570 Assignment4

Yifu He 10442277

###Q1:

Step1: after the linear regression (x and y represent xom and cvx), we can get the parameter of sigma and c.

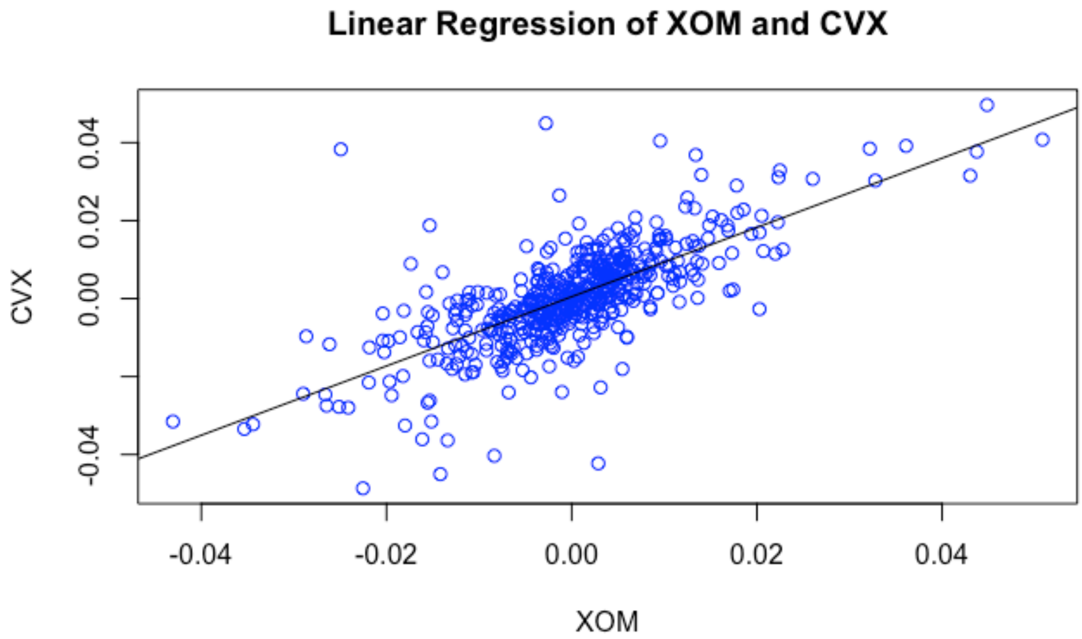
```
> relation

Call:
lm(formula = y ~ x)

Coefficients:
(Intercept)            x
  0.0004814    0.8882734
```

We can also get the summary of the linear regression,

```
> summary(relation)

Call:
lm(formula = y ~ x)

Residuals:
      Min        1Q    Median        3Q       Max
-0.045283 -0.004507  0.000057  0.004919  0.059951

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 0.0004814  0.0003980    1.21    0.227
x           0.8882734  0.0374877   23.70   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.008918 on 500 degrees of freedom
Multiple R-squared:  0.5289,    Adjusted R-squared:  0.528
F-statistic: 561.5 on 1 and 500 DF,  p-value: < 2.2e-16
```

p-value of each parameter is acceptable.

We can draw the plot of the regression.



