Instructor: **Stanislav Anatolyev / Yaroslav Korobka** — Discussion Session: **1**
Course: **Econometrics of Macro and Financial Data** — Date: October 10, 2025

1 Stationarity and ergodicity

Definition 1 (Strict stationarity). Series $\{z_t\}_{t=-\infty}^{\infty}$ is said to be strictly stationary if joint distribution of collection $(z_t, z_{t-1}, \dots, z_{t-k})$ does not depend on t for $\forall\, k$.

Definition 2 (Weak stationarity). Series $\{z_t\}_{t=-\infty}^{\infty}$ is said to be weakly stationary if $\mathbb{E}[z_t]$, var $[z_t]$ and cov $[z_t, z_{t-k}]$ for $\forall\, k$ exist and do not depend on t.

Remark 1. Strict stationarity does not imply weak stationarity (e.g. Cauchy).

Definition 3 (Ergodicity). Series $\{z_t\}_{t=-\infty}^{\infty}$ is said to be ergodic if cov $[g(z_t), h(z_{t+k})] \to 0$ as $k \to \infty$ for $\forall\, g$ and h.

Theorem 2 (Invariance to transformations). *If $\{z_t\}_{t=-\infty}^{\infty}$ is stationary and ergodic, then so is $\{f(z_t, z_{t-1}, \dots)\}_{t=-\infty}^{\infty}$ for $\forall$ measurable function f.*

Example 3. We list some examples of the series:

- non-stationary: $y_t = x_t + \delta \cdot \mathbb{1}\{t \geq t_0\}$, $\mathbb{E}[y_t] = \mathbb{E}[x_t]$ for $t < t_0$ and $\mathbb{E}[y_t] = \mathbb{E}[x_t] + \delta$ for $t \geq t_0$
- non-ergodic: $x_t = Z$, where $Z \sim \mathcal{N}(0,1)$, cov $[x_t, x_{t+k}] =$ var $[Z] = 1 \nrightarrow 0$ as $k \to \infty$
- strong white noise (SWN): $\{\varepsilon_t\}_{t=-\infty}^{\infty}$ is i.i.d. series, $\mathbb{E}[\varepsilon_t] = 0$, $\sigma^2 =$ var $[\varepsilon_t]$
- weak white noise (WWN): $\{\varepsilon_t\}_{t=-\infty}^{\infty}$ is serially uncorrelated, $\mathbb{E}[\varepsilon_t] = 0$, var $[\varepsilon_t] = \sigma^2$, cov $[\varepsilon_t, \varepsilon_{t-j}] = 0$ for $\forall\, j \neq 0$

Example 4. Consider the Bernoulli process $a_t \in \{-1, +1\}$ with $\mathbb{P}\{a_t = +1\} = 1 - \mathbb{P}\{a_t = -1\} = \frac{1}{2}$, and let $\{\theta_t\}_{t=-\infty}^{+\infty}$ be the standard normal white noise independent of $\{a_t\}_{t=-\infty}^{+\infty}$. Show that the process

$$z_t = (a_t - a_{t-1})^2 + \theta_{t+1}^2$$

is strictly stationary and ergodic. Determine its mean and order of serial correlation (you need not derive the whole ACF).

Solution: The process z_t is strictly stationary and ergodic because it is a measurable function of a jointly strictly stationary and ergodic vector process $(a_t, \theta_t)'$. The mean of both a_t and θ_t is zero, a_t is serially independent, and $a_t^2 = 1$ with probability 1. Hence,

$$\begin{aligned} \mathbb{E}[z_t] &= \mathbb{E}[(a_t - a_{t-1})^2] + \mathbb{E}[\theta_{t+1}^2] \\ &= \mathbb{E}[a_t^2] + \mathbb{E}[a_{t-1}^2] - 2\mathbb{E}[(a_t a_{t-1})] + \text{var}\,[(\theta_{t+1})] \\ &= 3. \end{aligned}$$

Because z_t and z_{t+2} are independent, the order of serial correlation cannot exceed 1. The serial correlation in z_t may come only from the a-part. Let us check if it is not zero:

$$\begin{aligned} \text{cov}\,[(z_t, z_{t+1})] &= \text{cov}\left(a_t^2 + a_{t-1}^2 - 2a_t a_{t-1}, a_t^2 + a_{t+1}^2 - 2a_t a_{t+1}\right) \\ &= \text{var}\left(a_t^2\right) = 0. \end{aligned}$$

This, despite the one-period overlap, there is in fact no serial correlation in the process.

