Lecture 6. Unit-Root Nonstationarity

1. Random Walk with Drift

Definition 1.1. A *random walk with drift* is defined by

$$x_t = \delta + x_{t-1} + \varepsilon_t,$$

where $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$. The constant term δ represents the time trend of x_t and is referred to as the *drift* of the model.

- Starting at the fixed value x_0, one obtains:

ε : white noise

$$\begin{aligned}
x_1 &= \delta + x_0 + \varepsilon_1 \\
x_2 &= \delta + x_1 + \varepsilon_2 = 2\delta + x_0 + \varepsilon_2 + \varepsilon_1 \\
x_3 &= \delta + x_2 + \varepsilon_3 = 3\delta + x_0 + \varepsilon_3 + \varepsilon_2 + \varepsilon_1 \\
&\vdots \\
x_t &= x_0 + \delta t + \varepsilon_t + \varepsilon_{t-1} + \cdots + \varepsilon_1.
\end{aligned}$$

So, the mean and variance are

$x_t \sim (x_0 + \delta t, \sigma_\varepsilon^2 t)$

time dependent

$$E[x_t] = x_0 + \delta t$$
$$Var[x_t] = \sigma_\varepsilon^2 t,$$

// $\neq$(not time invariant. time dependent

both of which are "time dependent" (i.e., no longer stationary). A positive drift δ ~~implies~~ that x_t eventually goes to ∞, while a negative drift implies that x_t converges to $-\infty$.

time trend

Remark 1.2. A random walk model with drift has been widely considered as a statistical model for the movement of logged stock prices:

$$p_t = \mu + p_{t-1} + \varepsilon_t,$$

where $p_t = \log(P_t)$ and $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$. Using $r_t = p_t - p_{t-1}$, the model is rewritten as

$$r_t = \mu + \varepsilon_t,$$

thereby meaning that a return series follows a white noise process with mean μ. Since the white noise process is featured with no autocorrelation, hence, the past returns $\{r_{t-1}, r_{t-2}, r_{t-3}, \ldots\}$ are not relevant to the current return r_t, which is exactly equivalent to the notion of a *weak-form market efficiency*.

Example 1.3. Simulate a random walk with drift when $\delta = 0.25$, $x_0 = 1$, and $\sigma_\varepsilon^2 = 1$

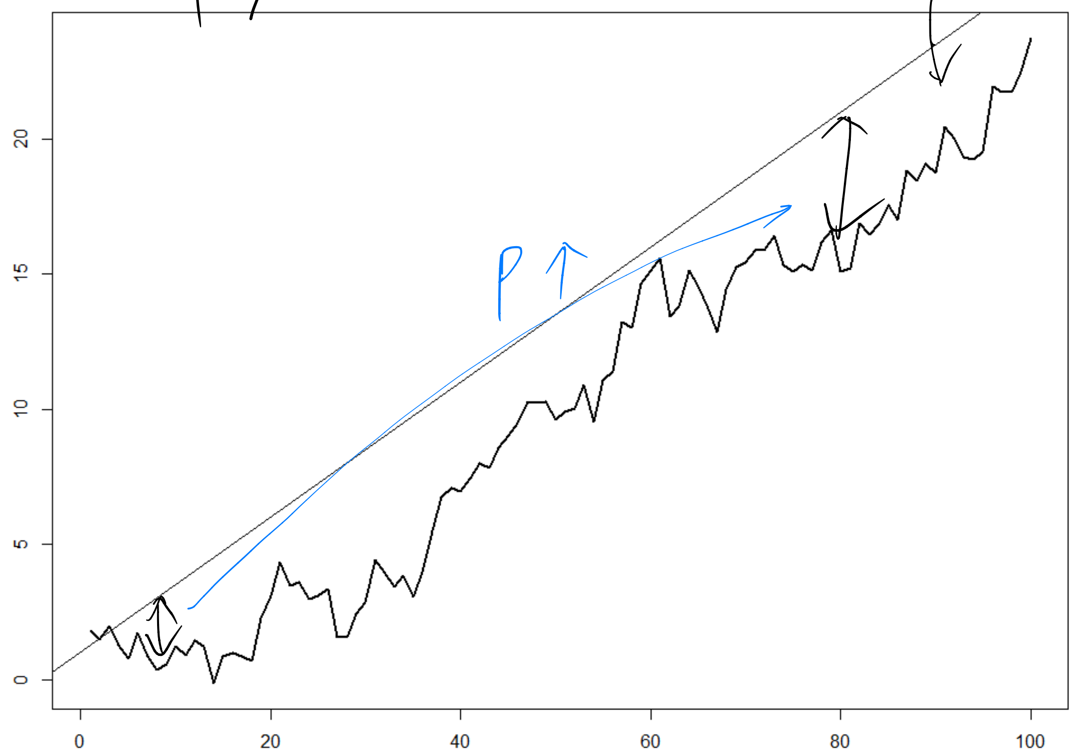
```
> e <- rnorm(n = 100)
> t <- 1:100
> x <- 1 + 0.25*t + cumsum(e)
> plot.ts(x, lwd = 2, xlab = "", ylab ="")
```

```
> abline(a = 1, b = 0.25)
```

$Var[x_t] = \sigma_\varepsilon^2 \cdot t \to \infty$ as $t \to \infty$

imply

$E[x_t]$...



