x1)^2
```

```r
beta1 = ((sumyx1 * sumx22) - (sumyx2 * sumx1x2))/((sumx12 * sumx22) - sumx2x1
_2)
beta2 = ((sumyx2 * sumx12) - (sumyx1 * sumx1x2))/((sumx12 * sumx22) - sumx2x1
_2)
sprintf('%s = %3.4f', c('beta1', 'beta2'),c(beta1,beta2))

## [1] "beta1 = 1.0935" "beta2 = 0.2168"
```

From the calculation

$$\beta_1 = 1.0935$$

$$\beta_2 = 0.2168$$

part (g)

```r
x1 = c(89.09,84.24,98.77,95.44,90.98,97.39,89.27,88.51,97.06,84.45)
x2 = c(78.48,70.56,93.52,86.72,79.20,91.36,80,76.96,92.56,66.40)
y =  c(113.27,109.77,130.08,120.45,115.09,125.37,116.22,112.08,127.85,107.61)
model = lm(formula = y ~ x1 + x2 - 1)
summary(model)

##
## Call:
## lm(formula = y ~ x1 + x2 - 1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.71757 -1.33676 -0.03497  1.54760  2.35368
##
## Coefficients:
##    Estimate Std. Error t value Pr(>|t|)
## x1   1.0935     0.1189   9.196 1.58e-05 ***
## x2   0.2168     0.1328   1.632    0.141
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.865 on 8 degrees of freedom
## Multiple R-squared:  0.9998, Adjusted R-squared:  0.9998
## F-statistic: 2.001e+04 on 2 and 8 DF,  p-value: 1.595e-15

beta1<-model$coefficient[1]
beta2<-model$coefficient[2]
sprintf('%s = %3.4f', c('beta1', 'beta2'),c(beta1,beta2))

