Financial Econometrics

Chapter 2: Linear Time Series Analysis

In Choi

Sogang University

Let $\{r_t\}$ be a time series for $t = 1, 2...$

- The mean function of $\{r_t\}$ is

$$\mu(t) = E(r_t).$$

- The autocovariance function of $\{r_t\}$ is

$$\gamma(t, s) = Cov(r_t, r_s) = E[(r_t - \mu(t))(r_s - \mu(s)].$$

- $\{r_t\}$ is weakly (second-order) stationary if

(i) $\mu(t)$ is a constant.
(ii) $\gamma(t, t - l)$ is independent of t for each l.

- Stationary processes vary around a fixed level within a finite range.
- The first two moments of future r_t are the same as those of the past.
- For a stationary process $\{r_t\}$, we may write $\gamma(t, t-l) = \gamma(l)$.

- Basic properties of $\gamma(\cdot)$ of a stationary process are:

$$(i)\ \gamma(0) \geq 0$$

$$(ii)\ |\gamma(l)| \leq \gamma(0) \text{ for all } l$$

$$(iii)\ \gamma(l) = \gamma(-l).$$

- The autocorrelation function of $\{r_t\}$is

$$\rho(l) = \frac{\gamma(l)}{\gamma(0)} = Corr(r_t, r_{t-l}), 0 \leq l < T - 1.$$

- For all l, $|\rho(l)| \leq 1$ and $\rho(l) = \rho(-l)$.

- Let $\{r_t\}_{t=1}^{T}$ be observations on a time series.

(i) Sample mean

$$\bar{r} = \frac{1}{T}\sum_{t=1}^{T} r_t.$$

This estimates μ.

(ii) Sample autocovariance function

$$\hat{\gamma}(l) = \frac{1}{T}\sum_{t=l+1}^{T} (r_t - \bar{r})(r_{t-l} - \bar{r}).$$

This estimates $\gamma(l)$.

(iii) Sample autocorrelation function

$$\hat{\rho}(l) = \hat{\gamma}(l)/\hat{\gamma}(0).$$

- For $H_o : \rho(l) = 0$, use the test statistic

$$\frac{\hat{\rho}(l)}{\sqrt{\left(1 + 2\sum_{i=1}^{l-1} \hat{\rho}(i)^2\right) / T}}.$$

When T is large, its distribution is standard normal.

- For $H_o : \rho(1) = 0$, use
$$\frac{\hat{\rho}(1)}{\sqrt{1/T}} \simeq N(0,1).$$
Reject H_o at the 5% level if $\left|\frac{\hat{\rho}(1)}{\sqrt{1/T}}\right| > 1.96$.
- For $H_o : \rho(1) = ... = \rho(m) = 0$, use the Ljung-Box statistic
$$Q(m) = T(T+2)\sum_{l=1}^{m}\frac{\hat{\rho}(l)^2}{T-l} \simeq \chi^2(m).$$
One needs to choose m in practice.

- A stochastic process $\{r_t\}$ is a white noise process if
  (i) $E(r_t) = 0$,
  (ii) $Var(r_t) = \sigma^2$
  (iii) $E(r_t r_{t-l}) = 0 \ (l \neq 0)$.
- The white noise process is stationary. We write

$$r_t \sim WN(0, \sigma^2).$$

The time series $\{r_t\}$ is a linear process if it has the representation

$$r_t = \sum_{j=-\infty}^{\infty} \psi_j a_{t-j}$$

for all t, where $\{a_t\}$ is a white noise process with variance σ^2 and $\{\psi_j\}$ is a sequence of constants with $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$. ($\psi$: sigh)

- Using the backward shift operator B, r_t can be written as

$$r_t = \psi(B) a_t$$

where $\psi(B) = \sum_{j=-\infty}^{\infty} \psi_j B^j$ and $B^j a_t = a_{t-j}$.

- $\{r_t\}$ is a moving average process of order q if

$$r_t = a_t + \theta_1 a_{t-1} + ... + \theta_q a_{t-q}$$

where $\{a_t\} \sim WN(0, \sigma^2)$ and $\theta_1, ..., \theta_q$ are constants.

