$$\frac{(n-m)\|P_{\mathscr{M}}(\mathbf{Y}-X\boldsymbol{\theta})\|^2}{m\|\mathbf{Y}-P_{\mathscr{M}}\mathbf{Y}\|^2}$$

has the F distribution with m and $(n-m)$ degree of freedom.

2.20. Suppose $(X, Z_1, \ldots, Z_n)'$ has a multivariate normal distribution. Show that

$$P_{\overline{\mathrm{sp}}\{1, Z_1, \ldots, Z_n\}}(X) = E_{\mathscr{M}(Z_1, \ldots, Z_n)}(X),$$

where the conditional expectation operator $E_{\mathscr{M}(Z_1, \ldots, Z_n)}$ is defined as in Section 2.7.

2.21. Suppose $\{X_t, t = 0, \pm 1, \ldots\}$ is a stationary process with mean zero and autocovariance function $\gamma(\cdot)$ which is absolutely summable (i.e. $\sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty$). Define f to be the function,

$$f(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \gamma(h) e^{-ih\lambda}, \qquad -\pi \le \lambda \le \pi,$$

and show that $\gamma(h) = \int_{-\pi}^{\pi} e^{ih\lambda} f(\lambda)\, d\lambda$.

2.22. (a) If $f \in L^2([-\pi, \pi])$, prove the Riemann–Lebesgue lemma: $\langle f, e_h \rangle \to 0$ as $h \to \infty$, where e_n was defined by (2.8.2).
(b) If $f \in L^2([-\pi, \pi])$ has a continuous derivative $f'(x)$ and $f(\pi) = f(-\pi)$, show that $\langle f, e_h \rangle = (ih)^{-1} \langle f', e_h \rangle$ and hence that $h\langle f, e_h \rangle \to 0$ as $h \to \infty$. Show also that $\sum_{h=-\infty}^{\infty} |\langle f, e_h \rangle| < \infty$ and conclude that $S_n f$ (see Section 2.8) converges uniformly to f.

2.23. Show that the space l^2 (Example 2.9.1) is a separable Hilbert space.

2.24. If $\mathscr{H}$ is any Hilbert space with orthonormal basis $\{e_n, n = 1, 2, \ldots\}$, show that the mapping defined by $Th = \{\langle h, e_n \rangle\}$, $h \in \mathscr{H}$, is an isomorphism of $\mathscr{H}$ onto l^2.

2.25.* Prove that $\mathscr{M}(Z)$ (see Definition 2.7.3) is closed.

P. Brockwell and R. Davis "Time Series: Theory and Methods". Springer-Verlag

CHAPTER 3

Stationary ARMA Processes

In this chapter we introduce an extremely important class of time series $\{X_t, t = 0, \pm 1, \pm 2, \ldots\}$ defined in terms of linear difference equations with constant coefficients. The imposition of this additional structure defines a parametric family of stationary processes, the autoregressive moving average or ARMA processes. For any autocovariance function $\gamma(\cdot)$ such that $\lim_{h\to\infty} \gamma(h) = 0$, and for any integer $k > 0$, it is possible to find an ARMA process with autocovariance function $\gamma_X(\cdot)$ such that $\gamma_X(h) = \gamma(h)$, $h = 0, 1, \ldots, k$. For this (and other) reasons the family of ARMA processes plays a key role in the modelling of time-series data. The linear structure of ARMA processes leads also to a very simple theory of linear prediction which is discussed in detail in Chapter 5.

§3.1 Causal and Invertible ARMA Processes

In many respects the simplest kind of time series $\{X_t\}$ is one in which the random variables X_t, $t = 0, \pm 1, \pm 2, \ldots$ are independently and identically distributed with zero mean and variance σ^2. From a second order point of view i.e. ignoring all properties of the joint distributions of $\{X_t\}$ except those which can be deduced from the moments $E(X_t)$ and $E(X_s X_t)$, such processes are identified with the class of all stationary processes having mean zero and autocovariance function

$$\gamma(h) = \begin{cases} \sigma^2 & \text{if } h = 0, \\ 0 & \text{if } h \ne 0. \end{cases} \tag{3.1.1}$$

Definition 3.1.1. The process $\{Z_t\}$ is said to be white noise with mean 0 and variance σ^2, written

$$\{Z_t\} \sim \mathrm{WN}(0,\sigma^2), \tag{3.1.2}$$

if and only if $\{Z_t\}$ has zero mean and covariance function (3.1.1).

If the random variables Z_t are independently and identically distributed with mean 0 and variance σ^2 then we shall write

$$\{Z_t\} \sim \mathrm{IID}(0,\sigma^2). \tag{3.1.3}$$

A very wide class of stationary processes can be generated by using white noise as the forcing terms in a set of linear difference equations. This leads to the notion of an autoregressive-moving average (ARMA) process.

Definition 3.1.2 (The ARMA (p,q) Process). The process $\{X_t, t = 0, \pm 1, \pm 2, \dots\}$ is said to be an ARMA(p,q) process if $\{X_t\}$ is stationary and if for every t,

$$X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q}, \tag{3.1.4}$$

where $\{Z_t\} \sim \mathrm{WN}(0,\sigma^2)$. We say that $\{X_t\}$ is an ARMA(p,q) process with mean μ if $\{X_t - \mu\}$ is an ARMA(p,q) process.

The equations (3.1.4) can be written symbolically in the more compact form

$$\phi(B)X_t = \theta(B)Z_t, \qquad t = 0, \pm 1, \pm 2, \dots, \tag{3.1.5}$$

where ϕ and θ are the p^{th} and q^{th} degree polynomials

$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p \tag{3.1.6}$$

and

$$\theta(z) = 1 + \theta_1 z + \cdots + \theta_q z^q \tag{3.1.7}$$

and B is the backward shift operator defined by

$$B^j X_t = X_{t-j}, \qquad j = 0, \pm 1, \pm 2, \dots. \tag{3.1.8}$$

The polynomials ϕ and θ will be referred to as the autoregressive and moving average polynomials respectively of the difference equations (3.1.5).

EXAMPLE 3.1.1 (The MA(q) Process). If $\phi(z) \equiv 1$ then

$$X_t = \theta(B)Z_t \tag{3.1.9}$$

and the process is said to be a moving-average process of order q (or MA(q)). It is quite clear in this case that the difference equations have the unique solution (3.1.9). Moreover the solution $\{X_t\}$ is a stationary process since (defining $\theta_0 = 1$ and $\theta_j = 0$ for $j > q$), we see that

$$EX_t = \sum_{j=0}^{q} \theta_j E Z_{t-j} = 0$$

and

$$\mathrm{Cov}(X_{t+h}, X_t) = \begin{cases} \sigma^2 \sum_{j=0}^{q-|h|} \theta_j \theta_{j+|h|} & \text{if } |h| \le q, \\ 0 & \text{if } |h| > q. \end{cases}$$

A realization of $\{X_1, \dots, X_{100}\}$ with $q = 1, \theta_1 = -.8$ and $Z_t \sim N(0,1)$ is shown in Figure 3.1(a). The autocorrelation function of the process is shown in Figure 3.1(b).

EXAMPLE 3.1.2 (The AR(p) Process). If $\theta(z) \equiv 1$ then

$$\phi(B)X_t = Z_t \tag{3.1.10}$$

and the process is said to be an autoregressive process of order p (or AR(p)). In this case (as in the general case to be considered in Theorems 3.1.1–3.1.3) the existence and uniqueness of a stationary solution of (3.1.10) needs closer investigation. We illustrate by examining the case $\phi(z) = 1 - \phi_1 z$, i.e.

$$X_t = Z_t + \phi_1 X_{t-1}. \tag{3.1.11}$$

Iterating (3.1.11) we obtain

$$\begin{aligned} X_t &= Z_t + \phi_1 Z_{t-1} + \phi_1^2 X_{t-2} \\ &= \cdots \\ &= Z_t + \phi_1 Z_{t-1} + \cdots + \phi_1^k Z_{t-k} + \phi_1^{k+1} X_{t-k-1}. \end{aligned}$$

If $|\phi_1| < 1$ and $\{X_t\}$ is stationary then $\|X_t\|^2 = E(X_t^2)$ is constant so that

$$\left\| X_t - \sum_{j=0}^{k} \phi_1^j Z_{t-j} \right\|^2 = \phi_1^{2k+2} \|X_{t-k-1}\|^2 \to 0 \quad \text{as } k \to \infty.$$

