Stochastic processes: basic notions [*]

Jean-Marie Dufour [†]
McGill University

First version: March 2002
Revised: September 2002, April 2004, September 2004, January 2005, July 2011, May 2016, July 2016
This version: July 2016
Compiled: January 17, 2022, 17:18

[*]This work was supported by the William Dow Chair in Political Economy (McGill University), the Bank of Canada (Research Fellowship), the Toulouse School of Economics (Pierre-de-Fermat Chair of excellence), the Universitad Carlos III de Madrid (Banco Santander de Madrid Chair of excellence), a Guggenheim Fellowship, a Konrad-Adenauer Fellowship (Alexander-von-Humboldt Foundation, Germany), the Canadian Network of Centres of Excellence [program on *Mathematics of Information Technology and Complex Systems* (MITACS)], the Natural Sciences and Engineering Research Council of Canada, the Social Sciences and Humanities Research Council of Canada, and the Fonds de recherche sur la société et la culture (Québec).

[†] William Dow Professor of Economics, McGill University, Centre interuniversitaire de recherche en analyse des organisations (CIRANO), and Centre interuniversitaire de recherche en économie quantitative (CIREQ). Mailing address: Department of Economics, McGill University, Leacock Building, Room 414, 855 Sherbrooke Street West, Montréal, Québec H3A 2T7, Canada. TEL: (1) 514 398 6071; FAX: (1) 514 398 4800; e-mail: `jean-marie.dufour@mcgill.ca`. Web page: `http://www.jeanmariedufour.com`

Contents

List of Definitions, Assumptions, Propositions and Theorems

1. Fundamental concepts

1.1. Probability space

Definition 1.1 PROBABLITY SPACE. *A* probability space *is a triplet* $(\Omega, \mathscr{A}, \mathbb{P})$ *where*

(1) Ω *is the set of all possible results of an experiment*;

(2) $\mathscr{A}$ *is a class of subsets of* Ω *(called* events*) forming a* $\sigma-$*algebra,* i.e.

- (i) $\Omega \in \mathscr{A}$,
- (ii) $A \in \mathscr{A} \Rightarrow A^c \in \mathscr{A}$,
- (iii) $\bigcup_{j=1}^{\infty} A_j \in \mathscr{A}$ *, for any sequence* $\{A_1, A_2, ...\} \subseteq \mathscr{A}$;

(3) $\mathbb{P} : \mathscr{A} \to [0,1]$ *is a function which assigns to each event* $A \in \mathscr{A}$ *a number* $\mathbb{P}(A) \in [0,1]$, *called the probability of A and such that*

- (i) $\mathbb{P}(\Omega) = 1$,
- (ii) *if* $\{A_j\}_{j=1}^{\infty}$ *is a sequence of disjoint events, then* $\mathbb{P}(\bigcup_{j=1}^{\infty} A_j) = \sum_{j=1}^{\infty} \mathbb{P}(A_j)$.

1.2. Real random variable

Definition 1.2 REAL RANDOM VARIABLE (HEURISTIC DEFINITION). *A* real random variable *X is a variable with real values whose behavior can be described by a probability distribution. Usually, this probability distribution is described by a distribution function:*

$$F_X(x) = \mathbb{P}[X \leq x] \ . \tag{1.1}$$

Definition 1.3 REAL RANDOM VARIABLE. *A* real random variable *X is a function* $X : \Omega \to \mathbb{R}$ *such that*

$$X^{-1}\big((-\infty, x]\big) := \{\omega \in \Omega : X(\omega) \leq x\} \in \mathscr{A}, \forall x \in \mathbb{R} \, .$$

X is a measurable function*. The probability distribution of X is defined by*

$$F_X(x) = \mathbb{P}[X^{-1}((-\infty, x])] \, . \tag{1.2}$$

1.3. Stochastic process

Definition 1.4 REAL STOCHASTIC PROCESS. *Let* $\mathbb{T}$ *be a non-empty set. A stochastic process on* $\mathbb{T}$ *is a collection of random variables* $X_t : \Omega \to \mathbb{R}$ *such that a random variable* X_t *is associated with each each element* $t \in \mathbb{T}$. *This stochastic process is denoted by* $\{X_t : t \in \mathbb{T}\}$, *or more simply by* X_t *when the definition of* $\mathbb{T}$ *is clear. If* $\mathbb{T} = \mathbb{R}$ *(real numbers),* $\{X_t : t \in \mathbb{T}\}$ *is a continuous time process. If* $\mathbb{T} = \mathbb{Z}$ *(integers) or* $\mathbb{T} \subseteq \mathbb{Z}$, $X_t : t \in \mathbb{T}\}$ *is discrete time process.*

The set $\mathbb{T}$ can be finite or infinite, though usually it is more convenient to assume that $\mathbb{T}$ is infinite. In the sequel, we focus on processes where $\mathbb{T}$ is a right-infinite interval of integers, *i.e.*

$$\mathbb{T} = (n_0,\ \infty) \text{ where } n_0 \in \mathbb{Z} \text{ or } n_0 = -\infty. \tag{1.3}$$

We can also consider random variables which take their values in more general spaces, *i.e.*

$$X_t : \Omega \to \Omega_0 \tag{1.4}$$

where Ω_0 is any non-empty set. Unless stated otherwise, we shall limit ourselves to the case where $\Omega_0 = \mathbb{R}$.

To observe a time series is equivalent to observing a realization of a process $\{X_t : t \in \mathbb{T}\}$ or a portion of such a realization: given $(\Omega, \mathscr{A}, \mathbb{P})$, $\omega \in \Omega$ is drawn first, and then the variables $X_t(\omega)$, $t \in \mathbb{T}$, are associated with it. Each realization is determined in one shot by ω.

The probability law of a stochastic process $\{X_t : t \in \mathbb{T}\}$ with $\mathbb{T} \subseteq \mathbb{R}$ can be described by specifying the joint distribution function of $(X_{t_1}, \ldots, X_{t_n})$ for each subset $\{t_1,\ t_2, \ldots, t_n\} \subseteq \mathbb{T}$ (where $n \geq 1$):

$$F(x_1, \ldots, x_n;\ t_1, \ldots, t_n) = \mathbb{P}[X_{t_1} \leq x_1, \ldots, X_{t_n} \leq x_n]\ . \tag{1.5}$$

This follows from Kolmogorov's theorem [see Brockwell and Davis (1991, Chapter 1)].

1.4. L_r spaces

Definition 1.5 L_r SPACE. *Let r be a real number.* L_r *is the set of real random variables X defined on* $(\Omega, \mathscr{A}, \mathbb{P})$ *such that* $\mathbb{E}[|X|^r] < \infty$.

The space L_r is always defined with respect to a probability space $(\Omega, \mathscr{A}, \mathbb{P})$. In other words, we have:

$$L_r = L_r(\Omega,\ \mathscr{A},\ \mathbb{P})\,. \tag{1.6}$$

L_2 is the set of random variables on $(\Omega, \mathscr{A}, \mathbb{P})$ whose second moments are finite (*square-integrable variables*). A stochastic process $\{X_t : t \in \mathbb{T}\}$ is in L_r iff $X_t \in L_r$, $\forall t \in \mathbb{T}$, *i.e.*

$$\mathbb{E}[|X_t|^r] < \infty\ , \forall t \in \mathbb{T}\ . \tag{1.7}$$

The properties of moments of random variables are summarized in Dufour (2022).

2. Stationary processes

In general, the variables of a process $\{X_t : t \in \mathbb{T}\}$ are not identically distributed nor independent. In particular, if we suppose that $\mathbb{E}(X_t^2) < \infty$, we have:

$$\mathbb{E}(X_t) = \mu_t\ , \tag{2.1}$$

$$\text{Cov}(X_{t_1}, X_{t_2}) = \mathbb{E}[(X_{t_1} - \mu_{t_1})(X_{t_2} - \mu_{t_2})] = C(t_1, t_2)\ . \tag{2.2}$$

The means, variances and covariances of the variables of the process depend on their position in the series. The behavior of X_t can change with time. The function $C : \mathbb{T} \times \mathbb{T} \to \mathbb{R}$ is called the *covariance function* of the process $\{X_t : t \in \mathbb{T}\}$.

In this section, we will focus on the case where $\mathbb{T}$ is a right-infinite interval of integers.

Assumption 2.1 PROCESS ON AN INTERVAL OF INTEGERS.

$$\mathbb{T} = \{t \in \mathbb{Z} : t > n_0\} \quad \text{where } n_0 \in \mathbb{Z} \cup \{-\infty\}. \tag{2.3}$$

Definition 2.1 STRICTLY STATIONARY PROCESS. *A stochastic process $\{X_t : t \in \mathbb{T}\}$ is strictly stationary (SS) iff the probability distribution of the vector $(X_{t_1+k}, X_{t_2+k}, \ldots, X_{t_n+k})'$ is identical with the one of $(X_{t_1}, X_{t_2}, \ldots, X_{t_n})'$, for any finite subset $\{t_1,\, t_2,\, \ldots,\, t_n\} \subseteq \mathbb{T}$ and any integer $k \geq 0$. To indicate that $\{X_t : t \in \mathbb{T}\}$ is SS, we write $\{X_t : t \in \mathbb{T}\} \sim SS$ or $X_t \sim SS$.*

Proposition 2.1 CHARACTERIZATION OF STRICT STATIONARITY FOR A PROCESS ON (n_0, ∞). *If the process $\{X_t : t \in \mathbb{T}\}$ is SS, then the probability distribution of the vector $(X_{t_1+k}, X_{t_2+k}, \ldots, X_{t_n+k})'$ is identical to the one of $(X_{t_1}, X_{t_2}, \ldots, X_{t_n})'$, for any finite subset $\{t_1,\, t_2,\, \ldots,\, t_n\} \subseteq \mathbb{T}$ and any integer $k > n_0 - \min\{t_1, \ldots, t_n\}$.*

For processes on the integers ($\mathbb{T} = \mathbb{Z}$), the above characterization can be formulated in a simpler way as follows.

Proposition 2.2 CHARACTERIZATION OF STRICT STATIONARITY FOR A PROCESS ON THE INTEGERS. *A process $\{X_t : t \in \mathbb{Z}\}$ is SS iff the probability distribution of $(X_{t_1+k},\, X_{t_2+k},\, \ldots,\, X_{t_n+k})'$ is identical with the probability distribution of $(X_{t_1},\, X_{t_2},\, \ldots,\, X_{t_n})'$, for any subset $\{t_1,\, t_2,\, \ldots,\, t_n\} \subseteq \mathbb{Z}$ and any integer k.*

Definition 2.2 SECOND-ORDER STATIONARY PROCESS. *A stochastic process $\{X_t : t \in \mathbb{T}\}$ is* second-order stationary *(S2) iff*

$$\begin{array}{ll} (1) & \mathbb{E}(X_t^2) < \infty, \forall t \in \mathbb{T}, \\ (2) & \mathbb{E}(X_s) = \mathbb{E}(X_t), \forall s, t \in \mathbb{T}, \\ (3) & \mathrm{Cov}(X_s, X_t) = \mathrm{Cov}(X_{s+k}, X_{t+k}), \forall s, t \in \mathbb{T}, \forall k \geq 0\ . \end{array} \tag{2.4}$$

If $\{X_t : t \in \mathbb{T}\}$ is S2, we write $\{X_t : t \in \mathbb{T}\} \sim S2$ or $X_t \sim S2$.

Remark 2.1 Instead of *second-order stationary*, we also say *weakly stationary* (WS).

Proposition 2.3 RELATION BETWEEN STRICT AND SECOND-ORDER STATIONARITY. *If the process $\{X_t : t \in \mathbb{T}\}$ is strictly stationary and $\mathbb{E}(X_t^2) < \infty$ for any $t \in \mathbb{T}$, then the process $\{X_t : t \in \mathbb{T}\}$ is second-order stationary.*

PROOF. Suppose $\mathbb{E}(X_t^2) < \infty$, for any $t \in \mathbb{T}$. If the process $\{X_t : t \in \mathbb{T}\}$ is SS, we have:

$$\mathbb{E}(X_s) = \mathbb{E}(X_t)\ , \forall s, t \in \mathbb{T}\ , \tag{2.5}$$

$$\mathbb{E}(X_s X_t) = \mathbb{E}(X_{s+k} X_{t+k}) \ , \forall s, t \in \mathbb{T}, \forall k \geq 0 \ . \tag{2.6}$$

Since

$$\text{Cov}(X_s, X_t) = \mathbb{E}(X_s X_t) - \mathbb{E}(X_s)\mathbb{E}(X_t) \ , \tag{2.7}$$

we see that

$$\text{Cov}(X_s, X_t) = \text{Cov}(X_{s+k}, X_{t+k}) \ , \forall s, t \in \mathbb{T} \ , \forall k \geq 0 \ , \tag{2.8}$$

so the conditions (2.5) - (2.8) are equivalent to the conditions (2.5) - (2.6). The mean of X_t is constant, and the covariance between any two variables of the process only depends on the distance between the variables, not their position in the series. □

Proposition 2.4 EXISTENCE OF AN AUTOCOVARIANCE FUNCTION. *If the process $\{X_t : t \in \mathbb{T}\}$ is second-order stationary, then there exists a function $\gamma : \mathbb{Z} \to \mathbb{R}$ such that*

$$\text{Cov}(X_s, X_t) = \gamma(t-s) \ , \forall s, t \in \mathbb{T}. \tag{2.9}$$

The function γ is called the autocovariance function of the process $\{X_t : t \in \mathbb{T}\}$, and $\gamma_k := \gamma(k)$ the lag-k autocovariance of the process $\{X_t : t \in \mathbb{T}\}$.

