NTHU STAT 5410 - Linear Models

Assignment 4 Report

105061110 周柏宇

1.

```
> gala <- read.table("C:/Users/Thomas/Downloads/Linear_mo
dels/hw4/E3.7.txt", header=T)
> y <- gala[,7]
> x1 <- gala[,2]
> x2 <- gala[,3]
> x3 <- gala[,4]
> x4 <- gala[,5]
> x5 <- gala[,6]
> fit <- lm(y ~ x1 + x2 + x3 + x4 + x5)
> summary(fit)

Call:
lm(formula = y ~ x1 + x2 + x3 + x4 + x5)

Residuals:
     Min       1Q   Median       3Q      Max 
-0.39447 -0.11847  0.00053  0.08313  0.56232 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)  
(Intercept) -2.156e+00  9.135e-01  -2.360   0.0333 *
x1          -9.012e-06  5.184e-04  -0.017   0.9864  
x2           1.316e-03  1.263e-03   1.041   0.3153  
x3           1.278e-04  7.690e-05   1.662   0.1188  
x4           7.899e-03  1.400e-02   0.564   0.5815  
x5           1.417e-04  7.375e-05   1.921   0.0754 .
```

```
>cv <- qt(0.975, fit$df)
# the 95% critical value, dfW=n-p=14
```

(a)

```
> c(-9.012e-06 - cv * 5.184e-04, -9.012e-06 +  cv * 5.184e-
04)
[1] -0.001120869  0.001102845
```

The 95% C.I. for β_1 is [-0.001120869, 0.001102845].

(b)

```
> A <- t(c(0, 0, 0, 1, 0, 2))
> y0 <- sum(A * fit$coef)
> y0
[1] 0.0004111073
> x <- model.matrix(fit) # the model matrix X
> xtxi <- solve(t(x)%*%x) # (XTX)-1
> bm <- sqrt(A%*%xtxi%*%t(A)) * summary(fit)$sigma
> bm
             [,1]
[1,] 0.0001641751
> cv <- qt(0.975, fit$df)
> c(y0 - cv * bm, y0  +  cv * bm)
[1] 5.898666e-05 7.632279e-04
```

The 95% C.I. for $\beta_3 + 2\beta_5$ is [0.00005898666, 0.0007632279].

2.

```
> gala <- read.table("C:/Users/Thomas/Downloads/Linear_mo
dels/hw4/set.txt", header=T)
> PRICE <- gala[,1]
> BDR <- gala[,2]
> FLR <- gala[,3]
> FP  <- gala[,4]
> RMS <- gala[,5]
> ST  <- gala[,6]
> LOT <- gala[,7]
> TAX <- gala[,8]
> BTH <- gala[,9]
> CON <- gala[,10]
> GAR <- gala[,11]
```

```
> CDN <- gala[,12]
> L1  <- gala[,13]
> L2  <- gala[,14]
```

Since the new data we try to predict does not contain all the factors in the data set, we only use the factors that the new data has to fit our model.

```
> fit <-
lm(PRICE ~ BDR + FLR + FP + RMS + ST + LOT + BTH + GAR)
> summary(fit)

Call:
lm(formula =
PRICE ~ BDR + FLR + FP + RMS + ST + LOT + BTH + GAR)

Residuals:
     Min       1Q   Median       3Q      Max
-10.3058  -2.8417  -0.1511   3.2882   7.9518

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 18.637664   5.240957   3.556 0.002429 **
BDR         -7.697444   1.829426  -4.208 0.000592 ***
FLR          0.017570   0.003235   5.431 4.49e-05 ***
FP           6.909765   3.083583   2.241 0.038680 *
RMS          3.904374   1.615617   2.417 0.027194 *
ST          10.818663   2.300203   4.703 0.000205 ***
LOT          0.263522   0.135109   1.950 0.067808 .
BTH          2.374591   2.557865   0.928 0.366221
GAR          1.770861   1.404310   1.261 0.224334
---
```

The fitted model is

$$\begin{aligned} PRICE = 18.637664 - 7.697444 *BDR &+ 0.017570 *FLR + 6.909765 *FP \\ &+ 3.904374 *RMS + 10.818663 *ST + 0.263522 *LOT \\ &+ 2.374591 *BTH + 1.770861 *GAR \end{aligned}$$

The price we try to predict for the new data

```
> predict(fit, data.frame(BDR=2, FLR=750, FP=1, RMS=5, ST=
1, LOT=25, BTH=1.5, GAR=1), se=T, interval="prediction")
```

```
$fit
       fit      lwr      upr
1 65.59152 52.89018 78.29287

$se.fit
[1] 3.740865

$df
[1] 17

$residual.scale
[1] 4.716756
```

The predicted price is 65.59152.

The 95% C.I. for the predicted price is [52.89018 78.29287].