$x_{t+s} = \delta + x_{t+s-1} + \varepsilon_{t+s}$

$x_{t+s-1} = \delta + x_{t+s-2} + \varepsilon_{t+s-1}$

$2\delta + x_{t+s-2} + \varepsilon_{t+s-1} + \varepsilon_{t+s}$.

Remark 1.4. The s-step ahead forecast is

$$
\begin{aligned}
\hat{x}_t[s] &= E[x_{t+s}|I_t] \\
&= E[\delta + x_{t+s-1} + \varepsilon_{t+s}|I_t] \\
&= E[2\delta + x_{t+s-2} + \varepsilon_{t+s} + \varepsilon_{t+s-1}|I_t] \\
&= E[\delta s + x_t + \varepsilon_{t+s} + \cdots + \varepsilon_{t+1}|I_t] \\
&= \delta s + x_t
\end{aligned}
$$

MA(1) expected value converges to μ.
($\because$ $\hat{x}_t(1) = \mu + \cdots$, $\hat{x}_t(2) = \mu$, $\hat{x}_t(3) = \mu$).

for any $s \geq 1$. So, the s-step ahead forecast is "not" mean reverting. The s-step ahead forecast error is

but this not conv.
$\hat{x}_t[s] \to \infty$ as $s \to \infty$
$= \delta s + x_t$

$$
\begin{aligned}
e_t[s] &= x_{t+s} - \hat{x}_t[s] \\
&= \varepsilon_{t+s} + \varepsilon_{t+s-1} + \cdots + \varepsilon_{t+1}
\end{aligned}
$$

and, hence, its variance is $Var[e_t[s]] = s\sigma_\varepsilon^2$. Since the variance of the s-step ahead forecast error explodes as $s \to \infty$, the precision of the forecast $\hat{x}_t[s]$ diminishes as s increases. This implies that the random walk model is "not" predictable.

$s\sigma_\varepsilon^2 \to \infty$ as $s \to \infty$

$\delta(t+s) + x_0 + \varepsilon_1 + \cdots + \varepsilon_t + \cdots + \varepsilon_{t+s}$

$x_{t+s} = \delta s + x_t + \varepsilon_{t+1} + \cdots + \varepsilon_{t+s}$

$= \delta t + \varepsilon_1 + \cdots + \varepsilon_t$

t

$\sum_{i=1} \varepsilon_i$

Remark 1.5. From the representation $x_t = x_0 + \delta t + \sum_{i=0}^{t-1} \varepsilon_{t-i}$, one obtains

time-trend non stationary

$$\frac{\partial x_{t+s}}{\partial \varepsilon_t} = 1$$

$(\because t \cdot \varepsilon_{t-i})$

variance : time dependent

for any $s \geq 0$. So, the impact of the shock ε_t on the random walk process x_{t+s} does not decay over time, or equivalently, the shock has a *permanent* effect.

current shock → feature value.

tempt. :

permanent. :

1 unit

2. Trend-Stationary Time Series

Definition 2.1. A time series $\{x_t\}$ is a *trend-stationary* process if it has the form

$$x_t = \underbrace{\beta_0 + \beta_1 t}_{\text{trend}} + \underbrace{u_t}_{\text{stationary}},$$

where u_t is a "stationary" time series.

- A trend-stationary model is closely related to a random walk model with drift, since both exhibit linear trend. In the trend-stationary model, one sees that

$$E[x_t] = \beta_0 + \beta_1 t + E[u_t],$$

⇒ time variant. not stationary.

which depends on time, and

$$Var[x_t] = Var[u_t],$$

⇒ time-invariant because u_t is stationary.

which is finite and time invariant. However, the mean and variance of the random walk model with drift are both time dependent.

random walk model : var : time dependent

- The trend-stationary series can be transformed into a stationary one by removing the time trend $\beta_0 + \beta_1 t$; i.e.,

$$x_t - \beta_0 + \beta_1 t = u_t$$

but trend-stationary var : time invariant.

is stationary.

Example 2.2. Suppose that u_t is a zero-mean stationary AR(1) time series:

$$x_t = \beta_0 + \beta_1 t + u_t$$
$$u_t = \phi_1 u_{t-1} + \varepsilon_t,$$

where $|\phi_1| < 1$ and $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$. The mean is

$$E[x_t] = \beta_0 + \beta_1 t + E[u_t] = \beta_0 + \beta_1 t,$$

which is time dependent and grows linearly in time with rate ~~δ~~ β_1, and the variance is

$$Var[x_t] = Var[u_t] = \frac{\sigma_\varepsilon^2}{1 - \phi_1^2},$$

→ time-invariant.

which is time invariant. The finite variance of x_t suggests that the trend-stationary x_t exhibits *trend reversion* in that x_t never deviates too far away from the deterministic trend $\beta_0 + \beta_1 t$. This is sharply contrast to the fact that the random walk model with drift has the time-varying variance (i.e., $Var[x_t] = \sigma_\varepsilon^2 t$) and so x_t does not exhibit the same trend reversion (i.e., as t increases, $Var[x_t]$ explodes).

