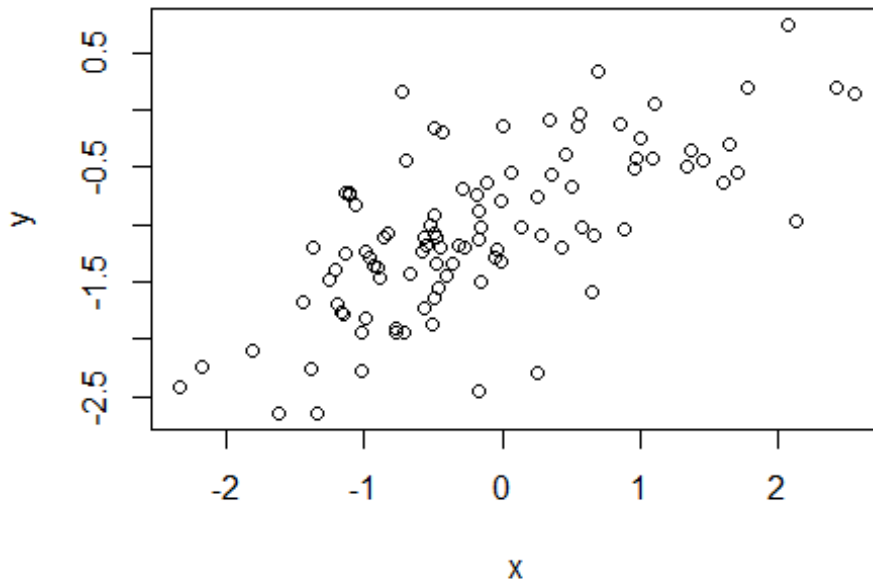# Homework 1

林双全

2017 年 3 月 4 日

13

```
set.seed(1234)
#(a)
x <- rnorm(100)
#(b)
e <- rnorm(100,0,0.5)
#(c)
y <- -1+0.5*x+e
length(y)

## [1] 100
```

In this linear model,β_0=-1,β_1=0.5.

```
#(d)
plot(x,y)
```

This scatterplot showing a positve correlation between x and y.

```
#(e)
lm.model <- lm(y~x)
summary(lm.model)

##
## Call:
## lm(formula = y ~ x)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.44313 -0.30700  0.00118  0.29322  1.49387
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.98142    0.05249 -18.698  < 2e-16 ***
## x            0.48696    0.05189   9.385 2.62e-15 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5186 on 98 degrees of freedom
## Multiple R-squared:  0.4733, Adjusted R-squared:  0.468
## F-statistic: 88.07 on 1 and 98 DF,  p-value: 2.62e-15

lm.model$coefficients

## (Intercept)           x
##  -0.9814230   0.4869575
```
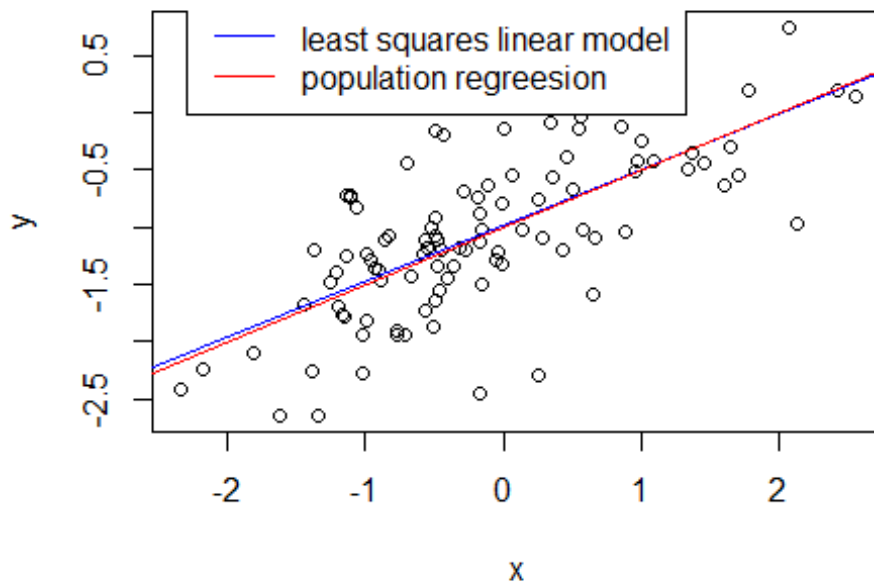
We can see that $\hat{\beta}_o$ and $\hat{\beta}_1$ are close to β_0 and β_1.

```
#(f)
plot(x,y)
abline(lm.model,col="blue")
abline(a=-1,b=0.5,col="red")
legend(-2.3,1,legend=c("least squares linear model","population regrees
ion"),col = c("blue","red"),lwd=1)
```

```
#(g)
polymodel <- lm(y~poly(x,2))
#polymodel <- lm(y~x+I(x^2))
summary(polymodel)

##
## Call:
## lm(formula = y ~ poly(x, 2))
##
## Residuals:
##      Min       1Q    Median       3Q      Max
## -1.49465 -0.32906 -0.00164  0.28410  1.47757
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -1.05776    0.05174 -20.444  < 2e-16 ***
## poly(x, 2)1  4.86651    0.51739   9.406 2.57e-15 ***
## poly(x, 2)2 -0.62128    0.51739  -1.201    0.233
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5174 on 97 degrees of freedom
## Multiple R-squared:  0.481,  Adjusted R-squared:  0.4703
## F-statistic: 44.96 on 2 and 97 DF,  p-value: 1.527e-14
```

