1

A Brief Overview of Time Series and Stochastic Processes

This chapter serves as a brief introduction to time series to readers unfamiliar with this topic. Knowledgeable readers may want to jump directly to Chapter 2, where the basics of long-range dependence and self-similarity are introduced. A number of references for the material of this chapter can be found in Section 1.6, below.

1.1 Stochastic Processes and Time Series

A *stochastic processes* $\{X(t)\}_{t\in T}$ is a collection of random variables $X(t)$ on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, indexed by the time parameter $t \in T$. In "*discrete time*," we typically choose for T,

$$\mathbb{Z} = \{\dots, -1, 0, 1, \dots\},\ \mathbb{Z}_+ = \{0, 1, \dots\},\ \{1, 2, \dots, N\}, \dots,$$

and denote t by n. In "*continuous time*," we will often choose for T,

$$\mathbb{R},\ \mathbb{R}_+ = [0, \infty),\ [0, N], \dots .$$

In some instances in this volume, the parameter space T will be a subset of $\mathbb{R}^q$, $q \geq 1$, and/or $X(t)$ will be a vector with values in $\mathbb{R}^p$, $p \geq 1$. But for the sake of simplicity, we suppose in this chapter that p and q equal 1. We also suppose that $X(t)$ is real-valued.

One way to think of a stochastic process is through its law. *The law of a stochastic process* $\{X(t)\}_{t\in T}$ is characterized by its *finite-dimensional distributions* (fdd, in short); that is, the probability distributions

$$\mathbb{P}(X(t_1) \leq x_1, \dots, X(t_n) \leq x_n), \quad t_i \in T, x_i \in \mathbb{R}, n \geq 1,$$

of the random vectors

$$\big(X(t_1), \dots, X(t_n)\big)', \quad t_i \in T, n \geq 1.$$

Here, the prime indicates transpose, and all vectors are column vectors throughout. Thus, the finite-dimensional distributions of a stochastic process fully characterize its law and, in particular, the dependence structure of the stochastic process. In order to check that two stochastic processes have the same law, it is therefore sufficient to verify that their finite-dimensional distributions are identical. Equality and convergence in distribution is denoted by $\overset{d}{=}$ and $\overset{d}{\to}$ respectively. Thus $\{X(t)\}_{t\in T} \overset{d}{=} \{Y(t)\}_{t\in T}$ means

$$\mathbb{P}(X(t_1) \leq x_1, \dots, X(t_n) \leq x_n) = \mathbb{P}(Y(t_1) \leq x_1, \dots, Y(t_n) \leq x_n), \quad t_i \in T, x_i \in \mathbb{R}, n \geq 1.$$

https://doi.org/10.1017/CBO9781139600347.003 Published online by Cambridge University Press

A tilde $\sim$ indicates equality in distribution: for example, we write $X \sim \mathcal{N}(\mu, \sigma^2)$ if X is Gaussian with mean μ and variance σ^2.

A stochastic process is often called a *time series*, particularly when it is in discrete time and the focus is on its mean and covariance functions.

1.1.1 Gaussian Stochastic Processes

A stochastic process $\{X(t)\}_{t\in T}$ is *Gaussian* if one of the following equivalent conditions holds:

(i) The finite-dimensional distributions $Z = (X(t_1), \ldots, X(t_n))'$ are multivariate Gaussian $\mathcal{N}(b, A)$ with mean $b = \mathbb{E}Z$ and covariance matrix $A = \mathbb{E}(Z - \mathbb{E}Z)(Z - \mathbb{E}Z)'$;
(ii) $a_1 X(t_1) + \cdots + a_n X(t_n)$ is a Gaussian random variable for any $a_i \in \mathbb{R}$, $t_i \in T$;
(iii) In the case when $\mathbb{E}X(t) = 0$, for any $a_i \in \mathbb{R}$, $t_i \in T$,

$$\begin{aligned}\mathbb{E}\exp\{i(a_1X(t_1) + \cdots + a_nX(t_n))\} &= \exp\left\{-\frac{1}{2}\mathbb{E}(a_1X(t_1) + \cdots + a_nX(t_n))^2\right\} \\ &= \exp\left\{-\frac{1}{2}\sum_{i,j=1}^{n} a_i a_j\, \mathbb{E}X(t_i)X(t_j)\right\}. \quad (1.1.1)\end{aligned}$$

The law of a Gaussian stochastic process with zero mean is determined by a *covariance function* $\mathrm{Cov}(X(t), X(s)) = \mathbb{E}X(t)X(s)$, $s, t \in T$. When the mean is not zero, the covariance function is defined as $\mathrm{Cov}(X(t), X(s)) = \mathbb{E}\big(X(t) - \mathbb{E}X(t)\big)\big(X(s) - \mathbb{E}X(s)\big)$, $s, t \in T$. Together with the mean function $\mathbb{E}X(t)$, $t \in T$, the covariance determines the law of the Gaussian stochastic process.

Example 1.1.1 (*Brownian motion*) *Brownian motion* (or *Wiener process*) is a Gaussian stochastic process $\{X(t)\}_{t\geq 0}$ with[1]

$$\mathbb{E}X(t) = 0, \quad \mathbb{E}X(t)X(s) = \sigma^2 \min\{t, s\}, \quad t, s \geq 0,\ \sigma > 0, \qquad (1.1.2)$$

or, equivalently, it is a Gaussian stochastic process with independent increments $X(t_k) - X(t_{k-1})$, $k = 1, \ldots, n$, with $t_0 \leq t_1 \leq \cdots \leq t_n$ such that $X(t) - X(s) \sim \sigma\mathcal{N}(0, t - s)$, $t \geq s > 0$. Brownian motion is often denoted $B(t)$ or $W(t)$. Brownian motion $\{B(t)\}_{t\in\mathbb{R}}$ on the real line is defined as $B(t) = B_1(t)$, $t \geq 0$, and $B(t) = B_2(-t)$, $t < 0$, where B_1 and B_2 are two independent Brownian motions on the half line.

