GR5205_HW3_Multilinear Regression

Jennifer Lieu

10/22/2018

Problem 3:

```
#load dataset
setwd("/Users/jenniferlieu/Desktop/Linear Regression Models Data 1")
hw3 <- read.table("HW3Problem3.txt", header=T)

#set dataset into dataframe
hw3[[1]] <- unlist(as.numeric(hw3[[1]]))
hw3[[2]] <- unlist(as.numeric(hw3[[2]]))
hw3[[3]] <- unlist(as.numeric(hw3[[3]]))
hw3[[4]] <- unlist(as.numeric(hw3[[4]]))
hw3[[5]] <- unlist(as.numeric(hw3[[5]]))
```

3.1

```
fullmodel <- lm(hw3[[1]]~hw3[[2]]+hw3[[3]]+hw3[[4]]+hw3[[5]])
#coefficients of linear model shown below
fullmodel
```

```
##
## Call:
## lm(formula = hw3[[1]] ~ hw3[[2]] + hw3[[3]] + hw3[[4]] + hw3[[5]])
##
## Coefficients:
## (Intercept)     hw3[[2]]     hw3[[3]]     hw3[[4]]     hw3[[5]]
##   1.220e+01   -1.420e-01    2.820e-01    6.193e-01    7.924e-06
```

3.2 and 3.3

```
summary(lm(hw3[[1]]~hw3[[2]]+hw3[[3]]+hw3[[4]]+hw3[[5]]))
```

```
##
## Call:
## lm(formula = hw3[[1]] ~ hw3[[2]] + hw3[[3]] + hw3[[4]] + hw3[[5]])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.1872 -0.5911 -0.0910  0.5579  2.9441
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.220e+01  5.780e-01  21.110  < 2e-16 ***
## hw3[[2]]    -1.420e-01  2.134e-02  -6.655 3.89e-09 ***
## hw3[[3]]     2.820e-01  6.317e-02   4.464 2.75e-05 ***
## hw3[[4]]     6.193e-01  1.087e+00   0.570     0.57
## hw3[[5]]     7.924e-06  1.385e-06   5.722 1.98e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.137 on 76 degrees of freedom
```

```
## Multiple R-squared:  0.5847, Adjusted R-squared:  0.5629
## F-statistic: 26.76 on 4 and 76 DF,  p-value: 7.272e-14
```

5847% of the variation of the rental rates were explained by this model. Due to the high P(>|t|) values for each beta, we conclude that all the marginal relationships between the response variable and covariate work except the one for x3.

3.4

```
reducedmodel1 <- lm(hw3[[1]]~1)
anova(reducedmodel1, fullmodel)
```

```
## Analysis of Variance Table
##
## Model 1: hw3[[1]] ~ 1
## Model 2: hw3[[1]] ~ hw3[[2]] + hw3[[3]] + hw3[[4]] + hw3[[5]]
##   Res.Df     RSS Df Sum of Sq      F    Pr(>F)
## 1     80 236.558
## 2     76  98.231  4    138.33 26.756 7.272e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Because 7.272e-14<.05, we reject the null hypothesis at the 5% significance level thatthere is an overall relationship between the rental rates and all of the covariates.

3.5

```
reducedmodel2 <- lm(hw3[[1]]~hw3[[3]]+hw3[[5]])
anova(reducedmodel2, fullmodel)
```

```
## Analysis of Variance Table
##
## Model 1: hw3[[1]] ~ hw3[[3]] + hw3[[5]]
## Model 2: hw3[[1]] ~ hw3[[2]] + hw3[[3]] + hw3[[4]] + hw3[[5]]
##   Res.Df     RSS Df Sum of Sq      F    Pr(>F)
## 1     78 159.491
## 2     76  98.231  2    61.261 23.698 1.003e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Because 1.003e-08<.05, we reject the null hypothesis at the 5% significance level that the coefficients of X1 and X3 are not zero.

3.6

```
reducedmodel3 <- lm(hw3[[1]]~hw3[[2]]+hw3[[3]]+hw3[[5]])
anova(reducedmodel3, fullmodel)
```

```
## Analysis of Variance Table
##
## Model 1: hw3[[1]] ~ hw3[[2]] + hw3[[3]] + hw3[[5]]
## Model 2: hw3[[1]] ~ hw3[[2]] + hw3[[3]] + hw3[[4]] + hw3[[5]]
##   Res.Df    RSS Df Sum of Sq      F Pr(>F)
## 1     77 98.650
## 2     76 98.231  1   0.41975 0.3248 0.5704
```

Because 0.5704>.05, cannot reject the null hypothesis at the 5% significance level that the coefficient of X3 are not zero.The p-values are the same between the marginal t test, and the f test.

