Assuming that these autocovariances are absolutely summable, the autocovariance-generating function is given by

$$g_Y(z) = \sum_{j=-\infty}^{\infty} \gamma_j z^j, \qquad [6.1.1]$$

where z denotes a complex scalar. If [6.1.1] is divided by 2π and evaluated at some z represented by $z = e^{-i\omega}$ for $i = \sqrt{-1}$ and ω a real scalar, the result is called the population spectrum of Y:

$$s_Y(\omega) = \frac{1}{2\pi} g_Y(e^{-i\omega}) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \gamma_j e^{-i\omega j}. \qquad [6.1.2]$$

Note that the spectrum is a function of ω: given any particular value of ω and a sequence of autocovariances $\{\gamma_j\}_{j=-\infty}^{\infty}$, we could in principle calculate the value of $s_Y(\omega)$.

De Moivre's theorem allows us to write $e^{-i\omega j}$ as

$$e^{-i\omega j} = \cos(\omega j) - i\cdot\sin(\omega j). \qquad [6.1.3]$$

Substituting [6.1.3] into [6.1.2], it appears that the spectrum can equivalently be written

$$s_Y(\omega) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \gamma_j[\cos(\omega j) - i\cdot\sin(\omega j)]. \qquad [6.1.4]$$

Note that for a covariance-stationary process, $\gamma_j = \gamma_{-j}$. Hence, [6.1.4] implies

$$\begin{aligned} s_Y(\omega) &= \frac{1}{2\pi}\gamma_0[\cos(0) - i\cdot\sin(0)] \\ &\quad + \frac{1}{2\pi}\left\{\sum_{j=1}^{\infty} \gamma_j[\cos(\omega j) + \cos(-\omega j) - i\cdot\sin(\omega j) - i\cdot\sin(-\omega j)]\right\}. \end{aligned} \qquad [6.1.5]$$

Next, we make use of the following results from trigonometry:[1]

$$\begin{aligned} \cos(0) &= 1 \\ \sin(0) &= 0 \\ \sin(-\theta) &= -\sin(\theta) \\ \cos(-\theta) &= \cos(\theta). \end{aligned}$$

Using these relations, [6.1.5] simplifies to

$$s_Y(\omega) = \frac{1}{2\pi}\left\{\gamma_0 + 2\sum_{j=1}^{\infty} \gamma_j \cos(\omega j)\right\}. \qquad [6.1.6]$$

Assuming that the sequence of autocovariances $\{\gamma_j\}_{j=-\infty}^{\infty}$ is absolutely summable, expression [6.1.6] implies that the population spectrum exists and that $s_Y(\omega)$ is a continuous, real-valued function of ω. It is possible to go a bit further and show that if the γ_j's represent autocovariances of a covariance-stationary process, then $s_Y(\omega)$ will be nonnegative for all ω.[2] Since $\cos(\omega j) = \cos(-\omega j)$ for any ω, the spectrum is symmetric around $\omega = 0$. Finally, since $\cos[(\omega + 2\pi k)\cdot j] = \cos(\omega j)$ for any integers k and j, it follows from [6.1.6] that $s_Y(\omega + 2\pi k) = s_Y(\omega)$ for any integer k. Hence, the spectrum is a periodic function of ω. If we know the value of $s_Y(\omega)$ for all ω between 0 and π, we can infer the value of $s_Y(\omega)$ for any ω.

[1]These are reviewed in Section A.1 of the Mathematical Review (Appendix A) at the end of the book.

[2]See, for example, Fuller (1976, p. 110).

Calculating the Population Spectrum for Various Processes

Let Y_t follow an $MA(\infty)$ process:

$$Y_t = \mu + \psi(L)\varepsilon_t, \tag{6.1.7}$$

where

$$\psi(L) = \sum_{j=0}^{\infty} \psi_j L^j$$

$$\sum_{j=0}^{\infty} |\psi_j| < \infty$$

$$E(\varepsilon_t\varepsilon_\tau) = \begin{cases} \sigma^2 & \text{for } t = \tau \\ 0 & \text{otherwise.} \end{cases}$$

Recall from expression [3.6.8] that the autocovariance-generating function for Y is given by

$$g_Y(z) = \sigma^2\psi(z)\psi(z^{-1}).$$

Hence, from [6.1.2], the population spectrum for an $MA(\infty)$ process is given by

$$s_Y(\omega) = (2\pi)^{-1}\cdot\sigma^2\psi(e^{-i\omega})\psi(e^{i\omega}). \tag{6.1.8}$$

For example, for a white noise process, $\psi(z) = 1$ and the population spectrum is a constant for all ω:

$$s_Y(\omega) = \sigma^2/2\pi. \tag{6.1.9}$$

Next, consider an $MA(1)$ process:

$$Y_t = \varepsilon_t + \theta\varepsilon_{t-1}.$$

Here, $\psi(z) = 1 + \theta z$ and the population spectrum is

$$\begin{aligned} s_Y(\omega) &= (2\pi)^{-1}\cdot\sigma^2(1 + \theta e^{-i\omega})(1 + \theta e^{i\omega}) \\ &= (2\pi)^{-1}\cdot\sigma^2(1 + \theta e^{-i\omega} + \theta e^{i\omega} + \theta^2). \end{aligned} \tag{6.1.10}$$

But notice that

$$e^{-i\omega} + e^{i\omega} = \cos(\omega) - i\cdot\sin(\omega) + \cos(\omega) + i\cdot\sin(\omega) = 2\cdot\cos(\omega), \tag{6.1.11}$$

so that [6.1.10] becomes

$$s_Y(\omega) = (2\pi)^{-1}\cdot\sigma^2[1 + \theta^2 + 2\theta\cdot\cos(\omega)]. \tag{6.1.12}$$

Recall that $\cos(\omega)$ goes from 1 to -1 as ω goes from 0 to π. Hence, when $\theta > 0$, the spectrum $s_Y(\omega)$ is a monotonically decreasing function of ω for ω in $[0, \pi]$, whereas when $\theta < 0$, the spectrum is monotonically increasing.

For an $AR(1)$ process

$$Y_t = c + \phi Y_{t-1} + \varepsilon_t,$$

we have $\psi(z) = 1/(1 - \phi z)$ as long as $|\phi| < 1$. Thus, the spectrum is

$$\begin{aligned} s_Y(\omega) &= \frac{1}{2\pi}\frac{\sigma^2}{(1 - \phi e^{-i\omega})(1 - \phi e^{i\omega})} \\ &= \frac{1}{2\pi}\frac{\sigma^2}{(1 - \phi e^{-i\omega} - \phi e^{i\omega} + \phi^2)} \\ &= \frac{1}{2\pi}\frac{\sigma^2}{[1 + \phi^2 - 2\phi\cdot\cos(\omega)]}. \end{aligned} \tag{6.1.13}$$

When $\phi > 0$, the denominator is monotonically increasing in ω over $[0, \pi]$, meaning that $s_Y(\omega)$ is monotonically decreasing. When $\phi < 0$, the spectrum $s_Y(\omega)$ is a monotonically increasing function of ω.

In general, for an $ARMA(p, q)$ process

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q},$$

the population spectrum is given by

$$s_Y(\omega) = \frac{\sigma^2}{2\pi} \frac{(1 + \theta_1 e^{-i\omega} + \theta_2 e^{-i2\omega} + \cdots + \theta_q e^{-iq\omega})}{(1 - \phi_1 e^{-i\omega} - \phi_2 e^{-i2\omega} - \cdots - \phi_p e^{-ip\omega})} \times \frac{(1 + \theta_1 e^{i\omega} + \theta_2 e^{i2\omega} + \cdots + \theta_q e^{iq\omega})}{(1 - \phi_1 e^{i\omega} - \phi_2 e^{i2\omega} - \cdots - \phi_p e^{ip\omega})}. \qquad [6.1.14]$$

If the moving average and autoregressive polynomials are factored as follows:

$$1 + \theta_1 z + \theta_2 z^2 + \cdots + \theta_q z^q = (1 - \eta_1 z)(1 - \eta_2 z) \cdots (1 - \eta_q z)$$
$$1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p = (1 - \lambda_1 z)(1 - \lambda_2 z) \cdots (1 - \lambda_p z),$$

then the spectral density in [6.1.14] can be written

$$s_Y(\omega) = \frac{\sigma^2 \prod_{j=1}^{q} [1 + \eta_j^2 - 2\eta_j \cdot \cos(\omega)]}{2\pi \prod_{j=1}^{p} [1 + \lambda_j^2 - 2\lambda_j \cdot \cos(\omega)]}.$$

Calculating the Autocovariances from the Population Spectrum

If we know the sequence of autocovariances $\{\gamma_j\}_{j=-\infty}^{\infty}$, in principle we can calculate the value of $s_Y(\omega)$ for any ω from [6.1.2] or [6.1.6]. The converse is also true: if we know the value of $s_Y(\omega)$ for all ω in $[0, \pi]$, we can calculate the value of the kth autocovariance γ_k for any given k. This means that the population spectrum $s_Y(\omega)$ and the sequence of autocovariances contain exactly the same information—neither one can tell us anything about the process that is not possible to infer from the other.

The following proposition (proved in Appendix 6.A at the end of this chapter) provides a formula for calculating any autocovariance from the population spectrum.

Proposition 6.1: Let $\{\gamma_j\}_{j=-\infty}^{\infty}$ be an absolutely summable sequence of autocovariances, and define $s_Y(\omega)$ as in [6.1.2]. Then

$$\int_{-\pi}^{\pi} s_Y(\omega) e^{i\omega k}\, d\omega = \gamma_k. \qquad [6.1.15]$$

Result [6.1.15] can equivalently be written as

$$\int_{-\pi}^{\pi} s_Y(\omega) \cos(\omega k)\, d\omega = \gamma_k. \qquad [6.1.16]$$

Interpreting the Population Spectrum

The following result obtains as a special case of Proposition 6.1 by setting $k = 0$:

$$\int_{-\pi}^{\pi} s_Y(\omega)\, d\omega = \gamma_0. \qquad [6.1.17]$$

In other words, the area under the population spectrum between $\pm\pi$ gives γ_0, the variance of Y_t.

More generally—since $s_Y(\omega)$ is nonnegative—if we were to calculate

$$\int_{-\omega_1}^{\omega_1} s_Y(\omega)\, d\omega$$

for any ω_1 between 0 and π, the result would be a positive number that we could interpret as the portion of the variance of Y_t that is associated with frequencies ω that are less than ω_1 in absolute value. Recalling that $s_Y(\omega)$ is symmetric, the claim is that

$$2 \cdot \int_{0}^{\omega_1} s_Y(\omega)\, d\omega \qquad [6.1.18]$$

represents the portion of the variance of Y that could be attributed to periodic random components with frequency less than or equal to ω_1.

