11 : CHAPTER

Stationary Stochastic Processes

A sequence is a function mapping from a set of integers, described as the index set, onto the real line or into a subset thereof. A time series is a sequence whose index corresponds to consecutive dates separated by a unit time interval.

In the statistical analysis of time series, the elements of the sequence are regarded as a set of random variables. Usually, no notational distinction is made between these random variables and their realised values. It is important nevertheless to bear the distinction in mind.

In order to analyse a statistical time series, it must be assumed that the structure of the statistical or stochastic process which generates the observations is essentially invariant through time. The conventional assumptions are summarised in the condition of stationarity, which is discussed in section 2.2. The condition has immediate implications for the autocovariance and the autocorrelation functions defined in section 2.3.

A fundamental process, from which many other stationary processes may be derived, is the so-called white-noise process which consists of a sequence of uncorrelated random variables, each with a zero mean and the same finite variance. By passing white noise through a linear filter, a sequence whose elements are serially correlated can be generated. The filtering operation is defined in section 2.4.

In fact, virtually every stationary stochastic process may be depicted as the product of a filtering operation applied to white noise. We demonstrate this result in section 2.10, after we have introduced the concepts underlying the spectral representation of a time series.

The spectral representation is rooted in the basic notion of Fourier analysis which is that well-behaved functions can be approximated over a finite interval, to any degree of accuracy, by a weighted combination of sine and cosine functions whose harmonically rising frequencies are integral multiples of a fundamental frequency. Such linear combinations are described as Fourier sums or Fourier series. Of course, the notion applies to sequences as well; for any number of well-behaved functions may be interpolated through the coordinates

of a finite sequence.

In section 2.6, we provide the Fourier representation of a finite sequence which is exact. This is extended, in section 2.7, to provide a representation of an infinite sequence in terms of an infinity of trigonometrical functions whose frequencies range continuously in the interval $[0, \pi]$. The trigonometrical functions and their weighting functions are gathered under a Fourier–Stieltjes integral. It is remarkable that, whereas a Fourier sum serves only to define a strictly periodic function, a Fourier integral suffices to represent an aperiodic time series generated by a stationary stochastic process.

The Fourier integral is also used to represent the underlying stochastic process. This is achieved by describing the stochastic processes which generate the weighting functions. There are two such weighting processes, associated respectively with the sine and cosine functions; and their common variance, which is a function $f(\omega)$, $\omega \in [0, \pi]$, is the so-called spectral density function.

The relationship between the spectral density function and the sequence of autocovariances, which is summarised in the Wiener –Khintchine theorem, provides a link between the time-domain and the frequency-domain analyses. In section 2.8, we show how the autocovariances may be obtained from the Fourier transform of the spectral density function. We demonstrate, at the end of section 2.10, that the spectral density function is, conversely, a Fourier transform of the sequence of autocovariances.

Finally, section 2.12 defines the periodogram as a cosine Fourier transform of the sequence of empirical autocovariances. The periodogram provides a basis for estimating the spectral density function.

2.2 Stationarity

Consider two vectors of $n + 1$ consecutive elements from the process $y(t)$:

$$\begin{aligned} y_t &= [y_t, y_{t+1}, \ldots, y_{t+n}]', \\ y_{t+k} &= [y_{t+k}, y_{t+k+1}, \ldots, y_{t+k+n}]'. \end{aligned} \tag{1}$$

Then $y(t)$ is strictly stationary if the joint probability density functions of the vectors y_t and y_{t+k} are the same for any value of k regardless of the size of n. On the assumption that the first and second-order moments of the distribution are finite, the condition of stationarity implies that all the elements of $y(t)$ have the same expected value and that the covariance between any pair of elements of the sequences is a function only of their temporal separation. Thus,

$$\begin{aligned} E(y_t) &= \mu, \\ C(y_{t+i}, y_{t+j}) &= C(y_i, y_j) \\ &= \gamma_{|i-j|}. \end{aligned} \tag{2}$$

On their own, the conditions of (2) constitute the conditions of weak stationarity.

A normal process is completely characterised by its mean and its autocovariances. Therefore, a normal process $y(t)$ which satisfies the conditions for weak stationarity is also stationary in the strict sense.

