TIME-SERIES MODELS

21.1 INTRODUCTION

For forecasting purposes, a simple model that *describes* the behavior of a variable (or a set of variables) in terms of past values, without the benefit of a well-developed theory, may well prove quite satisfactory. Researchers have observed that the large simultaneous equations macroeconomic models constructed in the 1960s frequently have poorer forecasting performance than fairly simple, univariate time-series models based on just a few parameters and compact specifications. It is just this observation that has raised to prominence the univariate time-series forecasting models pioneered by Box and Jenkins (1984).

In this chapter, we introduce some of the tools employed in the analysis of time-series data.[1] Section 21.2 describes stationary stochastic processes. We encountered this body of theory in Chapters 19 and 20, where we discovered that certain assumptions were required to ascribe familiar properties to a time series of data. We continue that discussion by defining several characteristics of a stationary time series. The recent literature in macroeconometrics has seen an explosion of studies of nonstationary time series. Nonstationarity mandates a revision of the standard inference tools we have used thus far. Chapter 22 introduces some extensions of the results of this chapter to nonstationary time series.

Some of the concepts to be discussed here were introduced in Section 19.2. Section 19.2 also contains a cursory introduction to the nature of time-series processes. It will be useful to review that material before proceeding with the rest of this chapter. Finally, Sections 13.6 on estimation and 13.9.2 and 20.4.3 on stability of dynamic models will be especially useful for the latter sections of this chapter.

[1] Each topic discussed here is the subject of a vast literature with articles and book-length treatments at all levels. For example, two survey papers on the subject of unit roots in economic time-series data, Diebold and Nerlove (1990) and Campbell and Perron (1991), cite between them more than 200 basic sources on the subject. The literature on unit roots and cointegration is almost surely the most rapidly moving target in econometrics. Stock's (1994) survey adds hundreds of references to those in the aforementioned surveys and brings the literature up to date as of then. Useful basic references on the subjects of this chapter are Box and Jenkins (1984); Judge et al. (1985); Mills (1990); Granger and Newbold (1996); Granger and Watson (1984); Hendry, Pagan, and Sargan (1984); Geweke (1984); and especially Harvey (1989, 1990); Enders (2004); Tsay (2005); Hamilton (1994); and Patterson (2000). There are also many survey style and pedagogical articles on these subjects. The aforementioned paper by Diebold and Nerlove is a useful tour guide through some of the literature. We recommend Dickey, Bell, and Miller (1986) and Dickey, Jansen, and Thorton (1991) as well. The latter is an especially clear introduction at a very basic level of the fundamental tools for empirical researchers.

21.2 STATIONARY STOCHASTIC PROCESSES

The essential building block for the models to be discussed in this chapter is the **white noise** time-series process,

$$\{\varepsilon_t\}, t = -\infty, +\infty,$$

where each element in the sequence has $E[\varepsilon_t] = 0$, $E[\varepsilon_t^2] = \sigma_\varepsilon^2$, and $\text{Cov}[\varepsilon_t, \varepsilon_s] = 0$ for all $s \neq t$. Each element in the series is a random draw from a population with zero mean and constant variance. It is occasionally assumed that the draws are independent or normally distributed, although for most of our analysis, neither assumption will be essential.

A **univariate time-series model** describes the behavior of a variable in terms of its own past values. Consider, for example, the autoregressive disturbance models introduced in Chapter 20.

$$u_t = \rho u_{t-1} + \varepsilon_t. \tag{21-1}$$

Autoregressive disturbances are generally the residual variation in a regression model built up from what may be an elaborate underlying theory, $y_t = \mathbf{x}_t'\boldsymbol{\beta} + u_t$. The theory usually stops short of stating what enters the disturbance. But the presumption that some time-series process generates $\mathbf{x}_t$ should extend equally to u_t. There are two ways to interpret this simple series. As stated, u_t equals the previous value of u_t plus an "innovation," ε_t. Alternatively, by manipulating the series, we showed that u_t could be interpreted as an aggregation of the entire history of the ε_t's.

Occasionally, statistical evidence is convincing that a more intricate process is at work in the disturbance. Perhaps a second-order **autoregression**,

$$u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \varepsilon_t, \tag{21-2}$$

better explains the movement of the disturbances in the regression. The model may not arise naturally from an underlying behavioral theory. But in the face of certain kinds of statistical evidence, one might conclude that the more elaborate model would be preferable.[2] This section will describe several alternatives to the AR(1) model that we have relied on in most of the preceding applications.

21.2.1 AUTOREGRESSIVE MOVING-AVERAGE PROCESSES

The variable y_t in the model

$$y_t = \mu + \gamma y_{t-1} + \varepsilon_t \tag{21-3}$$

is said to be **autoregressive** (or **self-regressive**) because under certain assumptions,

$$E[y_t \mid y_{t-1}] = \mu + \gamma y_{t-1}.$$

A more general pth-order autoregression or AR(p) process would be written

$$y_t = \mu + \gamma_1 y_{t-1} + \gamma_2 y_{t-2} + \cdots + \gamma_p y_{t-p} + \varepsilon_t. \tag{21-4}$$

[2] For example, the estimates of ε_t computed after a correction for first-order autocorrelation may fail tests of randomness such as the LM (Section 19.7.1) test.

The analogy to the classical regression is clear. Now consider the first-order moving average, or MA(1) specification[3]

$$y_t = \mu + \varepsilon_t - \theta\varepsilon_{t-1}. \tag{21-5}$$

By writing

$$y_t = \mu + (1 - \theta L)\varepsilon_t,$$

or

$$\frac{y_t}{1-\theta L} = \frac{\mu}{1-\theta} + \varepsilon_t,$$

we find that

$$y_t = \frac{\mu}{1-\theta} - \theta y_{t-1} - \theta^2 y_{t-2} - \cdots + \varepsilon_t.$$

Once again, the effect is to represent y_t as a function of its own past values.

An extremely general model that encompasses (21-4) and (21-5) is the **autoregressive moving average,** or ARMA(p, q), model:

$$y_t = \mu + \gamma_1 y_{t-1} + \gamma_2 y_{t-2} + \cdots + \gamma_p y_{t-p} + \varepsilon_t - \theta_1\varepsilon_{t-1} - \cdots - \theta_q\varepsilon_{t-q}. \tag{21-6}$$

Note the convention that the ARMA(p, q) process has p autoregressive (lagged dependent-variable) terms and q lagged **moving average** terms. Researchers have found that models of this sort with relatively small values of p and q have proved quite effective as forecasting models.