We can see that the adjusted R-squared of polymodel very close to the lm.model's, and the x^2 is insignificant because its p-value is bigger than 0.05, thus the quadratic can not improve the model fit.

```r
#(h)
set.seed(2234)
e2 <- rnorm(100,0,0.15)
y2 <- -1+0.5*x+e2
lm.model2 <- lm(y2~x)
summary(lm.model2)

##
## Call:
## lm(formula = y2 ~ x)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.30248 -0.08292 -0.00913  0.08444  0.42437
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.98989    0.01366  -72.44   <2e-16 ***
## x            0.50779    0.01351   37.59   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.135 on 98 degrees of freedom
## Multiple R-squared:  0.9351, Adjusted R-squared:  0.9345
## F-statistic:  1413 on 1 and 98 DF,  p-value: < 2.2e-16

plot(x,y2)
abline(lm.model2,col="blue")
abline(a=-1,b=0.5,col="red")
legend(-2.2,0.5,legend=c("least squares linear model","population regre
esion"),col = c("blue","red"),lwd=1)
```
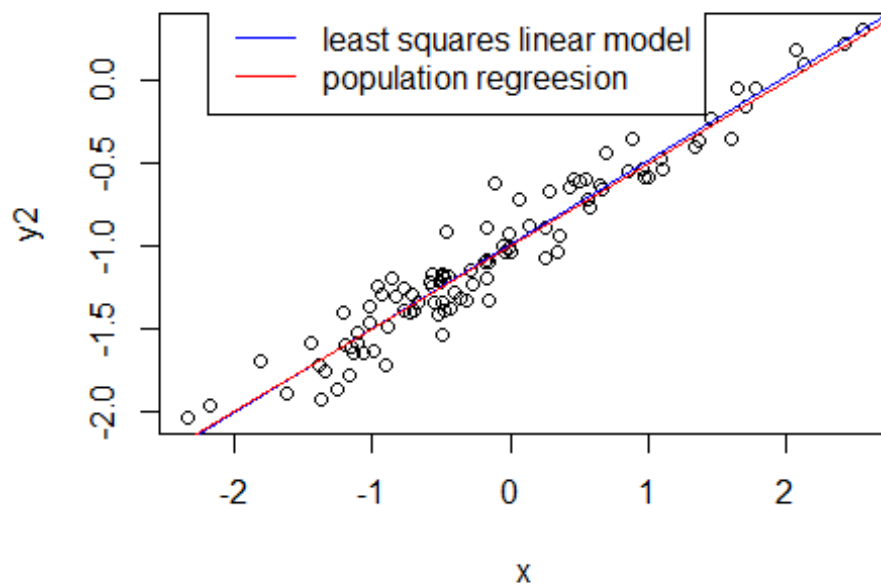
The adjustment R-squared of new model is far bigger than the previous model,and RSE is far less.Thus the new model has better fitting effect.

```
#(i)
set.seed(3234)
e3 <- rnorm(100,0,0.75)
y3 <- -1+0.5*x+e3
lm.model3 <- lm(y3~x)
summary(lm.model3)

##
## Call:
## lm(formula = y3 ~ x)
##
## Residuals:
##      Min      1Q   Median      3Q      Max
## -1.36199 -0.62823 -0.06472  0.65814  1.48505
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -1.05566    0.07441 -14.187  < 2e-16 ***
## x            0.47532    0.07356   6.462 4.03e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7351 on 98 degrees of freedom
```

```
## Multiple R-squared:  0.2988, Adjusted R-squared:  0.2916
## F-statistic: 41.76 on 1 and 98 DF,  p-value: 4.026e-09

plot(x,y3)
abline(lm.model3,col="blue")
abline(a=-1,b=0.5,col="red")
legend(-2.4,1.5,legend=c("least squares linear model","population regre
esion"),col = c("blue","red"),lwd=1)
```
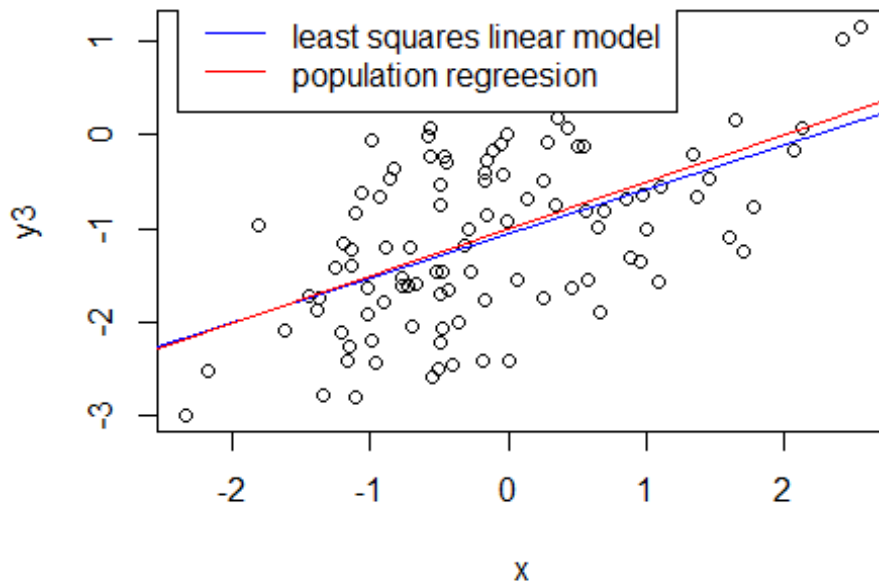

