# NTHU STAT 5410 - Linear Models Assignment 5 Report

105061110 周柏宇

1.

(a) Since , where and . The weighted least squares estimator is and its variance is . The reason why we should use WLS rather than OLS is that the variance of violets the assumption of equal variance from the Gauss-Markov theorem, which guarantees the variance of will be the smallest among all the linear unbiased estimators. However, we can fix the issue by fitting a transformed version of the original model , where . Then for the transformed model is exactly the for the original model.

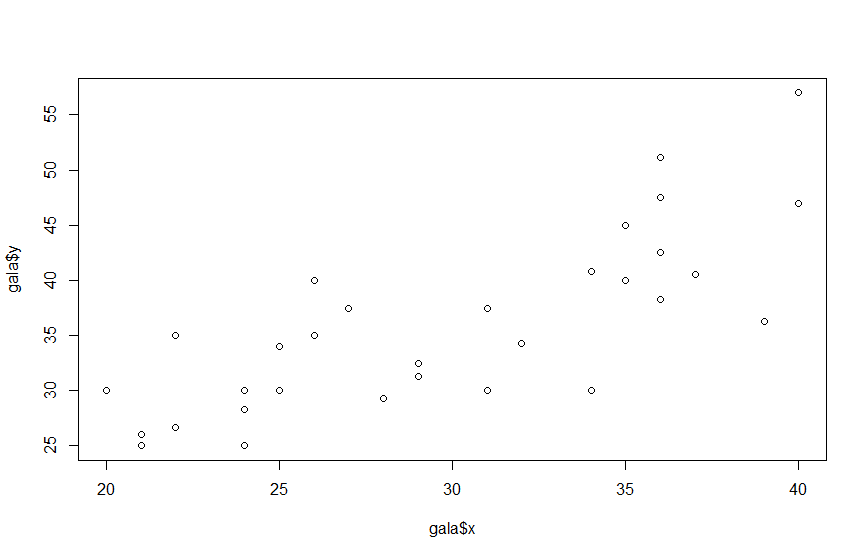
(b) In this case, , where and . The best linear unbiased estimator (BLUE) of β is and its variance is .

2.

(a)

> gala <- read.table("C:/Users/Thomas/Downloads/Linear\_models/hw5/E6.10.txt", header=T)

> plot(gala$x, gala$y)



Despite that the responses are averaged, we still can get a feeling of unequal variance among the observations due to their duplicates. If we assume the raw responses (i.e. ) are of constant variance . Then the covariance matrix for the sample mean will be

, where and . Therefore, .

> n <- gala[,2]

> x <- gala[,3]

> y <- gala[,4]

> fit1 <- lm(y ~ x, weights=n)

(b)

> fit2 <- lm(y ~ x)

> summary(fit2)

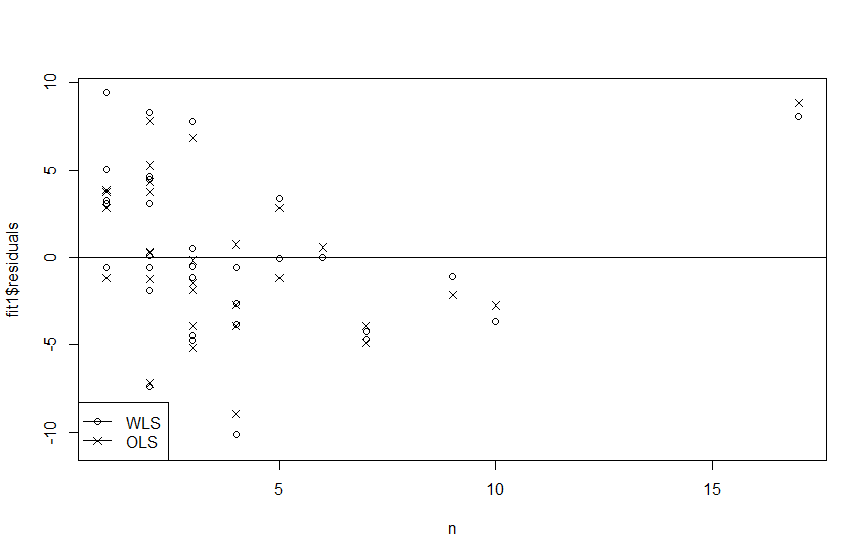
> plot(n, fit1$residuals, pch=1)

> points(n, fit2$residuals, pch=4)

> legend("bottomleft", legend = c("WLS", "OLS"),

pch = c(1, 4), lty = c(1, 1))

> abline(a=0, b=0)



(c)

> length(unique(x))

[1] 17

> length(x)

[1] 32

We can see that with 32 observations there are only 17 unique x. As the plot shows, we have several duplicates on some x.

3.