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

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

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

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

$$\begin{aligned}\tilde{\phi}_1 &= \frac{\sum_{t=2}^{T} y_t y_{t-1}}{\sum_{t=2}^{T} y_{t-1}^2} + \left(\tilde{\phi}_1 y_1^2 - \frac{\tilde{\sigma}^2 \tilde{\phi}_1}{1-\tilde{\phi}_1^2}\right) \Big/ \sum_{t=2}^{T} y_{t-1}^2 \\ &\approx \frac{\sum_{t=2}^{T} y_t y_{t-1}}{\sum_{t=2}^{T} y_{t-1}^2}, \qquad \text{when } T \text{ is large.}\end{aligned}$$

(b) 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.

OLSE of ϕ_1 is equal to MLE when T is large.

The following equations are utilized.

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

8. Asymptotic distribution of OLSE $\hat{\phi}_1$:

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

Proof:

$y_{t-1}\epsilon_t$, $t = 1, 2, \cdots, T$, are distributed with mean zero and variance $\dfrac{\sigma_\epsilon^4}{1 - \phi_1^2}$.

From the central limit theorem,

$$\frac{(1/T)\sum_{t=1}^{T} y_{t-1}\epsilon_t}{\sqrt{\sigma_\epsilon^4/(1 - \phi_1^2)}/\sqrt{T}} \longrightarrow N(0, 1)$$

Rewriting,

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T} y_{t-1}\epsilon_t \longrightarrow N(0, \frac{\sigma_\epsilon^4}{1 - \phi_1^2}).$$

Next,

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

yields:

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

9. Some formulas:

   (a) 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,

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

   (b) 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,

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

(c) 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:

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

10. **AR(1) +drift:** $y_t = \mu + \phi_1 y_{t-1} + \epsilon_t$

Mean:

Using the lag operator,

$$\phi(L)y_t = \mu + \epsilon_t$$

where $\phi(L) = 1 - \phi_1 L$.

Multiply $\phi(L)^{-1}$ on both sides. Then, when $|\phi_1| < 1$, we have:

$$y_t = \phi(L)^{-1}\mu + \phi(L)^{-1}\epsilon_t.$$

Taking the expectation on both sides,

$$\begin{aligned} \mathrm{E}(y_t) &= \phi(L)^{-1}\mu + \phi(L)^{-1}\mathrm{E}(\epsilon_t) \\ &= \phi(1)^{-1}\mu = \frac{\mu}{1 - \phi_1} \end{aligned}$$

Example: AR(2) Model: Consider $y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \epsilon_t$.

1. The stationarity condition is: two solutions of x from $\phi(x) = 1 - \phi_1 x - \phi_2 x^2 = 0$ are outside the unit circle.

2. Rewriting the AR(2) model,

$$(1 - \phi_1 L - \phi_2 L^2) y_t = \epsilon_t.$$

Let $1/\alpha_1$ and $1/\alpha_2$ be the solutions of $\phi(x) = 0$.

Then, the AR(2) model is written as:

$$(1 - \alpha_1 L)(1 - \alpha_2 L) y_t = \epsilon_t,$$

which is rewritten as:

$$y_t = \frac{1}{(1 - \alpha_1 L)(1 - \alpha_2 L)} \epsilon_t$$

$$= \left(\frac{\alpha_1/(\alpha_1 - \alpha_2)}{1 - \alpha_1 L} + \frac{-\alpha_2/(\alpha_1 - \alpha_2)}{1 - \alpha_2 L} \right) \epsilon_t$$

3. **Mean of AR(2) Model:**

   When y_t is stationary, i.e., α_1 and α_2 are inside the unit circle,

$$\mu = \mathrm{E}(y_t) = \mathrm{E}(\phi(L)\epsilon_t) = 0$$

4. **Autocovariance Function of AR(2) Model:**

$$\begin{aligned}
\gamma(\tau) &= \mathrm{E}((y_t - \mu)(y_{t-\tau} - \mu)) = \mathrm{E}(y_t y_{t-\tau}) \\
&= \mathrm{E}\Big((\phi_1 y_{t-1} + \phi_2 y_{t-2} + \epsilon_t) y_{t-\tau}\Big) \\
&= \phi_1 \mathrm{E}(y_{t-1} y_{t-\tau}) + \phi_2 \mathrm{E}(y_{t-2} y_{t-\tau}) + \mathrm{E}(\epsilon_t y_{t-\tau}) \\
&= \begin{cases} \phi_1 \gamma(\tau - 1) + \phi_2 \gamma(\tau - 2), & \text{for } \tau \neq 0, \\ \phi_1 \gamma(\tau - 1) + \phi_2 \gamma(\tau - 2) + \sigma_\epsilon^2, & \text{for } \tau = 0. \end{cases}
\end{aligned}$$