The adjustment R-squared of new model is smaller than the previous model,and RSE is bigger.Thus the new model has worse fitting effect.

```
#(j)
confint(lm.model)

##                   2.5 %      97.5 %
## (Intercept) -1.0855864 -0.8772595
## x            0.3839876  0.5899275

confint(lm.model2)

##                   2.5 %      97.5 %
## (Intercept) -1.0170116 -0.9627763
## x            0.4809878  0.5346017

confint(lm.model3)
```

```
##                2.5 %      97.5 %
## (Intercept) -1.203328 -0.9079986
## x            0.329347  0.6212923
```

Thus the data more noisy, the confidence intervel of coefficient more bigger.

15

```r
library(MASS)
summary(Boston)
```

```
##       crim                zn              indus             chas

##  Min.   : 0.00632   Min.   :  0.00   Min.   : 0.46   Min.    :0.00000

##  1st Qu.: 0.08204   1st Qu.:  0.00   1st Qu.: 5.19   1st Qu.:0.00000

##  Median : 0.25651   Median :  0.00   Median : 9.69   Median :0.00000

##  Mean   : 3.61352   Mean   : 11.36   Mean   :11.14   Mean    :0.06917

##  3rd Qu.: 3.67708   3rd Qu.: 12.50   3rd Qu.:18.10   3rd Qu.:0.00000

##  Max.   :88.97620   Max.   :100.00   Max.   :27.74   Max.    :1.00000

##       nox               rm              age              dis
##  Min.   :0.3850   Min.   :3.561   Min.   :  2.90   Min.   : 1.130
##  1st Qu.:0.4490   1st Qu.:5.886   1st Qu.: 45.02   1st Qu.: 2.100
##  Median :0.5380   Median :6.208   Median : 77.50   Median : 3.207
##  Mean   :0.5547   Mean   :6.285   Mean   : 68.57   Mean   : 3.795
##  3rd Qu.:0.6240   3rd Qu.:6.623   3rd Qu.: 94.08   3rd Qu.: 5.188
##  Max.   :0.8710   Max.   :8.780   Max.   :100.00   Max.   :12.127
##       rad              tax            ptratio           black
##  Min.   : 1.000   Min.   :187.0   Min.   :12.60   Min.   :  0.32
##  1st Qu.: 4.000   1st Qu.:279.0   1st Qu.:17.40   1st Qu.:375.38
##  Median : 5.000   Median :330.0   Median :19.05   Median :391.44
##  Mean   : 9.549   Mean   :408.2   Mean   :18.46   Mean   :356.67
##  3rd Qu.:24.000   3rd Qu.:666.0   3rd Qu.:20.20   3rd Qu.:396.23
##  Max.   :24.000   Max.   :711.0   Max.   :22.00   Max.   :396.90
##      lstat            medv
##  Min.   : 1.73   Min.   : 5.00
##  1st Qu.: 6.95   1st Qu.:17.02
##  Median :11.36   Median :21.20
##  Mean   :12.65   Mean   :22.53
##  3rd Qu.:16.95   3rd Qu.:25.00
##  Max.   :37.97   Max.   :50.00
```

```r
attach(Boston)
chas <- as.factor(chas)
```

```
#(a)
lm.zn <- lm(crim~zn)
summary(lm.zn)

##
## Call:
## lm(formula = crim ~ zn)
##
## Residuals:
##    Min    1Q Median    3Q    Max
## -4.429 -4.222 -2.620  1.250 84.523
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  4.45369    0.41722  10.675  < 2e-16 ***
## zn          -0.07393    0.01609  -4.594 5.51e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.435 on 504 degrees of freedom
## Multiple R-squared:  0.04019,    Adjusted R-squared:  0.03828
## F-statistic:   21.1 on 1 and 504 DF,  p-value: 5.506e-06

lm.indus <- lm(crim~indus)
summary(lm.indus)

##
## Call:
## lm(formula = crim ~ indus)
##
## Residuals:
##     Min     1Q  Median     3Q    Max
## -11.972  -2.698  -0.736   0.712  81.813
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -2.06374    0.66723  -3.093  0.00209 **