PROOF. To show the existence of the autocovariance function in (2.9), we need to prove the following implication:

$$t_2 - s_2 = t_1 - s_1 \ \Rightarrow \ \text{Cov}(X_{s_2}, X_{t_2}) = \text{Cov}(X_{s_1}, X_{t_1}) \tag{2.10}$$

for all pairs (s_1, t_1) and (s_2, t_2). Suppose that $t_2 - s_2 = t_1 - s_1$. Then, using the stationarity assumption,

$$\begin{aligned}
\text{Cov}(X_{s_2}, X_{t_2}) &= \text{Cov}(X_{s_2+(s_1-s_2)}, X_{t_2+(s_1-s_2)}) \\
&= \text{Cov}(X_{s_1}, X_{s_1+(t_2-s_2)}) \\
&= \text{Cov}(X_{s_1}, X_{s_1+(t_1-s_1)}) \\
&= \text{Cov}(X_{s_1}, X_{t_1}) \ . \qquad (2.11)
\end{aligned}$$

□

Proposition 2.5 PROPERTIES OF THE AUTOCOVARIANCE FUNCTION. *Let $\{X_t : t \in \mathbb{T}\}$ be a second-order stationary process. The autocovariance function $\gamma(k)$ of the process $\{X_t : t \in \mathbb{T}\}$ satisfies the following properties:*

(1) $\gamma(0) = V(X_t) \geq 0 \ , \ \forall t \in \mathbb{T}$;

(2) $\gamma(k) = \gamma(-k) \ , \ \forall k \in \mathbb{Z}$ (i.e., $\gamma(k)$ *is an even function of* k);

(3) $|\gamma(k)| \leq \gamma(0)$, $\forall k \in \mathbb{Z}$;

(4) *the function* $\gamma(k)$ *is positive semidefinite,* i.e.

$$\sum_{i=1}^{N}\sum_{j=1}^{N} a_i a_j \gamma(t_i - t_j) \geq 0, \tag{2.12}$$

for any positive integer N *and for all the vectors* $a = (a_1, \ldots, a_N)' \in \mathbb{R}^N$ *and* $(t_1, \ldots, t_N)' \in \mathbb{T}^N$;

(5) *any* $N \times N$ *matrix of the form*

$$\begin{aligned} \Gamma_N &= [\gamma(j-i)]_{i,\, j=1,\ldots,N} \\ &= \begin{bmatrix} \gamma(0) & \gamma(1) & \gamma(2) & \cdots & \gamma(N-1) \\ \gamma(1) & \gamma(0) & \gamma(1) & \cdots & \gamma(N-2) \\ \vdots & \vdots & \vdots & & \vdots \\ \gamma(N-1) & \gamma(N-2) & \gamma(N-3) & \cdots & \gamma(0) \end{bmatrix} \end{aligned} \tag{2.13}$$

is positive semidefinite.

Proposition 2.6 EXISTENCE OF AN AUTOCORRELATION FUNCTION. *If the process* $\{X_t : t \in \mathbb{T}\}$ *is second-order stationary, then there exists a function* $\rho : \mathbb{Z} \to [-1, 1]$ *such that*

$$\rho(t-s) = \mathrm{Corr}(X_s, X_t) = \gamma(t-s)/\gamma(0) \text{ , } \forall s, t \in \mathbb{T}, \tag{2.14}$$

where $0/0 := 1$. *The function* ρ *is called the* autocorrelation function *of the process* $\{X_t : t \in \mathbb{T}\}$, *and* $\rho_k := \rho(k)$ *the* lag-k autocorrelation *of the process* $\{X_t : t \in \mathbb{T}\}$.

Proposition 2.7 PROPERTIES OF THE AUTOCORRELATION FUNCTION. *Let* $\{X_t : t \in \mathbb{T}\}$ *be a second-order stationary process. The autocorrelation function* $\rho(k)$ *of the process* $\{X_t : t \in \mathbb{T}\}$ *satisfies the following properties:*

(1) $\rho(0) = 1$;

(2) $\rho(k) = \rho(-k)$, $\forall k \in \mathbb{Z}$;

(3) $|\rho(k)| \leq 1$, $\forall k \in \mathbb{Z}$;

(4) *the function* $\rho(k)$ *is positive semidefinite,* i.e.

$$\sum_{i=1}^{N}\sum_{j=1}^{N} a_i a_j \rho(t_i - t_j) \geq 0 \tag{2.15}$$

for any positive integer N *and for all the vectors* $a = (a_1, \ldots, a_N)' \in \mathbb{R}^N$ *and* $(t_1, \ldots, t_N)' \in \mathbb{T}^N$;

(5) *any $N \times N$ matrix of the form*

$$R_N = \frac{1}{\gamma_0}\Gamma_N = \begin{bmatrix} 1 & \rho(1) & \rho(2) & \cdots & \rho(N-1) \\ \rho(1) & 1 & \rho(1) & \cdots & \rho(N-2) \\ \vdots & \vdots & \vdots & & \vdots \\ \rho(N-1) & \rho(N-2) & \rho(N-3) & \cdots & 1 \end{bmatrix} \tag{2.16}$$

is positive semidefinite, where $\gamma(0) = V(X_t)$.

Theorem 2.8 CHARACTERIZATION OF AUTOCOVARIANCE FUNCTIONS. *An even function $\gamma : \mathbb{Z} \to \mathbb{R}$ is positive semidefinite iff $\gamma(\cdot)$ is the autocovariance function of a second-order stationary process $\{X_t : t \in \mathbb{Z}\}$.*

PROOF. See Brockwell and Davis (1991, Chapter 2). □

Corollary 2.9 CHARACTERIZATION OF AUTOCORRELATION FUNCTIONS. *An even function $\rho : \mathbb{Z} \to [-1, 1]$ is positive semidefinite iff ρ is the autocorrelation function of a second-order stationary process $\{X_t : t \in \mathbb{Z}\}$.*

Definition 2.3 DETERMINISTIC PROCESS. *Let $\{X_t : t \in \mathbb{T}\}$ be a stochastic process, $\mathbb{T}_1 \subseteq \mathbb{T}$ and $I_t = \{X_s : s \leq t\}$. We say that the process $\{X_t : t \in \mathbb{T}\}$ is* deterministic *on $\mathbb{T}_1$ iff there exists a collection of functions $\{g_t(I_{t-1}) : t \in \mathbb{T}_1\}$ such that $X_t = g_t(I_{t-1})$ with probability one, $\forall t \in \mathbb{T}_1$.*

A deterministic process can be perfectly predicted form its own past (at points where it is deterministic).

Proposition 2.10 CRITERION FOR A DETERMINISTIC PROCESS. *Let $\{X_t : t \in \mathbb{T}\}$ be a second-order stationary process, where $\mathbb{T} = \{t \in \mathbb{Z} : t > n_0\}$ and $n_0 \in \mathbb{Z} \cup \{-\infty\}$, and let $\gamma(k)$ its autocovariance function. If there exists an integer $N \geq 1$ such that the matrix Γ_N is singular [where Γ_N is defined in Proposition 2.5], then the process $\{X_t : t \in \mathbb{T}\}$ is deterministic for $t > n_0 + N - 1$. In particular, if $V(X_t) = \gamma(0) = 0$, the process is deterministic for $t \in \mathbb{T}$.*

For a second-order indeterministic stationary process at any $t \in \mathbb{T}$, all the matrices $\Gamma_N, N \geq 1$, are invertible.

Definition 2.4 STATIONARITY OF ORDER m. *Let m be a non-negative integer. A stochastic process $\{X_t : t \in \mathbb{T}\}$ is stationary of order m iff*

(1) $\mathbb{E}(|X_t|^m) < \infty, \forall t \in \mathbb{T}$,
and

(2) $\mathbb{E}[X_{t_1}^{m_1} X_{t_2}^{m_2} \cdots X_{t_n}^{m_n}] = \mathbb{E}[X_{t_1+k}^{m_1} X_{t_2+k}^{m_2} \cdots X_{t_n+k}^{m_n}]$
for any $k \geq 0$, and for any nonempty subset $\{t_1, \ldots, t_n\} \in \mathbb{T}^N$ and all the non-negative integers $m_1, \ldots, m_n$ such that $m_1 + m_2 + \cdots + m_n \leq m$.

If $m = 1$, the mean is constant, but not necessarily the other moments. If $m = 2$, the process is second-order stationary.

Definition 2.5 ASYMPTOTIC STATIONARITY OF ORDER *m*. *Let m a non-negative integer. A stochastic process* $\{X_t : t \in \mathbb{T}\}$ *is asymptotically stationary of order m iff*

(1) *there exists an integer N such that* $(|X_t|^m) < \infty$, *for* $t \geq N$,
and

(2) $\lim_{t_1 \to \infty} \left[\mathbb{E}\left(X_{t_1}^{m_1} X_{t_1+\Delta_2}^{m_2} \cdots X_{t_1+\Delta_n}^{m_n} \right) - \mathbb{E}\left(X_{t_1+k}^{m_1} X_{t_1+\Delta_2+k}^{m_2} \cdots X_{t_1+\Delta_n+k}^{m_n} \right) \right] = 0$
for any $k \geq 0,\ t_1 \in \mathbb{T}$, *all the positive integers* $\Delta_2, \Delta_3, \ldots, \Delta_n$ *such that* $\Delta_2 < \Delta_3 < \cdots < \Delta_n$, *and all non-negative integers* $m_1, \ldots, m_n$ *such that* $m_1 + m_2 + \cdots + m_n \leq m$.

3. Some important models

In this section, we will again assume that $\mathbb{T}$ is a right-infinite interval of integers (Assumption 2.1):

$$\mathbb{T} = \{t \in \mathbb{Z} : t > n_0\} \ \text{ where } n_0 \in \mathbb{Z} \cup \{-\infty\}. \tag{3.1}$$

3.1. Noise models

Definition 3.1 SEQUENCE OF INDEPENDENT RANDOM VARIABLES. *A process* $\{X_t : t \in \mathbb{T}\}$ *is a* sequence of independent *random variables iff the variables* X_t *are mutually independent. This is denoted by:*

$$\{X_t : t \in \mathbb{T}\} \sim IND \text{ or } \{X_t\} \sim IND. \tag{3.2}$$

Further, we write:

$$\{X_t : t \in \mathbb{T}\} \sim IND(\mu_t) \quad \text{if } \mathbb{E}(X_t) = \mu_t\,, \tag{3.3}$$

$$\{X_t : t \in \mathbb{T}\} \sim IND(\mu_t, \sigma_t^2) \quad \text{if } \mathbb{E}(X_t) = \mu_t \text{ and } \mathrm{V}(X_t) = \sigma_t^2\,.$$

Definition 3.2 RANDOM SAMPLE. *A* random sample *is a sequence of independent and identically distributed (*i.i.d.*) random variables. This is denoted by*

$$\{X_t : t \in \mathbb{T}\} \sim IID\,. \tag{3.4}$$

A random sample is a SS process. If $\mathbb{E}(X_t^2) < \infty$, for all $t \in \mathbb{T}$, the process is S2. In this case, we write:

$$\{X_t : t \in \mathbb{T}\} \sim IID(\mu, \sigma^2)\,, \quad \text{if } \mathbb{E}(X_t) = \mu \text{ and } \mathrm{V}(X_t) = \sigma^2. \tag{3.5}$$

Definition 3.3 WHITE NOISE. *A* white noise *is a sequence of random variables in* L_2 *with mean zero, the same variance and mutually uncorrelated,* i.e.

$$\mathbb{E}(X_t^2) < \infty, \forall t \in \mathbb{T}, \tag{3.6}$$

$$\mathbb{E}(X_t^2) = \sigma^2 \ , \forall t \in \mathbb{T} \, , \tag{3.7}$$

$$\text{Cov}(X_s, X_t) = 0 \, , \ \textit{if } s \neq t \, . \tag{3.8}$$

This is denoted by:

$$\{X_t : t \in \mathbb{T}\} \sim \text{WN}(0, \sigma^2) \textit{ or } \{X_t\} \sim \text{WN}(0, \sigma^2) \, . \tag{3.9}$$

Definition 3.4 HETEROSKEDASTIC WHITE NOISE. *A* heteroskedastic white noise *is a sequence of random variables in L_2 with mean zero and mutually uncorrelated,* i.e.

$$\mathbb{E}(X_t^2) < \infty, \forall t \in \mathbb{T} \, , \tag{3.10}$$

$$\mathbb{E}(X_t) = 0, \forall t \in \mathbb{T} \, , \tag{3.11}$$

$$\text{Cov}(X_t, X_s) = 0, \ \textit{if } s \neq t \, , \tag{3.12}$$

$$\mathbb{E}(X_t^2) = \sigma_t^2 \, , \ \forall t \in \mathbb{T} \, . \tag{3.13}$$

This is denoted by:

$$\{X_t : t \in \mathbb{Z}\} \sim \text{WN}(0, \sigma_t^2) \textit{ or } \{X_t\} \sim \text{WN}(0, \sigma_t^2) \, . \tag{3.14}$$

Each one of these four models will be called a *noise* process.

3.2. Harmonic processes

Many time series exhibit apparent periodic behavior. This suggests one to use periodic functions to describe them.

Definition 3.5 PERIODIC FUNCTION. *A function $f(t), t \in \mathbb{R}$, is periodic of period P on $\mathbb{R}$ iff*

$$f(t+P) = f(t) \, , \forall t, \tag{3.15}$$

and P is the lowest number such that (3.15) holds for all t. $\frac{1}{P}$ is the frequency associated with the function (number of cycles per unit of time).

Example 3.1 Sinus function:

$$\sin(t) = \sin(t+2\pi) = \sin(t+2\pi k), \forall k \in \mathbb{Z}. \tag{3.16}$$

For the sinus function, the period is $P = 2\pi$ and the frequency is $f = 1/(2\pi)$.

Example 3.2 Cosine function:

$$\cos(t) = \cos(t+2\pi) = \cos(t+2\pi k), \forall k \in \mathbb{Z}. \tag{3.17}$$

Example 3.3

$$\sin(\nu t) = \sin\left[\nu\left(t + \frac{2\pi}{\nu}\right)\right] = \sin\left[\nu\left(t + \frac{2\pi k}{\nu}\right)\right] \, , \forall k \in \mathbb{Z}. \tag{3.18}$$

Example 3.4

$$\cos(\nu t) = \cos\left[\nu\left(t + \frac{2\pi}{\nu}\right)\right] = \cos\left[\nu\left(t + \frac{2\pi k}{\nu}\right)\right] \;, \forall k \in \mathbb{Z}. \tag{3.19}$$

For $\sin(\nu t)$ and $\cos(\nu t)$, the period is $P = 2\pi/\nu$.

