Math Review Notes—Time Series

Gregory Faletto

Contents

Last updated December 9, 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.

Definition 1.1. 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 $\gamma(t_1, t_2) = \gamma(|t_1 - t_2|)$.

Definition 1.2. 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.

1.2.1 Worked Examples

Midterm Problem 2 part (1) (chapter 12 exercise 6)

Midterm Problem 2 part (2) (exercise 7 in chapter 12; similar to exercise 1 in chapter 14. Suppose $\{y_t\}$ has the following general linear process

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

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

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

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

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

Derive the conditions under which

(a) $\bar{y}_T$ 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 (Theorem **??**), $\bar{y}_T$ is a consistent estimator of μ as $T \to \infty$ if $\lim_{T \to \infty} \mathbb{E}(\bar{y}_T) = \mathbb{E}(y_T) = \mu$ and $\lim_{T \to \infty} \text{Var}(\bar{y}_T) = 0$. In this case in particular (MA(∞) process), we can write

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

Then we have

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

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

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

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) See section 1.4.2.

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:

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

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:

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

$$\gamma_h = \frac{\sigma^2\phi^h}{1 - \rho^2} = \phi^h\gamma(0) \ \ \forall \ h \geq 1$$

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

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

- 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

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

- 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:

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

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

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

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

- 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:

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

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

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

- 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 $\text{ARMA}(p,q)$ 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 **??**), $\bar{y}_T$ is a consistent estimator of μ as $T \to \infty$ if $\lim_{T \to \infty} \mathbb{E}(\bar{y}_T) = \mathbb{E}(y_T) = \mu$ and $\lim_{T \to \infty} \operatorname{Var}(\bar{y}_T) = 0$. We have

$$\mathbb{E}(\bar{y}_T) = \frac{1}{T}\mathbb{E}\left(\sum_{t=1}^{T} y_t\right) = \frac{1}{T}\sum_{t=1}^{T}\mathbb{E}(y_t) = \mu$$

$$\operatorname{Var}(\bar{y}_T) = \frac{1}{T^2}\operatorname{Var}\left(\sum_{t=1}^{T} y_t\right) = \frac{1}{T^2}\left(\sum_{t=1}^{T}\operatorname{Var}(y_t) + 2\sum_{0 \leq i < j \leq T}\operatorname{Cov}(y_i, y_j)\right)$$

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

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

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

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

Notice that

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

Therefore

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

is a sufficient condition for

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

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 (Theorem **??**), $\hat{\gamma}(h)$ is a consistent estimator of $\gamma(h)$ as $T \to \infty$ if $\lim_{T \to \infty}\mathbb{E}(\hat{\gamma}(h)) = \gamma(h)$ and $\lim_{T \to \infty}\text{Var}(\hat{\gamma}(h)) = 0$.

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

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

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

$$\vdots$$

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

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

and for any fixed h

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

it follows that

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

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

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

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

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

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

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

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

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 (Theorem **??**), we have

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

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

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

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

1.4.3   Worked examples

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

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

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 $\{y_t : t = 0, 1, \ldots, T\}$. 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) (in section 1.2.1), 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) (in section 1.2.1), 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 (Theorem (**??**)):

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 $\overline{\sigma}_T^2 = \mathrm{Var}(\sqrt{T}\overline{x}_T) = T^{-1}\sum_{t=1}^T \sigma_t^2$. 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) (in section 1.2.1), 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 Weak Law of Large Numbers (Theorem **??**) we have

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

Chapter 14 Question 3. The time series $\{y_t\}$ and $\{x_t\}$ are independently generated according to the following schemes:

$$y_t = \lambda y_{t-1} + \epsilon_{1t}, \quad |\lambda| < 1$$

$$x_t = \rho x_{t-1} + \epsilon_{2t}, \quad |\rho| < 1$$

for $t = 1, 2, \ldots, T$, where ϵ_{1t} and ϵ_{2t} are non-autocorrelated and distributed independently of each other with zero means and variances equal to σ_1^2 and σ_2^2 respectively. An investigator estimates the simple regression

$$y_t = \beta x_t + u_t \quad t = 1, 2, \ldots T$$

by the OLS method. Show that

(a) $\hat{\beta} \xrightarrow{P} 0$ as $T \to \infty$

(b)
$$t_{\hat{\beta}}^2 = \frac{\hat{\beta}^2}{\widehat{\text{Var}}(\hat{\beta})} = \frac{(T-1)r^2}{1 - r^2}$$

(c)
$$Tr^2 \xrightarrow{P} \frac{1 + \lambda\rho}{1 - \lambda\rho} \text{ as } T \to \infty$$

where $\hat{\beta}$ is the OLS estimator of β, $\widehat{\text{Var}}(\hat{\beta})$ is the estimated variance of $\hat{\beta}$, and r is the sample correlation coefficient between x and y, i.e.

$$r^2 = \left(\sum_{t=1}^{T} x_t y_t \right)^2 \bigg/ \left(\sum_{t=1}^{T} x_t^2 \sum_{t=1}^{T} y_t^2 \right)$$

What are the implications of these results for problems of spurious correlation in economic time series analysis?

Solution.

(a)
$$\hat{\beta} = \frac{\sum_{t=1}^{T} x_t y_t}{\sum_{t=1}^{T} x_t^2} = \frac{T^{-1} \sum_{t=1}^{T} x_t y_t}{T^{-1} \sum_{t=1}^{T} x_t^2}$$

By the Weak Law of Large Numbers (Theorem **??**), we have

$$T^{-1} \sum_{t=1}^{T} x_t y_t = T^{-1} \sum_{t=1}^{T} x_t y_t \xrightarrow{P} \mathbb{E}(x_t y_t) = \text{Cov}(x_t, y_t) = 0$$

because of the i.i.d. distributions of ϵ_1 and ϵ_{2t}. By the Law of Large Numbers (Theorem **??**),

$$T^{-1} \sum_{t=1}^{T} x_t^2 \xrightarrow{a.s.} \mathbb{E}(x_t^2) = \gamma_x(0) > 0$$

Therefore

$$\hat{\beta} \xrightarrow{P} \frac{0}{\gamma_x(0)} = 0$$

(b)

$$\widehat{\mathrm{Var}}(\hat{\beta}) = \frac{\hat{\sigma}^2}{S_{XX}}$$

where

$$\hat{\sigma}^2 = \frac{1}{T-1}\sum_{t=1}^{T}(y_t - \hat{\beta}x_t)^2 = \frac{1}{T-1}\sum_{t=1}^{T}\left(y_t^2 - 2y_t\hat{\beta}x_t + \hat{\beta}^2 x_t^2\right)$$

$$= \frac{1}{T-1}\sum_{t=1}^{T}y_t^2 - \frac{1}{T-1}\sum_{t=1}^{T}\left(2y_tx_t\frac{\sum_{t=1}^{T}x_ty_t}{\sum_{t=1}^{T}x_t^2}\right) + \frac{1}{T-1}\sum_{t=1}^{T}\left(\left[\frac{\sum_{t=1}^{T}x_ty_t}{\sum_{t=1}^{T}x_t^2}\right]^2 x_t^2\right)$$

$$= \frac{1}{T-1}\sum_{t=1}^{T}y_t^2 - 2\cdot\frac{1}{T-1}\frac{\left(\sum_{t=1}^{T}x_ty_t\right)^2}{\sum_{t=1}^{T}x_t^2} + \frac{1}{T-1}\frac{\left(\sum_{t=1}^{T}x_ty_t\right)^2}{\sum_{t=1}^{T}x_t^2}$$

$$= \frac{1}{T-1}\left(\sum_{t=1}^{T}y_t^2 - \frac{\left(\sum_{t=1}^{T}x_ty_t\right)^2}{\sum_{t=1}^{T}x_t^2}\right)$$

$$\implies \hat{t}^2 = \frac{\left(\sum_{t=1}^{T}x_ty_t\right)^2}{\left(\sum_{t=1}^{T}x_t^2\right)^2}\cdot\sum_{t=1}^{T}x_t^2\cdot(T-1)\bigg/\left(\sum_{t=1}^{T}y_t^2 - \frac{\left(\sum_{t=1}^{T}x_ty_t\right)^2}{\sum_{t=1}^{T}x_t^2}\right)$$

$$= \frac{\left(\sum_{t=1}^{T}x_ty_t\right)^2}{\sum_{t=1}^{T}x_t^2}\cdot(T-1)\bigg/\left(\sum_{t=1}^{T}y_t^2 - \frac{\left(\sum_{t=1}^{T}x_ty_t\right)^2}{\sum_{t=1}^{T}x_t^2}\right)$$

$$= \left(\sum_{t=1}^{T}x_ty_t\right)^2\cdot(T-1)\bigg/\left(\sum_{t=1}^{T}y_t^2\sum_{t=1}^{T}x_t^2 - \left(\sum_{t=1}^{T}x_ty_t\right)^2\right)$$

Note that

$$r^2 = \left(\sum_{t=1}^{T}x_ty_t\right)^2\bigg/\left(\sum_{t=1}^{T}x_t^2\sum_{t=1}^{T}y_t^2\right), \quad 1-r^2 = \left[\sum_{t=1}^{T}x_t^2\sum_{t=1}^{T}y_t^2 - \left(\sum_{t=1}^{T}x_ty_t\right)^2\right]\bigg/\left(\sum_{t=1}^{T}x_t^2\sum_{t=1}^{T}y_t^2\right)$$

$$\implies \frac{r^2}{1-r^2} = \left(\sum_{t=1}^{T}x_ty_t\right)^2\bigg/\left[\sum_{t=1}^{T}x_t^2\sum_{t=1}^{T}y_t^2 - \left(\sum_{t=1}^{T}x_ty_t\right)^2\right]$$

Therefore

$$\hat{t}^2 = \left(\sum_{t=1}^{T}x_ty_t\right)^2\cdot(T-1)\bigg/\left(\sum_{t=1}^{T}y_t^2\sum_{t=1}^{T}x_t^2 - \left(\sum_{t=1}^{T}x_ty_t\right)^2\right) = \boxed{\frac{(T-1)r^2}{1-r^2}}$$

(c)

$$Tr^2 = T\cdot\frac{\left(T^{-1}\sum_{t=1}^{T}x_ty_t\right)^2}{T^{-1}\sum_{t=1}^{T}x_t^2\cdot T^{-1}\sum_{t=1}^{T}y_t^2} = \frac{T^{-1}\left(2\sum_{1\le i<j\le T}x_ix_jy_iy_j + \sum_{t=1}^{T}x_t^2y_t^2\right)}{T^{-1}\sum_{t=1}^{T}x_t^2\cdot T^{-1}\sum_{t=1}^{T}y_t^2}$$

$$= \frac{2T^{-1}\sum_{1\le i<j\le T}x_ix_jy_iy_j + T^{-1}\sum_{t=1}^{T}x_t^2y_t^2}{T^{-1}\sum_{t=1}^{T}x_t^2\cdot T^{-1}\sum_{t=1}^{T}y_t^2}$$

