R 실습 4 : 불안정시계열

1. 시뮬레이션 : GDP of Korea
2. 시뮬레이션 : 가성회귀
3. 단위근 검정
4. 시뮬레이션 : 공적분시스템
5. Engle-Yoo 공적분 검정

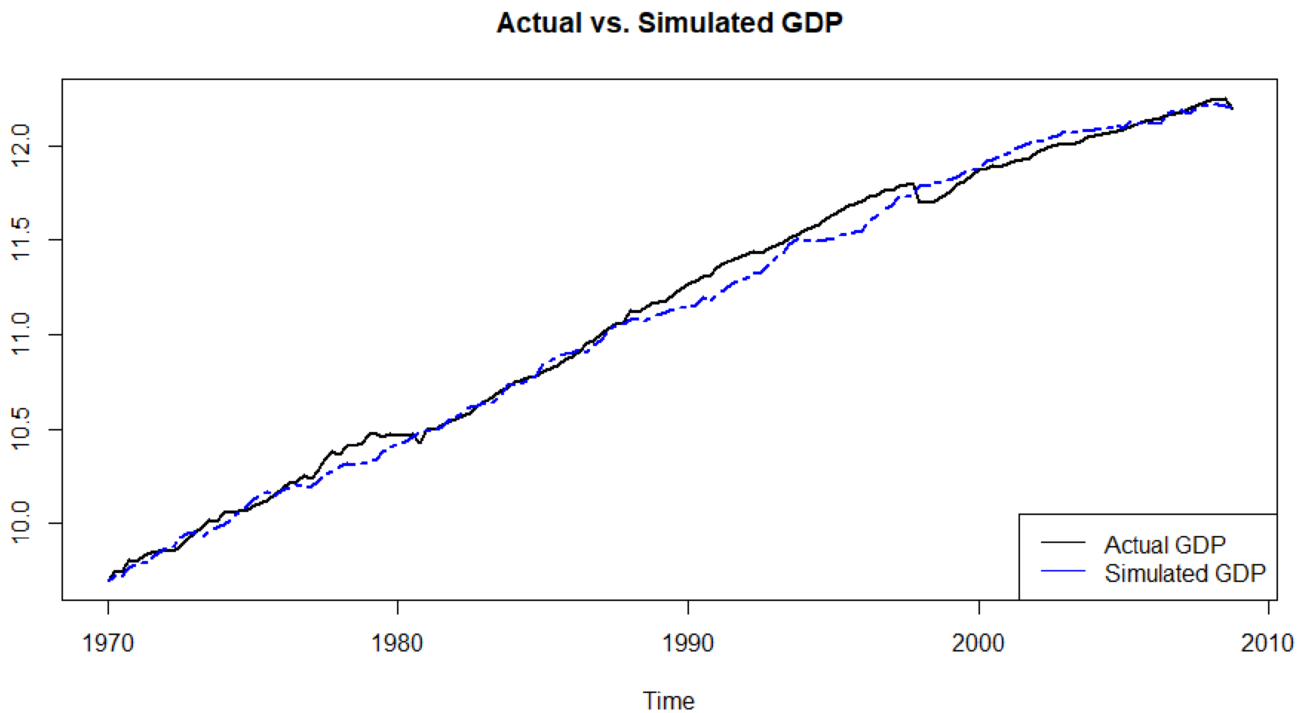
1. 시뮬레이션 : GDP of Korea

– (한국의 예 (1970:1–2008:4): 다음과 같은 가상적인 random walk process를 설정

$$y_t = 0.01607558 + y_{t-1} + e_t \ , \ e_t \sim iidN(0, 0.01846839)$$

– 위 모형으로 가상적인 시계열을 만들어 실제값과 비교해 보면 매우 유사한 변화를 보이고 있음천히 감소함

– 이는 한국경제의 GDP가 random walk process에 따른다는 하나의 증거가 됨



(chap8-R-1.R)

```
library(openxlsx)

data1<-read.xlsx("http://kanggc.iptime.org/time/R/lgdpq_sa(70-08).xlsx")
lgdp<-data1$lgdp_sa
lgdp_sa<-ts(lgdp, start=c(1970,1), frequency=4)
glgdp_sa<-lgdp_sa-lag(lgdp_sa, k=-1)

mean(glgdp_sa)
sqrt(var(glgdp_sa))

set.seed(1)
x1<-w<-rnorm(156, mean=0, sd=0.01846839)