## indus        0.50978    0.05102   9.991  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.866 on 504 degrees of freedom
## Multiple R-squared:  0.1653, Adjusted R-squared:  0.1637
## F-statistic: 99.82 on 1 and 504 DF,  p-value: < 2.2e-16

lm.chas <- lm(crim~chas)
summary(lm.chas)

##
## Call:
## lm(formula = crim ~ chas)
```

```
## 
## Residuals:
##     Min     1Q Median     3Q    Max
## -3.738 -3.661 -3.435  0.018 85.232
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.7444     0.3961   9.453   <2e-16 ***
## chas1        -1.8928     1.5061  -1.257    0.209
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 8.597 on 504 degrees of freedom
## Multiple R-squared:  0.003124,   Adjusted R-squared:  0.001146
## F-statistic: 1.579 on 1 and 504 DF,  p-value: 0.2094

lm.nox <- lm(crim~nox)
summary(lm.nox)

## 
## Call:
## lm(formula = crim ~ nox)
## 
## Residuals:
##     Min      1Q  Median      3Q     Max
## -12.371  -2.738  -0.974   0.559  81.728
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -13.720      1.699  -8.073 5.08e-15 ***
## nox           31.249      2.999  10.419  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 7.81 on 504 degrees of freedom
## Multiple R-squared:  0.1772, Adjusted R-squared:  0.1756
## F-statistic: 108.6 on 1 and 504 DF,  p-value: < 2.2e-16

lm.rm <- lm(crim~rm)
summary(lm.rm)

## 
## Call:
## lm(formula = crim ~ rm)
## 
## Residuals:
##    Min     1Q Median     3Q    Max
## -6.604 -3.952 -2.654  0.989 87.197
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept)    20.482        3.365    6.088 2.27e-09 ***
## rm             -2.684        0.532   -5.045 6.35e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.401 on 504 degrees of freedom
## Multiple R-squared:  0.04807,    Adjusted R-squared:  0.04618
## F-statistic: 25.45 on 1 and 504 DF,  p-value: 6.347e-07
```

```
lm.age <- lm(crim~age)
summary(lm.age)
```

```
##
## Call:
## lm(formula = crim ~ age)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -6.789 -4.257 -1.230  1.527 82.849
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -3.77791    0.94398  -4.002 7.22e-05 ***
## age          0.10779    0.01274   8.463 2.85e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.057 on 504 degrees of freedom
## Multiple R-squared:  0.1244, Adjusted R-squared:  0.1227
## F-statistic: 71.62 on 1 and 504 DF,  p-value: 2.855e-16
```

```
lm.dis <- lm(crim~dis)
summary(lm.dis)
```

```
##
## Call:
## lm(formula = crim ~ dis)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -6.708 -4.134 -1.527  1.516 81.674
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  9.4993     0.7304  13.006   <2e-16 ***
## dis         -1.5509     0.1683  -9.213   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.965 on 504 degrees of freedom
```

```
## Multiple R-squared:  0.1441, Adjusted R-squared:  0.1425
## F-statistic: 84.89 on 1 and 504 DF,  p-value: < 2.2e-16

lm.rad <- lm(crim~rad)
summary(lm.rad)

##
## Call:
## lm(formula = crim ~ rad)
##
## Residuals:
##     Min     1Q  Median     3Q     Max
## -10.164  -1.381  -0.141   0.660  76.433
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -2.28716    0.44348  -5.157 3.61e-07 ***