Note that

$$2T^{-1} \sum_{1 \le i < j \le T} (x_i x_j)(y_i y_j) = \sum_{h=1}^{T-1} 2T^{-1}(T-h)(x_t x_{t-h})(y_t y_{t-h})$$

By the Weak Law of Large Numbers (Theorem **??**) and the independence of x_t and y_t,

$$\sum_{h=1}^{T-1} 2T^{-1}(T-h)(x_t x_{t-h})(y_t y_{t-h}) \approx 2T^{-1} \sum_{h=1}^{T-1} (x_t x_{t-h})(y_t y_{t-h}) \xrightarrow{p} 2 \sum_{h=1}^{\infty} \gamma_x(h)\gamma_y(h)$$

Recall that for an AR(1) process with coefficient ϕ, $\gamma(h) = \phi^h \gamma(0)$. (See section 1.3.4).

$$= 2 \sum_{h=1}^{\infty} \rho^h \gamma_x(0) \lambda^h \gamma_y(0) = 2\gamma_x(0)\gamma_y(0) \sum_{h=1}^{\infty} (\rho\lambda)^h$$

By the Weak Law of Large Numbers (Theorem **??**) and the independence of x_t and y_t,

$$T^{-1} \sum_{t=1}^{T} x_t^2 y_t^2 \xrightarrow{a.s.} \mathbb{E}(x_t^2 y_t^2) = \mathbb{E}(x_t^2)\mathbb{E}(y_t^2) = \mathrm{Var}(x_t)\mathrm{Var}(y_t) = \sigma_1^2 \sigma_2^2 = \gamma_x(0)\gamma_y(0)$$

Therefore in the numerator, we have

$$2T^{-1} \sum_{1 \le i < j \le T} x_i x_j y_i y_j + T^{-1} \sum_{t=1}^{T} x_t^2 y_t^2 \xrightarrow{p} 2\gamma_x(0)\gamma_y(0) \sum_{h=1}^{\infty} (\rho\lambda)^h + \gamma_x(0)\gamma_y(0)$$

$$= \gamma_x(0)\gamma_y(0)\left(2\frac{\rho\lambda}{1-\rho\lambda} + 1\right) = \gamma_x(0)\gamma_y(0)\left(\frac{1+\rho\lambda}{1-\rho\lambda}\right)$$

Next we examine the denominator. By the Law of Large Numbers (Theorem **??**),

$$T^{-1} \sum_{t=1}^{T} x_t^2 \xrightarrow{a.s.} \mathbb{E}(x_t^2) = \mathrm{Var}(x_t) = \gamma_x(0) = \sigma_2^2$$

By the Law of Large Numbers (Theorem **??**),

$$T^{-1} \sum_{t=1}^{T} y_t^2 \xrightarrow{a.s.} \mathbb{E}(y_t^2) = \gamma_y(0) = \sigma_1^2$$

Therefore in the denominator, we have

$$T^{-1} \sum_{t=1}^{T} x_t^2 \cdot T^{-1} \sum_{t=1}^{T} y_t^2 \xrightarrow{a.s.} \gamma_x(0)\gamma_y(0) = \sigma_1^2 \sigma_2^2$$

This yields

$$\boxed{Tr^2 \xrightarrow{p} \frac{1+\rho\lambda}{1-\rho\lambda}}$$

Lidan's explanation: Because as $T \to \infty$

$$t_{\hat{\beta}} \to \sqrt{Tr^2} = \sqrt{\frac{1 + \lambda\rho}{1 - \lambda\rho}}$$

so if $\lambda\rho \approx 1$, at very high probability we will reject the null $\hat{\beta} = 0$ when in fact $\hat{\beta} \xrightarrow{p} 0$.

My original explanation: Because $\hat{\beta}$ converges in probability to 0 and $t_{\hat{\beta}}^2$ is proportional to r^2 (which we would expect to be close to 0), this suggests that a regression of uncorrelated variables should result in an insignificant $\hat{\beta}$, but if r^2 is high due to a spurious correlation, the $\hat{\beta}$ could be found to be statistically significant even if there is no meaningful relationship between x and y.

1.5 Chapter 15: Unit Root Processes

Need to review concepts in this chapter.

1.5.1 Worked Problems

Problem 3. Suppose that a time series of interest can be decomposed into a deterministic trend, a random walk component, and stationary errors:

$$y_t = \alpha + \delta t + \gamma_t + v_t \tag{1}$$

$$\gamma_t = \gamma_{t-1} + u_t$$

with

$$v_t \sim iid \ \mathcal{N}(0, \sigma_v^2), \quad u_t \sim iid \ \mathcal{N}(0, \sigma_u^2), \quad u_t \perp\!\!\!\perp v_t$$

Let $\lambda = \sigma_u^2 / \sigma_v^2$.

(a) Show that under $\lambda = 0$, y_t reduces to a trend stationary process.

(b) Alternatively, suppose that y_t follows and ARIMA(0,1,1) process of the form

$$y_t = \delta + y_{t-1} + w_t \tag{2}$$

$$w_t = \epsilon_t + \theta\epsilon_{t-1}$$

where ϵ_t are iid $\mathcal{N}(0, \sigma_\epsilon^2)$. In this case show that under $\theta = -1$, y_t is a trend stationary process.

(c) Derive a relation between λ and the MA(1) parameter θ, and hence or otherwise show that a test of $\theta = -1$ in (2) is equivalent to a test of $\lambda = 0$ in (1).

(d) Show that (2) as a characterization of (1) implies $\theta < 0$.

Solution.

(a) $\lambda = 0 \implies \sigma_u^2 = 0 \implies u_t = 0$ (constant)

$$\implies \gamma_t = \gamma_{t-1} + 0 \iff \gamma_t = \gamma_0$$

$$\implies y_t = \alpha + \delta t + \gamma_0 + v_t = (\alpha + \gamma_0) + \delta t + v_t$$

which is trend stationary because $d_t = (\alpha + \gamma_0) + \delta t$ is perfectly predictable, and $y_t - d_t = v_t$ is covariance stationary.

Also note that $\text{Var}(y_t - \delta t) = \sigma_v^2$, $\text{Cov}\big([y_t - \delta t][y_{t-h} - \delta(t-h)]\big) = 0$, which implies trend stationarity of y_t. (Recall Definition 1.2: "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.")

(b) $\theta = -1 \implies w_t = \epsilon_t - \epsilon_{t-1} = (1 - L)\epsilon_t$

$$\implies (1 - L)y_t = \delta + (1 - L)\epsilon_t \iff y_t = (1 - L)^{-1}\delta + \epsilon_t$$

which is trend stationary because $d_t = (1-L)^{-1}\delta$ is perfectly predictable, and $y_t - d_t = \epsilon_t$ is covariance stationary.

(c) **Lidan's solution:** From (1), let

$$z_t = y_t - y_{t-1} = \alpha + \delta t + \gamma_{t-1} + u_t + v_t - (\alpha + \delta(t - 1) + \gamma_{t-1} + v_{t-1})$$

$$= \delta + u_t + v_t - v_{t-1}$$

From (2), let

$$a_t = y_t - y_{t-1} = \delta + y_{t-1} + w_t - y_{t-1} = \delta + \epsilon_t + \theta\epsilon_{t-1}$$

Calculate the autocovariances for each and set them equal (using the serial independence of u_t, v_t, and ϵ_t as well as the independence of u_t, v_t, and ϵ_t for all t):

-
$$\gamma(0) : \text{Var}(z_t) = \text{Var}(a_t) \iff \text{Var}(u_t + v_t - v_{t-1}) = \text{Var}(\epsilon_t + \theta\epsilon_{t-1})$$

$$\iff \sigma_u^2 + 2\sigma_v^2 = \sigma_\epsilon^2(1 + \theta^2) \tag{3}$$

-
$$\gamma(1) : \text{Cov}(z_t, z_{t-1}) = \text{Cov}(a_t, a_{t-1}) \iff \text{Cov}(-v_{t-1}, v_{t-1}) = \text{Cov}(\theta\epsilon_{t-1}, \epsilon_{t-1})$$

$$\iff -\sigma_v^2 = \theta\sigma_\epsilon^2 \tag{4}$$

-
$$\gamma(h) \ (h \geq 2) : 0 = 0$$

Plugging (4) into (3) and using $\lambda = \sigma_u^2/\sigma_v^2$ we have

$$\sigma_u^2 + 2\sigma_v^2 = -\frac{\sigma_v^2}{\theta}(1 + \theta^2) \iff \sigma_u^2 = \sigma_v^2(-1/\theta - \theta - 2) \iff \lambda = -1/\theta - \theta - 2$$

$$\iff \theta^2 + (2 + \lambda)\theta + 1 = 0 \iff \theta = \frac{-(2+\lambda) \pm \sqrt{(2+\lambda)^2 - 4}}{2}$$

$$\iff \boxed{\theta = \frac{-(2+\lambda) \pm \sqrt{\lambda^2 + 4\lambda}}{2}}$$

Clearly if $\lambda = 0$ then $\theta = -1$. The reverse is also true:

$$\theta = -1 \implies \lambda = -1/(-1) - (-1) - 2 = 1 + 1 - 2 = 0$$

Therefore a test of $\theta = -1$ in (2) is equivalent to a test of $\lambda = 0$ in (1).

original solution:

$$y_t = \alpha + \delta t + \gamma_{t-1} + u_t + v_t$$

$$y_{t-1} = \alpha + \delta(t-1) + \gamma_{t-1} + v_{t-1}$$

$$\implies y_t = y_{t-1} + \delta + u_t - v_{t-1} + v_t$$

Comparing this to the second expression, $y_t = \delta + y_{t-1} + \epsilon_t + \theta\epsilon_{t-1}$, they match if $u_t - v_{t-1} + v_t = \epsilon_t + \theta\epsilon_{t-1}$. Since the distributions of ϵ_t and v_t are both i.i.d. normal, this is the case if $\lambda = 0$ (so that $\sigma_u^2 = 0$ and $u_t = 0 \ \forall \ t$), $\theta = -1$, and $\sigma_\epsilon^2 = \sigma_v^2$.

