YK_Final_P3

Yinan Kang

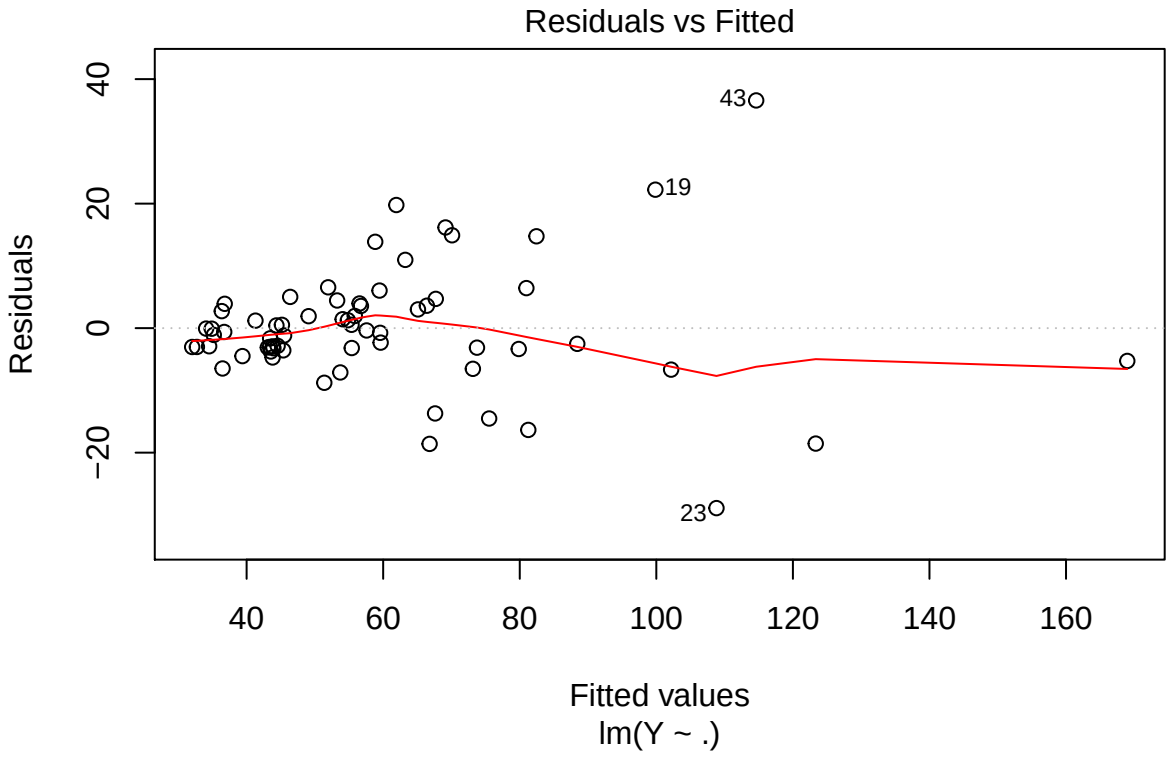
5/13/2019

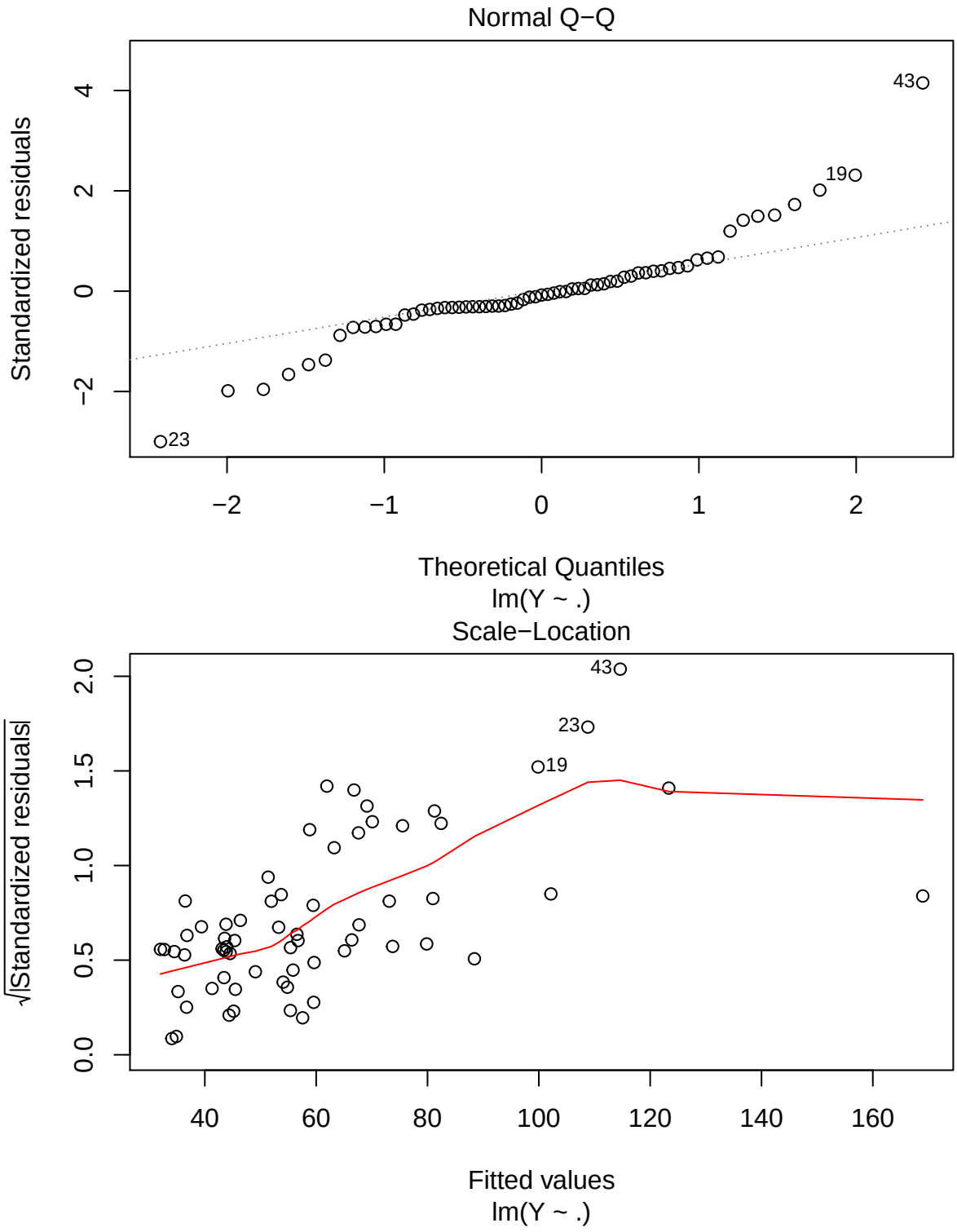
```
require(dplyr)
require(leaps)
require(neuralnet)
require(readxl)
require(caret)
require(lmtest)
```