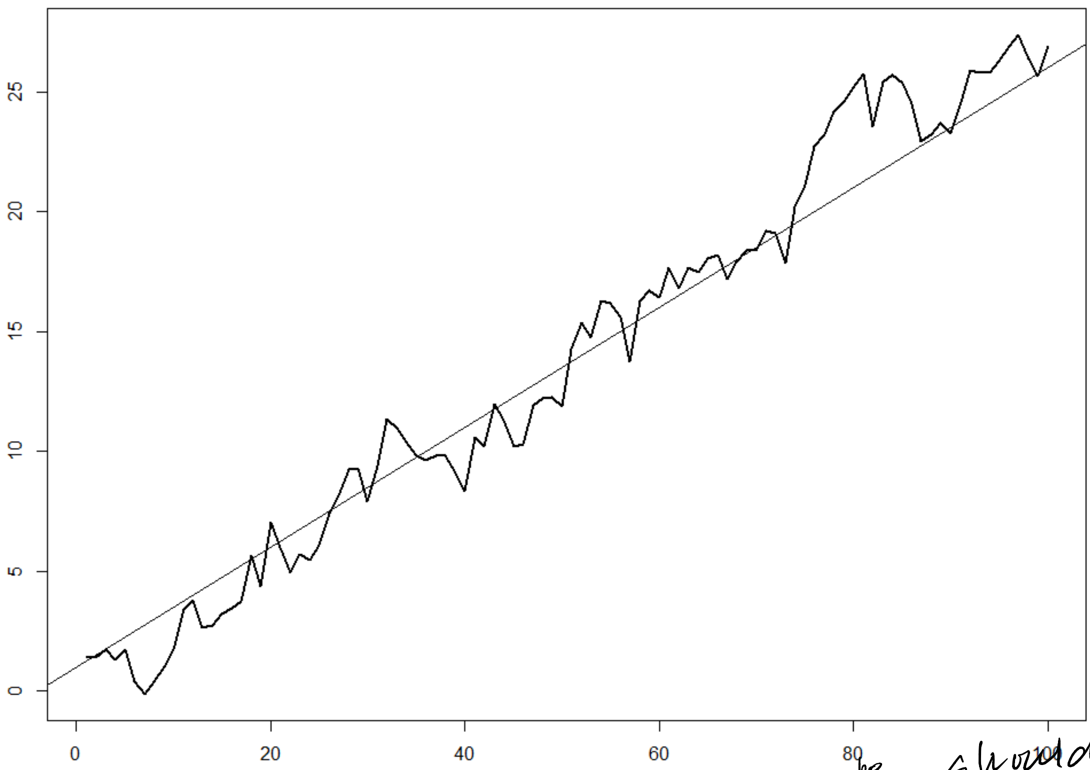
Example 2.3. Simulate the trend stationary AR(1) process x_t of the form

$$x_t = 1 + 0.25t + u_t$$
$$u_t = 0.75u_{t-1} + \varepsilon_t,$$

random walk : time-varying
trend-stationary : tred reversion

where $\varepsilon_t \sim WN(0,1)$ for $t = 1, \ldots, 100$.

```
> u <- arima.sim(list(order(1, 0, 0), ar = 0.75), sd = sqrt(1), n = 100)
> x <- 1 + 0.25*t + u
> plot.ts(x, lwd = 1, xlab = "", ylab ="")
> abline(a = 1, b = 0.25)
```



maybe should be random walk process"

Remark 2.4. Another difference between the trend-stationary and unit-root nonstationary models is about the permanence of shocks. Consider the trend-stationary AR(1) process. Since u_t is stationary (i.e., $|\phi_1| < 1$), one obtains

$$x_t = \beta_0 + \beta_1 t + \frac{\varepsilon_t}{1 - \phi_1 L} = \beta_0 + \beta_1 t + \sum_{i=0}^{\infty} \phi_1^i \varepsilon_{t-i}$$

stationary"

$\varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_1^2 \varepsilon_{t-2} + \phi_1^3 \varepsilon_{t-3} + \cdots$

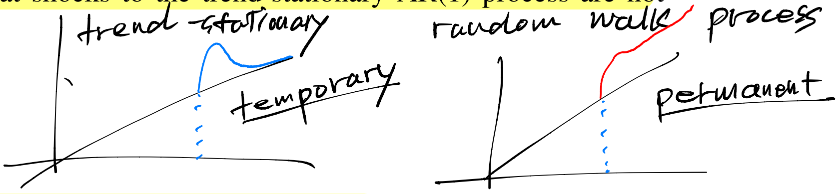
$\varepsilon_{t+s} + \phi_1 \varepsilon_{t+s-1} + \cdots + \phi_1^s \varepsilon_t + \cdots$

1) variance time - variant / invariant
2) shock . (permanent / temporary

and

$$\frac{\partial x_{t+s}}{\partial \varepsilon_t} = \phi_1^s$$

for any $s \geq 1$. So, the stationarity implies that shocks to the trend-stationary AR(1) process are not permanent but "temporary."



