Linear HW#6

Tonghao Zhang

2016/11/13

7.22

There is one possible explaination of this phenanimon, multicolinearity among the predictor variables. Because of multicolinearity, two variable can fit the data fairly well while other coefficients are not statsistically significant. But multicolinearity does not inhibit us from obtaining precise estimate value. With more predictor included, the predictions will be more precise.

7.24

a)

```
data7.24<-read.table("CH06PR05.txt")
colnames(data7.24)<-c("Y","X1","X2")
lm7.24a<-lm(Y~X1,data=data7.24)
lm7.24a
```

```
##
## Call:
## lm(formula = Y ~ X1, data = data7.24)
##
## Coefficients:
## (Intercept)           X1
##      50.775        4.425
```

Y $= 50.775 + 4.425\ X_1$

b)

```
lm6.5<-lm(Y~X2+X1,data=data7.24)
lm6.5
```

```
##
## Call:
## lm(formula = Y ~ X2 + X1, data = data7.24)
##
## Coefficients:
## (Intercept)           X2           X1
##      37.650        4.375        4.425
```

```
cor(data7.24$X1,data7.24$X2)
```

```
## [1] 0
```

The coefficent β_1 from part a is identical to that of 6.5.

c)

```r
anova(lm7.24a)
```

```
## Analysis of Variance Table
##
## Response: Y
##           Df  Sum Sq Mean Sq F value    Pr(>F)
## X1         1 1566.45 1566.45  54.751 3.356e-06 ***
## Residuals 14  400.55   28.61
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
anova(lm6.5)
```

```
## Analysis of Variance Table
##
## Response: Y
##           Df  Sum Sq Mean Sq F value    Pr(>F)
## X2         1  306.25  306.25  42.219 2.011e-05 ***
## X1         1 1566.45 1566.45 215.947 1.778e-09 ***
## Residuals 13   94.30    7.25
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Yes, SSR(X_1) equals to SSR($X_1|X_2$).

d)

The correlation between $X1$ and X_2 is 0, so the infomation in these two variable does not overlap, the two models from 6.5 and 7.24 produce the same β_1, SSR(X_1) equals to SSR($X_1|X_2$).

Problem 3

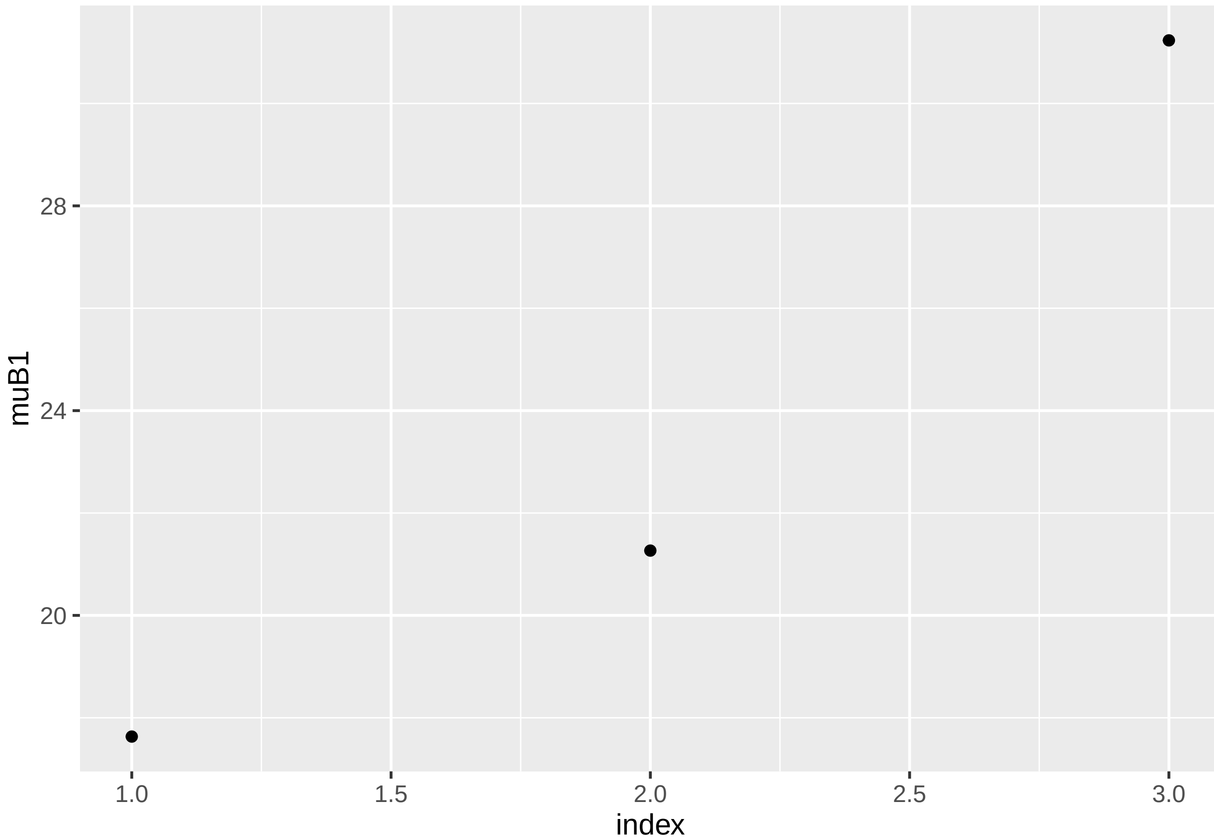
a)