(d) **Lidan's Solution:** By (4) (which follows from (2)),

$$-\sigma_v^2 = \theta\sigma_\epsilon^2 \implies \theta < 0$$

My solution: Again consider (2)

$$y_t = \delta + y_{t-1} + \epsilon_t + \theta\epsilon_{t-1}$$

which is clearly a unit root process with a drift, compared to a re-written version of (1)

$$y_t = \delta + y_{t-1} + u_t + v_t - v_{t-1}$$

which has a strong resemblance to a unit root process with a drift. These match up if $\epsilon_t = u_t + v_t$ and $\theta < 0$.

Problem 4

Homework 4 Problem 3. Let $\{u_t\}$ be an i.i.d. sequence with mean zero and variance σ^2, and let

$$y_t = u_1 + u_2 + \ldots + u_t$$

with $y_0 = 0$.

(a) Show that

$$T^{-3/2} \sum_{t=1}^{T} y_{t-1} = T^{-1/2} \sum_{t=1}^{T-1} u_t - T^{-3/2} \sum_{t=1}^{T-1} t u_t$$

(b) Show that

$$\begin{bmatrix} T^{-1/2} \sum_{t=1}^{T} u_t \\ T^{-3/2} \sum_{t=1}^{T} y_{t-1} \end{bmatrix} \xrightarrow{d} \mathcal{N}\left(0, \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1/3 \end{bmatrix}\right)$$

(c) Use the functional limit theorem (Donsker's Theorem, Theorem **??**) to show that

$$T^{-3/2} \sum_{t=1}^{T} y_{t-1} \xrightarrow{d} \sigma \int_0^1 W(r) dr$$

where $W(\cdot)$ is a standard Brownian motion process.

(d) Use parts (a) - (c) to show that

$$T^{-3/2} \sum_{t=1}^{T} t u_t \xrightarrow{d} \sigma \cdot W(a) - \sigma \int_0^1 W(r) dr$$

Solution.

$$y_t = u_1 + u_2 + \ldots + u_t = \sum_{i=1}^{t} u_i = \sum_{i=0}^{t} u_i$$

where the last equality follows because $y_0 = 0 \iff u_0 = 0$.

$$u_t \sim iid\ (0, \sigma^2), \quad y_0 = 0$$

(a)

$$y_T = \sum_{t=1}^{T} u_t$$

$$T^{-3/2} \sum_{t=1}^{T} y_{t-1} = T^{-3/2} \sum_{t=1}^{T} \left(\sum_{i=0}^{t-1} u_i \right) = T^{-3/2} \sum_{t=1}^{T} (T - (t-1)) u_{t-1} = T^{-3/2} \sum_{t'=0}^{T-1} (T - t') u_{t'}$$

$$= T^{-3/2} \cdot T \sum_{t'=0}^{T-1} u_{t'} - T^{-3/2} \sum_{t'=0}^{T-1} t' u_{t'} = T^{-1/2} \sum_{t'=1}^{T-1} u_{t'} - T^{-3/2} \sum_{t'=1}^{T-1} t' u_{t'}$$

$$\implies \boxed{T^{-3/2} \sum_{t=1}^{T} y_{t-1} = T^{-1/2} \sum_{t'=1}^{T-1} u_{t'} - T^{-3/2} \sum_{t'=1}^{T-1} t' u_{t'}} \tag{5}$$

(b) By the Central Limit Theorem (Theorem **??**), since $u_t \sim (0, \sigma^2)$, we have

$$\frac{1}{\sqrt{T\sigma^2}} \sum_{t=1}^{T} u_t \xrightarrow{d} \mathcal{N}(0, 1)$$

which implies

$$\boxed{\frac{1}{\sqrt{T}} \sum_{t=1}^{T} u_t \xrightarrow{d} \mathcal{N}(0, \sigma^2)} \tag{6}$$

The distribution of $T^{-3/2} \sum_{t=1}^{T} y_{t-1}$ is trickier. Note that from (5),

$$T^{-3/2} \sum_{t=1}^{T} y_{t-1} = T^{-1/2} \left(\sum_{t=0}^{T-1} u_t - \sum_{t=0}^{T-1} \frac{t u_t}{T} \right)$$

Note that $x_t = (t u_t)/T$ is a martingale difference process because

$$\mathbb{E}\left(\frac{t u_t}{T} \mid u_{t-1}, u_{t-2}, \dots \right) = 0$$

and $\mathrm{Var}(x_t) = \frac{1}{T^2} \mathrm{Var}(t u_t) = \frac{t^2}{T^2} \sigma^2 < \infty$. By the Central Limit Theorem in martingale difference processes (Theorem (**??**)):

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 $\bar{\sigma}_T^2 = \mathrm{Var}(\sqrt{T}\bar{x}_T) = T^{-1} \sum_{t=1}^{T} \sigma_t^2$. If $\mathbb{E}(|x_t|^r) < K < \infty$ for any $r > 2$ and for all t, and

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

then

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

---

since

$$T^{-1} \sum_{t=1}^{T} \sigma_t^2 = T^{-1} \sum_{t=1}^{T} \frac{t^2}{T^2} \sigma^2 = \frac{\sigma^2}{T^3} \cdot \frac{T(T+1)(2T+1)}{6} = \frac{\sigma^2}{6} \cdot \frac{2T^2 + 3T + 1}{T^2},$$

$$\mathbb{E}(|x_t|^r) = \mathbb{E}\left(\left| \frac{t u_t}{T} \right|^r \right) = \frac{t^r}{T^r} \mathbb{E}(|u_t|^r)$$

which is finite for $r = 4$ if u_t has finite fourth moment, and since by the Weak Law of Large Numbers (Theorem **??**)

$$T^{-1} \sum_{t=1}^{T} x_t^2 - T^{-1} \sum_{t=1}^{T} \frac{t^2}{T^2} \sigma^2 = T^{-1} \sum_{t=1}^{T} \frac{t^2}{T^2} (u_t^2 - \sigma^2)$$

$$\left| T^{-1} \sum_{t=1}^{T} \frac{t^2}{T^2} (u_t^2 - \sigma^2) \right| \leq \left| T^{-1} \sum_{t=1}^{T} (u_t^2 - \sigma^2) \right| \leq T^{-1} \sum_{t=1}^{T} |u_t^2 - \sigma^2| \xrightarrow{p} 0$$

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

we have (if u_t has finite fourth moment)

$$\sqrt{T} \cdot \frac{1}{T} \left(\sum_{t=1}^{T} \frac{t u_t}{T} \right) \Big/ \sqrt{\frac{\sigma^2}{6} \cdot \frac{2T^2 + 3T + 1}{T^2}} \xrightarrow{d} \mathcal{N}(0,1)$$

$$\iff \sqrt{T} \left(\sum_{t=1}^{T} t u_t \right) \cdot \sqrt{\frac{6}{\sigma^2} \cdot \frac{1}{2T^2 + 3T + 1}} \xrightarrow{d} \mathcal{N}(0,1) \implies \sqrt{T} \left(\sum_{t=1}^{T} t u_t \right) \cdot \sqrt{\frac{6}{\sigma^2} \cdot \frac{1}{2T^2}} \xrightarrow{d} \mathcal{N}(0,1)$$

$$\iff T^{-1/2} \left(\sum_{t=1}^{T} \frac{t u_t}{T} \right) \cdot \sqrt{\frac{3}{\sigma^2}} \xrightarrow{d} \mathcal{N}(0,1) \iff T^{-1/2} \frac{\sqrt{3}}{\sigma} \sum_{t=0}^{T-1} \frac{t u_t}{T} \xrightarrow{d} \mathcal{N}\left(0, 1\right)$$

$$\iff T^{-1/2} \sum_{t=0}^{T-1} \frac{t u_t}{T} \xrightarrow{d} \mathcal{N}\left(0, \frac{\sigma^2}{3}\right) \tag{7}$$

To get the covariance between these distributions, we have (using the serial independence of u_t)

$$\mathrm{Cov}\left(T^{-1/2} \sum_{t=0}^{T-1} u_t, T^{-3/2} \sum_{t=0}^{T-1} t u_t \right) = \mathbb{E}\left[T^{-1/2} \sum_{t=0}^{T-1} u_t \cdot T^{-3/2} \sum_{t=0}^{T-1} t u_t \right] = T^{-2} \cdot \mathbb{E}\left[\sum_{t=0}^{T-1} u_t \cdot \sum_{t=0}^{T-1} t u_t \right]$$

$$= T^{-2} \cdot \mathbb{E}\left[\sum_{t=0}^{T-1} t u_t^2 \right] = T^{-2} \sum_{t=0}^{T-1} t \mathbb{E}(u_t^2) = \frac{\sigma^2}{T^2} \cdot \frac{T(T-1)}{2} \rightarrow \boxed{\frac{\sigma^2}{2}} \tag{8}$$

Putting (6), (7), and (8) together, we have $T^{-3/2} \sum_{t=1}^{T} y_{t-1} \xrightarrow{d} \mathcal{N}(0, \sigma^2 + \sigma^2/3 - 2 \cdot \sigma^2/2) = \boxed{\mathcal{N}(0, \sigma^2/3)}$.

Lastly, to get the covariance between these distributions, we have (again using the serial independence of u_t)

$$\mathrm{Cov}\left(T^{-1/2} \sum_{t=1}^{T-1} u_t, T^{-3/2} \sum_{t=1}^{T} y_{t-1} \right) = T^{-2} \mathrm{Cov}\left(\sum_{t=1}^{T-1} u_t, \sum_{t=1}^{T} (T - (t-1)) u_{t-1} \right)$$

$$= T^{-2} \mathrm{Cov}\left(\sum_{t=1}^{T-1} u_t, \sum_{t'=0}^{T-1} (T - t') u_{t'} \right) = T^{-2} \mathbb{E}\left(\sum_{t=1}^{T-1} (T - t) u_t^2 \right) = \frac{(T-1)T - T(T-1)/2}{T^2} \sigma^2$$

$$= \frac{(T-1)T}{2T^2} \sigma^2 \rightarrow \boxed{\sigma^2/2} \tag{9}$$

which yields the result.