Since $\sum_{j=0}^{\infty} \phi_1^j Z_{t-j}$ is mean-square convergent (by the Cauchy criterion), we conclude that

$$X_t = \sum_{j=0}^{\infty} \phi_1^j Z_{t-j}. \tag{3.1.12}$$

Equation (3.1.12) is valid not only in the mean square sense but also (by Proposition 3.1.1 below) with probability one, i.e.

$$X_t(\omega) = \sum_{j=0}^{\infty} \phi_1^j Z_{t-j}(\omega) \quad \text{for all } \omega \notin E,$$

where E is a subset of the underlying probability space with probability zero. All the convergent series of random variables encountered in this chapter will (by Proposition 3.1.1) be both mean square convergent and absolutely convergent with probability one. Now $\{X_t\}$ defined by (3.1.12) is stationary since

$$EX_t = \sum_{j=0}^{\infty} \phi_1^j E Z_{t-j} = 0$$

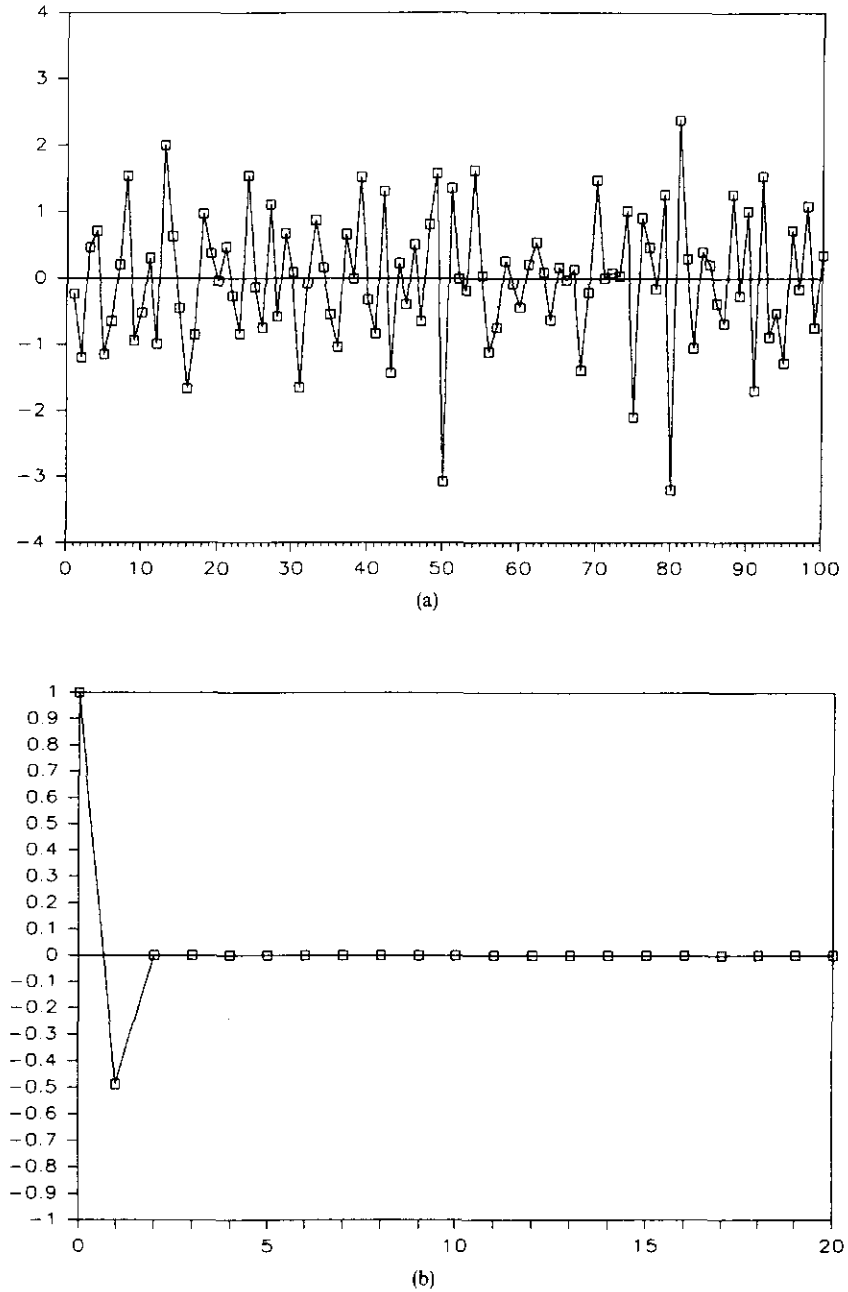


Figure 3.1. (a) 100 observations of the series $X_t = Z_t - .8Z_{t-1}$, Example 3.1.1. (b) The autocorrelation function of $\{X_t\}$.

and

$$\begin{aligned}\operatorname{Cov}(X_{t+h}, X_t) &= \lim_{n\to\infty} E\left[\left(\sum_{j=0}^{n} \phi_1^j Z_{t+h-j}\right)\left(\sum_{k=0}^{n} \phi_1^k Z_{t-k}\right)\right] \\ &= \sigma^2 \phi_1^{|h|} \sum_{j=0}^{\infty} \phi_1^{2j} \\ &= \sigma^2 \phi_1^{|h|}/(1-\phi_1^2).\end{aligned}$$

Moreover $\{X_t\}$ as defined by (3.1.12) satisfies the difference equations (3.1.11) and is therefore the unique stationary solution. A realization of the process with $\phi_1 = .9$ and $Z_t \sim N(0,1)$ is shown in Figure 3.2(a). The autocorrelation function of the same process is shown in Figure 3.2(b).

In the case when $|\phi_1| > 1$ the series (3.1.12) does not converge in L^2. However we can rewrite (3.1.11) in the form

$$X_t = -\phi_1^{-1} Z_{t+1} + \phi_1^{-1} X_{t+1}. \tag{3.1.13}$$

Iterating (3.1.13) gives

$$\begin{aligned}X_t &= -\phi_1^{-1} Z_{t+1} - \phi_1^{-2} Z_{t+2} + \phi_1^{-2} X_{t+2} \\ &= \cdots \\ &= -\phi_1^{-1} Z_{t+1} - \cdots - \phi_1^{-k-1} Z_{t+k+1} + \phi_1^{-k-1} X_{t+k+1},\end{aligned}$$

which shows, by the same arguments as in the preceding paragraph, that

$$X_t = -\sum_{j=1}^{\infty} \phi_1^{-j} Z_{t+j} \tag{3.1.14}$$

is the unique stationary solution of (3.1.11). This solution should not be confused with the non-stationary solution $\{X_t, t = 0, \pm 1, \ldots\}$ of (3.1.11) obtained when X_0 is any specified random variable which is uncorrelated with $\{Z_t\}$.

The stationary solution (3.1.14) is frequently regarded as unnatural since X_t as defined by (3.1.14) is correlated with $\{Z_s, s > t\}$, a property not shared by the solution (3.1.12) obtained when $|\phi| < 1$. It is customary therefore when modelling stationary time series to restrict attention to AR(1) processes with $|\phi_1| < 1$ for which X_t has the representation (3.1.12) in terms of $\{Z_s, s \le t\}$. Such processes are called causal or future-independent autoregressive processes. It should be noted that every AR(1) process with $|\phi_1| > 1$ can be reexpressed as an AR(1) process with $|\phi_1| < 1$ and a new white noise sequence (Problem 3.3). From a second-order point of view therefore, nothing is lost by eliminating AR(1) processes with $|\phi_1| > 1$ from consideration.

If $|\phi_1| = 1$ there is no stationary solution of (3.1.11) (Problem 3.4). Consequently there is no such thing as an AR(1) with $|\phi_1| = 1$ according to our Definition 3.1.2.

The concept of causality will now be defined for a general ARMA(p, q) process.

