Time Series Econometrics
Supplementary Lecture 5

Yukai Yang

Department of Statistics
Uppsala University

1 Exercise 11.5

Consider the following bivariate VAR:

$$
\begin{aligned}
y_{1,t} &= 0.3y_{1,t-1} + 0.8y_{2,t-1} + \varepsilon_{1,t} \\
y_{2,t} &= 0.9y_{1,t-1} + 0.4y_{2,t-1} + \varepsilon_{2,t},
\end{aligned}
$$

with $\boldsymbol{\varepsilon} = (\varepsilon_{1,t}, \varepsilon_{2,t})'$ *and*

$$E(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_\tau') = \mathbf{\Omega} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$

for $t = \tau$ *and* $\mathbf{0}$ *otherwise.*

(a) Is this system covariance-stationary?

(b) Calculate $\mathbf{\Psi}_s = \partial \mathbf{y}_{t+s} / \partial \boldsymbol{\varepsilon}_t'$ *for* $s = 0, 1$ *and* 2. *What is the limit as* $s \to \infty$*?*

(c) Calculate the fraction of the MSE of the two-period ahead forecast error for variable 1,

$$E[y_{1,t+2} - \hat{E}(y_{1,t+2} | y_t, y_{t-1}, \dots)]^2$$

that is due to $\varepsilon_{1,t+1}$ *and* $\varepsilon_{1,t+2}$.

1.1 (a) Stationarity

First off, the VAR in matrix form is

$$
\begin{aligned}
\begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix} &= \begin{pmatrix} 0.3 & 0.8 \\ 0.9 & 0.4 \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix} \\
\mathbf{y}_t &= \mathbf{\Phi} \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t
\end{aligned}
$$

The process is stationary if the eigenvalues of $\mathbf{\Phi}$ are inside the unit circle (p. 259).

$$
\begin{aligned}
|\mathbf{\Phi} - \lambda \mathbf{I}| = \begin{vmatrix} 0.3-\lambda & 0.8 \\ 0.9 & 0.4-\lambda \end{vmatrix} &= 0 \\
(0.3-\lambda)(0.4-\lambda) - 0.8 \times 0.9 &= 0 \\
\lambda^2 - 0.7\lambda - 0.6 &= 0.
\end{aligned}
$$

The quadratic formula gives us

$$
\begin{aligned}
\lambda &= \frac{0.7 \pm \sqrt{0.7^2 + 4 \times 0.6}}{2} = 0.35 \pm \frac{\sqrt{2.89}}{2} = 0.35 \pm 0.85 \\
&= \begin{cases} 1.2 \\ -0.5 \end{cases}
\end{aligned}
$$

One of the eigenvalues is 1.2, which is outside of the unit circle. Hence, this process is non-stationary.

1.2 (b) Impulse response

From p. 260 we know that

$$
\begin{aligned}
\mathbf{\Psi}_0 &= \mathbf{I} \\
\mathbf{\Psi}_1 &= \mathbf{\Phi}_1 \\
\mathbf{\Psi}_2 &= \mathbf{\Phi}_1 \mathbf{\Psi}_1 + \mathbf{\Phi}_2.
\end{aligned}
$$

In our case, we have no $\mathbf{\Phi}_2$, so for us

$$
\begin{aligned}
\mathbf{\Psi}_0 &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\
\mathbf{\Psi}_1 &= \begin{pmatrix} 0.3 & 0.8 \\ 0.9 & 0.4 \end{pmatrix} \\
\mathbf{\Psi}_2 &= \begin{pmatrix} 0.3 & 0.8 \\ 0.9 & 0.4 \end{pmatrix} \begin{pmatrix} 0.3 & 0.8 \\ 0.9 & 0.4 \end{pmatrix} = \begin{pmatrix} 0.3^2 + 0.8 \times 0.9 & 0.3 \times 0.8 + 0.8 \times 0.4 \\ 0.9 \times 0.3 + 0.4 \times 0.9 & 0.9 \times 0.8 + 0.4^2 \end{pmatrix} \\
&= \begin{pmatrix} 0.81 & 0.56 \\ 0.63 & 0.88 \end{pmatrix}
\end{aligned}
$$