Example 3.5 GENERAL COSINE FUNCTION.

$$\begin{aligned} f(t) &= C\ \cos(\nu t+\theta) = C[\cos(\nu t)\cos(\theta) - \sin(\nu t)\sin(\theta)] \\ &= A\ \cos(\nu t) + B\ \sin(\nu t) \end{aligned} \tag{3.20}$$

where $C \geq 0$, $A = C\ \cos(\theta)$ and $B = -C\ \sin\ \theta$. Further,

$$C = \sqrt{A^2+B^2}\,, \quad \tan(\theta) = -B/A \text{ (if } C \neq 0). \tag{3.21}$$

In the above function, the different parameters have the following names:

$$\begin{aligned} C &= \text{amplitude}\,; \\ \nu &= \text{angular frequency (radians/time unit)}\,; \\ P &= 2\pi/\nu = \text{period}\,; \\ \bar{\nu} &= \frac{1}{P} = \frac{\nu}{2\pi} = \text{ frequency (number of cycles per time unit)}\,; \\ \theta &= \text{phase angle (usually } 0 \leq \theta < 2\pi \text{ or } -\pi/2 < \theta \leq \pi/2)\,. \end{aligned}$$

Example 3.6

$$\begin{aligned} f(t) &= C\ \sin(\nu t+\theta) = C\ \cos(\nu t+\theta-\pi/2) \\ &= C\,[\sin(\nu t)\cos(\theta) + \cos(\nu t)\sin(\theta)] \\ &= A\ \cos(\nu t) + B\ \sin(\nu t) \end{aligned} \tag{3.22}$$

where

$$0 \leq \nu < 2\pi\,, \tag{3.23}$$

$$A = C\ \sin(\theta) = C\ \cos\left(\theta - \frac{\pi}{2}\right)\,, \tag{3.24}$$

$$B = C\ \cos(\theta) = -C\ \sin\left(\theta - \frac{\pi}{2}\right)\,. \tag{3.25}$$

Consider the model

$$\begin{aligned} X_t &= C\ \cos(\nu t+\theta) \\ &= A\ \cos(\nu t) + B\ \sin(\nu t),\, t \in \mathbb{Z}. \end{aligned} \tag{3.26}$$

If A and B are constants,

$$\mathbb{E}(X_t) = A\ \cos(\nu t) + B\ \sin(\nu t)\,,\ t \in \mathbb{Z}, \tag{3.27}$$

so the process X_t is non-stationary (since the mean is not constant). Suppose now that A and B are random variables such that

$$\mathbb{E}(A) = \mathbb{E}(B) = 0, \quad \mathbb{E}(A^2) = \mathbb{E}(B^2) = \sigma^2, \quad \mathbb{E}(AB) = 0\,. \tag{3.28}$$

A and B do not depend on t but are fixed for each realization of the process $[A = A(\omega), B = B(\omega)]$. In this case,

$$\mathbb{E}(X_t) = 0\,, \tag{3.29}$$

$$\begin{aligned}
\mathbb{E}(X_s X_t) &= \mathbb{E}(A^2)\cos(\nu s)\cos(\nu t) + \mathbb{E}(B^2)\sin(\nu s)\sin(\nu t) \\
&= \sigma^2[\cos(\nu s)\cos(\nu t) + \sin(\nu s)\sin(\nu t)] \\
&= \sigma^2\cos[\nu(t-s)]\,.
\end{aligned} \tag{3.30}$$

The process X_t is stationary of order 2 with the following autocovariance and autocorrelation functions:

$$\gamma_X(k) = \sigma^2\cos(\nu k)\,, \tag{3.31}$$

$$\rho_X(k) = \cos(\nu k)\,. \tag{3.32}$$

If we add m cyclic processes of the form (3.26), we obtain a *harmonic process* of order m.

Definition 3.6 HARMONIC PROCESS OF ORDER *m*. *We say the process* $\{X_t : t \in \mathbb{T}\}$ *is a harmonic process of order m if it can written in the form*

$$X_t = \sum_{j=1}^{m}[A_j\cos(\nu_j t) + B_j\sin(\nu_j t)],\ \forall t \in \mathbb{T}\,, \tag{3.33}$$

where $\nu_1, \ldots, \nu_m$ *are distinct constants in the interval* $[0, 2\pi)$.

If A_j, B_j , $j = 1, \ldots, m$, are random variables in L_2 such that

$$\mathbb{E}(A_j) = \mathbb{E}(B_j) = 0\,,\ j = 1, \ldots, m\,, \tag{3.34}$$

$$\mathbb{E}(A_j^2) = \mathbb{E}(B_j^2) = \sigma_j^2\,,\ j = 1, \ldots, m\,, \tag{3.35}$$

$$\mathbb{E}(A_j A_k) = \mathbb{E}(B_j B_k) = 0, \text{ for } j \neq k, \tag{3.36}$$

$$\mathbb{E}(A_j B_k) = 0, \forall j, k\,, \tag{3.37}$$

the harmonic process X_t is second-order stationary, with:

$$\mathbb{E}(X_t) = 0\,, \tag{3.38}$$

$$\mathbb{E}(X_s X_t) = \sum_{j=1}^{m} \sigma_j^2 \cos[\nu_j(t-s)] \,, \tag{3.39}$$

hence

$$\gamma_X(k) = \sum_{j=1}^{m} \sigma_j^2 \cos(\nu_j k) \,, \tag{3.40}$$

$$\rho_X(k) = \sum_{j=1}^{m} \sigma_j^2 \cos(\nu_j k) / \sum_{j=1}^{m} \sigma_j^2 \,. \tag{3.41}$$

If we add a white noise u_t to X_t in (3.33), we obtain again a second-order stationary process:

$$X_t = \sum_{j=1}^{m} [A_j \cos(\nu_j t) + B_j \sin(\nu_j t)] + u_t, \, t \in \mathbb{T} \,, \tag{3.42}$$

where the process $\{u_t : t \in \mathbb{T}\} \sim$ WN$(0, \sigma^2)$ is uncorrelated with A_j, B_j , $j = 1, \ldots, m$. In this case, $\mathbb{E}(X_t) = 0$ and

$$\gamma_X(k) = \sum_{j=1}^{m} \sigma_j^2 \cos(\nu_j k) + \sigma^2 \delta(k) \tag{3.43}$$

where

$$\begin{aligned} \delta(k) &= 1 \quad \text{if } k = 0 \\ &= 0 \quad \text{otherwise.} \end{aligned} \tag{3.44}$$

If a series can be described by an equation of the form (3.42), we can view it as a realization of a second-order stationary process.

3.3. Linear processes

Many stochastic processes with dependence are obtained through transformations of noise processes.

Definition 3.7 AUTOREGRESSIVE PROCESS. *The process $\{X_t : t \in \mathbb{T}\}$ is an autoregressive process of order p if it satisfies an equation of the form*

$$X_t = \bar{\mu} + \sum_{j=1}^{p} \varphi_j X_{t-j} + u_t \,, \; \forall t \in \mathbb{T} \,, \tag{3.45}$$

where $\{u_t : t \in \mathbb{Z}\} \sim$ WN$(0, \sigma^2)$. In this case, we denote

$$\{X_t : t \in \mathbb{T}\} \sim \text{AR}(p) \,. \tag{3.46}$$

Usually, $\mathbb{T} = \mathbb{Z}$ *or* $\mathbb{T} = \mathbb{Z}_+$ *(positive integers). If* $\sum_{j=1}^{p} \varphi_j \neq 1$*, we can define*

$$\mu := \bar{\mu}/(1 - \sum_{j=1}^{p} \varphi_j) \tag{3.47}$$

and write

$$\tilde{X}_t = \sum_{j=1}^{p} \varphi_j \tilde{X}_{t-j} + u_t\,, \; \forall t \in \mathbb{T}\,, \tag{3.48}$$

where $\tilde{X}_t := X_t - \mu$.

Definition 3.8 MOVING AVERAGE PROCESS. *The process* $\{X_t : t \in \mathbb{T}\}$ *is a moving average process of order q if it can written in the form*

$$X_t = \bar{\mu} + \sum_{j=0}^{q} \psi_j u_{t-j}\,, \forall t \in \mathbb{T}, \tag{3.49}$$

where $\{u_t : t \in \mathbb{Z}\} \sim \mathrm{WN}(0,\, \sigma^2)$*. In this case, we denote*

$$\{X_t : t \in \mathbb{T}\} \sim \mathrm{MA}(q). \tag{3.50}$$

Without loss of generality, we can set $\psi_0 = 1$ and $\psi_j = -\theta_j,\ j = 1, \ldots, q$:

$$X_t = \bar{\mu} + u_t - \sum_{j=1}^{q} \theta_j u_{t-j}\ , t \in \mathbb{T} \tag{3.51}$$

or, equivalently,

$$\tilde{X}_t = u_t - \sum_{j=1}^{q} \theta_j u_{t-j} \tag{3.52}$$

where $\tilde{X}_t := X_t - \bar{\mu}$.

Definition 3.9 AUTOREGRESSIVE-MOVING-AVERAGE PROCESS. *The process* $\{X_t : t \in \mathbb{T}\}$ *is an autoregressive-moving-average (ARMA) process of order* (p, q) *if it can be written in the form*

$$X_t = \bar{\mu} + \sum_{j=1}^{p} \varphi_j X_{t-j} + u_t - \sum_{j=1}^{q} \theta_j u_{t-j},\ \forall t \in \mathbb{T}\,, \tag{3.53}$$

where $\{u_t : t \in \mathbb{Z}\} \sim \mathrm{WN}(0, \sigma^2)$*. In this case, we denote*

$$\{X_t : t \in \mathbb{T}\} \sim \mathrm{ARMA}(p, q)\,. \tag{3.54}$$

If $\sum_{j=1}^{p} \varphi_j \neq 1$, we can also write

$$\tilde{X}_t = \sum_{j=1}^{p} \varphi_j \tilde{X}_{t-j} + u_t - \sum_{j=1}^{q} \theta_j u_{t-j} \tag{3.55}$$

where $\tilde{X}_t = X_t - \mu$ and $\mu = \bar{\mu}/(1 - \sum_{j=1}^{p} \varphi_j)$.

Definition 3.10 MOVING AVERAGE PROCESS OF INFINITE ORDER. *The process $\{X_t : t \in \mathbb{T}\}$ is a moving-average process of infinite order if it can be written in the form*

$$X_t = \bar{\mu} + \sum_{j=-\infty}^{+\infty} \psi_j u_{t-j}\,, \; \forall t \in \mathbb{Z}\,, \tag{3.56}$$

where $\{u_t : t \in \mathbb{Z}\} \sim \mathrm{WN}(0,\, \sigma^2)$. We also say that X_t is a weakly linear process. *In this case, we denote*

$$\{X_t : t \in \mathbb{T}\} \sim \mathrm{MA}(\infty). \tag{3.57}$$

In particular, if $\psi_j = 0$ for $j < 0$, i.e.

$$X_t = \bar{\mu} + \sum_{j=0}^{\infty} \psi_j u_{t-j}\,, \; \forall t \in \mathbb{Z}\,, \tag{3.58}$$

we say that X_t is a causal function of u_t (causal linear process).

Definition 3.11 AUTOREGRESSIVE PROCESS OF INFINITE ORDER. *The process $\{X_t : t \in \mathbb{T}\}$ is an autoregressive process of infinite order if it can be written in the form*

$$X_t = \bar{\mu} + \sum_{j=1}^{\infty} \varphi_j X_{t-j} + u_t\,, \; t \in \mathbb{T}\,, \tag{3.59}$$

where $\{u_t : t \in \mathbb{Z}\} \sim \mathrm{WN}(0,\, \sigma^2)$. In this case, we denote

$$\{X_t : t \in \mathbb{T}\} \sim \mathrm{AR}(\infty). \tag{3.60}$$

Definition 3.12 Remark 3.1 We can generalize the notions defined above by assuming that $\{u_t : t \in \mathbb{Z}\}$ is a noise. Unless stated otherwise, we will suppose $\{u_t\}$ is a white noise.

QUESTIONS :

(1) Under which conditions are the processes defined above stationary (strictly or in L_r)?

(2) Under which conditions are the processes $\mathrm{MA}(\infty)$ or $\mathrm{AR}(\infty)$ well defined (convergent series)?

(3) What are the links between the different classes of processes defined above?

(4) When a process is stationary, what are its autocovariance and autocorrelation functions?

3.4. Integrated processes

Definition 3.13 RANDOM WALK. *The process $\{X_t : t \in \mathbb{T}\}$ is a random walk if it satisfies an equation of the form*

$$X_t - X_{t-1} = v_t,\ \forall t \in \mathbb{T}, \tag{3.61}$$

where $\{v_t : t \in \mathbb{T}\} \sim IID$. To ensure that this process is well defined, we suppose that $n_0 \neq -\infty$. If $n_0 = -1$, we can write

$$X_t = X_0 + \sum_{j=1}^{t} v_j \tag{3.62}$$

hence the name "integrated process". If $\mathbb{E}(v_t) = \bar{\mu}$ or $\mathrm{Med}(v_t) = \bar{\mu}$, one often writes

$$X_t - X_{t-1} = \bar{\mu} + u_t \tag{3.63}$$

where $u_t := v_t - \bar{\mu} \sim IID$ and $\mathbb{E}(u_t) = 0$ or $\mathrm{Med}(u_t) = 0$ (depending on whether $\mathbb{E}(u_t) = 0$ or $\mathrm{Med}(u_t) = 0$). If $\bar{\mu} \neq 0$, we say the the random walk has a drift.

