Weekly Homework 4 570

Weiping Zhang
Financial Engineering

November 26,2018

Problem 1.

Solution.
Firstly, we estimate the co-intergrating realtion. We use y_t(the log return of Chevron) and x_t to do a linear regression. We first do the normal resgression y_t x_t, the result shows that the term of the intercept is not significant. Therefore, we fit the model agagin without the intercept term, the coeffient of x_t is significant. And the α here is 0.8882734, c here is 0.0004814.

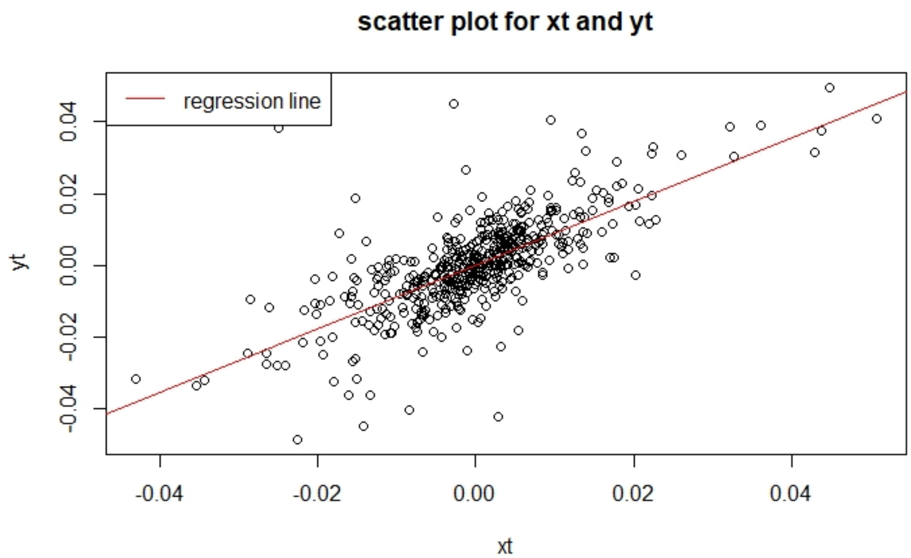
```
> fit<-lm(yt~xt)
> summary(fit)

Call:
lm(formula = yt ~ xt)

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    
xt          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
```



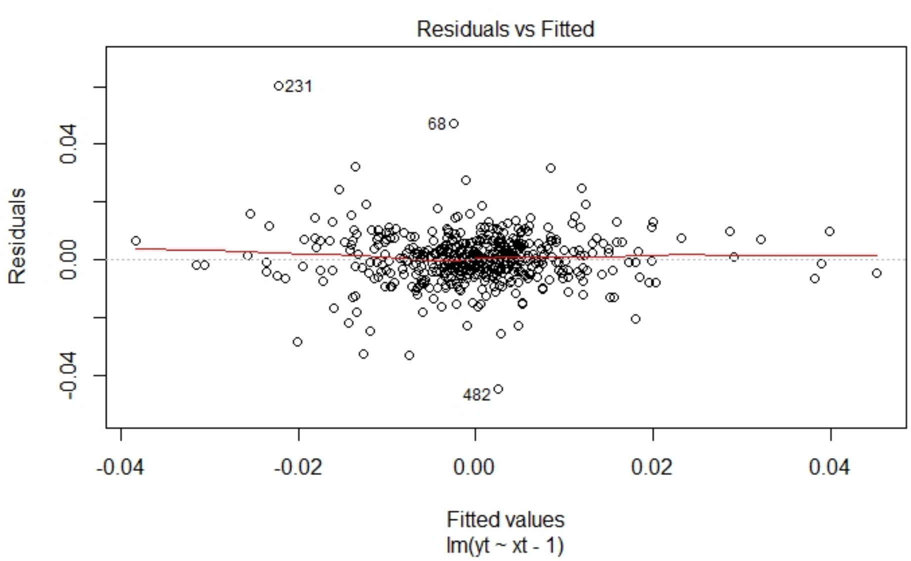
Following the procedure, we the need to test for stationarity of the residual z_t useing ADF unit root test. The result gave that the p-value is smaller than 0.01, therefore, under the cofidence level equals 0.01, we can reject the null hypothesis, which is that the series is non-stationary. So we can see that the residual is stationary in this case.

```
> adf.test(zt)

        Augmented Dickey-Fuller Test

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

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



Finally,we calculate Δ and follow the trading rules. When $|y_t - \alpha x_t - c + \Delta|$ is less than 0.001, we set the signal to 1. When $|y_t - \alpha x_t - c - \Delta|$ is less than 0.001, we set the signal to -1. Modify a new dataframe,we can get the result of long and short signals along with dates.

	Date	order	act
8	2015-12-10	8	short
9	2015-12-11	9	long
13	2015-12-17	13	long
62	2016-03-01	62	short
167	2016-07-29	167	short
363	2017-05-10	363	short
368	2017-05-17	368	long

Based on the signals, we can calculate the pnl. We assume that we long 1 porfolio is that we short 1 shares B and long ratio*1 shares A.We short 1 porfolio is that we buy 1 shares B and short ratio*1 shares A. Based on this rule, the pnl here is 12.23

R code attached:

```
install.packages("tseries")
library(tseries)
#Problem 1
xom<-read.csv((file.choose()))
head(xom)
cvx<-read.csv((file.choose()))
head(cvx)
xom_close<-xom$Close
cvx_close<-cvx$Close
n1<-length(xom_close)
n2<-length(cvx_close)
xt<-log(xom_close[-1]/xom_close[-n1])
yt<-log(cvx_close[-1]/cvx_close[-n2])
plot(xt,yt,main="scatter plot for xt and yt")
fit<-lm(yt~xt)
summary(fit)
coef<-fit$coefficients
coef
alpha<-as.vector(coef)[2]
alpha
const<-as.vector(coef)[1]
const
abline(fit,col="red")
legend("topleft", lty=c(1), col=c("red"),
       legend=c("regression line"))
zt<-fit$residuals
plot(fit,which=1)
#the non-linear trend is not obvious
adf.test(zt)
```