2.3 Autocovariance and autocorrelation functions

The covariance between two elements y_t and $y_{t+\tau}$ of a process $y(t)$ which are separated by τ intervals of time, is known as the autocovariance at lag τ and is denoted by γ_τ. The autocorrelation at lag τ, denoted by ρ_τ, is defined by

$$\rho_\tau = \frac{\gamma_\tau}{\gamma_0,} \tag{3}$$

where γ_0 is the variance of the process $y(t)$.

The stationarity conditions imply that the autocovariances of $y(t)$ satisfy the equality

$$\gamma_\tau = \gamma_{-\tau} \tag{4}$$

for all values of τ.

The autocovariance matrix of a stationary process corresponding to the n elements $y_0, y_1, \ldots, y_{n-1}$ is given by

$$\Gamma = \begin{bmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \cdots & \gamma_{n-1} \\ \gamma_1 & \gamma_0 & \gamma_1 & \cdots & \gamma_{n-2} \\ \gamma_2 & \gamma_1 & \gamma_0 & \cdots & \gamma_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \gamma_{n-1} & \gamma_{n-2} & \gamma_{n-3} & \cdots & \gamma_0 \end{bmatrix}. \tag{5}$$

The sequences $\{\gamma_\tau\}$ and $\{\rho_\tau\}$ are described as the autocovariance and autocorrelation functions respectively.

2.4 Filtering white noise

A white-noise process is a sequence $\varepsilon(t)$ of uncorrelated random variables with mean zero and common variance σ_ε^2. Thus

$$\begin{aligned} E(\varepsilon_t) &= 0, \quad \text{for all } t \\ E(\varepsilon_{t+i}\varepsilon_{t+j}) &= \begin{cases} \sigma_\varepsilon^2, & \text{if } i = j; \\ 0, & \text{if } i \neq j. \end{cases} \end{aligned} \tag{6}$$

By a process of linear filtering, a variety of time series may be constructed whose elements display complex interdependencies. A finite linear filter, also called a moving-average operator, is a polynomial in the lag operator of the

form $\mu(L) = \mu_0 + \mu_1 L + \cdots + \mu_q L^q$. The effect of this filter on $\varepsilon(t)$ is described by the equation

$$\begin{aligned} y(t) &= \mu(L)\varepsilon(t) \\ &= \mu_0 + \mu_1\varepsilon(t-1) + \mu_2\varepsilon(t-2) + \cdots + \mu_q\varepsilon(t-q) \\ &= \sum_{i=0}^{q} \mu_i\varepsilon(t-i). \end{aligned} \tag{7}$$

The operator $\mu(L)$ is also be described as the transfer function which maps the input sequence $\varepsilon(t)$ into the output sequence $y(t)$. Figure 2.1 gives a pictorial representation.

An operator with an indefinite number of terms in rising powers of L may also be considered: $\mu(L) = \{\mu_0 + \mu_1 L + \mu_2 L^2 + \cdots\}$. However, for this to be practical, the coefficients $\{\mu_0, \mu_1, \mu_2, \ldots\}$ must be functions of a limited number of fundamental parameters. In addition, it is required that

$$\sum_i |\mu_i| < \infty. \tag{8}$$

Given the value of $\sigma_\varepsilon^2 = V\{\varepsilon(t)\}$, the autocovariances of the filtered sequence $y(t) = \mu(L)\varepsilon(t)$ may be determined by evaluating the expression

$$\begin{aligned} \gamma_\tau &= E(y_t y_{t-\tau}) \\ &= E\Big(\sum_i \mu_i\varepsilon_{t-i} \sum_j \mu_j\varepsilon_{t-\tau-j}\Big) \\ &= \sum_i \sum_j \mu_i\mu_j E(\varepsilon_{t-i}\varepsilon_{t-\tau-j}). \end{aligned} \tag{9}$$

From equation (6), it follows that

$$\gamma_\tau = \sigma_\varepsilon^2 \sum_j \mu_j\mu_{j+\tau}; \tag{10}$$

and so the variance of the filtered sequence is

$$\gamma_0 = \sigma_\varepsilon^2 \sum_j \mu_j^2. \tag{11}$$

The condition under equation (8) guarantees that these quantities are finite, as is required by the condition of stationarity.