(c) Let $r \in [0, 1)$, $t \in [0, T]$. Define

$$R_T(r) = \frac{1}{\sigma\sqrt{T}} y_{[rT]}$$

where $[rT]$ denotes the largest integer part of rT and $y_{[rT]} = 0$ if $[rT] = 0$. That is,

$$R_T(r) = \begin{cases} 0 & 0 \le r < 1/T \\ \frac{y_1}{\sigma\sqrt{T}} & 1/T \le r < 2/T \\ \frac{y_2}{\sigma\sqrt{T}} & 2/T \le r < 3/T \\ \vdots & \vdots \\ \frac{y_{T-1}}{\sigma\sqrt{T}} & (T-1)/T \le r < 1 \end{cases}$$

We have

$$T^{-3/2} \sum_{t=1}^{T} y_{t-1} = T^{-3/2} \sum_{t'=0}^{T-1} y_{t'} = T^{-3/2} \sum_{t'=0}^{T-1} (y_{t'-1} + u_{t'}) = \sigma \sum_{t'=0}^{T-1} \int_{t'/T}^{(t'+1)/T} R_T(r) dr + o_p(1)$$

$$= \sigma \int_0^1 R_T(r) dr + o_p(1) \implies \sigma \int_0^1 W(r) dr \text{ as } T \to \infty$$

where the last step follows from Donsker's Theorem (Theorem **??**, the functional central limit theorem) and the continuous mapping theorem (Theorem **??**) and $W(r)$ is a standard Weiner process.

Donsker's Theorem, Theorem 43, p.335, Section 15.6.3. Let $a \in [0, 1)$, $t \in [0, T]$, and suppose $(J-1)/T \le a < J/T$, $J = 1, 2, \ldots, T$. Define

$$R_T(a) = \frac{1}{\sqrt{T}} s_{[Ta]}$$

where

$$s_{[Ta]} = \epsilon_1 + \epsilon_2 + \ldots + \epsilon_{[Ta]}$$

$[Ta]$ denotes the largest integer part of Ta and $s_{[Ta]} = 0$ if $[Ta] = 0$. Then $R_T(a)$ weakly converges to $w(a)$, i.e.,

$$R_T(a) \to w(a)$$

where $w(a)$ is a Wiener process. Note that when $a = 1$, $R_T(1) = 1/\sqrt{T} \cdot S_{[T]} = 1/\sqrt{T} \cdot (\epsilon_1 + \epsilon_2 + \ldots + \epsilon_T)$. Since ϵ_t's are IID, by the central limit theorem, $R_T(1) \to \mathcal{N}(0, 1)$.

Continuous Mapping Theorem (Theorem 44 of Pesaran in 15.6.3). Let $a \in [0, 1)$, $i \in [0, n]$, and suppose $(J-1)/n \le a < J/n$, $J = 1, 2, \ldots, n$. Define $R_n(a) = n^{-1/2} S_{[n \cdot a]}$. If $f(\cdot)$ is continuous over $[0, 1)$, then

$$f[R_n(a)] \xrightarrow{d} f[w(a)]$$

(d) We have $T^{-1/2} \sum_{t=1}^{T} u_t = \sigma \cdot R_T(1) \implies \sigma \cdot W(1)$ by Donsker's Theorem. Using that and the result from (c), we have

$$T^{-3/2} \sum_{t=1}^{T-1} t u_t = T^{-1/2} \sum_{t=1}^{T-1} u_t - T^{-3/2} \sum_{t=1}^{T} y_{t-1} \xrightarrow{d} \sigma \cdot W(1) - \sigma \int_0^1 W(r) dr$$

1.6   Chapter 17: Introduction to Forecasting

Feel pretty good on concepts

1.6.1   17.7: Iterated and direct multi-step AR methods

Suppose y_t follows the AR(1) model:

$$y_t = a + \phi y_{t-1} + \epsilon_t, \quad |\phi| < 1, \epsilon_t \sim iid(0, \sigma_\epsilon^2) \tag{10}$$

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

$$\iff y_t = a \left(\frac{1-\phi^h}{1-\phi} \right) + \phi^h y_{t-h} + \sum_{j=0}^{h-1} \phi^j \epsilon_{t-j} \tag{11}$$

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 (12) using all available data Ω_T. Then we use the form of Equation (11):

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

(2) **Direct method:** We directly calculate OLS estimates of the parameters in Equation (11) using all available data Ω_T:

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

Proposition 1.1. (Pesaran Chapter 17 Proposition 45.) Suppose data is generated by Equation (12). 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

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

1.6.2 Worked Problems

Problem 4 (Homework 5 Question 3—fine on all but part (c), which even Lidan is hazy on.

Consider the AR(1) model

$$y_t = \phi y_{t-1} + u_t, \quad u_t \sim iid(0, \sigma^2) \tag{12}$$

(a) Derive iterated and direct forecasts of y_{T+2} condition on y_T, and show that they can be estimated as

$$\text{Iterated: } \hat{y}^{(it)}_{T+2|T} = \hat{\phi}^2 y_T$$

$$\text{Direct: } \hat{y}^{(d)}_{T+2|T} = \hat{\phi}_2 y_T$$

where $\hat{\phi}$ and $\hat{\phi}_2$ are OLS coefficients in the regressions of y_t on y_{t-1} and y_{t-2}, respectively, using the M observations $y_{T-M+1}, y_{T-M+2}, \ldots, y_T$.

(b) Show that conditional on y_t,

$$\mathbb{E}(y_{T+2} - \hat{y}^{(it)}_{T+2|T})^2 = \mathbb{E}(\phi^2 - \hat{\phi}^2)^2 y_T^2 + (1 + \phi^2)\sigma^2$$

$$\mathbb{E}(y_{T+2} - \hat{y}^{(d)}_{T+2|T})^2 = \mathbb{E}(\phi^2 - \hat{\phi}_2)^2 y_T^2 + (1 + \phi^2)\sigma^2$$

(c) Hence, or otherwise, show that

$$\lim_{M \to \infty} \mathbb{E}(d_{T+2}) = 0$$

where d_{T+2} is the loss differential of the two forecasting methods, defined by

$$d_{T+2} = (y_{T+2} - \hat{y}^{(it)}_{T+2|T})^2 + (y_{T+2} - \hat{y}^{(d)}_{T+2|T})^2$$

Solution.

(a) Note that we can write

$$y_t = \phi\left[\phi y_{t-2} + u_{t-1}\right] + u_t = \phi^2 y_{t-2} + u_t + \phi u_{t-1}$$

$$\implies y_{T+2} = \phi^2 y_T + u_{T+2} + \phi u_{T+1}$$

$$\implies \mathbb{E}(y_{T+2} \mid y_{T-M+1}, y_{T-M+2}, \ldots, y_T) = \phi^2 y_T + \mathbb{E}(u_{T+2} \mid y_{T-M+1}, \ldots, y_T) + \phi\mathbb{E}(u_{T+1} \mid y_{T-M+1}, \ldots, y_T)$$

$$= \phi^2 y_T$$

$$\Longleftrightarrow \mathbb{E}(y_{T+2} \mid y_{T-M+1}, y_{T-M+2}, \ldots, y_T) = \phi^2 y_T \qquad (13)$$

If we calculate the OLS estimate $\hat{\phi}$ in Equation (12), we can substitute that into Equation (13) to obtain the iterated estimate of y_{T+2}:

$$\hat{y}_{T+2|T}^{(it)} = \hat{\phi}^2 y_T$$

Since we have no intercept term, the OLS estimate would be simply

$$\hat{\phi} = (x'x)^{-1}x'y = \frac{\sum_{t=T-M+2}^{T} y_t y_{t-1}}{\sum_{t=T-M+1}^{T-1} y_t^2}$$

Alternatively, we could directly calculate the OLS estimate $\hat{\phi}_2$ of ϕ^2 in Equation (13)

$$\hat{y}_{T+2|T}^{(d)} = \hat{\phi}_2 y_T$$

The OLS estimate would be simply

$$\hat{\phi}_2 = (x'x)^{-1}x'y = \frac{\sum_{t=T-M+3}^{T} y_t y_{t-2}}{\sum_{t=T-M+1}^{T-2} y_t^2}$$

(b) We have

$$\mathbb{E}(y_{T+2} - \hat{y}_{T+2|T}^{(it)})^2 = \mathbb{E}(y_{T+2}^2) + \mathbb{E}((\hat{y}_{T+2|T}^{(it)})^2) - 2\mathbb{E}(y_{T+2} \cdot \hat{y}_{T+2|T}^{(it)})$$

$$= \mathbb{E}((\phi^2 y_T + u_{T+2} + \phi u_{T+1})^2) + \mathbb{E}((\hat{\phi}^2 y_T)^2) - 2\mathbb{E}((\phi^2 y_T + u_{T+2} + \phi u_{T+1}) \cdot \hat{\phi}^2 y_T)$$

$$= \mathbb{E}(\phi^4 y_T^2 + (u_{T+2} + \phi u_{T+1})^2 + 2\phi^2 y_T(u_{T+2} + \phi u_{T+1})) + y_T^2 \mathbb{E}(\hat{\phi}^2) - 2\phi^2 y_T \mathbb{E}(\hat{\phi}^2 y_T)$$

$$= \mathbb{E}(\phi^4)y_T^2 + \mathbb{E}(u_{T+2}^2 + \phi^2 u_{T+1}^2) + y_T^2 \mathbb{E}(\hat{\phi}^2) - 2\phi^2 y_T^2 \mathbb{E}(\hat{\phi}^2) = \mathbb{E}(\phi^4)y_T^2 + (1+\phi^2)\sigma^2 + y_T^2 \mathbb{E}(\hat{\phi}^2) - 2\phi^2 y_T^2 \mathbb{E}(\hat{\phi}^2)$$

$$= y_T^2 \left[\mathbb{E}(\phi^4 + \hat{\phi}^2 - 2\phi^2 \hat{\phi}^2) \right] + (1+\phi^2)\sigma^2 = \boxed{\mathbb{E}(\phi^2 - \hat{\phi}^2)^2 y_T^2 + (1+\phi^2)\sigma^2}$$

$$\mathbb{E}(y_{T+2} - \hat{y}_{T+2|T}^{(d)})^2 = \mathbb{E}(y_{T+2}^2) + \mathbb{E}((\hat{y}_{T+2|T}^{(d)})^2) - 2\mathbb{E}(y_{T+2} \cdot \hat{y}_{T+2|T}^{(d)})$$

$$= \mathbb{E}((\phi^2 y_T + u_{T+2} + \phi u_{T+1})^2) + \mathbb{E}((\hat{\phi}_2 y_T)^2) - 2\mathbb{E}((\phi^2 y_T + u_{T+2} + \phi u_{T+1}) \cdot \hat{\phi}_2 y_T)$$

$$= \mathbb{E}(\phi^4 y_T^2 + (u_{T+2} + \phi u_{T+1})^2 + 2\phi^2 y_T(u_{T+2} + \phi u_{T+1})) + y_T^2 \mathbb{E}(\hat{\phi}_2) - 2\phi^2 y_T \mathbb{E}(\hat{\phi}_2 y_T)$$

$$= \mathbb{E}(\phi^4)y_T^2 + \mathbb{E}(u_{T+2}^2 + \phi^2 u_{T+1}^2) + y_T^2 \mathbb{E}(\hat{\phi}_2) - 2\phi^2 y_T^2 \mathbb{E}(\hat{\phi}_2) = \mathbb{E}(\phi^4)y_T^2 + (1+\phi^2)\sigma^2 + y_T^2 \mathbb{E}(\hat{\phi}_2) - 2\phi^2 y_T^2 \mathbb{E}(\hat{\phi}_2)$$

$$= y_T^2 \left[\mathbb{E}(\phi^4 + \hat{\phi}_2 - 2\phi^2 \hat{\phi}_2) \right] + (1+\phi^2)\sigma^2 = \boxed{\mathbb{E}(\phi^2 - \hat{\phi}_2)^2 y_T^2 + (1+\phi^2)\sigma^2}$$