But what is the limit $\lim_{s \to \infty} \mathbf{\Psi}_s$? Since this is a VAR(1), $\mathbf{\Psi}_s = \mathbf{\Phi}^s$. Recall that for a matrix $\mathbf{A}$ with distinct eigenvalues we can decompose it: $\mathbf{A} = \mathbf{T}\mathbf{\Lambda}\mathbf{T}^{-1}$, where $\mathbf{T}$ is a matrix with the eigenvectors of $\mathbf{A}$ and $\mathbf{\Lambda}$ is a diagonal matrix with its elements being the eigenvalues. In our case,

$$
\begin{aligned}
\mathbf{\Psi}_s = \mathbf{\Phi}^s = \mathbf{T}\mathbf{\Lambda}^s\mathbf{T}^{-1} &= \mathbf{T} \begin{pmatrix} 1.2 & 0 \\ 0 & -0.5 \end{pmatrix}^s \mathbf{T}^{-1} \\
&= \mathbf{T} \begin{pmatrix} 1.2^s & 0 \\ 0 & (-0.5)^s \end{pmatrix} \mathbf{T}^{-1}
\end{aligned}
$$

Clearly, $\lim_{s\to\infty} 1.2^s = \infty$, so the matrix $\mathbf{\Psi}_s$ will get larger and larger and thus exhibit an explosive behavior. The non-stationarity of the process means the VMA weights are not convergent, and it also means that they will not be absolute summable.

1.3 (c) Variance decomposition

In general, we have

$$MSE(\hat{\mathbf{y}}_{t+s|t}) = E\left[(\mathbf{y}_{t+s} - \hat{\mathbf{y}}_{t+s|t})(\mathbf{y}_{t+s} - \hat{\mathbf{y}}_{t+s|t})'\right] = \sum_{j=0}^{s-1} \mathbf{\Psi}_j \mathbf{\Omega} \mathbf{\Psi}_j'$$

In this case, $s = 2$ so

$$MSE(\hat{\mathbf{y}}_{t+2|t}) = \sum_{j=0}^{1} \mathbf{\Psi}_j \mathbf{\Omega} \mathbf{\Psi}_j' = \mathbf{\Omega} + \mathbf{\Psi}_1 \mathbf{\Omega} \mathbf{\Psi}_1' \tag{1}$$

For ease of notation, let

$$\mathbf{\Psi}_1 = \begin{pmatrix} \psi_{1,1} & \psi_{1,2} \\ \psi_{2,1} & \psi_{2,2} \end{pmatrix}, \quad \mathbf{\Omega} = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}.$$

Then the second term in (1) is

$$\begin{aligned}
\mathbf{\Psi}_1 \mathbf{\Omega} \mathbf{\Psi}_1' &= \begin{pmatrix} \psi_{1,1} & \psi_{1,2} \\ \psi_{2,1} & \psi_{2,2} \end{pmatrix} \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix} \begin{pmatrix} \psi_{1,1} & \psi_{2,1} \\ \psi_{1,2} & \psi_{2,2} \end{pmatrix} \\
&= \begin{pmatrix} \psi_{1,1} & \psi_{1,2} \\ \psi_{2,1} & \psi_{2,2} \end{pmatrix} \begin{pmatrix} \psi_{1,1}\sigma_1^2 & \psi_{2,1}\sigma_1^2 \\ \psi_{1,2}\sigma_2^2 & \psi_{2,2}\sigma_2^2 \end{pmatrix} \\
&= \begin{pmatrix} \psi_{1,1}^2\sigma_1^2 + \psi_{1,2}^2\sigma_2^2 & \psi_{2,1}\psi_{1,1}\sigma_1^2 + \psi_{2,2}\psi_{1,2}\sigma_2^2 \\ \psi_{2,1}\psi_{1,1}\sigma_1^2 + \psi_{2,2}\psi_{1,2}\sigma_2^2 & \psi_{2,1}^2\sigma_1^2 + \psi_{2,2}^2\sigma_2^2 \end{pmatrix}
\end{aligned}$$

To get (1), we add $\mathbf{\Omega}$ to this and get:

$$\mathbf{\Omega} + \mathbf{\Psi}_1 \mathbf{\Omega} \mathbf{\Psi}_1' = \begin{pmatrix} (1 + \psi_{1,1}^2)\sigma_1^2 + \psi_{1,2}^2\sigma_2^2 & \psi_{2,1}\psi_{1,1}\sigma_1^2 + \psi_{2,2}\psi_{1,2}\sigma_2^2 \\ \psi_{2,1}\psi_{1,1}\sigma_1^2 + \psi_{2,2}\psi_{1,2}\sigma_2^2 & \psi_{2,1}^2\sigma_1^2 + (1 + \psi_{2,2}^2)\sigma_2^2 \end{pmatrix}$$

The diagonal gives us the variances of the forecast errors two periods ahead. The (1, 1) element expresses the variance of the forecast error of variable 1 as a function of the error variances of variable 1 and variable 2. The first term in the (1, 1) element, $(1 + \psi_{1,1}^2)\sigma_1^2$, is the part of the MSE that is due to variable 1, and $\psi_{1,2}\sigma_2^2$ is the part that is due to variable 2. Hence, the fraction of the MSE of the two-period ahead forecast error for variable 1 that is due to $\varepsilon_{1,t+2}$ and $\varepsilon_{1,t+1}$ is

$$\frac{(1 + \psi_{1,1}^2)\sigma_1^2}{(1 + \psi_{1,1}^2)\sigma_1^2 + \psi_{1,2}^2\sigma_2^2} = \frac{(1 + 0.3^2) \times 1}{(1 + 0.3^2) \times 1 + 0.8^2 \times 2} = \frac{1.09}{2.37} = 0.46$$

Thus, 46 % of the variance of the forecast error of variable 1 two periods ahead is due to its own innovations.