2.5 The z-transform

In the subsequent analysis, it will prove helpful to present the results in the notation of the z-transform. The z-transform of the infinite sequence $y(t) = \{y_t; t = 0, \pm 1, \pm 2, \ldots\}$ is defined by

$$y(z) = \sum_{\tau=-\infty}^{\infty} y_t z^t. \tag{12}$$

Here z is usually taken to be a complex number on the perimeter of the unit circle. Thus $z = e^{-i\omega}$ with $\omega \in [0, 2\pi]$

If $y(t) = \mu(L)\varepsilon(t)$ is a moving-average process, then the z-transform is given by $y(z) = \mu(z)\varepsilon(z)$ where $\mu(z) = \{\mu_0 + \mu_1 z + \mu_2 z^2 + \cdots\}$ has the same form as the operator $\mu(L)$, and where $\varepsilon(z)$ is the z-transform of the white-noise sequence.

The z-transform of a sequence of autocovariances is called the autocovariance generating function. For the moving-average process, this is given by

$$\begin{aligned}
\gamma(z) &= \sigma_\varepsilon^2 \mu(z)\mu(z^{-1}) \\
&= \sigma_\varepsilon^2 \sum_i \mu_i z^i \sum_j \mu_j z^{-j} \\
&= \sigma_\varepsilon^2 \sum_i \sum_j \mu_i \mu_j z^{i-j} \\
&= \sum_\tau \left\{ \sigma_\varepsilon^2 \sum_j \mu_j \mu_{j+\tau} \right\} z^\tau \quad ; \quad \tau = i - j \\
&= \sum_{\tau=-\infty}^{\infty} \gamma_\tau z^\tau.
\end{aligned} \tag{13}$$

The final equality is by virtue of equation (10).

2.6 The Fourier representation of a sequence

According to the basic result of Fourier analysis, it is always possible to approximate an arbitrary analytic function defined over a finite interval of the real line, to any desired degree of accuracy, by a weighted sum of sine and cosine functions of harmonically increasing frequencies.

Similar results apply in the case of sequences, which may be regarded as functions mapping from the set of integers onto the real line. For a sample of T observations $y_0, \ldots, y_{T-1}$, it is possible to devise an expression in the form

$$y_t = \sum_{j=0}^{n} \big\{ \alpha_j \cos(\omega_j t) + \beta_j \sin(\omega_j t) \big\}, \tag{14}$$

wherein $\omega_j = 2\pi j/T$ is a multiple of the fundamental frequency $\omega_1 = 2\pi/T$. Thus, the elements of a finite sequence can be expressed exactly in terms of sines and cosines. This expression is called the Fourier decomposition of y_t and the set of coefficients $\{\alpha_j, \beta_j; j = 0, 1, \ldots, n\}$ are called the Fourier coefficients.

When T is even, we have $n = T/2$; and it follows that

$$\begin{aligned} \sin(\omega_0 t) &= \sin(0) = 0 \\ \cos(\omega_0 t) &= \cos(0) = 1 \\ \sin(\omega_n t) &= \sin(\pi t) = 0 \\ \cos(\omega_n t) &= \cos(\pi t) = (-1)^t. \end{aligned} \tag{15}$$

Therefore, equation (14) becomes

$$y_t = \alpha_0 + \sum_{j=1}^{n-1} \big\{\alpha_j \cos(\omega_j t) + \beta_j \sin(\omega_j t)\big\} + \alpha_n (-1)^t. \tag{16}$$

When T is odd, we have $n = (T-1)/2$; and then equation (14) becomes

$$y_t = \alpha_0 + \sum_{j=1}^{n} \big\{\alpha_j \cos(\omega_j t) + \beta_j \sin(\omega_j t)\big\}. \tag{17}$$

In both cases, there are T nonzero coefficients amongst the set $\{\alpha_j, \beta_j; j = 0, 1, \ldots, n\}$; and the mapping from the sample values to the coefficients constitutes a one-to-one invertible transformation.

In equation (16), the frequencies of the trigonometric functions range from $\omega_1 = 2\pi/T$ to $\omega_n = \pi$; whereas, in equation (17), they range from $\omega_1 = 2\pi/T$ to $\omega_n = \pi(T-1)/T$. The frequency π is the so-called Nyquist frequency.

Although the process generating the data may contain components of frequencies higher than the Nyquist frequency, these will not be detected when it is sampled regularly at unit intervals of time. In fact, the effects on the process of components with frequencies in excess of the Nyquist value will be confounded with those whose frequencies fall below it.