2 Lag operator

Definition 4 (Lag operator)**.** Lag operator L is defined as follows:

$$Lx_t = x_{t-1}, \quad LLx_t = x_{t-2}, \quad \ldots, \quad L^k x_t = x_{t-k}.$$

Definition 5 (Lag polynomial)**.** Lag polynomial $\Phi(L)$ of order k is defined as

$$\Phi(L) = 1 - \phi_1 L - \ldots - \phi_k L^k,$$

so when applied to x_t we have $\Phi(L)x_t = x_t - \phi_1 x_{t-1} - \ldots - \phi_k x_{t-k}$.

Theorem 5 (Fundamental theorem of algebra)**.** *$\Phi(L)$ of order k can be factorized as $\Phi(L) = \Pi_{i=1}^{k}(1 - \phi_i L)$.*

Example 6. Some examples follow:

- $\Phi(0) = 1$
- $\Phi(1) = 1 - \phi_1 - \ldots - \phi_k$
- $\Phi(L)\mu = \mu \cdot \Phi(1)$

3 Autocorrelation function (ACF)

Definition 6 (ACF)**.** We define the autocorrelation function as

$$\text{ACF}(j) = \frac{\text{cov}\,[x_t, x_{t+j}]}{\text{var}\,[x_t]}.$$

Remark 7. ACF makes sense only for stationary and ergodic series. Stationarity is used in the denominator, ergodicity in the numerator.

4 Standard linear processes

1. **autoregression of order 1, AR(1):**

$$x_t = \mu + \phi x_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim \text{WWN}.$$

   - $\phi = 1$: random walk (with drift $\mu \neq 0$, without drift $\mu = 0$) $\Rightarrow x_t = x_{t-1} + \varepsilon_t$ (non-stationarity, non-ergodic); can write as $x_t = x_0 + \varepsilon_1 + \ldots + \varepsilon_t \Rightarrow \text{var}\,[x_t] = \text{var}\,[x_0] + t\sigma^2 = \text{cov}\,[x_t, x_{t+k}]$ (check this); x_t is not measurable so it does not exist as a random variable, $x_t = \sum_{j=0}^{\infty} \varepsilon_{t-j}$.
   - $|\phi| < 1$ (necessary stationarity condition): moments are

$$m_x := \mathbb{E}[x_t] = \mathbb{E}[x_{t-1}]\phi + \mu \Rightarrow m_x = \frac{\mu}{1-\phi}$$

     for the mean,

$$\sigma_x^2 := \text{var}\,[x_t] = \text{var}\,[x_{t-1}]\phi^2 + \sigma^2 \Rightarrow \sigma_x^2 = \frac{\sigma^2}{1-\phi^2}$$

     for the variance, and

$$\gamma_x(1) := \text{cov}\,[x_t, x_{t+1}] = \text{cov}\,[x_t, \mu + \phi x_t + \varepsilon_{t+1}] = \phi\sigma_x^2$$

     covariances with $\gamma_x(j) = \text{cov}\,[x_t, x_{t+j}] = \phi^j \sigma_x^2$. ACF is then ϕ^j.

   We can also write AR(1) using the lag operator as

$$\Phi(L)x_t = \mu + \varepsilon_t, \quad \Phi(L) = 1 - \phi L.$$

   It follows that

$$\begin{aligned} x_t &= \Phi(L)^{-1}\mu + \Phi(L)^{-1}\varepsilon_t \\ &= \Phi(L)^{-1}\mu + \sum_{j=0}^{\infty} \phi^j \varepsilon_{t-j} \end{aligned}$$

   because $\Phi(L)^{-1} = \sum_{j=0}^{\infty} \phi^j L^j$ by Taylor.

2. **autoregression of order p, AR(p):**

$$x_t = \mu + \phi_1 x_{t-1} + \ldots + \phi_p x_{t-p} + \varepsilon_t, \quad \varepsilon_t \sim \text{WWN}.$$

Using lag operator we can write

$$\Phi(L)x_t = \mu + \varepsilon_t, \quad \Phi(L) = 1 - \phi_1 L - \ldots - \phi_p L^p$$

from which follows

$$x_t = \Phi(1)^{-1}\mu + \sum_{j=0}^{\infty} \phi^j \varepsilon_{t-j}.$$

Stationarity condition: roots of $\Phi(L)$ lie outside the unit circle. For example, for AR(1) we have $1 - \phi L = 0 \Rightarrow L = |\frac{1}{\phi}| \Rightarrow |\phi| < 1$.