## rad          0.61791    0.03433  17.998  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.718 on 504 degrees of freedom
## Multiple R-squared:  0.3913, Adjusted R-squared:   0.39
## F-statistic: 323.9 on 1 and 504 DF,  p-value: < 2.2e-16

lm.tax <- lm(crim~tax)
summary(lm.tax)

##
## Call:
## lm(formula = crim ~ tax)
##
## Residuals:
##     Min     1Q  Median     3Q     Max
## -12.513  -2.738  -0.194   1.065  77.696
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -8.528369   0.815809  -10.45   <2e-16 ***
## tax          0.029742   0.001847   16.10   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.997 on 504 degrees of freedom
## Multiple R-squared:  0.3396, Adjusted R-squared:  0.3383
## F-statistic: 259.2 on 1 and 504 DF,  p-value: < 2.2e-16

lm.ptratio <- lm(crim~ptratio)
summary(lm.ptratio)
```

```
## 
## Call:
## lm(formula = crim ~ ptratio)
## 
## Residuals:
##    Min     1Q Median     3Q    Max
## -7.654 -3.985 -1.912  1.825 83.353
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -17.6469     3.1473  -5.607 3.40e-08 ***
## ptratio       1.1520     0.1694   6.801 2.94e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 8.24 on 504 degrees of freedom
## Multiple R-squared:  0.08407,    Adjusted R-squared:  0.08225
## F-statistic: 46.26 on 1 and 504 DF,  p-value: 2.943e-11

lm.black <- lm(crim~black)
summary(lm.black)

## 
## Call:
## lm(formula = crim ~ black)
## 
## Residuals:
##     Min     1Q  Median     3Q    Max
## -13.756 -2.299  -2.095 -1.296 86.822
## 
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 16.553529   1.425903  11.609   <2e-16 ***
## black       -0.036280   0.003873  -9.367   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 7.946 on 504 degrees of freedom
## Multiple R-squared:  0.1483, Adjusted R-squared:  0.1466
## F-statistic: 87.74 on 1 and 504 DF,  p-value: < 2.2e-16

lm.lstat <- lm(crim~lstat)
summary(lm.lstat)

## 
## Call:
## lm(formula = crim ~ lstat)
## 
## Residuals:
##     Min     1Q  Median     3Q    Max
## -13.925 -2.822  -0.664  1.079 82.862
```
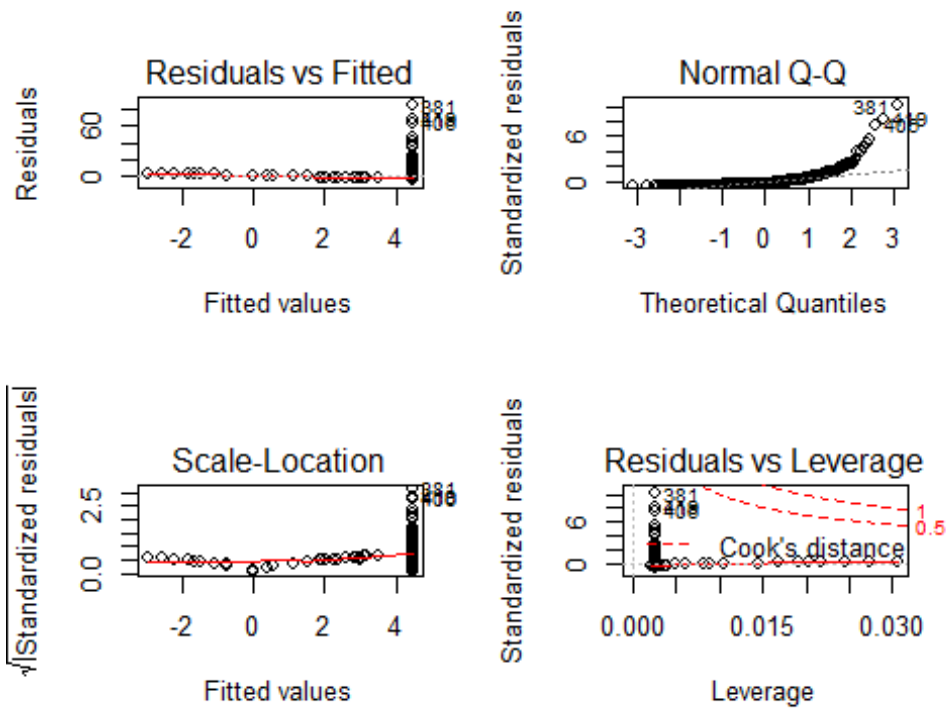
```
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -3.33054    0.69376  -4.801 2.09e-06 ***
## lstat        0.54880    0.04776  11.491  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 7.664 on 504 degrees of freedom
## Multiple R-squared:  0.2076, Adjusted R-squared:  0.206
## F-statistic:   132 on 1 and 504 DF,  p-value: < 2.2e-16

lm.medv <- lm(crim~medv)
summary(lm.medv)

## 
## Call:
## lm(formula = crim ~ medv)
## 
## Residuals:
##     Min     1Q Median     3Q    Max
## -9.071 -4.022 -2.343  1.298 80.957
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 11.79654    0.93419   12.63   <2e-16 ***
## medv        -0.36316    0.03839   -9.46   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 7.934 on 504 degrees of freedom
## Multiple R-squared:  0.1508, Adjusted R-squared:  0.1491
## F-statistic: 89.49 on 1 and 504 DF,  p-value: < 2.2e-16
```
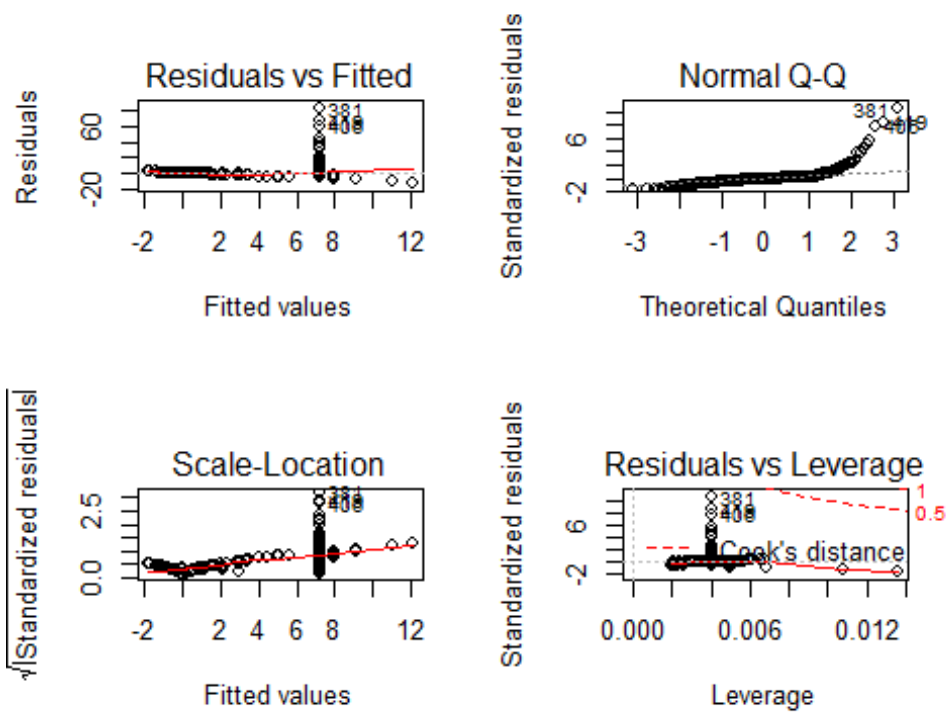
All variables except chas are significant.

```
par(mfrow=c(2,2))
plot(lm.zn)
```