To demonstrate this, consider the case where the process contains a component which is a pure cosine wave of unit amplitude and zero phase whose frequency ω lies in the interval $\pi < \omega < 2\pi$. Let $\omega^* = 2\pi - \omega$. Then

$$\begin{aligned} \cos(\omega t) &= \cos\{(2\pi - \omega^*)t\} \\ &= \cos(2\pi)\cos(\omega^* t) + \sin(2\pi)\sin(\omega^* t) \\ &= \cos(\omega^* t); \end{aligned} \tag{18}$$

which indicates that ω and ω^* are observationally indistinguishable. Here, $\omega^* < \pi$ is described as the alias of $\omega > \pi$.

2.7 The spectral representation of a stationary process

By allowing the value of n in the expression (14) to tend to infinity, it is possible to express a sequence of indefinite length in terms of a sum of sine and cosine functions. However, in the limit as $n \to \infty$, the coefficients α_j, β_j tend to vanish; and therefore an alternative representation in terms of differentials is called for.

By writing $\alpha_j = dA(\omega_j)$, $\beta_j = dB(\omega_j)$ where $A(\omega)$, $B(\omega)$ are step functions with discontinuities at the points $\{\omega_j; j = 0, \ldots, n\}$, the expression (14) can be rendered as

$$y_t = \sum_j \big\{ \cos(\omega_j t)dA(\omega_j) + \sin(\omega_j t)dB(\omega_j) \big\}. \tag{19}$$

In the limit, as $n \to \infty$, the summation is replaced by an integral to give the expression

$$y(t) = \int_0^\pi \big\{ \cos(\omega t)dA(\omega) + \sin(\omega t)dB(\omega) \big\}. \tag{20}$$

Here, $\cos(\omega t)$ and $\sin(\omega t)$, and therefore $y(t)$, may be regarded as infinite sequences defined over the entire set of positive and negative integers.

Since $A(\omega)$ and $B(\omega)$ are discontinuous functions for which no derivatives exist, one must avoid using $\alpha(\omega)d\omega$ and $\beta(\omega)d\omega$ in place of $dA(\omega)$ and $dB(\omega)$. Moreover, the integral in equation (20) is a Fourier–Stieltjes integral.

In order to derive a statistical theory for the process that generates $y(t)$, one must make some assumptions concerning the functions $A(\omega)$ and $B(\omega)$. So far, the sequence $y(t)$ has been interpreted as a realisation of a stochastic process. If $y(t)$ is regarded as the stochastic process itself, then the functions $A(\omega)$, $B(\omega)$ must, likewise, be regarded as stochastic processes defined over the interval $[0, \pi]$. A single realisation of these processes now corresponds to a single realisation of the process $y(t)$.

The first assumption to be made is that the functions $A(\omega)$ and $B(\omega)$ represent a pair of stochastic processes of zero mean which are indexed on the continuous parameter ω. Thus

$$E\{dA(\omega)\} = E\{dB(\omega)\} = 0. \tag{21}$$

The second and third assumptions are that the two processes are mutually uncorrelated and that non-overlapping increments within each process are uncorrelated. Thus

$$\begin{aligned} E\{dA(\omega)dB(\lambda)\} &= 0 \quad \text{for all} \quad \omega, \lambda, \\ E\{dA(\omega)dA(\lambda)\} &= 0 \quad \text{if} \quad \omega \neq \lambda, \\ E\{dB(\omega)dB(\lambda)\} &= 0 \quad \text{if} \quad \omega \neq \lambda. \end{aligned} \tag{22}$$

The final assumption is that the variance of the increments is given by

$$\begin{aligned} V\{dA(\omega)\} = V\{dB(\omega)\} &= 2dF(\omega) \\ &= 2f(\omega)d\omega. \end{aligned} \tag{23}$$

We can see that, unlike $A(\omega)$ and $B(\omega)$, $F(\omega)$ is a continuous differentiable function. The function $F(\omega)$ and its derivative $f(\omega)$ are the spectral distribution function and the spectral density function, respectively.

