5　Time Series Analysis (時系列分析)

5.1　Introduction

代表的テキスト：

・J.D. Hamilton (1994) *Time Series Analysis*

沖本・井上訳 (2006)『時系列解析 (上・下)』

・A.C. Harvey (1981) *Time Series Models*

国友・山本訳 (1985)『時系列モデル入門』

・沖本竜義 (2010)『経済・ファイナンスデータの計量時系列分析』

1. **Stationarity (定常性) :**

Let $y_1, y_2, \cdots, y_T$ be time series data.

(a) **Weak Stationarity (弱定常性) :**

$$
\begin{aligned}
&\mathrm{E}(y_t) = \mu, \\
&\mathrm{E}((y_t - \mu)(y_{t-\tau} - \mu)) = \gamma(\tau), \qquad \tau = 0, 1, 2, \cdots
\end{aligned}
$$

The first and second moments do not depend on time.

The second moment depends on time difference, not time itself.

(b) **Strong Stationarity (強定常性) :**

Let $f(y_{t_1}, y_{t_2}, \cdots, y_{t_r})$ be the joint distribution of $y_{t_1}, y_{t_2}, \cdots, y_{t_r}$.

$$
f(y_{t_1}, y_{t_2}, \cdots, y_{t_r}) = f(y_{t_1+\tau}, y_{t_2+\tau}, \cdots, y_{t_r+\tau})
$$

All the moments are same for all τ.

2. **Ergodicity (エルゴード性) :**

   As time difference between two randopm variables is large, the correlation between the two random variables becomes zero.

   $y_1, y_2, \cdots, y_T$ is said to be ergodic in mean when $\bar{y}$ converges in probability to $\mathrm{E}(y_t)$.

3. **Auto-covariance Function (自己共分散関数) :** In the case of stationary process,

$$\mathrm{E}((y_t - \mu)(y_{t-\tau} - \mu)) = \gamma(\tau), \qquad \tau = 0, 1, 2, \cdots$$

$$\gamma(\tau) = \gamma(-\tau)$$

4. **Auto-correlation Function (自己相関関数) :** In the case of stationary process,

$$\rho(\tau) = \frac{\mathrm{E}((y_t - \mu)(y_{t-\tau} - \mu))}{\sqrt{\mathrm{Var}(y_t)}\sqrt{\mathrm{Var}(y_{t-\tau})}} = \frac{\gamma(\tau)}{\gamma(0)}$$

Note that $\mathrm{Var}(y_t) = \mathrm{Var}(y_{t-\tau}) = \gamma(0)$.

Graph between τ and $\rho(\tau)$ (or $\hat{\rho}(\tau)$) $\longrightarrow$ **Correlogram (コレログラム)**

5. **Sample Mean (標本平均) :**

$$\hat{\mu} = \frac{1}{T}\sum_{t=1}^{T} y_t$$

6. **Sample Auto-covariance (標本自己共分散) :**

$$\hat{\gamma}(\tau) = \frac{1}{T}\sum_{t=\tau+1}^{T} (y_t - \hat{\mu})(y_{t-\tau} - \hat{\mu})$$

7. **Sample Auto-correlation (標本自己相関関数) :**

$$\hat{\rho}(\tau) = \frac{\hat{\gamma}(\tau)}{\hat{\gamma}(0)}$$

8. **Lag Operator (ラグ作要素) :**

$$L^{\tau} y_t = y_{t-\tau}, \qquad \tau = 1, 2, \cdots$$

9. **Likelihood Function (尤度関数)** — Innovation Form :

The joint distribution of $y_1, y_2, \cdots, y_T$ is written as:

$$\begin{aligned}
f(y_1, , y_2, \cdots, y_T) &= f(y_T | y_{T-1}, \cdots, y_1) f(y_{T-1}, \cdots, y_1) \\
&= f(y_T | y_{T-1}, \cdots, y_1) f(y_{T-1} | y_{T-2}, \cdots, y_1) f(y_{T-2}, \cdots, y_1) \\
&\ \ \vdots \\
&= f(y_T | y_{T-1}, \cdots, y_1) f(y_{T-1} | y_{T-2}, \cdots, y_1) \ \cdots \ f(y_2 | y_1) f(y_1) \\
&= f(y_1) \prod_{t=2}^{T} f(y_t | y_{t-1}, \cdots, y_1).
\end{aligned}$$