Example 1.1.2 (*Ornstein–Uhlenbeck process*) The *Ornstein–Uhlenbeck (OU) process* is a Gaussian stochastic process $\{X(t)\}_{t\in\mathbb{R}}$ with

$$\mathbb{E}X(t) = 0, \quad \mathbb{E}X(t)X(s) = \frac{\sigma^2}{2\lambda} e^{-\lambda(t-s)}, \quad t > s, \qquad (1.1.3)$$

[1] One imposes, sometimes, the additional condition that the process has continuous paths; but we consider here only the finite-dimensional distributions.

Figure 1.1 If the process has stationary increments, then, in particular, the increments taken over the bold intervals have the same distributions.

where $\lambda > 0, \sigma > 0$ are two parameters. The OU process is the only Gaussian stationary Markov process. It satisfies the Langevin stochastic differential equation

$$dX(t) = -\lambda X(t)dt + \sigma dW(t),$$

where $\{W(t)\}$ is a Wiener process. The term $-\lambda X(t)dt$ in the equation above adds a drift towards the origin.

1.1.2 Stationarity (of Increments)

A stochastic process $\{X(t)\}_{t\in T}$ is *(strictly) stationary* if $T = \mathbb{R}$ or $\mathbb{Z}$ or $\mathbb{R}_+$ or $\mathbb{Z}_+$, and for any $h \in T$,

$$\{X(t)\}_{t\in T} \stackrel{d}{=} \{X(t+h)\}_{t\in T}. \tag{1.1.4}$$

A stochastic process $\{X(t)\}_{t\in T}$ is said to have *(strictly) stationary increments* if $T = \mathbb{R}$ or $\mathbb{Z}$ or $\mathbb{R}_+$ or $\mathbb{Z}_+$, and for any $h \in T$,

$$\{X(t+h) - X(h)\}_{t\in T} \stackrel{d}{=} \{X(t) - X(0)\}_{t\in T}. \tag{1.1.5}$$

See Figure 1.1.

Example 1.1.3 (*The OU process*) The OU process in Example 1.1.2 is strictly stationary. Its finite-dimensional distributions are determined, for $t > s$, by

$$\mathbb{E}X(t)X(s) = \frac{\sigma^2}{2\lambda}e^{-\lambda(t-s)} = \frac{\sigma^2}{2\lambda}e^{-\lambda((t+h)-(s+h))} = \mathbb{E}X(t+h)X(s+h).$$

Thus, the law of the OU process is the same when shifted by $h \in \mathbb{R}$.

An example of a stochastic process with (strictly) stationary increments is Brownian motion in Example 1.1.1.

There is an obvious connection between stationarity and stationarity of increments. If $T = \mathbb{Z}$ and $\{X_t\}_{t\in\mathbb{Z}}$ has (strictly) stationary increments, then $\Delta X_t = X_t - X_{t-1}$, $t \in \mathbb{Z}$, is (strictly) stationary. Indeed, for any $h \in \mathbb{Z}$,

$$\begin{aligned}\{\Delta X_{t+h}\}_{t\in\mathbb{Z}} &= \{X_{t+h} - X_{t+h-1}\}_{t\in\mathbb{Z}} = \{X_{t+h} - X_h - (X_{t+h-1} - X_h)\}_{t\in\mathbb{Z}} \\ &\stackrel{d}{=} \{X_t - X_0 - (X_{t-1} - X_0)\}_{t\in\mathbb{Z}} = \{X_t - X_{t-1}\}_{t\in\mathbb{Z}} = \{\Delta X_t\}_{t\in\mathbb{Z}}.\end{aligned}$$

Conversely, if $\{Y_t\}_{t\in\mathbb{Z}}$ is (strictly) stationary, then $X_t = \sum_{k=1}^t Y_k$ can be seen easily to have (strictly) stationary increments.

If $T = \mathbb{R}$, then the *difference operator* Δ is replaced by the derivative when it exists and the sum is replaced by an integral.

https://doi.org/10.1017/CBO9781139600347.003 Published online by Cambridge University Press

1.1.3 Weak or Second-Order Stationarity (of Increments)

The probabilistic properties of (strictly) stationary processes do not change with time. In some circumstances, such as modeling, this is sometimes too strong a requirement. Instead of focusing on all probabilistic properties, one often requires instead that only second-order properties do not change with time. This leads to the following definition of (weak) stationarity.

A stochastic process $\{X(t)\}_{t\in T}$ is *(weakly or second-order) stationary* if $T = \mathbb{R}$ or $\mathbb{Z}$ and for any $t, s \in T$,

$$\mathbb{E}X(t) = \mathbb{E}X(0), \quad \mathrm{Cov}(X(t), X(s)) = \mathrm{Cov}(X(t-s), X(0)). \tag{1.1.6}$$

The time difference $t - s$ above is called the *time lag*. Weakly stationary processes are often called *time series*. Note that for Gaussian processes, weak stationarity is the same as strong stationarity. This, however, is not the case in general.