In order to express equation (20) in terms of complex exponentials, we may define a pair of conjugate complex stochastic processes:

$$\begin{aligned} dZ(\omega) &= \frac{1}{2}\{dA(\omega) - idB(\omega)\}, \\ dZ^*(\omega) &= \frac{1}{2}\{dA(\omega) + idB(\omega)\}. \end{aligned} \tag{24}$$

Also, we may extend the domain of the functions $A(\omega)$, $B(\omega)$ from $[0, \pi]$ to $[-\pi, \pi]$ by regarding $A(\omega)$ as an even function such that $A(-\omega) = A(\omega)$ and by regarding $B(\omega)$ as an odd function such that $B(-\omega) = -B(\omega)$. Then we have

$$dZ^*(\omega) = dZ(-\omega). \tag{25}$$

From conditions under (22), it follows that

$$\begin{aligned} E\{dZ(\omega)dZ^*(\lambda)\} &= 0 \quad \text{if} \quad \omega \neq \lambda, \\ E\{dZ(\omega)dZ^*(\omega)\} &= f(\omega)d\omega. \end{aligned} \tag{26}$$

These results may be used to reexpress equation (20) as

$$\begin{aligned} y(t) &= \int_0^\pi \left\{ \frac{(e^{i\omega t} + e^{-i\omega t})}{2} dA(\omega) - i\frac{(e^{i\omega t} - e^{-i\omega t})}{2} dB(\omega) \right\} \\ &= \int_0^\pi \left\{ e^{i\omega t} \frac{\{dA(\omega) - idB(\omega)\}}{2} + e^{-i\omega t} \frac{\{dA(\omega) + idB(\omega)\}}{2} \right\} \\ &= \int_0^\pi \left\{ e^{i\omega t} dZ(\omega) + e^{-i\omega t} dZ^*(\omega) \right\}. \end{aligned} \tag{27}$$

When the integral is extended over the range $[-\pi, \pi]$, this becomes

$$y(t) = \int_{-\pi}^{\pi} e^{i\omega t} dZ_y(\omega) \tag{28}$$

This is commonly described as the spectral representation of the process $y(t)$.

2.8 The autocovariances and the spectral density function

The sequence of the autocovariances of the process $y(t)$ may be expressed in terms of the spectrum of the process. From equation (28), it follows that the autocovariance y_t at lag $\tau = t - k$ is given by

$$\begin{aligned}\gamma_\tau = C(y_t, y_k) &= E\left\{\int_\omega e^{i\omega t} dZ_y(\omega) \int_\lambda e^{-i\lambda k} dZ_y(-\lambda)\right\} \\ &= \int_\omega \int_\lambda e^{i\omega t} e^{-i\lambda k} E\{dZ_y(\omega) dZ_y^*(\lambda)\} \\ &= \int_\omega e^{i\omega\tau} E\{dZ_y(\omega) dZ_y^*(\omega)\} \\ &= \int_\omega e^{i\omega\tau} f_y(\omega) d\omega.\end{aligned} \tag{29}$$

Here the final equalities are derived by using the results (24) and (25). This equation indicates that the Fourier transform of the spectrum is the autocovariance function.

Setting $\tau = 0$ in equation (29) gives

$$\gamma_0 = \int_\omega f_y(\omega) d\omega, \tag{30}$$

which is the variance of $y(t)$. This equation shows how the variance or 'power'of the process is distributed amongst the cyclical components of which it is composed.

It is easy to see that a flat spectrum corresponds to the autocovariance function which characterises a white-noise process $\varepsilon(t)$. Let $f_\varepsilon = f_\varepsilon(\omega)$ be the flat spectrum. Then, from equation (30), it follows that

$$\begin{aligned}\gamma_0 &= \int_{-\pi}^{\pi} f_\varepsilon(\omega) d\omega \\ &= 2\pi f_\varepsilon,\end{aligned} \tag{31}$$

and, from equation (29), it follows that

$$\begin{aligned}\gamma_\tau &= \int_{-\pi}^{\pi} f_\varepsilon(\omega) e^{i\omega\tau} d\omega \\ &= f_\varepsilon \int_{-\pi}^{\pi} e^{i\omega\tau} d\omega \\ &= 0.\end{aligned} \tag{32}$$