## [1] "beta1 = 1.0935" "beta2 = 0.2168"
```

From the summary

$$\beta_1 = 1.0935$$

$$\beta_2 = 0.2168$$

Problem 2

part (a)

The training MSE function is

$$M(\alpha) = \frac{1}{n}\sum_{i=1}^{n} (y_i - \alpha x_i)^2$$

From the God's model, we will get

$$M(\beta^*) = \frac{1}{n}\sum_{i=1}^{n} (y_i - \beta^* x_i)^2$$

$$E(M(\beta^*)) = E\,[\frac{1}{n}\sum_{i=1}^{n} (y_i - \beta^* x_i)^2]$$

$$E(M(\beta^*)) = \frac{1}{n}\sum_{i=1}^{n} \; E[y_i - \beta^* x_i]^2$$

$$E(M(\beta^*)) = \frac{1}{n}\sum_{i=1}^{n} \; E[\beta^* x_i + \epsilon - \beta^* x_i]^2$$

$$E(M(\beta^*)) = \frac{1}{n}\sum_{i=1}^{n} \; E(\epsilon^2)$$

$$E(M(\beta^*)) = \frac{1}{n}\sum_{i=1}^{n} \; \sigma^2$$

$$E(M(\beta^*)) = \frac{1}{n}\, n\sigma^2$$

$$E(M(\beta^*)) = \sigma^2$$

part (b)

As we know that we minimize the RSS to obtain the optimal value of the estimated parameters, we minimize the MSE function associated with the training data to obtain the optimal value of the estimated parameter $\hat{\beta}$

Thus, this estimator is less than the coefficient of MSE function as following:

$$M(\hat{\beta}) \leq M(\beta^*)$$

part (c)

The test MSE function is

$$\widetilde{M}(\alpha) = \frac{1}{n}\sum_{i=1}^{n}(\tilde{y}_i - \tilde{\alpha}x_i)^2$$

From the God's model, we will get

$$\widetilde{M}(\beta) = \frac{1}{n}\sum_{i=1}^{n}(\tilde{y}_i - \beta\tilde{x}_i)^2$$

$$E(\widetilde{M}(\hat{\beta})) = E[\frac{1}{n}\sum_{i=1}^{n}(\tilde{y}_i - \beta\tilde{x}_i)^2]$$

$$E(\widetilde{M}(\hat{\beta})) = \frac{1}{n}\sum_{i=1}^{n}E\,(\tilde{y}_i - \beta\tilde{x}_i)^2$$

Since we know God's model is

$$y = \beta_x^* + \epsilon$$

$$\tilde{y} = \beta^*\tilde{x}_i + \epsilon$$

Then,

$$E(\widetilde{M}(\hat{\beta})) = \frac{1}{n}\sum_{i=1}^{n}E\,(\beta^*\tilde{x}_i - \hat{\beta}\tilde{x}_i + \epsilon)^2$$

Expand equation

$$E(\widetilde{M}(\hat{\beta})) = \frac{1}{n}\sum_{i=1}^{n}E\,[(\beta^*\tilde{x}_i - \hat{\beta}\tilde{x}_i)^2 + 2((\beta^*\tilde{x}_i - \beta\tilde{x}_i)(\epsilon)) + \epsilon^2]$$

$$E(\widetilde{M}(\hat{\beta})) = \frac{1}{n}\sum_{i=1}^{n}[\,(\beta^*\tilde{x}_i - \hat{\beta}\tilde{x}_i)^2 + E(\epsilon^2)]$$

$$E(\widetilde{M}(\hat{\beta})) = \frac{1}{n}\sum_{i=1}^{n}[\,(\beta^*\tilde{x}_i - \hat{\beta}\tilde{x}_i)^2 + \sigma^2]$$

Therefore,

$$E(\widetilde{M}(\hat{\beta})) = \sigma^2 + \frac{1}{n}\sum_{i=1}^{n}(\beta^*\tilde{x}_i - \beta\tilde{x}_i)^2$$

part (d)

From (5),(7), (8), we know that

$$E(M(\hat{\beta})) \leq E(M(\beta^*))$$

Since

$$M(\hat{\beta}) \leq M(\beta^*)$$

and

$$E(M(\beta^*)) = \sigma^2$$

Therefore,

$$E(M(\hat{\beta})) \leq E(M(\beta^*)) = \sigma^2$$

From (8)

$$E(\widetilde{M}(\hat{\beta})) = \sigma^2 + \frac{1}{n}\sum_{i=1}^{n} \hat{\beta}\tilde{x}_i + \beta^*\tilde{x}_i)^2$$

As a result,

$$\sigma^2 \leq \sigma^2 + \frac{1}{n}\sum_{i=1}^{n} \hat{\beta}\tilde{x}_i + \beta^*\tilde{x}_i)^2$$

$$E(M(\hat{\beta})) \leq E(\widetilde{M}(\hat{\beta}))$$

Problem 3

```r
physicallaw <- read.csv("PhysicalLaw.csv", header=TRUE,  sep=",")
train <- physicallaw[1:200,]
test <- physicallaw[201:240,]
```

part (a)

```r
lm.fit <- lm(F ~ p1 + p2 +d, data = train)
summary(lm.fit)
```

```
##
## Call:
## lm(formula = F ~ p1 + p2 + d, data = train)
##
## Residuals:
```

```
##    Min      1Q Median      3Q     Max
## -53.57 -22.48 -11.82   4.22 456.90
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.017e+01  1.082e+01   0.940 0.348164
## p1           1.859e-05  7.403e-06   2.511 0.012845 *
## p2           1.494e-05  7.311e-06   2.043 0.042373 *
## d           -2.738e+00  7.000e-01  -3.912 0.000126 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 61.86 on 196 degrees of freedom
## Multiple R-squared:  0.1166, Adjusted R-squared:  0.1031
## F-statistic: 8.627 on 3 and 196 DF,  p-value: 2.09e-05
```

The fitted parameters

```
beta0<-lm.fit$coefficient[1]
beta1<-lm.fit$coefficient[2]
beta2<-lm.fit$coefficient[3]
beta3<-lm.fit$coefficient[4]
sprintf('%s = %3.6f', c('beta0', 'beta1','beta2','beta3'),c(beta0,beta1,beta2
,beta3))
```