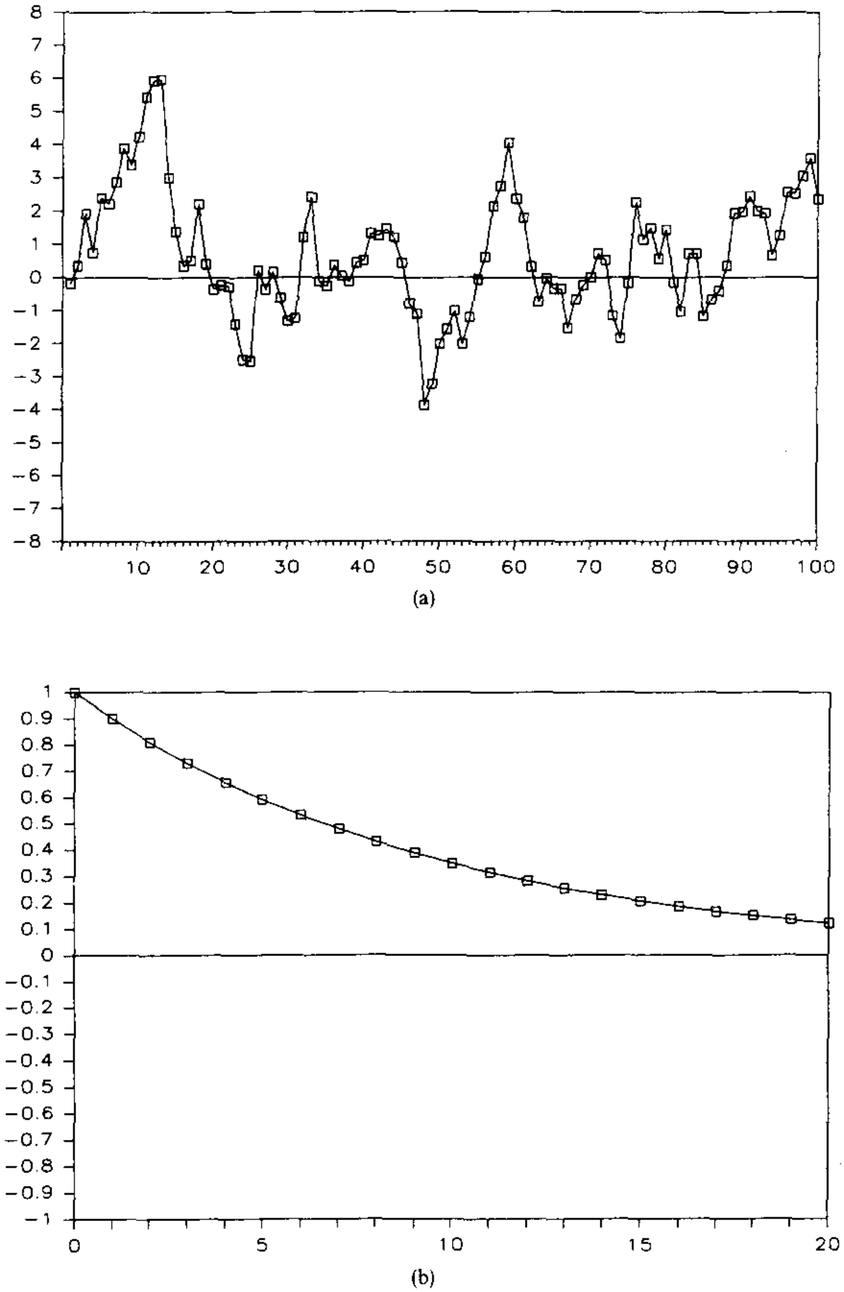


Figure 3.2. (a) 100 observations of the series $X_t - .9X_{t-1} = Z_t$, Example 3.1.2. (b) The autocorrelation function of $\{X_t\}$.

Definition 3.1.3. An ARMA(p, q) process defined by the equations $\phi(B)X_t = \theta(B)Z_t$ is said to be *causal* (or more specifically to be a causal function of $\{Z_t\}$) if there exists a sequence of constants $\{\psi_j\}$ such that $\sum_{j=0}^{\infty} |\psi_j| < \infty$ and

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}, \qquad t = 0, \pm 1, \ldots. \tag{3.1.15}$$

It should be noted that causality is a property not of the process $\{X_t\}$ alone but rather of the relationship between the two processes $\{X_t\}$ and $\{Z_t\}$ appearing in the defining ARMA equations. In the terminology of Section 4.10 we can say that $\{X_t\}$ is causal if it is obtained from $\{Z_t\}$ by application of a causal linear filter. The following proposition clarifies the meaning of the sum appearing in (3.1.15).

Proposition 3.1.1. *If $\{X_t\}$ is any sequence of random variables such that $\sup_t E|X_t| < \infty$, and if $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$, then the series*

$$\psi(B)X_t = \sum_{j=-\infty}^{\infty} \psi_j B^j X_t = \sum_{j=-\infty}^{\infty} \psi_j X_{t-j}, \tag{3.1.16}$$

converges absolutely with probability one. If in addition $\sup_t E|X_t|^2 < \infty$ then the series converges in mean square to the same limit.

PROOF. The monotone convergence theorem and finiteness of $\sup_t E|X_t|$ give

$$\begin{aligned} E\left(\sum_{j=-\infty}^{\infty} |\psi_j||X_{t-j}|\right) &= \lim_{n\to\infty} E\left(\sum_{j=-n}^{n} |\psi_j||X_{t-j}|\right) \\ &\le \lim_{n\to\infty}\left(\sum_{j=-n}^{n} |\psi_j|\right)\sup_t E|X_t| \\ &< \infty, \end{aligned}$$

from which it follows that $\sum_{j=-\infty}^{\infty} |\psi_j||X_{t-j}|$ and $\psi(B)X_t = \sum_{j=-\infty}^{\infty} \psi_j X_{t-j}$ are both finite with probability one.

If $\sup_t E|X_t|^2 < \infty$ and $n > m > 0$, then

$$\begin{aligned} E\left|\sum_{m<|j|\le n} \psi_j X_{t-j}\right|^2 &= \sum_{m<|j|\le n}\ \sum_{m<|k|\le n} \psi_j \bar{\psi}_k E(X_{t-j}\bar{X}_{t-k}) \\ &\le \sup_t E|X_t|^2 \left(\sum_{m<j\le n} |\psi_j|\right)^2 \\ &\to 0 \quad \text{as } m, n \to \infty, \end{aligned}$$

and so by the Cauchy criterion the series (3.1.16) converges in mean square. If S denotes the mean square limit, then by Fatou's lemma,

$$E|S - \psi(B)X_t|^2 = E \liminf_{n\to\infty} \left|S - \sum_{j=-n}^{n} \psi_j X_{t-j}\right|^2$$

$$\le \liminf_{n\to\infty} E\left|S - \sum_{j=-n}^{n} \psi_j X_{t-j}\right|^2$$
$$= 0,$$

showing that the limits S and $\psi(B)X_t$ are equal with probability one. □

Proposition 3.1.2. *If $\{X_t\}$ is a stationary process with autocovariance function $\gamma(\cdot)$ and if $\sum_{j=-\infty}^{\infty}|\psi_j| < \infty$, then for each $t \in \mathbb{Z}$ the series (3.1.16) converges absolutely with probability one and in mean square to the same limit. If*

$$Y_t = \psi(B)X_t$$

then the process $\{Y_t\}$ is stationary with autocovariance function

$$\gamma_Y(h) = \sum_{j,k=-\infty}^{\infty} \psi_j\psi_k\gamma(h - j + k).$$

Proof. The convergence assertions follow at once from Proposition 3.1.1 and the observation that if $\{X_t\}$ is stationary then

$$E|X_t| \le (E|X_t|^2)^{1/2} = c,$$

where c is finite and independent of t.

To check the stationarity of $\{Y_t\}$ we observe, using the mean square convergence of (3.1.16) and continuity of the inner product, that

$$EY_t = \lim_{n\to\infty} \sum_{j=-n}^{n} \psi_j EX_{t-j} = \left(\sum_{j=-\infty}^{\infty} \psi_j\right) EX_t,$$

and

$$E(Y_{t+h}Y_t) = \lim_{n\to\infty} E\left[\left(\sum_{j=-n}^{n} \psi_j X_{t+h-j}\right)\left(\sum_{k=-n}^{n} \psi_k X_{t-k}\right)\right]$$
$$= \sum_{j,k=-\infty}^{\infty} \psi_j\psi_k(\gamma(h - j + k) + (EX_t)^2).$$

Thus EY_t and $E(Y_{t+h}Y_t)$ are both finite and independent of t. The autocovariance function $\gamma_Y(\cdot)$ of $\{Y_t\}$ is given by

$$\gamma_Y(h) = E(Y_{t+h}Y_t) - EY_{t+h} \cdot EY_t = \sum_{j,k=-\infty}^{\infty} \psi_j\psi_k\gamma(h - j + k). \quad \square$$

It is an immediate corollary of Proposition 3.1.2 that operators such as $\psi(B) = \sum_{j=-\infty}^{\infty} \psi_j B^j$ with $\sum_{j=-\infty}^{\infty}|\psi_j| < \infty$, when applied to stationary processes, are not only meaningful but also inherit the algebraic properties of power series. In particular if $\sum_{j=-\infty}^{\infty}|\alpha_j| < \infty$, $\sum_{j=-\infty}^{\infty}|\beta_j| < \infty$, $\alpha(z) = \sum_{j=-\infty}^{\infty} \alpha_j z^j$, $\beta(z) = \sum_{j=-\infty}^{\infty} \beta_j z^j$ and

$$\alpha(z)\beta(z) = \psi(z), \qquad |z| \le 1,$$

then $\alpha(B)\beta(B)X_t$ is well-defined and

$$\alpha(B)\beta(B)X_t = \beta(B)\alpha(B)X_t = \psi(B)X_t.$$

The following theorem gives necessary and sufficient conditions for an ARMA process to be causal. It also gives an explicit representation of X_t in terms of $\{Z_s, s \le t\}$.