These are the same as the conditions under (6) which have served to define a white-noise process. When the variance is denoted by σ_ε^2, the expression for the spectrum of the white-noise process becomes

$$f_\varepsilon(\omega) = \frac{\sigma_\varepsilon^2}{2\pi}. \tag{33}$$

2.9 Canonical factorisation of the spectral density function

Let $y(t)$ be a stationary stochastic process whose spectrum is $f_y(\omega)$. Since $f_y(\omega) \geq 0$, it is always possible to find a complex function $\mu(\omega)$ such that

$$f_y(\omega) = \frac{1}{2\pi}\mu(\omega)\mu^*(\omega). \tag{34}$$

For a wide class of stochastic processes, the function $\mu(\omega)$ may be constructed in such a way that it can be expanded as a one-sided Fourier series:

$$\mu(\omega) = \sum_{j=0}^{\infty} \mu_j e^{-i\omega j}. \tag{35}$$

On defining

$$dZ_\varepsilon(\omega) = \frac{dZ_y(\omega)}{\mu(\omega)}, \tag{36}$$

the spectral representation of the process $y(t)$ given in equation (27), may be rewritten as

$$y(t) = \int_\omega e^{i\omega t}\mu(\omega)dZ_\varepsilon(\omega). \tag{37}$$

Expanding the expression of $\mu(\omega)$ and interchanging the order of integration and summation gives

$$\begin{aligned} y(t) &= \int_\omega e^{i\omega t}\Big(\sum_j \mu_j e^{-i\omega j}\Big)dZ_\varepsilon(\omega) \\ &= \sum_j \mu_j \Big\{\int_\omega e^{i\omega(t-j)}dZ_\varepsilon(\omega)\Big\} \\ &= \sum_j \mu_j \varepsilon(t-j), \end{aligned} \tag{38}$$

where we have defined

$$\varepsilon(t) = \int_\omega e^{i\omega t}dZ_\varepsilon(\omega). \tag{39}$$

The spectrum of $\varepsilon(t)$ is given by

$$\begin{aligned} E\{dZ_\varepsilon(\omega)dZ_\varepsilon^*(\omega)\} &= E\left\{\frac{dZ_y(\omega)dZ_y^*(\omega)}{\mu(\omega)\mu^*(\omega)}\right\} \\ &= \frac{f_y(\omega)}{\mu(\omega)\mu^*(\omega)} \\ &= \frac{1}{2\pi}. \end{aligned} \tag{40}$$

Hence $\varepsilon(t)$ is identified as a white-noise process with unit variance. Therefore equation (38) represents a moving-average process; and what our analysis implies is that virtually every stationary stochastic process can be represented in this way.

2.10 Filtering in the frequency domain

It is a straightforward matter to derive the spectrum of a process $y(t) = \mu(L)x(t)$ which is formed by mapping the process $x(t)$ through a linear filter.

Taking the spectral representation of the process $x(t)$ to be

$$x(t) = \int_\omega e^{i\omega t} dZ_x(\omega), \tag{41}$$

we have

$$\begin{aligned} y(t) &= \sum_j \mu_j x(t-j) \\ &= \sum_j \mu_j \left\{ \int_\omega e^{i\omega(t-j)} dZ_x(\omega) \right\} \\ &= \int_\omega e^{i\omega t} \left(\sum_j \mu_j e^{-i\omega j} \right) dZ_x(\omega). \end{aligned} \tag{42}$$

On writing $\sum \mu_j e^{-i\omega j} = \mu(\omega)$, this becomes

$$\begin{aligned} y(t) &= \int_\omega e^{i\omega t} \mu(\omega) dZ_x(\omega) \\ &= \int_\omega e^{i\omega t} dZ_y(\omega). \end{aligned} \tag{43}$$

It follows that the spectral density function $f_y(\omega)$ of the filtered process $y(t)$ is given by

$$\begin{aligned} f_y(\omega)d\omega &= E\{dZ_y(\omega)dZ_y^*(\omega)\} \\ &= \mu(\omega)\mu^*(\omega)E\{dZ_x(\omega)dZ_x^*(\omega)\} \\ &= |\mu(\omega)|^2 f_x(\omega)d\omega. \end{aligned} \tag{44}$$

In the case of the process defined in equation (7), where $y(t)$ is obtained by filtering a white-noise sequence, the result is specialised to give

$$\begin{aligned} f_y(\omega) &= |\mu(\omega)|^2 f_\varepsilon(\omega) \\ &= \frac{\sigma_\varepsilon^2}{2\pi}|\mu(\omega)|^2. \end{aligned} \tag{45}$$