Therefore, the log-likelihood function is given by:

$$\log f(y_1, y_2, \cdots, y_T) = \log f(y_1) + \sum_{t=2}^{T} \log f(y_t|y_{t-1}, \cdots, y_1).$$

Under the normality assumption, $f(y_t|y_{t-1}, \cdots, y_1)$ is given by the normal distribution with conditional mean $\mathrm{E}(y_t|y_{t-1}, \cdots, y_1)$ and conditional variance $\mathrm{Var}(y_t|y_{t-1}, \cdots, y_1)$.

5.2 Autoregressive Model (自己回帰モデル or AR モデル)

1. **AR(p) Model :**

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \epsilon_t,$$

which is rewritten as:

$$\phi(L) y_t = \epsilon_t,$$

where

$$\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p.$$

2. **Stationarity (定常性) :**

   Suppose that all the p solutions of x from $\phi(x) = 0$ are real numbers

   When the p solutions are greater than one, y_t is stationary.

   Suppose that the p solutions include imaginary numbers.

   When the p solutions are outside unit circle, y_t is stationary.

   Example: AR(1) Model: $y_t = \phi_1 y_{t-1} + \epsilon_t$

   (a) The stationarity condition is: the solution of $\phi(x) = 1 - \phi_1 x = 0$, i.e., $x = 1/\phi_1$, is greater than one in absolute value, or equivalently, $|\phi_1| < 1$.

(b) Rewriting the AR(1) model,

$$
\begin{aligned}
y_t &= \phi_1 y_{t-1} + \epsilon_t \\
&= \phi_1^2 y_{t-2} + \epsilon_t + \phi_1 \epsilon_{t-1} \\
&= \phi_1^3 y_{t-3} + \epsilon_t + \phi_1 \epsilon_{t-1} + \phi_1^2 \epsilon_{t-2} \\
&\ \ \vdots \\
&= \phi_1^s y_{t-s} + \epsilon_t + \phi_1 \epsilon_{t-1} + \cdots + \phi_1^{s-1} \epsilon_{t-s+1}.
\end{aligned}
$$

As s is large, ϕ_1^s approaches zero. $\Longrightarrow$ Stationarity condition

(c) For stationarity, $y_t = \phi_1 y_{t-1} + \epsilon_t$ is rewritten as:

$$
y_t = \epsilon_t + \phi_1 \epsilon_{t-1} + \phi_1^2 \epsilon_{t-2} + \cdots
$$

MA representation of AR model.

(d) (*) **MA (Moving Average，移動平均) Model:** MA(q)

$$y_t = \epsilon_t + \theta_1\epsilon_{t-1} + \theta_2\epsilon_{t-2} + \cdots + \theta_q\epsilon_{t-q},$$

which is rewritten as:

$$y_t = \theta(L)\epsilon_t,$$

where

$$\theta(L) = 1 + \theta_1 L + \theta_2 L^2 + \cdots + \theta_q L^q.$$

(e) Mean of AR(1) process, μ

$$\begin{aligned}\mu = \mathrm{E}(y_t) &= \mathrm{E}(\epsilon_t + \phi_1\epsilon_{t-1} + \phi_1^2\epsilon_{t-2} + \cdots) \\ &= \mathrm{E}(\epsilon_t) + \phi_1\mathrm{E}(\epsilon_{t-1}) + \phi_1^2\mathrm{E}(\epsilon_{t-2}) + \cdots = 0\end{aligned}$$