The initial condition is obtained by solving the following three equations:

$$\gamma(0) = \phi_1\gamma(1) + \phi_2\gamma(2) + \sigma_\epsilon^2,$$

$$\gamma(1) = \phi_1\gamma(0) + \phi_2\gamma(1),$$

$$\gamma(2) = \phi_1\gamma(1) + \phi_2\gamma(0).$$

Therefore, the initial conditions are given by:

$$\gamma(0) = \left(\frac{1-\phi_2}{1+\phi_2}\right)\frac{\sigma_\epsilon^2}{(1-\phi_2)^2-\phi_1^2},$$

$$\gamma(1) = \frac{\phi_1}{1-\phi_2}\gamma(0) = \left(\frac{\phi_1}{1-\phi_2}\right)\left(\frac{1-\phi_2}{1+\phi_2}\right)\frac{\sigma_\epsilon^2}{(1-\phi_2)^2-\phi_1^2}.$$

Given $\gamma(0)$ and $\gamma(1)$, we obtain $\gamma(\tau)$ as follows:

$$\gamma(\tau) = \phi_1\gamma(\tau-1) + \phi_2\gamma(\tau-2), \qquad \text{for } \tau = 2, 3, \cdots.$$

5. **Another solution for $\gamma(0)$:**

From $\gamma(0) = \phi_1\gamma(1) + \phi_2\gamma(2) + \sigma_\epsilon^2$,

$$\gamma(0) = \frac{\sigma_\epsilon^2}{1 - \phi_1\rho(1) - \phi_2\rho(2)}$$

where

$$\rho(1) = \frac{\phi_1}{1 - \phi_2}, \qquad \rho(2) = \phi_1\rho(1) + \phi_2 = \frac{\phi_1^2 + (1 - \phi_2)\phi_2}{1 - \phi_2}.$$

6. **Autocorrelation Function of AR(2) Model:**

Given $\rho(1)$ and $\rho(2)$,

$$\rho(\tau) = \phi_1\rho(\tau - 1) + \phi_2\rho(\tau - 2), \qquad \text{for } \tau = 3, 4, \cdots,$$

7. **$\boldsymbol{\phi_{k,k}}$ = Partial Autocorrelation Coefficient of AR(2) Process:**

$$\begin{pmatrix} 1 & \rho(1) & \cdots & \rho(k-2) & \rho(k-1) \\ \rho(1) & 1 & & \rho(k-3) & \rho(k-2) \\ \vdots & \vdots & & \vdots & \vdots \\ \rho(k-1) & \rho(k-2) & \cdots & \rho(1) & 1 \end{pmatrix} \begin{pmatrix} \phi_{k,1} \\ \phi_{k,2} \\ \vdots \\ \phi_{k,k-1} \\ \phi_{k,k} \end{pmatrix} = \begin{pmatrix} \rho(1) \\ \rho(2) \\ \vdots \\ \rho(k) \end{pmatrix},$$

for $k = 1, 2, \cdots$.

$$\phi_{k,k} = \frac{\begin{vmatrix} 1 & \rho(1) & \cdots & \rho(k-2) & \rho(1) \\ \rho(1) & 1 & & \rho(k-3) & \rho(2) \\ \vdots & \vdots & & \vdots & \vdots \\ \rho(k-1) & \rho(k-2) & \cdots & \rho(1) & \rho(k) \end{vmatrix}}{\begin{vmatrix} 1 & \rho(1) & \cdots & \rho(k-2) & \rho(k-1) \\ \rho(1) & 1 & & \rho(k-3) & \rho(k-2) \\ \vdots & \vdots & & \vdots & \vdots \\ \rho(k-1) & \rho(k-2) & \cdots & \rho(1) & 1 \end{vmatrix}}$$