2 Exercise 16.3

Let y_t be covariance stationary with mean zero and absolutely summable autocovariances:

$$\sum_{j=-\infty}^{\infty} |\gamma_j| < \infty$$

for $\gamma_j = E(y_t y_{t-j})$. Adapting the argument in expression [7.2.6], show that

$$\frac{1}{T}\sum_{t=1}^{T} \frac{t}{T} y_t \xrightarrow{m.s.} 0.$$

First, recall the concept of mean square convergence. We say that

$$X_T \xrightarrow{m.s.} c$$

if, for every $\varepsilon > 0$, there exists an N such that, for all $T \geq N$,

$$E(X_T - c)^2 < \varepsilon.$$

This is a mode of convergence which is stronger than convergence in probability (and hence stronger than convergence in distribution, too). From Proposition 7.2, the "generalized Chebyshev's inequality", it follows that if $X_T \xrightarrow{m.s.} c$, then

$$P(|X_T - c| > \delta) \leq \frac{E(X - c)^2}{\delta^2}.$$

But under the assumption that $X_T \xrightarrow{m.s.} c$, this means that

$$E(X_T - c)^2 < \varepsilon\delta^2$$

for $T \geq N$ for some N. Hence, we can say that

$$P(|X_T - c| > \delta) \leq \frac{E(X - c)^2}{\delta^2} < \varepsilon$$

Thus, mean square convergence implies

$$P(|X_T - c| > \delta) < \varepsilon$$

which means $X_T \xrightarrow{p} c$.

In the context of the exercise, the mean square convergence is

$$\frac{1}{T}\sum_{t=1}^{T} \frac{t}{T} y_t \xrightarrow{m.s.} 0 \quad \text{if} \quad E\left(\left|\frac{1}{T}\sum_{t=1}^{T} \frac{t}{T} y_t\right|^2\right) \to 0$$

This term arises in the derivation of [16.3.13], the asymptotic distribution for the estimators of the transformed regression.

$$
\begin{aligned}
E\left(\left|\frac{1}{T}\sum_{t=1}^{T}\frac{t}{T}y_t\right|^2\right) &= \frac{1}{T^4}E\left(\left|\sum_{t=1}^{T}ty_t\right|^2\right) = \frac{1}{T^2}E\left(\frac{\sum_{s=1}^{T}sy_s\sum_{t=1}^{T}ty_t}{T^2}\right) \\
&= \frac{1}{T^2}E\left(\frac{\sum_{t=1}^{T}t^2y_t^2}{T^2}+\frac{2\sum_{t=2}^{T}\sum_{a=1}^{t-1}tsy_ty_s}{T^2}\right) \\
&= \frac{1}{T^2}\left(\frac{\sum_{t=1}^{T}t^2E(y_t^2)}{T^2}+\frac{2\sum_{t=2}^{T}\sum_{s=1}^{t-1}tsE(y_ty_s)}{T^2}\right) \\
&= \frac{1}{T}\left(\frac{\gamma_0\sum_{t=1}^{T}t^2}{T^3}+\frac{2\sum_{t=2}^{T}\sum_{s=1}^{t-1}ts\gamma_{t-s}}{T^3}\right)
\end{aligned}
$$

Remember that $\sum_{t=1}^{T}t^2 = \dfrac{T(T+1)(2T+1)}{6}$, thus meaning that

$$
\begin{aligned}
\frac{\gamma_0\sum_{t=1}^{T}t^2}{T^3} &= \gamma_0\frac{T(T+1)(2T+1)}{6T^3} = \gamma_0\left(\frac{\left(1+\frac{1}{T}\right)\left(2+\frac{1}{T}\right)}{6}\right) \\
&\leq |\gamma_0|
\end{aligned}
$$

since the term in brackets is less than 1. We also have that

$$
\frac{2\sum_{t=2}^{T}\sum_{s=1}^{t-1}ts\gamma_{t-s}}{T^3} \leq \frac{2\sum_{t=2}^{T}\sum_{s=1}^{t-1}T(T-1)|\gamma_{t-s}|}{T^3} = \frac{2T(T-1)\sum_{t=2}^{T}\sum_{s=1}^{t-1}|\gamma_{t-s}|}{T^3}.
$$