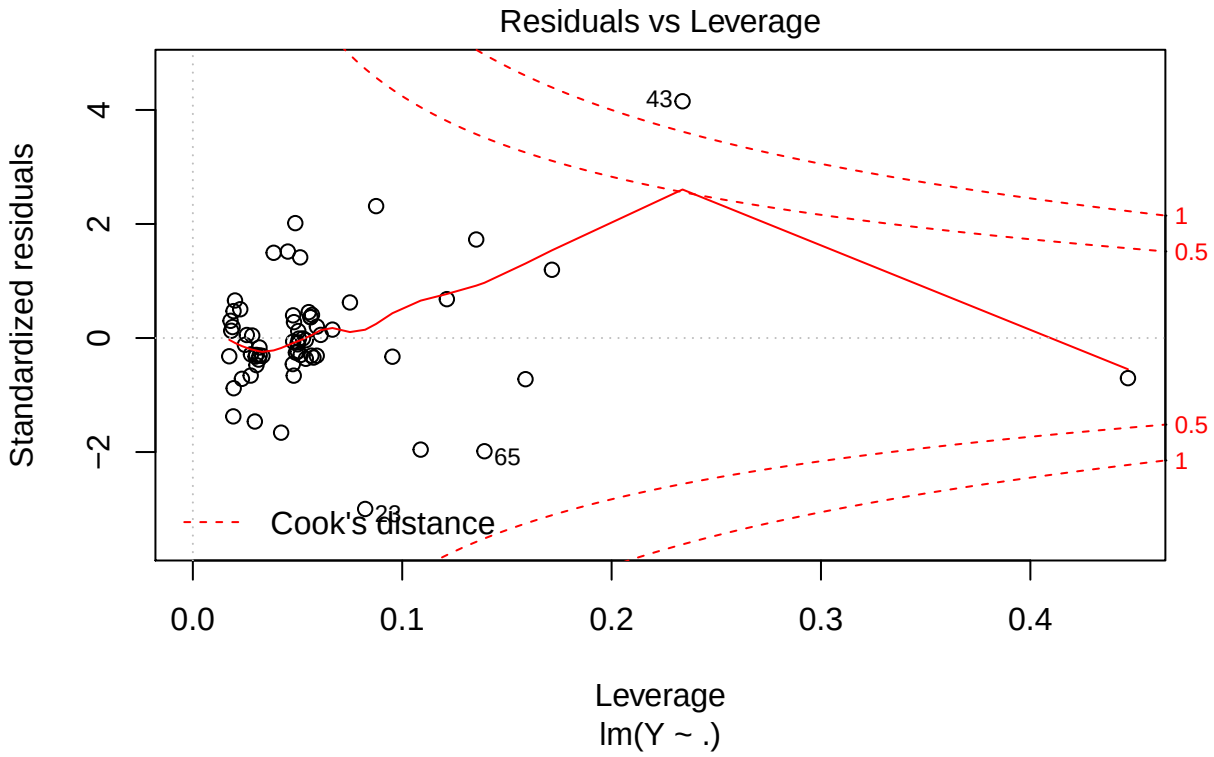
```
rm(list=ls())
df.3 <- read.csv("/cloud/project/Question 3.csv")
```

Fit Model

```
model.3 <- lm(Y~., data=df.3)
plot(model.3)
```




Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(Y ~ .)
43
19
23
−2
−1
0
1
2
4
Scale−Location
√|Standardized residuals|
Fitted values
lm(Y ~ .)
0.0
0.5
1.0
1.5
2.0
40
60
80
100
120
140
160