What does it mean to attribute a certain portion of the variance of Y to cycles with frequency less than or equal to ω_1? To explore this question, let us consider the following rather special stochastic process. Suppose that the value of Y at date t is determined by

$$Y_t = \sum_{j=1}^{M} [\alpha_j \cdot \cos(\omega_j t) + \delta_j \cdot \sin(\omega_j t)]. \qquad [6.1.19]$$

Here, α_j and δ_j are zero-mean random variables, meaning that $E(Y_t) = 0$ for all t. The sequences $\{\alpha_j\}_{j=1}^{M}$ and $\{\delta_j\}_{j=1}^{M}$ are serially uncorrelated and mutually uncorrelated:

$$E(\alpha_j \alpha_k) = \begin{cases} \sigma_j^2 & \text{for } j = k \\ 0 & \text{for } j \neq k \end{cases}$$

$$E(\delta_j \delta_k) = \begin{cases} \sigma_j^2 & \text{for } j = k \\ 0 & \text{for } j \neq k \end{cases}$$

$$E(\alpha_j \delta_k) = 0 \qquad \text{for all } j \text{ and } k.$$

The variance of Y_t is then

$$\begin{aligned} E(Y_t^2) &= \sum_{j=1}^{M} \left[E(\alpha_j^2) \cdot \cos^2(\omega_j t) + E(\delta_j^2) \cdot \sin^2(\omega_j t) \right] \\ &= \sum_{j=1}^{M} \sigma_j^2 \left[\cos^2(\omega_j t) + \sin^2(\omega_j t) \right] \qquad [6.1.20] \\ &= \sum_{j=1}^{M} \sigma_j^2, \end{aligned}$$

with the last line following from equation [A.1.12]. Thus, for this process, the portion of the variance of Y that is due to cycles of frequency ω_j is given by σ_j^2.

If the frequencies are ordered $0 < \omega_1 < \omega_2 < \cdots < \omega_M < \pi$, the portion of the variance of Y that is due to cycles of frequency less than or equal to ω_j is given by $\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_j^2$.

The kth autocovariance of Y is

$$\begin{aligned} E(Y_t Y_{t-k}) &= \sum_{j=1}^{M} \{E(\alpha_j^2)\cdot\cos(\omega_j t)\cdot\cos[\omega_j(t-k)] \\ &\quad + E(\delta_j^2)\cdot\sin(\omega_j t)\cdot\sin[\omega_j(t-k)]\} \\ &= \sum_{j=1}^{M} \sigma_j^2\{\cos(\omega_j t)\cdot\cos[\omega_j(t-k)] \\ &\quad + \sin(\omega_j t)\cdot\sin[\omega_j(t-k)]\}. \end{aligned} \qquad [6.1.21]$$

Recall the trigonometric identity[3]

$$\cos(A - B) = \cos(A)\cdot\cos(B) + \sin(A)\cdot\sin(B). \qquad [6.1.22]$$

For $A = \omega_j t$ and $B = \omega_j(t - k)$, we have $A - B = \omega_j k$, so that [6.1.21] becomes

$$E(Y_t Y_{t-k}) = \sum_{j=1}^{M} \sigma_j^2\cdot\cos(\omega_j k). \qquad [6.1.23]$$

Since the mean and the autocovariances of Y are not functions of time, the process described by [6.1.19] is covariance-stationary, although [6.1.23] implies that the sequence of autocovariances $\{\gamma_k\}_{k=0}^{\infty}$ is not absolutely summable.

We were able to attribute a certain portion of the variance of Y_t to cycles of less than a given frequency for the process in [6.1.19] because that is a rather special covariance-stationary process. However, there is a general result known as the *spectral representation theorem* which says that any covariance-stationary process Y_t can be expressed in terms of a generalization of [6.1.19]. For any fixed frequency ω in $[0, \pi]$, we define random variables $\alpha(\omega)$ and $\delta(\omega)$ and propose to write a stationary process with absolutely summable autocovariances in the form

$$Y_t = \mu + \int_0^{\pi} [\alpha(\omega)\cdot\cos(\omega t) + \delta(\omega)\cdot\sin(\omega t)]\, d\omega.$$

The random processes represented by $\alpha(\cdot)$ and $\delta(\cdot)$ have zero mean and the further properties that for any frequencies $0 < \omega_1 < \omega_2 < \omega_3 < \omega_4 < \pi$, the variable $\int_{\omega_1}^{\omega_2} \alpha(\omega)\, d\omega$ is uncorrelated with $\int_{\omega_3}^{\omega_4} \alpha(\omega)\, d\omega$ and the variable $\int_{\omega_1}^{\omega_2} \delta(\omega)\, d\omega$ is uncorrelated with $\int_{\omega_3}^{\omega_4} \delta(\omega)\, d\omega$, while for any $0 < \omega_1 < \omega_2 < \pi$ and $0 < \omega_3 < \omega_4 < \pi$, the variable $\int_{\omega_1}^{\omega_2} \alpha(\omega)\, d\omega$ is uncorrelated with $\int_{\omega_3}^{\omega_4} \delta(\omega)\, d\omega$. For such a process, one can calculate the portion of the variance of Y_t that is due to cycles with frequency less than or equal to some specified value ω_1 through a generalization of the procedure used to analyze [6.1.19]. Moreover, this magnitude turns out to be given by the expression in [6.1.18].

We shall not attempt a proof of the spectral representation theorem here; for details the reader is referred to Cramér and Leadbetter (1967, pp. 128–38). Instead, the next section provides a formal derivation of a finite-sample version of these results, showing the sense in which the sample analog of [6.1.18] gives the portion of the sample variance of an observed series that can be attributed to cycles with frequencies less than or equal to ω_1.

[3]See, for example, Thomas (1972, p. 176).

6.2. The Sample Periodogram

For a covariance-stationary process Y_t with absolutely summable autocovariances, we have defined the value of the population spectrum at frequency ω to be

$$s_Y(\omega) = \frac{1}{2\pi}\sum_{j=-\infty}^{\infty}\gamma_j e^{-i\omega j}, \qquad [6.2.1]$$

where

$$\gamma_j \equiv E(Y_t - \mu)(Y_{t-j} - \mu)$$

and $\mu = E(Y_t)$. Note that the population spectrum is expressed in terms of $\{\gamma_j\}_{j=0}^{\infty}$, which represents population second moments.

Given an observed sample of T observations denoted $y_1, y_2, \ldots, y_T$, we can calculate up to $T - 1$ sample autocovariances from the formulas

$$\hat{\gamma}_j = \begin{cases} T^{-1}\sum_{t=j+1}^{T}(y_t - \bar{y})(y_{t-j} - \bar{y}) & \text{for } j = 0, 1, 2, \ldots, T-1 \\ \hat{\gamma}_{-j} & \text{for } j = -1, -2, \ldots, -T+1, \end{cases} \qquad [6.2.2]$$

where $\bar{y}$ is the sample mean:

$$\bar{y} = T^{-1}\sum_{t=1}^{T} y_t. \qquad [6.2.3]$$

For any given ω we can then construct the sample analog of [6.2.1], which is known as the *sample periodogram*:

$$\hat{s}_y(\omega) = \frac{1}{2\pi}\sum_{j=-T+1}^{T-1}\hat{\gamma}_j e^{-i\omega j}. \qquad [6.2.4]$$

As in [6.1.6], the sample periodogram can equivalently be expressed as

$$\hat{s}_y(\omega) = \frac{1}{2\pi}\left[\hat{\gamma}_0 + 2\sum_{j=1}^{T-1}\hat{\gamma}_j\cos(\omega j)\right]. \qquad [6.2.5]$$

The same calculations that led to [6.1.17] can be used to show that the area under the periodogram is the sample variance of y:

$$\int_{-\pi}^{\pi}\hat{s}_y(\omega)\,d\omega = \hat{\gamma}_0.$$

Like the population spectrum, the sample periodogram is symmetric around $\omega = 0$, so that we could equivalently write

$$\hat{\gamma}_0 = 2\int_0^{\pi}\hat{s}_y(\omega)\,d\omega.$$

There also turns out to be a sample analog to the spectral representation theorem, which we now develop. In particular, we will see that given any T observations on a process $(y_1, y_2, \ldots, y_T)$, there exist frequencies $\omega_1, \omega_2, \ldots, \omega_M$ and coefficients $\hat{\mu}, \hat{\alpha}_1, \hat{\alpha}_2, \ldots, \hat{\alpha}_M, \hat{\delta}_1, \hat{\delta}_2, \ldots, \hat{\delta}_M$ such that the value for y at date t can be expressed as

$$y_t = \hat{\mu} + \sum_{j=1}^{M}\{\hat{\alpha}_j\cdot\cos[\omega_j(t-1)] + \hat{\delta}_j\cdot\sin[\omega_j(t-1)]\}, \qquad [6.2.6]$$

where the variable $\hat{\alpha}_j \cdot \cos[\omega_j(t-1)]$ is orthogonal in the sample to $\hat{\alpha}_k \cdot \cos[\omega_k(t-1)]$ for $j \neq k$, the variable $\hat{\delta}_j \cdot \sin[\omega_j(t-1)]$ is orthogonal to $\hat{\delta}_k \cdot \sin[\omega_k(t-1)]$ for $j \neq k$, and the variable $\hat{\alpha}_j \cdot \cos[\omega_j(t-1)]$ is orthogonal to $\hat{\delta}_k \cdot \sin[\omega_k(t-1)]$ for all j and k. The sample variance of y is $T^{-1}\Sigma_{t=1}^{T}(y_t - \bar{y})^2$, and the portion of this variance that can be attributed to cycles with frequency ω_j can be inferred from the sample periodogram $\hat{s}_y(\omega_j)$.

We will develop this claim for the case when the sample size T is an odd number. In this case y_t will be expressed in terms of periodic functions with $M \equiv (T-1)/2$ different frequencies in [6.2.6]. The frequencies $\omega_1, \omega_2, \ldots, \omega_M$ are specified as follows:

$$\begin{aligned} \omega_1 &= 2\pi/T \\ \omega_2 &= 4\pi/T \\ &\vdots \\ \omega_M &= 2M\pi/T. \end{aligned} \qquad [6.2.7]$$

Thus, the highest frequency considered is

$$\omega_M = \frac{2(T-1)\pi}{2T} < \pi.$$

Consider an OLS regression of the value of y_t on a constant and on the various cosine and sine terms,

$$y_t = \mu + \sum_{j=1}^{M} \{\alpha_j \cdot \cos[\omega_j(t-1)] + \delta_j \cdot \sin[\omega_j(t-1)]\} + u_t.$$

This can be viewed as a standard regression model of the form

$$y_t = \boldsymbol{\beta}'\mathbf{x}_t + u_t, \qquad [6.2.8]$$

where

$$\mathbf{x}_t = \begin{bmatrix} 1 & \cos[\omega_1(t-1)] & \sin[\omega_1(t-1)] & \cos[\omega_2(t-1)] & \sin[\omega_2(t-1)] \\ & & \cdots \cos[\omega_M(t-1)] & \sin[\omega_M(t-1)] \end{bmatrix}' \qquad [6.2.9]$$

$$\boldsymbol{\beta}' = [\mu \quad \alpha_1 \quad \delta_1 \quad \alpha_2 \quad \delta_2 \quad \cdots \quad \alpha_M \quad \delta_M]. \qquad [6.2.10]$$

Note that $\mathbf{x}_t$ has $(2M+1) = T$ elements, so that there are as many explanatory variables as observations. We will show that the elements of $\mathbf{x}_t$ are linearly independent, meaning that an OLS regression of y_t on $\mathbf{x}_t$ yields a perfect fit. Thus, the fitted values for this regression are of the form of [6.2.6] with no error term u_t. Moreover, the coefficients of this regression have the property that $\frac{1}{2}(\hat{\alpha}_j^2 + \hat{\delta}_j^2)$ represents the portion of the sample variance of y that can be attributed to cycles with frequency ω_j. This magnitude $\frac{1}{2}(\hat{\alpha}_j^2 + \hat{\delta}_j^2)$ further turns out to be proportional to the sample periodogram evaluated at ω_j. In other words, any observed series $y_1, y_2, \ldots, y_T$ can be expressed in terms of periodic functions as in [6.2.6], and the portion of the sample variance that is due to cycles with frequency ω_j can be found from the sample periodogram. These points are established formally in the following proposition, which is proved in Appendix 6.A at the end of this chapter.