(f) Variance of AR(1) process, $\gamma(0)$

$$\gamma(0) = \mathrm{V}(y_t) = \mathrm{V}(\epsilon_t + \phi_1\epsilon_{t-1} + \phi_1^2\epsilon_{t-2} + \cdots)$$

$$= \mathrm{V}(\epsilon_t) + \mathrm{V}(\phi_1 \epsilon_{t-1}) + \mathrm{V}(\phi_1^2 \epsilon_{t-2}) + \cdots$$

$$= \mathrm{V}(\epsilon_t) + \phi_1^2 \mathrm{V}(\epsilon_{t-1}) + \phi_1^4 \mathrm{V}(\epsilon_{t-2}) + \cdots$$

$$= \sigma^2 (1 + \phi_1^2 + \phi_1^4 + \cdots) = \frac{\sigma^2}{1 - \phi_1^2}$$

(g) Autocovariance and autocorrelation functions of the AR(1) process:

Rewriting the AR(1) process, we have:

$$y_t = \phi_1^\tau y_{t-\tau} + \epsilon_t + \phi_1 \epsilon_{t-1} + \cdots + \phi_1^{\tau-1} \epsilon_{t-\tau+1}.$$

Therefore, the autocovariance function of AR(1) process is:

$$\begin{aligned}
\gamma(\tau) &= \mathrm{E}((y_t - \mu)(y_{t-\tau} - \mu)) = \mathrm{E}(y_t y_{t-\tau}) \\
&= \mathrm{E}\Big((\phi_1^\tau y_{t-\tau} + \epsilon_t + \phi_1 \epsilon_{t-1} + \cdots + \phi_1^{\tau-1} \epsilon_{t-\tau+1}) y_{t-\tau}\Big) \\
&= \phi_1^\tau \mathrm{E}(y_{t-\tau} y_{t-\tau}) + \mathrm{E}(\epsilon_t y_{t-\tau}) + \phi_1 \mathrm{E}(\epsilon_{t-1} y_{t-\tau}) + \cdots + \phi_1^{\tau-1} \mathrm{E}(\epsilon_{t-\tau+1} y_{t-\tau}) \\
&= \phi_1^\tau \gamma(0) = \frac{\sigma^2 \phi_1^\tau}{1 - \phi_1^2}.
\end{aligned}$$

The autocorrelation function of AR(1) process is:

$$\rho(\tau) = \frac{\gamma(\tau)}{\gamma(0)} = \phi_1^{\tau}.$$

(h) **Another Derivation of $\gamma(\tau)$:**

Multiply $y_{t-\tau}$ on both sides of the AR(1) process and take the expectation:

$$\mathrm{E}(y_t y_{t-\tau}) = \phi_1 \mathrm{E}(y_{t-1} y_{t-\tau}) + \mathrm{E}(\epsilon_t y_{t-\tau})$$

$$\gamma(\tau) = \begin{cases} \phi_1 \gamma(\tau - 1), & \text{for } \tau \neq 0, \\ \phi_1 \gamma(\tau - 1) + \sigma^2, & \text{for } \tau = 0. \end{cases}$$

Using $\gamma(\tau) = \gamma(-\tau)$, $\gamma(\tau)$ for $\tau = 0$ is given by:

$$\gamma(0) = \phi_1 \gamma(1) + \sigma^2 = \phi_1^2 \gamma(0) + \sigma^2.$$

Note that $\gamma(1) = \phi_1 \gamma(0)$.

Autocovariance function $\gamma(\tau)$ is:

$$\gamma(\tau) = \phi_1 \gamma(\tau - 1) = \phi_1^2 \gamma(\tau - 2) = \cdots = \phi_1^\tau \gamma(0).$$