Properties

$$(i)\ E(r_t) = 0.$$

$$\begin{aligned}(ii)\ \gamma(l) &= E\left(\sum_{j=-\infty}^{\infty}\psi_j a_{t-j}\right)\left(\sum_{j=-\infty}^{\infty}\psi_j a_{t-l-j}\right)\\ &= E\left(\sum_{i,j=-\infty}^{\infty}\psi_i\psi_j a_{t-i}a_{t-l-j}\right)\\ &= \sum_{j=-\infty}^{\infty}\psi_{j+l}\psi_j E\left(a_{t-l-j}^2\right)\\ &= \sigma^2\sum_{j=-\infty}^{\infty}\psi_{j+l}\psi_j\end{aligned}$$

Thus, linear processes are weakly stationary.

Note that by the Cauchy-Schwarz inequality

$$\left|\sum_{j=-\infty}^{\infty} \psi_{j+l}\psi_j\right| \leq \sqrt{\sum_{j=-\infty}^{\infty} \psi_{j+l}^2 \sum_{j=-\infty}^{\infty} \psi_j^2} < \infty \; .$$

The second inequality follows because $\sum_{j=-\infty}^{\infty} \psi_j^2 \leq \left(\sum_{j=-\infty}^{\infty} \left|\psi_j\right|\right)^2 < \infty$.

Autoregressive models

- $\{r_t\}$ is an AR(p) process if for every t

$$r_t = \phi_1 r_{t-1} + \cdots + \phi_p r_{t-p} + a_t$$

  where $a_t \sim WN(0, \sigma^2)$, $\phi_p \neq 0$. (ϕ : fee)
- If $\{r_t\}$ has a non-zero mean, we use the model

$$r_t = \phi_0 + \phi_1 r_{t-1} + \cdots + \phi_p r_{t-p} + a_t$$

Autoregressive models

Example Let r_t be the number of new BMWs that are repaired in year t during their 2 year warranty periods. Suppose that approximately 10% of the cars repaired a year ago come back for repair. Then, r_t can be modelled as

$$r_t = \mu + 0.1 r_{t-1} + a_t.$$

Here a_t denotes the number of cars produced and repaired in year t.

Autoregressive models

- Consider the AR(1) model

$$r_t = \phi_0 + \phi_1 r_{t-1} + a_t.$$

If $|\phi_1| < 1$, this process can be written as

$$r_t = \frac{\phi_0}{1-\phi_1} + \sum_{j=0}^{\infty} \phi_1^j a_{t-j}.$$

Thus, it is weakly stationary if $|\phi_1| < 1$.

Autoregressive models

- The AR(1) process has mean and varaince

$$E(r_t) = \frac{\phi_0}{1 - \phi_1}$$

and

$$Var(r_t) = \frac{\sigma^2}{1 - \phi_1^2},$$

respectively. In addition, $\rho(k) = \phi_1^k$.

Autoregressive models

- Consider the AR(1) model

$$r_t = \phi_1 r_{t-1} + a_t; \ \phi_1 = 1, r_0 = 0.$$

Then,

$$r_t = a_1 + ... + a_t.$$

Since $Var(r_t) = t\sigma^2$, r_t is not stationary. It displays growing variance.

Autoregressive models

- For an AR(p) process $\{r_t\}$,

$$r_t = \phi_1 r_{t-1} + \cdots + \phi_p r_{t-p} + a_t$$

  consider the characteristic equation $1 - \phi_1 z - ... - \phi_p z^p = 0$. If all the roots of this equation is greater than one, the process is stationary. (For a proof, see chapter 2 of W. Fuller (1996).)
- Equivalently, if $\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p \neq 0$ for all $|z| \leq 1$, the AR process is stationary.

Moving average model of order 1 and invertibility

- The model for observation $\{r_t\}$

$$r_t = a_t + \theta a_{t-1}, \ a_t \sim WN(0, \sigma^2) \text{ for every } t$$

is called the moving average (MA) model of order 1.

- The model can be rewritten as

$$a_t = (1+\theta B)^{-1} r_t = (1 - \theta B + \ldots + (-\theta)^k B^k)(1-(-\theta)^{k+1} B^{k+1})^{-1} r_t,$$

which gives

$$r_t = \theta r_{t-1} - \theta^2 r_{t-2} + \ldots - (-\theta)^k r_{t-k} + a_t - (-\theta)^{k+1} a_{t-k-1}.$$

Moving average model of order 1 and invertibility

- If $|\theta| < 1$, we obtain the infinite series

$$r_t = \theta r_{t-1} - \theta^2 r_{t-2} + ... + a_t.$$

- If $|\theta| \geq 1$, r_t depends on $r_{t-1}, r_{t-2}, ..., r_{t-k}$ with weights that increase with k.We avoid this situation by requiring that $|\theta| < 1$.
- If $|\theta| < 1$, we say that the MA process is invertible. When the MA(1) process is invertible, it can be expressed as an $AR(\infty)$ process properly.