Proposition 6.2: Let T denote an odd integer and let $M \equiv (T - 1)/2$. Let $\omega_j = 2\pi j/T$ for $j = 1, 2, \ldots, M$, and let $\mathbf{x}_t$ be the $(T \times 1)$ vector in [6.2.9]. Then

$$\sum_{t=1}^{T} \mathbf{x}_t\mathbf{x}_t' = \begin{bmatrix} T & \mathbf{0}' \\ \mathbf{0} & (T/2)\cdot\mathbf{I}_{T-1} \end{bmatrix}. \qquad [6.2.11]$$

Furthermore, let $\{y_1, y_2, \ldots, y_T\}$ be any T numbers. Then the following are true:

(*a*) *The value of y_t can be expressed as*

$$y_t = \hat{\mu} + \sum_{j=1}^{M} \{\hat{\alpha}_j\cdot\cos[\omega_j(t-1)] + \hat{\delta}_j\cdot\sin[\omega_j(t-1)]\},$$

with $\hat{\mu} = \bar{y}$ (the sample mean from [6.2.3]) and

$$\hat{\alpha}_j = (2/T)\sum_{t=1}^{T} y_t\cdot\cos[\omega_j(t-1)] \qquad \text{for } j = 1, 2, \ldots, M \qquad [6.2.12]$$

$$\hat{\delta}_j = (2/T)\sum_{t=1}^{T} y_t\cdot\sin[\omega_j(t-1)] \qquad \text{for } j = 1, 2, \ldots, M. \qquad [6.2.13]$$

(*b*) *The sample variance of y_t can be expressed as*

$$(1/T)\sum_{t=1}^{T}(y_t - \bar{y})^2 = (1/2)\sum_{j=1}^{M}(\hat{\alpha}_j^2 + \hat{\delta}_j^2), \qquad [6.2.14]$$

and the portion of the sample variance of y that can be attributed to cycles of frequency ω_j is given by $\frac{1}{2}(\hat{\alpha}_j^2 + \hat{\delta}_j^2)$.

(*c*) *The portion of the sample variance of y that can be attributed to cycles of frequency ω_j can equivalently be expressed as*

$$(1/2)(\hat{\alpha}_j^2 + \hat{\delta}_j^2) = (4\pi/T)\cdot\hat{s}_y(\omega_j), \qquad [6.2.15]$$

where $\hat{s}_y(\omega_j)$ is the sample periodogram at frequency ω_j.

Result [6.2.11] establishes that $\Sigma_{t=1}^{T}\mathbf{x}_t\mathbf{x}_t'$ is a diagonal matrix, meaning that the explanatory variables contained in $\mathbf{x}_t$ are mutually orthogonal. The proposition asserts that any observed time series $(y_1, y_2, \ldots, y_T)$ with T odd can be written as a constant plus a weighted sum of $(T - 1)$ periodic functions with $(T - 1)/2$ different frequencies; a related result can also be developed when T is an even integer. Hence, the proposition gives a finite-sample analog of the spectral representation theorem. The proposition further shows that the sample periodogram captures the portion of the sample variance of y that can be attributed to cycles of different frequencies.

Note that the frequencies ω_j in terms of which the variance of y is explained all lie in $[0, \pi]$. Why aren't negative frequencies $\omega < 0$ employed as well? Suppose that the data were actually generated by a special case of the process in [6.1.19],

$$Y_t = \alpha\cdot\cos(-\omega t) + \delta\cdot\sin(-\omega t), \qquad [6.2.16]$$

where $-\omega < 0$ represents some particular negative frequency and where α and δ are zero-mean random variables. Since $\cos(-\omega t) = \cos(\omega t)$ and $\sin(-\omega t) = -\sin(\omega t)$, the process [6.2.16] can equivalently be written

$$Y_t = \alpha\cdot\cos(\omega t) - \delta\cdot\sin(\omega t). \qquad [6.2.17]$$

Thus there is no way of using observed data on y to decide whether the data are generated by a cycle with frequency $-\omega$ as in [6.2.16] or by a cycle with frequency

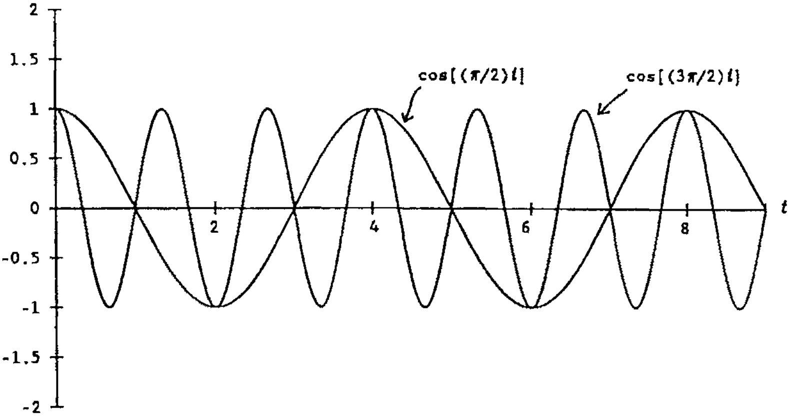

FIGURE 6.1 Aliasing: plots of $\cos[(\pi/2)t]$ and $\cos[(3\pi/2)t]$ as functions of t.

$+\omega$ as in [6.2.17]. It is simply a matter of convention that we choose to focus only on positive frequencies.

Why is $\omega = \pi$ the largest frequency considered? Suppose the data were generated from a periodic function with frequency $\omega > \pi$, say, $\omega = 3\pi/2$ for illustration:

$$Y_t = \alpha\cdot\cos[(3\pi/2)t] + \delta\cdot\sin[(3\pi/2)t]. \qquad [6.2.18]$$

Again, the properties of the sine and cosine function imply that [6.2.18] is equivalent to

$$Y_t = \alpha\cdot\cos[(-\pi/2)t] + \delta\cdot\sin[(-\pi/2)t]. \qquad [6.2.19]$$

Thus, by the previous argument, a representation with cycles of frequency $(3\pi/2)$ is observationally indistinguishable from one with cycles of frequency $(\pi/2)$.

To summarize, if the data-generating process actually includes cycles with negative frequencies or with frequencies greater than π, these will be imputed to cycles with frequencies between 0 and π. This is known as *aliasing*.

Another way to think about aliasing is as follows. Recall that the value of the function $\cos(\omega t)$ repeats itself every $2\pi/\omega$ periods, so that a frequency of ω is associated with a period of $2\pi/\omega$.[4] We have argued that the highest-frequency cycle that one can observe is $\omega = \pi$. Another way to express this conclusion is that the shortest-period cycle that one can observe is one that repeats itself every $2\pi/\pi = 2$ periods. If $\omega = 3\pi/2$, the cycle repeats itself every $\frac{4}{3}$ periods. But if the data are observed only at integer dates, the sampled data will exhibit cycles that are repeated every four periods, corresponding to the frequency $\omega = \pi/2$. This is illustrated in Figure 6.1, which plots $\cos[(\pi/2)t]$ and $\cos[(3\pi/2)t]$ as functions of t. When sampled at integer values of t, these two functions appear identical. Even though the function $\cos[(3\pi/2)t]$ repeats itself every time that t increases by $\frac{4}{3}$, one would have to observe y_t at four distinct dates $(y_t, y_{t+1}, y_{t+2}, y_{t+3})$ before one would see the value of $\cos[(3\pi/2)t]$ repeat itself for an integer value of t.

[4]See Section A.1 of the Mathematical Review (Appendix A) at the end of the book for a further discussion of this point.

Note that in a particular finite sample, the lowest frequency used to account for variation in y is $\omega_1 = 2\pi/T$, which corresponds to a period of T. If a cycle takes longer than T periods to repeat itself, there is not much that one could infer about it if one has only T observations available.

Result (c) of Proposition 6.2 indicates that the portion of the sample variance of y that can be attributed to cycles of frequency ω_j is proportional to the sample periodogram evaluated at ω_j, with $4\pi/T$ the constant of proportionality. Thus, the proposition develops the formal basis for the claim that the sample periodogram reflects the portion of the sample variance of y that can be attributed to cycles of various frequencies.

Why is the constant of proportionality in [6.2.15] equal to $4\pi/T$? The population spectrum $s_Y(\omega)$ could be evaluated at any ω in the continuous set of points between 0 and π. In this respect it is much like a probability density $f_X(x)$, where X is a continuous random variable. Although we might loosely think of the value of $f_X(x)$ as the "probability" that $X = x$, it is more accurate to say that the integral $\int_{x_1}^{x_2} f_X(x)\,dx$ represents the probability that X takes on a value between x_1 and x_2. As $x_2 - x_1$ becomes smaller, the probability that X will be observed to lie between x_1 and x_2 becomes smaller, and the probability that X would take on precisely the value x is effectively equal to zero. In just the same way, although we can loosely think of the value of $s_Y(\omega)$ as the contribution that cycles with frequency ω make to the variance of Y, it is more accurate to say that the integral

$$\int_{-\omega_1}^{\omega_1} s_Y(\omega)\,d\omega = \int_0^{\omega_1} 2s_Y(\omega)\,d\omega$$

represents the contribution that cycles of frequency less than or equal to ω_1 make to the variance of Y, and that $\int_{\omega_1}^{\omega_2} 2s_Y(\omega)\,d\omega$ represents the contribution that cycles with frequencies between ω_1 and ω_2 make to the variance of Y. Assuming that $s_Y(\omega)$ is continuous, the contribution that a cycle of any particular frequency ω makes is technically zero.

Although the population spectrum $s_Y(\omega)$ is defined at any ω in $[0, \pi]$, the representation in [6.2.6] attributes all of the sample variance of y to the particular frequencies $\omega_1, \omega_2, \ldots, \omega_M$. Any variation in Y that is in reality due to cycles with frequencies other than these M particular values is attributed by [6.2.6] to one of these M frequencies. If we are thinking of the regression in [6.2.6] as telling us something about the population spectrum, we should interpret $\frac{1}{2}(\hat{\alpha}_j^2 + \hat{\delta}_j^2)$ not as the portion of the variance of Y that is due to cycles with frequency exactly equal to ω_j, but rather as the portion of the variance of Y that is due to cycles with frequency near ω_j. Thus [6.2.15] is not an estimate of the height of the population spectrum, but an estimate of the area under the population spectrum.

This is illustrated in Figure 6.2. Suppose we thought of $\frac{1}{2}(\hat{\alpha}_j^2 + \hat{\delta}_j^2)$ as an estimate of the portion of the variance of Y that is due to cycles with frequency between ω_{j-1} and ω_j, that is, an estimate of 2 times the area under $s_Y(\omega)$ between ω_{j-1} and ω_j. Since $\omega_j = 2\pi j/T$, the difference $\omega_j - \omega_{j-1}$ is equal to $2\pi/T$. If $\hat{s}_y(\omega_j)$ is an estimate of $s_Y(\omega_j)$, then the area under $s_Y(\omega)$ between ω_{j-1} and ω_j could be approximately estimated by the area of a rectangle with width $2\pi/T$ and height $\hat{s}_y(\omega_j)$. The area of such a rectangle is $(2\pi/T)\cdot\hat{s}_j(\omega_j)$. Since $\frac{1}{2}(\hat{\alpha}_j^2 + \hat{\delta}_j^2)$ is an estimate of 2 times the area under $s_Y(\omega)$ between ω_{j-1} and ω_j, we have $\frac{1}{2}(\hat{\alpha}_j^2 + \hat{\delta}_j^2) = (4\pi/T)\cdot\hat{s}_y(\omega_j)$, as claimed in equation [6.2.15].