Therefore, $\gamma(0)$ is given by:

$$\gamma(0) = \frac{\sigma^2}{1 - \phi_1^2}$$

(i) Estimation of AR(1) model:

i. Likelihood function

$$\begin{aligned}
\log f(y_T, \cdots, y_1) &= \log f(y_1) + \sum_{t=2}^{T} \log f(y_t | y_{t-1}, \cdots, y_1) \\
&= -\frac{1}{2}\log(2\pi) - \frac{1}{2}\log\left(\frac{\sigma^2}{1 - \phi_1^2}\right) - \frac{1}{2\sigma^2/(1 - \phi_1^2)} y_1^2 \\
&\quad -\frac{T-1}{2}\log(2\pi) - \frac{T-1}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{t=2}^{T}(y_t - \phi_1 y_{t-1})^2
\end{aligned}$$

$$= -\frac{T}{2}\log(2\pi) - \frac{T}{2}\log(\sigma^2) - \frac{1}{2}\log\left(\frac{1}{1-\phi_1^2}\right)$$
$$-\frac{1}{2\sigma^2/(1-\phi_1^2)}y_1^2 - \frac{1}{2\sigma^2}\sum_{t=2}^{T}(y_t - \phi_1 y_{t-1})^2$$

Note as follows:

$$f(y_1) = \frac{1}{\sqrt{2\pi\sigma^2/(1-\phi_1^2)}}\exp\left(-\frac{1}{2\sigma^2/(1-\phi_1^2)}y_1^2\right)$$
$$f(y_t|y_{t-1},\cdots,y_1) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{1}{2\sigma^2}(y_t - \phi_1 y_{t-1})^2\right)$$

$$\frac{\partial \log f(y_T,\cdots,y_1)}{\partial\sigma^2} = -\frac{T}{2}\frac{1}{\sigma^2} + \frac{1}{2\sigma^4/(1-\phi_1^2)}y_1^2 + \frac{1}{2\sigma^4}\sum_{t=2}^{T}(y_t - \phi_1 y_{t-1})^2 = 0$$

$$\frac{\partial \log f(y_T,\cdots,y_1)}{\partial\phi_1} = -\frac{\phi_1}{1-\phi_1^2} + \frac{\phi_1}{\sigma^2}y_1^2 + \frac{1}{\sigma^2}\sum_{t=2}^{T}(y_t - \phi_1 y_{t-1})y_{t-1} = 0$$

The MLE of ϕ_1 and σ^2 satisfies the above two equation.

$$\tilde{\sigma}^2 = \frac{1}{T}\left((1-\tilde{\phi}_1^2)y_1^2 + \sum_{t=2}^{T}(y_t - \tilde{\phi}_1 y_{t-1})^2\right)$$

$$\tilde{\phi}_1 = \frac{\sum_{t=2}^{T} y_t y_{t-1}}{\sum_{t=2}^{T} y_{t-1}^2} + \left(\tilde{\phi}_1 y_1^2 - \frac{\tilde{\sigma}^2\tilde{\phi}_1}{1-\tilde{\phi}_1^2}\right)\Big/\sum_{t=2}^{T} y_{t-1}^2$$

ii. Ordinary Least Squares (OLS) Method

$$S(\phi_1) = \sum_{t=2}^{T}(y_t - \phi_1 y_{t-1})^2$$

is minimized with respect to ϕ_1.

$$\hat{\phi}_1 = \frac{\sum_{t=2}^{T} y_{t-1}y_t}{\sum_{t=2}^{T} y_{t-1}^2} = \phi_1 + \frac{\sum_{t=2}^{T} y_{t-1}\epsilon_t}{\sum_{t=2}^{T} y_{t-1}^2} = \phi_1 + \frac{(1/T)\sum_{t=2}^{T} y_{t-1}\epsilon_t}{(1/T)\sum_{t=2}^{T} y_{t-1}^2}$$

$$\longrightarrow \phi_1 + \frac{\mathrm{E}(y_{t-1}\epsilon_t)}{\mathrm{E}(y_{t-1}^2)} = \phi_1$$

OLSE of ϕ_1 is a consistent estimator.

The following equations are utilized.

$$\mathrm{E}(y_{t-1}\epsilon_t) = 0$$

$$\mathrm{E}(y_{t-1}^2) = \mathrm{Var}(y_{t-1}) = \gamma(0)$$

(j) Asymptotic distribution of OLSE $\hat{\phi}_1$:

$$\sqrt{T}(\hat{\phi}_1 - \phi_1) \longrightarrow N(0, 1 - \phi_1^2)$$

Proof:

$y_{t-1}\epsilon_t$ is distributed with mean zero and variance $\sigma^2\gamma(0) = \dfrac{\sigma^4}{1-\phi_1^2}$, where $\sigma^2 = \mathrm{V}(\epsilon_t)$.