ARMA(1,1) model

- The time series r_t is an $ARMA(1,1)$ process if it satisfies

$$r_t = \phi r_{t-1} + a_t + \theta a_{t-1}, \; a_t \sim WN(0, \sigma^2) \text{ for every } t.$$

- The $ARMA(1,1)$ process can be written more compactly as

$$\phi(B) r_t = \theta(B) a_t$$

where $\phi(B) = 1 - \phi B$ and $\theta(B) = 1 + \theta B$.

- If $\phi + \theta = 0$, $r_t = a_t$.

ARMA(1,1) model

- Suppose that $|\phi| < 1$. Then,

$$
\begin{aligned}
r_t &= \phi r_{t-1} + a_t + \theta a_{t-1} \\
\phi r_{t-1} &= \phi^2 r_{t-2} + \phi a_{t-1} + \phi\theta a_{t-2} \\
\phi^2 r_{t-2} &= \phi^3 r_{t-3} + \phi^2 a_{t-2} + \phi^2\theta a_{t-3} \\
&\vdots
\end{aligned}
$$

Adding all these equations, we obtain

$$
\begin{aligned}
r_t &= a_t + \phi a_{t-1} + \phi^2 a_{t-2} + \ldots + \theta a_{t-1} + \phi\theta a_{t-2} + \phi^2\theta a_{t-3} + \ldots \\
&= a_t + (\phi + \theta) \sum_{j=1}^{\infty} \phi^{j-1} a_{t-j}. \qquad (1)
\end{aligned}
$$

When $|\phi| < 1, \sum_{j=1}^{\infty} |\phi|^{j-1} = \frac{1}{1-|\phi|} < \infty$, and hence r_t is stationary.

ARMA(1,1) model

- If $|\phi| = 1$, $\{r_t\}$ is non-stationary.
- Write

$$a_t = -\theta a_{t-1} + r_t - \phi r_{t-1}$$

If $|\theta| < 1$,

$$a_t = r_t - (\phi + \theta) \sum_{j=1}^{\infty} (-\theta)^{j-1} r_{t-j}$$

The $ARMA(1,1)$ process in this case is said to be invertible since a_t can be expressed in terms of the present and past values of the process, $r_s, s \leq t$.

ARMA(1,1) model

Or we may write

$$r_t = a_t - (\phi + \theta) \sum_{j=1}^{\infty} (-\theta)^{j-1} r_{t-j}$$

which shows that r_t is a proper linear combination of a_t and the past observations $t_{t-1}, r_{t-2}, \ldots$

ARMA(1,1) model

- When $|\theta| \geq 1$, the *ARMA*(1,1) process is said to be noninvertible.

ARMA(p,q) model

- Why autoregressive moving average (*ARMA*) models?

1. Combination of AR and MA models
2. Parsimonious (not too many parameters): Recall that MA(1) model is AR(∞)
3. If $X_t \sim ARMA(p_1, q_1)$ and $Y_t \sim ARMA(p_2, q_2)$, $X_t + Y_t \sim ARMA(P, Q)$ where $P = p_1 + p_2$ and $Q = \max(p_1 + q_2, p_2 + q_1)$.
   For example, if $X_t \sim AR(1)$ and $Y_t \sim AR(1)$, $X_t + Y_t \sim ARMA(2, 1)$. An aggregation of AR processes results in an *ARMA* process.