Analysis: We see unequal variance (Megaphone shape) in the Residuals Plot

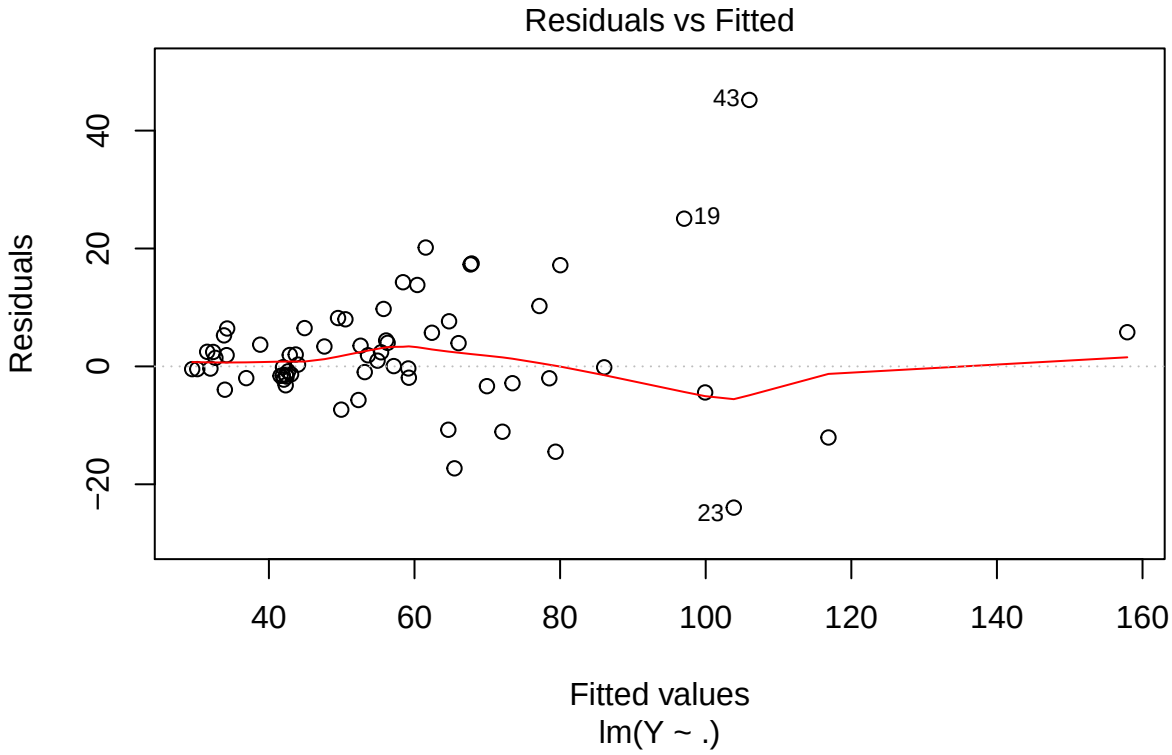
Using Weighted Least Squares

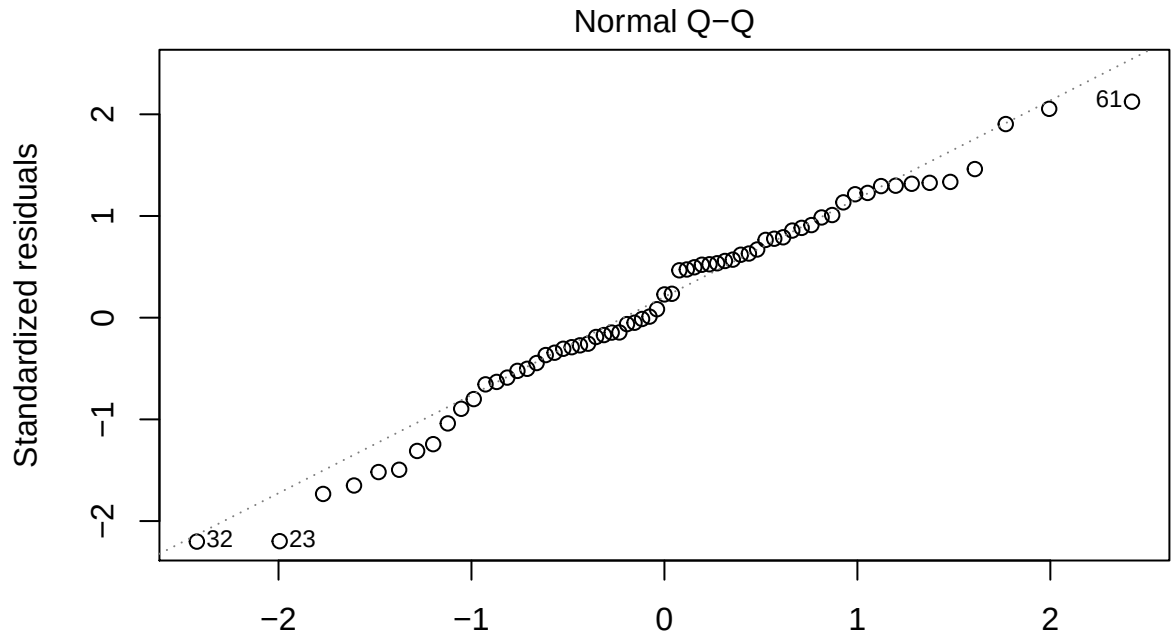
```
ei <- model.3$residuals
abs.ei <- abs(ei)
model.3.1 <- lm(abs.ei ~ ., data=df.3)

s <- model.3.1$fitted.values
wi <- 1/(s^2)

model.3.new <- lm(Y~., weights=wi, data=df.3)
summary(model.3.new)
```

