HW2 - GLM

Thomas Kwok

March 19, 2019

1. Use 6.5 brands data to compute the following.
2. Regress Y on Xl using simple linear regression model (2.1) and obtain the residuals.

setwd("C:/Users/thoma/Desktop/Spring 2019/Linear Model 2")  
brands <- read.csv("CH7Q30.csv")  
lm1 <- lm(Yi1~Xi2, data=brands)  
resid(lm1)

## 1 2 3 4 5 6 7 8 9   
## -13.375 -13.125 -16.375 -10.125 -5.375 -6.125 -6.375 -3.125 5.625   
## 10 11 12 13 14 15 16   
## 2.875 8.625 6.875 10.625 8.875 16.625 13.875

1. Regress X1 on X2 using simple linear regression model (2.1) and obtain the residuals.

lm2 <- lm(Xi1~Xi2, data=brands)  
resid(lm2)

## 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16   
## -3 -3 -3 -3 -1 -1 -1 -1 1 1 1 1 3 3 3 3

1. Calculate the coefficient of simple correlation between the two sets of residuals and show that it equals RY1|2

lm3 <- lm(Yi1~Xi1+Xi2, data=brands)  
ssr1 <- sum((fitted(lm3) - mean(brands$Yi1))^2)  
ssr2 <- sum((fitted(lm1) - mean(brands$Yi1))^2)  
sse1 <- sum((fitted(lm3) - brands$Yi1)^2)  
sse2 <- sum((fitted(lm1) - brands$Yi1)^2)  
  
cor(lm1$residuals, lm2$residuals)

## [1] 0.9711943

sqrt((ssr1-ssr2)/(sse2))

## [1] 0.9711943

1. Use Table 7.6 to solve the following problems
2. calculate xtx, xty, b, s{b} for the following problems.

library(matlib)

## Warning: package 'matlib' was built under R version 3.5.3

x1 <- c(4,4,4,4,6,6,6,6)  
x2 <- c(2,2,3,3,2,2,3,3)  
y1 <- c(42,39,48,51,49,53,61,60)  
  
x1m <- mean(x1)  
x1std <- sd(x1)  
x1s <- matrix(0, nrow = 8, ncol = 1)  
for (i in 1:8)  
{x1s[i,1]<-(1/sqrt(7)) \* ((x1[i] -x1m)/x1std)}  
  
  
x2m <- mean(x2)  
x2std <- sd(x2)  
x2s <- matrix(0, nrow = 8, ncol = 1)  
for (i in 1:8)  
{x2s[i,1]<-(1/sqrt(7)) \* ((x2[i] -x2m)/x2std)}  
  
y1m <- mean(y1)  
y1std <- sd(y1)  
y1s <- matrix(0, nrow = 8, ncol = 1)  
for (i in 1:8)  
{y1s[i,1]<-(1/sqrt(7)) \* ((y1[i] -y1m)/y1std)}  
  
x <- matrix(c(x1s, x2s), nrow = 8, ncol = 2)  
y <- matrix(c(y1s), nrow = 8, ncol = 1)  
  
"XTX Matrix"

## [1] "XTX Matrix"

xtx <- t(x) %\*% x  
xtx

## [,1] [,2]  
## [1,] 1 0  
## [2,] 0 1

"XTY Matrix"

## [1] "XTY Matrix"

xty <- t(x) %\*% y  
xty

## [,1]  
## [1,] 0.7419309  
## [2,] 0.6384057

"b"

## [1] "b"

b<- inv(xtx) %\*% xty  
b

## [,1]  
## [1,] 0.7419309  
## [2,] 0.6384057

"s{b}"

## [1] "s{b}"

sse <- (t(y) %\*% y) - (t(b) %\*% xty)  
mse <- sse/7  
s2b <- mse %x% xtx  
s2b

## [,1] [,2]  
## [1,] 0.005996683 0.000000000  
## [2,] 0.000000000 0.005996683

1. Show that the standardized regression coefficients obtained in part (a3) are related to the regression coefficients for the regression model in the original variables according to (7.53)

The original variables in 7.53 take the standardized b found in a3 and then multiply it by the the standard deviation of y divided by the standard deviation of x. The method is replicated below.

(y1std/x1std)%\*%(b[1,1])

## [,1]  
## [1,] 5.375

(y1std/x2std)%\*%(b[2,1])

## [,1]  
## [1,] 9.25