Theorem 3.1.1. *Let $\{X_t\}$ be an ARMA(p, q) process for which the polynomials $\phi(\cdot)$ and $\theta(\cdot)$ have no common zeroes. Then $\{X_t\}$ is causal if and only if $\phi(z) \ne 0$ for all $z \in \mathbb{C}$ such that $|z| \le 1$. The coefficients $\{\psi_j\}$ in (3.1.15) are determined by the relation*

$$\psi(z) = \sum_{j=0}^{\infty} \psi_j z^j = \theta(z)/\phi(z), \qquad |z| \le 1. \tag{3.1.17}$$

(The numerical calculation of the coefficients ψ_j is discussed in Section 3.3.)

Proof. First assume that $\phi(z) \ne 0$ if $|z| \le 1$. This implies that there exists $\varepsilon > 0$ such that $1/\phi(z)$ has a power series expansion,

$$1/\phi(z) = \sum_{j=0}^{\infty} \xi_j z^j = \xi(z), \qquad |z| < 1 + \varepsilon.$$

Consequently $\xi_j(1 + \varepsilon/2)^j \to 0$ as $j \to \infty$ so that there exists $K \in (0, \infty)$ for which

$$|\xi_j| < K(1 + \varepsilon/2)^{-j} \quad \text{for all } j = 0, 1, 2, \ldots.$$

In particular we have $\sum_{j=0}^{\infty}|\xi_j| < \infty$ and $\xi(z)\phi(z) \equiv 1$ for $|z| \le 1$. By Proposition 3.1.2 we can therefore apply the operator $\xi(B)$ to both sides of the equation $\phi(B)X_t = \theta(B)Z_t$ to obtain

$$X_t = \xi(B)\theta(B)Z_t.$$

Thus we have the desired representation,

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}$$

where the sequence $\{\psi_j\}$ is determined by (3.1.17).

Now assume that $\{X_t\}$ is causal, i.e. $X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}$ for some sequence $\{\psi_j\}$ such that $\sum_{j=0}^{\infty}|\psi_j| < \infty$. Then

$$\theta(B)Z_t = \phi(B)X_t = \phi(B)\psi(B)Z_t.$$

If we let $\eta(z) = \phi(z)\psi(z) = \sum_{j=0}^{\infty} \eta_j z^j$, $|z| \le 1$, we can rewrite this equation as

$$\sum_{j=0}^{q} \theta_j Z_{t-j} = \sum_{j=0}^{\infty} \eta_j Z_{t-j},$$

and taking inner products of each side with Z_{t-k} (recalling that $\{Z_t\} \sim \text{WN}(0, \sigma^2)$) we obtain $\eta_k = \theta_k$, $k = 0, \ldots, q$ and $\eta_k = 0$, $k > q$. Hence

$$\theta(z) = \eta(z) = \phi(z)\psi(z), \qquad |z| \le 1.$$

Since $\theta(z)$ and $\phi(z)$ have no common zeroes and since $|\psi(z)| < \infty$ for $|z| \le 1$, we conclude that $\phi(z)$ cannot be zero for $|z| \le 1$. □

Remark 1. If $\{X_t\}$ is an ARMA process for which the polynomials $\phi(\cdot)$ and $\theta(\cdot)$ have common zeroes, then there are two possibilities:

(a) none of the common zeroes lie on the unit circle, in which case (Problem 3.6) $\{X_t\}$ is the unique stationary solution of the ARMA equations with no common zeroes, obtained by cancelling the common factors of $\phi(\cdot)$ and $\theta(\cdot)$,

(b) at least one of the common zeroes lies on the unit circle, in which case the ARMA equations may have more than one stationary solution (see Problem 3.24).

Consequently ARMA processes for which $\phi(\cdot)$ and $\theta(\cdot)$ have common zeroes are rarely considered.

Remark 2. The first part of the proof of Theorem 3.1.1 shows that if $\{X_t\}$ is a stationary solution of the ARMA equations with $\phi(z) \ne 0$ for $|z| \le 1$, then we must have $X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}$ where $\{\psi_j\}$ is defined by (3.1.17). Conversely if $X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}$ then $\phi(B)X_t = \phi(B)\psi(B)Z_t = \theta(B)Z_t$. Thus the process $\{\psi(B)Z_t\}$ is the *unique* stationary solution of the ARMA equations if $\phi(z) \ne 0$ for $|z| \le 1$.

Remark 3. We shall see later (Problem 4.28) that if $\phi(\cdot)$ and $\theta(\cdot)$ have no common zeroes and if $\phi(z) = 0$ for some $z \in \mathbb{C}$ with $|z| = 1$, then there is no stationary solution of $\phi(B)X_t = \theta(B)Z_t$.

We now introduce another concept which is closely related to that of causality.

Definition 3.1.4. An ARMA(p, q) process defined by the equations $\phi(B)X_t = \theta(B)Z_t$ is said to be *invertible* if there exists a sequence of constants $\{\pi_j\}$ such that $\sum_{j=0}^{\infty} |\pi_j| < \infty$ and

$$Z_t = \sum_{j=0}^{\infty} \pi_j X_{t-j}, \qquad t = 0, \pm 1, \ldots. \tag{3.1.18}$$

Like causality, the property of invertibility is not a property of the process $\{X_t\}$ alone, but of the relationship between the two processes $\{X_t\}$ and $\{Z_t\}$ appearing in the defining ARMA equations. The following theorem gives necessary and sufficient conditions for invertibility and specifies the coefficients π_j in the representation (3.1.18).

Theorem 3.1.2. *Let $\{X_t\}$ be an ARMA(p, q) process for which the polynomials $\phi(\cdot)$ and $\theta(\cdot)$ have no common zeroes. Then $\{X_t\}$ is invertible if and only if $\theta(z) \ne 0$ for all $z \in \mathbb{C}$ such that $|z| \le 1$. The coefficients $\{\pi_j\}$ in (3.1.18) are determined by the relation*

$$\pi(z) = \sum_{j=0}^{\infty} \pi_j z^j = \phi(z)/\theta(z), \qquad |z| \le 1. \tag{3.1.19}$$

(The coefficients $\{\pi_j\}$ can be calculated from recursion relations analogous to those for $\{\psi_j\}$ (see Problem 3.7).)

Proof. First assume that $\theta(z) \ne 0$ if $|z| \le 1$. By the same argument as in the proof of Theorem 3.1.1, $1/\theta(z)$ has a power series expansion

$$1/\theta(z) = \sum_{j=0}^{\infty} \eta_j z^j = \eta(z), \qquad |z| < 1 + \varepsilon,$$

for some $\varepsilon > 0$. Since $\sum_{j=0}^{\infty} |\eta_j| < \infty$, Proposition 3.1.2 allows us to apply $\eta(B)$ to both sides of the equation $\phi(B)X_t = \theta(B)Z_t$ to obtain

$$\eta(B)\phi(B)X_t = \eta(B)\theta(B)Z_t = Z_t.$$

Thus we have the desired representation

$$Z_t = \sum_{j=0}^{\infty} \pi_j X_{t-j},$$

where the sequence $\{\pi_j\}$ is determined by (3.1.19).

Conversely if $\{X_t\}$ is invertible then $Z_t = \sum_{j=0}^{\infty} \pi_j X_{t-j}$ for some sequence $\{\pi_j\}$ such that $\sum_{j=0}^{\infty} |\pi_j| < \infty$. Then

$$\phi(B)Z_t = \pi(B)\phi(B)X_t = \pi(B)\theta(B)Z_t.$$

Setting $\xi(z) = \pi(z)\theta(z) = \sum_{j=0}^{\infty} \xi_j z^j$, $|z| \le 1$, we can rewrite this equation as

$$\sum_{j=0}^{p} \phi_j Z_{t-j} = \sum_{j=0}^{\infty} \xi_j Z_{t-j},$$

and taking inner products of each side with Z_{t-k} we obtain $\xi_k = \phi_k$, $k = 0, \ldots, p$ and $\xi_k = 0$, $k > p$. Hence

$$\phi(z) = \xi(z) = \pi(z)\theta(z), \qquad |z| \le 1.$$

Since $\phi(z)$ and $\theta(z)$ have no common zeroes and since $|\pi(z)| < \infty$ for $|z| \le 1$, we conclude that $\theta(z)$ cannot be zero for $|z| \le 1$. □

Remark 4. If $\{X_t\}$ is a stationary solution of the equations

$$\phi(B)X_t = \theta(B)Z_t, \qquad \{Z_t\} \sim \mathrm{WN}(0, \sigma^2), \tag{3.1.20}$$

and if $\phi(z)\theta(z) \ne 0$ for $|z| \le 1$, then

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}$$

and

$$Z_t = \sum_{j=0}^{\infty} \pi_j X_{t-j},$$

where $\sum_{j=0}^{\infty} \psi_j z^j = \theta(z)/\phi(z)$ and $\sum_{j=0}^{\infty} \pi_j z^j = \phi(z)/\theta(z)$, $|z| \le 1$.