Note as follows:

$$\mathrm{E}(y_{t-1}\epsilon_t) = 0, \qquad \mathrm{V}(y_{t-1}\epsilon_t) = \sigma^2\gamma(0)$$

ϵ_t is independent of $y_{t-1} = \epsilon_{t-1} + \phi_1\epsilon_{t-2} + \phi_1^2\epsilon_{t-3} + \cdots$.

$\mathrm{V}(y_{t-1}\epsilon_t) = \mathrm{E}(\epsilon_t^2 y_{t-1}^2) = \mathrm{E}(\epsilon_t^2)\mathrm{E}(y_{t-1}^2) = \sigma^2\gamma(0)$, using $\mathrm{Cov}(\epsilon_t^2,\ y_{t-1}^2) = \mathrm{E}(\epsilon_t^2 y_{t-1}^2) - \mathrm{E}(\epsilon_t^2)\mathrm{E}(y_{t-1}^2) = 0$.

Furthermore, $y_{t-1}\epsilon_t$ is independent of $y_{s-1}\epsilon_s$ for $t \neq s$.

$\mathrm{Cov}(y_{t-1}\epsilon_t,\ y_{s-1}\epsilon_s) = \mathrm{E}(\epsilon_t\epsilon_s y_{t-1}y_{s-1})$ because of $\mathrm{E}(y_{t-1}\epsilon_t) = 0$.

From $y_{t-1} = \sum_{i=1}^{\infty} \phi_1^{i-1}\epsilon_{t-i}$,

$\epsilon_t\epsilon_s y_{t-1}y_{s-1} = \epsilon_t\epsilon_s(\sum_{i=1}^{\infty} \phi_1^{i-1}\epsilon_{t-i})(\sum_{j=1}^{\infty} \phi_1^{j-1}\epsilon_{s-j}) = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} \phi_1^{i+j-2}\epsilon_t\epsilon_s\epsilon_{t-i}\epsilon_{s-j}$.

Thus, $\mathrm{Cov}(y_{t-1}\epsilon_t,\ y_{s-1}\epsilon_s) = 0$ from $\mathrm{E}(\epsilon_t\epsilon_s\epsilon_{t-i}\epsilon_{s-j}) = 0$ except $t = s$ and $i = j$ for $i, j = 1, 2, \cdots$.

From the central limit theorem,

$$\frac{(1/T)\sum_{t=1}^{T} y_{t-1}\epsilon_t}{\sqrt{\sigma^2\gamma(0)/T}} \longrightarrow N(0, 1)$$

Rewriting,

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T} y_{t-1}\epsilon_t \longrightarrow N\Big(0,\ \sigma^2\gamma(0)\Big).$$

Next,

$$\frac{1}{T}\sum_{t=1}^{T} y_{t-1}^2 \longrightarrow \mathrm{E}(y_{t-1}^2) = \gamma(0) = \frac{\sigma^2}{1-\phi_1^2}$$

yields:

$$\sqrt{T}(\hat{\phi}_1 - \phi_1) = \frac{(1/\sqrt{T})\sum_{t=1}^{T} y_{t-1}\epsilon_t}{(1/T)\sum_{t=1}^{T} y_{t-1}^2} \longrightarrow N\Big(0,\ \frac{\sigma^2}{\gamma(0)}\Big) = N(0, 1-\phi_1^2)$$

(k) Some formulas:

i. Central Limit Theorem

Random variables $x_1,\ x_2,\ \cdots,\ x_T$ are mutually independently distributed with mean μ and variance σ^2.

Define $\overline{x} = (1/T)\sum_{t=1}^{T} x_t$.