The disturbances ε_t are labeled the **innovations** in the model. The term is fitting because the only new information that enters the processes in period t is this innovation. Consider, then, the AR(1) process

$$y_t = \mu + \gamma y_{t-1} + \varepsilon_t. \tag{21-7}$$

Either by successive substitution or by using the lag operator, we obtain

$$(1 - \gamma L)y_t = \mu + \varepsilon_t,$$

or

$$y_t = \frac{\mu}{1-\gamma} + \sum_{i=0}^{\infty} \gamma^i \varepsilon_{t-i}.[4] \tag{21-8}$$

The observed series is a particular type of aggregation of the history of the innovations. The moving average, MA(q) model,

$$y_t = \mu + \varepsilon_t - \theta_1\varepsilon_{t-1} - \cdots - \theta_q\varepsilon_{t-q} = \mu + D(L)\varepsilon_t, \tag{21-9}$$

is yet another, particularly simple form of aggregation in that only information from the q most recent periods is retained. The general result is that many time-series processes can be viewed either as regressions on lagged values with additive disturbances or as aggregations of a history of innovations. They differ from one to the next in the form of that aggregation.

[3] The lag operator is discussed in Section 20.2.2. Because μ is a constant, $(1-\theta L)^{-1}\mu = \mu + \theta\mu + \theta^2\mu + \cdots = \mu/(1-\theta)$. The lag operator may be set equal to one when it operates on a constant.

[4] See Section 20.3 for discussion of models with infinite lag structures.

More involved processes can be similarly represented in either an autoregressive or moving-average form. (We will turn to the mathematical requirements later.) Consider, for example, the ARMA(2, 1) process,

$$y_t = \mu + \gamma_1 y_{t-1} + \gamma_2 y_{t-2} + \varepsilon_t - \theta \varepsilon_{t-1},$$

which we can write as

$$(1 - \theta L)\varepsilon_t = y_t - \mu - \gamma_1 y_{t-1} - \gamma_2 y_{t-2}.$$

If $|\theta| < 1$, then we can divide both sides of the equation by $(1 - \theta L)$ and obtain

$$\varepsilon_t = \sum_{i=0}^{\infty} \theta^i (y_{t-i} - \mu - \gamma_1 y_{t-i-1} - \gamma_2 y_{t-i-2}).$$

After some tedious manipulation, this equation produces the autoregressive form,

$$y_t = \frac{\mu}{1-\theta} + \sum_{i=1}^{\infty} \pi_i y_{t-i} + \varepsilon_t,$$

where

$$\pi_1 = \gamma_1 - \theta \quad \text{and} \quad \pi_j = -(\theta^j - \gamma_1 \theta^{j-1} - \gamma_2 \theta^{j-2}), \quad j = 2, 3, \ldots. \tag{21-10}$$

Alternatively, by similar (yet more tedious) manipulation, we can write

$$y_t = \frac{\mu}{1-\gamma_1-\gamma_2} + \left[\frac{1-\theta L}{1-\gamma_1 L - \gamma_2 L^2}\right]\varepsilon_t = \frac{\mu}{1-\gamma_1-\gamma_2} + \sum_{i=0}^{\infty} \delta_i \varepsilon_{t-i}. \tag{21-11}$$

In each case, the weights, π_i in the **autoregressive form** and δ_i in the **moving-average form,** are complicated functions of the original parameters. But nonetheless, each is just an alternative representation of the same time-series process that produces the current value of y_t. This result is a fundamental property of certain time series. We will return to the issue after we formally define the assumption that we have used at the preceding several steps that allows these transformations.

21.2.2 STATIONARITY AND INVERTIBILITY

At several points in the preceding, we have alluded to the notion of **stationarity,** either directly or indirectly by making certain assumptions about the parameters in the model. In Section 19.3.2, we characterized an AR(1) disturbance process

$$u_t = \rho u_{t-1} + \varepsilon_t,$$

as stationary if $|\rho| < 1$ and ε_t is **white noise.** Then

$$E[u_t] = 0 \quad \text{for all } t,$$

$$\text{Var}[u_t] = \frac{\sigma_\varepsilon^2}{1-\rho^2}, \tag{21-12}$$

$$\text{Cov}[u_t, u_s] = \frac{\rho^{|t-s|}\sigma_\varepsilon^2}{1-\rho^2}.$$

If $|\rho| \geq 1$, then the variance and covariances are undefined.

In the following, we use ε_t to denote the white noise innovations in the process. The ARMA(p, q) process will be denoted as in (21-6).

DEFINITION 21.1 Covariance Stationarity

A stochastic process y_t is weakly stationary or covariance stationary if it satisfies the following requirements:[5]

1. $E[y_t]$ *is independent of* t.
2. $\text{Var}[y_t]$ *is a finite, positive constant, independent of* t.
3. $\text{Cov}[y_t, y_s]$ *is a finite function of* $|t-s|$, *but not of* t *or* s.

The third requirement is that the covariance between observations in the series is a function only of how far apart they are in time, not the time at which they occur. These properties clearly hold for the AR(1) process shown earlier. Whether they apply for the other models we have examined remains to be seen.

We define the **autocovariance** at lag k as

$$\lambda_k = \text{Cov}[y_t, y_{t-k}].$$

Note that

$$\lambda_{-k} = \text{Cov}[y_t, y_{t+k}] = \lambda_k.$$

Stationarity implies that autocovariances are a function of k, but not of t. For example, in (21-12), we see that the autocovariances of the AR(1) process $y_t = \mu + \gamma y_{t-1} + \varepsilon_t$ are

$$\text{Cov}[y_t, y_{t-k}] = \frac{\gamma^k \sigma_\varepsilon^2}{1-\gamma^2}, \quad k = 0, 1, \ldots. \tag{21-13}$$

If $|\gamma| < 1$, then this process is stationary. For any MA(q) series,

$$\begin{aligned} y_t &= \mu + \varepsilon_t - \theta_1\varepsilon_{t-1} - \cdots - \theta_q\varepsilon_{t-q}, \\ E[y_t] &= \mu + E[\varepsilon_t] - \theta_1 E[\varepsilon_{t-1}] - \cdots - \theta_q E[\varepsilon_{t-q}] = \mu, \\ \text{Var}[y_t] &= \left(1 + \theta_1^2 + \cdots + \theta_q^2\right)\sigma_\varepsilon^2, \\ \text{Cov}[y_t, y_{t-1}] &= \left(-\theta_1 + \theta_1\theta_2 + \theta_2\theta_3 + \cdots + \theta_{q-1}\theta_q\right)\sigma_\varepsilon^2, \end{aligned} \tag{21-14}$$

and so on until

$$\begin{aligned} \text{Cov}[y_t, y_{t-(q-1)}] &= \left[-\theta_{q-1} + \theta_1\theta_q\right]\sigma_\varepsilon^2, \\ \text{Cov}[y_t, y_{t-q}] &= -\theta_q\sigma_\varepsilon^2, \end{aligned}$$

and, for lags greater than q, the autocovariances are zero. It follows, therefore, that finite moving-average processes are stationary regardless of the values of the parameters. The MA(1) process $y_t = \varepsilon_t - \theta\varepsilon_{t-1}$ is an important special case that has $\text{Var}[y_t] = (1+\theta^2)\sigma_\varepsilon^2$, $\lambda_1 = -\theta\sigma_\varepsilon^2$, and $\lambda_k = 0$ for $|k| > 1$.