Definition 3.14 WEAK RANDOM WALK. *The process $\{X_t : t \in \mathbb{T}\}$ is a* weak random walk *if X_t satisfies an equation of the form*

$$X_t - X_{t-1} = \bar{\mu} + u_t \tag{3.64}$$

where $\{u_t : t \in \mathbb{T}\} \sim \mathrm{WN}(0,\ \sigma^2)$, $\{u_t : t \in \mathbb{T}\} \sim \mathrm{WN}(0,\ \sigma_t^2)$, or $\{u_t : t \in \mathbb{T}\} \sim IND(0)]$.

Definition 3.15 INTEGRATED PROCESS. *The process $\{X_t : t \in \mathbb{T}\}$ is integrated of order d if it can be written in the form*

$$(1-B)^d X_t = Z_t\ ,\ \forall t \in \mathbb{T}, \tag{3.65}$$

where $\{Z_t : t \in \mathbb{T}\}$ is a stationary process (usually stationary of order 2) and d is a non-negative integer ($d = 0, 1, 2, \ldots$). In particular, if $\{Z_t : t \in \mathbb{T}\}$ is an $\mathrm{AR}MA(p, q)$ stationary process, $\{X_t : t \in \mathbb{T}\}$ is an $\mathrm{ARIMA}(p, d, q)$ process: $\{X_t : t \in \mathbb{T}\} \sim \mathrm{ARIMA}(p, d, q)$. We note

$$\begin{aligned}
B\,X_t &= X_{t-1}\ , && (3.66)\\
(1-B)X_t &= X_t - X_{t-1}\ , && (3.67)\\
(1-B)^2 X_t &= (1-B)(1-B)X_t = (1-B)(X_t - X_{t-1}) && (3.68)\\
&= X_t - 2X_{t-1} + X_{t-2}, && (3.69)\\
(1-B)^d X_t &= (1-B)(1-B)^{d-1} X_t,\ d = 1, 2, \ldots && (3.70)
\end{aligned}$$

where $(1-B)^0 = 1$.

3.5. Deterministic trends

Definition 3.16 DETERMINISTIC TREND. *The process $\{X_t : t \in \mathbb{T}\}$ follows a deterministic trend if it can be written in the form*

$$X_t = f(t) + Z_t \; , \forall t \in \mathbb{T}, \tag{3.71}$$

where $f(t)$ is a deterministic function of time and $\{Z_t : t \in \mathbb{T}\}$ is a noise or a stationary process.

Example 3.7 Important cases of deterministic trend:

$$X_t = \beta_0 + \beta_1 t + u_t, \tag{3.72}$$

$$X_t = \sum_{j=0}^{k} \beta_j t^j + u_t, \tag{3.73}$$

where $\{u_t : t \in \mathbb{T}\} \sim \mathrm{WN}(0, \sigma^2)$.

4. Transformations of stationary processes

Theorem 4.1 ABSOLUTE MOMENT SUMMABILITY CRITERION FOR CONVERGENCE OF A LINEAR TRANSFORMATION OF A STOCHASTIC PROCESS. *Let $\{X_t : t \in \mathbb{Z}\}$ be a stochastic process on the integers, $r \geq 1$ and $\{a_j : j \in \mathbb{Z}\}$ a sequence of real numbers. If*

$$\sum_{j=-\infty}^{\infty} |a_j| \, \mathbb{E}(|X_{t-j}|^r)^{1/r} < \infty \tag{4.1}$$

then, for any t, the random series $\sum_{j=-\infty}^{\infty} a_j X_{t-j}$ converges absolutely a.s. and in mean of order r to a random variable Y_t such that $\mathbb{E}(|Y_t|^r) < \infty$.

PROOF. See Dufour (2016). □

Theorem 4.2 ABSOLUTE SUMMABILITY CRITERION FOR CONVERGENCE OF A LINEAR TRANSFORMATION OF A WEAKLY STATIONARY PROCESS. *Let $\{X_t : t \in \mathbb{Z}\}$ be a second-order stationary process and $\{a_j : j \in \mathbb{Z}\}$ an sequence of real numbers absolutely convergent sequence of real numbers,* i.e.

$$\sum_{j=-\infty}^{\infty} |a_j| < \infty . \tag{4.2}$$

Then the random series $\sum_{j=-\infty}^{\infty} a_j X_{t-j}$ converges absolutely a.s. and in mean of order 2 to a random variable $Y_t \in L_2$, $\forall t$, and the process $\{Y_t : t \in \mathbb{Z}\}$ is second-order stationary with autocovariance *function*

$$\gamma_Y(k) = \sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} a_i a_j \gamma_X(k-i+j) \; . \tag{4.3}$$

PROOF. See Gouriéroux and Monfort (1997, Property 5.6). □

Theorem 4.3 NECESSARY AND SUFFICIENT CONDITION FOR CONVERGENCE OF LINEAR FILTERS OF ARBITRARY WEAKLY STATIONARY PROCESSES. *The series* $\sum_{j=-\infty}^{\infty} a_j X_{t-j}$ *converges absolutely a.s. for any second-order stationary process* $\{X_t : t \in \mathbb{Z}\}$ *iff*

$$\sum_{j=-\infty}^{\infty} |a_j| < \infty . \tag{4.4}$$

5. Infinite order moving averages

We study here random series of the form

$$\sum_{j=o}^{\infty} \psi_j u_{t-j},\, t \in \mathbb{Z} \tag{5.1}$$

and

$$\sum_{j=-\infty}^{\infty} \psi_j u_{t-j},\, t \in \mathbb{Z} \tag{5.2}$$

where $\{u_t : t \in \mathbb{Z}\} \sim \mathrm{WN}(0, \sigma^2)$.

5.1. Convergence conditions

Theorem 5.1 MEAN SQUARE CONVERGENCE OF AN INFINITE MOVING AVERAGE. *Let* $\{\psi_j : j \in \mathbb{Z}\}$ *be a sequence of fixed real constants and* $\{u_t : t \in \mathbb{Z}\} \sim \mathrm{WN}(0, \sigma^2)$.

(1) *If* $\sum_{j=1}^{\infty} \psi_j^2 < \infty$, $\sum_{j=1}^{\infty} \psi_j u_{t-j}$ *converges in q.m. to a random variable* X_{Ut} *in* L_2.

(2) *If* $\sum_{j=-\infty}^{0} \psi_j^2 < \infty$, $\sum_{j=-\infty}^{0} \psi_j u_{t-j}$ *converges in q.m. to a random variable* X_{Lt} *in* L_2.

(3) *If* $\sum_{j=-\infty}^{\infty} \psi_j^2 < \infty$, $\sum_{j=-\infty}^{\infty} \psi_j u_{t-j}$ *converges in q.m. to a random variable* X_t *in* L_2, *and* $\sum_{j=-n}^{n} \psi_j u_{t-j} \xrightarrow[n\to\infty]{2} X_t$.

PROOF. Suppose $\sum_{j=0}^{\infty} \psi_j^2 < \infty$. We can write

$$\sum_{j=1}^{\infty} \psi_j u_{t-j} = \sum_{j=1}^{\infty} Y_j(t), \quad \sum_{j=-\infty}^{0} \psi_j u_{t-j} = \sum_{j=-\infty}^{0} Y_j(t) \tag{5.3}$$

where $Y_j(t) := \psi_j u_{t-j}$,

$$\mathbb{E}\big[Y_j(t)^2\big] = \psi_j^2 \,\mathbb{E}(u_{t-j}^2) = \psi_j^2 \sigma^2 < \infty, \text{ for } t \in \mathbb{Z},$$

and the variables $Y_j(t)$, $t \in \mathbb{Z}$, are orthogonal. If $\sum_{j=1}^{\infty} \psi_j^2 < \infty$, the series $\sum_{j=1}^{\infty} Y_j(t)$ converges in q.m. to a random variable X_{Ut} such that $\mathbb{E}\big[X_{Ut}^2\big] < \infty$, i.e.

$$\sum_{j=1}^{n} Y_j(t) \underset{n\to\infty}{\overset{2}{\to}} X_{Ut} := \sum_{j=1}^{\infty} \psi_j u_{t-j}\,; \tag{5.4}$$

see Dufour (2016, Section on "Series of orthogonal variables"). By a similar argument, if $\sum_{j=-\infty}^{0} \psi_j^2 < \infty$, the series $\sum_{j=-\infty}^{0} Y_j(t)$ converges in q.m. to a random variable X_{Ut} such that $\mathbb{E}\big[X_{Lt}^2\big] < \infty$, i.e.

$$\sum_{j=-m}^{0} Y_j(t) \underset{m\to\infty}{\overset{2}{\to}} X_{Lt} := \sum_{j=-\infty}^{0} \psi_j u_{t-j}\,. \tag{5.5}$$

Finally, if $\sum_{j=-\infty}^{\infty} \psi_j^2 < \infty$, we must have $\sum_{j=1}^{\infty} \psi_j^2 < \infty$ and $\sum_{j=-\infty}^{0} \psi_j^2 < \infty$, hence

$$\sum_{j=-m}^{n} Y_j(t) = \sum_{j=-m}^{0} Y_j(t) + \sum_{j=1}^{n} Y_j(t) \underset{\substack{m\to\infty \\ n\to\infty}}{\overset{2}{\to}} X_{Lt} + X_{Ut} := X_t := \sum_{j=-\infty}^{\infty} \psi_j u_{t-j} \tag{5.6}$$

where, by the c_r-inequality [see Dufour (2022)],

$$\mathbb{E}\big[X_t^2\big] = \mathbb{E}\big[(X_{Lt} + X_{Ut})^2\big] \le 2\{\mathbb{E}\big[X_{Lt}^2\big] + \mathbb{E}\big[X_{Ut}^2\big]\} < \infty\,. \tag{5.7}$$

The random variable X_t is denoted:

$$X_t := \sum_{j=-\infty}^{\infty} \psi_j u_{t-j}\,. \tag{5.8}$$

The last statement on the convergence of $\sum_{j=-n}^{n} \psi_j u_{t-j}$ follows from the definition of mean-square convergence of $\sum_{j=-\infty}^{\infty} \psi_j u_{t-j}$. □

Corollary 5.2 ALMOST SURE CONVERGENCE OF AN INFINITE MOVING AVERAGE. *Let* $\{\psi_j : j \in \mathbb{Z}\}$ *be a sequence of fixed real constants, and* $\{u_t : t \in \mathbb{Z}\} \sim \mathrm{WN}(0, \sigma^2)$.

(1) *If* $\sum_{j=1}^{\infty} |\psi_j| < \infty$, $\sum_{j=1}^{\infty} \psi_j u_{t-j}$ *converges a.s. and in q.m. to a random variable* X_{Ut} *in* L_2.

(2) *If* $\sum_{j=-\infty}^{0} |\psi_j| < \infty$, $\sum_{j=-\infty}^{0} \psi_j u_{t-j}$ *converges a.s. and in q.m. to a random variable* X_{Lt} *in* L_2.

(3) *If* $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$, $\sum_{j=-\infty}^{\infty} \psi_j u_{t-j}$ *converges a.s. and in q.m. to a random variable* X_t *in* L_2, $\sum_{j=-n}^{n} \psi_j u_{t-j} \xrightarrow[n\to\infty]{a.s.} X_t$ *and* $\sum_{j=-n}^{n} \psi_j u_{t-j} \xrightarrow[n\to\infty]{2} X_t$.

PROOF. This result from Theorem 5.1 and the observation that

$$\sum_{j=-\infty}^{\infty} |\psi_j| < \infty \Rightarrow \sum_{j=-\infty}^{\infty} \psi_j^2 < \infty . \tag{5.9}$$

□

Theorem 5.3 ALMOST SURE CONVERGENCE OF AN INFINITE MOVING AVERAGE OF INDEPENDENT VARIABLES. *Let* $\{\psi_j : j \in \mathbb{Z}\}$ *be a sequence of fixed real constants, and* $\{u_t : t \in \mathbb{Z}\} \sim \mathrm{IID}(0, \sigma^2)$.

(1) *If* $\sum_{j=1}^{\infty} \psi_j^2 < \infty$, $\sum_{j=1}^{\infty} \psi_j u_{t-j}$ *converges a.s. and in q.m. to a random variable* X_{Ut} *in* L_2.

(2) *If* $\sum_{j=-\infty}^{0} \psi_j^2 < \infty$, $\sum_{j=-\infty}^{0} \psi_j u_{t-j}$ *converges a.s. and in q.m. to a random variable* X_{Lt} *in* L_2.

(3) *If* $\sum_{j=-\infty}^{\infty} \psi_j^2 < \infty$, $\sum_{j=-\infty}^{\infty} \psi_j u_{t-j}$ *converges a.s. and in q.m. to a random variable* X_t *in* L_2, $\sum_{j=-n}^{n} \psi_j u_{t-j} \xrightarrow[n\to\infty]{a.s.} X_t$ *and* $\sum_{j=-n}^{n} \psi_j u_{t-j} \xrightarrow[n\to\infty]{2} X_t$.

