Math Review Notes—Time Series

Gregory Faletto

Contents

Last updated November 6, 2018

1 Time Series

These notes are based on my notes from *Time Series and Panel Data Econometrics* (1st edition) by M. Hashem Pesaran as well as coursework for Economics 613: Economic and Financial Time Series I at USC.

1.1 Chapter 6: ARDL Models

In an ARDL model, if the error are serially correlated, then the coefficient estimates are biased (even as $T \to \infty$).

1.2 Chapters 12 and 13: Intro to Stochastic Processes and Spectral Analysis

Stationarity conditions: $\{X_t\}$ is **strictly stationary** if the joint distribution functions of $\{X_{t_1}, X_{t_2}, \ldots, X_{t_k}\}$ and $\{X_{t_1+h}, X_{t_2+h}, \ldots, X_{t_k+h}\}$ are identical for all values of $t_1, t_2, \ldots, t_k$ and h and all positive integers k.

X_t is **weakly (or covariance) stationary** if it has a constant mean and variance and its covariance function $\gamma(t_1, t_2)$ depends only on the absolute difference $|t_1 - t_2|$, namely $\gamma(t_1, t_2) = \gamma(|t_1 - t_2|)$.

X_t is said to be **trend stationary** if $y_t = X_t - d_t$ is covariance stationary, where d_t is the perfectly predictable component of X_t.

The process $\{\epsilon_t\}$ is said to be a **white noise process** if it has mean zero, a constant variance, and ϵ_t and ϵ_s are uncorrelated for all $s \neq t$.

Autocovariance generating function: The autocovariance generating function for the general linear stationary process $y_t = \sum_{i=0}^{\infty} a_i \epsilon_{t-i}$ is given by:

$$G(z) = \sigma^2 a(z) a(z^{-1})$$

where $a(z) = \sum_{i=0}^{\infty} a_i z^i$.

Wold's Decomposition (Theorem 42, p. 275, Section 12.5) Any trend-stationary process $\{y_t\}$ can be represented in the form of $y_t = d_t + \sum_{i=0}^{\infty} \alpha_i \epsilon_{t-i}$ where $\alpha_0 = 1$ and $\sum_{i=0}^{\infty} \alpha_i^2 < K < \infty$. The term d_t is a deterministic component, while $\{\epsilon_t\}$ is a serially uncorrelated process: $\epsilon_t = y_t - \mathbb{E}(y_t \mid y_{t-1}, y_{t-2}, \ldots)$.

Stationarity conditions for an ARMA(p,q) process: Consider the ARMA(p,q) process

$$y_t = \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{i=0}^{q} \theta_i \epsilon_{t-i}, \quad \theta_0 = 1$$

The MA part is stationary for any finite q. The AR part is stationary if the roots of the characteristic equation

$$\lambda^t = \sum_{i=1}^{p} \phi_i \lambda^{t-i}$$

lie strictly inside the unit circle. Alternatively, in terms of $z = \lambda^{-1}$, the process is stationary if the roots of

$$1 - \sum_{i=1}^{p} \phi_i z^i = 0$$

lie outside the unit circle. The ARMA process is **invertible** (so that y_t can be solved uniquely in terms of its past values) if all the roots of

$$1 - \sum_{i=1}^{p} \theta_i z^i = 0$$

fall outside the unit circle.

Spectral Density Function: Definition (Equation 13.3):

$$f(\omega) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \gamma(h) e^{ih\omega}, \omega \in (-\pi, \pi)$$

Equation (13.5):

$$f(\omega) = \frac{1}{2\pi} \left[\gamma(0) + 2 \sum_{h=1}^{\infty} \gamma(h) \cos(h\omega) \right], \quad \omega \in [0, \pi]$$

Can also be found using the autocovariance generating function. We have (Equation 13.6, section 13.3.1)

$$f(\omega) = \frac{1}{2\pi} G(e^{i\omega}) = \frac{\sigma^2}{2\pi} a(e^{i\omega}) a(e^{-i\omega})$$

Properties of spectral density function:

(1) $f(\omega)$ always exists and is bounded if $\gamma(h)$ is absolutely summable.

(2) $f(\omega)$ is symmetric.

(3) The spectrum of a stationary process is finite at zero frequency; that is, $f(0) < \infty$.

Linear (time-domain) processes don't have to be stationary, but to write something as a frequency-domain process, it must be stationary.

1.3 Some time series and their properties

1.3.1 White noise process:

$x_t = \epsilon_t$, $\epsilon_t \sim IID(0, \sigma^2)$

- Autocovariances:

$$\gamma(0) = \sigma^2$$

$$\gamma(h) = 0, \quad \forall h \neq 0$$

- Spectral density function:

$$f_x(\omega) = \frac{1}{2\pi} \cdot \sigma^2 = \frac{\sigma^2}{2\pi} \text{ (flat spectrum)}$$

1.3.2 MA(1) process:

$x_t = \epsilon_t + \theta \epsilon_{t-1}$ with $\epsilon_t \sim iid(0, \sigma^2)$, $|\rho| < 1$.