Note that the two sums will give us $T-1$ $|\gamma_1|$ terms, $T-2$ $|\gamma_2|$ terms and so on. If we let $t-s=j$, then we can write this as

$$
\begin{aligned}
\frac{2T(T-1)\sum_{j=1}^{T-1}(T-j)|\gamma_j|}{T^3} &\leq \frac{2T(T-1)\sum_{j=1}^{T-1}T|\gamma_j|}{T^3} \leq \frac{2T^3\sum_{j=1}^{T-1}|\gamma_j|}{T^3} \\
&= 2\sum_{j=1}^{T-1}|\gamma_j|
\end{aligned}
$$

Putting this together,

$$
\begin{aligned}
E\left(\left|\frac{1}{T}\sum_{t=1}^{T}\frac{t}{T}y_t\right|^2\right) &= \frac{1}{T}\left(\frac{\gamma_0\sum_{t=1}^{T}t^2}{T^3}+\frac{2\sum_{t=2}^{T}\sum_{s=1}^{t-1}ts\gamma_{t-s}}{T^3}\right) \\
&\leq \frac{1}{T}\left(|\gamma_0|+2\sum_{j=1}^{T-1}|\gamma_j|\right)
\end{aligned}
$$

Since $\left(\sum_{j=-\infty}^{\infty}|\gamma_j|\right)$ is finite by absolute summability, then we can say that $\left(|\gamma_0|+2\sum_{j=1}^{T-1}|\gamma_j|\right) < M$ for some finite constant M. Consequently,

$$E\left(\left|\frac{1}{T}\sum_{t=1}^{T}\frac{t}{T}y_t\right|^2\right) < \frac{M}{T} \to 0 \quad \text{as } T \to \infty,$$

meaning that

$$\frac{1}{T}\sum_{t=1}^{T}\frac{t}{T}y_t \xrightarrow{m.s.} 0.$$

3 Exercise 17.1

Let $\{u_t\}$ be an i.i.d. sequence with mean zero and variance σ^2, and let $y_t = u_1 + u_2 + \cdots + u_t$ with $y_0 = 0$. Deduce from [17.3.17] and [17.3.18] that

$$\begin{pmatrix} T^{-1/2}\sum_{t=1}^{T} u_t \\ T^{-3/2}\sum_{t=1}^{T} y_{t-1} \end{pmatrix} \xrightarrow{L} N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \sigma^2\begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} \end{pmatrix}\right).$$

Comparing this result with Proposition 17.1, argue that

$$\begin{pmatrix} W(1) \\ \int_0^1 W(r)dr \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} \end{pmatrix}\right).$$

First, consider the definition of a Brownian motion.

Definition. *Standard Brownian motion $W(\cdot)$ is a continuous-time stochastic process, associating each date $t \in [0,1]$ with the scalar $W(t)$ such that:*

(a) $W(0) = 0$*;*

(b) For any dates $0 \le t_1 < t_2 < \cdots < t_k \le 1$, the changes $[W(t_2) - W(t_1)], [W(t_3) - W(t_2)], \ldots, [W(t_k) - W(t_{k-1})]$ are independent multivariate Gaussian with $[W(s) - W(t)] \sim N(0, s-t)$;

(c) For any given realization, $W(t)$ is continuous in t with probability 1.

This is used a lot in dealing with unit root processes. With a unit root, we have a random walk defined by

$$y_t = y_{t-1} + u_t$$

where $\{u_t\}$ is an i.i.d. sequence, and $E(u_t) = 0$ and $E(u_t^2) = \sigma^2$. If $y_0 = 0$, then

$$y_t = u_1 + u_2 + \cdots + u_t.$$

Here, we can define a function

$$X_T(r) = \begin{cases} 0, & \text{for } 0 \le r < 1/T \\ y_1/T, & \text{for } 1/T \le r < 2/T \\ y_2/T, & \text{for } 2/T \le r < 3/T \\ \vdots & \\ y_T/T, & \text{for } r = 1. \end{cases}$$

Here, $r \in [0,1]$ denotes the fraction of the sample (which is of size T). Consider now the integral of this:

$$\begin{aligned} \int_0^1 X_T(r)dr &= \int_{1/T}^{2/T} \frac{y_1}{T}dr + \int_{2/T}^{3/T} \frac{y_2}{T}dr + \cdots + \int_{(T-1)/T}^{1} \frac{y_{T-1}}{T}dr + \int_1^1 \frac{y_T}{T}dr \\ &= \frac{y_1}{T^2} + \frac{y_2}{T^2} + \cdots + \frac{y_{T-1}}{T^2} \\ &= \sum_{t=1}^{T-1} \frac{y_t}{T^2}. \end{aligned}$$