3. Unit-Root Nonstationary Models

Definition 3.1. A time series$\{x_t\}$ is called a *unit-root nonstationary* time series if it has the form

$$x_t = \delta + x_{t-1} + \psi(L)\varepsilon_t,$$

where $\psi(L) = \psi_0 + \psi_1 L + \psi_2 L^2 + \cdots$ with $\psi(1) \neq 0$ and $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$. A random walk model with drift is a special case of the unit-root nonstatioanry model when $\psi(L) = 1$.

- Starting at the fixed value x_0, one obtains:

$$\begin{aligned}
x_1 &= \delta + x_0 + \psi_0 \varepsilon_1 \\
x_2 &= \delta + x_1 + \psi_0 \varepsilon_2 + \psi_1 \varepsilon_1 \\
&= \delta + (\delta + x_0 + \psi_0 \varepsilon_1) + \psi_0 \varepsilon_2 + \psi_1 \varepsilon_1 \\
&= 2\delta + x_0 + \psi_0 \varepsilon_2 + (\psi_0 + \psi_1)\varepsilon_1 \\
x_3 &= \delta + x_2 + \psi_0 \varepsilon_3 + \psi_1 \varepsilon_2 + \psi_2 \varepsilon_1 \\
&= \delta + (2\delta + x_0 + \psi_0 \varepsilon_2 + (\psi_0 + \psi_1)\varepsilon_1) + \psi_0 \varepsilon_3 + \psi_1 \varepsilon_2 + \psi_2 \varepsilon_1 \\
&= 3\delta + x_0 + \psi_0 \varepsilon_3 + (\psi_0 + \psi_1)\varepsilon_2 + (\psi_0 + \psi_1 + \psi_2)\varepsilon_1 \\
&\vdots \\
x_t &= \delta t + x_0 + \sum_{i=0}^{t-1} \varepsilon_{t-i} \left(\sum_{j=0}^{i} \psi_j \right). \qquad (3.1)
\end{aligned}$$

From (3.1), the mean is

$$E[x_t] = \delta t + x_0,$$

which is *time variant* if $\delta \neq 0$, and the variance is

$$Var[x_t] = \sigma_\varepsilon^2 \sum_{i=0}^{t-1} \left(\sum_{j=0}^{i} \psi_j \right)^2,$$

which is also *time variant.*

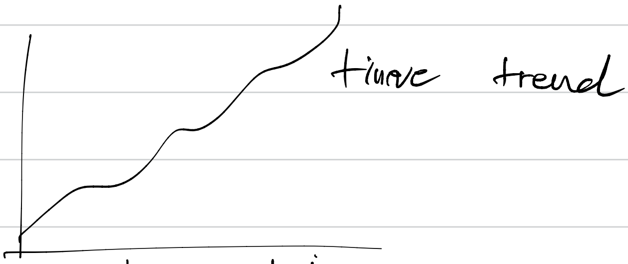
should be stationary

4. ARIMA Models

(

For modeling the trending data (which is nonstationary) in the ARMA form (which is applied to a stationary series), a trend should be removed before fitting the ARMA models to the data. Two common trend removal or de-trending procedures are *first differencing* and *time-trend regression*.

use trend-stationary to remove trend



two-choices

1) first diff

$$\Delta x_t = x_t - x_{t-1} \sim I(0) \Rightarrow \Delta x_t \sim ARMA(p, q).$$

2) time-trend regression

$$x_t = \beta_0 + \beta_1 x_{t-1} + u_t$$

$$\Rightarrow x_t - \beta_0 - \beta_1 x_{t-1} = u_t \sim I(0)$$

$$\Rightarrow x_t - \beta_0 - \beta_1 x_{t-1} \sim ARMA(p, q)$$

unit root nonstationary : first diff
trend - stationary : time trend regression

- For the unit-root nonstationary series x_t, one applies the ARMA model to Δx_t. For the trend-stationary series x_t, one applies the ARMA model augmented with an appropriate form of time trend to x_t.

Consider the following specification

$$x_t = \alpha + \delta t + u_t, \tag{4.1}$$

trend - stationary.

where u_t follows a zero-mean ARMA($p+1$, q), process: i.e.,

$$(1-\phi_1 L - \cdots - \phi_{p+1}L^{p+1})u_t = (1+\theta_1 L + \cdots + \theta_q L^q)\varepsilon_t,$$

where $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$.