```
## [1] "beta0 = 10.174819" "beta1 = 0.000019"  "beta2 = 0.000015"
## [4] "beta3 = -2.738264"
```

From the summary

$$\beta_0 = 10.174819$$

$$\beta_1 = 0.000019$$

$$\beta_2 = 0.000015$$

$$\beta_3 = -2.738264$$

95% Confidenet Interval

```
conf.lmfit <- round(confint(lm.fit), 6)
print(conf.lmfit)
```

```
##                  2.5 %    97.5 %
## (Intercept) -11.162853 31.512491
## p1            0.000004  0.000033
## p2            0.000001  0.000029
## d            -4.118821 -1.357707
```

R-squared and P-Value

```
R.squared = summary(lm.fit)$r.squared
sprintf('%s = %3.4f', 'R-squared',R.squared)

## [1] "R-squared = 0.1166"
```

From the summary, R-squared for this train model is 0.1166 , which means the model accounts for 11.66% of the variability in the data. In the other word, 11.66% of the variability in F has been explained by the linear relationship between independent and dependent variables.

Also, P-value of this test is 2.09e-05 or 0.0000209 and the p-values of each estimated parameter are as following:

```
p.valueBeta0 = coef(summary(lm.fit))[1, 4]
p.valueBeta1 = coef(summary(lm.fit))[2, 4]
p.valueBeta2 = coef(summary(lm.fit))[3, 4]
p.valueBeta3 = coef(summary(lm.fit))[4, 4]

sprintf('%s = %3.6f', c('P-value of Beta0', 'P-value of Beta1','P-value of Be
ta2','P-value of Beta3'),c(p.valueBeta0,p.valueBeta1,p.valueBeta2,p.valueBeta
3))

## [1] "P-value of Beta0 = 0.348164" "P-value of Beta1 = 0.012845"
## [3] "P-value of Beta2 = 0.042373" "P-value of Beta3 = 0.000126"
```

part (b)

From the p-values of each estimated parameter, there are three features that are significant with alpha = 0.05. p1 is significant since its p-value = 0.012845 is less than alpha 0.05. p2 is significant since its p-value = 0.042373 is less than alpha 0.05. d is significant since its p-value = 0.000126 is less than alpha 0.05.

Therefore, the features p1, p2, and d are significant problems with alpha = 0.05.

part (c)

```
Fpred <- predict(lm.fit,test)
Ftest <- test$F
RMSE <- sqrt(mean((Ftest-Fpred)^2))

sprintf('%s = %3.3f', 'The root mean squared of this model is',RMSE)

## [1] "The root mean squared of this model is = 29.902"
```

Problem 4

Read the data and split into train and test sets:

```
modelselection <- read.csv("ModelSelection.csv", header=TRUE,  sep=",")
train <- modelselection[1:1000,]
test <- modelselection[1001:1500,]
```

model selection

```
combs = combn(1:10, 3)
mse = c()
for (n in 1:ncol(combs)){
        fit.train = lm(y ~ I(x^combs[,n][1]) + I(x^combs[,n][2]) + I(x^combs[
,n][3]), data=train)
        ypred = predict(fit.train, newdata=test)
        ytest = test$y
        mse[n] = mean((ytest-ypred)^2)
}
print(mse)