lgdp_sa[1]
x1d<-data.frame(x1)
x1d[1,1]<-lgdp_sa[1]
x1d.ts<-ts(x1d)
x1d.ts[1]

mean(x1)
var(x1)
sqrt(var(x1))

for(t in 2:156) x1d.ts[t]=0.01607558+x1d.ts[t-1]+w[t]
x1d<-ts(x1d.ts,start=c(1970,1), frequency=4)

lgdp_sa
x1d

plot(lgdp_sa, lwd=2, type="l", ylab="", main="Actual vs. Simulated GDP")
lines(x1d, lwd=2, lty=6, col="blue")
legend("bottomright", legend=c("Actual GDP","Simulated GDP"),col=c("black", "blue"),lty=1)
```

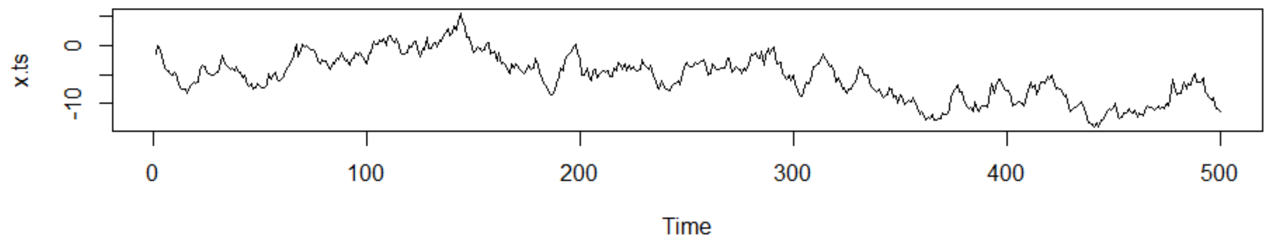
2. 시뮬레이션 : 가성회귀

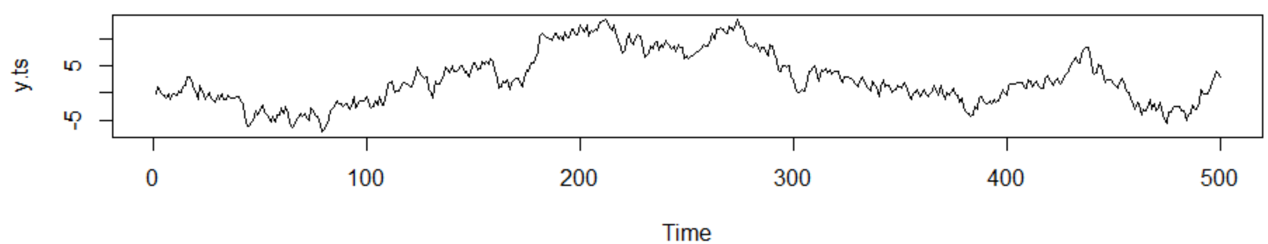
- 가성회귀란 시계열이 확률적 추세를 가졌음에도 불구하고 회귀분석방법으로 확정적 추세를 제거한 후 사용하면 표본의 수가 증가함에 따라 상관관계가 없는 변수 사이에도 마치 강한 상관관계가 있는 것으로 나타나는 것을 말함
- 다음의 두 임의보행 확률과정에서 관측치를 생성하면 두 계열은 불안정계열이며 서로 상관관계가 없음

