Financial Econometrics
HT Week 3 Assignment Answers

February 2021

Exercise 4.18

Consider an MA(2)

$$Y_t = \mu + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \varepsilon_t$$

1. **What is a minimal set of assumptions sufficient to ensure $\{Y_t\}$ is covariance stationary if $\{\varepsilon_t\}$ is an i.i.d. sequence?**
   ε_t needs to have mean 0 and finite variance so that it is a White Noise process.

2. **What are the values of the following quantities?**

   (a) $\mathrm{E}[Y_{t+1}]$, $\mathrm{E}_t[Y_{t+1}]$ **and** $\mathrm{E}_t[Y_{t+2}]$

$$\begin{aligned}
\mathrm{E}[Y_{t+1}] &= \mathrm{E}[\mu + \theta_1 \varepsilon_t + \theta_2 \varepsilon_{t-1} + \varepsilon_{t+1}] \\
&= \mu + \theta_1 \mathrm{E}[\varepsilon_t] + \theta_2 \mathrm{E}[\varepsilon_{t-1}] + \mathrm{E}[\varepsilon_{t+1}] \\
&= \mu + \theta_1 \times 0 + \theta_2 \times 0 + 0 \\
&= \mu
\end{aligned}$$

$$\begin{aligned}
\mathrm{E}_t[Y_{t+1}] &= \mathrm{E}_t[\mu + \theta_1 \varepsilon_t + \theta_2 \varepsilon_{t-1} + \varepsilon_{t+1}] \\
&= \mu + \theta_1 \mathrm{E}_t[\varepsilon_t] + \theta_2 \mathrm{E}_t[\varepsilon_{t-1}] + \mathrm{E}_t[\varepsilon_{t+1}] \\
&= \mu + \theta_1 \varepsilon_t + \theta_2 \times \varepsilon_{t-1} + 0
\end{aligned}$$

$$\begin{aligned}
\mathrm{E}_t[Y_{t+2}] &= \mathrm{E}_t[\mu + \theta_1 \varepsilon_{t+1} + \theta_2 \varepsilon_t + \varepsilon_{t+2}] \\
&= \mu + \theta_1 \mathrm{E}_t[\varepsilon_{t+1}] + \theta_2 \mathrm{E}_t[\varepsilon_t] + \mathrm{E}_t[\varepsilon_{t+2}] \\
&= \mu + \theta_1 \times 0 + \theta_2 \varepsilon_t + 0 \\
&= \mu + \theta_2 \varepsilon_t
\end{aligned}$$

   (b) $\mathrm{V}[Y_{t+1}]$, $\mathrm{V}_t[Y_{t+1}]$, $\mathrm{V}_t[Y_{t+2}]$

$$\begin{aligned}
Y_{t+1} - \mathrm{E}[Y_{t+1}] &= \mu + \theta_1 \varepsilon_t + \theta_2 \varepsilon_{t-1} + \varepsilon_{t+1} - \mu \\
&= \theta_1 \varepsilon_t + \theta_2 \varepsilon_{t-1} + \varepsilon_{t+1} \\
\mathrm{V}[Y_{t+1}] &= \theta_1^2 \mathrm{E}\left[\varepsilon_t^2\right] + \theta_2^2 \mathrm{E}\left[\varepsilon_{t-1}^2\right] + \mathrm{E}\left[\varepsilon_{t+1}^2\right] \\
&= \sigma^2\left(1 + \theta_1^2 + \theta_2^2\right)
\end{aligned}$$

since cross-products have 0 expectation due to WN assumption above.

$$\begin{aligned} Y_{t+1} - \mathrm{E}_t\left[Y_{t+1}\right] &= \mu + \theta_1 \varepsilon_t + \theta_2 \varepsilon_{t-1} + \varepsilon_{t+1} - \mu - \theta_1 \varepsilon_t - \theta_2 \varepsilon_{t-1} \\ &= \varepsilon_{t+1} \\ \mathrm{V}\left[Y_{t+1}\right] &= \mathrm{E}_t\left[\varepsilon_{t+1}^2\right] \\ &= \sigma^2 \textbf{ if homoskedastic} \end{aligned}$$