PROOF. This result from Theorem 5.1 and by applying results on the convergence of series of independent variables [Dufour (2016, Section on "Series of independent variables")]. □

5.2. Mean, variance and covariances

Let

$$S_n(t) = \sum_{j=-n}^{n} \psi_j u_{t-j}\,. \tag{5.10}$$

By Theorem 5.1, we have:

$$S_n(t) \xrightarrow[n\to\infty]{2} X_t \tag{5.11}$$

where $X_t \in L_2$, hence [using Dufour (2016, Section on "Convergence of functions of random variables")]

$$\mathbb{E}(X_t) = \lim_{n\to\infty} \mathbb{E}[S_n(t)] = 0\,, \tag{5.12}$$

$$\mathrm{V}(X_t) = \mathbb{E}(X_t^2) = \lim_{n\to\infty} \mathbb{E}[S_n(t)^2] = \lim_{n\to\infty} \sum_{j=-n}^{n} \psi_j^2 \sigma^2 = \sigma^2 \sum_{j=-\infty}^{\infty} \psi_j^2\,, \tag{5.13}$$

$$\begin{aligned}
\mathrm{Cov}(X_t, X_{t+k}) &= \mathbb{E}(X_t X_{t+k}) = \lim_{n\to\infty} \mathbb{E}\left[\left(\sum_{i=-n}^{n} \psi_i u_{t-i}\right)\left(\sum_{j=-n}^{n} \psi_j u_{t+k-j}\right)\right] \\
&= \lim_{n\to\infty} \sum_{i=-n}^{n}\sum_{j=-n}^{n} \psi_i \psi_j \mathbb{E}(u_{t-i}u_{t+k-j}) \\
&= \begin{cases} \lim_{n\to\infty} \sum_{i=-n}^{n-k} \psi_i \psi_{i+k} \sigma^2 = \sigma^2 \sum_{i=-\infty}^{\infty} \psi_i \psi_{i+k}, & \text{if } k \geq 1, \\ \lim_{n\to\infty} \sum_{j=-n}^{n} \psi_j \psi_{j+|k|} \sigma^2 = \sigma^2 \sum_{j=-\infty}^{\infty} \psi_j \psi_{j+|k|}\,, & \text{if } k \leq -1, \end{cases}
\end{aligned} \tag{5.14}$$

since $t-i = t+k-j \Rightarrow j = i+k$ and $i = j-k$. For any $k \in \mathbb{Z}$, we can write

$$\mathrm{Cov}(X_t, X_{t+k}) = \sigma^2 \sum_{j=-\infty}^{\infty} \psi_j \psi_{j+|k|}\,, \tag{5.15}$$

$$\mathrm{Corr}(X_t, X_{t+k}) = \sum_{j=-\infty}^{\infty} \psi_j \psi_{j+|k|} \Big/ \sum_{j=-\infty}^{\infty} \psi_j^2\,. \tag{5.16}$$

The series $\sum_{j=-\infty}^{\infty} \psi_j \psi_{j+k}$ converges absolutely, for

$$\left|\sum_{j=-\infty}^{\infty} \psi_j \psi_{j+k}\right| \leq \sum_{j=-\infty}^{\infty} |\psi_j \psi_{j+k}| \leq \left[\sum_{j=-\infty}^{\infty} \psi_j^2\right]^{\frac{1}{2}} \left[\sum_{j=-\infty}^{\infty} \psi_{j+k}^2\right]^{\frac{1}{2}} < \infty\,. \tag{5.17}$$

If $X_t = \mu + X_t = \mu + \sum_{j=-\infty}^{+\infty} \psi_j u_{t-j}$, then

$$\mathbb{E}(X_t) = \mu\,,\ \mathrm{Cov}(X_t, X_{t+k}) = \mathrm{Cov}(X_t, X_{t+k})\,. \tag{5.18}$$

In the case of a causal MA(∞) process causal, we have

$$X_t = \mu + \sum_{j=0}^{\infty} \psi_j u_{t-j} \tag{5.19}$$

where $\{u_t : t \in \mathbb{Z}\} \sim \mathrm{WN}(0, \sigma^2)$,

$$\mathrm{Cov}(X_t, X_{t+k}) = \sigma^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+|k|} , \tag{5.20}$$

$$\mathrm{Corr}(X_t, X_{t+k}) = \sum_{j=0}^{\infty} \psi_j \psi_{j+|k|} / \sum_{j=0}^{\infty} \psi_j^2 . \tag{5.21}$$

5.3. Stationarity

The process

$$X_t = \mu + \sum_{j=-\infty}^{\infty} \psi_j u_{t-j} , t \in \mathbb{Z}, \tag{5.22}$$

where $\{u_t : t \in \mathbb{Z}\} \sim \mathrm{WN}(0, \sigma^2)$ and $\sum_{j=-\infty}^{\infty} \psi_j^2 < \infty$, is second-order stationary, for $\mathbb{E}(X_t)$ and $\mathrm{Cov}(X_t, X_{t+k})$ do not depend on t. If we suppose that $\{u_t : t \in \mathbb{Z}\} \sim$IID, with $\mathbb{E}|u_t| < \infty$ and $\sum_{j=-\infty}^{\infty} \psi_j^2 < \infty$, the process is strictly stationary.

5.4. Operational notation

We can denote the process MA(∞)

$$X_t = \mu + \psi(B)u_t = \mu + \left(\sum_{j=-\infty}^{\infty} \psi_j B^j \right) u_t \tag{5.23}$$

where $\psi(B) = \sum_{j=-\infty}^{\infty} \psi_j B^j$ and $B^j u_t = u_{t-j}$.

6. Finite order moving averages

The MA(q) process can be written

$$X_t = \mu + u_t - \sum_{j=1}^{q} \theta_j u_{t-j} \tag{6.1}$$

where $\theta(B) = 1 - \theta_1 B - \cdots - \theta_q B^q$. This process is a special case of the MA(∞) process with

$$\begin{aligned} \psi_0 &= 1\,,\ \psi_j = -\theta_j\,,\ \text{for } 1 \le j \le q\,, \\ \psi_j &= 0\,,\ \text{for } j < 0 \text{ or } j > q. \end{aligned} \tag{6.2}$$

This process is clearly second-order stationary, with

$$\mathbb{E}(X_t) = \mu\,, \tag{6.3}$$

$$\mathrm{V}(X_t) = \sigma^2 \left(1 + \sum_{j=1}^{q} \theta_j^2\right), \tag{6.4}$$

$$\gamma(k) := \mathrm{Cov}(X_t, X_{t+k}) = \sigma^2 \sum_{j=-\infty}^{\infty} \psi_j \psi_{j+|k|}\,. \tag{6.5}$$

On defining $\theta_0 := -1$, we then see that

$$\begin{aligned} \gamma(k) &= \sigma^2 \sum_{j=0}^{q-k} \theta_j \theta_{j+k} \\ &= \sigma^2 \left[-\theta_k + \sum_{j=1}^{q-k} \theta_j \theta_{j+k}\right] \qquad (6.6) \\ &= \sigma^2 [-\theta_k + \theta_1 \theta_{k+1} + \ldots + \theta_{q-k} \theta_q]\,,\ \text{for } 1 \le k \le q, \qquad (6.7) \end{aligned}$$

$$\begin{aligned} &\gamma(k) = 0\,,\ \text{for } k \ge q+1, \\ &\gamma(-k) = \gamma(k)\,,\ \text{for } k < 0\,. \end{aligned} \tag{6.8}$$

The autocorrelation function of X_t is thus

$$\begin{aligned} \rho(k) &= \left(-\theta_k + \sum_{j=1}^{q-k} \theta_j \theta_{j+k}\right) \Big/ \left(1 + \sum_{j=1}^{q} \theta_j^2\right), \quad \text{for } 1 \le k \le q \\ &= 0, \qquad \text{for } k \ge q+1 \end{aligned} \tag{6.9}$$

The autocorrelations are zero for $k \ge q+1$.

For $q = 1$,

$$\begin{aligned} \rho(k) &= -\theta_1/(1+\theta_1^2), \quad \text{if } k = 1\,, \\ &= 0, \qquad \text{if } k \ge 2, \end{aligned} \tag{6.10}$$

hence $|\rho(1)| \le 0.5$. For $q = 2$,

$$\begin{aligned} \rho(k) &= (-\theta_1 + \theta_1\theta_2)/(1+\theta_1^2+\theta_2^2)\,, \quad \text{if } k = 1\,, \\ &= -\theta_2/(1+\theta_1^2+\theta_2^2)\,, \quad \text{if } k = 2\,, \\ &= 0\,, \qquad \text{if } k \ge 3\,, \end{aligned} \tag{6.11}$$

hence $|\rho(2)| \le 0.5$.

For any MA(q) process,

$$\rho(q) = -\theta_q/(1+\theta_1^2+\ldots+\theta_q^2)\,, \tag{6.12}$$

hence $|\rho(q)| \leq 0.5$.

There are general constraints on the autocorrelations of an MA(q) process:

$$|\rho(k)| \leq \cos(\pi/\{[q/k]+2\}) \tag{6.13}$$

where $[x]$ is the largest integer less than or equal to x. From the latter formula, we see:

$$\begin{array}{ll} \text{for } q=1\,, & |\rho(1)| \leq \cos(\pi/3) = 0.5, \\ \text{for } q=2\,, & |\rho(1)| \leq \cos(\pi/4) = 0.7071, \\ & |\rho(2)| \leq \cos(\pi/3) = 0.5, \\ \text{for } q=3\,, & |\rho(1)| \leq \cos(\pi/5) = 0.809, \\ & |\rho(2)| \leq \cos(\pi/3) = 0.5, \\ & |\rho(3)| \leq \cos(\pi/3) = 0.5. \end{array} \tag{6.14}$$

See Chanda (1962), and Kendall, Stuart, and Ord (1983, p. 519).

7. Autoregressive processes

Consider a process $\{X_t : t \in \mathbb{Z}\}$ which satisfies the equation:

$$X_t = \bar{\mu} + \sum_{j=1}^{p} \varphi_j X_{t-j} + u_t\,,\ \forall t \in \mathbb{Z}, \tag{7.1}$$

where $\{u_t : t \in \mathbb{Z}\} \sim \text{WN}(0,\, \sigma^2)$. In symbolic notation,

$$\varphi(B)X_t = \bar{\mu} + u_t\,,\ t \in \mathbb{Z}, \tag{7.2}$$

where $\varphi(B) = 1 - \varphi_1 B - \cdots - \varphi_p B^p$.

7.1. Stationarity

Consider the process AR(1)

$$X_t = \varphi_1 X_{t-1} + u_t\,,\ \varphi_1 \neq 0. \tag{7.3}$$

If X_t is S2 ,

$$\mathbb{E}(X_t) = \varphi_1 \mathbb{E}(X_{t-1}) = \varphi_1 \mathbb{E}(X_t) \tag{7.4}$$

hence $\mathbb{E}(X_t) = 0$. By successive substitutions,

$$\begin{aligned} X_t &= \varphi_1[\varphi_1 X_{t-2} + u_{t-1}] + u_t \\ &= u_t + \varphi_1 u_{t-1} + \varphi_1^2 X_{t-2} \end{aligned}$$

$$= \sum_{j=0}^{N-1} \varphi_1^j u_{t-j} + \varphi_1^N X_{t-N} \,. \tag{7.5}$$

If we suppose that X_t is S2 with $\mathbb{E}(X_t^2) \neq 0$, we see that

$$\mathbb{E}\left[\left(X_t - \sum_{j=0}^{N-1} \varphi_1^j u_{t-j}\right)^2\right] = \varphi_1^{2N}\mathbb{E}(X_{t-N}^2) = \varphi_1^{2N}\mathbb{E}(X_t^2) \underset{N\to\infty}{\to} 0 \Leftrightarrow |\varphi_1| < 1. \tag{7.6}$$

The series $\sum_{j=0}^{\infty} \varphi_1^j u_{t-j}$ converges in *q.m.* to

$$X_t = \sum_{j=0}^{\infty} \varphi_1^j u_{t-j} := (1-\varphi_1 B)^{-1} u_t = \frac{1}{1-\varphi_1 B}\, u_t \tag{7.7}$$

where

$$(1-\varphi_1 B)^{-1} = \sum_{j=0}^{\infty} \varphi_1^j B^j \,. \tag{7.8}$$

Since

$$\sum_{j=0}^{\infty} \mathbb{E}|\varphi_1^j u_{t-j}| \leq \sigma \sum_{j=0}^{\infty} |\varphi_1|^j = \frac{\sigma}{1-|\varphi_1|} < \infty \tag{7.9}$$

when $|\varphi_1| < 1$, the convergence is also a.s. The process $X_t = \sum_{j=0}^{\infty} \varphi_1^j u_{t-j}$ is S2.

When $|\varphi_1| < 1$, the difference equation

$$(1-\varphi_1 B) X_t = u_t \tag{7.10}$$

has a unique stationary solution which can be written

$$X_t = \sum_{j=0}^{\infty} \varphi_1^j u_{t-j} = (1-\varphi_1 B)^{-1} u_t. \tag{7.11}$$

The latter is thus a causal MA(∞) process.

This condition is sufficient (but non necessary) for the existence of a unique stationary solution. The stationarity condition is often expressed by saying that the polynomial $\varphi(z) = 1 - \varphi_1 z$ has all its roots outside the unit circle $|z| = 1$:

$$1 - \varphi_1 z_* = 0 \Leftrightarrow z_* = \frac{1}{\varphi_1} \tag{7.12}$$

where $|z_*| = 1/|\varphi_1| > 1$. In this case, we also have $\mathbb{E}(X_{t-k} u_t) = 0, \; \forall k \geq 1$. The same conclusion holds if we consider the general process

$$X_t = \bar{\mu} + \varphi_1 X_{t-1} + u_t \ . \tag{7.13}$$

For the AR(p) process,

$$X_t = \bar{\mu} + \sum_{j=1}^{p} \varphi_j X_{t-j} + u_t \tag{7.14}$$

or

$$\varphi(B) X_t = \bar{\mu} + u_t, \tag{7.15}$$

the stationarity condition is the following:

$$\begin{gathered}\text{if the polynomial } \varphi(z) = 1 - \varphi_1 z - \cdots - \varphi_p z^p \text{ has all its roots outside the unit circle,} \\ \text{the equation (7.14) has one and only one weakly stationary solution.}\end{gathered} \tag{7.16}$$

$\varphi(z)$ is a polynomial of order p with no root equal to zero. It can be written in the form

$$\varphi(z) = (1 - G_1 z)(1 - G_2 z) \ldots (1 - G_p z) \,, \tag{7.17}$$

so the roots of $\varphi(z)$ are

$$z_1^* = 1/G_1, \ldots, z_p^* = 1/G_p \,, \tag{7.18}$$

and the stationarity condition have the equivalent form:

$$|G_j| < 1, \ j = 1, \ldots, p. \tag{7.19}$$

The stationary solution can be written

$$X_t = \varphi(B)^{-1} \bar{\mu} + \varphi(B)^{-1} u_t = \mu + \varphi(B)^{-1} u_t \tag{7.20}$$

where

$$\mu = \bar{\mu} \Big/ \left(1 - \sum_{j=1}^{p} \varphi_j \right) , \tag{7.21}$$

$$\begin{aligned} \varphi(B)^{-1} &= \prod_{j=1}^{p} (1 - G_j B)^{-1} = \prod_{j=1}^{p} \left(\sum_{k=0}^{\infty} G_j^k B^k \right) \\ &= \sum_{j=1}^{p} \frac{K_j}{1 - G_j B} \end{aligned} \tag{7.22}$$

and $K_1, \ldots, K_p$ are constants (expansion in partial fractions). Consequently,

$$X_t = \mu + \sum_{j=1}^{p} \left(\frac{K_j}{1 - G_j B} \right) u_t$$

$$= \quad \mu + \sum_{k=0}^{\infty} \psi_k u_{t-k} = \mu + \psi(B) u_t \tag{7.23}$$

where $\psi_k = \sum_{j=1}^{p} K_j G_j^k$. Thus

$$\mathbb{E}(X_{t-j} u_t) = 0, \forall j \geq 1. \tag{7.24}$$

For the process AR(1) and AR(2), the stationarity conditions can be written as follows.