- Autocovariances: By Equation (12.2), the autocovariance function is

$$\text{Cov}(u_t, u_{t-h}) = \gamma(h) = \sigma^2 \sum_{i=0}^{1-|h|} a_i a_{i+|h|} \text{ if } 0 \leq |h| \leq 1$$

$$\implies \mathbb{E}(x_t^2) = \gamma(0) = (1 + \theta^2)\sigma^2$$

$$\mathbb{E}(x_t x_{t-1}) = \gamma(1) = \theta \sigma^2$$

$$\gamma(h) = 0 \ \ \forall \ |h| > 1$$

So the covariance matrix is

$$\begin{pmatrix} \sigma^2(1+\theta^2) & \sigma^2\theta & 0 & 0 & \cdots & 0 \\ \sigma^2\theta & \sigma^2(1+\theta^2) & \sigma^2\theta & 0 & \cdots & 0 \\ 0 & \sigma^2\theta & \sigma^2(1+\theta^2) & \sigma^2\theta & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \sigma^2\theta & \sigma^2(1+\theta^2) & \sigma^2\theta \\ 0 & 0 & \cdots & 0 & \sigma^2\theta & \sigma^2(1+\theta^2) \end{pmatrix}$$

$$= \sigma^2(1+\theta^2) I_T + \sigma^2 \theta A$$

where A is defined as in section 14.3.2 (p. 304).

- Spectral density function:

$$f(\omega) = \frac{\sigma^2}{2\pi} \big[1 + 2\theta \cos(\omega) + \rho^2\big], \quad \omega \in [0, \pi]$$

1.3.3 MA(∞) process:

This process is covariance stationary.

- Autocovariances:

1.3.4 AR(1) process:

$x_t = \phi x_{t-1} + \epsilon_t$, $|\phi| < 1$, $\epsilon_t \sim IID(0, \sigma^2)$.

- Yule-Walker Equations:

$$\mathbb{E}[x_t x_{t-h}] = \mathbb{E}[\phi x_{t-1} x_{t-h}] + \mathbb{E}[\epsilon x_{t-h}]$$

$$\gamma_h = \phi \gamma_{h-1} + \mathbb{E}[\epsilon x_{t-h}]$$

$$\implies \gamma_0 = \phi \gamma_1 + \sigma^2, \quad \gamma_h = \phi \gamma_{h-1} \; \forall h \geq 1$$

- Autocovariances:

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

$$\gamma_h = \frac{\sigma^2 \phi^h}{1 - \rho^2} \quad \forall \; h \geq 1$$

$$\implies \text{Cov}(x) =$$

$$\begin{pmatrix} \sigma^2/(1-\phi^2) & \sigma^2\phi/(1-\phi^2) & \sigma^2\phi^2/(1-\phi^2) & \sigma^2\phi^3/(1-\phi^2) & \cdots & \sigma^2\phi^{T-1}/(1-\phi^) \\ \sigma^2\phi^/(1-\phi) & \sigma^2/(1-\phi^2) & \sigma^2\phi/(1-\phi^2) & \sigma^2\phi^2/(1-\phi^2) & \cdots & \sigma^2\phi^{T-2}/(1-\phi^2) \\ \sigma^2\phi^2/(1-\phi^2) & \sigma^2\phi/(1-\phi^2) & \sigma^2/(1-\phi^2) & \sigma^2\phi/(1-\phi^2) & \cdots & \sigma^2\phi^{T-3}/(1-\phi^2) \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ \sigma^2\phi^{T-2}/(1-\phi^2) & \sigma^2\phi^{T-3}/(1-\phi^2) & \cdots & \sigma^2\phi/(1-\phi^2) & \sigma^2/(1-\phi^2) & \sigma^2\phi/(1-\phi^2) \\ \sigma^2\phi^{T-1}/(1-\phi^2) & \sigma^2\phi^{T-2}/(1-\phi^2) & \cdots & \sigma^2\phi^2/(1-\phi^2) & \sigma^2\phi/(1-\phi^2) & \sigma^2/(1-\phi^2) \end{pmatrix}$$

- If stationary, can be written as an infinite MA process with absolutely summable coefficients

$$x_t = \sum_{i=0}^{\infty} \phi^i \epsilon_{t-i} = \left(\frac{1}{1 - \phi L} \right) \epsilon_t$$