ARMA(p,q) model

Example

Let a_t be a number of new, overnight patients that arrive on day t and assume that it is a white noise process. Typically 10% stay just one day, 50% two days, 30% three days and 10% four days. If r_t is the number of patients leaving the hospital on day t, we may model it as

$$r_t = \mu + 0.1a_{t-1} + 0.5a_{t-2} + 0.3a_{t-3} + 0.1a_{t-4}.$$

ARMA(p,q) model

- $\{r_t\}$ is an ARMA(p,q) process if for every t

$$r_t - \phi_1 r_{t-1} - \cdots - \phi_p r_{t-p} = a_t + \theta_1 a_{t-1} + \cdots + \theta_q a_{t-q}$$

where $a_t \sim WN(0, \sigma^2)$, $\phi_p \neq 0$, $\theta_q \neq 0$ and the polynomials

$$\left(1 - \phi_1 z - \cdots \phi_p z^p\right)$$

and

$$(1 + \theta_1 z + \cdots \theta_q z^q)$$

have no common factors.

ARMA(p,q) model

- $$\left(1-4z\right)\left(1-5z\right)$$

  and

  $$\left(1-z\right)\left(1+2z\right)$$

  have no common factors. But

  $$\left(1-4z\right)\left(1-5z\right)$$

  and

  $$\left(1-4z\right)\left(1-6z\right)$$

  have the common factor $\left(1-4z\right)$.

ARMA(p,q) model

- The requirement of no common factor is to ensure that there are no redundant parameters in the model. For example, if

$$r_t - 0.5r_{t-1} = a_t - 0.5a_{t-1},$$

it is better to write

$$r_t = a_t.$$

ARMA(p,q) model

- The *ARMA* model can also be written as

$$\phi(B)\, r_t = \theta(B)\, a_t$$

where

$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p$$

and

$$\theta(z) = 1 + \theta_1 z + \cdots + \theta_q z^q$$

where

$$B^j r_t = r_{t-j} \text{ and } B^j a_t = a_{t-j}.$$

ARMA(p,q) model

- **A useful fact:** Let $\{Y_t\}$ be a weakly stationary time series with zero mean. If

$$\sum_{j=-\infty}^{\infty} |\psi_j| < \infty,$$

the time series

$$X_t = \sum_{j=-\infty}^{\infty} \psi_j Y_{t-j}$$

is also weakly stationary with zero mean.

ARMA(p,q) model

Example

Consider the $ARMA(2, q)$

$$\begin{aligned}\left(1-\phi_1 B-\phi_2 B^2\right) r_t &= \theta(B) a_t \\ &= u_t .\end{aligned}$$

Suppose that

$$\left(1-\phi_1 z-\phi_2 z^2\right)=(1-\alpha_1 z)(1-\alpha_2 z)$$

where $|\alpha_1| < 1$ and $|\alpha_2| < 1$ or equivalently,

$$1-\phi_1 z-\phi_2 z^2 \neq 0 \text{ for } |z| \leq 1.$$

Let

$$(1-\alpha_2 B) r_t = W_t .$$

Example (continued)

Then

$$\left(1 - \phi_1 B - \phi_2 B^2\right) r_t = (1 - \alpha_1 B) W_t = u_t$$

Because u_t is stationary and $|\alpha_1| < 1$, W_t is stationary. We may write

$$r_t - \alpha_2 r_{t-1} = W_t,$$

where W_t is stationary. Since $|\alpha_2| < 1$, r_t is stationary.

ARMA(p,q) model

- An $ARMA(p,q)$ process $\{r_t\}$ is stationary if

$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p \neq 0 \text{ for all } |z| \leq 1.$$

ARMA(p,q) model

- An $ARMA(p, q)$ process $\{r_t\}$ is said to be invertible if there exist constants $\{\pi_j\}$ such that

$$\sum_{j=0}^{\infty} |\pi_j| < \infty$$

and

$$a_t = \sum_{j=1}^{\infty} \pi_j r_{t-j} \text{ for all } t.$$

- The coefficients $\{\pi_j\}$ are determined by the relation

$$\pi(z) = \sum_{j=0}^{\infty} \pi_j z^j = \phi(z)/\theta(z).$$

ARMA(p,q) model

- Invertibility is equivalent to the condition:

$$\theta(z) = 1 + \theta_1 z + \cdots + \theta_q z^q \neq 0 \text{ for all } |z| \leq 1.$$

ARMA(p,q) model

Example

If

$$r_t - 0.5r_{t-1} = a_t + 0.4a_{t-1}, a_t \sim WN(0, \sigma^2),$$

$$\phi(z) = 1 - 0.5z = 0 \Rightarrow z = 2$$

$$\theta(z) = 1 + 0.4z = 0 \Rightarrow z = -\frac{5}{2}.$$

r_t is stationary and invertible.