(c) Begin by expanding $\hat{\phi}^2$ in a first order Taylor series about ϕ:

$$\hat{\phi} = \phi + \mathcal{O}_p(M^{-1/2}) \implies \hat{\phi}^2 = \phi^2 + 2\phi(\hat{\phi} - \phi) + \mathcal{O}_p(M^{-1})$$

Then

$$\mathbb{E}\left[(\hat{\phi}^2 - \phi^2)^2 \mid y_T\right] = \mathbb{E}\left[\left(2\phi(\hat{\phi} - \phi) + \mathcal{O}_p(M^{-1})\right)^2 \mid y_T\right] = 4\phi^2 \mathbb{E}\left[\left((\hat{\phi} - \phi) + \mathcal{O}_p(M^{-1})\right)^2 \mid y_T\right]$$

$$= 4\phi^2 \mathbb{E}\left[\left((\hat{\phi} - \phi) + \mathcal{O}_p(M^{-1})\right)^2 \mid y_T\right]$$

1.7 Chapter 18: Measurement and Modeling of Volatility

Maybe review a little, but feel pretty okay on concepts

GARCH(1, 1) model (Pesaran Equation 18.5):

$$h_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \phi_1 h_{t-1}^2, \quad \alpha_0 > 0 \tag{14}$$

- This process is unconditionally stationary if $|\alpha_1 + \phi_1| < 1$.

- The unconditional variance exists and is fixed if $|\alpha_1 + \phi_1| < 1$.

- The case where $\alpha_1 + \phi_1 = 1$ is known as the Integrated GARCH(1,1), or IGARCH(1,1) for short. The RiskMetrics exponentially weighted formulation of h_t^2 for large H is a special case of the IGARCH(1,1) model where α_0 is set to 0. RiskMetrics formulation avoids the variance non-existence problem by focusing on H fixed.

1.7.1 Higher order GARCH models (Pesaran Section 18.4.2)

The various members of the GARCH and GARCH-M class of models can be written compactly as

$$y_t = \boldsymbol{\beta}' \boldsymbol{x}_{t-1} + \gamma h_t^2 + \epsilon_t \tag{15}$$

where

$$h_t^2 = \text{Var}(\epsilon_t \mid \Omega_{t-1}) = \mathbb{E}(\epsilon_t^2 \mid \Omega_{t-1}) = \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{p} \phi_i h_{t-i}^2 \tag{16}$$

and Ω_{t-1} is the information set at time $t-1$ containing at least $(\boldsymbol{x}_{t-1}, \boldsymbol{x}_{t-2}, \ldots, y_{t-1}, y_{t-2}, \ldots)$. The unconditional variance of ϵ_t is determined by

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \sigma_{t-i}^2 + \sum_{i=1}^{p} \phi_i \sigma_{t-i}^2$$

and yields a stationary outcome if all the roots of

$$1 - \sum_{i=1}^{q} \alpha_i \lambda^i + \sum_{i=1}^{p} \phi_i \lambda^i = 0$$

like outside the unit circle. In that case

$$\mathrm{Var}(\epsilon_t) = \sigma^2 = \frac{\alpha_0}{1 - \sum_{i=1}^{q} \alpha_i - \sum_{i=1}^{p} \phi_i} > 0 \tag{17}$$

Clearly the necessary condition for (16) to be covariance stationary is given by

$$\sum_{i=1}^{q} \alpha_i + \sum_{i=1}^{p} \phi_i < 1$$

1.7.2   Testing for GARCH effects (Pesaran Section 18.5.1)

If we consider y_t as periodic data defined by $y_t = r_t - \bar{r}$ with r_t representing, say, asset return and $\bar{r}$ representing the unconditional mean, then we have the GARCH(1,1) representation of volatility:

$$\mathrm{Var}(y_t \mid \Omega_{t-1}) = h_t^2 = \bar{\sigma}^2(1 - \alpha - \beta) + \alpha y_{t-1}^2 + \beta h_{t-1}^2$$

Then the test for GARCH effects would test

$$H_0 : \alpha = 0$$

against

$$H_1 : \alpha \neq 0$$

GARCH(1,1) can be approximated by ARCH(q):

$$\mathrm{Var}(y_t \mid \Omega_{t-1}) = \frac{\bar{\sigma}^2(1 - \alpha - \beta)}{1 - \beta} + \alpha y_{t-1}^2 + \alpha\beta y_{t-2}^2 + \ldots + \alpha\beta^{q-1} y_{t-q}^2$$

$$= \tilde{\alpha}_0 + \tilde{\alpha}_1 y_{t-1}^2 + \tilde{\alpha}_2 z_{t-2}^2 + \ldots + \tilde{\alpha}_q z_{t-q}^2$$

which means that we can approximate this hypothesis test by instead using the Lagrange multiplier test proposed by Engle:

$$H_0 : \tilde{\alpha}_1 = \tilde{\alpha}_2 = \ldots = \tilde{\alpha}_q$$

against

$$H_1 : \tilde{\alpha}_1 \neq 0, \tilde{\alpha}_2 \neq 0, \ldots, \tilde{\alpha}_q \neq 0$$

1.7.3   Worked Problems

Problem 1. Consider the generalized autoregressive heteroskedastic model

$$y_t = h_t z_t$$

where

$$z_t \mid \Omega_{t-1} \sim IIDN(0,1) \tag{18}$$

$$h_t^2 = \mathrm{Var}(y_t \mid \Omega_{t-1}) = \mathbb{E}(y_t^2 \mid \Omega_{t-1}) = \bar{\sigma}^2(1 - \alpha - \beta) + \alpha y_{t-1}^2 + \beta h_{t-1}^2 \tag{19}$$

and Ω_t is the information set that contains at least y_t and its lagged values.

(a) Derive the conditions under which $\{y_t\}$ is a stationary process.

(b) Are the observations $\{y_t\}$ serially independent and/or serially uncorrelated?

(c) Develop a test of the GARCH effect and discuss the estimation of the above model by the maximum likelihood method.

(d) Discuss the relevance of GARCH models for the analysis of financial time series data.

Solution.

(a) Note that (using the fact that y_t and h_t are conditionally independent given Ω_{t-1})

$$\mathbb{E}(y_t) = \mathbb{E}(h_t z_t) = \mathbb{E}\big[\mathbb{E}(h_t z_t \mid \Omega_{t-1})\big] = \mathbb{E}\big[\mathbb{E}(h_t \mid \Omega_{t-1})\mathbb{E}(z_t \mid \Omega_{t-1})\big] = \mathbb{E}\big[\mathbb{E}(h_t \mid \Omega_{t-1}) \cdot 0\big] = 0$$

We have

$$\mathrm{Var}(y_t) = \mathbb{E}\big[\big(y_t - \mathbb{E}(y_t)\big)^2\big] = \mathbb{E}\big[y_t^2\big] = \mathbb{E}\big[\mathbb{E}(y_t^2 \mid \Omega_{t-1})\big] = \mathbb{E}(h_t^2) = \mathbb{E}(\bar{\sigma}^2(1 - \alpha - \beta) + \alpha y_{t-1}^2 + \beta h_{t-1}^2)$$

$$= \mathbb{E}\big(\bar{\sigma}^2(1 - \alpha - \beta)\big) + \alpha\mathbb{E}(y_{t-1}^2) + \beta\mathbb{E}(h_{t-1}^2) = \bar{\sigma}^2(1 - \alpha - \beta) + (\alpha + \beta)\mathrm{Var}(y_{t-1})$$

Therefore in order for this to be a stationary process, we require $\boxed{|\alpha + \beta| < 1}$.

more stuff I did on original homework

By Definition 1.1 $\{y_t\}$ is a stationary process if it has constant mean and its covariance function depends only on the absolute difference $|t_1 - t_2|$; that is,

$$\text{Cov}(y_{t_1}, y_{t_2}) = \gamma(t_1, t_2) = \gamma(|t_1 - t_2|) \text{ for all } t_1, t_2$$

$$\vdots$$

In order for this to be true, we must first show that h_t is finite for all t. It is sufficient to find the conditions that make h_t stationary. Using Equation (16)

$$h_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{p} \phi_i h_{t-i}^2$$

From equation (19) we have

$$h_t^2 = \bar{\sigma}^2(1 - \alpha - \beta) + \alpha y_{t-1}^2 + \beta h_{t-1}^2$$

Therefore we note that $p = q = 1$ and we have a GARCH(1,1) model. From Equation (14):

$$h_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \phi_1 h_{t-1}^2, \quad \alpha_0 > 0$$

is unconditionally stationary if $|\alpha_1 + \phi_1| < 1$. In this case, we require $\boxed{|\alpha + \beta| < 1}$ for stationarity of h_t.

$$\vdots$$

Assume without loss of generality that $t_2 \geq t_1$. (Note that this implies $z_{t_2} \perp\!\!\!\perp z_{t_1}, h_{t_2}, h_{t_1} \mid \Omega_{t_2-1}$.)

$$\text{Cov}(y_{t_1}, y_{t_2}) = \mathbb{E}\left[\left(y_{t_1} - \mathbb{E}(y_{t_1})\right)\left(y_{t_2} - \mathbb{E}(y_{t_2})\right)\right] = \mathbb{E}\left[y_{t_1} y_{t_2}\right] = \mathbb{E}\left[h_{t_1} h_{t_2} z_{t_1} z_{t_2}\right]$$

$$= \mathbb{E}\left[\mathbb{E}(h_{t_1} h_{t_2} z_{t_1} z_{t_2} \mid \Omega_{t_2-1})\right] = \mathbb{E}\left[\mathbb{E}(h_{t_1} h_{t_2} z_{t_1} \mid \Omega_{t_2-1})\mathbb{E}(z_{t_2} \mid \Omega_{t_2-1})\right]$$

$$= \mathbb{E}\left[\mathbb{E}(h_{t_1} h_{t_2} z_{t_1} \mid \Omega_{t_2-1}) \cdot 0\right] = \boxed{0}$$

Therefore given that h_t is finite (that is, given $|\alpha + \beta| < 1$), we have that $\mathbb{E}(y_t) = 0$, $\text{Cov}(y_{t_1}, y_{t_2}) = 0 \ \forall \ t_1, t_2$ which implies that under these conditions y_t is stationary.