- If u_t is stationary (i.e., all of the solutions of the AR characteristic equation are greater than 1 in modulus), then x_t follows a trend-stationary ARMA($p+1$, q) process. In this case, one applies the ARMA($p+1$, q) model to $u_t = x_t - \alpha - \delta t$:

$x_t - \alpha - \delta t \sim ARMA(p+1, q)$

$$(1-\phi_1 L - \cdots - \phi_{p+1}L^{p+1})(x_t - \alpha - \delta t) = (1+\theta_1 L + \cdots + \theta_q L^q)\varepsilon_t.$$

- What if u_t is nonstationary? In specific, what if one of the solutions of the AR characteristic equation is equal to 1 in modulus while other p solutions are greater than 1 in modulus? In this case, one applies the ARMA(p, q) model to Δx_t.

Theorem 4.1. *Consider $u_t \sim ARMA(p+1,q)$. If one of the solutions of the AR characteristic equation is a unit root and other p solutions are greater than 1 in modulus, then $\Delta u_t \sim ARMA(p,q)$ and Δu_t is stationary.*

Proof. One can factor $\rho(L) = 1 - \phi_1 L - \cdots - \phi_{p+1}L^{p+1}$ as

$x_t = \alpha + \delta t + u_t$

$$\rho(L) = \left(\frac{1}{\lambda_1} - L\right)\left(\frac{1}{\lambda_2} - L\right)\cdots\left(\frac{1}{\lambda_{p+1}} - L\right).$$

$u_t = x_t - \alpha - \delta t$
$u_{t-1} = x_{t-1} - \alpha - \delta(t-1)$
$= x_{t-1} - \alpha - \delta t + \delta$
$\Delta u_t = x_t - x_{t-1} - \delta$
$= \Delta x_t - \delta$
$\Delta x_t = \Delta u_t + \delta$

Let λ_{p+1} set equal to one. Then, it shows

$$\underbrace{\left(\frac{1}{\lambda_1} - L\right)\left(\frac{1}{\lambda_2} - L\right)\cdots\left(\frac{1}{\lambda_p} - L\right)}\underbrace{(1-L)u_t}_{\Delta u_t} = \theta(L)\varepsilon_t$$

or

$$\phi(L)\Delta u_t = \theta(L)\varepsilon_t.$$

Thus, $\Delta u_t \sim \text{ARMA}(p,q)$. □

If $u_t \sim \text{ARMA}(p+1,q)$ and the AR characteristic equation of u_t has a unit root, the first difference of (4.1), i.e., $\Delta x_t = \delta + \Delta u_t$, is stationary since Δu_t is stationary.

- Since $\Delta u_t = \Delta x_t - \delta$ follows an ARMA(p, q) process, Δx_t is a stationary ARMA(p, q) process with mean δ: i.e.,

$$\phi(L)(\Delta x_t - \delta) = \theta(L)\varepsilon_t,$$

where $\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p$ and $\theta(L) = 1 + \theta_1 L + \theta_2 L^2 + \cdots + \theta_q L^q$.

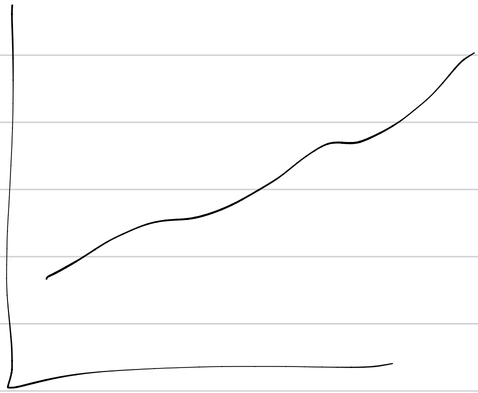
$$\phi(L)\cdot(1-L)u_t = \theta(L)\varepsilon_t$$

$$\phi(L)\cdot(u_t - u_{t-1}) = \theta(L)\varepsilon_t$$

$$\phi(L)\cdot\Delta u_t = \theta(L)\varepsilon_t$$

$\therefore$ $\Delta u_t \sim$ stationary ARMA(p,q).

$$\phi(L)\cdot(1-L)(1-L)u_t = \theta(L)\varepsilon_t$$
$$\phi(L)\cdot(1-L)(u_t - u_{t-1}) = \theta(L)\varepsilon_t$$
$$\phi(L)(u_t - u_{t-1} - (u_{t-1} - u_{t-2}))$$
$$\phi(L)(u_t - 2u_{t-1} + u_{t-2})$$
$$\phi(L)(\Delta u_t - \Delta u_{t-1}) = \theta(L)\varepsilon_t$$



1) $x_t - \alpha - \delta t = u_t \sim I(0)$

$\Rightarrow \phi(L)(x_t - \alpha - \delta t) = \theta(L) \cdot \varepsilon_t$

2) $u_t \sim ARMA(p+1, q)$

unit root case $\Rightarrow$ non-stationary

$\Rightarrow \Delta u_t \sim ARMA(p, q)$ & stationary

$x_t = \alpha + \delta t + u_t$

$\delta > 0$.