However, from a previous result (in [17.3.8]) it is known that

$$\sqrt{T}X_T(\cdot) \xrightarrow{L} \sigma W(\cdot).$$

Integrating both sides over r allows for an application of the continuous mapping theorem, such that

$$\int_0^1 \sqrt{T}X_T(r)dr \xrightarrow{L} \sigma \int_0^1 W(r)dr.$$

If we go back to our previous equation and multiply by $\sqrt{T}$, we get

$$\int_0^1 \sqrt{T}X_T(r)dr = \sum_{t=1}^{T-1} \frac{y_t}{T^{3/2}} \xrightarrow{L} \sigma \int_0^1 W(r)dr.$$

We are to deduce the convergence from [17.3.17] and [17.3.18], where [17.3.18] is

$$\begin{pmatrix} T^{-1/2}\sum_{t=1}^T u_t \\ T^{-3/2}\sum_{t=1}^T tu_t \end{pmatrix} \xrightarrow{L} N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \sigma^2 \begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} \end{pmatrix}\right).$$

With the help of [17.3.17] we can see that

$$
\begin{aligned}
T^{-3/2}\sum_{t=1}^{T} y_{t-1} &= T^{-3/2}\sum_{t=1}^{T}\sum_{s=1}^{t-1} u_s \\
&= T^{-3/2}\sum_{t=1}^{T}[u_1+u_2+u_3+\cdots+u_{t-1}] \\
&= T^{-3/2}\left[(u_1)+(u_1+u_2)+(u_1+u_2+u_3)+\cdots+(u_1+u_2+\cdots+u_{T-1})\right] \\
&= T^{-3/2}\left[(T-1)u_1+(T-2)u_2+(T-3)u_3+\cdots+(T-(T-1))u_{T-1}\right] \\
&= T^{-3/2}\sum_{t=1}^{T}(T-t)u_t \\
&= T^{-1/2}\sum_{t=1}^{T}u_t - T^{-3/2}\sum_{t=1}^{T}tu_t.
\end{aligned}
$$

The distributions of these two terms are given above in [17.3.18]. Thus, asymptotically, $T^{-3/2}\sum_{t=1}^{T} y_{t-1}$ will be normal with zero mean and variance equal to

$$
V\left(T^{-3/2}\sum_{t=1}^{T} y_{t-1}\right) \xrightarrow{p} \sigma^2(1+1/3-2\times 1/2) = \frac{\sigma^2}{3}.
$$

Thus, from [17.3.17] and [17.3.18] we have arrived at

$$
\begin{pmatrix} T^{-1/2}\sum_{t=1}^{T} u_t \\ T^{-3/2}\sum_{t=1}^{T} y_{t-1} \end{pmatrix} \xrightarrow{L} N\left(\begin{pmatrix}0\\0\end{pmatrix}, \sigma^2\begin{pmatrix}1 & \frac{1}{2}\\ \frac{1}{2} & \frac{1}{3}\end{pmatrix}\right).
$$

We have also already seen a kind of motivation for the connection to Brownian motions, but as per the question we will use Proposition 17.1 to come to this conclusion a little more formally.

Proposition (17.1)**.** *Suppose that ξ_t follows a random walk without drift,*

$$
\xi_t = \xi_{t-1} + u_t,
$$

where $\xi_0 = 0$ and $\{u_t\}$ is an i.i.d. sequence with mean zero and variance σ^2. Then

(a) $T^{-1/2}\sum_{t=1}^{T} u_t \xrightarrow{L} \sigma W(1)$;

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

Since we now have established convergence to two "different" asymptotic distributions, this must mean that they are the same. Hence,

$$
\begin{pmatrix} T^{-1/2}\sum_{t=1}^{T} u_t \\ T^{-3/2}\sum_{t=1}^{T} y_{t-1} \end{pmatrix} \xrightarrow{L} \begin{pmatrix}\sigma W(1) \\ \sigma\int_0^1 W(r)dr\end{pmatrix} \sim N\left(\begin{pmatrix}0\\0\end{pmatrix}, \sigma^2\begin{pmatrix}1 & \frac{1}{2}\\ \frac{1}{2} & \frac{1}{3}\end{pmatrix}\right)
$$