- Autocovariance generating function:

$$G(z) = \left(\frac{\sigma^2}{1 - \phi^2} \right) \left(1 + \sum_{h=1}^{\infty} \phi^h (z^h + z^{-h}) \right)$$

- Spectral density function:

$$f(\omega) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \frac{\sigma^2 \phi^{|h|}}{(1-\phi^2)} (e^{i\omega})^h = \frac{1}{2\pi} \frac{\sigma^2}{(1-\phi e^{i\omega})(1-\phi e^{-i\omega})} = \frac{1}{2\pi} \frac{\sigma^2}{1 - 2\phi\cos(\omega) + \phi^2}$$

1.3.5 AR(2) process:

$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \epsilon$, $|\phi_1| < 1$, $|\phi_2| < 1$, $\epsilon_t \sim IID(0, \sigma^2)$.

Can be written as

$$x_t = \frac{1}{1-\phi L}\epsilon_t = \epsilon_t + \phi\epsilon_{t-1} + \phi^2\epsilon_{t-2} + \ldots$$

- Yule-Walker equations:

$$\mathbb{E}[x_t x_{t-h}] = \mathbb{E}[\phi_1 x_{t-1} x_{t-h}] + \mathbb{E}[\phi_2 x_{t-2} x_{t-h}] + \mathbb{E}[\epsilon x_{t-h}]$$

$$\gamma_h = \phi_1 \gamma_{h-1} + \phi_2 \gamma_{h-2} + \mathbb{E}[\epsilon x_{t-h}]$$

$$\implies \boxed{\gamma_0 = \phi_1\gamma_1 + \phi_2\gamma_2 + \sigma^2, \quad \gamma_1 = \phi_1\gamma_0 + \phi_2\gamma_1, \quad \gamma_2 = \phi_1\gamma_1 + \phi_2\gamma_0}$$

- Autocovariances:

1.3.6 AR(p) process:

$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \ldots + \phi_p x_{t-p} + \epsilon$, $|\phi_i| < 1$, $\epsilon_t \sim IID(0, \sigma^2)$.

- Stationary if the eigenvalues of $\mathbf{\Phi}$ lie inside the unit circle, which is equivalent to all the roots of

  $$\phi(z) = 1 - \phi_1 z - \phi_2 z^2 - \ldots - \phi_p z^p = 0$$

  being strictly larger than unity. Under this condition the AR process has the infinite-order MA representation'

  $$x_t = \sum_{i=0}^{\infty} \alpha_i \epsilon_{t-i}$$

  where $\alpha_i = \phi_1 \alpha_{i-1} + \ldots + \phi_p \alpha_{i-p}$.

- Autocovariance generating function:

  $$G(z) = \frac{\sigma^2}{\phi(z)\phi(z^{-1})}$$

1.3.7 ARMA(1, 1) process:

$x_t = \phi x_{t-1} + \epsilon_t + \theta\epsilon_{t-1}$, with $|\phi| < 1$ (implying stationarity), $\mathbb{E}(\epsilon_t^2) = \sigma^2$, $\mathbb{E}(\epsilon_t\epsilon_s) = 0$ for $t \neq s$.

- Yule-Walker Equations:

$$\gamma(0) = \phi\gamma(1) + \sigma^2(1 + \theta^2)$$

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

$$\gamma(h) = \phi\gamma(h-1) \;\; \forall\; h \geq 2$$

- Autocovariances:

$$\gamma(0) = \sigma^2\left(1_{\frac{(\phi+\theta)^2}{1-\phi^2}}\right)$$

$$\gamma(1) = \sigma^2\left(\phi + \theta + \frac{(\phi+\theta)^2\phi}{1-\phi^2}\right)$$

$$\gamma(2) = \phi^{h-1}\gamma(1) \;\forall\; h \geq 2$$

- Autocorrelation function:

$$\rho(h) = \begin{cases} 1 & h = 0 \\ \frac{(\phi+\theta)(1+\phi\theta)}{1+2\phi\theta+\theta^2} & h = 1 \\ \phi^{h-1}\rho(1) & h \geq 2 \end{cases}$$

- Autocovariance generating function: the autocovariance function of an ARMA(p, q) process $\phi(L)y_t = \theta(L)\epsilon_t$ is given by

$$f(\omega) = \sigma^2\frac{\theta(z)\theta(z^{-1})}{\phi(z)\phi(z^{-1})}$$

  Plugging in for the ARMA(1,1) case yields (**double-check**)

$$f(\omega) = \sigma^2\frac{(1+\theta)^2}{(1-\rho)^2}$$

- Spectral Density Function: the spectral density function of an ARMA(p, q) process $\phi(L)y_t = \theta(L)\epsilon_t$ is given by

$$f(\omega) = \frac{\sigma^2}{2\pi}\frac{\theta(e^{i\omega})\theta(e^{-i\omega})}{\phi(e^{i\omega})\phi(e^{-i\omega})}, \quad \omega \in [0, 2\pi]$$