(b) y_t is serially uncorrelated because $\text{Cov}(y_{t_1}, y_{t_2}) = 0 \ \forall \ t_1, t_2$. However, it is clear that y_t is not serially independent since past values of y_t affect $h_t^2 = \text{Var}(y_t)$. Observe that

$$\Pr(y_t \leq y \mid y_{t-1}) = \Pr(h_t z_t \leq y \mid y_{t-1}) = \Pr\left(z_t \sqrt{\bar{\sigma}^2(1 - \alpha - \beta) + \alpha y_{t-1}^2 + \beta h_{t-1}^2} \leq y \mid y_{t-1}\right)$$

In other words, even though the mean of y_t remains constant, because y_{t-1} affects the variance of y_t, it affects the heaviness of the tails of y_t, changing the probability distribution of y_t. Therefore the conditional cumulative distribution function of y_t given y_{t-1} is not equal to the unconditional cdf, so y_t and y_{t-1} are not independent.

Lidan's Explanation: From (a) we know that $\{y_t\}$ is serially uncorrelated, but they are not necessarily independent because from Equation (19) we know that h_t is not independent from h_{t-1}.

(c) See section 1.7.2.

(d) **Lidan:** Usually financial time series data are fat-tailed. So the series may not look stationary, and therefore the local variance would be clustered in some very low and very high values. To capture this serial correlation and heterogeneity in volatility, we need to use a GARCH model.

Book: In financial econometrics, ARCH and GARCH are fundamental tools for analyzing the time-variation of conditional variance. In many applications in finance, the assumption that the conditional variance of the disturbances is constant over time is not valid. GARCH models allow for time variation in volatility, relating (unobserved) volatility to squares of past innovations in price changes. However, this approach only partly overcomes the deficiency of the historical measure and continues to respond very slowly when volatility undergoes rapid changes.

1.8   Chapter 21: Vector Autoregressive Models

Feel ok on concepts.

1.8.1   Worked Problems

Problem 1. Consider the bivariate autoregressive model:

$$Y_t = \mathbf{\Phi} Y_{t-1} + U_t, \qquad U_t \sim IIDN(0, \Sigma) \tag{20}$$

where

$$Y_t = (y_{1t}, y_{2t})', \ Y_{t-1} = (y_{1,t-1}, y_{2,t-1})', \ U_t = (u_{1t}, u_{2t})'$$

and

$$\mathbf{\Phi} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix}, \quad \Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix} \succ 0$$

(a) Derive the conditional mean and variance of y_{1t} with respect to y_{2t} and lagged values of y_{1t} and y_{2t}.

(b) Show that the univariate representation of y_{1t} is an ARMA(2,1) process.

Solution

(a) Note that

$$y_{1t} = \phi_{11}y_{1,t-1} + \phi_{12}y_{2,t-1} + u_{1t} \tag{21}$$

$$y_{2t} = \phi_{21}y_{1,t-1} + \phi_{22}y_{2,t-1} + u_{2t} \tag{22}$$

First,

$$\mathbb{E}(y_{1t} \mid y_{2t}, y_{2,t-1}, y_{2,t-2}, \ldots, y_{1,t-1}, y_{1,t-2}, \ldots) = \mathbb{E}(y_{1t} \mid y_{2t}, \Omega_{t-1})$$

$$= \phi_{11}y_{1,t-1} + \phi_{12}y_{2,t-1} + \mathbb{E}(u_{1t} \mid y_{2t}) = \phi_{11}y_{1,t-1} + \phi_{12}y_{2,t-1} + \mathbb{E}(u_{1t} \mid \phi_{21}y_{1,t-1} + \phi_{22}y_{2,t-1} + u_{2t})$$

$$= \phi_{11}y_{1,t-1} + \phi_{12}y_{2,t-1} + \mathbb{E}(u_{1t} \mid u_{2t})$$

Recall Proposition **??**: for a bivariate normal distribution with mean 0, the conditional distribution of u_{1t} given u_{2t} is

$$u_{1t} \mid u_{2t} \sim \mathcal{N}\left(\rho \frac{\sqrt{\sigma_{11}}}{\sqrt{\sigma_{22}}} u_{2t}, (1 - \rho^2)\sigma_{11} \right)$$

where $\rho = \sigma_{12}/\sqrt{\sigma_{11}}\sqrt{\sigma_{22}}$.

$$\iff u_{1t} \mid u_{2t} \sim \mathcal{N}\left(\frac{\sigma_{12}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{22}}} \cdot \frac{\sqrt{\sigma_{11}}}{\sqrt{\sigma_{22}}} u_{2t}, \left[1 - \left(\frac{\sigma_{12}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{22}}} \right)^2 \right] \sigma_{11} \right)$$

$$\iff u_{1t} \mid u_{2t} \sim \mathcal{N}\left(\frac{\sigma_{12}}{\sigma_{22}} u_{2t}, \sigma_{11} - \frac{\sigma_{12}^2}{\sigma_{22}} \right)$$

$$\implies \mathbb{E}(u_{1t} \mid u_{2t}) = \frac{\sigma_{12}}{\sigma_{22}} u_{2t}, \ \mathrm{Var}(u_{1t} \mid u_{2t}) = \sigma_{11} - \frac{\sigma_{12}^2}{\sigma_{22}}$$

$$\implies \mathbb{E}(y_{1t} \mid y_{2t}, \Omega_{t-1}) = \phi_{11}y_{1,t-1} + \phi_{12}y_{2,t-1} + \frac{\sigma_{12}}{\sigma_{22}} u_{2t}$$

$$= \phi_{11}y_{1,t-1} + \phi_{12}y_{2,t-1} + \frac{\sigma_{12}}{\sigma_{22}}(y_{2t} - \phi_{21}y_{1,t-1} - \phi_{22}y_{2,t-1})$$

$$= \boxed{\left(\phi_{11} - \frac{\sigma_{12}}{\sigma_{22}}\phi_{21} \right)y_{1,t-1} + \left(\phi_{12} - \frac{\sigma_{12}}{\sigma_{22}}\phi_{22} \right)y_{2,t-1} + \frac{\sigma_{12}}{\sigma_{22}}y_{2t}}$$

Second,

$$\mathrm{Var}(y_{1t} \mid y_{2t}, \Omega_{t-1}) = \phi_{11}^2 \mathrm{Var}(y_{1,t-1} \mid y_{2t}, \Omega_{t-1}) + \phi_{12}^2 \mathrm{Var}(y_{2,t-1} \mid y_{2t}, \Omega_{t-1}) + \mathrm{Var}(u_{1t} \mid y_{2t}, \Omega_{t-1})$$

$$= \mathrm{Var}(u_{1t} \mid u_{2t}) = \boxed{\sigma_{11} - \frac{\sigma_{12}^2}{\sigma_{22}}}$$

(b) Again, using equations (21) and (22),

$$y_{1t} = \phi_{11}y_{1,t-1} + \phi_{12}y_{2,t-1} + u_{1t}, \quad y_{2t} = \phi_{21}y_{1,t-1} + \phi_{22}y_{2,t-1} + u_{2t}$$

$$\implies (1 - \phi_{11}L)y_{1t} = \phi_{12}Ly_{2t} + u_{1t} \tag{23}$$

$$(1 - \phi_{22}L)y_{2t} = \phi_{21}Ly_{1t} + u_{2t} \tag{24}$$

We can multiply $(1 - \phi_{22}L)$ on both sides of (23) to yield

$$(1 - \phi_{22}L)(1 - \phi_{11}L)y_{1t} = \phi_{12}L(1 - \phi_{22}L)y_{2t} + (1 - \phi_{22}L)u_{1t}$$

Using (24), we have

$$(1 - (\phi_{22} + \phi_{11})L + \phi_{11}\phi_{22}L^2)y_{1t} = \phi_{12}L\big[\phi_{21}Ly_{1t} + u_{2t}\big] + (1 - \phi_{22}L)u_{1t}$$

$$\iff (1 - (\phi_{22} + \phi_{11})L + (\phi_{11}\phi_{22} - \phi_{12}\phi_{21})L^2)y_{1t} = \phi_{12}Lu_{2t} + (1 - \phi_{22}L)u_{1t} \tag{25}$$

If we can show that the left side of (25) is an AR(2) process and the right side is an MA(1) process, we are done. **Lidan: "It is obvious that the left hand side of (25) is an AR(2) process."** The left side is a stationary AR(2) process if the absolute values of all the roots of

$$1 - (\phi_{22} + \phi_{11})z + (\phi_{11}\phi_{22} - \phi_{12}\phi_{21})z^2 = 0$$

are greater than 1. The roots are

$$z = \frac{\phi_{22} + \phi_{11} \pm \sqrt{(\phi_{22} + \phi_{11})^2 - 4(\phi_{11}\phi_{22} - \phi_{12}\phi_{21})}}{2(\phi_{11}\phi_{22} - \phi_{12}\phi_{21})} = \frac{\phi_{22} + \phi_{11} \pm \sqrt{\phi_{22}^2 + \phi_{11}^2 - 2\phi_{22}\phi_{11} + 4\phi_{12}\phi_{21}}}{2(\phi_{11}\phi_{22} - \phi_{12}\phi_{21})}$$

$$= \frac{\phi_{22} + \phi_{11} \pm \sqrt{(\phi_{22} - \phi_{11})^2 + 4\phi_{12}\phi_{21}}}{2(\phi_{11}\phi_{22} - \phi_{12}\phi_{21})}$$

$$\vdots$$

To check if the right side of (25) is MA(1), we will write $x_t = \phi_{12}Lu_{2t} + (1 - \phi_{22}L)u_{1t}$ as an MA(1) process; that is,

$$x_t = \phi_{12}u_{2,t-1} + u_{1t} - \phi_{22}u_{1,t-1} = \xi_t + \theta\xi_{t-1}$$

with $\xi_t \sim iid(0, \sigma_\xi^2)$, $|\theta| < 1$. If x_t is an MA(1) process, it must satisfy

$$\gamma(0) = \mathbb{E}(x_t^2) = (1 + \theta^2)\sigma_\xi^2, \quad \gamma(1) = \mathbb{E}(x_t x_{t-1}) = \theta\sigma_\xi^2$$