Remark 5. If $\{X_t\}$ is any ARMA process, $\phi(B)X_t = \theta(B)Z_t$, with $\phi(z)$ non-zero for all z such that $|z| = 1$, then it is possible to find polynomials $\tilde{\phi}(\cdot)$, $\tilde{\theta}(\cdot)$ and a white noise process $\{Z_t^*\}$ such that $\tilde{\phi}(B)X_t = \tilde{\theta}(B)Z_t^*$ and such that $\{X_t\}$ is a causal function of $\{Z_t^*\}$. If in addition $\theta(z)$ is non-zero when $|z| = 1$ then $\tilde{\theta}(\cdot)$ can be chosen in such a way that $\{X_t\}$ is also an invertible function of $\{Z_t^*\}$, i.e. such that $\tilde{\theta}(z)$ is non-zero for $|z| \le 1$ (see Proposition 3.5.1). If $\{Z_t\} \sim \text{IID}(0, \sigma^2)$ it is not true in general that $\{Z_t^*\}$ is independent (Breidt and Davis (1990)). It is true, however, if $\{Z_t\}$ is Gaussian (see Problem 3.18).

Remark 6. Theorem 3.1.2 can be extended to include the case when the moving average polynomial has zeroes on the unit circle if we extend the definition of invertibility to require only that $Z_t \in \overline{\text{sp}}\{X_s, -\infty < s \le t\}$. Under this definition, an ARMA process is invertible if and only if $\theta(z) \ne 0$ for all $|z| < 1$ (see Problem 3.8 and Propositions 4.4.1 and 4.4.3).

In view of Remarks 4 and 5 we shall focus attention on causal invertible ARMA processes except when the contrary is explicitly indicated. We conclude this section however with a discussion of the more general case when causality and invertibility are not assumed. Recall from Remark 3 that if $\phi(\cdot)$ and $\theta(\cdot)$ have no common zeroes and if $\phi(z) = 0$ for some $z \in \mathbb{C}$ with $|z| = 1$, then there is no stationary solution of $\phi(B)X_t = \theta(B)Z_t$. If on the other hand $\phi(z) \ne 0$ for all $z \in \mathbb{C}$ such that $|z| = 1$, then a well-known result from complex analysis guarantees the existence of $r > 1$ such that

$$\theta(z)\phi(z)^{-1} = \sum_{j=-\infty}^{\infty} \psi_j z^j = \psi(z), \qquad r^{-1} < |z| < r, \tag{3.1.21}$$

the Laurent series being absolutely convergent in the specified annulus (see e.g. Ahlfors (1953)). The existence of this Laurent expansion plays a key role in the proof of the following theorem.

Theorem 3.1.3. *If $\phi(z) \ne 0$ for all $z \in \mathbb{C}$ such that $|z| = 1$, then the* ARMA *equations $\phi(B)X_t = \theta(B)Z_t$ have the unique stationary solution,*

$$X_t = \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j}, \tag{3.1.22}$$

where the coefficients ψ_j are determined by (3.1.21).

Proof. By Proposition 3.1.2, $\{X_t\}$ as defined by (3.1.22) is a stationary process. Applying the operator $\phi(B)$ to each side of (3.1.22) and noting, again by Proposition 3.1.2, that $\phi(B)\psi(B)Z_t = \theta(B)Z_t$, we obtain

$$\phi(B)X_t = \theta(B)Z_t. \tag{3.1.23}$$

Hence $\{X_t\}$ is a stationary solution of the ARMA equations.

To prove the converse let $\{X_t\}$ be any stationary solution of (3.1.23). Since $\phi(z) \ne 0$ for all $z \in \mathbb{C}$ such that $|z| = 1$, there exists $\delta > 1$ such that the series $\sum_{j=-\infty}^{\infty} \xi_j z^j = \phi(z)^{-1} = \xi(z)$ is absolutely convergent for $\delta^{-1} < |z| < \delta$. We can therefore apply the operator $\xi(B)$ to each side of (3.1.23) to get

$$\xi(B)\phi(B)X_t = \xi(B)\theta(B)Z_t,$$

or equivalently

$$X_t = \psi(B)Z_t. \qquad \square$$

§3.2 Moving Average Processes of Infinite Order

In this section we extend the notion of MA(q) process introduced in Section 3.1 by allowing q to be infinite.

Definition 3.2.1. If $\{Z_t\} \sim \text{WN}(0, \sigma^2)$ then we say that $\{X_t\}$ is a moving average (MA(∞)) of $\{Z_t\}$ if there exists a sequence $\{\psi_j\}$ with $\sum_{j=0}^{\infty} |\psi_j| < \infty$ such that

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}, \qquad t = 0, \pm 1, \pm 2, \ldots. \tag{3.2.1}$$

Example 3.2.1. The MA(q) process defined by (3.1.9) is a moving average of $\{Z_t\}$ with $\psi_j = \theta_j$, $j = 0, 1, \ldots, q$ and $\psi_j = 0$, $j > q$.

Example 3.2.2. The AR(1) process with $|\phi| < 1$ is a moving average of $\{Z_t\}$ with $\psi_j = \phi^j$, $j = 0, 1, 2, \ldots$.

Example 3.2.3. By Theorem 3.1.1 the causal ARMA(p,q) process $\phi(B)X_t = \theta(B)Z_t$ is a moving average of $\{Z_t\}$ with $\sum_{j=0}^{\infty} \psi_j z^j = \theta(z)/\phi(z)$, $|z| \le 1$.

It should be emphasized that in the definition of MA(∞) of $\{Z_t\}$ it is required that X_t should be expressible in terms of Z_s, $s \le t$, only. It is for this reason that we need the assumption of causality in Example 3.2.3. However, even for non-causal ARMA processes, it is possible to find a white noise sequence $\{Z_t^*\}$ such that X_t is a moving average of $\{Z_t^*\}$ (Proposition 3.5.1). Moreover, as we shall see in Section 5.7, a large class of stationary processes have MA(∞) representations. We consider a special case in the following proposition.

Proposition 3.2.1. *If $\{X_t\}$ is a zero-mean stationary process with autocovariance function $\gamma(\cdot)$ such that $\gamma(h) = 0$ for $|h| > q$ and $\gamma(q) \ne 0$, then $\{X_t\}$ is an* MA(q)

process, i.e. there exists a white noise process $\{Z_t\}$ *such that*

$$X_t = Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q}. \tag{3.2.2}$$

PROOF. For each t, define the subspace $\mathcal{M}_t = \overline{\text{sp}}\{X_s, -\infty < s \le t\}$ of L^2 and set

$$Z_t = X_t - P_{\mathcal{M}_{t-1}} X_t. \tag{3.2.3}$$

Clearly $Z_t \in \mathcal{M}_t$, and by definition of $P_{\mathcal{M}_{t-1}}$, $Z_t \in \mathcal{M}_{t-1}^{\perp}$. Thus if $s < t$, $Z_s \in \mathcal{M}_s \subset \mathcal{M}_{t-1}$ and hence $EZ_sZ_t = 0$. Moreover, by Problem 2.18

$$P_{\overline{\text{sp}}\{X_s, s=t-n,\dots,t-1\}} X_t \xrightarrow{\text{m.s.}} P_{\mathcal{M}_{t-1}} X_t \quad \text{as } n \to \infty,$$

so that by stationarity and the continuity of the L^2 norm,

$$\begin{aligned} \|Z_{t+1}\| &= \|X_{t+1} - P_{\mathcal{M}_t} X_{t+1}\| \\ &= \lim_{n\to\infty} \|X_{t+1} - P_{\overline{\text{sp}}\{X_s, s=t+1-n,\dots,t\}} X_{t+1}\| \\ &= \lim_{n\to\infty} \|X_t - P_{\overline{\text{sp}}\{X_s, s=t-n,\dots,t-1\}} X_t\| \\ &= \|X_t - P_{\mathcal{M}_{t-1}} X_t\| = \|Z_t\|. \end{aligned}$$

Defining $\sigma^2 = \|Z_t\|^2$, we conclude that $\{Z_t\} \sim \text{WN}(0, \sigma^2)$.