(a) AR(1) – For $(1 - \varphi_1 B) X_t = \bar{\mu} + u_t$,

$$|\varphi_1| < 1 \tag{7.25}$$

(b) AR(2) – For $(1 - \varphi_1 B - \varphi_2 B^2) X_t = \bar{\mu} + u_t$,

$$\varphi_2 + \varphi_1 < 1 \tag{7.26}$$

$$\varphi_2 - \varphi_1 < 1 \tag{7.27}$$

$$-1 < \varphi_2 < 1 \tag{7.28}$$

7.2. Mean, variance and autocovariances

Suppose:

a) the autoregressive process X_t is second-order stationary with $\sum_{j=1}^{p} \varphi_j \neq 1$
and
b) $\mathbb{E}(X_{t-j} u_t) = 0$, $\forall j \geq 1$, (7.29)

i.e., we assume that X_t is a weakly stationary solution of the equation (7.14) such that $\mathbb{E}(X_{t-j} u_t) = 0$, $\forall j \geq 1$.

By the stationarity assumption, we have: $\mathbb{E}(X_t) = \mu$, $\forall t$, hence

$$\mu = \bar{\mu} + \sum_{j=1}^{p} \varphi_j \mu \tag{7.30}$$

and

$$\mathbb{E}(X_t) = \mu = \bar{\mu} \Big/ \left(1 - \sum_{j=1}^{p} \varphi_j\right) . \tag{7.31}$$

For stationarity to hold, it is necessary that $\sum_{j=1}^{p} \varphi_j \neq 1$. Let us rewrite the process in the form

$$\tilde{X}_t = \sum_{j=1}^{p} \varphi_j \tilde{X}_{t-j} + u_t \tag{7.32}$$

where $\tilde{X}_t = X_t - \mu$, $\mathbb{E}(\tilde{X}_t) = 0$. Then, for $k \geq 0$,

$$\tilde{X}_{t+k} = \sum_{j=1}^{p} \varphi_j \tilde{X}_{t+k-j} + u_{t+k}\,, \tag{7.33}$$

$$\mathbb{E}(\tilde{X}_{t+k}\,\tilde{X}_t) = \sum_{j=1}^{p} \varphi_j \mathbb{E}(\tilde{X}_{t+k-j}\tilde{X}_t) + \mathbb{E}(u_{t+k}\tilde{X}_t)\,, \tag{7.34}$$

$$\gamma(k) = \sum_{j=1}^{p} \varphi_j \gamma(k-j) + \mathbb{E}(u_{t+k}\,\tilde{X}_t)\,, \tag{7.35}$$

where

$$\begin{aligned} \mathbb{E}(u_{t+k}\,\tilde{X}_t) &= \sigma^2, \quad \text{if } k = 0, \\ &= 0\,, \quad \text{if } k \geq 1. \end{aligned} \tag{7.36}$$

Thus

$$\rho(k) = \sum_{j=1}^{p} \varphi_j \rho(k-j),\, k \geq 1. \tag{7.37}$$

These formulae are called the “Yule-Walker equations”. If we know $\rho(0), \ldots, \rho(p-1)$, we can easily compute $\rho(k)$ for $k \geq p+1$. We can also write the Yule-Walker equations in the form:

$$\varphi(B)\rho(k) = 0, \text{ for } k \geq 1, \tag{7.38}$$

where $B^j \rho(k) := \rho(k-j)$. To obtain $\rho(1), \ldots, \rho(p-1)$ for $p > 1$, it is sufficient to solve the linear equation system:

$$\begin{aligned} \rho(1) &= \varphi_1 + \varphi_2 \rho(1) + \cdots + \varphi_p \rho(p-1) \\ \rho(2) &= \varphi_1 \rho(1) + \varphi_2 + \cdots + \varphi_p \rho(p-2) \\ &\vdots \\ \rho(p-1) &= \varphi_1 \rho(p-2) + \varphi_2 \rho(p-3) + \cdots + \varphi_p \rho(1) \end{aligned} \tag{7.39}$$

where we use the identity $\rho(-j) = \rho(j)$. The other autocorrelations may then be obtained by recurrence:

$$\rho(k) = \sum_{j=1}^{p} \varphi_j \rho(k-j),\; k \geq p\,. \tag{7.40}$$

To compute $\gamma(0) = V(X_t)$, we solve the equation

$$\begin{aligned} \gamma(0) &= \sum_{j=1}^{p} \varphi_j \gamma(-j) + \mathbb{E}(u_t\,\tilde{X}_t) \\ &= \sum_{j=1}^{p} \varphi_j \gamma(j) + \sigma^2 \end{aligned} \tag{7.41}$$

hence, using $\gamma(j) = \rho(j)\gamma(0)$,

$$\gamma(0)\left[1 - \sum_{j=1}^{p} \varphi_j \rho(j)\right] = \sigma^2 \tag{7.42}$$

and

$$\gamma(0) = \frac{\sigma^2}{1 - \sum_{j=1}^{p} \varphi_j \rho(j)}. \tag{7.43}$$

7.3. Special cases

1. AR(1) – If

$$\tilde{X}_t = \varphi_1 \tilde{X}_{t-1} + u_t \tag{7.44}$$

we have:

$$\begin{aligned}
\rho(1) &= \varphi_1\,, &(7.45)\\
\rho(k) &= \varphi_1 \rho(k-1)\,,\ \text{for } k \geq 1\,, &(7.46)\\
\rho(2) &= \varphi_1 \rho(1) = \varphi_1^2\,, &(7.47)\\
\rho(k) &= \varphi_1^k,\ k \geq 1\,, &(7.48)\\
\gamma(0) &= \mathrm{V}(X_t) = \frac{\sigma^2}{1 - \varphi_1^2}\,. &(7.49)
\end{aligned}$$

These is no constraint on $\rho(1)$, but there are constraints on $\rho(k)$ for $k \geq 2$.

2. AR(2) – If

$$X_t = \varphi_1 \tilde{X}_{t-1} + \varphi_2 \tilde{X}_{t-2} + u_t\,, \tag{7.50}$$

we have:

$$\begin{aligned}
\rho(1) &= \varphi_1 + \varphi_2 \rho(1)\,, &(7.51)\\
\rho(1) &= \frac{\varphi_1}{1 - \varphi_2}\,, &(7.52)\\
\rho(2) &= \frac{\varphi_1^2}{1 - \varphi_2} + \varphi_2 = \frac{\varphi_1^2 + \varphi_2 (1 - \varphi_2)}{1 - \varphi_2}\,, &(7.53)\\
\rho(k) &= \varphi_1 \rho(k-1) + \varphi_2 \rho(k-2),\ \text{for } k \geq 2\,. &(7.54)
\end{aligned}$$

Constraints on $\rho(1)$ and $\rho(2)$ are entailed by the stationarity of the AR(2) model:

$$\begin{aligned}
|\rho(1)| &< 1,\ |\rho(2)| < 1\,, &(7.55)\\
\rho(1)^2 &< \frac{1}{2}[1 + \rho(2)]\,; &(7.56)
\end{aligned}$$

see Box and Jenkins (1976, p. 61).

7.4. Explicit form for the autocorrelations

The autocorrelations of an AR(p) process satisfy the equation

$$\rho(k) = \sum_{j=1}^{p} \varphi_j \rho(k-j), k \geq 1, \tag{7.57}$$

where $\rho(0) = 1$ and $\rho(-k) = \rho(k)$, or equivalently

$$\varphi(B)\rho(k) = 0\,,\ k \geq 1. \tag{7.58}$$

The autocorrelations can be obtained by solving the homogeneous difference equation (7.57).

The polynomial $\varphi(z)$ has m distinct non-zero roots $z_1^*, \ldots, z_m^*$ (where $1 \leq m \leq p$) with multiplicities $p_1, \ldots, p_m$ (where $\sum\limits_{j=1}^{m} p_j = p$), so that $\varphi(z)$ can be written

$$\varphi(z) = (1 - G_1 z)^{p_1}(1 - G_2 z)^{p_2} \cdots (1 - G_m z)^{p_m} \tag{7.59}$$

where $G_j = 1/z_j^*$, $j = 1, \ldots, m$. The roots are real or complex numbers. If z_j^* is a complex (non real) root, its conjugate $\bar{z}_j^*$ is also a root. Consequently, the solutions of equation (7.57) have the general form

$$\rho(k) = \sum_{j=1}^{m} \left(\sum_{\ell=0}^{p_j - 1} A_{j\ell} k^{\ell} \right) G_j^k, k \geq 1, \tag{7.60}$$

where the $A_{j\ell}$ are (possibly complex) constants which can be determined from the values p autocorrelations. We can easily find $\rho(1), \ldots, \rho(p)$ from the Yule-Walker equations.

If we write $G_j = r_j e^{i\theta_j}$, where $i = \sqrt{-1}$ while r_j and θ_j are real numbers $(r_j > 0)$, we see that

$$\begin{aligned}
\rho(k) &= \sum_{j=1}^{m} \left(\sum_{\ell=0}^{p_j - 1} A_{j\ell} k^{\ell} \right) r_j^k e^{i\theta_j k} \\
&= \sum_{j=1}^{m} \left(\sum_{\ell=0}^{p_j - 1} A_{j\ell} k^{\ell} \right) r_j^k [\cos(\theta_j k) + i\ \sin(\theta_j k)] \\
&= \sum_{j=1}^{m} \left(\sum_{\ell=0}^{p_j - 1} A_{j\ell} k^{\ell} \right) r_j^k \cos(\theta_j k).
\end{aligned} \tag{7.61}$$

By stationarity, $0 < |G_j| = r_j < 1$ so that $\rho(k) \to 0$ when $k \to \infty$. The autocorrelations decrease at an exponential rate with oscillations.

7.5. MA(∞) representation of an AR(p) process

We have seen that a weakly stationary process which satisfies the equation

$$\varphi(B)\tilde{X}_t = u_t \tag{7.62}$$

where $\varphi(B) = 1 - \varphi_1 B - \cdots - \varphi_p B^p$, can be written as

$$\tilde{X}_t = \psi(B) u_t \tag{7.63}$$

with

$$\psi(B) = \varphi(B)^{-1} = \sum_{j=0}^{\infty} \psi_j B^j \tag{7.64}$$

To compute the coefficients ψ_j, it is sufficient to note that

$$\varphi(B)\psi(B) = 1. \tag{7.65}$$

Setting $\psi_j = 0$ for $j < 0$, we see that

$$\begin{aligned} \left(1 - \sum_{k=1}^{p} \varphi_k B^k\right)\left(\sum_{j=-\infty}^{\infty} \psi_j B^j\right) &= \sum_{j=-\infty}^{\infty} \psi_j \left(B^j - \sum_{k=1}^{p} \varphi_k B^{j+k}\right) \\ &= \sum_{j=-\infty}^{\infty} \left(\psi_j - \sum_{k=1}^{p} \varphi_k \psi_{j-k}\right) B^j = \sum_{j=-\infty}^{\infty} \tilde{\psi}_j B^j = 1. \end{aligned} \tag{7.66}$$

Thus $\tilde{\psi}_j = 1$, if $j = 0$, and $\tilde{\psi}_j = 0$, if $j \neq 0$. Consequently,

$$\begin{aligned} \varphi(B)\psi_j = \psi_j - \sum_{k=1}^{p} \varphi_k \psi_{j-k} \quad &= 1\,,\ \text{if } j = 0 \\ &= 0\,,\ \text{if } j \neq 0, \end{aligned} \tag{7.67}$$

where $B^k \psi_j := \psi_{j-k}$. Since $\psi_j = 0$ for $j < 0$, we see that:

$$\begin{aligned} \psi_0 &= 1\,, \\ \psi_j &= \sum_{k=1}^{p} \varphi_k \psi_{j-k}\,,\ \text{for } j \geq 1. \end{aligned} \tag{7.68}$$

More explicitly,

$$\begin{aligned} \psi_0 &= 1\,, \\ \psi_1 &= \varphi_1 \psi_0 = \varphi_1\,, \\ \psi_2 &= \varphi_1 \psi_1 + \varphi_2 \psi_0 = \varphi_1^2 + \varphi_2\,, \\ \psi_3 &= \varphi_1 \psi_2 + \varphi_2 \psi_1 + \varphi_3 = \varphi_1^3 + 2\ \varphi_2 \varphi_1 + \varphi_3\,, \\ &\vdots \\ \psi_p &= \sum_{k=1}^{p} \varphi_k \psi_{j-k}\,, \\ \psi_j &= \sum_{k=1}^{p} \varphi_k \psi_{j-k},\ j \geq p+1\,. \end{aligned} \tag{7.69}$$

Under the stationarity condition *i.e.*, the roots of $\varphi(z) = 0$ are outside the unit circle], the coefficients ψ_j decline at an exponential rate as $j \to \infty$, possibly with oscillations.