Proposition 6.2 also provides a convenient formula for calculating the value of the sample periodogram at frequency $\omega_j = 2\pi j/T$ for $j = 1, 2, \ldots, (T - 1)/2$,

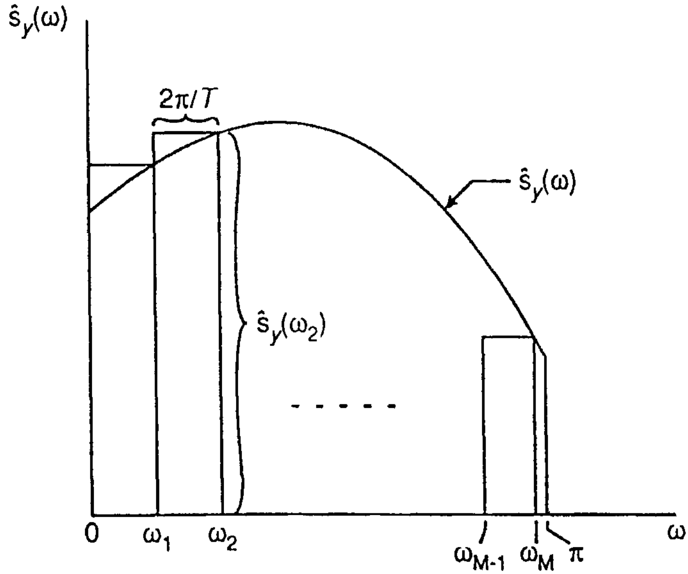

FIGURE 6.2 The area under the sample periodogram and the portion of the variance of y attributable to cycles of different frequencies.

namely,

$$\hat{s}_y(\omega_j) = [T/(8\pi)](\hat{\alpha}_j^2 + \hat{\delta}_j^2),$$

where

$$\hat{\alpha}_j = (2/T) \sum_{t=1}^{T} y_t \cdot \cos[\omega_j(t - 1)]$$

$$\hat{\delta}_j = (2/T) \sum_{t=1}^{T} y_t \cdot \sin[\omega_j(t - 1)].$$

That is,

$$\hat{s}_y(\omega_j) = \frac{1}{2\pi T}\left\{\left[\sum_{t=1}^{T} y_t \cdot \cos[\omega_j(t - 1)]\right]^2 + \left[\sum_{t=1}^{T} y_t \cdot \sin[\omega_j(t - 1)]\right]^2\right\}.$$

6.3. *Estimating the Population Spectrum*

Section 6.1 introduced the population spectrum $s_Y(\omega)$, which indicates the portion of the population variance of Y that can be attributed to cycles of frequency ω. This section addresses the following question: Given an observed sample $\{y_1, y_2, \ldots, y_T\}$, how might $s_Y(\omega)$ be estimated?

Large-Sample Properties of the Sample Periodogram

One obvious approach would be to estimate the population spectrum $s_Y(\omega)$ by the sample periodogram $\hat{s}_y(\omega)$. However, this approach turns out to have some serious limitations. Suppose that

$$Y_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j},$$

where $\{\psi_j\}_{j=0}^{\infty}$ is absolutely summable and where $\{\varepsilon_t\}_{t=-\infty}^{\infty}$ is an i.i.d. sequence with $E(\varepsilon_t) = 0$ and $E(\varepsilon_t^2) = \sigma^2$. Let $s_Y(\omega)$ be the population spectrum defined in [6.1.2], and suppose that $s_Y(\omega) > 0$ for all ω. Let $\hat{s}_y(\omega)$ be the sample periodogram defined in [6.2.4]. Fuller (1976, p. 280) showed that for $\omega \neq 0$ and a sufficiently large sample size T, twice the ratio of the sample periodogram to the population spectrum has approximately the following distribution:

$$\frac{2 \cdot \hat{s}_y(\omega)}{s_Y(\omega)} \approx \chi^2(2). \qquad [6.3.1]$$

Moreover, if $\lambda \neq \omega$, the quantity

$$\frac{2 \cdot \hat{s}_y(\lambda)}{s_Y(\lambda)} \qquad [6.3.2]$$

also has an approximate $\chi^2(2)$ distribution, with the variable in [6.3.1] approximately independent of that in [6.3.2].

Since a $\chi^2(2)$ variable has a mean of 2, result [6.3.1] suggests that

$$E\left[\frac{2 \cdot \hat{s}_y(\omega)}{s_Y(\omega)}\right] \cong 2,$$

or since $s_Y(\omega)$ is a population magnitude rather than a random variable,

$$E[\hat{s}_y(\omega)] \cong s_Y(\omega).$$

Thus, if the sample size is sufficiently large, the sample periodogram affords an approximately unbiased estimate of the population spectrum.

Note from Table B.2 that 95% of the time, a $\chi^2(2)$ variable will fall between 0.05 and 7.4. Thus, from [6.3.1], $\hat{s}_y(\omega)$ is unlikely to be as small as 0.025 times the true value of $s_Y(\omega)$, and $\hat{s}_y(\omega)$ is unlikely to be any larger than 3.7 times as big as $s_Y(\omega)$. Given such a large confidence interval, we would have to say that $\hat{s}_y(\omega)$ is not an altogether satisfactory estimate of $s_Y(\omega)$.

Another feature of result [6.3.1] is that the estimate $\hat{s}_y(\omega)$ is not getting any more accurate as the sample size T increases. Typically, one expects an econometric estimate to get better and better as the sample size grows. For example, the variance for the sample autocorrelation coefficient $\hat{\rho}_j$ given in [4.8.8] goes to zero as $T \to \infty$, so that given a sufficiently large sample, we would be able to infer the true value of ρ_j with virtual certainty. The estimate $\hat{s}_y(\omega)$ defined in [6.2.4] does not have this property, because we have tried to estimate as many parameters ($\gamma_0, \gamma_1, \ldots, \gamma_{T-1}$) as we had observations ($y_1, y_2, \ldots, y_T$).

Parametric Estimates of the Population Spectrum

Suppose we believe that the data could be represented with an $ARMA(p, q)$ model,

$$\begin{aligned} Y_t = \mu + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} \\ + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q}, \end{aligned} \qquad [6.3.3]$$

where ε_t is white noise with variance σ^2. Then an excellent approach to estimating the population spectrum is first to estimate the parameters $\mu, \phi_1, \ldots, \phi_p, \theta_1,$

$\ldots, \theta_q$ and σ^2 by maximum likelihood as described in the previous chapter. The maximum likelihood estimates $(\hat{\phi}_1, \ldots, \hat{\phi}_p, \hat{\theta}_1, \ldots, \hat{\theta}_q, \hat{\sigma}^2)$ could then be plugged into a formula such as [6.1.14] to estimate the population spectrum $s_Y(\omega)$ at any frequency ω. If the model is correctly specified, the maximum likelihood estimates $(\hat{\phi}_1, \ldots, \hat{\phi}_p, \hat{\theta}_1, \ldots, \hat{\theta}_q, \hat{\sigma}^2)$ will get closer and closer to the true values as the sample size grows; hence, the resulting estimate of the population spectrum should have this same property.

Even if the model is incorrectly specified, if the autocovariances of the true process are reasonably close to those for an $ARMA(p, q)$ specification, this procedure should provide a useful estimate of the population spectrum.

Nonparametric Estimates of the Population Spectrum

The assumption in [6.3.3] is that Y_t can be reasonably approximated by an $ARMA(p, q)$ process with p and q small. An alternative assumption is that $s_Y(\omega)$ will be close to $s_Y(\lambda)$ when ω is close to λ. This assumption forms the basis for another class of estimates of the population spectrum known as *nonparametric* or *kernel estimates*.

If $s_Y(\omega)$ is close to $s_Y(\lambda)$ when ω is close to λ, this suggests that $s_Y(\omega)$ might be estimated with a weighted average of the values of $\hat{s}_y(\lambda)$ for values of λ in a neighborhood around ω, where the weights depend on the distance between ω and λ. Let $\hat{s}_Y(\omega)$ denote such an estimate of $s_Y(\omega)$ and let $\omega_j = 2\pi j/T$. The suggestion is to take

$$\hat{s}_Y(\omega_j) = \sum_{m=-h}^{h} \kappa(\omega_{j+m}, \omega_j) \cdot \hat{s}_y(\omega_{j+m}). \qquad [6.3.4]$$

Here, h is a *bandwidth* parameter indicating how many different frequencies $\{\omega_{j\pm1}, \omega_{j\pm2}, \ldots, \omega_{j\pm h}\}$ are viewed as useful for estimating $s_Y(\omega_j)$. The *kernel* $\kappa(\omega_{j+m}, \omega_j)$ indicates how much weight each frequency is to be given. The kernel weights sum to unity:

$$\sum_{m=-h}^{h} \kappa(\omega_{j+m}, \omega_j) = 1.$$

One approach is to take $\kappa(\omega_{j+m}, \omega_j)$ to be proportional to $h + 1 - |m|$. One can show that[5]

$$\sum_{m=-h}^{h} [h + 1 - |m|] = (h + 1)^2.$$

Hence, in order to satisfy the property that the weights sum to unity, the proposed kernel is

$$\kappa(\omega_{j+m}, \omega_j) = \frac{h + 1 - |m|}{(h + 1)^2} \qquad [6.3.5]$$

[5]Notice that

$$\begin{aligned}\sum_{m=-h}^{h} [h + 1 - |m|] &= \sum_{m=-h}^{h} (h + 1) - \sum_{m=-h}^{h} |m| \\ &= (h + 1) \sum_{m=-h}^{h} 1 - 2\sum_{s=0}^{h} s \\ &= (2h + 1)(h + 1) - 2h(h + 1)/2 \\ &= (h + 1)^2.\end{aligned}$$

and the estimator [6.3.4] becomes

$$\hat{s}_Y(\omega_j) = \sum_{m=-h}^{h} \left[\frac{h + 1 - |m|}{(h + 1)^2}\right] \hat{s}_y(\omega_{j+m}). \qquad [6.3.6]$$

For example, for $h = 2$, this is

$$\hat{s}_Y(\omega_j) = \tfrac{1}{9}\hat{s}_y(\omega_{j-2}) + \tfrac{2}{9}\hat{s}_y(\omega_{j-1}) + \tfrac{3}{9}\hat{s}_y(\omega_j) + \tfrac{2}{9}\hat{s}_y(\omega_{j+1}) + \tfrac{1}{9}\hat{s}_y(\omega_{j+2}).$$

Recall from [6.3.1] and [6.3.2] that the estimates $\hat{s}_y(\omega)$ and $\hat{s}_y(\lambda)$ are approximately independent in large samples for $\omega \neq \lambda$. Because the kernel estimate averages over a number of different frequencies, it should give a much better estimate than does the periodogram.