For the AR(1) process, the stationarity requirement is that $|\gamma| < 1$, which in turn, implies that the variance of the moving average representation in (21-8) is finite.

[5] *Strong* stationarity requires that the joint distribution of all sets of observations $(y_t, y_{t-1}, \ldots)$ be invariant to when the observations are made. For practical purposes in econometrics, this statement is a theoretical fine point. Although weak stationary suffices for our applications, we would not normally analyze weakly stationary time series that were not strongly stationary as well. Indeed, we often go even beyond this step and assume joint normality.

Consider the AR(2) process

$$y_t = \mu + \gamma_1 y_{t-1} + \gamma_2 y_{t-2} + \varepsilon_t.$$

Write this equation as

$$C(L)y_t = \mu + \varepsilon_t,$$

where

$$C(L) = 1 - \gamma_1 L - \gamma_2 L^2.$$

Then, if it is possible, we invert this result to produce

$$y_t = [C(L)]^{-1}(\mu + \varepsilon_t).$$

Whether the inversion of the polynomial in the lag operator leads to a convergent series depends on the values of γ_1 and γ_2. If so, then the moving-average representation will be

$$y_t = \sum_{i=0}^{\infty} \delta_i(\mu + \varepsilon_{t-i}),$$

so that

$$\text{Var}[y_t] = \sum_{i=0}^{\infty} \delta_i^2 \sigma_\varepsilon^2.$$

Whether this result is finite or not depends on whether the series of δ_is is exploding or converging. For the AR(2) case, the series converges if $|\gamma_2| < 1$, $\gamma_1 + \gamma_2 < 1$, and $\gamma_2 - \gamma_1 < 1$.[6]

For the more general case, the autoregressive process is stationary if the roots of the **characteristic equation,**

$$C(z) = 1 - \gamma_1 z - \gamma_2 z^2 - \cdots - \gamma_p z^p = 0,$$

have modulus greater than one, or "lie outside the **unit circle.**"[7] It follows that if a stochastic process is stationary, it has an infinite moving-average representation (and, if not, it does not). The AR(1) process is the simplest case. The characteristic equation is

$$C(z) = 1 - \gamma z = 0,$$

and its single root is $1/\gamma$. This root lies outside the unit circle if $|\gamma| < 1$, which we saw earlier.

Finally, consider the inversion of the moving-average process in (21-9). Whether this inversion is possible depends on the coefficients in $D(L)$ in the same fashion that stationarity hinges on the coefficients in $C(L)$. This counterpart to stationarity of an autoregressive process is called **invertibility.** For it to be possible to invert a moving-average process to produce an autoregressive representation, the roots of $D(L) = 0$ must be outside the unit circle. Notice, for example, that in (21-5), the inversion of the

[6] This requirement restricts (γ_1, γ_2) to within a triangle with points at $(2, -1)$, $(-2, -1)$, and $(0, 1)$.

[7] The roots may be complex. (See Sections ~~13.9.2 and~~ 20.4.3.) They are of the form $a \pm bi$, where $i = \sqrt{-1}$. The unit circle refers to the two-dimensional set of values of a and b defined by $a^2 + b^2 = 1$, which defines a circle centered at the origin with radius 1.

moving-average process is possible only if $|\theta| < 1$. Because the characteristic equation for the MA(1) process is $1 - \theta L = 0$, the root is $1/\theta$, which must be larger than one.

If the roots of the characteristic equation of a moving-average process all lie outside the unit circle, then the series is said to be invertible. Note that invertibility has no bearing on the stationarity of a process. All moving-average processes with finite coefficients are stationary. Whether an ARMA process is stationary or not depends only on the AR part of the model.

21.2.3 AUTOCORRELATIONS OF A STATIONARY STOCHASTIC PROCESS

The function

$$\lambda_k = \text{Cov}[y_t, y_{t-k}]$$

is called the **autocovariance function** of the process y_t. The **autocorrelation function,** or **ACF,** is obtained by dividing by the variance, λ_0, to obtain

$$\rho_k = \frac{\lambda_k}{\lambda_0}, \quad -1 \leq \rho_k \leq 1.$$

For a stationary process, the ACF will be a function of k and the parameters of the process. The ACF is a useful device for describing a time-series process in much the same way that the moments are used to describe the distribution of a random variable. One of the characteristics of a stationary stochastic process is an autocorrelation function that either abruptly drops to zero at some finite lag or eventually tapers off to zero. The AR(1) process provides the simplest example, because

$$\rho_k = \gamma^k,$$

which is a geometric series that either declines monotonically from $\rho_0 = 1$ if γ is positive or with a damped sawtooth pattern if γ is negative. Note as well that for the process $y_t = \gamma y_{t-1} + \varepsilon_t$,

$$\rho_k = \gamma \rho_{k-1}, \quad k \geq 1,$$

which bears a noteworthy resemblance to the process itself.

For higher-order autoregressive series, the autocorrelations may decline monotonically or may progress in the fashion of a damped sine wave.[8] Consider, for example, the second-order autoregression, where we assume without loss of generality that $\mu = 0$ (because we are examining second moments in deviations from the mean):

$$y_t = \gamma_1 y_{t-1} + \gamma_2 y_{t-2} + \varepsilon_t.$$

If the process is stationary, then $\text{Var}[y_t] = \text{Var}[y_{t-s}]$ for all s. Also, $\text{Var}[y_t] = \text{Cov}[y_t, y_t]$, and $\text{Cov}[\varepsilon_t, y_{t-s}] = 0$ if $s > 0$. These relationships imply that

$$\lambda_0 = \gamma_1 \lambda_1 + \gamma_2 \lambda_2 + \sigma_\varepsilon^2.$$

[8] The behavior is a function of the roots of the characteristic equation. This aspect is discussed in Section ~~13.9 and especially 13.9.2.~~ 21.4.3.