3.7

```r
fullmodel1 <- lm(RentalRates~Age+OperatingExpense+VacancyRates+SquareFootage, data=hw3)

X.h <- data.frame(Age=5.0, OperatingExpense=8.25, VacancyRates=0, SquareFootage=250000)
predict(fullmodel1,X.h,interval = "confidence", level= .9875)
```

```
##        fit      lwr      upr
## 1 15.79813 15.08664 16.50962
```

```r
X.h <- data.frame(Age=6.0, OperatingExpense=8.50, VacancyRates=0.23, SquareFootage=270000)
predict(fullmodel1,X.h,interval = "confidence", level= .9875)
```

```
##        fit      lwr      upr
## 1 16.02754 15.42391 16.63116
```

```r
X.h <- data.frame(Age=14.0, OperatingExpense=11.50, VacancyRates=0.11, SquareFootage=300000)
predict(fullmodel1,X.h,interval = "confidence", level= .9875)
```

```
##        fit      lwr      upr
## 1 15.90072 15.33232 16.46913
```

```r
X.h <- data.frame(Age=12.0, OperatingExpense=10.25, VacancyRates=0, SquareFootage=310000)
predict(fullmodel1,X.h,interval = "confidence", level= .9875)
```

```
##        fit     lwr      upr
## 1 15.84339 15.1804 16.50638
```

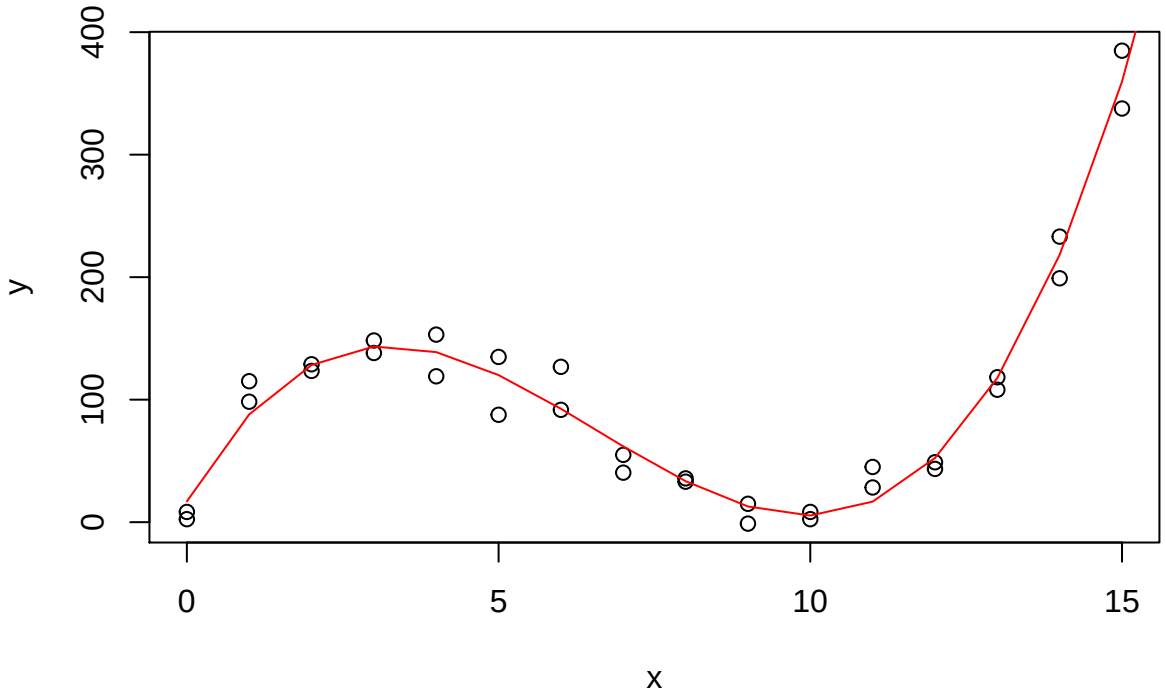
Problem 4:

```r
hw4 <- read.table("HW3Problem4.txt")
x<-hw4$x
y<-hw4$y
x1<-x
x2<-x^2
x3<-x^3
ymodel<- lm(y~x1+x2+x3)
ymodel
```

```
##
## Call:
## lm(formula = y ~ x1 + x2 + x3)
##
## Coefficients:
## (Intercept)           x1           x2           x3
##     16.9729      88.1660     -18.0874       0.9155
```

```r
z<- c(0:16)
plot(y~x)
lines(0:16,y=16.9729 + 88.1660*z - 18.0874*z*z +  0.9155*z*z*z, col="Red", add=TRUE)
```

```
## Warning in plot.xy(xy.coords(x, y), type = type, ...): "add" is not a
## graphical parameter
```


y
0
100
200
300
400
0
5
10
15
x