$$\begin{aligned} Y_{t+2} - \mathrm{E}_t\left[Y_{t+2}\right] &= \mu + \theta_1 \varepsilon_{t+1} + \theta_2 \varepsilon_t + \varepsilon_{t+2} - \mu - \theta_2 \varepsilon_t \\ &= \theta_1 \varepsilon_{t+1} + \varepsilon_{t+2} \\ \mathrm{V}_t\left[Y_{t+2}\right] &= \mathrm{E}_t\left[\left(\theta_1 \varepsilon_{t+1} + \varepsilon_{t+2}\right)^2\right] \\ &= \theta_1^2 \mathrm{E}_t\left[\varepsilon_{t+1}^2\right] + \mathrm{E}_t\left[\varepsilon_{t+2}^2\right] \textbf{ since WN} \\ &= \sigma^2\left(1+\theta_1^2\right) \textbf{ if homoskedastic} \end{aligned}$$

(c) ρ_h**,** $h = 1, 2, 3, 4, \ldots$

$$\begin{aligned} \rho_1 &= \frac{\theta_1 + \theta_1 \theta_2}{\left(1 + \theta_1^2 + \theta_2^2\right)} \\ \rho_2 &= \frac{\theta_2}{\left(1 + \theta_1^2 + \theta_2^2\right)} \\ \rho_j &= j,\ j \geq 3 \end{aligned}$$

The autocovariances (numerators) are

$$\begin{aligned} \mathrm{Cov}\left(Y_t, Y_{t-1}\right) &= \mathrm{Cov}\left(\mu + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \varepsilon_t, \mu + \theta_1 \varepsilon_{t-2} + \theta_2 \varepsilon_{t-3} + \varepsilon_{t-1}\right) \\ &= \mathrm{Cov}\left(\theta_1 \varepsilon_{t-1}, \varepsilon_{t-1}\right) + \mathrm{Cov}\left(\theta_2 \varepsilon_{t-2}, \theta_1 \varepsilon_{t-2}\right) \text{ where crosses drop due to WN} \\ &= \theta_1 \sigma^2 + \theta_1 \theta_2 \sigma^2 \end{aligned}$$

$$\begin{aligned} \mathrm{Cov}\left(Y_t, Y_{t-2}\right) &= \mathrm{Cov}\left(\mu + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \varepsilon_t, \mu + \theta_1 \varepsilon_{t-3} + \theta_2 \varepsilon_{t-4} + \varepsilon_{t-2}\right) \\ &= \mathrm{Cov}\left(\theta_2 \varepsilon_{t-2}, \varepsilon_{t-2}\right) \text{ where crosses drop due to WN} \\ &= \theta_2 \sigma^2 \end{aligned}$$

and higher order are clearly 0.

Exercise 4.34

1. **Outline the steps needed to perform a unit root test on as time-series of FX rates. Be sure to detail the any important considerations that may affect the test.**

   The goal of a unit root test, or any other test, is to provide a powerful test (large chance of rejecting if the alternative is true) while maintaining good size properties (knowing the critical values to control the probability of rejecting the null when it is in fact true). Unit root testing has some unique issues but general principals of testing apply:

   - Exclude irrelevant regressors to avoid losing power.
   - Include relevant regressors to lower the error variance and ensure that the errors are white noise.

   Beyond these two generic testing considerations, unit root tests are very sensitive to the inclusion of deterministic terms. Consider these two versions of an ADF:

   $$\Delta y_t = \gamma y_{t-1} + \pi_1 \Delta y_{t-1} + \ldots + \pi_P \Delta y_{t-P} + \varepsilon_t$$

   and

   $$\Delta y_t = \pi_0 + \gamma y_{t-1} + \pi_1 \Delta y_{t-1} + \ldots + \pi_P \Delta y_{t-P} + \varepsilon_t$$

   The second only differs from the inclusion of a constant. In standard regression problems, this change makes no difference. Unit root testing is different and critical values depend on whether the null model contains a constant, and including deterministic terms lowers the critical value. Therefore including a constant require larger t-statistics to reject. This results in a loss of power and leads a researcher to be hesitant to include a constant in an ADF test.

   The other side of the "constant" choice is what happens when a constant is excluded but needed. In this case the test has asymptotically no power, a very bad thing. The intuition behind this result is simple. If the process has a time trend but not a unit root, excluding the time trend forces the process to behave as if it were a unit root.

   To balance between these two considerations, it is usually a good idea to use a loose selection criteria when deciding to include deterministic trends. Once the deterministic trend decision has been made, a loose selection criteria on the number of included lags of Δy can be used to ensure the errors are as close to white noise as possible without killing power.