ARMA(p,q) model

Example

Let

$$r_t = a_t - a_{t-1}$$

$$\theta\left(z\right) = 1 - z = 0 \Rightarrow z = 1$$

r_t is not invertible.

Example

$$\left(1 - B\right)\left(1 - 0.5B\right) r_t = a_t$$

$$\phi\left(z\right) = \left(1 - z\right)\left(1 - 0.5z\right) = 0 \Rightarrow z = 1, 2$$

So r_t is not stationary.

The ACF and PACF of an ARMA(p,q) process

- Methods for calculating autocovariance function (ACF)

$$\phi(B)\, r_t = \theta(B)\, a_t$$

1. Use the linear process representation of r_t.
2. Multiply each side of the equation

$$r_t - \phi_1 r_{t-1} - \cdots \phi_p r_{t-p} = a_t + \theta_1 a_{t-1} + \cdots + \theta_q a_{t-q}$$

by r_{t-h} $(h = 0, 1, \cdots)$ and take expectation. This provides a difference equation for $\gamma(\cdot)$.

The ACF and PACF of an ARMA(p,q) process

Example

The $ARMA(1,1)$ process

$$\begin{aligned} r_t - \phi r_{t-1} &= a_t + \theta a_{t-1}, a_t \sim WN(0, \sigma^2) \\ &\Rightarrow \quad r_t = a_t + (\phi + \theta) \sum_{j=1}^{\infty} \phi^{j-1} a_{t-j}. \end{aligned}$$

$$\begin{aligned} E(r_t^2) - \phi E(r_t r_{t-1}) &= E\left(r_t (a_t + \theta a_{t-1})\right) \\ &\text{or} \\ \gamma(0) - \phi\gamma(1) &= \sigma^2 \left(1 + \theta(\phi + \theta)\right) \end{aligned} \tag{2}$$

$$\begin{aligned} E(r_{t-1} r_t) - \phi E\left(r_{t-1}^2\right) &= \sigma^2 \theta \\ &\text{or} \\ \gamma(1) - \phi\gamma(0) &= \sigma^2 \theta \end{aligned} \tag{3}$$

Example (Continued)

$$
\begin{aligned}
E(r_{t-h}r_t) - \phi E(r_{t-h}r_t) &= 0 \qquad \text{for } h \geq 2 \\
&\text{or} \\
\gamma(h) - \phi\gamma(h-1) &= 0 \qquad \text{for } h \geq 2
\end{aligned} \tag{4}
$$

Solving (2) and (3), we obtain

$$
\begin{aligned}
\gamma(0) &= \frac{\sigma^2[1 + 2\theta\phi + \theta^2]}{1-\phi^2} \\
\gamma(1) &= \sigma^2\left[\theta + \frac{\phi(1 + 2\theta\phi + \theta^2)}{1-\phi^2}\right] \\
\gamma(h) &= \phi^{h-1}\gamma(1), \;\; h \geq 2.
\end{aligned}
$$

The ACF and PACF of an ARMA(p,q) process

- Suppose that we wish to estimate the correlation between r_t and r_{t+h} excluding the effects of the intervening variables $r_{t+1}, \cdots r_{t+h-1}$. The estimate of this is called the partial autocorrelation between r_t and r_{t+h}. We denote this as ω_h.

The ACF and PACF of an ARMA(p,q) process

- Consider the OLS regression

$$r_t = \hat{\alpha}_1 r_{t+1} + \cdots + \hat{\alpha}_{h-1} r_{t+h-1} + \hat{\alpha}_h r_{t+h} + \hat{u}_t.$$

The partial autocorrelation ω_h is approximately equal to $\hat{\alpha}_h$ in large samples.

The ACF and PACF of an ARMA(p,q) process

- More intuitively, consider the two regressions

$$r_t = \hat{\beta}_1 r_{t+1} + \cdots + \hat{\beta}_{h-1} r_{t+h-1} + \hat{r}_t$$

and

$$r_{t+h} = \hat{\zeta}_1 r_{t+1} + \cdots + \hat{\zeta}_{h-1} r_{t+h-1} + \hat{r}_{t+h}.$$

In large samples, the OLS regression coefficient from regressing $\hat{r}_t$ on $\hat{r}_{t+h}$ is exactly equal to $\hat{\alpha}_h$.