Now, using additional lags, we find that

$$\lambda_1 = \gamma_1\lambda_0 + \gamma_2\lambda_1,$$

and

$$\lambda_2 = \gamma_1\lambda_1 + \gamma_2\lambda_0. \qquad (22\text{-}15)$$

These three equations provide the solution:

$$\lambda_0 = \sigma_\varepsilon^2 \frac{[(1-\gamma_2)/(1+\gamma_2)]}{(1-\gamma_2)^2 - \gamma_1^2}.$$

The variance is unchanging, so we can divide throughout by λ_0 to obtain the relationships for the autocorrelations,

$$\rho_1 = \gamma_1\rho_0 + \gamma_2\rho_1.$$

Because $\rho_0 = 1$, $\rho_1 = \gamma_1/(1-\gamma_2)$. Using the same procedure for additional lags, we find that

$$\rho_2 = \gamma_1\rho_1 + \gamma_2,$$

so $\rho_2 = \gamma_1^2/(1-\gamma_2) + \gamma_2$. Generally, then, for lags of two or more,

$$\rho_k = \gamma_1\rho_{k-1} + \gamma_2\rho_{k-2}.$$

Once again, the autocorrelations follow the same difference equation as the series itself. The behavior of this function depends on γ_1, γ_2, and k, although not in an obvious way. The inherent behavior of the autocorrelation function can be deduced from the characteristic equation.[9] For the second-order process we are examining, the autocorrelations are of the form

$$\rho_k = \phi_1(1/z_1)^k + \phi_2(1/z_2)^k,$$

where the two roots are[10]

$$1/z = \tfrac{1}{2}\left[\gamma_1 \pm \sqrt{\gamma_1^2 + 4\gamma_2}\right].$$

If the two roots are real, then we know that their reciprocals will be less than one in absolute value, so that ρ_k will be the sum of two terms that are decaying to zero. If the two roots are complex, then ρ_k will be the sum of two terms that are oscillating in the form of a damped sine wave.

Applications that involve autoregressions of order greater than two are relatively unusual. Nonetheless, higher-order models can be handled in the same fashion. For the AR(p) process

$$y_t = \gamma_1 y_{t-1} + \gamma_2 y_{t-2} + \cdots + \gamma_p y_{t-p} + \varepsilon_t,$$

the autocovariances will obey the **Yule–Walker equations**

$$\lambda_0 = \gamma_1\lambda_1 + \gamma_2\lambda_2 + \cdots + \gamma_p\lambda_p + \sigma_\varepsilon^2,$$
$$\lambda_1 = \gamma_1\lambda_0 + \gamma_2\lambda_1 + \cdots + \gamma_p\lambda_{p-1},$$

[9] The set of results that we would use to derive this result are exactly those we used in Section 20.4.3 to analyze the stability of a dynamic equation, which makes sense, of course, because the equation linking the autocorrelations is a simple difference equation.

[10] We used the device in Section 20.4.3 to find the characteristic roots. For a second-order equation, the quadratic is easy to manipulate.

22-9

and so on. The autocorrelations will once again follow the same difference equation as the original series,

$$\rho_k = \gamma_1\rho_{k-1} + \gamma_2\rho_{k-2} + \cdots + \gamma_p\rho_{k-p}.$$

The ACF for a moving-average process is very simple to obtain. For the first-order process,

$$\begin{aligned} y_t &= \varepsilon_t - \theta\varepsilon_{t-1}, \\ \lambda_0 &= (1+\theta^2)\sigma_\varepsilon^2, \\ \lambda_1 &= -\theta\sigma_\varepsilon^2, \end{aligned}$$

then $\lambda_k = 0$ for $k > 1$. Higher-order processes appear similarly. For the MA(2) process, by multiplying out the terms and taking expectations, we find that

$$\begin{aligned} \lambda_0 &= \left(1+\theta_1^2+\theta_2^2\right)\sigma_\varepsilon^2, \\ \lambda_1 &= (-\theta_1 + \theta_1\theta_2)\sigma_\varepsilon^2, \\ \lambda_2 &= -\theta_2\sigma_\varepsilon^2, \\ \lambda_k &= 0, \quad k > 2. \end{aligned}$$

The pattern for the general MA(q) process $y_t = \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \cdots - \theta_q\varepsilon_{t-q}$ is analogous. The signature of a moving-average process is an autocorrelation function that abruptly drops to zero at one lag past the order of the process. As we will explore later, this sharp distinction provides a statistical tool that will help us distinguish between these two types of processes empirically.

The mixed process, ARMA(p, q), is more complicated because it is a mixture of the two forms. For the ARMA(1, 1) process

$$y_t = \gamma y_{t-1} + \varepsilon_t - \theta\varepsilon_{t-1},$$

the Yule–Walker equations are

$$\begin{aligned} \lambda_0 &= E[y_t(\gamma y_{t-1} + \varepsilon_t - \theta\varepsilon_{t-1})] = \gamma\lambda_1 + \sigma_\varepsilon^2 - \sigma_\varepsilon^2(\theta\gamma - \theta^2), \\ \lambda_1 &= \gamma\lambda_0 - \theta\sigma_\varepsilon^2, \end{aligned}$$

and

$$\lambda_k = \gamma\lambda_{k-1}, \quad k > 1.$$

The general characteristic of ARMA processes is that when the moving-average component is of order q, then in the series of autocorrelations there will be an initial q terms that are complicated functions of both the AR and MA parameters, but after q periods,

$$\rho_k = \gamma_1\rho_{k-1} + \gamma_2\rho_{k-2} + \cdots + \gamma_p\rho_{k-p}, \quad k > q.$$

22

21.2.4 PARTIAL AUTOCORRELATIONS OF A STATIONARY STOCHASTIC PROCESS

The autocorrelation function ACF(k) gives the gross correlation between y_t and y_{t-k}. But as we saw in our analysis of the classical regression model in Section 3.4, a gross correlation such as this one can mask a completely different underlying relationship. In

this setting, we observe, for example, that a correlation between y_t and y_{t-2} could arise primarily because both variables are correlated with y_{t-1}. Consider the AR(1) process $y_t = \gamma y_{t-1} + \varepsilon_t$, where $E[\varepsilon_t] = 0$ so $E[y_t] = E[y_t]/(1-\gamma) = 0$. The second gross autocorrelation is $\rho_2 = \gamma^2$. But in the same spirit, we might ask what is the correlation between y_t and y_{t-2} *net of the intervening effect of* y_{t-1}? In this model, if we remove the effect of y_{t-1} from y_t, then only ε_t remains, and this disturbance is uncorrelated with y_{t-2}. We would conclude that the **partial autocorrelation** between y_t and y_{t-2} in this model is zero.