  Plugging in for the ARMA(1,1) case yields

$$f_x(\omega) = \frac{\sigma^2}{2\pi}\frac{(e^{i\omega} - \theta e^{i\omega})(e^{-i\omega} - \theta e^{-i\omega})}{(e^{i\omega} - \phi e^{i\omega})(e^{-i\omega} - \phi e^{-i\omega})} = \frac{\sigma^2}{2\pi}\frac{1 - 2\theta + \theta^2}{1 - 2\phi + \phi^2}$$

- If $\phi = \theta$, the ARMA(1,1) process becomes a white noise process. We can see this two ways. The ARMA(1, 1) process can be represented in the following way:

$$(1 - \phi L)y_t = (1 - \theta L)\epsilon_t$$

Therefore $\phi(L) = \theta(L)$ yields $y_t = \epsilon_t$.

We can also see that when $\phi = \theta$, an ARMA(1,1) process is equivalent to a white noise process as follows. Plugging in $\phi = \theta$ to the spectral density function, we have

$$f_x(\omega) = \frac{\sigma^2}{2\pi}\frac{1 - 2\theta + \theta^2}{1 - 2\theta + \theta^2} = \frac{\sigma^2}{2\pi}$$

showing that if $\theta = \phi$, the spectral density function is constant and independent of θ and ϕ. We can see that it in fact is a white noise process. Since a white noise process has the following covariances:

$$\gamma(0) = \sigma^2$$

$$\gamma(h) = 0, \quad \forall h \neq 0$$

for a white noise process we have

$$f_x(\omega) = \frac{1}{2\pi} \cdot \sigma^2 = \frac{\sigma^2}{2\pi}$$

1.4 Chapter 14: Estimation of Stationary Time Series Processes

1.4.1 Sufficient conditions for ergodicity of mean. (Book section 14.2.1)

By Chebyshev's Inequality (see section **??**), $\overline{y}_T$ is a consistent estimator of μ as $T \to \infty$ if $\lim_{T\to\infty} \mathbb{E}(\overline{y}_T) = \mathbb{E}(y_T) = \mu$ and $\lim_{T\to\infty} \text{Var}(\overline{y}_T) = 0$. We have

$$\mathbb{E}(\overline{y}_T) = \frac{1}{T}\mathbb{E}\bigg(\sum_{t=1}^{T} y_t\bigg) = \frac{1}{T}\sum_{t=1}^{T}\mathbb{E}(y_t) = \mu$$

$$\text{Var}(\overline{y}_T) = \frac{1}{T^2}\text{Var}\bigg(\sum_{t=1}^{T} y_t\bigg) = \frac{1}{T^2}\bigg(\sum_{t=1}^{T}\text{Var}(y_t) + 2\sum_{0\le i<j\le T}\text{Cov}(y_i, y_j)\bigg)$$

$$= \frac{1}{T^2}\bigg(\sum_{t=1}^{T}\gamma(0) + 2\sum_{0\le i<j\le T}\gamma(j-i)\bigg) = \frac{1}{T^2}\bigg(T\gamma(0) + 2\sum_{h=1}^{T-1}(T-h)\gamma(h)\bigg)$$

$$= \frac{1}{T}\bigg[\gamma(0) + 2\sum_{h=1}^{T-1}\bigg(1 - \frac{h}{T}\bigg)\gamma(h)\bigg] = \frac{1}{T^2}\mathbf{1}\text{Var}(\boldsymbol{y})\mathbf{1}'$$

where $\mathbf{1}$ is a vector of ones and

$$\mathrm{Var}(\boldsymbol{y}) = \begin{pmatrix} \gamma(0) & \gamma(1) & \cdots & \gamma(T-2) & \gamma(T-1) \\ \gamma(1) & \gamma(0) & \cdots & \gamma(T-3) & \gamma(T-2) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \gamma(T-2) & \gamma(T-3) & \cdots & \gamma(0) & \gamma(1) \\ \gamma(T-1) & \gamma(T-2) & \cdots & \gamma(1) & \gamma(0) \end{pmatrix}$$

Notice that

$$\left| \gamma(0) + 2\sum_{h=1}^{T-1} \left(1 - \frac{h}{T}\right)\gamma(h) \right| < \left| 2\sum_{h=0}^{T-1} \gamma(h) \right| \leq 2\sum_{h=0}^{T-1} |\gamma(h)|$$

Therefore

$$\sum_{h=0}^{T-1} |\gamma(h)| < \infty$$

is a sufficient condition for

$$\lim_{T\to\infty} \mathrm{Var}(\overline{y}_T) = \lim_{T\to\infty} \frac{1}{T}\left[\gamma(0) + 2\sum_{h=1}^{T-1}\left(1 - \frac{h}{T}\right)\gamma(h)\right] = 0$$