Averaging $\hat{s}_y(\omega)$ over different frequencies can equivalently be represented as multiplying the jth sample autocovariance $\hat{\gamma}_j$ for $j > 0$ in the formula for the sample periodogram [6.2.5] by a weight κ_j^*. For example, consider an estimate of the spectrum at frequency ω that is obtained by taking a simple average of the value of $\hat{s}_y(\lambda)$ for λ between $\omega - \nu$ and $\omega + \nu$:

$$\hat{s}_Y(\omega) = (2\nu)^{-1} \int_{\omega-\nu}^{\omega+\nu} \hat{s}_y(\lambda)\, d\lambda. \qquad [6.3.7]$$

Substituting [6.2.5] into [6.3.7], such an estimate could equivalently be expressed as

$$\begin{aligned} \hat{s}_Y(\omega) &= (4\nu\pi)^{-1} \int_{\omega-\nu}^{\omega+\nu} \left[\hat{\gamma}_0 + 2\sum_{j=1}^{T-1} \hat{\gamma}_j \cos(\lambda j)\right] d\lambda \\ &= (4\nu\pi)^{-1}(2\nu)\hat{\gamma}_0 + (2\nu\pi)^{-1} \sum_{j=1}^{T-1} \hat{\gamma}_j (1/j)\cdot\big[\sin(\lambda j)\big]_{\lambda=\omega-\nu}^{\omega+\nu} \\ &= (2\pi)^{-1}\hat{\gamma}_0 + (2\nu\pi)^{-1} \sum_{j=1}^{T-1} \hat{\gamma}_j (1/j)\cdot\{\sin[(\omega + \nu)j] - \sin[(\omega - \nu)j]\}. \end{aligned} \qquad [6.3.8]$$

Using the trigonometric identity[6]

$$\sin(A + B) - \sin(A - B) = 2\cdot\cos(A)\cdot\sin(B), \qquad [6.3.9]$$

expression [6.3.8] can be written

$$\begin{aligned} \hat{s}_Y(\omega) &= (2\pi)^{-1}\hat{\gamma}_0 + (2\nu\pi)^{-1} \sum_{j=1}^{T-1} \hat{\gamma}_j (1/j)\cdot[2\cdot\cos(\omega j)\cdot\sin(\nu j)] \\ &= (2\pi)^{-1}\left\{\hat{\gamma}_0 + 2\sum_{j=1}^{T-1}\left[\frac{\sin(\nu j)}{\nu j}\right]\hat{\gamma}_j \cos(\omega j)\right\}. \end{aligned} \qquad [6.3.10]$$

Notice that expression [6.3.10] is of the following form:

$$\hat{s}_Y(\omega) = (2\pi)^{-1}\left\{\hat{\gamma}_0 + 2\sum_{j=1}^{T-1} \kappa_j^* \hat{\gamma}_j \cos(\omega j)\right\}, \qquad [6.3.11]$$

where

$$\kappa_j^* = \left[\frac{\sin(\nu j)}{\nu j}\right]. \qquad [6.3.12]$$

The sample periodogram can be regarded as a special case of [6.3.11] when $\kappa_j^* = 1$. Expression [6.3.12] cannot exceed 1 in absolute value, and so the estimate [6.3.11] essentially downweights $\hat{\gamma}_j$ relative to the sample periodogram.

[6]See, for example, Thomas (1972, pp. 174–75).

Recall that $\sin(\pi j) = 0$ for any integer j. Hence, if $\nu = \pi$, then $\kappa_j^* = 0$ for all j and [6.3.11] becomes

$$\hat{s}_Y(\omega) = (2\pi)^{-1}\hat{\gamma}_0. \qquad [6.3.13]$$

In this case, all autocovariances other than $\hat{\gamma}_0$ would be shrunk to zero. When $\nu = \pi$, the estimate [6.3.7] is an unweighted average of $\hat{s}_y(\lambda)$ over all possible values of λ, and the resulting estimate would be the flat spectrum for a white noise process.

Specification of a kernel function $\kappa(\omega_{j+m}, \omega_j)$ in [6.3.4] can equivalently be described in terms of a weighting sequence $\{\kappa_j^*\}_{j=1}^{T-1}$ in [6.3.11]. Because they are just two different representations for the same idea, the weight κ_j^* is also sometimes called a kernel. Smaller values of κ_j^* impose more smoothness on the spectrum. Smoothing schemes may be chosen either because they provide a convenient specification for $\kappa(\omega_{j+m}, \omega_j)$ or because they provide a convenient specification for κ_j^*.

One popular estimate of the spectrum employs the modified *Bartlett kernel*, which is given by

$$\kappa_j^* = \begin{cases} 1 - \dfrac{j}{q+1} & \text{for } j = 1, 2, \ldots, q \\ 0 & \text{for } j > q. \end{cases} \qquad [6.3.14]$$

The Bartlett estimate of the spectrum is thus

$$\hat{s}_Y(\omega) = (2\pi)^{-1}\left\{\hat{\gamma}_0 + 2\sum_{j=1}^{q}[1 - j/(q+1)]\hat{\gamma}_j\cos(\omega j)\right\}. \qquad [6.3.15]$$

Autocovariances γ_j for $j > q$ are treated as if they were zero, or as if Y_t followed an $MA(q)$ process. For $j \leq q$, the estimated autocovariances $\hat{\gamma}_j$ are shrunk toward zero, with the shrinkage greater the larger the value of j.

How is one to choose the bandwidth parameter h in [6.3.6] or q in [6.3.15]? The periodogram itself is asymptotically unbiased but has a large variance. If one constructs an estimate based on averaging the periodogram at different frequencies, this reduces the variance but introduces some bias. The severity of the bias depends on the steepness of the population spectrum and the size of the bandwidth. One practical guide is to plot an estimate of the spectrum using several different bandwidths and rely on subjective judgment to choose the bandwidth that produces the most plausible estimate.

6.4. *Uses of Spectral Analysis*

We illustrate some of the uses of spectral analysis with data on manufacturing production in the United States. The data are plotted in Figure 6.3. The series is the Federal Reserve Board's seasonally unadjusted monthly index from January 1947 to November 1989. Economic recessions in 1949, 1954, 1958, 1960, 1970, 1974, 1980, and 1982 appear as roughly year-long episodes of falling production. There are also strong seasonal patterns in this series; for example, production almost always declines in July and recovers in August.

The sample periodogram for the raw data is plotted in Figure 6.4, which displays $\hat{s}_y(\omega_j)$ as a function of j where $\omega_j = 2\pi j/T$. The contribution to the sample variance of the lowest-frequency components (j near zero) is several orders of magnitude larger than the contributions of economic recessions or the seasonal factors. This is due to the clear upward trend of the series in Figure 6.3. Let y_t

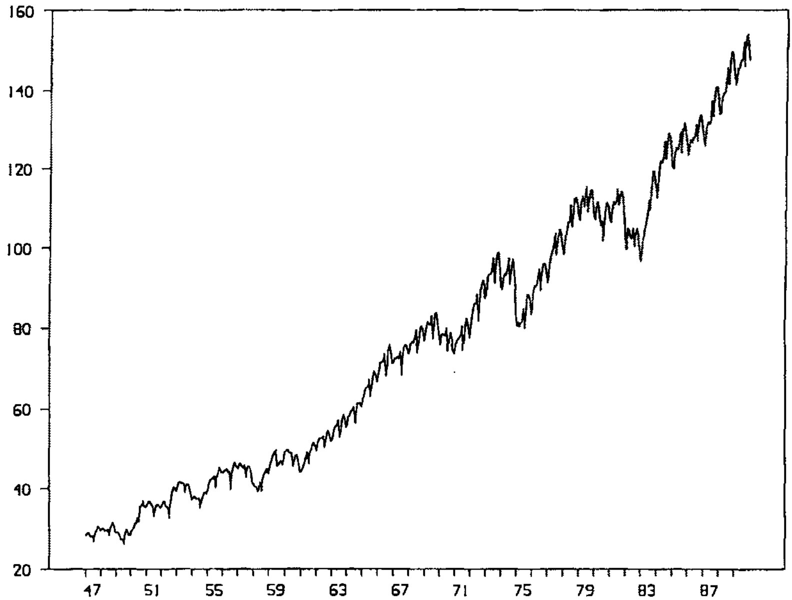

FIGURE 6.3 Federal Reserve Board's seasonally unadjusted index of industrial production for U.S. manufacturing, monthly 1947:1 to 1989:11.

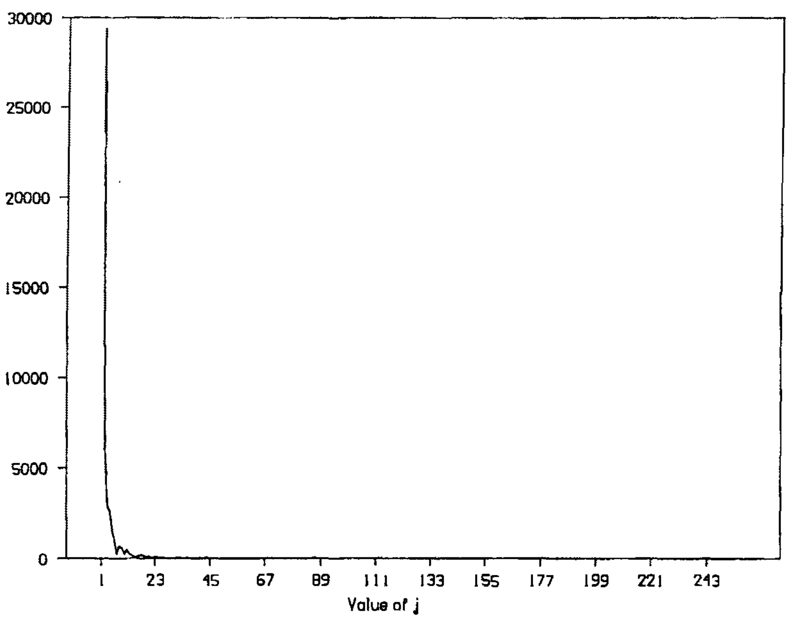

FIGURE 6.4 Sample periodogram for the data plotted in Figure 6.3. The figure plots $\hat{s}_y(\omega_j)$ as a function of j, where $\omega_j = 2\pi j/T$.

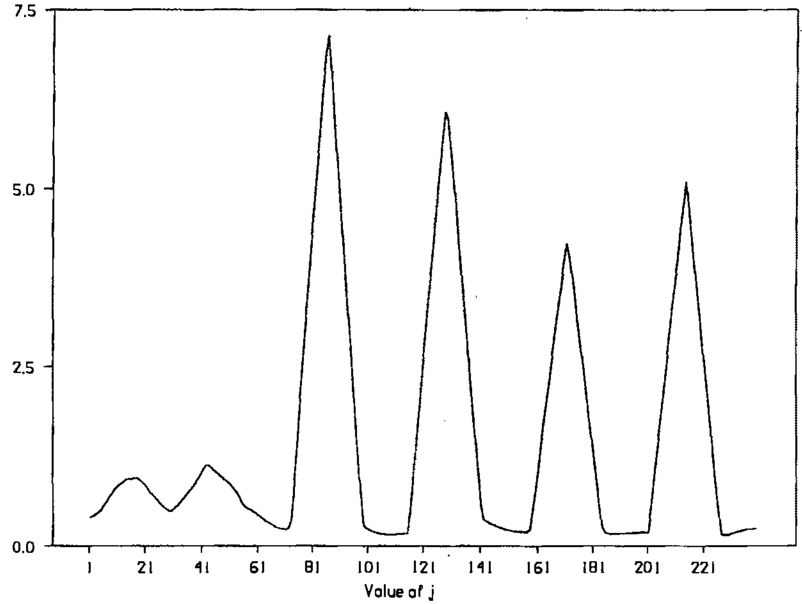

FIGURE 6.5 Estimate of the spectrum for monthly growth rate of industrial production, or spectrum of 100 times the first difference of the log of the series in Figure 6.3.

represent the series plotted in Figure 6.3. If one were trying to describe this with a sine function

$$y_t = \delta \cdot \sin(\omega t),$$

the presumption would have to be that ω is so small that even at date $t = T$ the magnitude ωT would still be less than $\pi/2$. Figure 6.4 thus indicates that the trend or low-frequency components are by far the most important determinants of the sample variance of y.