$\phi(L)(\Delta x_t - \delta) = \theta(L)\,\varepsilon_t$

$\Uparrow$

$\rightarrow \quad x_t - x_{t-1} = \alpha + \delta t + u_t - \alpha - \delta(t-1) - u_{t-1}$

$\Delta x_t = \delta + \Delta u_t$

$\Rightarrow \Delta x_t - \delta = \Delta u_t \sim ARMA(p, q)$

Definition 4.2. A unit-root nonstationary process is modeled with an ARMA model with a unit root in the AR characteristic equation. Such a model is called an *Autoregressive Integrated Moving Average* (ARIMA) model. Specifically, a time series x_t is said to follow ARIMA(p, 1, q) process if Δx_t follows a stationary and invertible ARMA(p, q) model:

$$\phi(L)(\Delta x_t - \delta) = \theta(L)\varepsilon_t,$$

where $\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p$, $\theta(L) = 1 + \theta_1 L + \theta_2 L^2 + \cdots + \theta_q L^q$, and $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$.

5. Unit-Root Tests

In practice, one needs to test the null hypothesis that a given nonstationary time series is a unit-root nonstationary time series. If the null hypothesis is not rejected, one can apply the ARIMA(p, 1, q) model to the nonstationary data. Otherwise, the nonstationary data should be analyzed with the ARMA(p, q) model augmented with a time trend.

random walk : $x_t = \delta + x_{t-1} + \varepsilon_t$
trend-stationary : $x_t = \beta_0 + \beta_1 t + u_t$
unit root nonstationary : $x_t = \delta + x_{t-1} + \psi \varepsilon_t$

5.1. Dickey-Fuller Unit-Root Test

Consider an AR(1) model of the form

$$x_t = c(t) + \phi_1 x_{t-1} + \varepsilon_t, \tag{5.1}$$

where $c(t)$ is a deterministic function of the time index t and $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$.

- If $\phi_1 = 1$, x_t follows a unit-root nonstationary process with the deterministic term $c(t)$.
- If $|\phi_1| < 1$, x_t follows a trend-stationary AR(1) process with the deterministic term $c(t)$. This is because

$$x_t = c(t) + (1 - \phi_1 L)^{-1}\varepsilon_t$$

is equivalent to

$$\begin{aligned} x_t &= c(t) + u_t \\ u_t &= \phi_1 u_{t-1} + \varepsilon_t. \end{aligned}$$

Theorem 5.1. *Consider $H_0: \phi_1 = 1$ versus $H_1: |\phi_1| < 1$ in (5.1). A test statistic is given by*

$$t = \frac{\hat{\phi}_1 - 1}{se(\hat{\phi}_1)},$$

where $\hat{\phi}_1$ is the OLS estimate of ϕ_1 and $se(\hat{\phi}_1)$ is its standard error. Dickey and Fuller (1979) show that the test statistic has a Dickey-Fuller (DF) distribution under H_0 and has an asymptotic standard normal distribution under H_1.

In testing unit-roots, it is important to specify $c(t)$, so that the null and alternative hypotheses appropriately reflect the trend properties of the observed data. Two cases emerge as follows:

$$x_t = \alpha + \delta t + u_t \qquad C_t$$

$$u_t = \phi_1 x_{t-1} + \varepsilon_t$$

$$c(t) = \alpha$$

$$c(t) = \alpha + \delta t$$

$$u_t = \phi_1 u_{t-1} + \varepsilon_t .$$

$$x_t - \phi_1 x_{t-1} = c(t) + \varepsilon_t$$

$$x_t(1 - \phi_1 L) = c(t) + \varepsilon_t$$

$$x_t = \frac{c(t)}{1-\phi_1 L} + \frac{\varepsilon_t}{1-\phi_1 L}$$

$c(t)$

$$c(t) = \alpha + \delta t$$

$$\varepsilon_t = (1-\phi_1 L) u_t$$

$$L \Rightarrow x_t \to x_{t-1}$$

$$u_t = \phi_1 u_{t-1} + \varepsilon_t$$

$$L x_t = x_{t-1}$$

$\Rightarrow u_t \sim AR(1)$ & stat

$$|\phi_1| < 1$$

$H_0: \phi_1 = 1 \quad \Rightarrow$ unit-root nonstationary

$H_1: |\phi_1| < 1. \quad \Rightarrow$ trend-stationary AR(1)

random walk $\quad x_t = \delta + x_{t-1} + \varepsilon_t$

trend-stationary $\quad x_t = \beta_0 + \beta_1 t + u_t$

unit-root nonstationary $\quad x_t = \delta + x_{t-1} + \varphi \varepsilon_t$

1. Case1 (constant only): With $c(t) = \alpha$, the test regression is

$$x_t = \alpha + \phi_1 x_{t-1} + \varepsilon_t.$$

Under $H_0 : \phi_1 = 1$, x_t is a random walk with drift α. Under $H_1 : |\phi_1| < 1$, x_t is an AR(1) process.