1.4.2 Estimation of autocovariances (Book section 14.2.2).

A moment estimator of $\gamma(h) = \mathbb{E}[(y_t - \mu)(y_{t-h} - \mu)]$ is

$$\hat{\gamma}(h) = \frac{1}{T}\sum_{t=h+1}^{T} (y_t - \overline{y}_T)(y_{t-h} - \overline{y}_T)$$

By Chebyshev's Inequality (see section **??**), $\hat{\gamma}(h)$ is a consistent estimator of $\gamma(h)$ as $T \to \infty$ if $\lim_{T\to\infty} \mathbb{E}(\hat{\gamma}(h)) = \gamma(h)$ and $\lim_{T\to\infty} \mathrm{Var}(\hat{\gamma}(h)) = 0$.

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

$$= \frac{1}{T}\sum_{t=h+1}^{T} (y_t - \mu)(y_{t-h} - \mu) + (y_t - \mu)(\mu - \overline{y}_T) + (\mu - \overline{y}_T)(y_{t-h} - \mu) + (\mu - \overline{y}_T)^2$$

$$= \frac{1}{T}\sum_{t=h+1}^{T} (y_t - \mu)(y_{t-h} - \mu) + (\mu - \overline{y}_T)\frac{1}{T}\sum_{t=h+1}^{T} (y_t - \mu) + (\mu - \overline{y}_T)\frac{1}{T}\sum_{t=h+1}^{T} (y_{t-h} - \mu) + \frac{1}{T}(T-h)(\mu - \overline{y}_T)^2$$

$$\vdots$$

Because **where does this line come from? on page 300 of book/331 of pdf.**

$$\overline{y}_T = \mu + \mathcal{O}_p(T^{-1/2})$$

and for any fixed h

$$T^{-1/2} \sum_{t=h+1}^{T} (y_t - \mu) = \mathcal{O}_p(1)$$

it follows that

$$(\mu - \overline{y}_T)\frac{1}{T}\sum_{t=h+1}^{T}(y_t - \mu) = \frac{\mu}{T}\sum_{t=h+1}^{T}(y_t - \mu) - \frac{\overline{y}_T}{\sqrt{T}} \cdot \frac{1}{\sqrt{T}}\sum_{t=h+1}^{T}(y_t - \mu) = \mathcal{O}_p(T^{-1})$$

$$(\mu - \overline{y}_T)\frac{1}{T}\sum_{t=h+1}^{T}(y_{t-h} - \mu) = \mathcal{O}_p(T^{-1})$$

$$\frac{1}{T}(T-h)(\mu - \overline{y}_T)^2 = (\mu - \overline{y}_T)^2 - \frac{h}{T}(\mu - \overline{y}_T)^2 = \mathcal{O}_p(T^{-1})$$

$$\implies \hat{\gamma}(h) = \frac{1}{T}\sum_{t=h+1}^{T}(y_t - \mu)(y_{t-h} - \mu) + \mathcal{O}_p(T^{-1})$$

which implies that $\lim_{T\to\infty} \mathbb{E}(\hat{\gamma}(h)) = \gamma(h)$. Also using results in Bartlett (1946) **where? do we need to know how to do this?** we have

$$\lim_{T\to\infty} \text{Var}(\hat{\gamma}_T(h) - \gamma(h)) = 0$$

under the assumption that

$$\lim_{H\to\infty} H^{-1}\sum_{h=1}^{H}\gamma_h^2 \to 0$$

1.4.3 Worked examples

Midterm Problem 2 part (2) (similar to exercise 1 in chapter 14. Suppose $\{y_t\}$ has the following general linear process

$$y_t = \mu + \alpha(L)\epsilon_t, \qquad \epsilon_t \sim i.i.d.\ (0, \sigma^2)$$

where $\alpha(L) = \alpha_0 + \alpha_1 L + \alpha_2 L^2 + \ldots;\ \alpha_0 = 1$. Let

$$\overline{y}_T = \frac{1}{T}\sum_{t=1}^{T} y_t$$

$$\gamma(h) = \mathbb{E}[(y_t - \mu)(y_{t-h} - \mu)]$$

$$\hat{\gamma}(h) = \frac{1}{T}\sum_{t=h+1}^{T} (y_t - \overline{y}_T)(y_{t-h} - \overline{y}_T)$$

Derive the conditions under which

(a) $\overline{y}_T$ is a consistent estimator of μ as $T \to \infty$

(b) For fixed h, $\hat{\gamma}(h)$ is a consistent estimator of $\gamma(h)$ as $T \to \infty$.

Solution.