The ACF and PACF of an ARMA(p,q) process

- For the $AR(p)$ process

$$r_t - \phi_1 r_{t-1} - \cdots \phi_p r_{t-p} = a_t,$$

$$\omega_p = \phi_p \text{ and } \omega_h = 0 \text{ for } h > p.$$

Thus, *PACF* is used for the *AR* order selection.

Estimation of AR and ARMA models

- For an AR(p) process $\{r_t\}$,

$$r_t = \phi_1 r_{t-1} + \cdots + \phi_p r_{t-p} + a_t,$$

use OLS for the estimation of $\phi_1, \cdots, \phi_p$. When r_t is stationary, the OLS estimators can be used as in standard linear regression.

Estimation of AR and ARMA models

- For an $ARMA(p,q)$ process $\{r_t\}$,

$$r_t = \phi_1 r_{t-1} + \cdots + \phi_p r_{t-p} + a_t + \theta_1 a_{t-1} + \ldots + \theta_q a_{t-q},$$

use nonlinear least squares that minimizes $\sum_{t=1}^{T} a_t^2$ with respect to the unknown coefficients. When r_t is stationary and invertible, the nonlinear least squares can be used as in standard linear regression.

Forecasting

.

- For an AR(p) process $\{r_t\}$,

$$r_t = \phi_1 r_{t-1} + \cdots + \phi_p r_{t-p} + a_t, (t = 1, ..., T),$$

1-step ahead forecast at time T is

$$\hat{r}_{T+1} = \phi_1 r_T + \cdots + \phi_p r_{T+1-p}.$$

In practice, we use the OLS estimators of $\phi_1, \cdots, \phi_p$.

Forecasting

- 1-step ahead forecast error is

$$e_T(1) = r_{T+1} - \hat{r}_{T+1} = a_{T+1}.$$

a_{T+1} is the unpredictable part of r_{T+1}. Moreover,

$$Var(e_T(1)) = \sigma^2.$$

Forecasting

- 2-step ahead forecast is

$$\hat{r}_{T+2} = \phi_1 \hat{r}_{T+1} + \cdots + \phi_p r_{T+2-p}.$$

Its forecast error is

$$\begin{aligned} e_T(1) &= r_{T+2} - \hat{r}_{T+2} \\ &= \phi_1 r_{T+1} + \cdots + \phi_p r_{T+2-p} + a_{T+2} \\ &\quad - \left(\phi_1 \hat{r}_{T+1} + \cdots + \phi_p r_{T+2-p}\right) \\ &= \phi_1 (r_{T+1} - \hat{r}_{T+1}) + a_{T+2} \\ &= \phi_1 a_{T+1} + a_{T+2}. \end{aligned}$$

Its variance is $(1 + \phi_1^2)\sigma^2$.

Model selection

- Choose a model which minimizes

$$AIC\,(p,q) = \ln \frac{\sum_{t=1}^{T} \hat{a}_t^2}{T} + \frac{2(p+q)}{T} \text{ (Akaike's information criteria)}$$

or

$$BIC\,(p,q) = \ln \frac{\sum_{t=1}^{T} \hat{a}_t^2}{T} + \frac{(p+q)\ln T}{T} \text{ (Bayesian information criteria)}$$

Choose a model that minimize the value of an information criterion.

Model selection

- The first term indicates how well the selected model fits the data. The smaller it is, the better fit we observe. It tends to become smaller as we have more variables in the model.
- The second term is a penalty term that prevents selecting too large a model to obtain a good fit.

Model selection

- Model selections based on information criteria seek a balance between model fit and size of the model.

Trend and seasonality

- Many time series data contain seasonal and/or trend component.

Example

Number of accidents, visitors to Korea.

- Classical decomposition

$$Y_t = TR_t + S_t + X_t$$

Y_t: observed time series
TR_t: trend component
S_t: seasonal component
X_t : random component

Trend and seasonality

- Linear trend model
  - $$TR_t = \alpha + \beta t$$

    α and β are unknown coefficients that can be estimated by the least squares method. We call t linear time trend.
  - Suppose that there is no seasonal component. If

    $$\ln(Y_t) = \alpha + \beta t + X_t,$$

    β denotes the growth rate of Y_t.