Let $\mu(z) = \sum \mu_j z^j$ denote the z-transform of the sequence $\{\mu_j\}$. Then

$$\begin{aligned} |\mu(z)|^2 &= \mu(z)\mu(z^{-1}) \\ &= \sum_\tau \sum_j \mu_j \mu_{j+\tau} z^\tau. \end{aligned} \tag{46}$$

It follows that, when $z = e^{-i\omega}$, equation (45) can be written as

$$\begin{aligned} f_y(\omega) &= \frac{\sigma_\varepsilon^2}{2\pi}\mu(z)\mu(z^{-1}) \\ &= \frac{1}{2\pi}\sum_\tau \left\{ \sigma_\varepsilon^2 \sum_j \mu_j \mu_{j+\tau} \right\} z^\tau. \end{aligned} \tag{47}$$

But, according to equation (10), $\gamma_\tau = \sigma_\varepsilon^2 \sum_j \mu_j \mu_{j+\tau}$ is the autocovariance of lag τ of the process $y(t)$. Therefore, the function $f_y(\omega)$ can be written as

$$\begin{aligned} f_y(\omega) &= \frac{1}{2\pi}\sum_{\tau=-\infty}^{\infty} e^{-i\omega\tau}\gamma_\tau \\ &= \frac{1}{2\pi}\left\{ \gamma_0 + 2\sum_{\tau=1}^{\infty} \gamma_\tau \cos(\omega\tau) \right\}, \end{aligned} \tag{48}$$

which indicates that the spectral density function is the Fourier transform of the autocovariance function of the filtered sequence. This is known as the Wiener–Khintchine theorem. The importance of this theorem is that it provides a link between the time domain and the frequency domain.

2.11 The Gain and Phase

Reference to Appendix A1.3 shows that the complex-valued function $\mu(\omega)$, which is entailed in the process of linear filtering, can be written as

$$\mu(\omega) = |\mu(\omega)|e^{-i\theta(\omega)}. \tag{49}$$

where

$$\begin{aligned} |\mu(\omega)|^2 &= \left\{ \sum_{j=0}^{\infty} \mu_j \cos(\omega j) \right\}^2 + \left\{ \sum_{j=0}^{\infty} \mu_j \sin(\omega j) \right\}^2 \\ \theta(\omega) &= \arctan\left\{ \frac{\sum \mu_j \sin(\omega j)}{\sum \mu_j \cos(\omega j)} \right\}. \end{aligned} \tag{50}$$

The function $|\mu(\omega)|$, which is described as the gain of the filter, indicates the extent to which the amplitude of the cyclical components of which $x(t)$ is composed are altered in the process of filtering.

The function $\theta(\omega)$, which is described as the phase displacement and which gives a measure in radians, indicates the extent to which the cyclical components are displaced along the time axis.

The substitution of expression (49) in equation (43) gives

$$y(t) = \int_{-\pi}^{\pi} e^{i\{\omega t - \theta(\omega)\}} |\mu(\omega)| dZ_x(\omega). \tag{51}$$

The importance of this equation is that it summarises the two effects of the filter.

2.12 Periodogram

Given a set of T observations $y_0, y_1, \ldots, y_{T-1}$ of the process $y(t)$ one may define a sequence of empirical autocovariances $c_0, c_1, \ldots, c_{T-1}$ whose generic element is

$$c_\tau = \frac{1}{T} \sum_{t=\tau}^{T-1} (y_{t-\tau} - \bar{y})(y_t - \bar{y}). \tag{52}$$

If the process $y(t)$ is stationary and if $\lim(\tau \to \infty)\gamma_\tau = 0$, then the empirical autocovariances represent consistent estimates of the corresponding autocovariances $\gamma_0, \gamma_1, \ldots, \gamma_{T-1}$ of the process. Therefore, it would appear that the spectral density function of the process which is given by equation (48) should be estimated by a sample spectrum in the form of

$$\hat{f}(\omega) = \frac{1}{2\pi}\left\{ c_0 + 2 \sum_{\tau=1}^{T-1} c_\tau \cos(\omega\tau) \right\}. \tag{53}$$