```r
strength<-c(17.8,18.2,16.9,21.4,20.1,22.3,30.6,32.1,31.0,6.5,5.2,5.9,10.4,11.7,12.6,14.6,14.3,15.9)
A<-c(1,1,1,2,2,2,3,3,3,1,1,1,2,2,2,3,3,3)
B<-c(rep(1,9),rep(2,9))
group<-c(1,1,1,2,2,2,3,3,3,4,4,4,5,5,5,6,6,6)
data3<-data.frame(strength=strength,group=group,A=A,B=B)
lm3.a<-lm(strength~as.factor(group),data=data3)
lm3.a
```

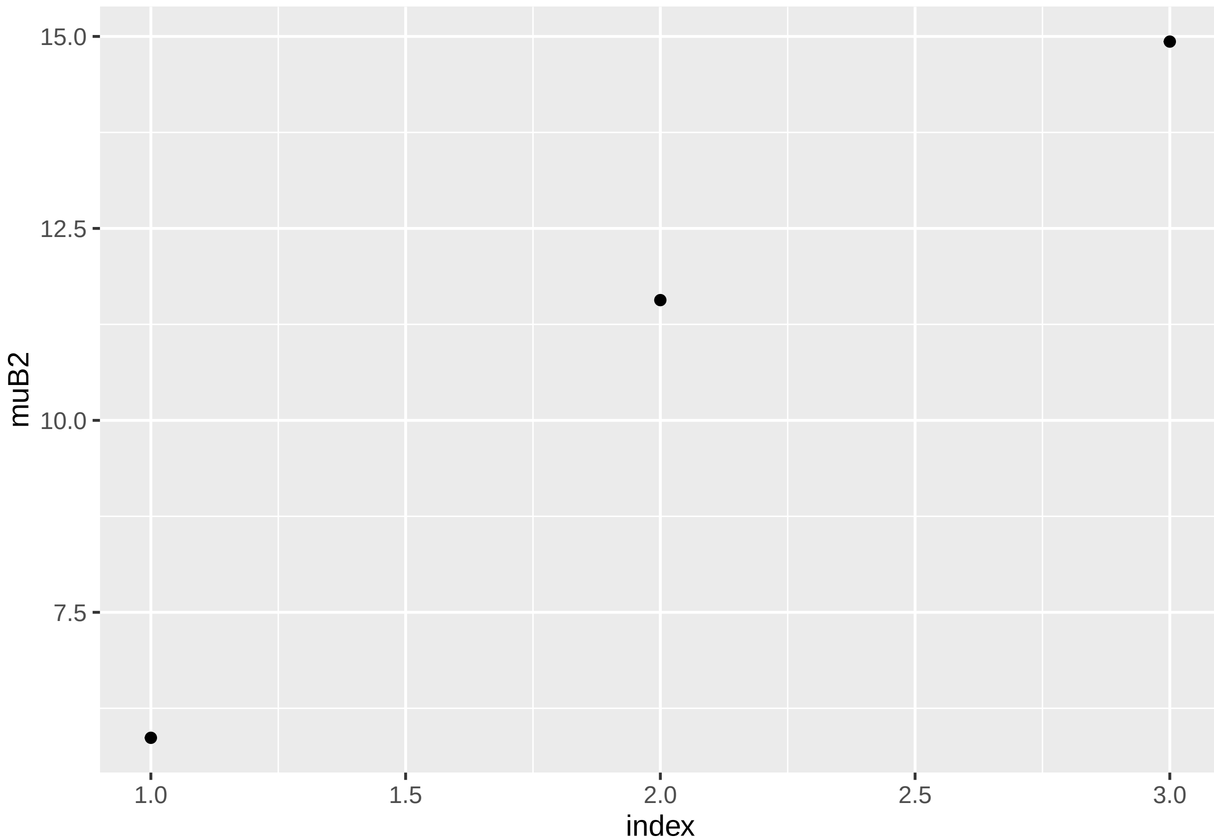
```
##
## Call:
## lm(formula = strength ~ as.factor(group), data = data3)
##
## Coefficients:
##       (Intercept)  as.factor(group)2  as.factor(group)3
##            17.633              3.633             13.600
## as.factor(group)4  as.factor(group)5  as.factor(group)6
##           -11.767             -6.067             -2.700
```