$$y_t = y_{t-1} + u_t \ , \ u_t \sim iidN(0, \sigma_u^2)$$

$$x_t = x_{t-1} + v_t \ , \ v_t \sim iidN(0, \sigma_v^2)$$

- y_t를 x_t에 대해 회귀모형을 추정하면 회귀계수가 통계적으로 유의하게 나옴





(chap8-R-2.R)

```
set.seed(123)
x1<-w1<-rnorm(500)
for(t in 2:500) x1[t]=x1[t-1]+w1[t]
(x1.ts<-ts(x1))

set.seed(1234)
y1<-z1<-rnorm(500)
for(t in 2:500) y1[t]=y1[t-1]+z1[t]
y1.ts<-ts(y1)

par(mfrow=c(2,1))
plot(x1.ts); plot(y1.ts)

lm(y1~x1)
ols1<-lm(y1~x1)
summary(ols1)

set.seed(123345)
x2<-w2<-rnorm(500)
for(t in 2:500) x2[t]=0.8*x2[t-1]+w2[t]
(x2.ts<-ts(x2))

set.seed(123456)
y2<-z2<-rnorm(500)
for(t in 2:500) y2[t]=0.5*y2[t-1]+z2[t]
y2.ts<-ts(y2)

plot(x2.ts); plot(y2.ts)

lm(y2~x2)
ols2<-lm(y2~x2)
summary(ols2)
```

3. 단위근 검정

(1) DF 검정

$$\Delta y_t = (\rho - 1)y_{t-1} + e_t,\ e_t \sim iidN(0, \sigma_e^2) \rightarrow \tau \text{ 를 이용}$$

$$\Delta y_t = \alpha + (\rho - 1)y_{t-1} + e_t,\ e_t \sim iidN(0, \sigma_e^2) \rightarrow \tau_\mu \text{를 이용}$$

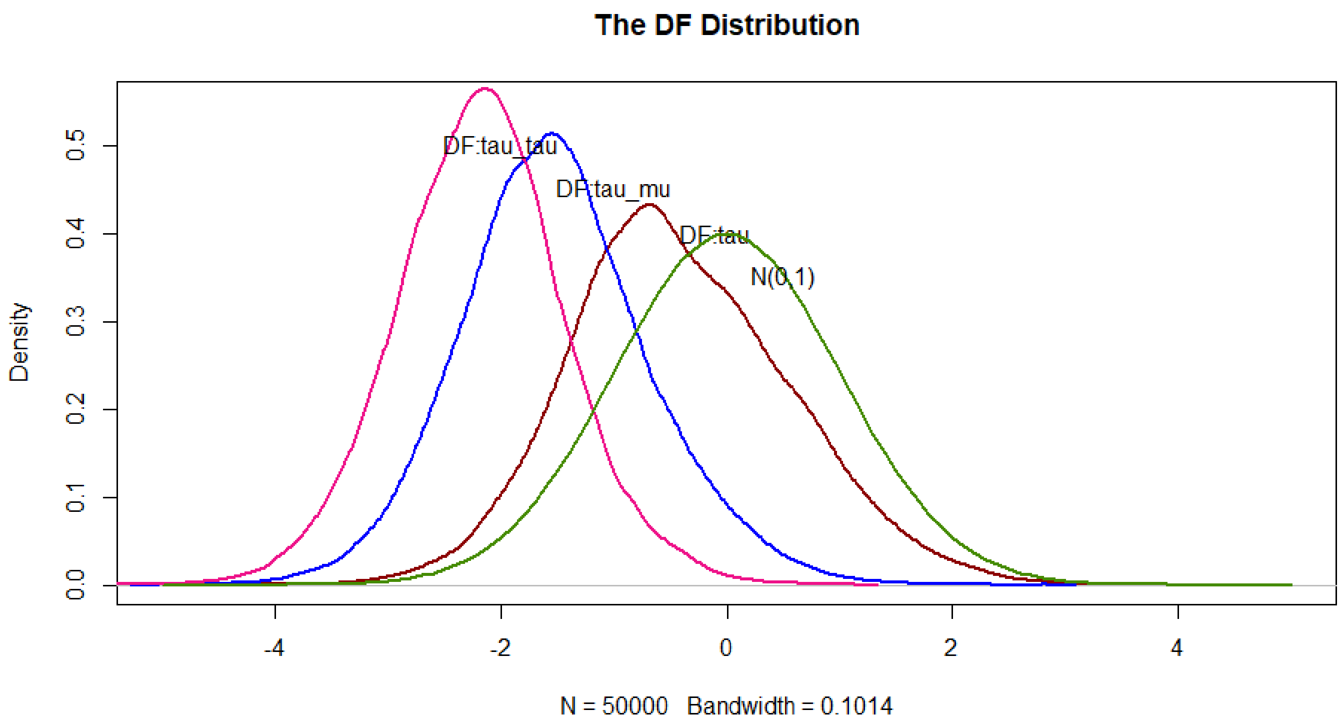
$$\Delta y_t = \alpha + \beta t + (\rho - 1)y_{t-1} + e_t,\ e_t \sim iidN(0, \sigma_e^2) \rightarrow \tau_\tau \text{를 이용}$$