Autocovariance Functions:

$$\gamma(1) = \phi_1\gamma(0) + \phi_2\gamma(1),$$

$$\gamma(2) = \phi_1\gamma(1) + \phi_2\gamma(0),$$

$$\gamma(\tau) = \phi_1\gamma(\tau-1) + \phi_2\gamma(\tau-2), \qquad \text{for } \tau = 3, 4, \cdots.$$

Autocorrelation Functions:

$$\rho(1) = \phi_1 + \phi_2\rho(1) = \frac{\phi_1}{1-\phi_2},$$

$$\rho(2) = \phi_1\rho(1) + \phi_2 = \frac{\phi_1^2}{1-\phi_2} + \phi_2,$$

$$\rho(\tau) = \phi_1\rho(\tau-1) + \phi_2\rho(\tau-2), \qquad \text{for } \tau = 3, 4, \cdots.$$

$$\phi_{1,1} = \rho(1) = \frac{\phi_1}{1 - \phi_2}$$

$$\phi_{2,2} = \frac{\begin{vmatrix} 1 & \rho(1) \\ \rho(1) & \rho(2) \end{vmatrix}}{\begin{vmatrix} 1 & \rho(1) \\ \rho(1) & 1 \end{vmatrix}} = \frac{\rho(2) - \rho(1)^2}{1 - \rho(1)^2} = \phi_2$$

$$\phi_{3,3} = \frac{\begin{vmatrix} 1 & \rho(1) & \rho(1) \\ \rho(1) & 1 & \rho(2) \\ \rho(2) & \rho(1) & \rho(3) \end{vmatrix}}{\begin{vmatrix} 1 & \rho(1) & \rho(2) \\ \rho(1) & 1 & \rho(1) \\ \rho(2) & \rho(1) & 1 \end{vmatrix}}$$

$$= \frac{(\rho(3) - \rho(1)\rho(2)) - \rho(1)^2(\rho(3) - \rho(1)) + \rho(2)\rho(1)(\rho(2) - 1)}{(1 - \rho(1)^2) - \rho(1)^2(1 - \rho(2)) + \rho(2)(\rho(1)^2 - \rho(2))} = 0.$$

8. **Log-Likelihood Function — Innovation Form:**

$$\log f(y_T, \cdots, y_1) = \log f(y_2, y_1) + \sum_{t=3}^{T} \log f(y_t | y_{t-1}, \cdots, y_1)$$

where

$$f(y_2, y_1) = \frac{1}{2\pi} \begin{vmatrix} \gamma(0) & \gamma(1) \\ \gamma(1) & \gamma(0) \end{vmatrix}^{-1/2} \exp\left(-\frac{1}{2}(y_1 \; y_2) \begin{pmatrix} \gamma(0) & \gamma(1) \\ \gamma(1) & \gamma(0) \end{pmatrix}^{-1} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right),$$

$$f(y_t | y_{t-1}, \cdots, y_1) = \frac{1}{\sqrt{2\pi\sigma_\epsilon^2}} \exp\left(-\frac{1}{2\sigma_\epsilon^2}(y_t - \phi_1 y_{t-1} - \phi_2 y_{t-2})^2 \right).$$

Note as follows:

$$\begin{pmatrix} \gamma(0) & \gamma(1) \\ \gamma(1) & \gamma(0) \end{pmatrix} = \gamma(0) \begin{pmatrix} 1 & \rho(1) \\ \rho(1) & 1 \end{pmatrix} = \gamma(0) \begin{pmatrix} 1 & \phi_1/(1 - \phi_2) \\ \phi_1/(1 - \phi_2) & 1 \end{pmatrix}.$$

9. **AR(2) +drift:** $y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \epsilon_t$

Mean:

Rewriting the AR(2)+drift model,

$$\phi(L)y_t = \mu + \epsilon_t$$

where $\phi(L) = 1 - \phi_1 L - \phi_2 L^2$.

Under the stationarity assumption, we can rewrite the AR(2)+drift model as follows:

$$y_t = \phi(L)^{-1}\mu + \phi(L)^{-1}\epsilon_t.$$

Therefore,

$$\mathrm{E}(y_t) = \phi(L)^{-1}\mu + \phi(L)^{-1}\mathrm{E}(\epsilon_t) = \phi(1)^{-1}\mu = \frac{\mu}{1 - \phi_1 - \phi_2}$$