4 Exercise 17.2

Phillips (1987) generalization of case 1. Suppose that data are generated from the process $y_t = y_{t-1} + u_t$, where $u_t = \psi(L)\varepsilon_t$, $\sum_{j=0}^{\infty} j \cdot |\psi_j| < \infty$, and ε_t are i.i.d with mean zero and variance σ^2, and finite fourth moment. Consider OLS estimation of the autoregression $y_t = \rho y_{t-1} + u_t$. Let $\hat{\rho}_T = \left(\sum y_{t-1}^2\right)^{-1}\left(\sum y_{t-1} y_t\right)$ be the OLS estimate of ρ, $s_T^2 = (T-1)^{-1} \times \sum \hat{u}_t^2$ the OLS estimate of the variance of the regression error, $\hat{\sigma}^2_{\hat{\rho}_T} = s_T^2 \left(\sum y_{t-1}^2\right)^{-1}$ the OLS estimate of the variance of $\hat{\rho}_T$ and $t_T = (\hat{\rho}_T - 1)/\hat{\sigma}_{\hat{\rho}_T}$ the OLS t test of $\rho = 1$ and define $\lambda = \sigma\psi(1)$. Use Proposition 17.3 to show that

(a)
$$T(\hat{\rho}_T - 1) \xrightarrow{L} \frac{\frac{1}{2}(\lambda^2 [W(1)]^2 - \gamma_0)}{\lambda^2 \int_0^1 [W(r)]^2 dr}$$

(b)
$$T^2 \hat{\sigma}^2_{\hat{\rho}_T} \xrightarrow{L} \frac{\gamma_0}{\lambda^2 \int [W(r)]^2 dr}$$

(c)
$$t_T = \frac{T(\hat{\rho}_T - 1)}{\sqrt{T^2 \hat{\sigma}^2_{\hat{\rho}_T}}} \xrightarrow{L} \left(\frac{\lambda^2}{\gamma_0}\right)^{1/2} \left\{ \frac{\frac{1}{2}\left([W(1)]^2 - 1\right)}{\left(\int_0^1 [W(r)]^2 dr\right)^{1/2}} + \frac{\frac{1}{2}(\lambda^2 - \gamma_0)}{\lambda^2 \left(\int_0^1 [W(r)]^2 dr\right)^{1/2}} \right\}$$

(d)
$$T(\hat{\rho}_T - 1) - \frac{1}{2}\frac{T^2 \hat{\sigma}_{\hat{\rho}_T}}{s_T^2}(\lambda^2 - \gamma_0) \xrightarrow{L} \frac{\frac{1}{2}\left([W(1)]^2 - 1\right)}{\int_0^1 [W(r)]^2 dr}$$

(e)
$$\left(\frac{\gamma_0}{\lambda^2}\right)^{1/2} t_T - \frac{\frac{1}{2}(\lambda^2 - \gamma_0)}{\lambda}\frac{T\hat{\sigma}_{\hat{\rho}_T}}{s_T} \xrightarrow{L} \frac{\frac{1}{2}\left([W(1)]^2 - 1\right)}{\left(\int_0^1 [W(r)]^2 dr\right)^{1/2}}$$

4.1 The meaning of the convergences

Before starting, it's worth trying to understand what the point with these are or else they will simply be complicated expressions. Remember that we now are considering a case with some form of serially correlated error terms. In Section 17.4, we studied "case 1", where there was no constant or time trend and the true process was a random walk **without** serially correlated errors. Call the estimator in this case ρ_T^* (see p. 488). This led to

$$T(\hat{\rho}_T^* - 1) \xrightarrow{L} \frac{\frac{1}{2}\left([W(1)]^2 - 1\right)}{\int_0^1 [W(r)]^2 dr} \tag{2}$$

The distribution of $T(\hat{\rho}_T^* - 1)$ is tabulated in the appendix, table B.5. Compare this to the limit of $T(\hat{\rho} - 1)$. Note that this limit depends on λ and γ_0, which is something that is specific for each type of serial correlation. The distribution in the case of serial correlation thus depends on the specific type of correlation structure, and we cannot use this test as some sort of general approach. If u_t would be an MA(4) we would have one limit distribution, but if u_t was an AR(2), we would have another limit distribution.

Next, have a look at the "t-test" in the (c) part of the question. This is the ratio of the term in (a) and the square root of the term in (b). So to find the limit of (c), we first need (b). For case 1 again, call the t statistic in that case t_T^*. This was found to have the limit distribution (p. 489):

$$t_T^* \xrightarrow{L} \frac{\frac{1}{2}\left([W(1)]^2 - 1\right)}{\left(\int_0^1 [W(r)]^2 dr\right)^{1/2}} \tag{3}$$

This can also be tabulated, and has been so in the appendix, table B.6.

Lastly, we have the two complicated expressions in (d) and (e). But if we look at their limits, we can note that the limit in (d) is (2), and the limit in (e) is (3). So in the end, what we can see is that if we have serially correlated errors in a random walk like in this question, we can "correct" the terms (a) and (c), which were used in the case of no serial correlation, in such a way that we can use the same limit distributions. Of course, we do not know λ and γ_0 which are used in the correction terms, but we can consistently estimate them so that is no problem (see p. 511).