1.2 Time Domain Perspective

Consider a (weakly) stationary time series $X = \{X_n\}_{n\in\mathbb{Z}}$. In the time domain, one focuses on the functions

$$\gamma_X(h) = \mathrm{Cov}(X_h, X_0) = \mathrm{Cov}(X_{n+h}, X_n), \quad \rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)}, \quad h, n \in \mathbb{Z}, \tag{1.2.1}$$

called the *autocovariance function* (ACVF, in short) and *autocorrelation function* (ACF, in short) of the series X, respectively. ACVF and ACF are measures of dependence in time series. Sample counterparts of ACVF and ACF are the functions

$$\widehat{\gamma}_X(h) = \frac{1}{N}\sum_{n=1}^{N-|h|}(X_{n+|h|} - \overline{X})(X_n - \overline{X}), \quad \widehat{\rho}_X(h) = \frac{\widehat{\gamma}_X(h)}{\widehat{\gamma}_X(0)}, \quad |h| \le N-1,$$

where $\overline{X} = \frac{1}{N}\sum_{n=1}^{N} X_n$ is the sample mean. The following are basic properties of the ACF:

- *Symmetry:* $\rho_X(h) = \rho_X(-h)$, $h \in \mathbb{Z}$.
- *Range:* $|\rho_X(h)| \le 1$, $h \in \mathbb{Z}$.
- *Interpretation:* $\rho_X(h)$ close to -1, 0 and 1 correspond to strong negative, weak and strong positive correlations, respectively, of the time series X at lag h.

Statistical properties of the sample AVCF and ACF are delicate, for example, $\widehat{\rho}(h)$ and $\widehat{\rho}(h-1)$ have a nontrivial (asymptotic) dependence structure.

1.2.1 Representations in the Time Domain

"Representing" a stochastic process is expressing it in terms of simpler processes. The following examples involve *representations in the time domain.*

Example 1.2.1 (*White Noise*) A time series $X_n = Z_n$, $n \in \mathbb{Z}$, is called a *White Noise*, denoted $\{Z_n\} \sim \mathrm{WN}(0, \sigma_Z^2)$, if $\mathbb{E}Z_n = 0$ and

$$\gamma_Z(h) = \begin{cases} \sigma_Z^2, & h = 0, \\ 0, & h \neq 0, \end{cases} \qquad \rho_Z(h) = \begin{cases} 1, & h = 0, \\ 0, & h \neq 0. \end{cases}$$

Example 1.2.2 (*MA(1) series*) A time series $\{X_n\}_{n\in\mathbb{Z}}$ is called a *Moving Average of order one* (MA(1) for short) if it is given by

$$X_n = Z_n + \theta Z_{n-1}, \ n \in \mathbb{Z},$$

where $\{Z_n\} \sim \mathrm{WN}(0, \sigma_Z^2)$. Observe that

$$\gamma_X(h) = \mathbb{E}X_h X_0 = \mathbb{E}\big(Z_h + \theta Z_{h-1}\big)\big(Z_0 + \theta Z_{-1}\big) = \begin{cases} \sigma_Z^2(1+\theta^2), & h = 0, \\ \sigma_Z^2\theta, & h = 1, \\ 0, & h \geq 2, \end{cases}$$

and hence

$$\rho_X(h) = \begin{cases} 1, & h = 0, \\ \frac{\theta}{1+\theta^2}, & h = 1, \\ 0, & h \geq 2. \end{cases}$$

Example 1.2.3 (*AR*(1) *series*) A (weakly) stationary time series $\{X_n\}_{n\in\mathbb{Z}}$ is called *Autoregressive of order one* (AR(1) for short) if it satisfies the AR(1) equation

$$X_n = \varphi X_{n-1} + Z_n, \quad n \in \mathbb{Z},$$

where $\{Z_n\} \sim \mathrm{WN}(0, \sigma_Z^2)$. To see that AR(1) time series exists at least for some values of φ, suppose that $|\varphi| < 1$. Then, we expect that

$$\begin{aligned} X_n &= \varphi^2 X_{n-2} + \varphi Z_{n-1} + Z_n \\ &= \varphi^m X_{n-m} + \varphi^{m-1} Z_{n-(m-1)} + \cdots + Z_n = \sum_{m=0}^{\infty} \varphi^m Z_{n-m}. \end{aligned}$$

The time series $\{X_n\}_{n\in\mathbb{Z}}$ above is well-defined in the $L^2(\Omega)$–sense[2] because

$$\mathbb{E}\Big(\sum_{m=n_1}^{n_2} \varphi^m Z_{n-m}\Big)^2 = \sum_{m=n_1}^{n_2} \varphi^{2m}\sigma_Z^2 \to 0, \quad \text{as } n_1, n_2 \to \infty,$$

for $|\varphi| < 1$. One can easily see that it satisfies the AR(1) equation and is (weakly) stationary. Hence, the time series $\{X_n\}_{n\in\mathbb{Z}}$ is AR(1).

When $|\varphi| > 1$, AR(1) time series is obtained by reversing the AR(1) equation as

$$X_n = \varphi^{-1} X_{n+1} - \varphi^{-1} Z_{n+1}, \quad n \in \mathbb{Z},$$

and performing similar substitutions as above to obtain

[2] A random variable is well-defined in the $L^2(\Omega)$–sense if $\mathbb{E}|X|^2 < \infty$. A series $\sum_{n=1}^{\infty} X_n$ is well-defined in the $L^2(\Omega)$–sense if $\sum_{n=1}^{N} X_n$ converges in $L^2(\Omega)$–sense as $N \to \infty$, that is, $\mathbb{E}|\sum_{n=N_1}^{N_2} X_n|^2 \to 0$ as $N_1, N_2 \to \infty$.

https://doi.org/10.1017/CBO9781139600347.003 Published online by Cambridge University Press

$$X_n = \varphi^{-1}X_{n+1} - \varphi^{-1}Z_{n+1}$$
$$= \varphi^{-2}X_{n+2} - \varphi^{-2}Z_{n+2} - \varphi^{-1}Z_{n+1} = -\sum_{m=0}^{\infty} \varphi^{-(m+1)}Z_{n+m+1}.$$

When $|\varphi| = 1$, there is no (weakly) stationary solution to the AR(1) equation. When $\varphi = 1$, the AR(1) equation becomes $X_n - X_{n-1} = Z_n$ and the non-stationary (in fact, stationary increment) time series satisfying this equation is called *Integrated of order one* (I(1) for short). When Z_n are i.i.d., this time series is known as a *random walk*.