```r
library(ggplot2)
muB1<-c(17.633,17.633+3.633,17.633+13.600)
muB2<-c(17.633-11.767,17.633-6.067,17.633-2.7)
```

```
index<-c(1,2,3)
data3.plot<-data.frame(muB1=muB1,muB2=muB2,index=index)
ggplot(data3.plot,aes(x=index,y=muB1))+geom_point()
```




```
ggplot(data3.plot,aes(x=index,y=muB2))+geom_point()
```



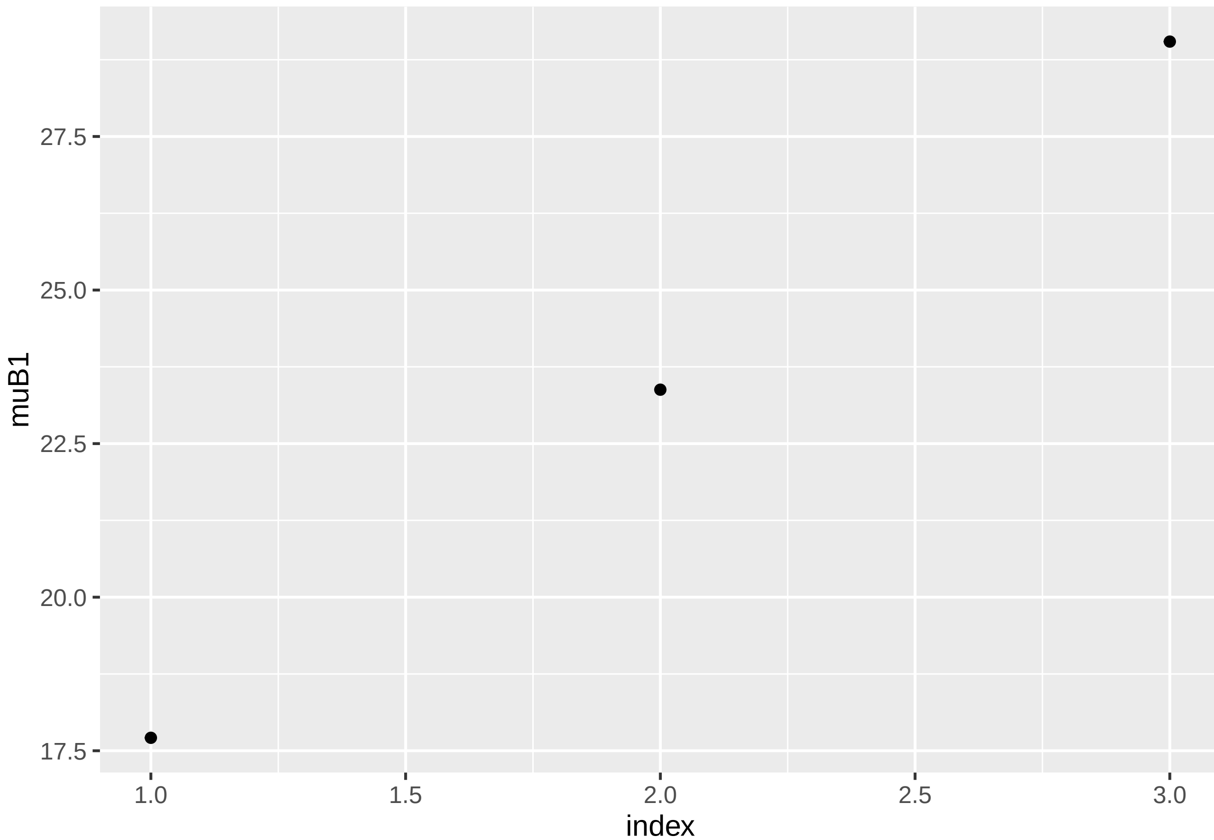
β_0 = 17.633, β_1 = 3.633, β_2 = 13.600, β_3 = -11.767, β_4 = -6.067, β_5 = -2.700.

b)

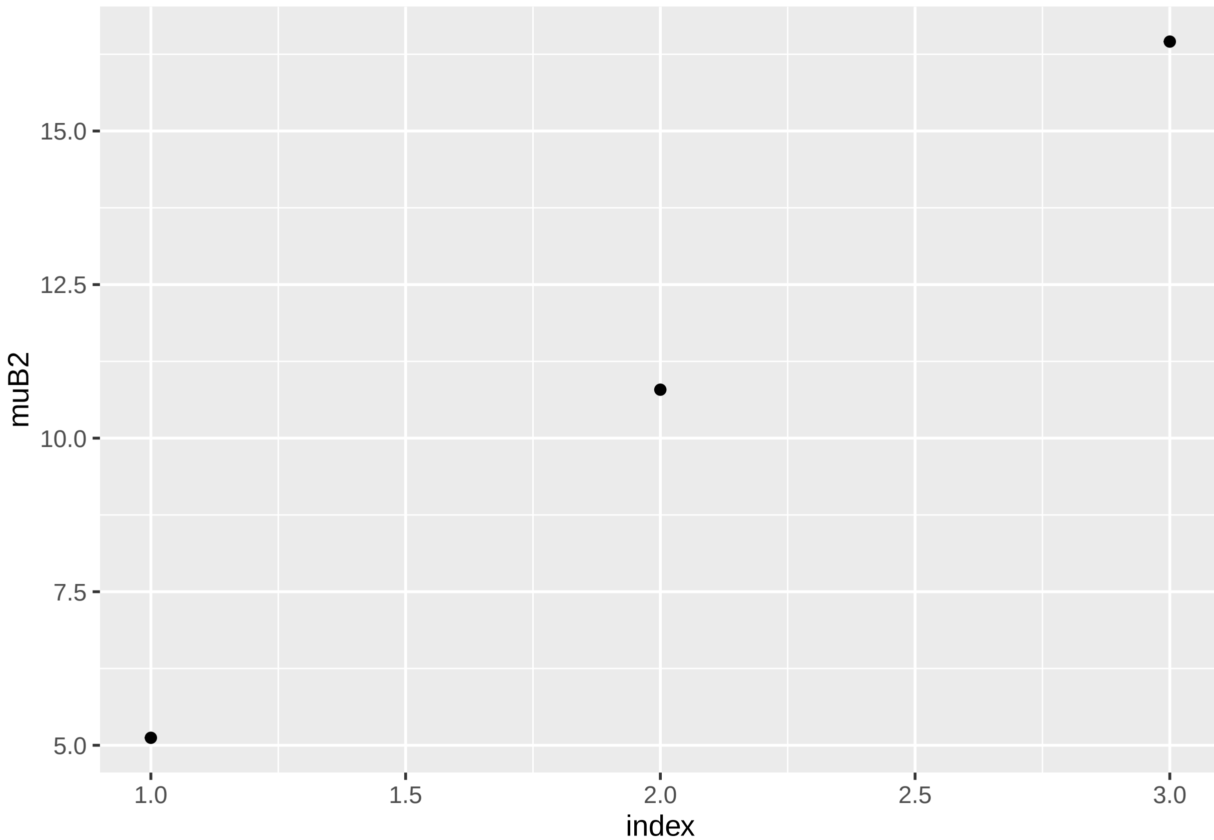
```
lm3.b<-lm(strength~A+as.factor(B),data=data3)
lm3.b
```