Now by (3.2.3), it follows that

$$\begin{aligned} \mathcal{M}_{t-1} &= \overline{\text{sp}}\{X_s, s < t-1, Z_{t-1}\} \\ &= \overline{\text{sp}}\{X_s, s < t-q, Z_{t-q}, \dots, Z_{t-1}\} \end{aligned}$$

and consequently $\mathcal{M}_{t-1}$ can be decomposed into the two orthogonal subspaces, $\mathcal{M}_{t-q-1}$ and $\overline{\text{sp}}\{Z_{t-q}, \dots, Z_{t-1}\}$. Since $\gamma(h) = 0$ for $|h| > q$, it follows that $X_t \perp \mathcal{M}_{t-q-1}$ and so by Proposition 2.3.2 and Theorem 2.4.1,

$$\begin{aligned} P_{\mathcal{M}_{t-1}} X_t &= P_{\mathcal{M}_{t-q-1}} X_t + P_{\overline{\text{sp}}\{Z_{t-q},\dots,Z_{t-1}\}} X_t \\ &= 0 + \sigma^{-2} E(X_t Z_{t-1}) Z_{t-1} + \cdots + \sigma^{-2} E(X_t Z_{t-q}) Z_{t-q} \\ &= \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q} \end{aligned}$$

where $\theta_j := \sigma^{-2} E(X_t Z_{t-j})$, which by stationarity is independent of t for $j = 1, \dots, q$. Substituting for $P_{\mathcal{M}_{t-1}} X_t$ in (3.2.3) gives (3.2.2). □

Remark. If $\{X_t\}$ has the same autocovariance function as that of an ARMA(p, q) process, then $\{X_t\}$ is also an ARMA(p, q) process. In other words, there exists a white noise sequence $\{Z_t\}$ and coefficients $\phi_1, \dots, \phi_p$, $\theta_1, \dots, \theta_q$ such that

$$X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = Z_t + \theta_1 + \cdots + \theta_q Z_{t-q}$$

(see Problem 3.19).

The following theorem is an immediate consequence of Proposition 3.1.2.

Theorem 3.2.1. *The* MA(∞) *process defined by* (3.2.1) *is stationary with mean zero and autocovariance function*

$$\gamma(k) = \sigma^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+|k|}. \tag{3.2.4}$$

Notice that Theorem 3.2.1 together with Example 3.2.3 completely determines the autocovariance function γ of any causal ARMA(p, q) process. We shall discuss the calculation of γ in more detail in Section 3.3.

The notion of AR(p) process introduced in Section 3.1 can also be extended to allow p to be infinite. In particular we note from Theorem 3.1.2 that any invertible ARMA(p, q) process satisfies the equations

$$X_t + \sum_{j=1}^{\infty} \pi_j X_{t-j} = Z_t, \qquad t = 0, \pm 1, \pm 2, \dots$$

which have the same form as the AR(p) equations (3.1.10) with $p = \infty$.

§3.3 Computing the Autocovariance Function of an ARMA(p, q) Process

We now give three methods for computing the autocovariance function of an ARMA process. In practice, the third method is the most convenient for obtaining numerical values and the second is the most convenient for obtaining a solution in closed form.

First Method. The autocovariance function γ of the causal ARMA(p, q) process $\phi(B)X_t = \theta(B)Z_t$ was shown in Section 3.2 to satisfy

$$\gamma(k) = \sigma^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+|k|}, \tag{3.3.1}$$

where

$$\psi(z) = \sum_{j=0}^{\infty} \psi_j z^j = \theta(z)/\phi(z) \qquad \text{for } |z| \le 1, \tag{3.3.2}$$

$\theta(z) = 1 + \theta_1 z + \cdots + \theta_q z^q$ and $\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p$. In order to determine the coefficients ψ_j we can rewrite (3.3.2) in the form $\psi(z)\phi(z) = \theta(z)$ and equate coefficients of z^j to obtain (defining $\theta_0 = 1$, $\theta_j = 0$ for $j > q$ and $\phi_j = 0$ for $j > p$),

$$\psi_j - \sum_{0<k\le j} \phi_k \psi_{j-k} = \theta_j, \qquad 0 \le j < \max(p, q+1) \tag{3.3.3}$$

and

$$\psi_j - \sum_{0<k\le p} \phi_k \psi_{j-k} = 0, \qquad j \ge \max(p, q+1). \tag{3.3.4}$$

These equations can easily be solved successively for $\psi_0, \psi_1, \psi_2, \ldots$. Thus

$$\begin{aligned} \psi_0 &= \theta_0 = 1, \\ \psi_1 &= \theta_1 + \psi_0\phi_1 = \theta_1 + \phi_1, \\ \psi_2 &= \theta_2 + \psi_0\phi_2 + \psi_1\phi_1 = \theta_2 + \phi_2 + \theta_1\phi_1 + \phi_1^2, \\ &\ldots \end{aligned} \tag{3.3.5}$$

Alternatively the general solution (3.3.4) can be written down, with the aid of Section 3.6 as

$$\psi_n = \sum_{i=1}^{k} \sum_{j=0}^{r_i-1} \alpha_{ij} n^j \xi_i^{-n}, \qquad n \ge \max(p, q+1) - p, \tag{3.3.6}$$

where ξ_i, $i = 1, \ldots, k$ are the distinct zeroes of $\phi(z)$ and r_i is the multiplicity of ξ_i (so that in particular we must have $\sum_{i=1}^{k} r_i = p$). The p constants α_{ij} and the coefficients ψ_j, $0 \le j < \max(p, q+1) - p$, are then determined uniquely by the $\max(p, q+1)$ boundary conditions (3.3.3). This completes the determination of the sequence $\{\psi_j\}$ and hence, by (3.3.1), of the autocovariance function γ.

EXAMPLE 3.3.1. $(1 - B + \frac{1}{4}B^2)X_t = (1 + B)Z_t$. The equations (3.3.3) take the form

$$\psi_0 = \theta_0 = 1,$$

$$\psi_1 = \theta_1 + \psi_0\phi_1 = \theta_1 + \phi_1 = 2,$$

and (3.3.4) becomes

$$\psi_j - \psi_{j-1} + \tfrac{1}{4}\psi_{j-2} = 0, \qquad j \ge 2.$$

The general solution of (3.3.4) is (see Section 3.6)

$$\psi_n = (\alpha_{10} + n\alpha_{11})2^{-n}, \qquad n \ge 0.$$

The constants α_{10} and α_{11} are found from the boundary conditions $\psi_0 = 1$ and $\psi_1 = 2$ to be

$$\alpha_{10} = 1 \quad \text{and} \quad \alpha_{11} = 3.$$

Hence

$$\psi_n = (1 + 3n)2^{-n}, \qquad n = 0, 1, 2, \ldots.$$

Finally, substituting in (3.3.1), we obtain for $k \ge 0$

$$\begin{aligned} \gamma(k) &= \sigma^2 \sum_{j=0}^{\infty} (1 + 3j)(1 + 3j + 3k)2^{-2j-k} \\ &= \sigma^2 2^{-k} \sum_{j=0}^{\infty} [(3k+1)4^{-j} + 3(3k+2)j4^{-j} + 9j^2 4^{-j}] \\ &= \sigma^2 2^{-k}[\tfrac{4}{3}(3k+1) + \tfrac{12}{9}(3k+2) + \tfrac{180}{27}] \\ &= \sigma^2 2^{-k}[\tfrac{32}{3} + 8k]. \end{aligned}$$

Second Method. An alternative method for computing the autocovariance function $\gamma(\cdot)$ of the causal ARMA(p, q)

$$\phi(B)X_t = \theta(B)Z_t, \tag{3.3.7}$$

is based on the difference equations for $\gamma(k)$, $k = 0, 1, 2, \ldots$, which are obtained by multiplying each side of (3.3.7) by X_{t-k} and taking expectations, namely

$$\gamma(k) - \phi_1\gamma(k-1) - \cdots - \phi_p\gamma(k-p) = \sigma^2 \sum_{k\le j\le q} \theta_j\psi_{j-k}, \qquad 0 \le k < \max(p, q+1), \tag{3.3.8}$$

and

$$\gamma(k) - \phi_1\gamma(k-1) - \cdots - \phi_p\gamma(k-p) = 0, \qquad k \ge \max(p, q+1). \tag{3.3.9}$$

(In evaluating the right-hand sides of these equations we have used the representation $X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}$.)