The definition of the population spectrum in equation [6.1.2] assumed that the process is covariance-stationary, which is not a good assumption for the data in Figure 6.3. We might instead try to analyze the monthly growth rate defined by

$$x_t = 100 \cdot [\log(y_t) - \log(y_{t-1})]. \qquad [6.4.1]$$

Figure 6.5 plots the estimate of the population spectrum of X as described in equation [6.3.6] with $h = 12$.

In interpreting a plot such as Figure 6.5 it is often more convenient to think in terms of the period of a cyclic function rather than its frequency. Recall that if the frequency of a cycle is ω, the period of the cycle is $2\pi/\omega$. Thus, a frequency of $\omega_j = 2\pi j/T$ corresponds to a period of $2\pi/\omega_j = T/j$. The sample size is $T = 513$ observations, and the first peak in Figure 6.5 occurs around $j = 18$. This corresponds to a cycle with a period of $513/18 = 28.5$ months, or about $2\frac{1}{2}$ years. Given the dates of the economic recessions noted previously, this is sometimes described as a "business cycle frequency," and the area under this hill might be viewed as telling us how much of the variability in monthly growth rates is due to economic recessions.

The second peak in Figure 6.5 occurs at $j = 44$ and corresponds to a period of $513/44 = 11.7$ months. This is natural to view as a 12-month cycle associated with seasonal effects. The four subsequent peaks correspond to cycles with periods of 6, 4, 3, and 2.4 months, respectively, and again seem likely to be picking up seasonal and calendar effects.

Since manufacturing typically falls temporarily in July, the growth rate is negative in July and positive in August. This induces negative first-order serial correlation to the series in [6.4.1] and a variety of calendar patterns for x_t that may account for the high-frequency peaks in Figure 6.5. An alternative strategy for detrending would use year-to-year growth rates, or the percentage change between y_t and its value for the corresponding month in the previous year:

$$w_t = 100 \cdot [\log(y_t) - \log(y_{t-12})]. \qquad [6.4.2]$$

The estimate of the sample spectrum for this series is plotted in Figure 6.6. When the data are detrended in this way, virtually all the variance that remains is attributed to components associated with the business cycle frequencies.

Filters

Apart from the scale parameter, the monthly growth rate x_t in [6.4.1] is obtained from $\log(y_t)$ by applying the filter

$$x_t = (1 - L)\log(y_t), \qquad [6.4.3]$$

where L is the lag operator. To discuss such transformations in general terms, let Y_t be any covariance-stationary series with absolutely summable autocovariances.

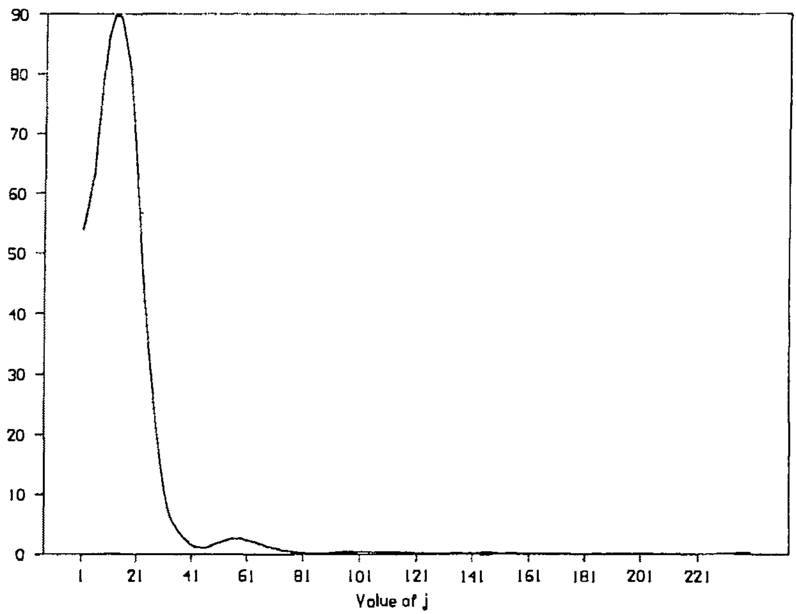

FIGURE 6.6 Estimate of the spectrum for year-to-year growth rate of monthly industrial production, or spectrum of 100 times the seasonal difference of the log of the series in Figure 6.3.

Denote the autocovariance-generating function of Y by $g_Y(z)$, and denote the population spectrum of Y by $s_Y(\omega)$. Recall that

$$s_Y(\omega) = (2\pi)^{-1} g_Y(e^{-i\omega}). \tag{6.4.4}$$

Suppose we transform Y according to

$$X_t = h(L)Y_t,$$

where

$$h(L) = \sum_{j=-\infty}^{\infty} h_j L^j$$

and

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

Recall from equation [3.6.17] that the autocovariance-generating function of X can be calculated from the autocovariance-generating function of Y using the formula

$$g_X(z) = h(z)h(z^{-1})g_Y(z). \tag{6.4.5}$$

The population spectrum of X is thus

$$s_X(\omega) = (2\pi)^{-1} g_X(e^{-i\omega}) = (2\pi)^{-1} h(e^{-i\omega})h(e^{i\omega})g_Y(e^{-i\omega}). \tag{6.4.6}$$

Substituting [6.4.4] into [6.4.6] reveals that the population spectrum of X is related to the population spectrum of Y according to

$$s_X(\omega) = h(e^{-i\omega})h(e^{i\omega})s_Y(\omega). \tag{6.4.7}$$

Operating on a series Y_t with the filter $h(L)$ has the effect of multiplying the spectrum by the function $h(e^{-i\omega})h(e^{i\omega})$.

For the difference operator in [6.4.3], the filter is $h(L) = 1 - L$ and the function $h(e^{-i\omega})h(e^{i\omega})$ would be

$$\begin{aligned} h(e^{-i\omega})h(e^{i\omega}) &= (1 - e^{-i\omega})(1 - e^{i\omega}) \\ &= 1 - e^{-i\omega} - e^{i\omega} + 1 \\ &= 2 - 2\cdot\cos(\omega), \end{aligned} \tag{6.4.8}$$

where the last line follows from [6.1.11]. If $X_t = (1 - L)Y_t$, then, to find the value of the population spectrum of X at any frequency ω, we first find the value of the population spectrum of Y at ω and then multiply by $2 - 2\cdot\cos(\omega)$. For example, the spectrum at frequency $\omega = 0$ is multiplied by zero, the spectrum at frequency $\omega = \pi/2$ is multiplied by 2, and the spectrum at frequency $\omega = \pi$ is multiplied by 4. Differencing the data removes the low-frequency components and accentuates the high-frequency components.

Of course, this calculation assumes that the original process Y_t is covariance-stationary, so that $s_Y(\omega)$ exists. If the original process is nonstationary, as appears to be the case in Figure 6.3, the differenced data $(1 - L)Y_t$ in general would not have a population spectrum that is zero at frequency zero.

The seasonal difference filter used in [6.4.2] is $h(L) = 1 - L^{12}$, for which

$$\begin{aligned} h(e^{-i\omega})h(e^{i\omega}) &= (1 - e^{-12i\omega})(1 - e^{12i\omega}) \\ &= 1 - e^{-12i\omega} - e^{12i\omega} + 1 \\ &= 2 - 2\cdot\cos(12\omega). \end{aligned}$$

This function is equal to zero when $12\omega = 0, 2\pi, 4\pi, 6\pi, 8\pi, 10\pi$, or 12π; that is, it is zero at frequencies $\omega = 0, 2\pi/12, 4\pi/12, 6\pi/12, 8\pi/12, 10\pi/12$, and π. Thus, seasonally differencing not only eliminates the low-frequency ($\omega = 0$) components of a stationary process, but further eliminates any contribution from cycles with periods of 12, 6, 4, 3, 2.4, or 2 months.

Composite Stochastic Processes

Let X_t be covariance-stationary with absolutely summable autocovariances, autocovariance-generating function $g_X(z)$, and population spectrum $s_X(\omega)$. Let W_t be a different covariance-stationary series with absolutely summable autocovariances, autocovariance-generating function $g_W(z)$, and population spectrum $s_W(\omega)$, where X_t is uncorrelated with W_τ for all t and τ. Suppose we observe the sum of these two processes,

$$Y_t = X_t + W_t.$$

Recall from [4.7.19] that the autocovariance-generating function of the sum is the sum of the autocovariance-generating functions:

$$g_Y(z) = g_X(z) + g_W(z).$$

It follows from [6.1.2] that the spectrum of the sum is the sum of the spectra:

$$s_Y(\omega) = s_X(\omega) + s_W(\omega). \qquad [6.4.9]$$

For example, if a white noise series W_t with variance σ^2 is added to a series X_t and if X_t is uncorrelated with W_τ for all t and τ, the effect is to shift the population spectrum everywhere up by the constant $\sigma^2/(2\pi)$. More generally, if X has a peak in its spectrum at frequency ω_1 and if W has a peak in its spectrum at ω_2, then typically the sum $X + W$ will have peaks at both ω_1 and ω_2.