Residuals vs Fitted

Normal Q-Q

Scale-Location

Residuals vs Leverage

```
par(mfrow=c(2,2))
plot(lm.indus)
```

Residuals vs Fitted

Normal Q-Q

Scale-Location

Residuals vs Leverage

Such that.

```
allmodel <- lm(crim~.,data=Boston)
summary(allmodel)

##
## Call:
## lm(formula = crim ~ ., data = Boston)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -9.924 -2.120 -0.353  1.019 75.051
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  17.033228   7.234903   2.354 0.018949 *
## zn            0.044855   0.018734   2.394 0.017025 *
## indus        -0.063855   0.083407  -0.766 0.444294
## chas         -0.749134   1.180147  -0.635 0.525867
## nox         -10.313535   5.275536  -1.955 0.051152 .
## rm            0.430131   0.612830   0.702 0.483089
## age           0.001452   0.017925   0.081 0.935488
## dis          -0.987176   0.281817  -3.503 0.000502 ***
## rad           0.588209   0.088049   6.680 6.46e-11 ***
## tax          -0.003780   0.005156  -0.733 0.463793
## ptratio      -0.271081   0.186450  -1.454 0.146611
## black        -0.007538   0.003673  -2.052 0.040702 *
## lstat         0.126211   0.075725   1.667 0.096208 .
## medv         -0.198887   0.060516  -3.287 0.001087 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.439 on 492 degrees of freedom
## Multiple R-squared:  0.454,  Adjusted R-squared:  0.4396
## F-statistic: 31.47 on 13 and 492 DF,  p-value: < 2.2e-16
```