The general solution of (3.3.9) has the same form as (3.3.6), viz.

$$\gamma(h) = \sum_{i=1}^{k} \sum_{j=0}^{r_i-1} \beta_{ij} h^j \xi_i^{-h}, \qquad h \ge \max(p, q+1) - p, \tag{3.3.10}$$

where the p constants β_{ij} and the covariances $\gamma(j)$, $0 \le j < \max(p, q+1) - p$, are uniquely determined from the boundary conditions (3.3.8) after first computing $\psi_0, \psi_1, \ldots, \psi_q$ from (3.3.5).

EXAMPLE 3.3.2. $(1 - B + \frac{1}{4}B^2)X_t = (1 + B)Z_t$. The equations (3.3.9) become

$$\gamma(k) - \gamma(k-1) + \tfrac{1}{4}\gamma(k-2) = 0, \qquad k \ge 2,$$

with general solution

$$\gamma(n) = (\beta_{10} + \beta_{11}n)2^{-n}, \qquad n \ge 0. \tag{3.3.11}$$

The boundary conditions (3.3.8) are

$$\gamma(0) - \gamma(1) + \tfrac{1}{4}\gamma(2) = \sigma^2(\psi_0 + \psi_1),$$

$$\gamma(1) - \gamma(0) + \tfrac{1}{4}\gamma(1) = \sigma^2\psi_0,$$

where from (3.3.5), $\psi_0 = 1$ and $\psi_1 = \theta_1 + \phi_1 = 2$. Replacing $\gamma(0)$, $\gamma(1)$ and $\gamma(2)$ in accordance with the general solution (3.3.11) we obtain

$$3\beta_{10} - 2\beta_{11} = 16\sigma^2,$$

$$-3\beta_{10} + 5\beta_{11} = 8\sigma^2,$$

whence $\beta_{11} = 8\sigma^2$ and $\beta_{10} = 32\sigma^2/3$. Finally therefore we obtain the solution

$$\gamma(k) = \sigma^2 2^{-k}[\tfrac{32}{3} + 8k],$$

as found in Example 3.3.1 using the first method.

EXAMPLE 3.3.3 (The Autocovariance Function of an MA(q) Process). By Theorem 3.2.1 the autocovariance function of the process

$$X_t = \sum_{j=0}^{q} \theta_j Z_{t-j}, \qquad \{Z_t\} \sim \text{WN}(0, \sigma^2),$$

has the extremely simple form

$$\gamma(k) = \begin{cases} \sigma^2 \sum_{j=0}^{q} \theta_j \theta_{j+|k|}, & |k| \le q, \\ 0, & |k| > q. \end{cases} \tag{3.3.12}$$

where θ_0 is defined to be 1 and $\theta_j, j > q$, is defined to be zero.

EXAMPLE 3.3.4 (The Autocovariance Function of an AR(p) Process). From (3.3.10) we know that the causal AR(p) process

$$\phi(B)X_t = Z_t,$$

has an autocovariance function of the form

$$\gamma(h) = \sum_{i=1}^{k} \sum_{j=0}^{r_i - 1} \beta_{ij} h^j \xi_i^{-h}, \qquad h \ge 0, \tag{3.3.13}$$

where ξ_i, $i = 1, \ldots, k$, are the zeroes (possibly complex) of $\phi(z)$, and r_i is the multiplicity of ξ_i. The constants β_{ij} are found from (3.3.8).

By changing the autoregressive polynomial $\phi(\cdot)$ and allowing p to be arbitrarily large it is possible to generate a remarkably large variety of covariance functions $\gamma(\cdot)$. This is extremely important when we attempt to find a process whose autocovariance function "matches" the sample autocovariances of a given data set. The general problem of finding a suitable ARMA process to represent a given set of data is discussed in detail in Chapters 8 and 9. In particular we shall prove in Section 8.1 that if $\gamma(\cdot)$ is *any* covariance function such that $\gamma(h) \to 0$ as $h \to \infty$, then for any k there is a causal AR(k) process whose autocovariance function at lags 0, 1, ..., k, coincides with $\gamma(j)$, $j = 0, 1, \ldots, k$.

We note from (3.3.13) that the rate of convergence of $\gamma(n)$ to zero as $n \to \infty$ depends on the zeroes ξ_i which are closest to the unit circle. (The causality condition guarantees that $|\xi_i| > 1$, $i = 1, \ldots, k$.) If $\phi(\cdot)$ has a zero close to the unit circle then the corresponding term or terms of (3.3.13) will decay in absolute value very slowly. Notice also that simple real zeroes of $\phi(\cdot)$ contribute terms to (3.3.13) which decrease geometrically with h. A pair of complex conjugate zeroes together contribute a geometrically damped sinusoidal term. We shall illustrate these possibilities numerically in Example 3.3.5 with reference to an AR(2) process.

EXAMPLE 3.3.5 (An Autoregressive Process with $p = 2$). For the causal AR(2),

$$(1 - \xi_1^{-1}B)(1 - \xi_2^{-1}B)X_t = Z_t, \qquad |\xi_1|, |\xi_2| > 1, \xi_1 \ne \xi_2,$$

we easily find from (3.3.13) and (3.3.8), using the relations

$$\phi_1 = \xi_1^{-1} + \xi_2^{-1},$$

and

$$\phi_2 = -\xi_1^{-1}\xi_2^{-1},$$

that

$$\gamma(h) = \frac{\sigma^2 \xi_1^2 \xi_2^2}{(\xi_1 \xi_2 - 1)(\xi_2 - \xi_1)} [(\xi_1^2 - 1)^{-1} \xi_1^{1-h} - (\xi_2^2 - 1)^{-1} \xi_2^{1-h}]. \tag{3.3.14}$$

Figure 3.3 illustrates some of the possible forms of $\gamma(\cdot)$ for different values of ξ_1 and ξ_2. Notice that if $\xi_1 = re^{i\theta}$ and $\xi_2 = re^{-i\theta}$, $0 < \theta < \pi$, then we can rewrite (3.3.14) in the more illuminating form,

$$\gamma(h) = \frac{\sigma^2 r^4 \cdot r^{-h} \sin(h\theta + \psi)}{(r^2 - 1)(r^4 - 2r^2 \cos 2\theta + 1)^{1/2} \sin\theta}, \tag{3.3.15}$$

where

$$\tan\psi = \frac{r^2 + 1}{r^2 - 1} \tan\theta \tag{3.3.16}$$

and $\cos\psi$ has the same sign as $\cos\theta$.

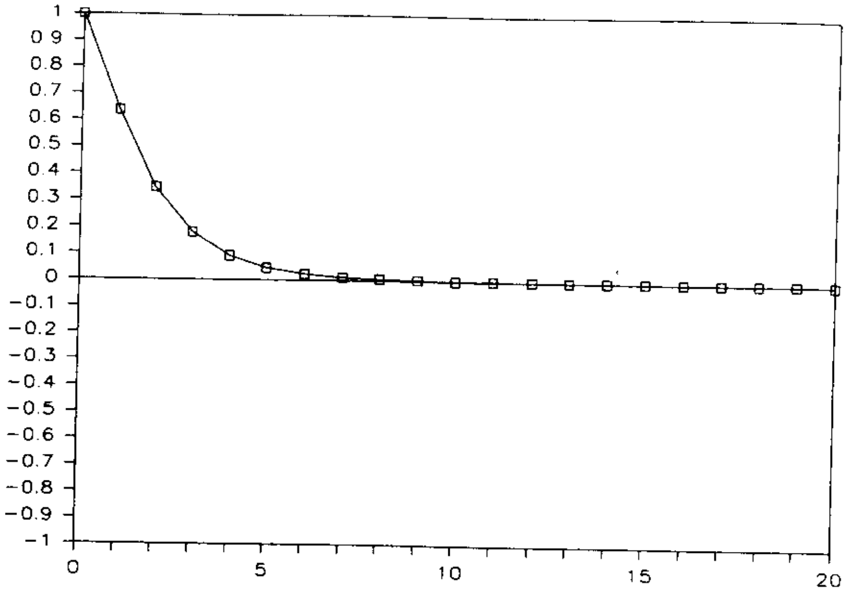
Third Method. The numerical determination of the autocovariance function $\gamma(\cdot)$ from equations (3.3.8) and (3.3.9) can be carried out readily by first finding $\gamma(0), \ldots, \gamma(p)$ from the equations with $k = 0, 1, \ldots, p$, and then using the subsequent equations to determine $\gamma(p+1), \gamma(p+2), \ldots$ recursively.