22

DEFINITION 21.2 Partial Autocorrelation Coefficient

The partial correlation between y_t and y_{t-k} is the simple correlation between y_{t-k} and y_t minus that part explained linearly by the intervening lags. That is,

$$\rho_k^* = \text{Corr}[y_t - E^*(y_t \mid y_{t-1}, \ldots, y_{t-k+1}), y_{t-k}],$$

where $E^*(y_t \mid y_{t-1}, \ldots, y_{t-k+1})$ is the minimum mean-squared error predictor of y_t by $y_{t-1}, \ldots, y_{t-k+1}$.

The function $E^*(.)$ might be the linear regression if the conditional mean happened to be linear, but it might not. The optimal *linear* predictor *is* the linear regression, however, so what we have is

$$\rho_k^* = \text{Corr}[y_t - \beta_1 y_{t-1} - \beta_2 y_{t-2} - \cdots - \beta_{k-1} y_{t-k+1}, y_{t-k}],$$

where $\boldsymbol{\beta} = [\beta_1, \beta_2, \ldots, \beta_{k-1}] = \{\text{Var}[y_{t-1}, y_{t-2}, \ldots, y_{t-k+1}]\}^{-1} \times \text{Cov}[y_t, (y_{t-1}, y_{t-2}, \ldots, y_{t-k+1})]'$. This equation will be recognized as a vector of regression coefficients. As such, what we are computing here (of course) is the correlation between a vector of residuals and y_{t-k}. There are various ways to formalize this computation [see, e.g., Enders (2004)]. One intuitively appealing approach is suggested by the equivalent definition (which is also a prescription for computing it), as follows.

22

DEFINITION 21.3 Partial Autocorrelation Coefficient

The partial correlation between y_t and y_{t-k} is the last coefficient in the linear projection of y_t on $[y_{t-1}, y_{t-2}, \ldots, y_{t-k}]$,

$$\begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_{k-1} \\ \rho_k^* \end{bmatrix} = \begin{bmatrix} \lambda_0 & \lambda_1 & \cdots & \lambda_{k-2} & \lambda_{k-1} \\ \lambda_1 & \lambda_0 & \cdots & \lambda_{k-3} & \lambda_{k-2} \\ & \cdots & \vdots & \cdots & \\ \lambda_{k-1} & \lambda_{k-2} & \cdots & \lambda_1 & \lambda_0 \end{bmatrix}^{-1} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_k \end{bmatrix}.$$

22 -11

As before, there are some distinctive patterns for particular time-series processes. Consider first the autoregressive processes,

$$y_t = \gamma_1 y_{t-1} + \gamma_2 y_{t-2} + \cdots + \gamma_p y_{t-p} + \varepsilon_t.$$

We are interested in the last coefficient in the projection of y_t on y_{t-1}, then on $[y_{t-1}, y_{t-2}]$, and so on. The first of these is the simple regression coefficient of y_t on y_{t-1}, so

$$\rho_1^* = \frac{\text{Cov}[y_t, y_{t-1}]}{\text{Var}[y_{t-1}]} = \frac{\lambda_1}{\lambda_0} = \rho_1.$$

The first partial autocorrelation coefficient for any process equals the first autocorrelation coefficient.

Without doing the messy algebra, we also observe that for the AR(p) process, ρ_1^* is a mixture of all the γ coefficients. Of course, if p equals 1, then $\rho_1^* = \rho_1 = \gamma$. For the higher-order processes, the autocorrelations are likewise mixtures of the autoregressive coefficients until we reach ρ_p^*. In view of the form of the AR(p) model, the last coefficient in the linear projection on p lagged values is γ_p. Also, we can see the signature pattern of the AR(p) process, any additional partial autocorrelations must be zero, because they will be simply $\rho_k^* = \text{Corr}[\varepsilon_t, y_{t-k}] = 0$ if $k > p$.

Combining results thus far, we have the characteristic pattern for an autoregressive process. The ACF, ρ_k, will gradually decay to zero, either monotonically if the characteristic roots are real or in a sinusoidal pattern if they are complex. The PACF, ρ_k^*, will be irregular out to lag p, when they abruptly drop to zero and remain there.

The moving-average process has the mirror image of this pattern. We have already examined the ACF for the MA(q) process; it has q irregular spikes, then it falls to zero and stays there. For the PACF, write the model as

$$y_t = (1 - \theta_1 L - \theta_2 L^2 - \cdots - \theta_q L^q)\varepsilon_t.$$

If the series is invertible, which we will assume throughout, then we have

$$\frac{y_t}{1 - \theta_1 L - \cdots - \theta_q L^q} = \varepsilon_t,$$

or

$$\begin{aligned} y_t &= \pi_1 y_{t-1} + \pi_2 y_{t-2} + \cdots + \varepsilon_t \\ &= \sum_{i=1}^{\infty} \pi_i y_{t-i} + \varepsilon_t. \end{aligned}$$

The autoregressive form of the MA(q) process has an infinite number of terms, which means that the PACF will not fall off to zero the way that the PACF of the AR process does. Rather, the PACF of an MA process will resemble the ACF of an AR process. For example, for the MA(1) process $y_t = \varepsilon_t - \theta\varepsilon_{t-1}$, the AR representation is

$$y_t = \theta y_{t-1} + \theta^2 y_{t-2} + \cdots + \varepsilon_t,$$

which is the familiar form of an AR(1) process. Thus, the PACF of an MA(1) process is identical to the ACF of an AR(1) process, $\rho_k^* = \theta^k$.

The ARMA(p, q) is a mixture of the two types of processes, so its ACF and PACF are likewise mixtures of the two forms discussed above. Generalities are difficult to draw, but normally, the ACF of an ARMA process will have a few distinctive spikes in the early lags corresponding to the number of MA terms, followed by the characteristic

smooth pattern of the AR part of the model. High-order MA processes are relatively uncommon in general, and high-order AR processes (greater than two) seem primarily to arise in the form of the nonstationary processes described in the next section. For a stationary process, the workhorses of the applied literature are the (2, 0) and (1, 1) processes. For the ARMA(1, 1) process, both the ACF and the PACF will display a distinctive spike at lag 1 followed by an exponentially decaying pattern thereafter.