Then,

$$\frac{\overline{x} - \mathrm{E}(\overline{x})}{\sqrt{\mathrm{V}(\overline{x})}} = \frac{\overline{x} - \mu}{\sigma/\sqrt{T}} \longrightarrow N(0, 1)$$

ii. Central Limit Theorem II

Random variables $x_1, x_2, \cdots, x_T$ are distributed with mean μ and variance σ^2.

Define $\overline{x} = (1/T)\sum_{t=1}^{T} x_t$.

Then,

$$\frac{\overline{x} - \mathrm{E}(\overline{x})}{\sqrt{\mathrm{V}(\overline{x})}} \longrightarrow N(0, 1)$$

iii. Let x and y be random variables.

y converges in distribution to a distribution, and x converges in probability to a fixed value.

Then, xy converges in distribution.

For example, consider:

$$y \longrightarrow N(\mu, \sigma^2), \qquad x \longrightarrow c.$$

Then, we obtain:

$$xy \longrightarrow N(c\mu, c^2\sigma^2)$$

5.3 Unit Root (単位根) Test (Dickey-Fuller (DF) Test)

1. **The Case of $|\phi_1| < 1$:**

   $$y_t = \phi_1 y_{t-1} + \epsilon_t, \qquad \epsilon_t \sim \text{i.i.d. } N(0, \sigma^2), \qquad y_0 = 0, \qquad t = 1, \cdots, T$$

   Then, OLSE of ϕ_1 is:

   $$\hat{\phi}_1 = \frac{\sum_{t=1}^{T} y_{t-1} y_t}{\sum_{t=1}^{T} y_{t-1}^2}.$$

   We have the following asymptotic distribution:

   $$\sqrt{T}(\hat{\phi}_1 - \phi_1) = \frac{\frac{1}{\sqrt{T}} \sum_{t=1}^{T} y_{t-1}\epsilon_t}{\frac{1}{T} \sum_{t=1}^{T} y_{t-1}^2} \longrightarrow N\left(0, \frac{\sigma^2}{\gamma(0)}\right) = N\left(0, 1 - \phi_1^2\right).$$

Note that $\gamma(0) = \dfrac{\sigma^2}{1-\phi_1^2}$.

2. In the case of $\phi_1 = 1$, as expected, we have:

$$\sqrt{T}(\hat{\phi}_1 - 1) \longrightarrow 0.$$

That is, $\hat{\phi}_1$ has the distribution which converges in probability to $\phi_1 = 1$ (i.e., degenerated distribution).

Is this true?

3. **The Case of $\boldsymbol{\phi_1 = 1}$:** $\Longrightarrow$ Random Walk Process

$y_t = y_{t-1} + \epsilon_t$ with $y_0 = 0$ is written as:

$$y_t = \epsilon_t + \epsilon_{t-1} + \epsilon_{t-2} + \cdots + \epsilon_1.$$

Therefore, we can obtain:

$$y_t \sim N(0, \sigma^2 t), \qquad \Longrightarrow \qquad \frac{y_t}{\sigma\sqrt{t}} \sim N(0,\ 1).$$

The variance of y_t depends on time t. $\Longrightarrow$ y_t is nonstationary.

4. Remember that $\hat{\phi}_1 = \phi_1 + \dfrac{\sum y_{t-1}\epsilon_t}{\sum y_{t-1}^2}$.

   (a) First, consider the numerator $\sum y_{t-1}\epsilon_t$.

   We have $y_t^2 = (y_{t-1} + \epsilon_t)^2 = y_{t-1}^2 + 2y_{t-1}\epsilon_t + \epsilon_t^2$.

   Therefore, we obtain:

$$y_{t-1}\epsilon_t = \frac{1}{2}(y_t^2 - y_{t-1}^2 - \epsilon_t^2).$$

   Taking into account $y_0 = 0$, we have:

$$\sum_{t=1}^{T} y_{t-1}\epsilon_t = \frac{1}{2}y_T^2 - \frac{1}{2}\sum_{t=1}^{T}\epsilon_t^2.$$