(2) ADF 검정

$$\Delta y_t = (\rho - 1)y_{t-1} + \sum_{j=1}^{p} \gamma_j \Delta y_{t-j} + e_t,\ e_t \sim iidN(0, \sigma_e^2) \rightarrow \tau \text{ 를 이용}$$

$$\Delta y_t = \alpha + (\rho - 1)y_{t-1} + \sum_{j=1}^{p} \gamma_j \Delta y_{t-j} + e_t,\ e_t \sim iidN(0, \sigma_e^2) \rightarrow \tau_\mu \text{를 이용}$$

$$\Delta y_t = \alpha + \beta t + (\rho - 1)y_{t-1} + \sum_{j=1}^{p} \gamma_j \Delta y_{t-j} + e_t,\ e_t \sim iidN(0, \sigma_e^2) \rightarrow \tau_\tau \text{를 이용}$$



(chap8-R-3.R)

```
library(tseries)
library(urca)

set.seed(123)
e=rnorm(500)
x=cumsum(e)
plot(x,type="l")

# DF Unit Root Test

dx=diff(x)
n=length(dx)
lagx=x[1:n]
tr=1:n

summary(lm(dx~0+lagx))
qnorm(c(0.01, 0.05, 0.1)/2)

df<-ur.df(x, type="none", lags=0)
summary(df)

summary(lm(dx~lagx))

df.d<-ur.df(x, type="drift", lags=0)
summary(df.d)

summary(lm(dx~tr+lagx))

df.t<-ur.df(x, type="trend", lags=0)
summary(df.t)
```

(chap8-R-4.R)

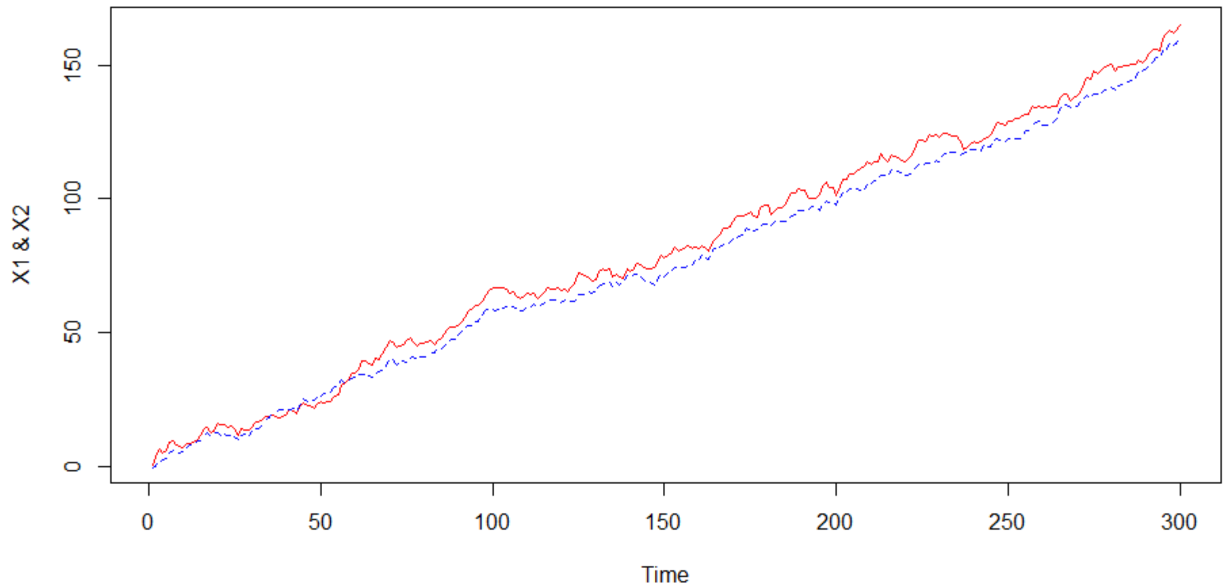
```
library(urca)

set.seed(123)
e=rnorm(500)
(x=cumsum(e))
plot(x,type="l")


# ADF Unit Root Test

(dx=diff(x))
n=(length(dx))
n
(lagx=x[2:n-1])
tr=1:498
ndx=dx[2:n]
lagdx=dx[1:n-1]

summary(lm(ndx~0+lagx+lagdx))

df<-ur.df(x, type="none", lags=1)
summary(df)

summary(lm(ndx~lagx+lagdx))

df.d<-ur.df(x, type="drift", lags=1)
summary(df.d)

summary(lm(ndx~tr+lagx+lagdx))

df.t<-ur.df(x, type="trend", lags=1)
summary(df.t)
```