The periodogram $I(\omega_j)$ of the sample is defined as the function

$$I(\omega_j) = \frac{T}{2}\{\alpha_j^2 + \beta_j^2\}, \tag{54}$$

where α_j and β_j are the Fourier coefficients defined in section 2.5. Given that these coefficients are

$$\alpha_j = (2/T) \sum_t y_t \cos(\omega_j t) \quad \text{and} \quad \beta_j = (2/T) \sum_t y_t \sin(\omega_j t),$$

it follows that

$$\begin{aligned} I(\omega_j) &= \frac{2}{T}\left[\left\{ \sum_{t=0}^{T-1} y_t \cos(\omega_j t) \right\}^2 + \left\{ \sum_{t=0}^{T-1} y_t \sin(\omega_j t) \right\}^2 \right] \\ &= \frac{2}{T}\left[\left\{ \sum_{t=0}^{T-1} (y_t - \bar{y}) \cos(\omega_j t) \right\}^2 + \left\{ \sum_{t=0}^{T-1} (y_t - \bar{y}) \sin(\omega_j t) \right\}^2 \right]. \end{aligned} \tag{55}$$

The second equality follows from the fact that, by construction, $\sum \cos(\omega_j t) = 0$ for all j. Expanding the expression gives

$$\begin{aligned} I(\omega_j) =& \frac{2}{T}\left\{\sum_t\sum_s \cos(\omega_j t)\cos(\omega_j s)(y_t - \bar{y})(y_s - \bar{y})\right\} \\ &+ \frac{2}{T}\left\{\sum_t\sum_s \sin(\omega_j t)\sin(\omega_j s)(y_t - \bar{y})(y_s - \bar{y})\right\}; \end{aligned} \tag{56}$$

and, by using the identity $\cos(A)\cos(B) + \sin(A)\sin(B) = \cos(A - B)$, we can rewrite this as

$$I(\omega_j) = \frac{2}{T}\left\{\sum_t\sum_s \cos(\omega_j[t - s])(y_t - \bar{y})(y_s - \bar{y})\right\}. \tag{57}$$

Next, on defining $\tau = t - s$ and writing $c_\tau = \sum_t (y_t - \bar{y})(y_{t-\tau} - \bar{y})/T$, we can reduce the latter expression to

$$I(\omega_j) = 2\left\{c_0 + 2\sum_{t=1}^{T-1} c_\tau \cos(\omega_j \tau)\right\}, \tag{58}$$

and it is clear that the sample spectrum is a scaled version of the periodogram.

In fact, the sample spectrum, as it is defined in equation (53), is not an adequate estimator of the spectral density function. To understand this, one must note that the periodogram has a number of ordinates equal to half the number of elements in the sample. Therefore, the periodogram inherits much of the volatility which one witnesses in the sequence of sample values $y_0, \ldots, y_{T-1}$; and this feature may be in distinct contrast to the smoothness of the spectral density function.

To obtain a reasonable estimate of the spectrum from the periodogram or the sample spectrum, it is necessary to smooth the periodogram by aggregateing adjacent ordinates. Various smoothing operators are available which may be applied directly to the ordinates of the periodogram. However, the same effect may be achieved by applying a differential weighting scheme to the empirical autocovariances which enter the sample spectrum. Thus, a class of spectral estimators may be defined which take the form of

$$f^w(\omega) = \frac{1}{2\pi}\left\{w_0 c_0 + 2\sum_{\tau=1}^{R} w_\tau c_\tau \cos(\omega\tau)\right\}, \tag{59}$$

where $\{w_0 c_0, w_1 c_1, \ldots, w_R c_R\}$ is the sequence of weighted autocovariances. Amongst such weighting scheme is Parzen's window defined by

$$\begin{aligned} w_\tau &= 1 - 6(\tau/R)^2 + 6(\tau/R)^3 \quad \text{for} \quad 0 \le \tau \le R/2, \\ w_\tau &= 2\{1 - (\tau/R)\}^3 \quad \text{for} \quad R/2 \le \tau \le R. \end{aligned} \tag{60}$$

This scheme takes only the autocovariances up to lag R and discards the remainder. The degree of smoothing is related inversely to the value of R.

Parzen's window is employed in subsequent chapters where the smoothed spectra obtained by weighting the autocovariance function are compared with parametric spectra derived by fitting autoregressive and moving average models.