Chapter 3 Part I Linear Regression (Problem 12)

kartheek raj mulasa

12/21/2019

Problem 12: This problem involves simple linear regression without an intercept.

(a) Recall that the coecient estimate ^ for the linear regression of Y onto X without an intercept is given by (3.38). Under what circumstance is the coecient estimate for the regression of X onto Y the same as the coecient estimate for the regression of Y onto X?

Answer

In generic linear equation y=mx+c,C=0 as intercept is zero then equation y=mx viceversa x onto y gives equation x=my both linear equations have the same slope occurs when y=x.

Given the linear regression equation in (3.38) we can say denmoniator term must be equal as numerator same for both events $\beta = \sum \{i\}x\{i\}y\{i\}$

$$\sum_{i`=1}^{n} x_{i`}^2 = \sum_{i`=1}^{n} y_{i`}^2$$

(b) Generate an example in R with n = 100 observations in which the coecient estimate for the regression of X onto Y is dierent from the coecient estimate for the regression of Y onto X.

Answer

creating the obersvations

```
set.seed(1)
x<-sample(1:1000,100,replace = T)
y<-x*2+sample(1:1000,100,replace = T)
length(x)
```

```
## [1] 100
```

```
length(y)
```

```
## [1] 100
```

Building linear models x onto y and y onto x.

```
slm<-lm(y~x+0)
slm1<-lm(x~y+0)
summary(slm)
```

```
##
## Call:
## lm(formula = y ~ x + 0)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -641.5 -122.1  138.6  299.6  707.6
##
## Coefficients:
```

```
##   Estimate Std. Error t value Pr(>|t|)
## x  2.74064    0.05449    50.3   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 330.8 on 99 degrees of freedom
## Multiple R-squared:  0.9623, Adjusted R-squared:  0.962
## F-statistic:  2530 on 1 and 99 DF,  p-value: < 2.2e-16
```

```
summary(slm1)
```

```
##
## Call:
## lm(formula = x ~ y + 0)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -243.14  -92.29  -23.92   73.22  260.33
##
## Coefficients:
##   Estimate Std. Error t value Pr(>|t|)
## y 0.351137   0.006981    50.3   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 118.4 on 99 degrees of freedom
## Multiple R-squared:  0.9623, Adjusted R-squared:  0.962
## F-statistic:  2530 on 1 and 99 DF,  p-value: < 2.2e-16
```

Generated linear models yields different coefficents

(c) Generate an example in R with n = 100 observations in which the coecient estimate for the regression of X onto Y is the same as the coecient estimate for the regression of Y onto X.

Answer

Generating the 100 obervations

```
x<-100:200
x<-x^2
y<-200:100
y<-y^2
```

Building linear models x onto y and y onto x.

```
slm2<-lm(x~y+0)
slm3<-lm(y~x+0)
summary(slm2)
```

```
##
## Call:
## lm(formula = x ~ y + 0)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -20166  -7470   5532  18842  32459
##
## Coefficients:
```

```
##   Estimate Std. Error t value Pr(>|t|)
## y  0.75414    0.06567   11.48   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 16460 on 100 degrees of freedom
## Multiple R-squared:  0.5687, Adjusted R-squared:  0.5644
## F-statistic: 131.9 on 1 and 100 DF,  p-value: < 2.2e-16
```

```
summary(slm3)
```

```
##
## Call:
## lm(formula = y ~ x + 0)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -20166  -7470   5532  18842  32459
##
## Coefficients:
##   Estimate Std. Error t value Pr(>|t|)
## x  0.75414    0.06567   11.48   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 16460 on 100 degrees of freedom
## Multiple R-squared:  0.5687, Adjusted R-squared:  0.5644
## F-statistic: 131.9 on 1 and 100 DF,  p-value: < 2.2e-16
```

Both models yield same coefficent.