Given the representation

$$\tilde{X}_t = \psi(B)u_t = \sum_{j=0}^{\infty} \psi_j u_{t-j}\,, \tag{7.70}$$

we can easily compute the autocovariances and autocorrelations of X_t :

$$\begin{aligned} \mathrm{Cov}(X_t, X_{t+k}) &= \sigma^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+|k|}\,, & (7.71) \\ \mathrm{Corr}(X_t, X_{t+k}) &= \left(\sum_{j=0}^{\infty} \psi_j \psi_{j+|k|}\right) / \left(\sum_{j=0}^{\infty} \psi_j^2\right). & (7.72) \end{aligned}$$

However, this has the drawback of requiring one to compute limits of series.

7.6. Partial autocorrelations

The Yule-Walker equations allow one to determine the autocorrelations from the coefficients $\varphi_1, \ldots, \varphi_p$. In the same way we can determine $\varphi_1, \ldots, \varphi_p$ from the autocorrelations

$$\rho(k) = \sum_{j=1}^{p} \varphi_j \rho(k-j),\; k = 1, 2, 3, \ldots \tag{7.73}$$

Taking into account the fact that $\rho(0) = 1$ and $\rho(-k) = \rho(k)$, we see that

$$\begin{bmatrix} 1 & \rho(1) & \rho(2) & \ldots & \rho(p-1) \\ \rho(1) & 1 & \rho(1) & \ldots & \rho(p-2) \\ \vdots & \vdots & \vdots & & \vdots \\ \rho(p-1) & \rho(p-2) & \rho(p-3) & \ldots & 1 \end{bmatrix} \begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_p \end{bmatrix} = \begin{bmatrix} \rho(1) \\ \rho(2) \\ \vdots \\ \rho(p) \end{bmatrix} \tag{7.74}$$

or, equivalently,

$$R(p)\,\bar{\varphi}(p) = \bar{\rho}(p) \tag{7.75}$$

where

$$R(p) = \begin{bmatrix} 1 & \rho(1) & \rho(2) & \ldots & \rho(p-1) \\ \rho(1) & 1 & \rho(1) & \ldots & \rho(p-2) \\ \vdots & \vdots & \vdots & & \vdots \\ \rho(p-1) & \rho(p-2) & \rho(p-3) & \ldots & 1 \end{bmatrix}, \tag{7.76}$$

$$\bar{\rho}(p) = \begin{bmatrix} \rho(1) \\ \rho(2) \\ \vdots \\ \rho(p) \end{bmatrix}, \quad \bar{\varphi}(p) = \begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_p \end{bmatrix}. \tag{7.77}$$

Consider now the sequence of equations

$$R(k)\ \bar{\varphi}(k) = \bar{\rho}(k),\ k = 1, 2, 3, ... \tag{7.78}$$

where

$$R(k) = \begin{bmatrix} 1 & \rho(1) & \rho(2) & \dots & \rho(k-1) \\ \rho(1) & 1 & \rho(1) & \dots & \rho(k-2) \\ \vdots & \vdots & \vdots & & \vdots \\ \rho(k-1) & \rho(k-2) & \rho(k-3) & \dots & 1 \end{bmatrix}, \tag{7.79}$$

$$\bar{\rho}(k) = \begin{bmatrix} \rho(1) \\ \rho(2) \\ \vdots \\ \rho(k) \end{bmatrix}, \quad \bar{\varphi}(k) = \begin{bmatrix} \varphi(1|k) \\ \varphi(2|k) \\ \vdots \\ \varphi(k|k) \end{bmatrix},\ k = 1, 2, 3, ... \tag{7.80}$$

so that we can solve for $\bar{\varphi}(k)$:

$$\bar{\varphi}(k) = R(k)^{-1}\,\bar{\rho}(k)\,. \tag{7.81}$$

[If $\sigma^2 > 0$, we can show that $R(k)^{-1}$ exists, $\forall k \geq 1$]. On using (7.75), we see easily that:

$$\varphi_k(k) = 0 \text{ for } k \geq p+1\,. \tag{7.82}$$

The coefficients φ_{kk} are called the lag- k *partial autocorrelations*.

In particular,

$$\varphi_1(|1) = \rho(1), \tag{7.83}$$

$$\varphi_2(2|2) = \frac{\begin{vmatrix} 1 & \rho(1) \\ \rho(1) & \rho(2) \end{vmatrix}}{\begin{vmatrix} 1 & \rho(1) \\ \rho(1) & 1 \end{vmatrix}} = \frac{\rho(2) - \rho(1)^2}{1 - \rho(1)^2}, \tag{7.84}$$

$$\varphi_3(3|3) = \frac{\begin{vmatrix} 1 & \rho(1) & \rho(1) \\ \rho(1) & 1 & \rho(2) \\ \rho(2) & \rho(1) & \rho(3) \end{vmatrix}}{\begin{vmatrix} 1 & \rho(1) & \rho(2) \\ \rho(1) & 1 & \rho(1) \\ \rho(2) & \rho(1) & 1 \end{vmatrix}}. \tag{7.85}$$

The partial autocorrelations may be computed using the following recursive formulae:

$$\varphi(k+1|k+1) = \frac{\rho(k+1) - \sum_{j=1}^{k} \varphi(j|k)\,\rho(k+1-j)}{1 - \sum_{j=1}^{k} \varphi(j|k)\,\rho(j)}, \tag{7.86}$$

$$\varphi(j\,|\,k+1) \;=\; \varphi(j\,|\,k) - \varphi(k+1\,|\,k+1)\,\varphi(k+1-j\,|\,k)\,, \;\; j = 1, 2, \ldots, k. \tag{7.87}$$

Given $\rho(1), \ldots, \rho(k+1)$ and $\varphi_1(k), \ldots, \varphi_k(k)$, we can compute $\varphi_j(k+1)\,,\; j = 1, \ldots, k+1$. The expressions (7.86) - (7.87) are called the *Durbin-Levinson formulae*; see Durbin (1960) and Box and Jenkins (1976, pp. 82-84).

8. Mixed processes

Consider a process $\{X_t : t \in \mathbb{Z}\}$ which satisfies the equation:

$$X_t = \bar{\mu} + \sum_{j=1}^{p} \varphi_j\, X_{t-j} + u_t - \sum_{j=1}^{q} \theta_j\, u_{t-j} \tag{8.1}$$

where $\{u_t : t \in \mathbb{Z}\} \sim \mathrm{WN}(0, \sigma^2)$. Using operational notation, this can written

$$\varphi(B)\, X_t = \bar{\mu} + \theta(B)\, u_t. \tag{8.2}$$

8.1. Stationarity conditions

If the polynomial $\varphi(z) = 1 - \varphi_1 z - \cdots - \varphi_p z^p$ has all its roots outside the unit circle, the equation (8.1) has one and only one weakly stationary solution, which can be written:

$$X_t = \mu + \frac{\theta\,(B)}{\varphi\,(B)} u_t = \mu + \sum_{j=0}^{\infty} \psi_j u_{t-j} \tag{8.3}$$

where

$$\mu = \bar{\mu}/\varphi(B) = \bar{\mu}/(1 - \sum_{j=1}^{p} \varphi_j)\,, \tag{8.4}$$

$$\frac{\theta\,(B)}{\varphi\,(B)} := \psi(B) = \sum_{j=0}^{\infty} \psi_j B^j\,. \tag{8.5}$$

The coefficients ψ_j are obtained by solving the equation

$$\varphi(B)\psi(B) = \theta(B). \tag{8.6}$$

In this case, we also have:

$$\mathbb{E}(X_{t-j}\, u_t) = 0, \forall j \geq 1. \tag{8.7}$$

The ψ_j coefficients may be computed in the following way (setting $\theta_0 = -1$):

$$\left(1 - \sum_{k=1}^{p} \varphi_k B^k\right) \left(\sum_{j=0}^{\infty} \psi_j B^j\right) = 1 - \sum_{j=1}^{q} \theta_j B^j = -\sum_{j=1}^{q} \theta_j B^j \tag{8.8}$$

hence

$$\varphi(B)\psi_j \begin{array}{ll} = -\theta_j & \text{for } j = 0, 1, \ldots, q \\ = 0 & \text{for } j \geq q+1, \end{array} \tag{8.9}$$

where $\psi_j = 0$, for $j < 0$. Consequently,

$$\begin{aligned} \psi_j &= \sum_{k=1}^{p} \varphi_k \psi_{j-k} - \theta_j, \quad \text{for } j = 0, 1, \ldots, q \\ &= \sum_{k=1}^{p} \varphi_k \psi_{j-k}\,, \qquad\quad \text{for } j \geq q+1, \end{aligned} \tag{8.10}$$

and

$$\begin{aligned} \psi_0 &= 1\,, \\ \psi_1 &= \varphi_1\psi_0 - \theta_1 = \varphi_1 - \theta_1\,, \\ \psi_2 &= \varphi_1\psi_1 + \varphi_2\psi_0 - \theta_2 = \varphi_1\psi_1 + \varphi_2 - \theta_2 = \varphi_1^2 - \varphi_1\theta_1 + \varphi_2 - \theta_2\,, \\ &\vdots \\ \psi_j &= \sum_{k=1}^{p} \varphi_k \psi_{j-k}, \; j \geq q+1\,. \end{aligned} \tag{8.11}$$

The ψ_j coefficients behave like the autocorrelations of an AR(p) process, except for the initial coefficients $\psi_1, \ldots, \psi_q$.

8.2. Autocovariances and autocorrelations

Suppose:

$$\begin{array}{l} \text{a) the process } X_t \text{ is second-order stationary with } \sum_{j=1}^{p} \varphi_j \neq 1\,; \\ \text{b) } \mathbb{E}(X_{t-j}u_t) = 0\,, \forall j \geq 1\,. \end{array} \tag{8.12}$$

By the stationarity assumption,

$$\mathbb{E}(X_t) = \mu, \forall t, \tag{8.13}$$

hence

$$\mu = \bar{\mu} + \sum_{j=1}^{p} \varphi_j \mu \tag{8.14}$$

and

$$\mathbb{E}(X_t) = \mu = \bar{\mu} \Big/ \left(1 - \sum_{j=1}^{p} \varphi_j\right). \tag{8.15}$$

The mean is the same as in the case of a pure AR(p) process. The MA(q) component of the model has no effect on the mean. Let us now rewrite the process in the form

$$\tilde{X}_t = \sum_{j=1}^{p} \varphi_j \tilde{X}_{t-j} + u_t - \sum_{j=1}^{q} \theta_j u_{t-j} \tag{8.16}$$

where $\tilde{X}_t = X_t - \mu$. Consequently,

$$\begin{aligned} \tilde{X}_{t+k} &= \sum_{j=1}^{p} \varphi_j \tilde{X}_{t+k-j} + u_{t+k} - \sum_{j=1}^{q} \theta_j u_{t+k-j}\,, & (8.17) \\ \mathbb{E}(\tilde{X}_t \tilde{X}_{t+k}) &= \sum_{j=1}^{p} \varphi_j \mathbb{E}(\tilde{X}_t \tilde{X}_{t+k-j}) + \mathbb{E}(\tilde{X}_t u_{t+k}) - \sum_{j=1}^{q} \theta_j \mathbb{E}(\tilde{X}_t u_{t+k-j})\,, & (8.18) \\ \gamma(k) &= \sum_{j=1}^{p} \varphi_j \gamma(k-j) + \gamma_{xu}(k) - \sum_{j=1}^{q} \theta_j \gamma_{xu}(k-j)\,, & (8.19) \end{aligned}$$

where

$$\begin{aligned} \gamma_{xu}(k) &= \mathbb{E}(\tilde{X}_t u_{t+k}) = 0\,, \quad \text{if } k \geq 1\,, \\ &\neq 0\,, \qquad\qquad\qquad \text{if } k \leq 0\,, \\ \gamma_{xu}(0) &= \mathbb{E}(\tilde{X}_t u_t) = \sigma^2. \end{aligned} \tag{8.20}$$

For $k \geq q+1$,

$$\begin{aligned} \gamma(k) &= \sum_{j=1}^{p} \varphi_j \gamma(k-j)\,, & (8.21) \\ \rho(k) &= \sum_{j=1}^{p} \varphi_j \rho(k-j)\,. & (8.22) \end{aligned}$$

The variance is given by

$$\gamma(0) = \sum_{j=1}^{p} \varphi_j \gamma(j) + \sigma^2 - \sum_{j=1}^{q} \theta_j \gamma_{xu}(-j)\,, \tag{8.23}$$

hence

$$\gamma(0) = \left[\sigma^2 - \sum_{j=1}^{q} \theta_j \gamma_{xu}(-j)\right] / \left[1 - \sum_{j=1}^{p} \varphi_j \rho(j)\right]. \tag{8.24}$$

In operational notation, the autocovariances satisfy the equation

$$\varphi(B)\gamma(k) = \theta(B)\gamma_{xu}(k)\,,\ k \geq 0, \tag{8.25}$$

where $\gamma(-k) = \gamma(k)$, $B^j \gamma(k) := \gamma(k-j)$ and $B^j \gamma_{xu}(k) := \gamma_{xu}(k-j)$. In particular,

$$\varphi(B)\gamma(k) = 0\,,\ \text{for } k \geq q+1, \tag{8.26}$$

$$\varphi(B)\rho(k) \quad = \quad 0\ ,\ \text{for}\ k \geq q+1. \tag{8.27}$$

To compute the autocovariances, we can solve the equations (8.19) for $k = 0, 1, \ldots, p$, and then apply (8.21). The autocorrelations of an process ARMA(p, q) process behave like those of an AR(p) process, except that initial values are modified.