Trend and seasonality

- Quadratic trend

$$TR_t = \alpha + \beta t + \gamma t^2$$

Trend and seasonality

- Smoothing data
  - Purpose: discern trend element of the series without specifying the model for the trend element
  - Moving average filter

$$\text{Two-sided} \quad : \quad \tilde{Y}_t = (2m+1)^{-1} \sum_{j=-m}^{m} Y_{t-j}$$

$$\text{One-sided} \quad : \quad \tilde{Y}_t = (m+1)^{-1} \sum_{j=0}^{m} Y_{t-j}$$

  - Exponential moving averages

$$\tilde{Y}_t = \sum_{j=0}^{m} \alpha(1-\alpha)^j Y_{t-j}$$

Trend and seasonality

- Linear time trend is eliminated by differencing

$$\Delta Y_t = Y_t - Y_{t-1}$$

For example, if $Y_t = \beta_0 + \beta_1 t + X_t$, $\Delta Y_t = \beta_1 + \Delta X_t$. Thus ΔY_t has no trend. But analyzing Y_t and ΔY_t sometimes serves different purposes. For example, if Y_t denotes log GDP, ΔY_t is the GDP growth rate.

Trend and seasonality

- Seasonal elements may change over time due to random changes (e.g., weather and housing starts), variations in the calendar (e.g., Lunar New year) and factors related to economic decisions (e.g., e-commerce and retail sale).

Trend and seasonality

- Estimating seasonal component assuming it does not change over time
  - Regress Y_t on $\{D_{1t}, D_{2t}, ..., D_{d,t}\}$ where d is the number of seasons and

    $$D_{i,t} = \begin{cases} 1 & \text{if } t \text{ corresponds to season } i \\ 0 & \text{otherwise} \end{cases},$$

    and obtain

    $$Y_t = \hat{\alpha}_1 D_{1t} + ... + \hat{\alpha}_d D_{d,t} + \hat{Y}_t.$$

    Here, $\hat{Y}_t$ is the deseasonalized time series. Notice that we use no intercept to avoid the problem of multicollinearity.

Trend and seasonality

- There are two methods.
  - If there is also a linear trend element in the series, regress Y_t on $\{t, D_{1t}, D_{2t}, ..., D_{d,t}\}$ and obtain

$$Y_t = \hat{a}_1 t + \hat{\alpha}_1 D_{1t} + ... + \hat{\alpha}_d D_{d,t} + \hat{Y}_t.$$

    Here, $\hat{Y}_t$ is the detrended and deseasonalized time series. Notice that we use no intercept to avoid the problem of multicollinearity.
  - Seasonal differencing

$$\Delta_d Y_t = Y_t - Y_{t-d}$$

    If

$$Y_t = S_t + X_t$$

    with $S_t = S_{t+d}$,

$$\Delta_d Y_t = X_t - X_{t-d}.$$

Trend and seasonality

- X-12-ARIMA method
  - An official program for seasonal adjustment made by the US Census Bureau
  - Regressors that account for shifts in the mean, outliers, holiday effects, and the residuals are modelled by the seasonal ARIMA model.

Autoregressive integrated moving average (ARIMA) model

- Popularized by Box and Jenkins (1976).
- If d is nonnegative integer, $\{X_t\}$ is an $ARIMA(p, d, q)$ process if

$$r_t = (1 - B)^d X_t$$

is an $ARMA(p, q)$ process.

Autoregressive integrated moving average (ARIMA) model

- Many economic time series are well represented by the $ARIMA(p, 1, q)$ model (See Nelson and Plosser, 1982, Journal of Monetary Economics). Examples are GNP,CPI,interest rate,exchange rate, etc.
- $\{r_t\}$ is said to have a stochastic trend. This is because $\{r_t\}$ does not show quickly fluctuating behavior.

Autoregressive integrated moving average (ARIMA) model

- How do we know that $d = 1$? Perform unit root tests.
- Consider the AR(1) model

$$r_t = \phi r_{t-1} + a_t, \;\; a_t \sim WN(0, \sigma^2).$$

Let

$$\widehat{\phi} = \sum_{t=2}^{T} r_t r_{t-1} / \sum_{t=2}^{T} r_{t-1}^2$$

When $|\phi| < 1$

$$\widehat{\phi} \simeq N(\phi, \frac{1-\phi^2}{T})$$

or

$$\sqrt{T}(\widehat{\phi} - \phi) \simeq N(0, 1-\phi^2)$$

for large T.