3. **moving average MA(1):**

$$x_t = \varepsilon_t - \theta\varepsilon_{t-1}, \quad \varepsilon_t \sim \text{WWN}.$$

Always stationary and ergodic process, $\theta \in (-\infty, \infty)$. Moments are

$$\mathbb{E}[x_t] = 0, \quad \text{var}\,[x_t] = (1+\theta)^2\sigma^2, \quad \text{cov}\,[x_t, x_{t+1}] = -\theta\sigma^2,$$

and $\forall\, |k| > 1$ covariances are 0. If $|\theta| > 1 \Rightarrow$ non-invertible representation of MA(1). That is,

$$x_t = \varepsilon_t - \theta\varepsilon_{t-1} \Rightarrow \varepsilon_t = (1-\theta L)^{-1}x_t = \sum_{j=0}^{\infty} \theta^j x_{t-j} \text{ does not converge.}$$

Solution: find an invertible representation (see Hamilton (1994)).

4. **moving average MA(q):**

$$x_t = \Theta(L)\varepsilon_t, \quad \varepsilon_t \sim \text{WWN}, \quad \Theta(L) := 1 - \theta_1 L - \ldots - \theta_q L^q.$$

Always stationary. Invertible if roots of $\Theta(L)$ lie outside the unit circle. ε_t is called *innovation* if MA(q) is invertible.

5. **ARMA(p, q):**

$$\Phi(L)x_t = \mu + \Theta(L)\varepsilon_t, \quad \varepsilon_t \sim \text{WWN}.$$

Stationarity condition: roots of $\Phi(L)$ should be outside the unit circle. Invertibility condition: roots of $\Theta(L)$ should be outside the unit circle. Non-reducability condition: no common roots of $\Phi(L)$ and $\Theta(L)$.

Example 8. Sum of two independent MA(1) processes is MA(1), that is, $\text{MA}(1) + \text{MA}(1) = \text{MA}(1)$ (see Hamilton (1994) for the proof).

Example 9. Sum of two independent AR(1) processes

- with equal coefficients is AR(1). That is, after summing up

$$(1-\pi L)x_t = u_t$$
$$(1-\rho L)w_t = \eta_t$$

we have $(1-\pi L)(x_t + w_t) = u_t + \eta_t$, which is equivalent to $(1-\pi L)y_t = \varepsilon_t$, that is, AR(1) process.

- with different coefficients is ARMA(2, 1). That is, after summing up

$$(1-\pi L)(1-\rho L)x_t = u_t(1-\rho L)$$
$$(1-\pi L)(1-\rho L)w_t = \eta_t(1-\pi L)$$

we have $(1-\pi L)(1-\rho L)(x_t + w_t) = u_t(1-\rho L) + \eta_t(1-\pi L)$. We have two independent MA(1) processes on the right-hand side which is equal to MA(1) due to Example 8. Using the fact that $(1-\pi L)(1-\rho L) = (1 - \phi_1 L - \phi_2 L^2)$, we have

$$(1 - \phi_1 L - \phi_2 L^2)y_t = \varepsilon_t(1-\gamma L)$$

which is ARMA(2, 1) process.

Example 10. In general, AR(p) + AR(q) = ARMA($p+q$, $\max\{p, q\}$).

5 Wold decomposition

Suppose $\{x_t\}_{t=-\infty}^{\infty}$ is weakly stationary. Then it can be decomposed as

$$x_t = d_t + \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i},$$

where d_t is a deterministic part, $\psi_0 = 1$, $\sum_{i=0}^{\infty} \psi_i^2 < \infty$, ε_t is WWN. We call $\varepsilon_t = x_t - \text{Proj}\{x_t | x_{t-1}, \dots\}$ the Wold innovation; d_t is perfectly predictable from the past, $d_t = \text{Proj}\{d_t | d_{t-1}, \dots\}$.

Example 11. Some examples of the Wold decomposition:

- white noise: $\eta_t \Rightarrow d_t = 0$, $\psi_0 = 1$, $\psi_j = 0 \; \forall \, j \geq 1$,
- random variable: $x_t = Z$, $Z \sim \mathcal{N}(0, 1) \Rightarrow d_t = Z$, $\varepsilon_t = 0$,
- AR(1) process: $(1 - \phi L) x_t = \mu + \varepsilon_t$, $|\phi| < 1 \Rightarrow x_t = (1 - \phi L)^{-1}(\mu + \varepsilon_t) = (1 - \phi)^{-1}\mu + \sum_{j=0}^{\infty} \phi^j \varepsilon_{t-j}$; here, $d_t = (1 - \phi)^{-1}\mu$ and $\psi_j = \phi^j$, $j \geq 0$.