##   [1] 0.0532580687 0.0738257056 0.0367787651 0.0736271285 0.0293425487
##   [6] 0.0748933637 0.0268965815 0.0766328290 0.0246967278 0.2995054271
##  [11] 0.0321220311 0.2986597550 0.0500998827 0.2989995104 0.0700327475
##  [16] 0.0104782849 0.0738964686 0.0043540149 0.0740135057 0.0025419055
##  [21] 0.0741489380 0.0192798660 0.2993542505 0.0379471887 0.3001328463
##  [26] 0.0581485274 0.0138603110 0.0749334867 0.0123047713 0.0753392851
##  [31] 0.0328262707 0.3011127573 0.0530948771 0.0313277249 0.0774199707
##  [36] 0.0515529538 0.0727583507 0.0239042925 0.0728138870 0.0252053635
##  [41] 0.0743402081 0.0285823262 0.0762949880 0.0639127647 0.0751951729
##  [46] 0.0551556816 0.0752906398 0.0484637175 0.0753819294 0.0642100356
##  [51] 0.0315719409 0.0658910292 0.0367861046 0.0679293955 0.0556230231
##  [56] 0.0752863146 0.0490375127 0.0753193858 0.0573742046 0.0437315512
##  [61] 0.0594220671 0.0508171688 0.0753882920 0.0528511248 0.0001034802
##  [66] 0.0727686239 0.0017661182 0.0728604262 0.0049791809 0.0729660978
##  [71] 0.0104342083 0.3003194754 0.0297917551 0.3006830981 0.0501857549
##  [76] 0.0121581505 0.0732594061 0.0151177104 0.0734849434 0.0314672023
##  [81] 0.3016834551 0.0517723249 0.0342288326 0.0751149625 0.0543919711
##  [86] 0.0638984994 0.0078362995 0.0639634110 0.0124212960 0.0640363539
##  [91] 0.0551335472 0.0753405932 0.0484386548 0.0753767050 0.0551799302
##  [96] 0.0183931466 0.0552308623 0.0484742200 0.0753766538 0.0485127112
## [101] 0.0179049063 0.0640855423 0.0219880846 0.0642370316 0.0369900375
## [106] 0.3045933401 0.0571531104 0.0407775217 0.0657133443 0.0607828096
## [111] 0.0551979859 0.0273081608 0.0553397439 0.0484603020 0.0756106800
## [116] 0.0486126211 0.0457653356 0.0568194392 0.0656327405 0.0501347682

min(mse)

## [1] 0.0001034802

n1<- combs[,65][1]
n2<- combs[,65][2]
n3<- combs[,65][3]
sprintf('%s = %1.0f', c('n1','n2','n3'),c(n1,n2,n3))

## [1] "n1 = 3" "n2 = 4" "n3 = 5"
```

Therefore, we choose n1 = 3, n2 = 4, and n3 = 5 to create model.

```r
best.model <- lm(y ~ I(x^3) + I(x^4) + I(x^5), data=train)
summary(best.model)

##
## Call:
## lm(formula = y ~ I(x^3) + I(x^4) + I(x^5), data = train)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.037467 -0.006586  0.000005  0.006643  0.030882
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.0000549  0.0003896  2566.7   <2e-16 ***
## I(x^3)      -2.9967670  0.0036997  -810.0   <2e-16 ***
## I(x^4)      -1.9997593  0.0011742 -1703.1   <2e-16 ***
## I(x^5)       3.9949471  0.0046391   861.1   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.009864 on 996 degrees of freedom
## Multiple R-squared:  0.9997, Adjusted R-squared:  0.9997
## F-statistic: 1.22e+06 on 3 and 996 DF,  p-value: < 2.2e-16

beta0<-best.model$coefficient[1]
beta1<-best.model$coefficient[2]
beta2<-best.model$coefficient[3]
beta3<-best.model$coefficient[4]
sprintf('%s = %3.6f', c('beta0', 'beta1','beta2','beta3'),
c(beta0,beta1,beta2,beta3))

## [1] "beta0 = 1.000055"  "beta1 = -2.996767" "beta2 = -1.999759"
## [4] "beta3 = 3.994947"
```

From the results

$$\beta_0 = 1.0000549$$

$$\beta_1 = -2.9967670$$

$$\beta_2 = -1.9997593$$

$$\beta_3 = 3.9949471$$

```r
ypred = predict(best.model, newdata=test)
ytest = test$y
MSE = mean((ytest - ypred)^2)
sprintf('%s = %3.10f', 'The mean squared of this model is',MSE)

## [1] "The mean squared of this model is = 0.0001034802"
```