For $|\varphi| < 1$, for example, observe that for $h \geq 0$,

$$\gamma_X(h) = \mathbb{E}X_hX_0 = \mathbb{E}\big(Z_h + \cdots + \varphi^hZ_0 + \varphi^{h+1}Z_{-1} + \cdots\big)\big(Z_0 + \varphi Z_{-1} + \cdots\big)$$
$$= \sigma_Z^2(\varphi^h + \varphi^{h+2} + \varphi^{h+4} + \cdots) = \sigma_Z^2 \frac{\varphi^h}{1-\varphi^2}$$

and hence, since $\rho_X(-h) = \rho_X(h)$, we get for $h \in \mathbb{Z}$,

$$\rho_X(h) = \varphi^{|h|}.$$

Example 1.2.4 (*ARMA (p,q) series*) A (weakly) stationary time series $\{X_n\}_{n\in\mathbb{Z}}$ is called *Autoregressive moving average of orders p and q* (ARMA(p,q), for short) if it satisfies the equation

$$X_n - \varphi_1X_{n-1} - \cdots - \varphi_pX_{n-p} = Z_n + \theta_1Z_{n-1} + \cdots + \theta_qZ_{n-q},$$

where $\{Z_n\} \sim \mathrm{WN}(0, \sigma_Z^2)$.

ARMA(p,q) time series exists if the so-called characteristic polynomial $1 - \varphi_1z - \cdots - \varphi_pz^p = 0$ does not have root on the unit circle $\{z : |z| = 1\}$. This is consistent with the AR(1) equation discussed above where the root $z = 1/\varphi_1$ of the polynomial $1 - \varphi_1z = 0$ is on the unit circle when $|\varphi_1| = 1$.

Example 1.2.5 (*Linear time series*) A time series is called *linear* if it can be written as

$$X_n = \sum_{k=-\infty}^{\infty} a_kZ_{n-k}, \tag{1.2.2}$$

where $\{Z_n\} \sim \mathrm{WN}(0,\sigma_Z^2)$ and $\sum_{k=-\infty}^{\infty}|a_k|^2 < \infty$. Time series in Examples 1.2.1–1.2.3 are, in fact, linear. Observe that

$$\mathbb{E}X_{n+h}X_n = \mathbb{E}\Big(\sum_{k=-\infty}^{\infty} a_kZ_{n+h-k}\Big)\Big(\sum_{k=-\infty}^{\infty} a_kZ_{n-k}\Big)$$
$$= \mathbb{E}\Big(\sum_{k'=-\infty}^{\infty} a_{k'+h}Z_{n-k'}\Big)\Big(\sum_{k=-\infty}^{\infty} a_kZ_{n-k}\Big)$$
$$= \sigma_Z^2 \sum_{k=-\infty}^{\infty} a_{k+h}a_k = \sigma_Z^2 \sum_{k=-\infty}^{\infty} a_{h-k}a_{-k} = \sigma_Z^2(a * a^{\vee})_h,$$

where $*$ stands for the usual convolution, and $a^\vee$ denotes the time reversal of a. Since $\mathbb{E}X_{n+h}X_h$ depends only on h and $\mathbb{E}X_n = 0$, linear time series are (weakly) stationary. The variables Z_n entering the linear series (1.2.2) are known as innovations, especially when they are i.i.d.

Remark 1.2.6 Some of the notions above extend to continuous-time stationary processes $\{X(t)\}_{t\in\mathbb{R}}$. For example, such process is called linear when it can be represented as

$$X(t) = \int_{\mathbb{R}} a(t-u)Z(du), \quad t \in \mathbb{R}, \tag{1.2.3}$$

where $Z(du)$ is a real-valued random measure on $\mathbb{R}$ with orthogonal increments and control measure $\mathbb{E}(Z(du))^2 = du$ (see Appendix B.1, as well as Section 1.4 below), and $a \in L^2(\mathbb{R})$ is a deterministic function.

1.3 Spectral Domain Perspective

We continue considering (weakly) stationary time series $X = \{X_n\}_{n\in\mathbb{Z}}$. The material of this section is also related to the Fourier series and transform discussed in Appendix A.1.1.

1.3.1 Spectral Density

In the spectral domain, the focus is on the function

$$f_X(\lambda) = \frac{1}{2\pi}\sum_{h=-\infty}^{\infty} e^{-ih\lambda}\gamma_X(h), \quad \lambda \in (-\pi, \pi], \tag{1.3.1}$$

called the *spectral density* of the time series X. The variable λ is restricted to the domain $(-\pi, \pi]$ since the spectral density $f_X(\lambda)$ is 2π–periodic: that is, $f_X(\lambda + 2\pi k) = f_X(\lambda)$, $k \in \mathbb{Z}$. Observe also that the spectral density f_X is well-defined pointwise when $\gamma_X \in \ell^1(\mathbb{Z})$.

The variable λ enters $f_X(\lambda)$ through $e^{-ih\lambda}$, or sines and cosines. When λ is close to 0, we will talk about *low frequencies (long waves)*, and when λ is close to π, we will have *high frequencies (short waves)* in mind. Graphically, the association is illustrated in Figure 1.2.