As another example, suppose that

$$Y_t = c + \sum_{j=-\infty}^{\infty} h_j X_{t-j} + \varepsilon_t,$$

where X_t is covariance-stationary with absolutely summable autocovariances and spectrum $s_X(\omega)$. Suppose that the sequence $\{h_j\}_{j=-\infty}^{\infty}$ is absolutely summable and that ε_t is a white noise process with variance σ^2 where ε is uncorrelated with X at all leads and lags. It follows from [6.4.7] that the random variable $\sum_{j=-\infty}^{\infty} h_j X_{t-j}$ has spectrum $h(e^{-i\omega})h(e^{i\omega})s_X(\omega)$, and so, from [6.4.9], the spectrum of Y is

$$s_Y(\omega) = h(e^{-i\omega})h(e^{i\omega})s_X(\omega) + \sigma^2/(2\pi).$$

APPENDIX 6.A. *Proofs of Chapter 6 Propositions*

■ **Proof of Proposition 6.1.** Notice that

$$\begin{aligned}\int_{-\pi}^{\pi} s_Y(\omega)e^{i\omega k}\,d\omega &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{j=-\infty}^{\infty}\gamma_j e^{-i\omega j}e^{i\omega k}\,d\omega \\ &= \frac{1}{2\pi}\sum_{j=-\infty}^{\infty}\gamma_j\int_{-\pi}^{\pi} e^{i\omega(k-j)}\,d\omega \\ &= \frac{1}{2\pi}\sum_{j=-\infty}^{\infty}\gamma_j\int_{-\pi}^{\pi}\{\cos[\omega(k-j)] + i\cdot\sin[\omega(k-j)]\}\,d\omega.\end{aligned} \qquad [6.A.1]$$

Consider the integral in [6.A.1]. For $k = j$, this would be

$$\int_{-\pi}^{\pi} \{\cos[\omega(k - j)] + i\cdot\sin[\omega(k - j)]\}\, d\omega = \int_{-\pi}^{\pi} \{\cos(0) + i\cdot\sin(0)\}\, d\omega$$
$$= \int_{-\pi}^{\pi} d\omega \qquad [6.A.2]$$
$$= 2\pi.$$

For $k \neq j$, the integral in [6.A.1] would be

$$\int_{-\pi}^{\pi} \{\cos[\omega(k - j)] + i\cdot\sin[\omega(k - j)]\}\, d\omega$$
$$= \left.\frac{\sin[\omega(k - j)]}{k - j}\right|_{\omega=-\pi}^{\pi} - i\cdot\left.\frac{\cos[\omega(k - j)]}{k - j}\right|_{\omega=-\pi}^{\pi} \qquad [6.A.3]$$
$$= (k - j)^{-1}\{\sin[\pi(k - j)] - \sin[-\pi(k - j)] - i\cdot\cos[\pi(k - j)] + i\cdot\cos[-\pi(k - j)]\}.$$

But the difference between the frequencies $\pi(k - j)$ and $-\pi(k - j)$ is $2\pi(k - j)$, which is an integer multiple of 2π. Since the sine and cosine functions are periodic, the magnitude in [6.A.3] is zero. Hence, only the term for $j = k$ in the sum in [6.A.1] is nonzero, and using [6.A.2], this sum is seen to be

$$\int_{-\pi}^{\pi} s_Y(\omega)e^{i\omega k}\, d\omega = \frac{1}{2\pi}\gamma_k \int_{-\pi}^{\pi} [\cos(0) + i\cdot\sin(0)]\, d\omega = \gamma_k,$$

as claimed in [6.1.15].

To derive [6.1.16], notice that since $s_Y(\omega)$ is symmetric around $\omega = 0$,

$$\int_{-\pi}^{\pi} s_Y(\omega)e^{i\omega k}\, d\omega = \int_{-\pi}^{0} s_Y(\omega)e^{i\omega k}\, d\omega + \int_{0}^{\pi} s_Y(\omega)e^{i\omega k}\, d\omega$$
$$= \int_{0}^{\pi} s_Y(-\omega)e^{-i\omega k}\, d\omega + \int_{0}^{\pi} s_Y(\omega)e^{i\omega k}\, d\omega$$
$$= \int_{0}^{\pi} s_Y(\omega)(e^{-i\omega k} + e^{i\omega k})\, d\omega$$
$$= \int_{0}^{\pi} s_Y(\omega)\cdot 2\cdot\cos(\omega k)\, d\omega,$$

where the last line follows from [6.1.11]. Again appealing to the symmetry of $s_Y(\omega)$,

$$\int_{0}^{\pi} s_Y(\omega)\cdot 2\cdot\cos(\omega k)\, d\omega = \int_{-\pi}^{\pi} s_Y(\omega)\cos(\omega k)\, d\omega,$$

so that

$$\int_{-\pi}^{\pi} s_Y(\omega)e^{i\omega k}\, d\omega = \int_{-\pi}^{\pi} s_Y(\omega)\cos(\omega k)\, d\omega,$$

as claimed. ■

■ **Derivation of Equation [6.2.11] in Proposition 6.2.** We begin by establishing the following result:

$$\sum_{t=1}^{T} \exp[i(2\pi s/T)(t - 1)] = \begin{cases} T & \text{for } s = 0 \\ 0 & \text{for } s = \pm 1, \pm 2, \ldots, \pm(T - 1). \end{cases} \qquad [6.A.4]$$

That [6.A.4] holds for $s = 0$ is an immediate consequence of the fact that $\exp(0) = 1$. To see that it holds for the other cases in [6.A.4], define

$$z \equiv \exp[i(2\pi s/T)]. \qquad [6.A.5]$$

Then the expression to be evaluated in [6.A.4] can be written

$$\sum_{t=1}^{T} \exp[i(2\pi s/T)(t-1)] = \sum_{t=1}^{T} z^{(t-1)}. \qquad [6.A.6]$$

We now show that for any N,

$$\sum_{t=1}^{N} z^{(t-1)} = \frac{1-z^N}{1-z}, \qquad [6.A.7]$$

provided that $z \neq 1$, which is the case whenever $0 < |s| < T$. Expression [6.A.7] can be verified by induction. Clearly, it holds for $N = 1$, for then

$$\sum_{t=1}^{N} z^{(t-1)} = z^{(0)} = 1.$$

Given that [6.A.7] holds for N, we see that

$$\begin{aligned}
\sum_{t=1}^{N+1} z^{(t-1)} &= \sum_{t=1}^{N} z^{(t-1)} + z^N \\
&= \frac{1-z^N}{1-z} + z^N \\
&= \frac{1-z^N + z^N(1-z)}{1-z} \\
&= \frac{1-z^{N+1}}{1-z},
\end{aligned}$$

as claimed in [6.A.7].

Setting $N = T$ in [6.A.7] and substituting the result into [6.A.6], we see that

$$\sum_{t=1}^{T} \exp[i(2\pi s/T)(t-1)] = \frac{1-z^T}{1-z} \qquad [6.A.8]$$

for $0 < |s| < T$. But it follows from the definition of z in [6.A.5] that

$$\begin{aligned}
z^T &= \exp[i(2\pi s/T)\cdot T] \\
&= \exp[i(2\pi s)] \\
&= \cos(2\pi s) + i\cdot\sin(2\pi s) \\
&= 1 \qquad \text{for } s = \pm 1, \pm 2, \ldots, \pm(T-1).
\end{aligned} \qquad [6.A.9]$$

Substituting [6.A.9] into [6.A.8] produces

$$\sum_{t=1}^{T} \exp[i(2\pi s/T)(t-1)] = 0 \qquad \text{for } s = \pm 1, \pm 2, \ldots, \pm(T-1),$$

as claimed in [6.A.4].

To see how [6.A.4] can be used to deduce expression [6.2.11], notice that the first column of $\Sigma_{t=1}^{T} \mathbf{x}_t \mathbf{x}_t'$ is given by

$$\begin{bmatrix}
T \\
\sum \cos[\omega_1(t-1)] \\
\sum \sin[\omega_1(t-1)] \\
\vdots \\
\sum \cos[\omega_M(t-1)] \\
\sum \sin[\omega_M(t-1)]
\end{bmatrix}, \qquad [6.A.10]$$

where Σ indicates summation over t from 1 to T. The first row of $\Sigma_{t=1}^{T} \mathbf{x}_t \mathbf{x}_t'$ is the transpose of [6.A.10]. To show that all the terms in [6.A.10] other than the first element are zero, we must show that

$$\sum_{t=1}^{T} \cos[\omega_j(t-1)] = 0 \qquad \text{for } j = 1, 2, \ldots, M \tag{6.A.11}$$

and

$$\sum_{t=1}^{T} \sin[\omega_j(t-1)] = 0 \qquad \text{for } j = 1, 2, \ldots, M \tag{6.A.12}$$

for ω_j the frequencies specified in [6.2.7]. But [6.A.4] establishes that

$$\begin{aligned} 0 &= \sum_{t=1}^{T} \exp[i(2\pi j/T)(t-1)] \\ &= \sum_{t=1}^{T} \cos[(2\pi j/T)(t-1)] + i\cdot\sum_{t=1}^{T} \sin[(2\pi j/T)(t-1)] \end{aligned} \tag{6.A.13}$$

for $j = 1, 2, \ldots, M$. For [6.A.13] to equal zero, both the real and the imaginary component must equal zero. Since $\omega_j = 2\pi j/T$, results [6.A.11] and [6.A.12] follow immediately from [6.A.13].

Result [6.A.4] can also be used to calculate the other elements of $\Sigma_{t=1}^{T}\mathbf{x}_t\mathbf{x}_t'$. To see how, note that

$$\begin{aligned} \frac{1}{2}[e^{i\theta} + e^{-i\theta}] &= \frac{1}{2}[\cos(\theta) + i\cdot\sin(\theta) + \cos(\theta) - i\cdot\sin(\theta)] \\ &= \cos(\theta) \end{aligned} \tag{6.A.14}$$

and similarly

$$\begin{aligned} \frac{1}{2i}[e^{i\theta} - e^{-i\theta}] &= \frac{1}{2i}\{\cos(\theta) + i\cdot\sin(\theta) - [\cos(\theta) - i\cdot\sin(\theta)]\} \\ &= \sin(\theta). \end{aligned} \tag{6.A.15}$$

Thus, for example, the elements of $\Sigma_{t=1}^{T}\mathbf{x}_t\mathbf{x}_t'$ corresponding to products of the cosine terms can be calculated as

$$\begin{aligned} &\sum_{t=1}^{T} \cos[\omega_j(t-1)]\cdot\cos[\omega_k(t-1)] \\ &\quad = \frac{1}{4}\sum_{t=1}^{T} \{\exp[i\omega_j(t-1)] + \exp[-i\omega_j(t-1)]\} \\ &\qquad \times \{\exp[i\omega_k(t-1)] + \exp[-i\omega_k(t-1)]\} \\ &\quad = \frac{1}{4}\sum_{t=1}^{T} \{\exp[i(\omega_j + \omega_k)(t-1)] + \exp[i(-\omega_j + \omega_k)(t-1)] \\ &\qquad + \exp[i(\omega_j - \omega_k)(t-1)] + \exp[i(-\omega_j - \omega_k)(t-1)]\} \\ &\quad = \frac{1}{4}\sum_{t=1}^{T} \{\exp[i(2\pi/T)(j+k)(t-1)] + \exp[i(2\pi/T)(k-j)(t-1)] \\ &\qquad + \exp[i(2\pi/T)(j-k)(t-1)] + \exp[i(2\pi/T)(-j-k)(t-1)]\}. \end{aligned} \tag{6.A.16}$$

For any $j = 1, 2, \ldots, M$ and any $k = 1, 2, \ldots, M$ where $k \neq j$, expression [6.A.16] is zero by virtue of [6.A.4]. For $k = j$, the first and last sums in the last line of [6.A.16] are zero, so that the total is equal to

$$(1/4)\sum_{t=1}^{T} (1 + 1) = T/2.$$

Similarly, elements of $\Sigma_{t=1}^{T}\mathbf{x}_t\mathbf{x}_t'$ corresponding to cross products of the sine terms can be found from

$$\sum_{t=1}^{T}\sin[\omega_j(t-1)]\cdot\sin[\omega_k(t-1)]$$

$$= -\frac{1}{4}\sum_{t=1}^{T}\{\exp[i\omega_j(t-1)] - \exp[-i\omega_j(t-1)]\}$$
$$\times\{\exp[i\omega_k(t-1)] - \exp[-i\omega_k(t-1)]\}$$
$$= -\frac{1}{4}\sum_{t=1}^{T}\{\exp[i(2\pi/T)(j+k)(t-1)] - \exp[i(2\pi/T)(k-j)(t-1)]$$
$$- \exp[i(2\pi/T)(j-k)(t-1)] + \exp[i(2\pi/T)(-j-k)(t-1)]\}$$
$$= \begin{cases} T/2 & \text{for } j = k \\ 0 & \text{otherwise.}\end{cases}$$