We have (using the serial independence of the u_{1t}, u_{2t})

$$\mathbb{E}(x_t^2) = \mathbb{E}\big(\big[\phi_{12}u_{2,t-1} + u_{1t} - \phi_{22}u_{1,t-1}\big]^2\big) = \mathbb{E}\big(\phi_{12}^2 u_{2,t-1}^2 - 2\phi_{12}\phi_{22}u_{2,t-1}u_{1,t-1} + \phi_{22}^2 u_{1,t-1}^2 + u_{1t}^2\big)$$

$$= \phi_{12}^2\mathbb{E}\big(u_{2,t-1}^2\big) - 2\phi_{12}\phi_{22}\mathbb{E}\big(u_{2,t-1}u_{1,t-1}\big) + \phi_{22}^2\mathbb{E}\big(u_{1,t-1}^2\big) + \mathbb{E}\big(u_{1t}^2\big)$$

$$= \phi_{12}^2\sigma_{22} - 2\phi_{12}\phi_{22}\sigma_{12} + \phi_{22}^2\sigma_{11} + \sigma_{11} = (1 + \phi_{22}^2)\sigma_{11} - 2\phi_{12}\phi_{22}\sigma_{12} + \phi_{12}^2\sigma_{22}$$

$$\mathbb{E}(x_t x_{t-1}) = \mathbb{E}\big(\big[\phi_{12}u_{2,t-1} + u_{1t} - \phi_{22}u_{1,t-1}\big]\big[\phi_{12}u_{2,t-2} + u_{1,t-1} - \phi_{22}u_{1,t-2}\big]\big)$$

$$= \mathbb{E}\big(\phi_{12}u_{2,t-1}u_{1,t-1} - \phi_{22}u_{1,t-1}^2\big) = \phi_{12}\sigma_{12} - \phi_{22}\sigma_{11}$$

Therefore we require

$$(1 + \theta^2)\sigma_\xi^2 = (1 + \phi_{22}^2)\sigma_{11} - 2\phi_{12}\phi_{22}\sigma_{12} + \phi_{12}^2\sigma_{22} \tag{26}$$

$$\theta\sigma_\xi^2 = \phi_{12}\sigma_{12} - \phi_{22}\sigma_{11} \tag{27}$$

Dividing (26) by (27) we have

$$\frac{1 + \theta^2}{\theta} = \frac{(1 + \phi_{22}^2)\sigma_{11} - 2\phi_{12}\phi_{22}\sigma_{12} + \phi_{12}^2\sigma_{22}}{\phi_{12}\sigma_{12} - \phi_{22}\sigma_{11}} \tag{28}$$

For simplicity, let the right side of (28) be $A \in \mathbb{R}$. Then we have

$$\theta^2 - A\theta + 1 = 0 \iff \boxed{\theta_1 = \frac{1}{2}\big(A + \sqrt{A^2 - 4}\big), \theta_2 = \frac{1}{2}\big(A + \sqrt{A^2 - 4}\big)}$$

Then the corresponding σ_ξ^2 are

$$\boxed{\sigma_{\xi,1}^2 = \sigma_{\xi,2}^2 = \frac{\phi_{12}\sigma_{12} - \phi_{22}\sigma_{11}}{\theta_1}}$$

As a double check,

$$\mathbb{E}(x_t x_{t-2}) = \mathbb{E}\big(\big[\phi_{12}u_{2,t-1} + u_{1t} - \phi_{22}u_{1,t-1}\big]\big[\phi_{12}u_{2,t-3} + u_{1,t-2} - \phi_{22}u_{1,t-3}\big]\big) = 0$$

as expected for an MA(1) process. Therefore the right side of (25) is an MA(1) process (provided that $0 < |\theta_1| < 1$ and/or $0 < |\theta_2| < 1$). Since the left side is an AR(2) process, this proves that the univariate representation (25) of y_{1t} is an ARMA(2,1) process.

Problem 2. Consider the VAR(2) model in the m-dimensional vector Y_t:

$$Y_t = \mu + \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + U_t, \quad U_t \sim (0, \Sigma) \tag{29}$$

where μ is an m-vector of fixed constants.

(a) Derive the conditions under which the VAR(2) model defined in (29) is stationary.

(b) Derive the error correction form of (29) and discuss what is meant by the process Y_t being cointegrated.

(c) Suppose now that one or more elements of Y_t is I(1). Derive suitable restrictions on the intercepts μ such that despite the I(1) nature of the variables in (29), Y_t has a fixed mean. Discuss the importance of such restrictions for the analysis of cointegration.

Solution.

(a) Let

$$Y_t^* = Y_t - (I - \Phi_1 - \Phi_2)^{-1}\mu. \tag{30}$$

Then

$$Y_t^* = \mu + \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + U_t - (I - \Phi_1 - \Phi_2)^{-1}\mu$$

$$= \mu + \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + U_t - [(I - \Phi_1 - \Phi_2) + \Phi_1 + \Phi_2](I - \Phi_1 - \Phi_2)^{-1}\mu$$

$$= \mu - (I - \Phi_1 - \Phi_2)(I - \Phi_1 - \Phi_2)^{-1}\mu + \Phi_1(Y_{t-1} - (I - \Phi_1 - \Phi_2)^{-1}\mu) + \Phi_2(Y_{t-2} - (I - \Phi_1 - \Phi_2)^{-1}\mu) + U_t$$

$$= \Phi_1 Y_{t-1}^* + \Phi_2 Y_{t-2}^* + U_t \tag{31}$$

Equation (31) can be rewritten in the companion form as follows:

$$\begin{bmatrix} Y_t^* \\ Y_{t-1}^* \end{bmatrix} = \begin{bmatrix} \Phi_1 & \Phi_2 \\ I & 0 \end{bmatrix} \begin{bmatrix} Y_{t-1}^* \\ Y_{t-2}^* \end{bmatrix} + \begin{bmatrix} U_t \\ 0 \end{bmatrix} \tag{32}$$

which is a VAR(1) model. If Y_{-M+1}, Y_{-M+2} are given, equation (32) can be solved iteratively from $t = -M + 2$ to obtain

$$\begin{bmatrix} Y_t^* \\ Y_{t-1}^* \end{bmatrix} = \begin{bmatrix} \Phi_1 & \Phi_2 \\ I & 0 \end{bmatrix}^{t+M-2} \begin{bmatrix} Y_{-M+2}^* \\ Y_{-M+1}^* \end{bmatrix} + \sum_{j=0}^{t+M-3} \begin{bmatrix} \Phi_1 & \Phi_2 \\ I & 0 \end{bmatrix}^j \begin{bmatrix} U_{t-j} \\ 0 \end{bmatrix} \tag{33}$$

Then the condition for (33) to be covariance stationary is for all the eigenvalues of

$$\Phi = \begin{bmatrix} \Phi_1 & \Phi_2 \\ I & 0 \end{bmatrix}$$

to lie inside the unit circle; that is, the solutions of

$$|\Phi - \lambda I_4| = 0$$

must satisfy $|\lambda| < 1$. Equivalently, the stability condition can be written in terms of the roots of the determinantal equation

$$|I_2 - \Phi_1 z - \Phi_2 z^2| = 0$$

in which case the process Y_t^* will be stationary if all the roots lie outside the unit circle ($|z| > 1$). Then if Y_t^* is stationary, so is Y_t, so either of these conditions (plus the invertibility of $(I - \Phi_1 - \Phi_2)$) ensure stationarity of Y_t.

(b) From (31) we have

$$Y_t^* = \Phi_1 Y_{t-1}^* + \Phi_2 Y_{t-2}^* + U_t$$

Note that (letting $\Delta Y_t^* = Y_t^* - Y_{t-1}^*$)

$$Y_t^* - Y_{t-1}^* + Y_{t-1}^* = \Phi_1 Y_{t-1}^* + \Phi_2 \big(Y_{t-2}^* - Y_{t-1}^* + Y_{t-1}^*\big) + U_t$$

$$\iff \Delta Y_t^* + Y_{t-1}^* = \Phi_1 Y_{t-1}^* + \Phi_2 \big(Y_{t-1}^* - \Delta Y_{t-1}^*\big) + U_t$$

$$\iff \Delta Y_t^* = -(I - \Phi_1 - \Phi_2) Y_{t-1}^* - \Phi_2 \Delta Y_{t-1}^* + U_t \iff \boxed{\Delta Y_t^* = -\Pi Y_{t-1}^* + \Gamma \Delta Y_{t-1}^* + U_t}$$

where $\Pi = (I - \Phi_1 - \Phi_2)$ and $\Gamma = -\sum_{i=2}^2 \Phi_i = -\Phi_2$. Because

$$\Delta Y_t^* = Y_t^* - Y_{t-1}^* = Y_t - (I - \Phi_1 - \Phi_2)^{-1} \mu - \big[Y_{t-1} - (I - \Phi_1 - \Phi_2)^{-1} \mu\big] = Y_t - Y_{t-1} = \Delta Y_t,$$

this can also be written as

$$\Delta Y_t = -\Pi \big[Y_{t-1} - (I - \Phi_1 - \Phi_2)^{-1} \mu\big] + \Gamma \Delta Y_{t-1} + U_t$$

$$= -\Pi \big[Y_{t-1} - \Pi^{-1} \mu\big] + \Gamma \Delta Y_{t-1} + U_t = \Pi \cdot \Pi^{-1} \mu - \Pi Y_{t-1} + \Gamma \Delta Y_{t-1} + U_t$$

$$\implies \boxed{\Delta Y_t = \mu - \Pi Y_{t-1} + \Gamma \Delta Y_{t-1} + U_t} \tag{34}$$

For the definition of cointegration, see Definition 1.3 and Section 1.9.1. In this particular case, if $Y_{t-1}^* \sim I(1)$ and the linear combinations ΠY_{t-1}^* of Y_{t-1}^* are covariance stationary (that is, $\Pi Y_{t-1}^* \sim I(0)$), we say Y_t^* is cointegrated (and therefore so is Y_t).

(c) Since (34) is I(0), for Y_t to have fixed mean, take expectations on both sides of (34):

$$\mu - (I_m - \Phi_1 - \Phi_2) \mathbb{E}(Y_{t-1}) = 0$$

If this restriction is violated, then (34) becomes a stationary process with a drift, which implies (29) has a time trend.