22.2.5 MODELING UNIVARIATE TIME SERIES

The preceding discussion is largely descriptive. There is no underlying economic theory that states *why* a compact ARMA(p, q) representation should adequately describe the movement of a given economic time series. Nonetheless, as a methodology for building forecasting models, this set of tools and its empirical counterpart have proved as good as and even superior to much more elaborate specifications (perhaps to the consternation of the builders of large macroeconomic models).[11] Box and Jenkins (1984) pioneered a forecasting framework based on the preceding that has been used in a great many fields and that has, certainly in terms of numbers of applications, largely supplanted the use of large integrated econometric models.

Box and Jenkins's approach to modeling a stochastic process can be motivated by the following.

THEOREM 22.1 Wold's Decomposition Theorem

Every zero-mean covariance stationary stochastic process can be represented in the form

$$y_t = E^*[y_t \mid y_{t-1}, y_{t-2}, \ldots, y_{t-p}] + \sum_{i=0}^{\infty} \pi_i \varepsilon_{t-i},$$

where ε_t is white noise, $\pi_0 = 1$, and the weights are square summable—that is,

$$\sum_{i=1}^{\infty} \pi_i^2 < \infty$$

—$E^*[y_t \mid y_{t-1}, y_{t-2}, \ldots, y_{t-p}]$ is the optimal linear predictor of y_t based on its lagged values, and the predictor E_t^* is uncorrelated with ε_{t-i}.

Thus, the theorem decomposes the process generating y_t into

$$E_t^* = E^*[y_t \mid y_{t-1}, y_{t-2}, \ldots, y_{t-p}] = \text{the linearly deterministic component}$$

and

$$\sum_{i=0}^{\infty} \pi_i \varepsilon_{t-i} = \text{the linearly indeterministic component.}$$

[11] This observation can be overstated. Even the most committed advocate of the Box–Jenkins methods would concede that an ARMA model of, for example, housing starts will do little to reveal the link between the interest rate policies of the Federal Reserve and their variable of interest. That is, the *covariation* of economic variables remains as interesting as ever.

The theorem states that for any stationary stochastic process, for a given choice of p, there is a Wold representation of the stationary series

$$y_t = \sum_{i=1}^{p} \gamma_i y_{t-i} + \sum_{i=0}^{\infty} \pi_i \varepsilon_{t-i}.$$

Note that for a specific ARMA(P, Q) process, if $p \geq P$, then $\pi_i = 0$ for $i > Q$. For practical purposes, the problem with the Wold representation is that we cannot estimate the infinite number of parameters needed to produce the full right-hand side, and, of course, P and Q are unknown. The compromise, then, is to base an estimate of the representation on a model with a finite number of moving-average terms. We can seek the one that best fits the data in hand.

It is important to note that neither the ARMA representation of a process nor the Wold representation is unique. In general terms, suppose that the process generating y_t is

$$\Gamma(L)y_t = \Theta(L)\varepsilon_t.$$

We assume that $\Gamma(L)$ is finite but $\Theta(L)$ need not be. Let $\Phi(L)$ be some other polynomial in the lag operator with roots that are outside the unit circle. Then

$$\left[\frac{\Phi(L)}{\Gamma(L)}\right]\Gamma(L)y_t = \left[\frac{\Phi(L)}{\Gamma(L)}\right]\Theta(L)\varepsilon_t,$$

or

$$\Phi(L)y_t = \Pi(L)\varepsilon_t.$$

The new representation is fully equivalent to the old one, but it might have a different number of autoregressive parameters, which is exactly the point of the Wold decomposition. The implication is that part of the model-building process will be to determine the lag structures. Further discussion on the methodology is given by Box and Jenkins (1984).

The Box–Jenkins approach to modeling stochastic processes consists of the following steps:

1. Satisfactorily transform the data so as to obtain a stationary series. This step will usually mean taking first differences, logs, or both to obtain a series whose autocorrelation function eventually displays the characteristic exponential decay of a stationary series.
2. Estimate the parameters of the resulting ARMA model, generally by nonlinear least squares.
3. Generate the set of residuals from the estimated model and verify that they satisfactorily resemble a white noise series. If not, respecify the model and return to step 2.
4. The model can now be used for forecasting purposes.

Space limitations prevent us from giving a full presentation of the set of techniques. Because this methodology has spawned a mini-industry of its own, however, there is no shortage of book length analyses and prescriptions to which the reader may refer. Five to consider are the canonical source, Box and Jenkins (1984), Granger and Newbold

(1996), Mills (1993), Enders (2004), and Patterson (2000). Some of the aspects of the estimation and analysis steps do have broader relevance for our work here, so we will continue to examine them in some detail.

21.2.6 ESTIMATION OF THE PARAMETERS OF A UNIVARIATE TIME SERIES

The broad problem of regression estimation with time-series data, which carries through to all the discussions of this chapter, is that the consistency and asymptotic normality results that we derived based on random sampling will no longer apply. For example, for a stationary series, we have assumed that $\text{Var}[y_t] = \lambda_0$ regardless of t. But we have yet to establish that an estimated variance,

$$c_0 = \frac{1}{T-1}\sum_{t=1}^{T}(y_t - \bar{y})^2,$$

will converge to λ_0, or anything else for that matter. It is necessary to assume that the process is **ergodic.** (We first encountered this assumption in Section 19.4.1—see Definition 19.3.) Ergodicity is a crucial element of our theory of estimation. When a time series has this property (with stationarity), then we can consider estimation of parameters in a meaningful sense. If the process is stationary and ergodic, then, by the Ergodic theorem (Theorems 19.1 and 19.2), moments such as $\bar{y}$ and c_0 converge to their population counterparts μ and λ_0.[12] The essential component of the condition is one that we have met at many points in this discussion, that autocovariances must decline sufficiently rapidly as the separation in time increases. It is possible to construct theoretical examples of processes that are stationary but not ergodic, but for practical purposes, a stationarity assumption will be sufficient for us to proceed with estimation. For example, in our models of stationary processes, if we assume that $\varepsilon_t \sim N[0, \sigma^2]$, which is common, then the stationary processes are ergodic as well.

Estimation of the parameters of a time-series process must begin with a determination of the type of process that we have in hand. (Box and Jenkins label this the **identification** step. But identification is a term of art in econometrics, so we will steer around that admittedly standard name.) For this purpose, the empirical estimates of the autocorrelation and partial autocorrelation functions are useful tools.