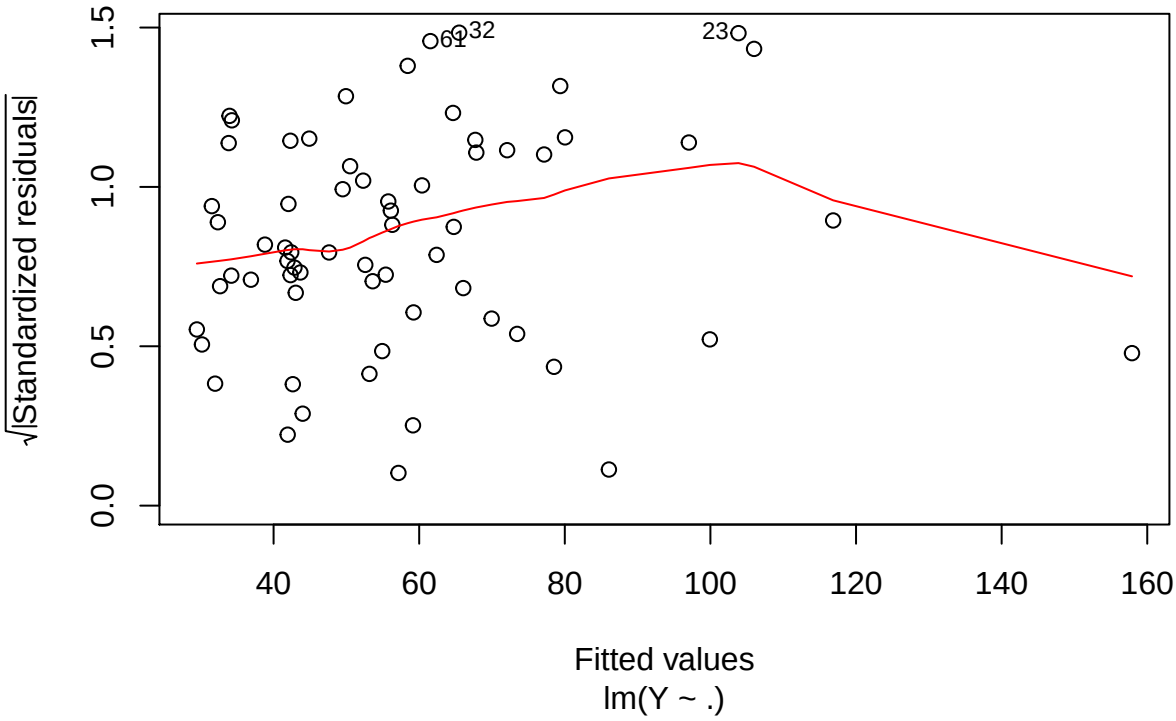
```
##
## Call:
## lm(formula = Y ~ ., data = df.3, weights = wi)
##
## Weighted Residuals:
##     Min      1Q  Median      3Q     Max
## -2.3041 -0.4933  0.2369  0.9425  2.3723
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  16.2356     1.5265  10.636 1.60e-15 ***
## X1           12.4000     0.7930  15.636  < 2e-16 ***
## X2            1.2718     0.1861   6.833 4.49e-09 ***
## X3            1.5929     0.3179   5.011 4.93e-06 ***
```