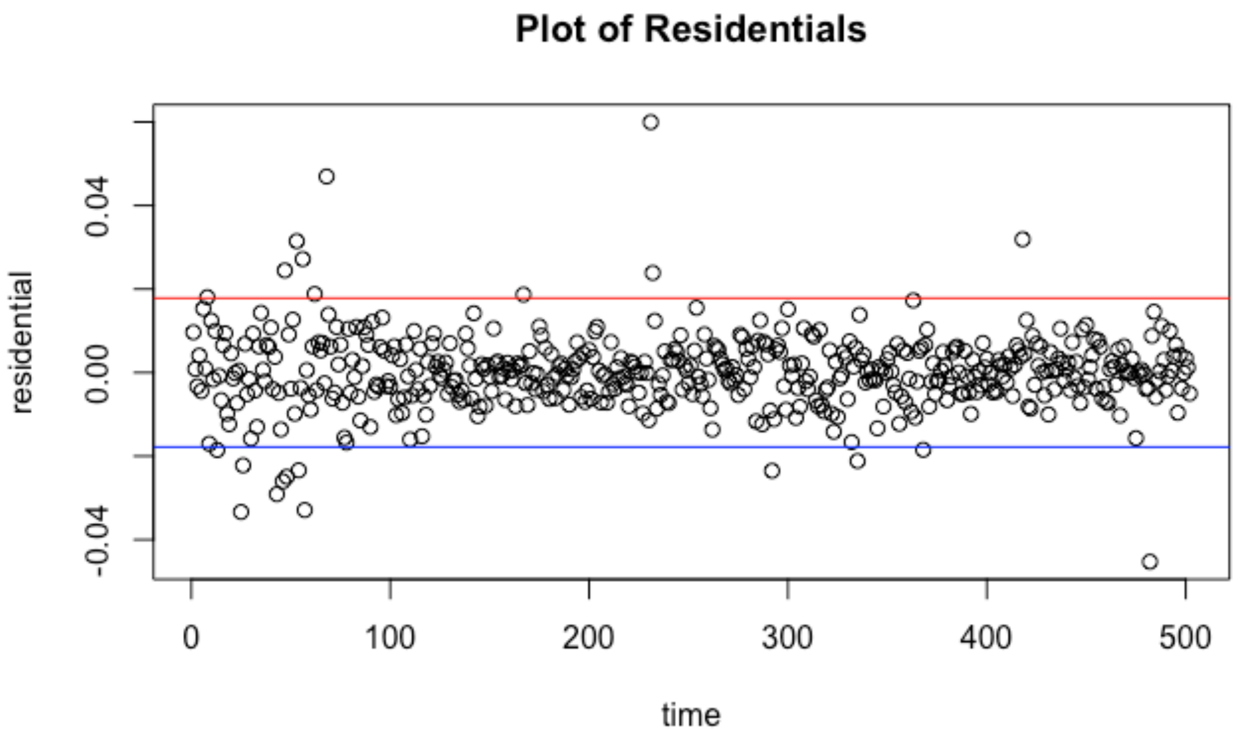
Step2:

```
> adf.test(z)

        Augmented Dickey-Fuller Test

data:  z
Dickey-Fuller = -8.5793, Lag order = 7, p-value = 0.01
alternative hypothesis: stationary

Warning message:
In adf.test(z) : p-value smaller than printed p-value
```

The p-value of is smaller than 0.01, so we reject the null hypothesis and accept the alternative hypothesis that its stationary.

Step3:

Draw a plot of the residentials and the cross point of delta and minus delta.



This is the table of transaction signals:

```
> transaction
                  1    1
Date9   2015-12-10 sell
Date10  2015-12-11  buy
Date14  2015-12-17  buy
Date63  2016-03-01 sell
Date168 2016-07-29 sell
Date364 2017-05-10 sell
Date369 2017-05-17  buy
```

##Question2

Question 2:

$a_t \sim N(0, \sqrt{0.02})$ $\quad$ $r_t = 0.01 + 0.2 r_{t-2} + a_t$

Because, the time series sequence r_t is weakly stationary.

$E[r_t] = E[r_{t-2}] = \mu$, $\quad$ $E(r_t) = 0.01 + 0.2\, E[r_{t-2}] + E[a_t]$

$\mu = 0.01 + 0.2\mu + 0$.

$$\mu = \frac{\phi_0}{1-\phi_2} = \frac{0.01}{0.8} = 0.0125.$$

$Var[r_t] = Var[\phi_0 + \phi_2 r_{t-2} + a_t]$

$\because$ r_{t-2} is independent of a_t, $\therefore$ $COV(r_{t-2}, a_t) = 0$.