The sample counterpart to the ACF is the **correlogram,**

$$r_k = \frac{\sum_{t=k+1}^{T}(y_t - \bar{y})(y_{t-k} - \bar{y})}{\sum_{t=1}^{T}(y_t - \bar{y})^2}.$$

A plot of r_k against k provides a description of a process and can be used to help discern what type of process is generating the data. The sample PACF is the counterpart to the ACF, but net of the intervening lags; that is,

$$r_k^* = \frac{\sum_{t=k+1}^{T} y_t^* y_{t-k}^*}{\sum_{t=k+1}^{T}(y_{t-k}^*)^2},$$

[12] The formal conditions for ergodicity are quite involved; see Davidson and MacKinnon (1993) or Hamilton (1994, Chapter 7).

where y_t^* and y_{t-k}^* are residuals from the regressions of y_t and y_{t-k} on $[1, y_{t-1}, y_{t-2}, \ldots, y_{t-k+1}]$. We have seen this at many points before; r_k^* is simply the last linear least squares regression coefficient in the regression of y_t on $[1, y_{t-1}, y_{t-2}, \ldots, y_{t-k+1}, y_{t-k}]$. Plots of the ACF and PACF of a series are usually presented together. Because the sample estimates of the autocorrelations and partial autocorrelations are not likely to be identically zero even when the population values are, we use diagnostic tests to discern whether a time series appears to be nonautocorrelated.[13] Individual sample autocorrelations will be approximately distributed with mean zero and variance $1/T$ under the hypothesis that the series is white noise. The Box–Pierce (1970) statistic

$$Q = T\sum_{k=1}^{p} r_k^2$$

is commonly used to test whether a series is white noise. Under the null hypothesis that the series is white noise, Q has a limiting chi-squared distribution with p degrees of freedom. A refinement that appears to have better finite-sample properties is the Ljung–Box (1979) statistic,

$$Q' = T(T+2)\sum_{k=1}^{p} \frac{r_k^2}{T-k}.$$

The limiting distribution of Q' is the same as that of Q.

The process of finding the appropriate specification is essentially trial and error. An initial specification based on the sample ACF and PACF can be found. The parameters of the model can then be estimated by least squares. For pure AR(p) processes, the estimation step is simple. The parameters can be estimated by linear least squares. If there are moving-average terms, then linear least squares is inconsistent, but the parameters of the model can be fit by nonlinear least squares. Once the model has been estimated, a set of residuals is computed to assess the adequacy of the specification. In an AR model, the residuals are just the deviations from the regression line.

The adequacy of the specification can be examined by applying the foregoing techniques to the estimated residuals. If they appear satisfactorily to mimic a white noise process, then analysis can proceed to the forecasting step. If not, a new specification should be considered.

Example 21.1 ACF and PACF for a Series of Bond Yields

Appendix Table F21.1 lists five years of monthly averages of the yield on a Moody's Aaa-rated corporate bond. (Note: In previous editions of this text, the second observation in the data file was incorrectly recorded as 9.72. The correct value is 9.22. Computations to follow are based on the corrected value.) The series is plotted in Figure 21.1. From the figure, it would appear that stationarity may not be a reasonable assumption. We will return to this question below. The ACF and PACF for the original series are shown in Table 21.1, with the diagnostic statistics discussed earlier.

Based on the two spikes in the PACF, the results appear to be consistent with an AR(2) process, although the ACF at longer lags seems a bit more persistent than might have been expected. Once again, this condition may indicate that the series is not stationary. Maintaining that assumption for the present, we computed the residuals from an AR(2) model and

[13] The LM test discussed in Section 19.7.1 is one of these.

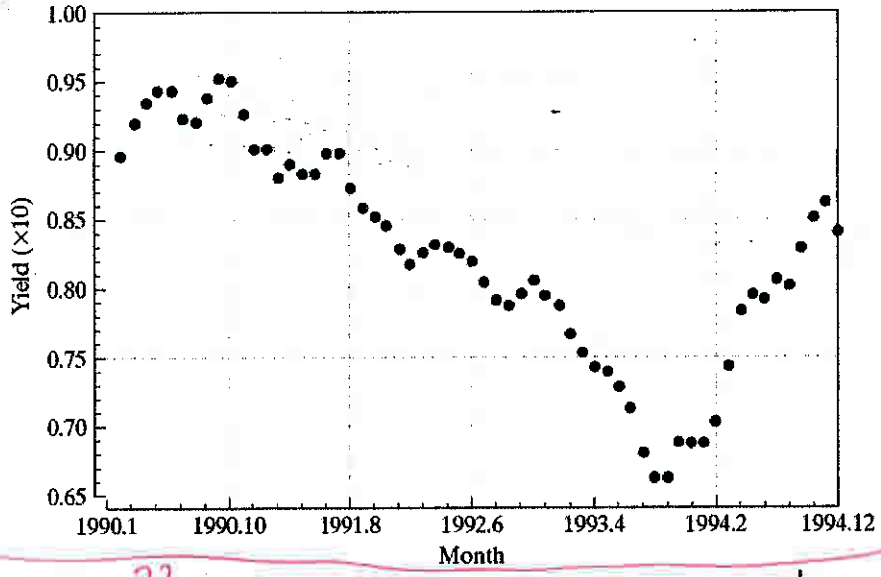

FIGURE 21.1 Monthly Data on Bond Yields, observations

subjected them to the same test as the original series. To compute the regression, we first subtracted the overall mean from all 60 obserations. We then fit the AR(2) without the first two observations. The coefficients of the AR(2) process are 1.4970 and −0.4986, which also satisfy the restrictions for stationarity given in Section 21.2.2. Despite the earlier suggestions, the residuals do appear to resemble a stationary process. (See Table 21.2.)

1.497701
−.51553

TABLE 21.1 ACF and PACF for Bond Yields

Time-series identification for YIELD
Box–Pierce statistic = 326.0507
Degrees of freedom = 14
Significance level = 0.0000
Box–Ljung statistic = 364.6475
Degrees of freedom = 14
Significance level = 0.0000
* ⇒ |coefficient| > 2/sqrt(*N*) or > 95% significant.
PACF is computed using Yule–Walker equations.

Lag	*Autocorrelation Function*			*Box/Prc*	*Partial Autocorrelations X*		
1	0.973*		**********	56.81*	0.973*		**********
2	0.922*		*********	107.76*	−0.477*	******	
3	0.863*		*********	152.47*	0.057		*
4	0.806*		*********	191.43*	0.021		*
5	0.745*		********	224.71*	−0.186	***	
6	0.679*		*******	252.39*	−0.046	*	
7	0.606*		*******	274.44*	−0.174	***	
8	0.529*		******	291.22*	−0.039	*	
9	0.450*		*****	303.37*	−0.049	*	
10	0.379*		****	311.98*	0.146		**
11	0.316*		***	317.95*	−0.023	*	
12	0.259*		***	321.97*	−0.001	*	
13	0.205		**	324.49*	−0.018	*	
14	0.161		**	326.05*	0.185		***

Note, *s in first column and bars in the right-hand panel have changed from earlier edition.