```
##
## Call:
## lm(formula = strength ~ A + as.factor(B), data = data3)
##
## Coefficients:
##   (Intercept)              A  as.factor(B)2
##        12.044          5.667        -12.589
```

```
data3.plot$muB1<-c(12.044+5.667,12.044+5.667*2,12.044+5.667*3)
data3.plot$muB2<-c(12.044+5.667,12.044+5.667*2,12.044+5.667*3)-12.589
ggplot(data3.plot,aes(x=index,y=muB1))+geom_point()
```




```
ggplot(data3.plot,aes(x=index,y=muB2))+geom_point()
```

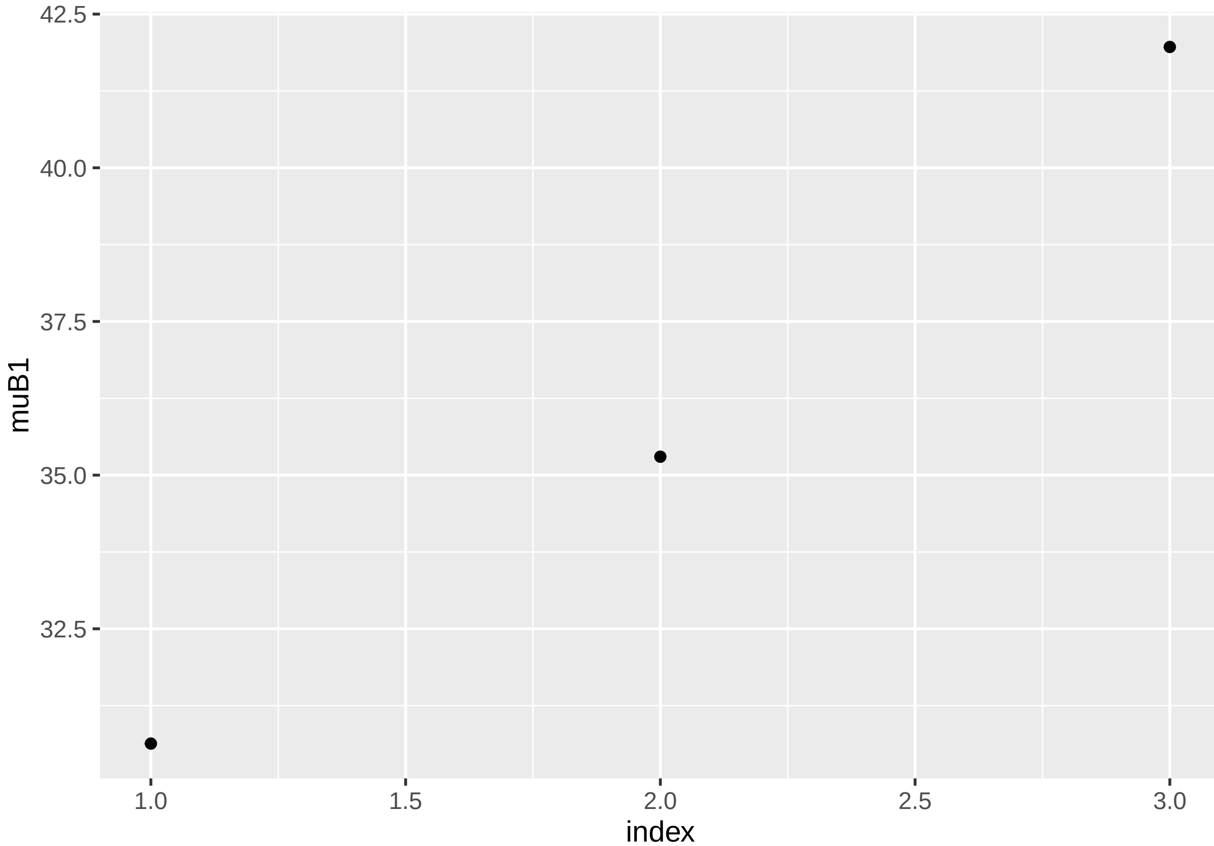


c)

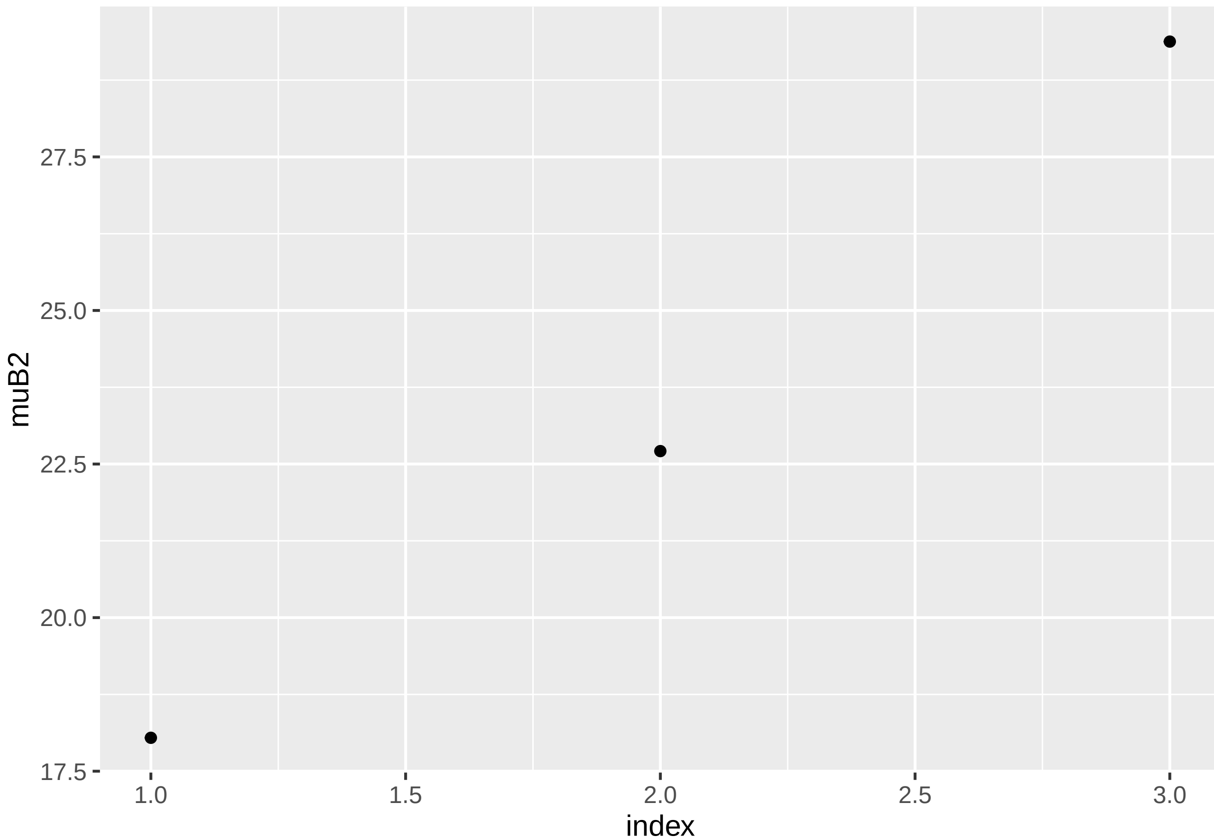
```
lm3.c<-lm(strength~as.factor(A)+B,data = data3)
lm3.c
```

```
##
## Call:
## lm(formula = strength ~ as.factor(A) + B, data = data3)
##
## Coefficients:
##   (Intercept)  as.factor(A)2  as.factor(A)3             B
##        30.633          4.667         11.333       -12.589
```

```
data3.plot$muB1<-c(30.633,30.633+4.667,30.633+11.333)
data3.plot$muB2<-c(30.633,30.633+4.667,30.633+11.333)-12.589
ggplot(data3.plot,aes(x=index,y=muB1))+geom_point()
```




```r
ggplot(data3.plot,aes(x=index,y=muB2))+geom_point()
```



8.6

```
data8.6<-read.table("CH08PR06.txt")
colnames(data8.6)<-c("Y","X")
```