Example 8.1 Consider the ARMA(1, 1) model:

$$X_t = \bar{\mu} + \varphi_1 X_{t-1} + u_t - \theta_1 u_{t-1}\ ,\ |\varphi_1| < 1 \tag{8.28}$$

$$\tilde{X}_t - \varphi_1\, \tilde{X}_{t-1} = u_t - \theta_1 u_{t-1} \tag{8.29}$$

where $\tilde{X}_t = X_t - \mu$. We have

$$\begin{aligned}\gamma(0) &= \varphi_1\gamma(1) + \gamma_{xu}(0) - \theta_1\gamma_{xu}(-1), && (8.30)\\ \gamma(1) &= \varphi_1\gamma(0) + \gamma_{xu}(1) - \theta_1\gamma_{xu}(0) && (8.31)\end{aligned}$$

and

$$\begin{aligned}\gamma_{xu}(1) &= 0, && (8.32)\\ \gamma_{xu}(0) &= \sigma^2, && (8.33)\\ \gamma_{xu}(-1) &= \mathbb{E}(\tilde{X}_t u_{t-1}) = \varphi_1\mathbb{E}(\tilde{X}_{t-1}u_{t-1}) + \mathbb{E}(u_t u_{t-1}) - \theta_1\mathbb{E}(u_{t-1}^2)\\ &= \varphi_1\gamma_{xu}(0) - \theta_1\sigma^2 = (\varphi_1 - \theta_1)\sigma^2 && (8.34)\end{aligned}$$

Thus,

$$\begin{aligned}\gamma(0) &= \varphi_1\gamma(1) + \sigma^2 - \theta_1(\varphi_1 - \theta_1)\sigma^2\\ &= \varphi_1\gamma(1) + [1 - \theta_1(\varphi_1 - \theta_1)]\sigma^2, && (8.35)\end{aligned}$$

$$\begin{aligned}\gamma(1) &= \varphi_1\gamma(0) - \theta_1\sigma^2\\ &= \varphi_1\{\varphi_1\gamma(1) + [1 - \theta_1(\varphi_1 - \theta_1)]\sigma^2\} - \theta_1\sigma^2\ , && (8.36)\end{aligned}$$

hence

$$\begin{aligned}\gamma(1) &= \{\varphi_1[1 - \theta_1(\varphi_1 - \theta_1)] - \theta_1\}\sigma^2/(1 - \varphi_1^2)\\ &= \{\varphi_1 - \theta_1\varphi_1^2 + \varphi_1\theta_1^2 - \theta_1\}\sigma^2/(1 - \varphi_1^2)\\ &= (1 - \theta_1\varphi_1)(\varphi_1 - \theta_1)\sigma^2/(1 - \varphi_1^2)\,. && (8.37)\end{aligned}$$

Similarly,

$$\begin{aligned}\gamma(0) &= \varphi_1\gamma(1) + [1 - \theta_1(\varphi_1 - \theta_1)]\sigma^2\\ &= \varphi_1\frac{(1 - \theta_1\varphi_1)(\varphi_1 - \theta_1)\sigma^2}{1 - \varphi_1^2} + [1 - \theta_1(\varphi_1 - \theta_1)]\sigma^2\end{aligned}$$

$$
\begin{aligned}
&= \frac{\sigma^2}{1-\varphi_1^2}\{\varphi_1(1-\theta_1\varphi_1)(\varphi_1-\theta_1)+(1-\varphi_1^2)[1-\theta_1(\varphi_1-\theta_1)]\} \\
&= \frac{\sigma^2}{1-\varphi_1^2}\{\varphi_1^2-\theta_1\varphi_1^3+\varphi_1^2\theta_1^2-\varphi_1\theta_1+1-\varphi_1^2-\theta_1\varphi_1+\theta_1\varphi_1^3+\theta_1^2-\varphi_1^2\theta_1^2\} \\
&= \frac{\sigma^2}{1-\varphi_1^2}\{1-2\,\varphi_1\theta_1+\theta_1^2\}. \qquad (8.38)
\end{aligned}
$$

Thus,

$$
\begin{aligned}
\gamma(0) &= (1-2\,\varphi_1\theta_1+\theta_1^2)\sigma^2/(1-\varphi_1^2), \qquad (8.39)\\
\gamma(1) &= (1-\theta_1\varphi_1)(\varphi_1-\theta_1)\sigma^2/(1-\varphi_1^2), \qquad (8.40)\\
\gamma(k) &= \varphi_1\gamma(k-1),\ \text{for } k\geq 2. \qquad (8.41)
\end{aligned}
$$

9. Invertibility

A second-order stationary AR(p) process in MA(∞) form. Similarly, any second-order stationary ARMA(p,q) process can also be expressed as MA(∞) process. By analogy, it is natural to ask the question: can an MA(q) or ARMA(p,q) process be represented in a autoregressive form?

Consider the $MA(1)$ process

$$X_t = u_t - \theta_1 u_{t-1},\, t\in\mathbb{Z}\,, \qquad (9.1)$$

where $\{u_t : t\in\mathbb{Z}\}\sim \text{WN}(0,\,\sigma^2)$ and $\sigma^2>0$. We see easily that

$$
\begin{aligned}
u_t &= X_t+\theta_1u_{t-1}\\
&= X_t+\theta_1(X_{t-1}+\theta_1u_{t-2})\\
&= X_t+\theta_1X_{t-1}+\theta_1^2u_{t-2}\\
&= \sum_{j=0}^{n}\theta_1^jX_{t-j}+\theta_1^{n+1}u_{t-n-1} \qquad (9.2)
\end{aligned}
$$

and

$$\mathbb{E}\left[\left(\sum_{j=0}^{n}\theta_1^jX_{t-j}-u_t\right)^2\right]=\mathbb{E}\left[\left(\theta_1^{n+1}u_{t-n-1}\right)^2\right]=\theta_1^{2(n+1)}\sigma^2\underset{n\to\infty}{\to}0 \qquad (9.3)$$

provided $|\theta_1|<1$. Consequently, the series $\sum_{j=0}^{n}\theta_1^jX_{t-j}$ converges in $q.m.$ to u_t if $|\theta_1|<1$. In other words, when $|\theta_1|<1$, we can write

$$\sum_{j=0}^{\infty}\theta_1^jX_{t-j}=u_t,\, t\in\mathbb{Z}\,, \qquad (9.4)$$

or

$$(1-\theta_1 B)^{-1}X_t = u_t, t\in\mathbb{Z} \tag{9.5}$$

where $(1-\theta_1 B)^{-1} = \sum_{j=0}^{\infty}\theta_1^j B^j$. The condition $|\theta_1|<1$ is equivalent to having the roots of the equation $1-\theta_1 z = 0$ outside the unit circle. If $\theta_1 = 1$,

$$X_t = u_t - u_{t-1} \tag{9.6}$$

and the series

$$(1-\theta_1 B)^{-1}X_t = \sum_{j=0}^{\infty}\theta_1^j X_{t-j} = \sum_{j=0}^{\infty} X_{t-j} \tag{9.7}$$

does not converge, for $\mathbb{E}(X_{t-j}^2)$ does not converge to 0 as $j\to\infty$. Similarly, if $\theta_1 = -1$,

$$X_t = u_t + u_{t-1} \tag{9.8}$$

and the series

$$(1-\theta_1 B)^{-1}X_t = \sum_{j=0}^{\infty}(-1)^j X_{t-j} \tag{9.9}$$

does not converge either. These models are not invertible.

Theorem 9.1 INVERTIBILITY CONDITION FOR A MA PROCESS. *Let $\{X_t : t\in\mathbb{Z})$ be a second-order stationary process such that*

$$X_t = \mu + \theta(B)u_t \tag{9.10}$$

where $\theta(B) = 1-\theta_1 B - \cdots - \theta_q B^q$. Then the process X_t satisfies an equation of the form

$$\sum_{j=0}^{\infty}\bar{\phi}_j X_{t-j} = \bar{\mu} + u_t \tag{9.11}$$

iff the roots of the polynomial $\theta(z)$ are outside the unit circle. Further, when the representation (9.11) exists, we have:

$$\bar{\phi}(B) = \theta(B)^{-1}, \bar{\mu} = \theta(B)^{-1}\mu = \mu \Big/ \left(1-\sum_{j=1}^{q}\theta_j\right). \tag{9.12}$$

Corollary 9.2 INVERTIBILITY CONDITON FOR AN ARMA PROCESS. *Let $\{X_t : t\in\mathbb{Z}\}$ be a second-order stationary ARMA process that satisfies the equation*

$$\varphi(B)X_t = \bar{\mu} + \theta(B)u_t \tag{9.13}$$

where $\varphi(B) = 1-\varphi_1 B - \cdots - \varphi_p B^p$ and $\theta(B) = 1-\theta_1 B - \cdots - \theta_q B^q$. Then the process X_t satisfies an equation of the form

$$\sum_{j=0}^{\infty}\bar{\phi}_j X_{t-j} = \bar{\bar{\mu}} + u_t \tag{9.14}$$

iff the roots du polynomial $\theta(z)$ are outside the unit circle. Further, when the representation (9.14) exists, we have:

$$\bar{\phi}(B) = \theta(B)^{-1}\varphi(B)\,, \quad \bar{\bar{\mu}} = \theta(B)^{-1}\bar{\mu} = \mu \,/\left(1 - \sum_{j=1}^{q}\theta_j\right). \tag{9.15}$$

10. Wold representation

We have seen that all second-order ARMA processes can be written in a causal MA(∞) form. This property indeed holds for all second-order stationary processes.

Theorem 10.1 WOLD REPRESENTATION OF WEAKLY STATIONARY PROCESSES. *Let $\{X_t,\, t \in \mathbb{Z}\}$ be a second-order stationary process such that $\mathbb{E}(X_t) = \mu$. Then X_t can be written in the form*

$$X_t = \mu + \sum_{j=0}^{\infty}\psi_j u_{t-j} + v_t \tag{10.1}$$

where $\{u_t : t \in \mathbb{Z}\} \sim \mathrm{WN}(0,\ \sigma^2)$, $\sum_{j=0}^{\infty}\psi_j^2 < \infty$, $\mathbb{E}(u_t X_{t-j}) = 0,\ \forall j \geq 1,$ and $\{v_t : t \in \mathbb{Z}\}$ is a deterministic process such that $\mathbb{E}(v_t) = 0$ and $\mathbb{E}(u_s v_t) = 0,\ \forall s, t$. Further, if $\sigma^2 > 0$, the sequences $\{\psi_j\}$ and $\{u_t\}$ are unique, and

$$u_t = \tilde{X}_t - P(\tilde{X}_t | \tilde{X}_{t-1}, \tilde{X}_{t-2}, \ldots) \tag{10.2}$$

where $\tilde{X}_t = X_t - \mu$.

PROOF. See Anderson (1971, Section 7.6.3, pp. 420-421) and Hannan (1970, Chapter III, Section 2, Theorem 2, pp. 136-137). □

If $\mathbb{E}(u_t^2) > 0$ in Wold representation, we say the process X_t is *regular*. v_t is called the deterministic *component of* the process while $\sum_{j=0}^{\infty}\psi_j u_{t-j}$ is its *indeterministic component*. When $v_t = 0, \forall t$, the process X_t is said to be *strictly indeterministic*.

Corollary 10.2 FORWARD WOLD REPRESENTATION OF WEAKLY STATIONARY PROCESSES. *Let $\{X_t : t \in \mathbb{Z}\}$ be second-order a stationary process such that $\mathbb{E}(X_t) = \mu$. Then X_t can be written in the form*

$$X_t = \mu + \sum_{j=0}^{\infty}\bar{\psi}_j \bar{u}_{t+j} + \bar{v}_t \tag{10.3}$$

where $\{\bar{u}_t : t \in \mathbb{Z}\} \sim \mathrm{WN}(0,\ \bar{\sigma}^2)$, $\sum_{j=0}^{\infty}\bar{\psi}_j^2 < \infty$, $\mathbb{E}(\bar{u}_t X_{t+j}) = 0$, $\forall j \geq 1,$ and $\{\bar{v}_t : t \in \mathbb{Z}\}$ is a deterministic (with respect to $\bar{v}_{t+1},\ \bar{v}_{t+2}$, ...) such that $\mathbb{E}(\bar{v}_t) = 0$ and $\mathbb{E}(\bar{u}_s \bar{v}_t) = 0, \forall s,\, t$. Further, if

$\bar{\sigma}^2 > 0$, *the sequences* $\{\bar{\psi}_j\}$ *and* $\{\bar{u}_t\}$ *are uniquely defined, and*

$$\bar{u}_t = \tilde{X}_t - P(\tilde{X}_t | \tilde{X}_{t+1}, \tilde{X}_{t+2}, ...) \tag{10.4}$$

where $\tilde{X}_t = X_t - \mu$.

PROOF. The result follows on applying Wold theorem to the process $Y_t := X_{-t}$ which is also second-order stationary. □

References

ANDERSON, O. D. (1975): "On a Paper by Davies, Pete and Frost Concerning Maximum Autocorrelations for Moving Average Processes," *Australian Journal of Statistics*, 17, 87.

ANDERSON, T. W. (1971): *The Statistical Analysis of Time Series*. John Wiley & Sons, New York.

BOX, G. E. P., AND G. M. JENKINS (1976): *Time Series Analysis: Forecasting and Control*. Holden-Day, San Francisco, second edn.

BROCKWELL, P. J., AND R. A. DAVIS (1991): *Time Series: Theory and Methods*. Springer-Verlag, New York, second edn.

CHANDA, K. C. (1962): "On Bounds of Serial Correlations," *Annals of Mathematical Statistics*, 33, 1457.

DUFOUR, J.-M. (2016): "Notions of Stochastic Asymptotic Theory," Lecture notes, Department of Economics, McGill University, Montréal, Canada.

——— (2022): "Properties of Moments of Random Variables," Lecture notes, Department of Economics, McGill University, Montréal, Canada.

DURBIN, J. (1960): "Estimation of Parameters in Time Series Regression Models," *Journal of the Royal Statistical Society, Series A*, 22, 139–153.

GOURIÉROUX, C., AND A. MONFORT (1997): *Time Series and Dynamic Models*. Cambridge University Press, Cambridge, U.K.

HANNAN, E. J. (1970): *Multiple Time Series*. John Wiley & Sons, New York.

KENDALL, M., A. STUART, AND J. K. ORD (1983): *The Advanced Theory of Statistics. Volume 3: Design and Analysis and Time Series*. Macmillan, New York, fourth edn.

SPANOS, A. (1999): *Probability Theory and Statistical Inference: Econometric Modelling with Observational Data*. Cambridge University Press, Cambridge, UK.