TABLE 21.2 ACF and PACF for Residuals

Time series identification for U
Box-Pierce statistic = 10.7650
Degrees of freedom = 14
Significance level = 0.7044
Box-Ljung statistic = 12.4641
Degrees of freedom = 14
Significance level = 0.5691
* ⇒ |coefficient| > 2/sqrt(N) or > 95% significant.
PACF is computed using Yule-Walker equations.

Lag	*Autocorrelation Function*			*Box/Prc*	*Partial Autocorrelations X*		
1	0.084		**	0.41	0.084		*
2	−0.120	**		1.25	−0.138	**	
3	−0.241	***		4.61	−0.242	***	
4	0.095		*	5.12	0.137		**
5	0.137		**	6.22	0.104		*
6	0.121		*	7.06	0.102		*
7	−0.084	*		7.46	−0.048	*	
8	0.049		*	7.60	0.184		***
9	−0.169	**		9.26	−0.327*	****	
10	0.025		*	9.30	0.025		*
11	−0.005	*		9.30	0.037		*
12	0.003		*	9.30	−0.100	**	
13	−0.137	**		10.39	−0.203	***	
14	−0.081	*		10.77	−0.167	***	

Note, * in third column and bars in both panels have changed from earlier edition.

replace all values. Correct figure & values on next page

msp 22-18

21.3 THE FREQUENCY DOMAIN

For the analysis of macroeconomic flow data such as output and consumption, and aggregate economic index series such as the price level and the rate of unemployment, the tools described in the previous sections have proved quite satisfactory. The low frequency of observation (yearly, quarterly, or, occasionally, monthly) and very significant aggregation (both across time and of individuals) make these data relatively smooth and straightforward to analyze. Much contemporary economic analysis, especially financial econometrics, has dealt with more disaggregated, microlevel data, observed at far greater frequency. Some important examples are stock market data for which daily returns data are routinely available, and exchange rate movements, which have been tabulated on an almost continuous basis. In these settings, analysts have found that the tools of spectral analysis, and the frequency domain, have provided many useful results and have been applied to great advantage. This section introduces a small amount of the terminology of spectral analysis to acquaint the reader with a few basic features of the technique. For those who desire further detail, Fuller (1976), Granger and Newbold (1996), Hamilton (1994), Chatfield (1996), Shumway (1988), and Hatanaka (1996) (among many others with direct application in economics) are excellent introductions. Most of the following is based on Chapter 6 of Hamilton (1994).

In this framework, we view an observed time series as a weighted sum of underlying series that have different cyclical patterns. For example, aggregate retail sales and construction data display several different kinds of cyclical variation, including a regular seasonal pattern and longer frequency variation associated with variation in the economy as a whole over the business cycle. The total variance of an observed time series

Insert in TB 22.2 on msp 22-17 as indicated

Time series identification for U
Box-Pierce Statistic = 10.6480 Box-Ljung Statistic = 12.3380
Degrees of freedom = 14 Degrees of freedom = 14
Significance level = .7134 Significance level = .5792
* => |coefficient| > 2/sqrt(N) or > 95% significant.
PACF is computed using Yule-Walker equations.

Lag	Autocorrelation Function	Box/Prc	Partial Autocorrelations	X
1	.063 |*	.23	.063 |*	X
2	-.119 | *	1.06	-.133 | *	X
3	-.235 | ***	4.27	-.241 | ***	X
4	.108 |*	4.95	.142 |**	X
5	.142 |**	6.11	.113 |*	X
6	.117 |*	6.91	.108 |*	X
7	-.091 | *	7.39	-.047 | *	X
8	.058 |*	7.58	.189 |**	X
9	-.167 | **	9.19	-.321* | ****	X
10	.034 |*	9.25	.021 |*	X
11	-.004 | *	9.25	.043 |*	X
12	.013 |*	9.26	-.072 | *	X
13	-.134 | *	10.31	-.179 | **	X
14	-.076 | *	10.65	-.114 | *	X

end of insert

22-19
End 22

computing time.) For series involving multiple thousands of observations, such as daily market returns, or far more, such as in recorded exchange rates and forward premiums, the amount of computation could become prohibitive. However, the computation can be done using an important tool, the fast Fourier transform (FFT), that reduces the computational level to $O(T \log_2 T)$, which is many orders of magnitude less than T^2. The FFT is programmed in some econometric software packages, such as RATS and Matlab. [See Press et al. (1986) for further discussion.]

21.4 SUMMARY AND CONCLUSIONS

This chapter has developed the standard tools for analyzing a stationary time series. The analysis takes place in one of two frameworks, the time domain or the frequency domain. The analysis in the time domain focuses on the different representations of the series in terms of autoregressive and moving-average components. This interpretation of the time series is closely related to the concept of regression—though in this case it is "auto" or self-regression, that is, on past values of the random variable itself. The autocorrelations and partial autocorrelations are the central tools for characterizing a time series in this framework. Constructing a time series in this fashion allows the analyst to construct forecasting equations that exploit the internal structure of the time series. (We have left for additional courses and the many references on the subject the embellishments of these models in terms of seasonal patterns, differences, and so on, that are the staples of Box–Jenkins model building.) The other approach to time-series analysis is in the frequency domain. In this representation, the time series is decomposed into a sum of components that vary cyclically with different periods (and frequencies). The different components that vary at different frequencies provide a view of how the different components contribute to the overall variation of the series.

The analysis in this chapter, of modern economic time-series analysis, is a prelude to the analysis of nonstationary series in the next chapter. Nonstationarity is, in large measure, the norm in recent time-series modeling.

All Terms with blue checks were not bold KTS in text

Key Terms and Concepts

- Autocorrelation function
- Autocovariance
- Autocovariance function
- Autocovariance generating function
- Autoregression
- Autoregressive
- Autoregressive form
- Autoregressive moving average
- Characteristic equation
- Correlogram
- Covariance stationarity
- Discrete Fourier transform
- Ergodic
- Frequency domain
- Identification
- Innovations
- Invertibility
- Lag window
- Linearly deterministic component
- Linearly indeterministic component
- Moving average
- Moving-average form
- Nonstationarity
- Partial autocorrelation
- Periodogram
- Sample periodogram
- Self-regressive
- Spectral density function
- Square summable
- Stationarity
- Strong stationarity
- Unit circle
- Univariate time-series model
- Weak stationarity
- White noise
- Wold's decomposition theorem
- Yule–Walker equations