2. Case 2 (constant and time trend): With $c(t) = \alpha + \delta t$, the test regression is

$$x_t = \alpha + \delta t + \phi_1 x_{t-1} + \varepsilon_t.$$

Under $H_0 : \phi_1 = 1$, x_t is a random walk with drift α and time trend δ. Under $H_1 : |\phi_1| < 1$, x_t is an AR(1) process with deterministic $\alpha + \delta t$.

5.2. *Augmented Dickey-Fuller Unit-Root Test*

Contrary to the simple AR(1) assumption of the Dickey-Fuller test, a typical time series has a complicated dynamic structure. Said and Dickey (1984) augment the basic Dickey-Fuller test to accommodate general AR(p) models. Their test is referred to as the *augmented Dickey-Fuller* (ADF) test.

Theorem 5.2. *(Dickey-Fuller transformation) An AR(p) model*

$$x_t = \phi_0 + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \phi_p x_{t-p} + \varepsilon_t$$

can be written as

$$x_t = \phi_0 + \rho x_{t-1} + \psi_1 \Delta x_{t-1} + \psi_2 \Delta x_{t-2} + \cdots + \psi_{p-1} \Delta x_{t-p+1} + \varepsilon_t,$$

where $\rho = \phi_1 + \cdots + \phi_p$ and ψ_j is a linear combination of ϕ_is.

- In the Dickey-Fuller transformation, the unit-root test is equivalent to check whether or not ρ is equal to one.

Example 5.3. The Dickey-Fuller transformation of AR(2) is

$$\begin{aligned} x_t &= \phi_0 + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \varepsilon_t \\ &= \phi_0 + \phi_1 x_{t-1} + \phi_2 x_{t-1} - \phi_2 x_{t-1} + \phi_2 x_{t-2} + \varepsilon_t \\ &= \phi_0 + (\phi_1 + \phi_2) x_{t-1} - \phi_2 (x_{t-1} - x_{t-2}) + \varepsilon_t \\ &= \phi_0 + \rho x_{t-1} + \psi_1 \Delta x_{t-1} + \varepsilon_t. \end{aligned}$$

$\psi_1 = \phi_2 = 0 \cdot \phi_1 + 1 \cdot \phi_2$

$= [\phi_1\ \phi_2]\begin{bmatrix}0\\1\end{bmatrix}$

$= \rho' \phi$

The ADF test is based on estimating a test regression of the form

$$x_t = c(t) + \rho x_{t-1} + \sum_{j=1}^{p-1} \psi_j \Delta x_{t-j} + \varepsilon_t, \tag{5.2}$$

where $c(t) = \alpha$ (i.e., constant only) or $c(t) = \alpha + \delta t$ (i.e., constant and time trend).

- Alternatively, the ADF test regression in (5.2) can be stated as

trend-stationary process

$$\Delta x_t = c(t) + \pi x_{t-1} + \sum_{j=1}^{p-1} \psi_j \Delta x_{t-j} + \varepsilon_t, \qquad (5.3)$$

H_1 : skip
complicate.

  where $\pi = \rho - 1$. In this alternative formulation, $H_0 : \pi = 0$ is equivalent to $H_0 : \rho = 1$, and the ADF t-statistic for testing a unit root is the usual t-statistic for testing $\pi = 0$. Due to this convenience, the test regression (5.3) is often used in practice and R follows this.

$\pi = 0 \Rightarrow$ unit root nonstationary.

In (5.2) and (5.3), the p lagged difference terms Δx_{t-j} approximate the AR(p) structure of the errors, and the value of p should be set so that ε_t is serially uncorrelated.

- If p is too small, then the remaining serial correlation in the errors will bias the test. If p is too large, then the power of test (i.e., the probability that H_0 is rejected when H_1 is true) will suffer.

- Ng and Perron (1995) suggest the *data dependent lag length selection procedure* for choosing p for the ADF test. The Ng and Perron procedure consists of the following two steps:

  1. Step 1: Set an upper bound p_{max} for p. To determine p_{max}, Schwert (1989) suggests

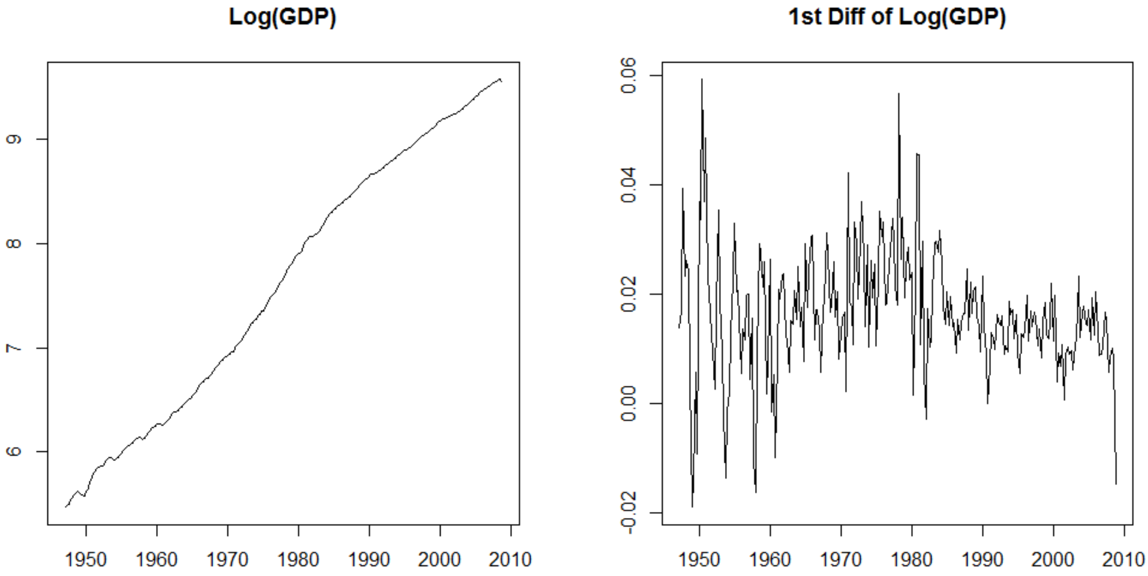
  $$p_{max} = \left[12 \cdot \left(\frac{T}{100} \right)^{1/4} \right],$$

     where $[x]$ denotes the integer part of x.

  2. Step 2: Estimate the ADF test regression with $p = p_{max}$. If the absolute value of the t-ratio of the the last lagged difference Δx_{t-p-1} is greater than 1.6, then set $p = p_{max}$ and perform the unit-root test. Otherwise, reduce the lag length by one and repeat the process .

Example 5.4. Consider the log series of US quarterly GDP from 1Q 1947 to 4Q 2008.

```
> mydat <- read.table("data1.txt", header = T)
> head(mydat)
  year mon day   gdp
1 1947   1   1 237.2
...
6 1948   4   1 267.3
> gdp <- ts(log(mydat$gdp), start = c(1947, 1), freq = 4)
> gdp.d <- diff(gdp)
> par(mfrow = c(1, 2))
> plot(gdp, xlab = "", ylab = "", main = "Log(GDP)")
> plot(gdp.d, xlab = "", ylab = "", main = "1st Diff of Log(GDP)")
```




```
> library(urca)
> pmax <- floor(12*(length(gdp)/100)^(1/4))
> output <- ur.df(gdp, type = "trend", lags = pmax)
> summary(output)
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.792e-02  2.825e-02   0.988   0.3242
z.lag.1     -4.300e-03  5.443e-03  -0.790   0.4304
tt           6.969e-05  9.896e-05   0.704   0.4821
z.diff.lag1  3.617e-01  6.952e-02   5.202  4.6e-07
...
z.diff.lag15 3.750e-02  6.422e-02   0.584   0.5599
> output <- ur.df(gdp, type = "trend", lags = c(pmax - 1))
> summary(output)
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.491e-02  2.796e-02   0.891   0.3739
z.lag.1     -3.644e-03  5.374e-03  -0.678   0.4984
tt           5.701e-05  9.763e-05   0.584   0.5598
z.diff.lag1  3.780e-01  6.713e-02   5.631  5.55e-08
```

>1.6

```
...
z.diff.lag14 1.329e-01  6.340e-02   2.096   0.0372
```

stop

- The absolute value of the t-ratio of the last lagged difference Δx_{t-14} (i.e., `z.diff.lag14`) is greater than 1.6, so $p = 14$.

prope p would be 14

unit-root or trend stationary ?

```
> library(fUnitRoots)
> adfTest(gdp, lags = c(pmax - 1), type = "ct")
Title:
 Augmented Dickey-Fuller Test
Test Results:
 PARAMETER:
  Lag Order: 14
 STATISTIC:
  Dickey-Fuller: -0.6781
 P VALUE:
  0.9716
```

by ADF

unit-root

- The unit-root hypothesis cannot be rejected at the 5% level.

```
> reg <- arima(gdp, order = c(0, 1, 1), method = "ML")
> reg$aic
[1] -1362.158
...
> reg <- arima(gdp, order = c(2, 1, 2), method = "ML")
> reg$aic
[1] -1541.513
```

- Using the AIC, one chooses the ARIMA(1, 1, 2) model for x_t.

minimum value of aic, ⇒ proper

```
> (reg <- arima(diff(gdp), order = c(1, 0, 2), method = "ML"))
Call:
arima(x = diff(gdp), order = c(1, 0, 2), method = "ML")
Coefficients:
         ar1     ma1     ma2  intercept
      0.4062  0.0113  0.1910     0.0164
s.e.  0.1407  0.1431  0.0742     0.0013
sigma^2 estimated as 9.467e-05: log likelihood = 793.6, aic = -1577.21
```

ARIMA 1 0 2 for Δx_t

ARMA(1,2) for Δx_t

- Finally, the fitted model is

$$(1-0.4062L)(\Delta x_t - 0.0164) = (1+0.0113L+0.1910L^2)\varepsilon_t.$$

1) L 이용

2) first diff. : unit root nonstationary.