4.2 (a)

To show this, we can simply apply Proposition 17.3 appropriately. First, rewrite the estimator as

$$\begin{aligned}
T(\hat{\rho}_T - 1) &= T\left(\frac{\sum_{t=1}^T y_{t-1}y_t}{\sum_{t=1}^T y_{t-1}^2} - 1\right) \\
&= T\left(\frac{\sum_{t=1}^T y_{t-1}(y_{t-1} + u_t)}{\sum_{t=1}^T y_{t-1}^2} - 1\right) \\
&= T\left(\frac{\sum_{t=1}^T y_{t-1}^2}{\sum_{t=1}^T y_{t-1}^2} + \frac{\sum_{t=1}^T y_{t-1}u_t}{\sum_{t=1}^T y_{t-1}^2} - 1\right) \\
&= \frac{T^{-1}}{T^{-2}} \frac{\sum_{t=1}^T y_{t-1}u_t}{\sum_{t=1}^T y_{t-1}^2}.
\end{aligned}$$

Next, note that the numerator is the $j = 0$ case from Proposition17.3(e),

$$T^{-1}\sum_{t=1}^T \xi_{t-1}u_t \xrightarrow{L} \frac{1}{2}\left(\lambda^2[W(1)]^2 - \gamma_0\right) \tag{Prop 17.3(e)}$$

where $\xi_t = y_t$. Similarly, the denominator is

$$T^{-2}\sum_{t=1}^T y_{t-1}^2 \xrightarrow{L} \lambda^2 \int_0^1 [W(r)]^2 dr \tag{Prop 17.3(h)}$$

So together we get

$$T(\hat{\rho}_T - 1) = \frac{T^{-1}}{T^{-2}} \frac{\sum_{t=1}^{T} y_{t-1} u_t}{\sum_{t=1}^{T} y_{t-1}^2} \xrightarrow{L} \frac{\frac{1}{2}(\lambda^2 [W(1)]^2 - \gamma_0)}{\lambda^2 \int_0^1 [W(r)]^2 dr}$$

4.3 (b)

Rewrite it

$$T^2 \hat{\sigma}^2_{\hat{\rho}_T} = T^2 \frac{s_T^2}{\sum_{t=1}^{T} y_{t-1}^2} = \frac{s_T^2}{T^{-2} \sum_{t=1}^{T} y_{t-1}^2}.$$

The numerator converges according to [17.6.10] to [1]

$$s_T^2 \xrightarrow{p} E(u^2) = \gamma_0,$$

and the denominator again by Proposition 17.3(h) so that we get

$$T^2 \hat{\sigma}^2_{\hat{\rho}_T} = \frac{(T-1)^{-1} \sum_{t=1}^{T} \hat{u}_t^2}{T^{-2} \sum_{t=1}^{T} y_{t-1}^2} \xrightarrow{L} \frac{\gamma_0}{\lambda^2 \int_0^1 [W(r)]^2 dr}$$

4.4 (c)

Now we have:

$$\begin{aligned}
t_T = \frac{T(\hat{\rho}_T - 1)}{\sqrt{T^2 \hat{\sigma}^2_{\hat{\rho}_T}}} &\xrightarrow{L} \frac{\frac{\frac{1}{2}(\lambda^2 [W(1)]^2 - \gamma_0)}{\lambda^2 \int_0^1 [W(r)]^2 dr}}{\sqrt{\frac{\gamma_0}{\lambda^2 \int_0^1 [W(r)]^2 dr}}} = \left(\frac{1}{\gamma_0}\right)^{1/2} \frac{\frac{1}{2}(\lambda^2 [W(1)]^2 - \gamma_0)}{\sqrt{\lambda^2 \int_0^1 [W(r)]^2 dr}} \\
&= \left(\frac{\lambda^2}{\gamma_0}\right)^{1/2} \frac{\frac{1}{2}(\lambda^2 [W(1)]^2 - \lambda^2 + \lambda^2 - \gamma_0)}{\lambda^2 \sqrt{\int_0^1 [W(r)]^2 dr}} \\
&= \left(\frac{\lambda^2}{\gamma_0}\right)^{1/2} \left[\frac{\frac{1}{2}\left([W(1)]^2 - 1\right)}{\sqrt{\int_0^1 [W(r)]^2 dr}} + \frac{\frac{1}{2}(\lambda^2 - \gamma_0)}{\lambda^2 \sqrt{\int_0^1 [W(r)]^2 dr}} \right]
\end{aligned}$$

Note now that the left term in the brackets is the limit distribution of t_T^* as discussed before.