(a) This is an MA(∞) process. By Chebyshev's Inequality, $\overline{y}_T$ is a consistent estimator of μ as $T \to \infty$ if $\lim_{T\to\infty} \mathbb{E}(\overline{y}_T) = \mathbb{E}(y_T) = \mu$ and $\lim_{T\to\infty} \operatorname{Var}(\overline{y}_T) = 0$. In this case in particular (MA(∞) process), we can write

$$\overline{y}_T = \frac{1}{T}\sum_{t=1}^{T}\big(\mu + \alpha(L)\epsilon_t\big) = \frac{1}{T}\cdot T\mu + \frac{1}{T}\sum_{t=1}^{T}\alpha(L)\epsilon_t = \mu + \frac{1}{T}\sum_{t=1}^{T}\alpha(L)\epsilon_t$$

Then we have

$$\mathbb{E}(\overline{y}_T) = \mu + \frac{1}{T}\mathbb{E}\left(\sum_{t=1}^{T}\alpha(L)\epsilon_t\right) = \mu + \frac{1}{T}\sum_{t=1}^{T}\mathbb{E}(\alpha(L)\epsilon_t) = \mu$$

$$\operatorname{Var}(\overline{y}_T) = 0 + \frac{1}{T^2}\operatorname{Var}\left(\sum_{t=1}^{T}\alpha(L)\epsilon_t\right) = \frac{1}{T^2}\sum_{t=1}^{T}\operatorname{Var}[\alpha(L)\epsilon_t] = \frac{1}{T^2}\sum_{t=1}^{T}\mathbb{E}[\alpha(L)\epsilon_t]^2 = \frac{1}{T}\alpha(1)^2\mathbb{E}[\epsilon_t]^2$$

$$= \frac{\sigma^2}{T}\alpha(1)^2$$

Therefore a sufficient condition for consistency is

$$\lim_{T\to\infty}\frac{\sigma^2}{T}\alpha(1)^2 = 0 \iff \alpha(1)^2 < \infty \impliedby \boxed{\sum_{i=0}^{\infty}\alpha_i = 0}$$

(b) **ask about the derivation** Per the derivation in section 1.4.2, we have that

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

For $\hat{\gamma}(h)$ to be consistent, we need

$$\frac{1}{T}\sum_{t=h_1}^{T}(y_t-\mu)(y_{t-h}-\mu) \xrightarrow{p} \gamma(h) \iff \lim_{T\to\infty} \Pr(|\hat{\gamma}(h)-\gamma(h)|<\epsilon)=1, \text{ for every } \epsilon>0$$

First we show that $(y_t-\mu)(y_{t-h}-\mu)$ is a martingale difference process:

$$\mathbb{E}[(y_t-\mu)(y_{t-h}-\mu) \mid F_{t-h}] = (y_{t-h}-\mu)\mathbb{E}[y_t-\mu \mid F_{t-h}] = 0$$

why? which theorem is being used to prove this result? ask We need to show that

$$\mathbb{E}[(y_t-\mu)^2(y_{t-h}-\mu)^2] = \mathbb{E}\left[\left(\sum_{i=0}^{\infty}\alpha_i\epsilon_{t-i}\right)^2\left(\sum_{j=0}^{\infty}\alpha_j\epsilon_{t-h-j}\right)^2\right] < \infty$$

By the Cauchy-Schwarz Inequality, we have

$$\mathbb{E}\left|\left[\left(\sum_{i=0}^{\infty}\alpha_i\epsilon_{t-i}\right)^2\left(\sum_{j=0}^{\infty}\alpha_j\epsilon_{t-h-j}\right)^2\right]\right|^2 \le \mathbb{E}\left[\left(\sum_{i=0}^{\infty}\alpha_i\epsilon_{t-i}\right)^2\right]^2\mathbb{E}\left[\left(\sum_{j=0}^{\infty}\alpha_j\epsilon_{t-h-j}\right)^2\right]^2$$

$$<\infty \iff \mathbb{E}\left[\sum_{i=0}^{\infty}\alpha_i\epsilon_{t-i}\right]^4<\infty, \quad \mathbb{E}\left[\sum_{j=0}^{\infty}\alpha_j\epsilon_{t-h-j}\right]^4<\infty$$

These conditions hold if $\mathbb{E}(\epsilon_t^4)<\infty$ and $\sum_{i=0}^{\infty}|\alpha_i|<\infty$. Then $\mathbb{E}[(y_t-\mu)^2(y_{t-h}-\mu)^2]<\infty$ holds and

$$\hat{\gamma} \xrightarrow{p} \gamma(h)$$

Midterm Problem 3 parts (3) and (4) (similar to 14.7 and 14.8 material. Consider the following ARMA(1, 1) model

$$y_t = \phi y_{t-1} + u_t + \theta u_{t-1}, \text{ for } t=-\infty,\dots,-1,0,1,\dots$$

where $|\theta|<1$, $|\phi|<1$, and u_t is i.i.d. with mean zero and variance σ_u^2, $\mathbb{E}(u_t^4)<\infty$.