Thereforce,we can find that these predictor are significant under $\alpha < 0.05$: zn,dis,rad,black,medv.

```
#(c)
x = c(coefficients(lm.zn)[2],coefficients(lm.indus)[2],coefficients(lm.
chas)[2],coefficients(lm.nox)[2],
      coefficients(lm.rm)[2],coefficients(lm.age)[2],coefficients(lm.di
s)[2],coefficients(lm.rad)[2],
      coefficients(lm.tax)[2],coefficients(lm.ptratio)[2],coefficients
(lm.black)[2],coefficients(lm.lstat)[2],
      coefficients(lm.medv)[2])
y = coefficients(allmodel)[2:14]
plot(x, y)
```
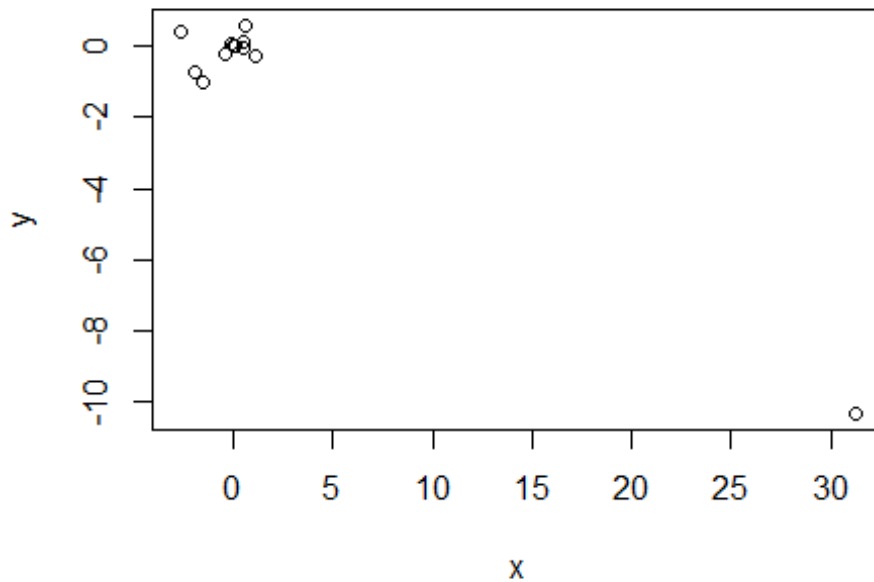
```
#(d)
#Because "chas" is dummy variable,thus remove it.

lmp.zn <- lm(crim~poly(zn,3))
summary(lmp.zn)

##
## Call:
## lm(formula = crim ~ poly(zn, 3))
##
## Residuals:
##    Min    1Q Median    3Q    Max
## -4.821 -4.614 -1.294  0.473 84.130
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.6135     0.3722   9.709  < 2e-16 ***
## poly(zn, 3)1 -38.7498     8.3722  -4.628  4.7e-06 ***
## poly(zn, 3)2  23.9398     8.3722   2.859  0.00442 **
## poly(zn, 3)3 -10.0719     8.3722  -1.203  0.22954
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.372 on 502 degrees of freedom
## Multiple R-squared:  0.05824,    Adjusted R-squared:  0.05261
## F-statistic: 10.35 on 3 and 502 DF,  p-value: 1.281e-06
```

```
lmp.indus <- lm(crim~poly(indus,3))
summary(lmp.indus)

##
## Call:
## lm(formula = crim ~ poly(indus, 3))
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -8.278 -2.514  0.054  0.764 79.713
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)        3.614      0.330  10.950  < 2e-16 ***
## poly(indus, 3)1   78.591      7.423  10.587  < 2e-16 ***
## poly(indus, 3)2  -24.395      7.423  -3.286  0.00109 **
## poly(indus, 3)3  -54.130      7.423  -7.292  1.2e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.423 on 502 degrees of freedom
## Multiple R-squared:  0.2597, Adjusted R-squared:  0.2552
## F-statistic: 58.69 on 3 and 502 DF,  p-value: < 2.2e-16

lmp.nox <- lm(crim~poly(nox,3))
summary(lmp.nox)

##
## Call:
## lm(formula = crim ~ poly(nox, 3))
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -9.110 -2.068 -0.255  0.739 78.302
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)      3.6135     0.3216  11.237  < 2e-16 ***
## poly(nox, 3)1   81.3720     7.2336  11.249  < 2e-16 ***
## poly(nox, 3)2  -28.8286     7.2336  -3.985 7.74e-05 ***
## poly(nox, 3)3  -60.3619     7.2336  -8.345 6.96e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.234 on 502 degrees of freedom
## Multiple R-squared:  0.297,  Adjusted R-squared:  0.2928
## F-statistic: 70.69 on 3 and 502 DF,  p-value: < 2.2e-16

lmp.rm <- lm(crim~poly(rm,3))
summary(lmp.rm)
```

```
## 
## Call:
## lm(formula = crim ~ poly(rm, 3))
## 
## Residuals:
##     Min      1Q  Median      3Q     Max
## -18.485  -3.468  -2.221  -0.015  87.219
## 
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.6135     0.3703   9.758  < 2e-16 ***
## poly(rm, 3)1 -42.3794     8.3297  -5.088 5.13e-07 ***
## poly(rm, 3)2  26.5768     8.3297   3.191  0.00151 **
## poly(rm, 3)3  -5.5103     8.3297  -0.662  0.50858
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 8.33 on 502 degrees of freedom
## Multiple R-squared:  0.06779,    Adjusted R-squared:  0.06222
## F-statistic: 12.17 on 3 and 502 DF,  p-value: 1.067e-07

lmp.age <- lm(crim~poly(age,3))
summary(lmp.age)

## 
## Call:
## lm(formula = crim ~ poly(age, 3))
## 
## Residuals:
##    Min     1Q Median     3Q    Max
## -9.762 -2.673 -0.516  0.019 82.842
## 
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)     3.6135     0.3485  10.368  < 2e-16 ***
## poly(age, 3)1  68.1820     7.8397   8.697  < 2e-16 ***
## poly(age, 3)2  37.4845     7.8397   4.781 2.29e-06 ***
## poly(age, 3)3  21.3532     7.8397   2.724  0.00668 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 7.84 on 502 degrees of freedom
## Multiple R-squared:  0.1742, Adjusted R-squared:  0.1693
## F-statistic: 35.31 on 3 and 502 DF,  p-value: < 2.2e-16

lmp.dis <- lm(crim~poly(dis,3))
summary(lmp.dis)

## 
## Call:
## lm(formula = crim ~ poly(dis, 3))
```

```
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -10.757  -2.588   0.031   1.267  76.378
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)      3.6135     0.3259  11.087  < 2e-16 ***