Example 1.3.1 (*Spectral density of white noise*) If $\{Z_n\} \sim \mathrm{WN}(0, \sigma_Z^2)$, then

$$f_Z(\lambda) = \frac{\sigma_Z^2}{2\pi},$$

that is, the spectral density $f_Z(\lambda)$, $\lambda \in (-\pi, \pi]$, is constant.

Figure 1.2 Low (left) and high (right) frequencies.

Example 1.3.2 (*Spectral density of AR(1) series*) If $\{X_n\}_{n\in\mathbb{Z}}$ is AR(1) time series with $|\varphi| < 1$ and $\gamma_X(h) = \sigma_Z^2\varphi^{|h|}/(1-\varphi^2)$, then

$$f_X(\lambda) = \frac{\sigma_Z^2}{2\pi(1-\varphi^2)}\Big(1 + \sum_{h=1}^{\infty}(e^{-ih\lambda} + e^{ih\lambda})\varphi^h\Big)$$
$$= \frac{\sigma_Z^2}{2\pi(1-\varphi^2)}\Big(1 + \frac{\varphi e^{-i\lambda}}{1-\varphi e^{-i\lambda}} + \frac{\varphi e^{i\lambda}}{1-\varphi e^{i\lambda}}\Big) = \frac{\sigma_Z^2}{2\pi}\frac{1}{|1-\varphi e^{-i\lambda}|^2}.$$

The spectral density has the following properties:

- *Symmetry:* $f_X(\lambda) = f_X(-\lambda)$. This follows from $\gamma_X(h) = \gamma_X(-h)$. In particular, we can focus only on $\lambda \in [0,\pi]$.
- *Nonnegativeness:* $f_X(\lambda) \geq 0$. For $\gamma_X \in \ell^1(\mathbb{Z})$, this follows from

$$0 \leq \frac{1}{N}\mathbb{E}\Big|\sum_{r=1}^{N} X_r e^{-ir\lambda}\Big|^2 = \frac{1}{N}\mathbb{E}\Big(\sum_{r,s=1}^{N} X_r X_s e^{-i(r-s)\lambda}\Big)$$
$$= \frac{1}{N}\sum_{|h|<N}(N-|h|)\gamma_X(h)e^{-ih\lambda} \to \sum_{h=-\infty}^{\infty}\gamma_X(h)e^{-ih\lambda} = 2\pi f_X(\lambda),$$

  as $N \to \infty$, by using dominated convergence.
- *Inverse relation:*

$$\gamma_X(h) = \int_{-\pi}^{\pi} e^{ih\lambda} f_X(\lambda)d\lambda, \quad h \in \mathbb{Z}. \tag{1.3.2}$$

1.3.2 Linear Filtering

The following is a useful result. If

$$Y_n = \sum_{k=-\infty}^{\infty} a_k X_{n-k} \tag{1.3.3}$$

with a (weakly) stationary time series $\{X_n\}_{n\in\mathbb{Z}}$ and $a = \{a_k\}_{k\in\mathbb{Z}} \in \ell^1(\mathbb{Z})$, then

$$f_Y(\lambda) = |\widehat{a}(\lambda)|^2 f_X(\lambda), \tag{1.3.4}$$

where $\widehat{a}(\lambda) = \sum_{j=-\infty}^{\infty} a_j e^{-ij\lambda}$ is the Fourier transform of a (see Appendix A.1.1). This follows from observing that

$$\gamma_Y(h) = \mathbb{E}Y_hY_0 = \sum_{j,k=-\infty}^{\infty} a_j a_k \gamma_X(h+k-j)$$
$$= \sum_{j,k} a_j a_k \int_{-\pi}^{\pi} e^{i(h+k-j)\lambda} f_X(\lambda)d\lambda = \int_{-\pi}^{\pi} e^{ih\lambda}\Big|\sum_{j=-\infty}^{\infty} a_j e^{-ij\lambda}\Big|^2 f_X(\lambda)d\lambda.$$

Relation (1.3.3) transforms the series $\{X_n\}$ into the series $\{Y_n\}$ by means of a linear filter.

Example 1.3.3 (*Spectral density of AR*(1) *series, cont'd*) Applying (1.3.3)–(1.3.4) to the AR(1) equation $X_n - \varphi X_{n-1} = Z_n$ with $\{Z_n\} \sim \mathrm{WN}(0, 1)$ yields

$$|1 - \varphi e^{-i\lambda}|^2 f_X(\lambda) = f_Z(\lambda) = \frac{\sigma_Z^2}{2\pi}$$

or

$$f_X(\lambda) = \frac{\sigma_Z^2}{2\pi} \frac{1}{|1 - \varphi e^{-i\lambda}|^2},$$

the result obtained directly in Example 1.3.2.

1.3.3 Periodogram

A sample counterpart to the spectral density is defined by

$$\frac{1}{2\pi} \sum_{|h|<N} \widehat{\gamma}_X(h) e^{-ih\lambda} = \frac{1}{2\pi N} \Big| \sum_{n=1}^{N} X_n e^{-in\lambda} \Big|^2 =: \frac{I_X(\lambda)}{2\pi}, \tag{1.3.5}$$

with the first relation holding only at the so-called *Fourier frequencies*

$$\lambda = \lambda_k = \frac{2\pi k}{N} \text{with } k = -\Big[\frac{N-1}{2}\Big], \dots, \Big[\frac{N}{2}\Big]. \tag{1.3.6}$$

$I_X(\lambda)$ is known as the *periodogram*, and has the following properties:

- *Computational speed:* $I_X(\lambda_k)$ can be computed efficiently by Fast Fourier Transform (FFT) in $O(N \log N)$ steps, supposing N can be factored out in many factors.
- *Statistical properties:* $I_X(\lambda)$ is not a consistent estimator for $2\pi f_X(\lambda)$, but is asymptotically unbiased. The periodogram needs to be smoothed to become consistent.