(1) Suppose that we have the data $\{y_t : t=0,1,\dots,T\}$. Consider the following estimator of ϕ:

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

Show that $\hat{\phi}$ is a consistent estimator of ϕ and derive the asymptotic distribution of $\sqrt{T}(\hat{\phi}_T-\phi)$. Comment on the case where $\theta=\phi$.

(2) Suppose that $\sigma_u^2=1$ is known. Show that θ can be consistently estimated by

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

Solution.

(1) From the results in Question 2 part 2(b), since $\mathbb{E}(y_t) = \mathbb{E}(y_{t-1}) = \mathbb{E}(y_{t-2}) = 0$, we know that

$$\hat{\phi}_T = \frac{\sum_{t=2}^T y_t y_{t-2}}{\sum_{t=2}^T y_{t-1} y_{t-2}} = \frac{T^{-1}\sum_{t=2}^T y_t y_{t-2}}{T^{-1}\sum_{t=2}^T y_{t-1} y_{t-2}} \xrightarrow{p} \frac{\gamma(2)}{\gamma(1)}$$

By the result from Question 3 part (2), we have $\gamma(h) = \phi\gamma(h-1)$ for $h \geq 2$. Therefore $\gamma(2)/\gamma(1) = \phi$, so $\hat{\phi}_T$ is a consistent estimator for ϕ. To obtain the asymptotic distribution, note that

$$\sqrt{T}(\hat{\phi}_T - \phi) = \sqrt{T}\left(\frac{T^{-1}\sum_{t=2}^T y_t y_{t-2}}{T^{-1}\sum_{t=2}^T y_{t-1} y_{t-2}} - \phi\right)$$

$$= \frac{T^{-1/2}\sum_{t=2}^T (\phi y_{t-1} + u_t + \theta u_{t-1}) y_{t-2}}{T^{-1}\sum_{t=2}^T y_{t-1} y_{t-2}} - \frac{\phi T^{-1/2}\sum_{t=2}^T y_{t-1} y_{t-2}}{T^{-1}\sum_{t=2}^T y_{t-1} y_{t-2}}$$

$$= \frac{T^{-1/2}\sum_{t=2}^T (u_t + \theta u_{t-1}) y_{t-2}}{T^{-1}\sum_{t=2}^T y_{t-1} y_{t-2}}$$

In Question 2 part 2(b), we showed that

$$\frac{1}{T}\sum_{t=h_1}^T (y_t - \mu)(y_{t-h} - \mu) \xrightarrow{p} \gamma(h)$$

Therefore in the denominator, since $\mathbb{E}(y_{t-1}) = \mathbb{E}(y_{t-h})_{=}0$, we have

$$T^{-1}\sum_{t=2}^T y_{t-1} y_{t-2} \xrightarrow{p} \gamma(1)$$

In the numerator,

$$T^{-1/2}\sum_{t=2}^T (u_t + \theta u_{t-1}) y_{t-2} = \frac{1}{\sqrt{T}}\sum_{t=2}^T \left[u_t y_{t-2} + \theta u_{t-1} y_{t-2}\right]$$

$$= \frac{1}{\sqrt{T}}\sum_{t=2}^T u_t y_{t-2} + \frac{1}{\sqrt{T}}\sum_{t=2}^T \theta u_{t-1} y_{t-2} = \frac{1}{\sqrt{T}}\left(\sum_{t=2}^{T-1} u_t y_{t-2} + u_T y_{T-2}\right) + \frac{1}{\sqrt{T}}\sum_{t'=1}^{T-1} \theta u_{t'} y_{t'-1}$$

$$= \frac{1}{\sqrt{T}}\left(\sum_{t=2}^{T-1} u_t y_{t-2} + u_T y_{T-2}\right) + \frac{1}{\sqrt{T}}\left(\theta u_1 y_0 + \sum_{t=2}^{T-1} \theta u_t y_{t-1}\right) = \frac{1}{\sqrt{T}}\left(\sum_{t=2}^{T-1} u_t (y_{t-2} + \theta y_{t-1}) + \theta u_1 y_0 + u_T y_{T-2}\right)$$

Since $\mathbb{E}(u_t(y_{t-2} + \theta y_{t-1}) \mid F_{t-1}) = 0$. Further, $T^{-1/2}(\theta u_1 y_0 + u_{T-1} y_{T-2}) = o_p(1)$. Then by the Central Limit Theorem in martingale difference processes (see section **??**):