## poly(dis, 3)1 -73.3886     7.3315 -10.010  < 2e-16 ***
## poly(dis, 3)2  56.3730     7.3315   7.689 7.87e-14 ***
## poly(dis, 3)3 -42.6219     7.3315  -5.814 1.09e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.331 on 502 degrees of freedom
## Multiple R-squared:  0.2778, Adjusted R-squared:  0.2735
## F-statistic: 64.37 on 3 and 502 DF,  p-value: < 2.2e-16

lmp.rad <- lm(crim~poly(rad,3))
summary(lmp.rad)

##
## Call:
## lm(formula = crim ~ poly(rad, 3))
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -10.381  -0.412  -0.269   0.179  76.217
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)      3.6135     0.2971  12.164  < 2e-16 ***
## poly(rad, 3)1 120.9074     6.6824  18.093  < 2e-16 ***
## poly(rad, 3)2  17.4923     6.6824   2.618  0.00912 **
## poly(rad, 3)3   4.6985     6.6824   0.703  0.48231
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.682 on 502 degrees of freedom
## Multiple R-squared:    0.4, Adjusted R-squared:  0.3965
## F-statistic: 111.6 on 3 and 502 DF,  p-value: < 2.2e-16

lmp.tax <- lm(crim~poly(tax,3))
summary(lmp.tax)

##
## Call:
## lm(formula = crim ~ poly(tax, 3))
##
## Residuals:
##     Min      1Q  Median      3Q     Max
```

```
## -13.273   -1.389    0.046    0.536   76.950
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)      3.6135     0.3047  11.860  < 2e-16 ***
## poly(tax, 3)1 112.6458      6.8537  16.436  < 2e-16 ***
## poly(tax, 3)2  32.0873      6.8537   4.682 3.67e-06 ***
## poly(tax, 3)3  -7.9968      6.8537  -1.167    0.244
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.854 on 502 degrees of freedom
## Multiple R-squared:  0.3689, Adjusted R-squared:  0.3651
## F-statistic:  97.8 on 3 and 502 DF,  p-value: < 2.2e-16

lmp.ptratio <- lm(crim~poly(ptratio,3))
summary(lmp.ptratio)

##
## Call:
## lm(formula = crim ~ poly(ptratio, 3))
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -6.833 -4.146 -1.655  1.408 82.697
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)           3.614      0.361  10.008  < 2e-16 ***
## poly(ptratio, 3)1    56.045      8.122   6.901 1.57e-11 ***
## poly(ptratio, 3)2    24.775      8.122   3.050  0.00241 **
## poly(ptratio, 3)3   -22.280      8.122  -2.743  0.00630 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.122 on 502 degrees of freedom
## Multiple R-squared:  0.1138, Adjusted R-squared:  0.1085
## F-statistic: 21.48 on 3 and 502 DF,  p-value: 4.171e-13

lmp.black <- lm(crim~poly(black,3))
summary(lmp.black)

##
## Call:
## lm(formula = crim ~ poly(black, 3))
##
## Residuals:
##     Min     1Q  Median     3Q    Max
## -13.096  -2.343  -2.128  -1.439  86.790
##
## Coefficients:
```

```
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)       3.6135     0.3536  10.218   <2e-16 ***
## poly(black, 3)1 -74.4312     7.9546  -9.357   <2e-16 ***
## poly(black, 3)2   5.9264     7.9546   0.745    0.457
## poly(black, 3)3  -4.8346     7.9546  -0.608    0.544
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.955 on 502 degrees of freedom
## Multiple R-squared:  0.1498, Adjusted R-squared:  0.1448
## F-statistic: 29.49 on 3 and 502 DF,  p-value: < 2.2e-16

lmp.lstat <- lm(crim~poly(lstat,3))
summary(lmp.lstat)

##
## Call:
## lm(formula = crim ~ poly(lstat, 3))
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -15.234  -2.151  -0.486   0.066  83.353
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)       3.6135     0.3392  10.654   <2e-16 ***
## poly(lstat, 3)1  88.0697     7.6294  11.543   <2e-16 ***
## poly(lstat, 3)2  15.8882     7.6294   2.082   0.0378 *
## poly(lstat, 3)3 -11.5740     7.6294  -1.517   0.1299
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.629 on 502 degrees of freedom
## Multiple R-squared:  0.2179, Adjusted R-squared:  0.2133
## F-statistic: 46.63 on 3 and 502 DF,  p-value: < 2.2e-16

lmp.medv <- lm(crim~poly(medv,3))
summary(lmp.medv)

##
## Call:
## lm(formula = crim ~ poly(medv, 3))
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -24.427  -1.976  -0.437   0.439  73.655
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)       3.614      0.292  12.374  < 2e-16 ***
## poly(medv, 3)1  -75.058      6.569 -11.426  < 2e-16 ***
```

```
## poly(medv, 3)2    88.086        6.569  13.409  < 2e-16 ***
## poly(medv, 3)3   -48.033        6.569  -7.312 1.05e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.569 on 502 degrees of freedom
## Multiple R-squared:  0.4202, Adjusted R-squared:  0.4167
## F-statistic: 121.3 on 3 and 502 DF,  p-value: < 2.2e-16
```

Thereforce, except "chas" and "black", other variables are non-linear.