$Var[r_t] = Var[r_{t-2}] \cdot \phi_2^2 + Var[a_t]$ $\quad$ $\gamma_0 = \frac{\sigma^2}{1-\phi_2^2}$

$\sigma^2 = 0.04\sigma^2 + 0.02$

$\sigma^2 \approx 0.021$.

lag-1. ACF: $\cancel{\rho_1 = \phi_1 \rho_{l-1}}$ $\quad$ $\phi_0 = \mu(1-\phi_2)$

$(r_t - \mu) = \phi_2 (r_{t-2} - \mu) + a_t$. $\quad$ $\gamma_0 = Var(r_t) = \sigma^2 = \frac{\sigma_a^2}{1-\phi_2^2}$

$\gamma_l = E[(r_t-\mu)(r_{t-l}-\mu)] = E[a_t(r_{t-l}-\mu) + \phi_2(r_{t-2}-\mu)(r_{t-l}-\mu)]$

~~lag-1~~

When lag-1, $\gamma_1 = 0.2 \cdot \gamma_1 \Rightarrow \gamma_1 = 0$.

$\rho_1 = \phi_1 + \phi_2 \cdot \rho_1 \Rightarrow \rho_1 = 0$.

when lag-2.

$\gamma_2 = E[(r_t-\mu)(r_{t-2}-\mu)] = E[\phi_1(r_{t-1}-\mu) + \phi_2(r_{t-2}-\mu)^2 + a_t(r_{t-2}-\mu)]$

$= \phi_2 Var(r_t)$

$\rho_2 = \frac{\gamma_2}{\gamma_0} = \phi_2 = \cancel{0.04}\ 0.2$

1-step forcast. at time t=100.

$\hat{r}_{100}(1) = E(r_{101} | F_{100}) = \phi_0 + \phi_2 r_{99} = 0.014.$

2-step forcast.

$\hat{r}_{100}(2) = E(r_{102} | F_{100}) = \phi_0 + \phi_2 r_{100} = 0.008.$

1-step forcast.

$e_{100}(1) = r_{101} - \hat{r}_{100}(1) = a_{101}.$ $Var[e_{101}(1)] = \sigma_a^2 = 0.02$

$Std = \sigma_a = 0.14$

2-step forcast.

Similarly.

$e_{100}(2) = r_{102} - \hat{r}_{100}(2) = a_{102} + \phi_1 a_{101} = a_{102}.$ $Var[a_{100}(2)] = \sigma_a^2 = \cancel{0.021}\ 0.02$

$std = \sigma_a = 0.14$

Code:

#homework4 Yifu He 10442277

invoke the package I will use
library(tseries)

#Question 1

```
##read the data into R
getwd()
setwd("/Users/yifuhe/Desktop")
cvx_ori <-read.csv("CVX.csv")
xom_ori <-read.csv("XOM.csv")
date<- unlist(cvx_ori[1])
cvx <-unlist(cvx_ori[5])
xom <-unlist(xom_ori[5])
y <-diff(log(cvx))
x <-diff(log(xom))
z<-c(NA)

##1.1
relation <-lm(y~x)
constant<-relation$coefficients[1]
sigma <- relation$coefficients[2]
summary(relation)
#png(file = "linearregression of XOM and CVX.png")
plot(x,y,col = "blue",main = "Linear Regression of XOM and CVX",xlab="XOM",ylab="CVX")
abline(lm(y~x),cex = 1.3,pch = 16)

##1.2
z<-relation$residuals
adf.test(z)

##1.3
stdz<-sd(z)
delta <-2*stdz
plot(z,ylab="residential",xlab="time",main="Plot of Residentials")
abline(h=-delta,col="blue")
abline(h=delta,col="red")
transaction <-data.frame(row.names=c("date","signals"))
for(i in 1:length(z)){
  if(abs(z[i]-delta)<=0.001){
  temp <- data.frame(date[i+1],"sell")
  names(temp)<-data.frame("date","signals")
  transaction <-rbind(transaction,temp)
  }
  if(abs(z[i]+delta)<=0.001){
    temp <- data.frame(date[i+1],"buy")
    names(temp)<-data.frame("date","signals")
    transaction <-rbind(transaction,temp)
  }
}
transaction
```