Warning: Two definitions of the periodogram I_X are commonly found in the literature. One definition appears in (1.3.5). The other popular definition is to set the whole left-hand side of (1.3.5) for the periodogram; that is, to incorporate the denominator 2π into the periodogram. Since the two definitions are different, it is important to check which convention is used in a given source. With the definition (1.3.5), we follow the convention used in Brockwell and Davis [186].

1.3.4 Spectral Representation

A (weakly) stationary, zero mean time series $X = \{X_n\}_{n\in\mathbb{Z}}$ with a spectral density f_X can be represented as

$$X_n = \int_{-\pi}^{\pi} e^{in\lambda} Z_X(d\lambda), \tag{1.3.7}$$

where $Z_X(d\lambda)$ is a complex-valued *random measure* such that $Z_X(-d\lambda) = \overline{Z_X(d\lambda)}$; that is, Z_X is Hermitian. Moreover,

$$\mathbb{E} Z_X(d\lambda) \overline{Z_X(d\lambda')} = 0 \tag{1.3.8}$$

https://doi.org/10.1017/CBO9781139600347.003 Published online by Cambridge University Press

when $d\lambda \neq d\lambda'$ (i.e., having orthogonal increments), and

$$\mathbb{E}|Z_X(d\lambda)|^2 = f_X(\lambda)d\lambda. \qquad (1.3.9)$$

As with most integrals, this is interpreted through discrete sums as

$$X_n \approx \sum_k e^{in\lambda_k} Z_X(d\lambda_k),$$

where $Z_X(d\lambda_k)$ are uncorrelated with variances $f_X(\lambda_k)d\lambda_k$. See Appendix B.1 for a more rigorous treatment. Thus, at frequency λ with larger value of $f_X(\lambda)$, the variance of the random coefficient at $e^{in\lambda}$ is larger. These terms dominate in the representation.

Remark 1.3.4 In writing the spectral representation (1.3.7) we assumed implicitly that the series X has a spectral density. Spectral representations, however, exist for all (weakly) stationary time series. They are written more generally as

$$X_n = \int_{(-\pi,\pi]} e^{in\lambda} Z_X(d\lambda),$$

where $Z_X(d\lambda)$ is a complex-valued random measure as above with the only difference that the property (1.3.9) is replaced by

$$\mathbb{E}|Z_X(d\lambda)|^2 = F_X(d\lambda)$$

for the so-called spectral measure F_X on $(-\pi, \pi]$. When the spectral measure F_X has a density f_X (with respect to the Lebesgue measure), f_X is the spectral density of the series X and the relation (1.3.9) holds.

Example 1.3.5 (*Spectral density of AR*(1) *series, cont'd*) The spectral density of AR(1) series was derived in Examples 1.3.2 and 1.3.3. Typical plots of AR(1) time series and their

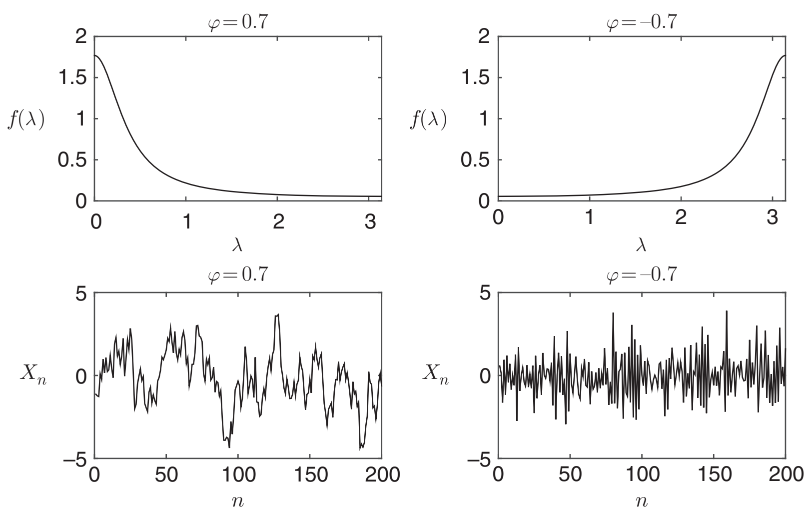

Figure 1.3 Typical plots of AR(1) spectral densities and of their sample paths in the cases $\varphi > 0$ and $\varphi < 0$.

https://doi.org/10.1017/CBO9781139600347.003 Published online by Cambridge University Press

spectral densities in the cases $\varphi > 0$ and $\varphi < 0$ are given in Figure 1.3. These plots are consistent with the idea behind spectral representation described above.

Remark 1.3.6 If $\{X_n\}_{n\in\mathbb{Z}}$ is given by its spectral representation, then by using (1.3.8) and (1.3.9),

$$\gamma_X(h) = \mathbb{E}X_h X_0 = \mathbb{E}\int_{-\pi}^{\pi} e^{ih\lambda} Z_X(d\lambda)\overline{\int_{-\pi}^{\pi} e^{i0\lambda'} Z_X(d\lambda')} = \int_{-\pi}^{\pi} e^{ih\lambda} f_X(\lambda) d\lambda,$$

which is the relation (1.3.2) connecting the spectral density to the ACVF.