4.5 (d)

Here we have

$$T(\hat{\rho}_T - 1) - \frac{1}{2} \frac{T^2 \hat{\sigma}^2_{\hat{\rho}_T}}{s_T^2} (\lambda^2 - \gamma_0)$$

---

[1] You can also argue that, since $\hat{\rho}_T$ is consistent for ρ, we may write $\hat{\rho}_T = \rho + o_p(1)$. Then, $s_T^2 = (T-1)^{-1} \times \sum_{t=1}^{T} (y_t - \hat{\rho}_T y_{t-1})^2 = (T-1)^{-1} \times \sum_{t=1}^{T} (y_t - \rho_T y_{t-1})^2 + o_p(1) = (T-1)^{-1} \sum_{t=1}^{T} u_t^2 + o_p(1) \xrightarrow{p} E(u^2) = \gamma_0$ by ergodicity of u_t.

where we can start with the second term:

$$\frac{1}{2}\frac{T^2\hat{\sigma}^2_{\hat{\rho}_T}}{s_T^2}(\lambda^2-\gamma_0) = \frac{1}{2}\frac{T^2\frac{s_T^2}{\sum_{t=1}^T y_{t-1}^2}}{s_T^2}(\lambda^2-\gamma_0) = \frac{(\lambda^2-\gamma_0)}{2}\frac{1}{T^{-2}\sum_{t=1}^T y_{t-1}^2}.$$

The first term is a constant, and the denominator in the second converges to

$$T^{-2}\sum_{t=1}^T y_{t-1}^2 \xrightarrow{L} \lambda^2\int_0^1 [W(1)]^2 dr$$

by Proposition 17.3(h). All together:

$$\begin{aligned}
T(\hat{\rho}_T-1) - \frac{1}{2}\frac{T^2\hat{\sigma}^2_{\hat{\rho}_T}}{s_T^2}(\lambda^2-\gamma_0) \xrightarrow{L} & \frac{\frac{1}{2}(\lambda^2[W(1)]^2-\gamma_0)}{\lambda^2\int_0^1[W(r)]^2dr} - \frac{(\lambda^2-\gamma_0)}{2}\frac{1}{\lambda^2\int_0^1[W(1)]^2dr} \\
&= \frac{\lambda^2[W(1)]^2-\gamma_0-\lambda^2+\gamma_0}{2\lambda^2\int_0^1[W(r)]^2dr} \\
&= \frac{\lambda^2\left([W(1)]^2-1\right)}{2\lambda^2\int_0^1[W(r)]^2dr} \\
&= \frac{\frac{1}{2}\left([W(1)]^2-1\right)}{\int_0^1[W(r)]^2dr}
\end{aligned}$$

which is what we also had in (2), and hence this adjusted test has a limiting non-standard distribution, but it has been tabulated in the appendix.

4.6 (e)

This time we have

$$\left(\frac{\gamma_0}{\lambda^2}\right)^{1/2} t_T - \frac{\frac{1}{2}(\lambda^2-\gamma_0)}{\lambda}\frac{T\hat{\sigma}_{\hat{\rho}_T}}{s_T}$$

where we get the first term from (c) and the second term is:

$$\begin{aligned}
\frac{\frac{1}{2}(\lambda^2-\gamma_0)}{\lambda}\frac{T\hat{\sigma}_{\hat{\rho}_T}}{s_T} \xrightarrow{L} & \frac{\frac{1}{2}(\lambda^2-\gamma_0)}{\lambda}\frac{\sqrt{\frac{\gamma_0}{\lambda^2\int_0^1[W(r)]^2dr}}}{\sqrt{\gamma_0}} \\
&= \frac{\frac{1}{2}(\lambda^2-\gamma_0)}{\lambda^2\int_0^1[W(r)]^2dr}
\end{aligned}$$

because of $s_T^2 \xrightarrow{p} \gamma_0$ as discussed before and the rest then follows from (the square root of) what we got for (b). So putting all together

$$\begin{aligned}
\left(\frac{\gamma_0}{\lambda^2}\right)^{1/2} t_T - \frac{\frac{1}{2}(\lambda^2-\gamma_0)}{\lambda}\frac{T\hat{\sigma}_{\hat{\rho}_T}}{s_T} \xrightarrow{L} & \left[\frac{\frac{1}{2}\left([W(1)]^2-1\right)}{\sqrt{\int_0^1 [W(r)]^2 dr}} + \frac{\frac{1}{2}(\lambda^2-\gamma_0)}{\lambda^2\sqrt{\int_0^1 [W(r)]^2 dr}}\right] \\
& - \frac{\frac{1}{2}(\lambda^2-\gamma_0)}{\lambda^2 \int_0^1 [W(r)]^2 dr} \\
& = \frac{\frac{1}{2}\left([W(1)]^2-1\right)}{\sqrt{\int_0^1 [W(r)]^2 dr}}
\end{aligned}$$

which is the limiting distribution in (3). Thus, for this one we can use Table B6 in the appendix.