Theorem 28 (Central limit theorem for martingale difference sequences). Let $\{x_t\}$ be a martingale difference sequence with respect to the information set Ω_t. Let $\overline{\sigma}_T^2 = \text{Var}(\sqrt{T}\overline{x}_T) = T^{-1}\sum_{t=1}^T \sigma_t^2$. If $\mathbb{E}(|x_t|^r) < K < \infty$, $r > 2$ and for all t, and

$$\frac{1}{T}\sum_{t=1}^T x_t^2 - \overline{\sigma}_t^2 \xrightarrow{p} 0$$

then

$$\sqrt{T} \cdot \frac{\overline{x}_T}{\overline{\sigma}_T} \xrightarrow{d} \mathcal{N}(0,1)$$

we have

$$\sqrt{T} \cdot \frac{\overline{x}_T}{T^{-1/2}\sqrt{\sum_{t=1}^T \sigma_t^2}} \xrightarrow{d} \mathcal{N}(0,1)$$

$$\vdots$$

$$\frac{1}{\sigma^2}\frac{\gamma(1)^2}{(1+\theta)^2)\gamma(0) + 2\theta\gamma(1)}\sqrt{T}(\hat{\phi}_T - \phi) \xrightarrow{d} \mathcal{N}(0,1)$$

$$\iff \sqrt{T}(\hat{\phi}_T - \phi) \xrightarrow{d} \mathcal{N}\left(0, \sigma^2 \frac{(1+\theta)^2)\gamma(0) + 2\theta\gamma(1)}{\gamma(1)^2}\right)$$

(2) From the results of Question 2 part 2(b), where we showed that

$$\frac{1}{T}\sum_{t=h_1}^T (y_t - \mu)(y_{t-h} - \mu) \xrightarrow{p} \gamma(h)$$

(and since $\mathbb{E}(y_{t-1}) = \mathbb{E}(y_{t-h})$=0,)

$$T^{-1}\sum_{t=2}^T y_t y_{t-1} \xrightarrow{p} \gamma(1), \quad T^{-1}\sum_{t=2}^T y_{t-1}^2 \xrightarrow{p} \gamma(0)$$

and by the law of large numbers (see section **??**) **(why?)**, we have

$$\hat{\theta}_T = \frac{1}{T}\sum_{t=1}^T y_t y_{t-1} - \frac{\hat{\phi}_T}{T}\sum_{t=1}^T y_{t-1}^2 \xrightarrow{p} \gamma(1) - \phi\gamma(0) = \phi\gamma(0) + \theta\sigma^2 - \phi\gamma(0) = \theta$$

1.5 Chapter 17: Introduction to Forecasting

1.5.1 17.7: Iterated and direct multi-step AR methods

Suppose y_t follows the AR(1) model:

$$y_t = a + \phi y_{t-1} + \epsilon_t, \quad |\phi| < 1, \epsilon_t \sim iid(0, \sigma_\epsilon^2) \tag{1}$$

$$\iff y_t = \frac{a}{1-\phi} + \sum_{i=0}^{\infty} \phi^i \epsilon_{t-i}$$

$$\iff y_t = a\left(\frac{1-\phi^h}{1-\phi}\right) + \phi^h y_{t-h} + \sum_{j=0}^{h-1} \phi^j \epsilon_{t-j} \tag{2}$$

We have two methods for forecasting y_{t+h} $h > 1$ steps ahead.

(1) **Iterated method:** In this method, we first calculate the OLS estimates of $\hat{a}_T$ and $\hat{\phi}_T$ in Equation (1) using all available data Ω_T. Then we use the form of Equation (2):

$$\hat{y}^*_{T+h|T} = \hat{a}_T\left(\frac{1-\hat{\phi}_T^h}{1-\hat{\phi}_T}\right) + \hat{\phi}_T^h y_T$$

(2) **Direct method:** We directly calculate OLS estimates of the parameters in Equation (2) using all available data Ω_T:

$$\tilde{y}^*_{T+h|T} = \tilde{a}_{h,T} + \tilde{\phi}_{h,T} y_T$$

Proposition 45. Suppose data is generated by Equation (1). If $u_t = \sum_{i=0}^{\infty} \phi^i \epsilon_{t-i}$ and $v_t = \sum_{j=0}^{h-1} \phi^j \epsilon_{t-j}$ are symmetrically distributed around zero and have finite second moments, and if $\mathbb{E}(\hat{\phi}_T$ and $\mathbb{E}(\tilde{\phi}_{h,T})$ exist, then for any finite T and h we have

$$\mathbb{E}(\hat{y}^*_{T+h|T} - y_{T+h}) = \mathbb{E}(\tilde{y}^*_{T+h|T} - y_{T+h}) = 0$$