Remark 1.3.7 Suppose that $f_X(\lambda) = |g_X(\lambda)|^2$ with $\overline{g_X(\lambda)} = g_X(-\lambda)$, which happens in many examples. Then,

$$X_n = \int_{-\pi}^{\pi} e^{in\lambda} Z_X(d\lambda) = \int_{-\pi}^{\pi} e^{in\lambda} g_X(\lambda)\widetilde{Z}(d\lambda),$$

where the random measure $\widetilde{Z}(d\lambda)$ satisfies

$$\mathbb{E}\widetilde{Z}(d\lambda)\overline{\widetilde{Z}(d\lambda')} = \begin{cases} d\lambda, & d\lambda = d\lambda', \\ 0, & \text{otherwise.} \end{cases}$$

Remark 1.3.8 Many notions above extend to continuous-time stationary processes $\{X(t)\}_{t\in\mathbb{R}}$. The spectral density of such a process is defined as

$$f_X(\lambda) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-i\lambda h}\gamma_X(h) dh, \quad \lambda \in \mathbb{R},$$

with the difference from (1.3.1) that it is a function for $\lambda \in \mathbb{R}$, and the sum is replaced by an integral. The inverse relation is

$$\gamma_X(h) = \int_{\mathbb{R}} e^{ih\lambda} f_X(\lambda) d\lambda, \quad h \in \mathbb{R},$$

(cf. (1.3.2)). The spectral representation reads

$$X(t) = \int_{\mathbb{R}} e^{it\lambda} Z_X(d\lambda), \quad t \in \mathbb{R},$$

where $Z_X(d\lambda)$ is a complex-valued random measure on $\mathbb{R}$ with orthogonal increments and control measure $\mathbb{E}|Z_X(d\lambda)|^2 = f_X(\lambda)d\lambda$.

1.4 Integral Representations Heuristics

The spectral representation (1.3.7) has components which are summarized through the first two columns of Table 1.1; that is, the dependence structure of X_n is transferred into the deterministic functions $e^{in\lambda}$. But one can also think of more general (integral) representations

$$X(t) = \int_E h_t(u) M(du), \tag{1.4.1}$$

where the components appearing in the last column of Table 1.1 are interpreted similarly.

https://doi.org/10.1017/CBO9781139600347.003 Published online by Cambridge University Press

Table 1.1 *Components in representations of time series and stochastic processes.*

Components in representation	Time series	Stochastic process
Dependent time series (process)	X_n	$X(t)$
Underlying space	$(-\pi, \pi]$	E
Deterministic functions	$e^{in\lambda}$	$h_t(u)$
Uncorrelated (or independent) random measure	$Z_X(d\lambda)$	$M(du)$

Example 1.4.1 (*Linear time series*) The linear time series in (1.2.2) is in fact defined through an integral representation since

$$X_n = \sum_{k=-\infty}^{\infty} a_{n-k} Z_k = \int_{\mathbb{Z}} h_n(k) M(dk) = \int_{\mathbb{Z}} h(n-k) M(dk),$$

where $h_n(k) = a_{n-k}$, $h(k) = a_k$ and $M(\{k\}) = Z_k$ are uncorrelated.

Various random measures and integral representations are defined and discussed in Appendix B.2. The following example provides a heuristic explanation of Gaussian random measures and their integrals.

Example 1.4.2 (*Gaussian random measure*) Suppose $E = \mathbb{R}$ and set $M(du) = B(du)$, viewing $B(du)$ as the increment on an infinitesimal interval du of a standard Brownian motion $\{B(u)\}_{u\in\mathbb{R}}$. Since Brownian motion has stationary and independent increments, and $\mathbb{E}B^2(u) = |u|$, one can think of the random measure $B(du)$ as satisfying

$$\mathbb{E}B(du_1)B(du_2) = \begin{cases} 0, & \text{if } du_1 \neq du_2, \\ du, & \text{if } du_1 = du_2 = du. \end{cases}$$

Thus, $B(du) \sim \mathcal{N}(0, du)$. The nonrandom measure $m(du) = \mathbb{E}B^2(du)$ is called the *control measure*. Here m is the Lebesgue control measure since $m(du) = du$. In the integral, $\int_{\mathbb{R}} h(u)B(du)$, each $B(du)$ is weighted by the nonrandom factor $h(u)$, and since the $B(du)$s are independent on disjoint intervals, one expects that

$$\int_{\mathbb{R}} h(u)B(du) \sim \mathcal{N}\Big(0, \int_{\mathbb{R}} h^2(u)du\Big).$$

Formally, the integral $I(h) = \int_E h(u)M(du)$ is defined first for simple functions h and then, by approximation, for all functions satisfying $\int_E h^2(u)du < \infty$.

1.4.1 Representations of a Gaussian Continuous-Time Process

Let $h \in L^2(\mathbb{R}, du)$; that is, $\int_{\mathbb{R}} h^2(u)du < \infty$. The Fourier transform of h is $\widehat{h}(x) = \int_{\mathbb{R}} e^{iux} h(x)du$ with the inverse formula $h(u) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-iux}\widehat{h}(x)dx$ (see Appendix A.1.2). It is complex-valued and Hermitian; that is, $\overline{\widehat{h}(dx)} = \widehat{h}(-dx)$. Introduce a similar transformation on $B(du)$, namely, let $\overline{\widehat{B}(dx)} = \widehat{B}(-dx)$ be complex-valued, with $\widehat{B}(dx) =$

$B_1(dx) + iB_2(dx)$, where $B_1(dx)$ and $B_2(dx)$ are real-valued, independent $\mathcal{N}(0, dx/2)$, and require $\widehat{B}$ to be Hermitian; that is, $\overline{\widehat{B}(dx)} = \widehat{B}(-dx)$. Then, $\mathbb{E}|\widehat{B}(dx)|^2 = dx$ and

$$I(h) := \int_{\mathbb{R}} h(u)B(du) \stackrel{d}{=} \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \widehat{h}(x)\widehat{B}(dx) =: \widehat{I}(\widehat{h}).$$

See Appendix B.1 for more details.