```
#p-value less than 0.01,rejected the null hypothesis. Under alpha=0.01, it's static
delta<-2*sd(zt)
date<-as.Date(xom$Date[-1])

yt_axt<-yt-alpha*xt
table<-as.data.frame(cbind(xt,yt,yt_axt))
table$Date=date
table<-table[,c('Date','xt','yt','yt_axt')]
table$signal=c(rep(0,nrow(table)))
table[abs((table$yt_axt+delta-const))<=0.001,]$signal<-1
table[abs((table$yt_axt-delta-const))<=0.001,]$signal<--1
order<-c(1:nrow(table))
table$order=order
table$act=c(rep("act",nrow(table)))
table[table$signal==1,]$act<-"long"
table[table$signal==-1,]$act<-"short"
trade<-subset(table[,c(-2,-3,-4,-5)],table$act!="act")
trade
#calculate the pnl at the accuracy of 0.001
cal_xp<-xom_close[(trade$order+1)]
cal_cp<-cvx_close[(trade$order+1)]
#Assuming that we long 1 porfolio is that we short 1 shares B and long ratio shars
#we short 1 porfolio is that we buy 1 shares B and short ratio shares A
#calculate the pnl under the accuracy of 0.01
sig<-trade$act
sig
sig[sig=="short"]=1
sig[sig=="long"]=-1
sig<-as.numeric(sig)
sig_A<-sig*alpha
sig_A[length(sig_A)]=-2*alpha
sig_A
sig_B<-sig*(-1)
sig_B[length(sig_B)]=2
sig_B
pnl_A<-sig_A%*%cal_xp
pnl_B<-sig_B%*%cal_cp
pnl_A
pnl_B
cal_xp
pnl<-pnl_A+pnl_B
pnl
```

□

Problem 2.

Solution.

The time series r_t model follows the model:

$$r_t = 0.01 + 0.2r_{t-2} + a_t$$

Take expectation on both sides:

$$\begin{aligned} Er_t &= 0.01 + 0.2Er_{t-2} + a_t \\ \mu &= 0.01 + 0.2\mu + 0 \\ (1-0.2)\mu &= 0.01(*) \\ \mu &= 0.0125 \end{aligned}$$

Using (*), the model can be rewritten as:

$$r_t - \mu = 0.2\left(r_{t-2} - \mu\right) + a_t \tag{1}$$

By repeated subsitutions, the prior equation implies that:

$$\begin{aligned} r_t - \mu &= a_t + 0.2r_{t-2} + 0.2^2 r_{t-4} + ... \\ &= \sum_{i=0}^{\infty} 0.2^i a_{t-2i} \end{aligned}$$

Therefore, we have $E[a_t\left(r_t - \mu\right)] = \sigma_a^2$, $E[a_t\left(r_{t-1} - \mu\right)] = 0$, $E[a_t\left(r_{t-2} - \mu\right)] = 0$.
Multiplying (1) by $(r_{t-2} - \mu), (r_t - \mu), (r_{t-1} - \mu)$. Then take expectation on both sides, we have:

$$\gamma_2 = 0.2\gamma_0 \tag{2}$$
$$\gamma_0 = 0.2\gamma_2 + \sigma_a^2 \tag{3}$$
$$\gamma_1 = 0.2\gamma_1 \tag{4}$$

From (4), we know that γ_1=0. From (2) and (3), we can solve that

$$\begin{aligned} \gamma_0 &= \frac{1}{48} \\ \gamma_2 &= \frac{1}{240} \end{aligned}$$

Dividing γ_1 and γ_2 by γ_0, we then have the lag-1 and lag-2 correlations or rt:

$$\begin{aligned} \rho_1 &= 0 \\ \rho_2 &= \frac{1}{5} \end{aligned}$$

1-Step Ahead Forecast:
The point forecast of r_{t+1} given $F_t = r_t, rt-1, ...$ is the conditional expectation

$$\begin{aligned} \hat{r}_t\left(1\right) &= E\left(r_{t+1} | \mathscr{F}_t\right) \\ &= 0.01 + 0.2r_{t-1} \end{aligned}$$

Therefore, the 1-step forecast is 0.01+0.2*0.02=0.014
and the associated forecast error is:

$$e_t\left(1\right) = r_{t+1} - \hat{r}_t\left(1\right) = a_{t+1}$$

Consequently, the associated standard deviations of the 1-step forecast error is

$$\sqrt{Var[e_t\left(1\right)]} = \sqrt{Var(a_{t+1})} = \sigma_a = 0.14142$$

The point forecast of r_{t+2} given $F_t = r_t, rt - 1, ...$ is the contional expectation

$$\begin{aligned}\hat{r}_t\left(2\right) &= E\left(r_{t+2}|\mathscr{F}_t\right)\\ &= 0.01 + 0.2r_t\end{aligned}$$

Therefore, the 2-step forecast is 0.01+0.2*(-0.01)=0.008
and the associated forecast error is:

$$e_t\left(2\right) = r_{t+2} - \hat{r}_t\left(2\right) = a_{t+2}$$

Consequently, the associated standard deviations of the 2-step forecast error is

$$\sqrt{Var[e_t\left(2\right)]} = \sqrt{Var(a_{t+2})} = \sigma_a = 0.14142$$

□