4. 시뮬레이션 : 공적분시스템

$x_{2t} = 0.5 + x_{2t-1} + u_{2t},\ u_{2t} \sim N(0,1)$ (임의보행과정)

$u_{1t} = 0.9u_{1t-1} + v_t,\ v_t \sim N(0,1)$ (AR(1)오차)

$x_{1t} = 4.5 + x_{2t} + u_{1t}$ (공적분관계)



(chap8-R-5.R)

```
set.seed(123)
x2<-u2<-rnorm(300)
for(t in 2:300) x2[t]=0.5+x2[t-1]+u2[t]
x2.ts<-ts(x2)

set.seed(1234)
u1<-v<-rnorm(300)
for(t in 2:300) u1[t]=0.9*u1[t-1]+v[t]
u1.ts<-ts(u1)

set.seed(12345)
x1<-z<-rnorm(300)
for(t in 2:300) x1[t]=4.5+x2[t]+u1[t]
x1.ts<-ts(x1)

plot(x1, type="l", col="red", xlab="Time", ylab="X1 & X2")
lines(x2, lty=2, col="blue")
```

5. Engle-Yoo 공적분 검정

$y_t = \alpha + \beta x_t + \varepsilon_t$ (공적분 회귀식)

$\Delta\varepsilon_t = \mu + (\rho - 1)\varepsilon_{t-1} + \sum_{j=1}^{4} \gamma_j \Delta\varepsilon_{t-j} + e_t$ (단위근 검정)

공적분 검정통계량의 임계치
(ADF검정통계량인 경우)

모형내 변수	표본크기	유의 수준	
		0.05	0.10
2	50	-3.29	-2.90
	100	-3.17	-2.91
	200	-3.25	-2.98
3	50	-3.75	-3.36
	100	-3.62	-3.32
	200	-3.78	-3.51
4	50	-3.98	-3.67
	100	-4.02	-3.71
	200	-4.13	-3.83
5	50	-4.15	-3.85
	100	-4.36	-4.06
	200	-4.43	-4.14

(chap8-R-6.R)

```
library(urca)

set.seed(123)
x2<-u2<-rnorm(300)
for(t in 2:300) x2[t]=0.5+x2[t-1]+u2[t]

set.seed(1234)
u1<-v<-rnorm(300)
for(t in 2:300) u1[t]=0.9*u1[t-1]+v[t]

set.seed(12345)
x1<-z<-rnorm(300)
for(t in 2:300) x1[t]=4.5+x2[t]+u1[t]

n=length(x1); tr=1:n

# Unit Root Test
x1.t<-ur.df(x1, type="trend", lags=1)
summary(x1.t)

x2.t<-ur.df(x2, type="trend", lags=1)
summary(x2.t)

# Cointegration Regression
#cr<-lm(x2~x1)
cr<-lm(x2~tr+x1)
summary(cr)
res<-(cr$resid)
```

(계속)

```
# Engle-Yoo Cointegration Test
res1.t<-ur.df(res, type="drift", lags=1)
summary(res1.t)

res2.t<-ur.df(res, type="drift", lags=2)
summary(res2.t)

res3.t<-ur.df(res, type="drift", lags=3)
summary(res3.t)

res4.t<-ur.df(res, type="drift", lags=4)
summary(res4.t)

resaic.t<-ur.df(res, type="drift",
selectlags=c("AIC"))
summary(resaic.t)

resbic.t<-ur.df(res, type="drift",
selectlags=c("BIC"))
summary(resbic.t)
```