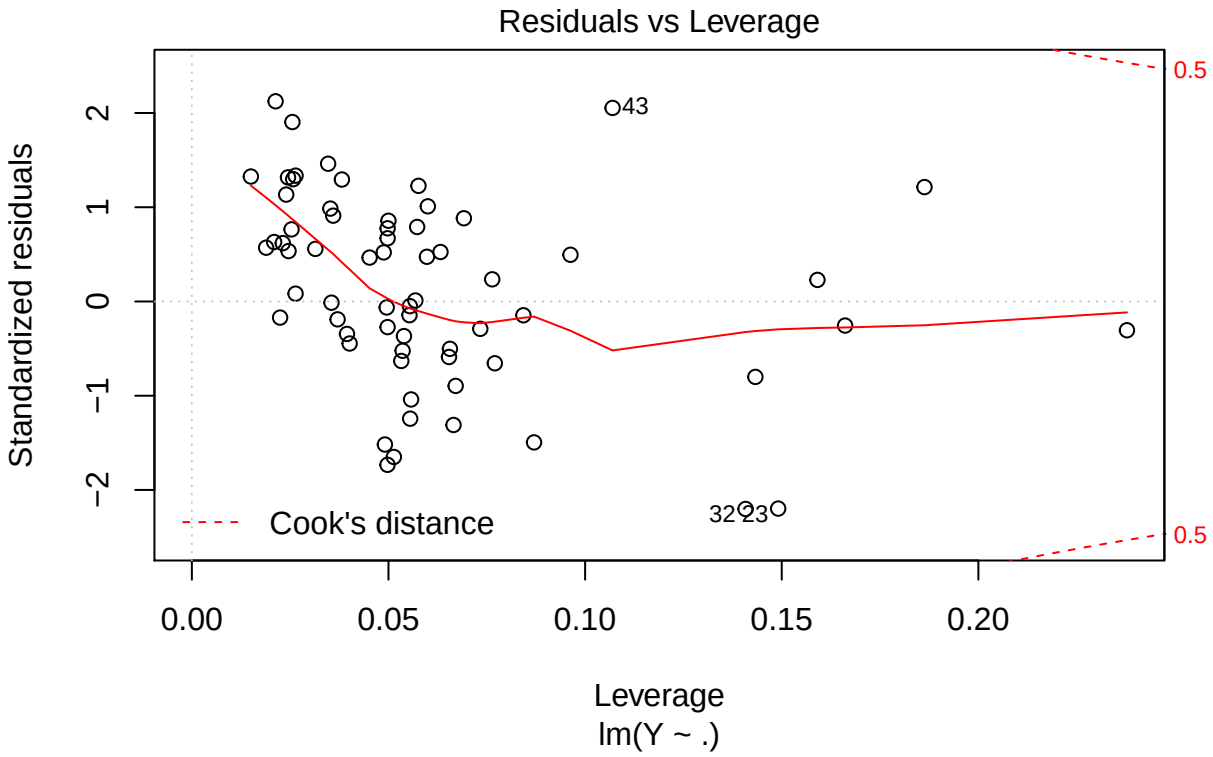
```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.129 on 61 degrees of freedom
## Multiple R-squared:  0.8878, Adjusted R-squared:  0.8823
## F-statistic: 160.9 on 3 and 61 DF,  p-value: < 2.2e-16
```

```
plot(model.3.new)
```




Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(Y ~ .)
61
32
23
−2
−1
0
1
2


Scale−Location
√|Standardized residuals|
Fitted values
lm(Y ~ .)
61
32
23
0.0
0.5
1.0
1.5
40
60
80
100
120
140
160


Residuals vs Leverage
Standardized residuals
Leverage
lm(Y ~ .)
Cook's distance
0.5
0.5
43
32
23
0.00
0.05
0.10
0.15
0.20
−2
−1
0
1
2