Autoregressive integrated moving average (ARIMA) model

- Thus,

$$t(\phi) = \frac{\widehat{\phi} - \phi}{\sqrt{\hat{\sigma}^2 \left(\sum r_{t-1}^2\right)^{-1}}} \simeq N(0,1),$$

where $\hat{\sigma}^2 = \frac{1}{T-1} \sum_{t=2}^{T} (r_t - \widehat{\phi} r_{t-1})^2$.

Autoregressive integrated moving average (ARIMA) model

- However, when $\phi = 1$, $T(\widehat{\phi} - 1) \simeq$ a nonnormal distribution and

$$t(1) = \frac{\widehat{\phi} - 1}{\sqrt{\hat{\sigma}^2 (\sum r_{t-1}^2)^{-1}}} \simeq \text{ a nonnormal distribution}$$

- The distribution of $T\left(\hat{\phi} - 1\right)$ and $t(1)$ are tabulated in Wayne Fuller's "Introduction to Statistical Time Series" (1976,Wiley). These are known as Dickey-Fuller test statistics for a unit root. Critical values of these tests are taken from the LHS tails of the distributions.

Autoregressive integrated moving average (ARIMA) model

- Alternatively, we may write the model as

$$\Delta r_t = \lambda r_{t-1} + a_t, \;\; a_t \sim WN(0, \sigma^2)$$

and test the null hypothesis $H_0 : \lambda = 0$. The test statistics are

$$T\hat{\lambda} \text{ and } \frac{\widehat{\lambda}}{\sqrt{\hat{\sigma}^2 (\sum r_{t-1}^2)^{-1}}}$$

- When

$$r_t - \mu = \phi(r_t - \mu) + u_t,$$

or

$$r_t = \mu(1-\phi) + \phi r_{t-1} + u_t,$$

$\widehat{\phi}$ also has a nonnormal distribution in the limit if $\phi = 1$. The Dickey-Fuller test statistics for this model are:

$$T(\widehat{\phi} - 1) \ (\widehat{\phi} = \frac{\sum_{t=2}^{T}(r_{t-1} - \bar{r}_-)(r_t - \bar{r})}{\sum_{t=2}^{T}(r_{t-1} - \bar{r}_-)^2})$$

$$\frac{\widehat{\phi} - 1}{\sqrt{\hat{\sigma}^2(\sum_{t=2}^{T}(r_{t-1} - \bar{r}_-)^2)^{-1}}}.$$

Autoregressive integrated moving average (ARIMA) model

- An AR(p) model

$$r_t = \phi_1 r_{t-1} + \ldots + \phi_p r_{t-p} + a_t, \;\; a_t \sim WN(0, \sigma^2)$$

can be written as

$$\Delta r_t = \lambda r_{t-1} + \sum_{j=2}^{p} w_j \Delta r_{t-j+1} + a_t, \;\; a_t \sim WN(0, \sigma^2)$$

where the values of $\lambda = \phi_1 + \ldots + \phi_p - 1$ and $w_j = -\sum_{k=j}^{p} \phi_k$.

- When there is a unit root, $\phi_1 + \ldots + \phi_p = 1$.
- The null of a unit root can be tested by using the t-test for the null hypothesis $\lambda = 0$ (the augmented Dickey-Fuller test).
- It has the same asymptotic distribution as the t-test for the AR(1) model.

Seasonal ARIMA model

- If d and D are non negative integers, $\{r_t\}$ is said to be a seasonal $ARIMA(p, d, q) \times (P, D, Q)_s$ process with period s if the differenced process $Y_t = (1 - B)^d (1 - B^s)^D r_t$ is an $ARMA$ process

$$\phi(B)\Phi(B^s)Y_t = \theta(B)\Theta(B^s)a_t, \; a_t \sim WN(0, \sigma^2)$$

where

$$\begin{aligned} \phi(z) &= 1 - \phi_1 z - \cdots - \phi_p z^p, \\ \Phi(z) &= 1 - \Phi_1 z - \cdots - \Phi_p z^P, \\ \theta(z) &= 1 + \theta_1 z + \cdots + \theta_q z^q \end{aligned}$$

and

$$\Theta(z) = 1 + \Theta_1 z + \cdots + \Theta_Q z^Q.$$