Example 1.4.3 (*The OU process*) Consider a stochastic process

$$X(t) = \sigma \int_{-\infty}^{t} e^{-(t-u)\lambda} B(du), \quad t \in \mathbb{R},$$

where $\sigma > 0$, $\lambda > 0$ are parameters, and $B(du)$ is a Gaussian random measure on $\mathbb{R}$ with the Lebesgue control measure du. The process $X(t)$ is Gaussian with zero mean and covariance function (for $t > s > 0$)

$$\mathbb{E}X(t)X(s) = \sigma^2 \int_{-\infty}^{s} e^{-(t-u)\lambda} e^{-(s-u)\lambda} du$$
$$= \sigma^2 e^{-(t+s)\lambda} \int_{-\infty}^{s} e^{2u\lambda} du = \frac{\sigma^2}{2\lambda} e^{-(t-s)\lambda}.$$

The process $X(t)$ is thus the OU process (see Example 1.1.2).

The integral representation above is in the time domain; that is, $I(h) = \int_{\mathbb{R}} h_t(u)B(du)$ with $h_t(u) = \sigma e^{-(t-u)\lambda} 1_{\{u<t\}}$. Observe that

$$\widehat{h_t}(x) = \int_{\mathbb{R}} e^{ixu} h_t(u)du = \sigma \int_{-\infty}^{t} e^{ixu} e^{-(t-u)\lambda} du$$
$$= \sigma e^{-t\lambda} \int_{-\infty}^{t} e^{(ix+\lambda)u} du = \frac{\sigma e^{-t\lambda} e^{(ix+\lambda)u}}{ix+\lambda}\Big|_{u=-\infty}^{t} = \frac{e^{ixt}\sigma}{ix+\lambda}.$$

Hence, by switching to the spectral domain, the OU process can be represented as

$$X(t) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{ixt} \frac{\sigma}{ix+\lambda} \widehat{B}(dx).$$

Example 1.4.4 (*Brownian motion*) Brownian motion can be represented as

$$B(t) = \int_{\mathbb{R}} 1_{[0,t)}(u)B(du) = \int_{\mathbb{R}} \Big(1_{(0,\infty)}(t-u) - 1_{(0,\infty)}(-u)\Big) B(du),$$

where $B(du)$ is a Gaussian random measure with the control measure du. With $h_t(u) = 1_{[0,t)}(u)$,

$$\widehat{h_t}(x) = \int_{\mathbb{R}} e^{ixu} h_t(u)du = \int_{0}^{t} e^{ixu} du = \frac{e^{ixt}-1}{ix}.$$

Then, switching to the spectral domain, Brownian motion can also be represented as

$$B(t) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \frac{e^{ixt}-1}{ix} \widehat{B}(dx).$$

A reader wishing to learn more about time series analysis and stochastic processes could consult the references given in Section 1.6 below. These references are helpful in understanding better the material presented in subsequent chapters.

1.5 A Heuristic Overview of the Next Chapter

In the next chapter, we introduce basic concepts and results involving long-range dependence and self-similarity. Because the precise definitions can be rather technical, we provide here a brief and heuristic overview.

There are several definitions of long-range dependence which are, in general, not equivalent. Basically, a stationary series $\{X_n\}_{n\in\mathbb{Z}}$ is long-range dependent if its autocovariance function $\gamma_X(k) = \mathbb{E}X_kX_0 - \mathbb{E}X_k\mathbb{E}X_0$ behaves like k^{2d-1} as $k \to \infty$, where $0 < d < 1/2$. This range of d ensures that $\sum_{k=-\infty}^{\infty} \gamma_X(k) = \infty$. From a spectral domain perspective, the spectral density $f_X(\lambda)$ of $\{X_n\}_{n\in\mathbb{Z}}$ behaves as λ^{-2d} as the frequency $\lambda \to 0$. Since $d > 0$, note that the spectral density diverges as $\lambda \to 0$. A typical example is FARIMA$(0, d, 0)$ series introduced in Section 2.4.1.

We also define the related notion of self-similarity. A process $\{Y(t)\}_{t\in\mathbb{R}}$ is H–self-similar if, for any constant c, the finite-dimensional distributions of $\{Y(ct\}_{t\in\mathbb{R}}$ are the same as those of $\{c^H Y(t)\}_{t\in\mathbb{R}}$, where H is a parameter often related to d. In fact, if the process $\{Y(t)\}$ has stationary increments, then $X_n = Y(n) - Y(n-1)$, $n \in \mathbb{Z}$, has long-range dependence with $H = d + 1/2$. Conversely, we can obtain $Y(t)$ from X_n by using a limit theorem.

Fractional Brownian motion is a typical example of $Y(t)$. It is Gaussian, H–self-similar, and has stationary increments. We provide both time-domain and spectral-domain representations for fractional Brownian motion. We also give additional examples of non-Gaussian self-similar processes, such as the Rosenblatt process and also processes with infinite variance defined through their integral representations, for instance, linear fractional stable motion and the Telecom process.

1.6 Bibliographical Notes

There are a number of excellent textbooks on time series and their analysis. The monograph by Brockwell and Davis [186] provides a solid theoretical foundation. The classic by Priestley [833] has served generations of scientists interested in the spectral analysis of time series. For more applied and computational aspects of the time series analysis, see Cryer and Chan [271], Shumway and Stoffer [909]. Nonlinear time series are treated in Douc, Moulines, and Stoffer [327].

On the side of stochastic processes, Lindgren [635] provides an up-to-date treatment of stationary stochastic processes. The basics of Brownian motion and related stochastic calculus are treated in Karatzas and Shreve [549], Mörters and Peres [733]. A number of other facts used in this monograph are discussed in Appendices B and C.