Finally, elements of $\Sigma_{t=1}^{T}\mathbf{x}_t\mathbf{x}_t'$ corresponding to cross products of the sine and cosine terms are given by

$$\sum_{t=1}^{T}\cos[\omega_j(t-1)]\cdot\sin[\omega_k(t-1)]$$

$$= \frac{1}{4i}\sum_{t=1}^{T}\{\exp[i\omega_j(t-1)] + \exp[-i\omega_j(t-1)]\}$$
$$\times\{\exp[i\omega_k(t-1)] - \exp[-i\omega_k(t-1)]\}$$
$$= \frac{1}{4i}\sum_{t=1}^{T}\{\exp[i(2\pi/T)(j+k)(t-1)] + \exp[i(2\pi/T)(k-j)(t-1)]$$
$$- \exp[i(2\pi/T)(j-k)(t-1)] - \exp[i(2\pi/T)(-j-k)(t-1)]\},$$

which equals zero for all j and k. This completes the derivation of [6.2.11]. ■

■ **Proof of Proposition 6.2(a).** Let **b** denote the estimate of $\boldsymbol{\beta}$ based on *OLS* estimation of the regression in [6.2.8]:

$$\mathbf{b} = \left\{\sum_{t=1}^{T}\mathbf{x}_t\mathbf{x}_t'\right\}^{-1}\left\{\sum_{t=1}^{T}\mathbf{x}_t y_t\right\}$$
$$= \begin{bmatrix} T & \mathbf{0}' \\ \mathbf{0} & (T/2)\cdot\mathbf{I}_{T-1}\end{bmatrix}^{-1}\left\{\sum_{t=1}^{T}\mathbf{x}_t y_t\right\} \qquad [6.A.17]$$
$$= \begin{bmatrix} T^{-1} & \mathbf{0}' \\ \mathbf{0} & (2/T)\cdot\mathbf{I}_{T-1}\end{bmatrix}\left\{\sum_{t=1}^{T}\mathbf{x}_t y_t\right\}.$$

But the definition of $\mathbf{x}_t$ in [6.2.9] implies that

$$\sum_{t=1}^{T}\mathbf{x}_t y_t = \Big[\sum y_t \quad \sum y_t\cdot\cos[\omega_1(t-1)] \quad \sum y_t\cdot\sin[\omega_1(t-1)]$$
$$\sum y_t\cdot\cos[\omega_2(t-1)] \quad \sum y_t\cdot\sin[\omega_2(t-1)] \quad \cdots \qquad [6.A.18]$$
$$\sum y_t\cdot\cos[\omega_M(t-1)] \quad \sum y_t\cdot\sin[\omega_M(t-1)]\Big]',$$

where Σ again denotes summation over t from 1 to T. Substituting [6.A.18] into [6.A.17] produces result (a) of Proposition 6.2. ■

■ **Proof of Proposition 6.2(b).** Recall from expression [4.A.6] that the residual sum of squares associated with *OLS* estimation of [6.2.8] is

$$\sum_{t=1}^{T}\hat{u}_t^2 = \sum_{t=1}^{T}y_t^2 - \left[\sum_{t=1}^{T}y_t\mathbf{x}_t'\right]\left[\sum_{t=1}^{T}\mathbf{x}_t\mathbf{x}_t'\right]^{-1}\left[\sum_{t=1}^{T}\mathbf{x}_t y_t\right]. \qquad [6.A.19]$$

Since there are as many explanatory variables as observations and since the explanatory variables are linearly independent, the OLS residuals $\hat{u}_t$ are all zero. Hence, [6.A.19] implies that

$$\sum_{t=1}^{T} y_t^2 = \left[\sum_{t=1}^{T} y_t \mathbf{x}_t'\right]\left[\sum_{t=1}^{T} \mathbf{x}_t \mathbf{x}_t'\right]^{-1}\left[\sum_{t=1}^{T} \mathbf{x}_t y_t\right]. \qquad [6.A.20]$$

But [6.A.17] allows us to write

$$\sum_{t=1}^{T} \mathbf{x}_t y_t = \begin{bmatrix} T & \mathbf{0}' \\ \mathbf{0} & (T/2)\cdot\mathbf{I}_{T-1} \end{bmatrix}\mathbf{b}. \qquad [6.A.21]$$

Substituting [6.A.21] and [6.2.11] into [6.A.20] establishes that

$$\begin{aligned}
\sum_{t=1}^{T} y_t^2 &= \mathbf{b}'\begin{bmatrix} T & \mathbf{0}' \\ \mathbf{0} & (T/2)\cdot\mathbf{I}_{T-1} \end{bmatrix}\begin{bmatrix} T & \mathbf{0}' \\ \mathbf{0} & (T/2)\cdot\mathbf{I}_{T-1} \end{bmatrix}^{-1}\begin{bmatrix} T & \mathbf{0}' \\ \mathbf{0} & (T/2)\cdot\mathbf{I}_{T-1} \end{bmatrix}\mathbf{b} \\
&= \mathbf{b}'\begin{bmatrix} T & \mathbf{0}' \\ \mathbf{0} & (T/2)\cdot\mathbf{I}_{T-1} \end{bmatrix}\mathbf{b} \\
&= T\cdot\hat{\mu}^2 + (T/2)\sum_{j=1}^{M}(\hat{\alpha}_j^2 + \hat{\delta}_j^2)
\end{aligned}$$

so that

$$(1/T)\sum_{t=1}^{T} y_t^2 = \hat{\mu}^2 + (1/2)\sum_{j=1}^{M}(\hat{\alpha}_j^2 + \hat{\delta}_j^2). \qquad [6.A.22]$$

Finally, observe from [4.A.5] and the fact that $\hat{\mu} = \bar{y}$ that

$$(1/T)\sum_{t=1}^{T} y_t^2 - \hat{\mu}^2 = (1/T)\sum_{t=1}^{T}(y_t - \bar{y})^2,$$

allowing [6.A.22] to be written as

$$(1/T)\sum_{t=1}^{T}(y_t - \bar{y})^2 = (1/2)\sum_{j=1}^{M}(\hat{\alpha}_j^2 + \hat{\delta}_j^2),$$

as claimed in [6.2.14]. Since the regressors are all orthogonal, the term $\frac{1}{2}(\hat{\alpha}_j^2 + \hat{\delta}_j^2)$ can be interpreted as the portion of the sample variance that can be attributed to the regressors $\cos[\omega_j(t-1)]$ and $\sin[\omega_j(t-1)]$. ■

■ **Proof of Proposition 6.2(c).** Notice that

$$(\hat{\alpha}_j^2 + \hat{\delta}_j^2) = (\hat{\alpha}_j + i\cdot\hat{\delta}_j)(\hat{\alpha}_j - i\cdot\hat{\delta}_j). \qquad [6.A.23]$$

But from result (a) of Proposition 6.2,

$$\hat{\alpha}_j = (2/T)\sum_{t=1}^{T} y_t\cdot\cos[\omega_j(t-1)] = (2/T)\sum_{t=1}^{T}(y_t - \bar{y})\cdot\cos[\omega_j(t-1)], \qquad [6.A.24]$$

where the second equality follows from [6.A.11]. Similarly,

$$\hat{\delta}_j = (2/T)\sum_{t=1}^{T}(y_t - \bar{y})\cdot\sin[\omega_j(t-1)]. \qquad [6.A.25]$$

It follows from [6.A.24] and [6.A.25] that

$$\begin{aligned}
\hat{\alpha}_j + i\cdot\hat{\delta}_j &= (2/T)\left\{\sum_{t=1}^{T}(y_t - \bar{y})\cdot\cos[\omega_j(t-1)]\right. \\
&\qquad \left. + i\cdot\sum_{t=1}^{T}(y_t - \bar{y})\cdot\sin[\omega_j(t-1)]\right\} \qquad [6.A.26] \\
&= (2/T)\sum_{t=1}^{T}(y_t - \bar{y})\cdot\exp[i\omega_j(t-1)].
\end{aligned}$$

Similarly,

$$\hat{\alpha}_j - i\cdot\hat{\delta}_j = (2/T)\sum_{\tau=1}^{T}(y_\tau - \bar{y})\cdot\exp[-i\omega_j(\tau - 1)]. \qquad [6.A.27]$$

Substituting [6.A.26] and [6.A.27] into [6.A.23] produces

$$\begin{aligned}
\hat{\alpha}_j^2 + \hat{\delta}_j^2 &= (4/T^2)\left\{\sum_{t=1}^{T}(y_t - \bar{y})\cdot\exp[i\omega_j(t-1)]\right\} \\
&\quad \times \left\{\sum_{\tau=1}^{T}(y_\tau - \bar{y})\cdot\exp[-i\omega_j(\tau-1)]\right\} \\
&= (4/T^2)\sum_{t=1}^{T}\sum_{\tau=1}^{T}(y_t - \bar{y})(y_\tau - \bar{y})\cdot\exp[i\omega_j(t-\tau)] \\
&= (4/T^2)\left\{\sum_{t=1}^{T}(y_t - \bar{y})^2 + \sum_{t=1}^{T-1}(y_t - \bar{y})(y_{t+1} - \bar{y})\cdot\exp[-i\omega_j]\right. \\
&\quad + \sum_{t=2}^{T}(y_t - \bar{y})(y_{t-1} - \bar{y})\cdot\exp[i\omega_j] \\
&\quad + \sum_{t=1}^{T-2}(y_t - \bar{y})(y_{t+2} - \bar{y})\cdot\exp[-2i\omega_j] \\
&\quad + \sum_{t=3}^{T}(y_t - \bar{y})(y_{t-2} - \bar{y})\cdot\exp[2i\omega_j] + \cdots \qquad [6.A.28] \\
&\quad + (y_1 - \bar{y})(y_T - \bar{y})\cdot\exp[-(T-1)i\omega_j] \\
&\quad \left. + (y_T - \bar{y})(y_1 - \bar{y})\cdot\exp[(T-1)i\omega_j]\right\} \\
&= (4/T)\left\{\hat{\gamma}_0 + \hat{\gamma}_1\cdot\exp[-i\omega_j] + \hat{\gamma}_{-1}\cdot\exp[i\omega_j]\right. \\
&\quad + \hat{\gamma}_2\cdot\exp[-2i\omega_j] + \hat{\gamma}_{-2}\cdot\exp[2i\omega_j] + \cdots \\
&\quad \left. + \hat{\gamma}_{T-1}\cdot\exp[-(T-1)i\omega_j] + \hat{\gamma}_{-T+1}\cdot\exp[(T-1)i\omega_j]\right\} \\
&= (4/T)(2\pi)\hat{s}_y(\omega_j),
\end{aligned}$$

from which equation [6.2.15] follows. ■

Chapter 6 Exercises

6.1. Derive [6.1.12] directly from expression [6.1.6] and the formulas for the autocovariances of an $MA(1)$ process.

6.2. Integrate [6.1.9] and [6.1.12] to confirm independently that [6.1.17] holds for white noise and an $MA(1)$ process.

Chapter 6 References

Anderson, T. W. 1971. *The Statistical Analysis of Time Series*. New York: Wiley.

Bloomfield, Peter. 1976. *Fourier Analysis of Time Series: An Introduction*. New York: Wiley.

Cramér, Harald, and M. R. Leadbetter. 1967. *Stationary and Related Stochastic Processes*. New York: Wiley.

Fuller, Wayne A. 1976. *Introduction to Statistical Time Series*. New York: Wiley.

Thomas, George B., Jr. 1972. *Calculus and Analytic Geometry*, alternate ed. Reading, Mass.: Addison-Wesley.