1.9   Chapter 22: Cointegration Analysis

Feel pretty good except for long run effects, examples we went over in class, the restrictions, and the 5 cases. Don't need to understand SURE.

1.9.1   22.4 Cointegrating VAR: multiple cointegrating relations and 22.5: Identification of long-run effects

Definition 1.3. We say that the m variables in Y_t are *cointegrated* if they are individually integrated (or have a random walk component) but there exist linear combinations of them which are stationary. That is, $y_{it} \sim I(1)$ for $i = 1, 2, \ldots, m$, but there exists an $m \times r$ matrix β such that $\beta' Y_t = \xi_t \sim I(0)$.

- In this case r denotes the number of cointegrating vectors, also known as the dimension of the cointegration space.

- The cointegrating relations summarized in the $r \times 1$ vector $\beta' Y_t$ are also known as long-run relations.

- $r = \mathbf{rank}(\Pi)$ is the dimension of the cointegration space.

- Cointegration is present if Π is rank-deficient; that is, $r < m$.

When $\mathbf{rank}(\Pi) = r < m$, we can write Π as

$$\Pi = \alpha\beta' \tag{35}$$

where α and β are $m \times r$ matrices of full column rank. Then

$$\Pi y_{t-1} = \alpha\beta' y_{t-1} \sim I(0)$$

and the VECM can be written as

$$\Delta y_t = -\alpha\beta' y_{t-1} + \sum_{j=1}^{p-1} \Gamma_j \Delta y_{t-j} + u_t \tag{36}$$

Since α is full rank, we have

$$\beta' y_{t-1} \sim I(0)$$

where $\beta' y_t$ is the r-vector of cointegrating relations, also known as the long-run relations.

However, β as defined above is not uniquely determined. Consider a linear transformation of β by a non-singular $r \times r$ matrix Q: $\tilde{\beta} = \beta Q$. Then since $\mathbf{rank}(\Pi) = r < m$, Π can be expressed as $\Pi = \alpha\beta'$ where α is a rank r $m \times r$ matrix. But consider that

$$\Pi = \alpha\beta = (\alpha Q'^{-1})(Q'\beta') = \tilde{\alpha}\tilde{\beta}'$$

with $\tilde{\alpha} = \alpha Q'^{-1}$. Therefore β is not uniquely determined without r^2 exact- or just-identifying restrictions, r restrictions on each of the r cointegrating relations.

1.9.2 Worked Problems

Problem 2: See Chapter 21 Problem 2.

1.10 Chapter 23: VARX Modeling

Lidan says will not be on final.

1.10.1 Worked Problems

1.11 Chapter 24: Impulse Response Analysis

Feel ok on concepts, should review and do an example problem.

1.11.1 Worked Problems

Chapter 24 Problem 1. Consider the VAR(2) model

$$x_t = \Phi_1 x_{t-1} + \Phi_2 x_{t-2} + \epsilon_t, \quad \epsilon_t \sim IID(0, \Sigma)$$

in the $m \times 1$ vectof of random variables x_t and Σ is the covariance matrix of the errors with typical element σ_{ij}.

(a) Derive the conditions under which this process is stationary, and show that it has the following moving average representation:

$$x_t = \sum_{j=0}^{\infty} A_j \epsilon_{t-j} \tag{37}$$

(b) Derive the coefficient matrices A_j in terms of Φ_1 and Φ_2.

(c) Using the above result, write down the orthogonalized (OIR) and generalize impulse (GIR) response functions of one standard error shock (i.e. $\sqrt{\sigma_{ii}}$ to the error of the ith equation, $\epsilon_{it} = s_i' \epsilon_t$, where s_i is am $m \times 1$ selection vector.

(d) What are the main differences betwen OIR and GIR functions?

Chapter 24 Problem 1 Solution.

(a) For stationarity conditions, see Ch. 21 Problem 2 in Section 1.8.1. To get the MA representation

$$x_t = \sum_{j=0}^{\infty} A_j \epsilon_{t-j} \tag{38}$$

of (37) note that

$$(I - \Phi_1 L - \Phi_2 L^2) x_t = \epsilon_t$$

$$\implies x_t = (I - \Phi_1 L - \Phi_2 L^2)^{-1} \epsilon_t = \sum_{j=0}^{\infty} A_j \epsilon_{t-j}$$

for some $\{A_j\}$.

(b) Now we seek to evaluate $(I - \Phi_1 L - \Phi_2 L^2)^{-1}$. To do this, let $(I - \Phi_1 L - \Phi_2 L^2)^{-1} = A_0 + A_1 L + A_2 L^2 + A_3 L^3 + \ldots$ and note that

$$(I - \Phi_1 L - \Phi_2 L^2)(I - \Phi_1 L - \Phi_2 L)^{-1} = I$$

$$\iff (I - \Phi_1 L - \Phi_2 L^2)(A_0 + A_1 L + A_2 L^2 + A_3 L^3 + \ldots + A_j L^j + \ldots) = I$$

$$\iff A_0 + (A_1 - \Phi_1) L + (A_2 - \Phi_1 A_1 - \Phi_2) L^2 + (A_3 - \Phi_1 A_2 - \Phi_2 A_1) L^3 + \ldots$$

$$+ (A_j - \Phi_1 A_{j-1} + \Phi_2 A_{j-2}) L^j + \ldots = I$$

In order for this equation to hold, all the lag terms must equal zero and the constant matrix A_0 must equal I.

$$\implies \boxed{A_0 = I}$$

$$A_1 - \Phi_1 = 0 \iff \boxed{A_1 = \Phi_1}$$

$$A_2 - \Phi_1 A_1 - \Phi_2 = 0 \iff A_2 - \Phi_1^2 - \Phi_2 = 0 \iff \boxed{A_2 = \Phi_1^2 + \Phi_2}$$

$$A_3 - \Phi_1 A_2 - \Phi_2 A_1 = 0 \iff A_3 - \Phi_1^3 - \Phi_1 \Phi_2 - \Phi_2 \Phi_1 = 0 \iff \boxed{A_3 = \Phi_1^3 + \Phi_1 \Phi_2 + \Phi_2 \Phi_1}$$

and, in general,

$$\boxed{A_j = \Phi_1 A_{j-1} + \Phi_2 A_{j-2}}$$

(c)   • **OIR:** We employ the Cholesky decomposition of Σ:

$$\Sigma = PP' \tag{39}$$

where P is a lower-triangular matrix. Then the MA representation (38) can be written as

$$x_t = \sum_{j=0}^{\infty} (A_j P)(P^{-1} u_{t-j}) = \sum_{j=0}^{\infty} B_j \eta_{t-j} \tag{40}$$

where $B_j = A_j P$, $\eta_t = P^{-1} u_t$, so we have

TIME SERIES

$$\mathbb{E}(\eta_t \eta_t') = P^{-1}\mathbb{E}(u_t u_t')(P^{-1})' = P^{-1}\Sigma(P^{-1})' = P^{-1}PP'(P^{-1})' = I_m$$

so the new errors $\eta_{1t}, \eta_{2t}, \ldots, \eta_{mt}$ are contemporaneously uncorrelated. Then the orthogonalized impact of a unit shock at time t to the ith equation on y at time $t + n$ is given by

$$B_n e_i, n = 0, 1, \ldots \tag{41}$$

where e_i is an $m \times 1$ selection vector. Written more compactly, the orthogonalized impulse response function of a unit (one standar error) shock to the ith variable on the jth variable is given by

$$OI_{ij,n} = e_j' A_n P e_i, \quad i, j = 1, 2, \ldots, m \tag{42}$$

These orthogonalized impulse responses are not unique and depend on the particular ordering of the variables in the VAR. The orthogonalized responses are invariant to the ordering of the variables only if Σ is diagonal.

- **GIR:** If the VAR model is perturbed by a shock of size $\delta_i = \sqrt{\sigma_{ii}}$ to its ith equation at time t, by the definition of the generalized IR function we have

$$GI_y(n, \delta_i, \Omega_{t-1}^0) = \mathbb{E}(y_t \mid u_{it} = \delta_i, \Omega_{t-1}^0) - \mathbb{E}(y_t \mid \Omega_{t-1}^0) \tag{43}$$

Once again using the MA(∞) representation (38 we obtain

$$GI_y(n, \delta_i, \Omega_{t-1}^0) = A_n \mathbb{E}(u_t \mid u_{it} = \delta_i) \tag{44}$$

which is history invariant (i.e. does not depend on $\Omega_{t=1}^0$. The computation of the conditional expectations $\mathbb{E}(u_t \mid u_{it} = \delta_i)$ depends on the nature of the multivariate distribution assumed for the disturbances u_t. In the case where $u_t \sim IID\mathcal{N}(0, \Sigma)$, we hae

$$\mathbb{E}(u_t \mid u_{it} = \delta_i) = \begin{bmatrix} \sigma_{1i}/\sigma_{ii} \\ \sigma_{2i}/\sigma_{ii} \\ \vdots \\ \sigma_{mi}/\sigma_{ii} \end{bmatrix} \delta_i \tag{45}$$

where as before $\Sigma = [\sigma_{ij}]$. Hence for a unit shock $\delta_i = \sqrt{\sigma_{ii}}$ we have

$$GI_y(n, \delta_i = \sqrt{\sigma_{ii}}, \Omega_{t-1}^0) = \frac{A_n \Sigma e_i}{\sqrt{\sigma_{ii}}}, \quad i, j = 1, 2, \ldots, m \tag{46}$$

The GIRF of a unit shock to the ith equation in the VAR(p) model

$$y_t = \Phi_1 y_{t-1} + \Phi_2 y_{t-2} + \ldots + \Phi_p y_{t-p} + u_t, \quad u_t \sim IID(0, \Sigma) \tag{47}$$

on the jth variable at horizon n is given by the jth element of (46), expressed more compactly by

$$GI_y(n, \delta_i = \sqrt{\sigma_{ii}}, \Omega_{t-1}^0) = \frac{e_j' A_n \Sigma e_i}{\sqrt{\sigma_{ii}}}, \quad i, j = 1, 2, \ldots, m \tag{48}$$

(d) The GIRF circumvents the problem of the dependence of the orthogonalized impulse responses to the ordering of the variables in the VAR. Unlike the OIR responses in (42), the GIR responses in (48) are invariant to the ordering of the variables in the VAR. The two responses coincide only for the first variable in the VAR or when Σ is diagonal.

Chapter 24 Problem 4.

Chapter 24 Problem 4 Solution.

1.12   Chapter 33: Theory and Practice of GVAR Modeling

Lidan says will not be on final.