EXAMPLE 3.3.6. For the process considered in Examples 3.3.1 and 3.3.2 the equations (3.3.8) and (3.3.9) with $k = 0, 1, 2$ are

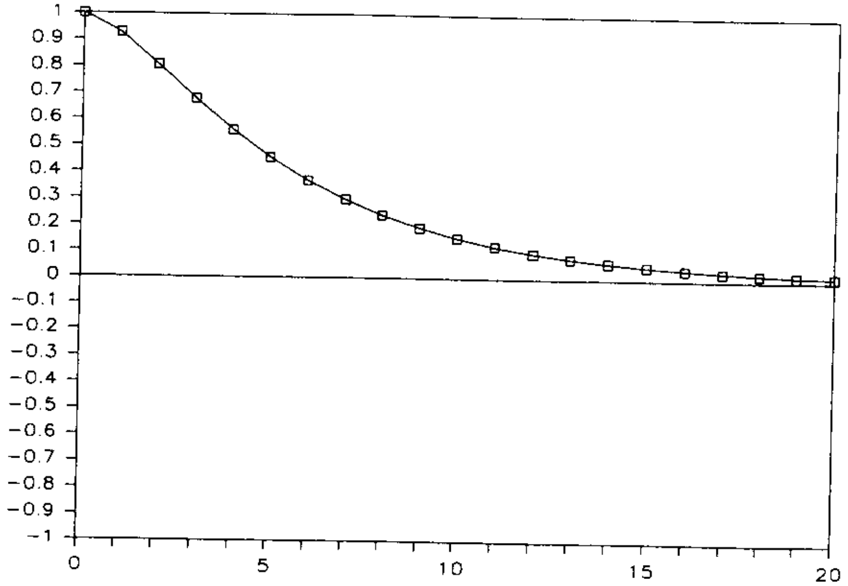
$$\gamma(0) - \gamma(1) + \tfrac{1}{4}\gamma(2) = 3\sigma^2,$$
$$\gamma(1) - \gamma(0) + \tfrac{1}{4}\gamma(1) = \sigma^2,$$
$$\gamma(2) - \gamma(1) + \tfrac{1}{4}\gamma(0) = 0,$$

with solution $\gamma(0) = 32\sigma^2/3$, $\gamma(1) = 28\sigma^2/3$, $\gamma(2) = 20\sigma^2/3$. The higher lag autocovariances can now easily be found recursively from the equations

$$\gamma(k) = \gamma(k-1) - \tfrac{1}{4}\gamma(k-2), \qquad k = 3, 4, \ldots.$$

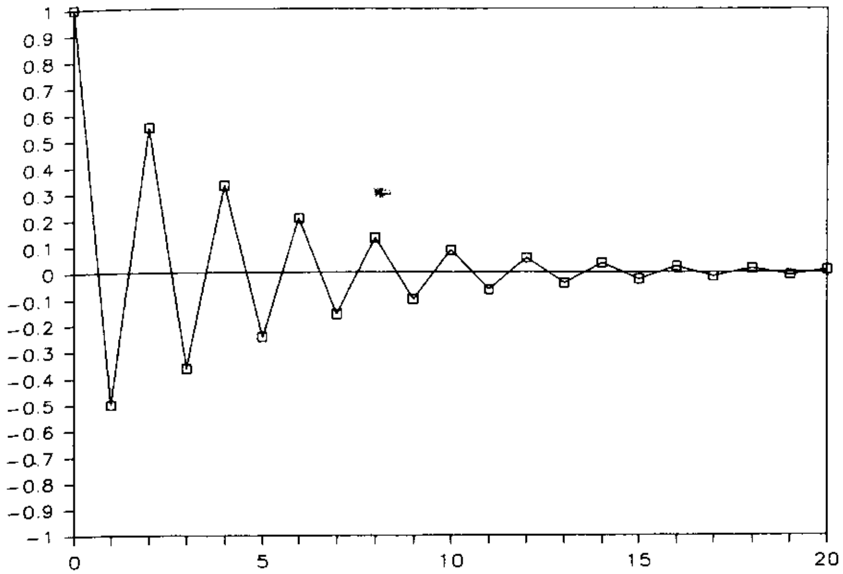


(a)

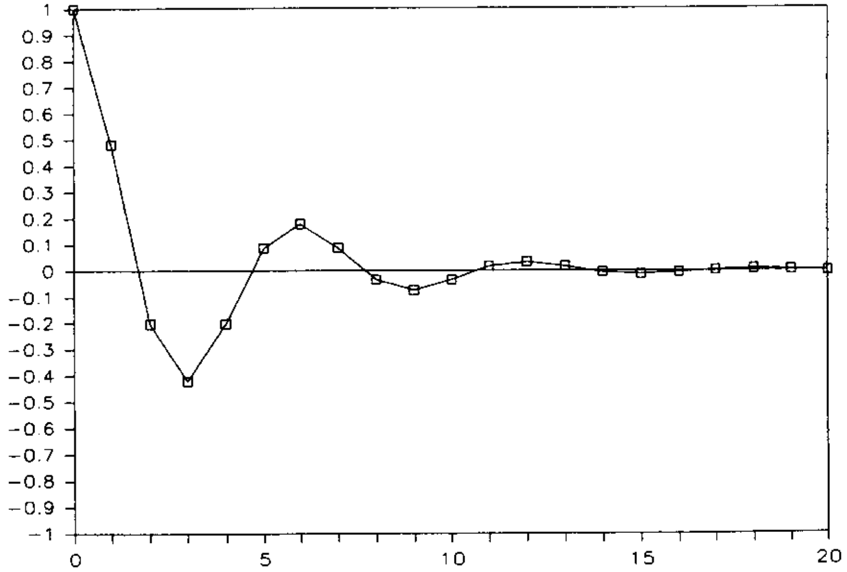


(b)

Figure 3.3. Autocorrelation functions $\gamma(h)/\gamma(0)$, $h = 0, \ldots, 20$, of the AR(2) process $(1 - \xi_1^{-1}B)(1 - \xi_2^{-1}B)X_t = Z_t$ when (a) $\xi_1 = 2$ and $\xi_2 = 5$, (b) $\xi_1 = \frac{10}{9}$ and $\xi_2 = 2$, (c) $\xi_1 = -\frac{10}{9}$ and $\xi_2 = 2$, (d) $\xi_1, \xi_2 = 2(1 \pm i\sqrt{3})/3$.



(c)



(d)

Figure 3.3 *Continued*

§3.4 The Partial Autocorrelation Function

The partial autocorrelation function, like the autocorrelation function, conveys vital information regarding the dependence structure of a stationary process. Like the autocorrelation function it also depends only on the second order properties of the process. The partial autocorrelation $\alpha(k)$ at lag k may be regarded as the correlation between X_1 and X_{k+1}, adjusted for the intervening observations $X_2, \ldots, X_k$. The idea is made precise in the following definition.

Definition 3.4.1. The partial autocorrelation function (pacf) $\alpha(\cdot)$ of a stationary time series is defined by

$$\alpha(1) = \operatorname{Corr}(X_2, X_1) = \rho(1),$$

and

$$\alpha(k) = \operatorname{Corr}(X_{k+1} - P_{\overline{\text{sp}}\{1, X_2, \ldots, X_k\}} X_{k+1}, X_1 - P_{\overline{\text{sp}}\{1, X_2, \ldots, X_k\}} X_1), \qquad k \geq 2,$$

where the projections $P_{\overline{\text{sp}}\{1, X_2, \ldots, X_k\}} X_{k+1}$ and $P_{\overline{\text{sp}}\{1, X_2, \ldots, X_k\}} X_1$ can be found from (2.7.13) and (2.7.14). The value $\alpha(k)$ is known as the partial autocorrelation at lag k.

The partial autocorrelation $\alpha(k)$, $k \geq 2$, is thus the correlation of the two residuals obtained after regressing X_{k+1} and X_1 on the intermediate observations $X_2, \ldots, X_k$. Recall that if the stationary process has zero mean then $P_{\overline{\text{sp}}\{1, X_2, \ldots, X_k\}}(\cdot) = P_{\overline{\text{sp}}\{X_2, \ldots, X_k\}}(\cdot)$ (see Problem 2.8).

Example 3.4.1. Let $\{X_t\}$ be the zero mean AR(1) process

$$X_t = .9X_{t-1} + Z_t.$$

For this example

$$\begin{aligned}\alpha(1) &= \operatorname{Corr}(X_2, X_1) \\ &= \operatorname{Corr}(.9X_1 + Z_2, X_1) \\ &= .9\end{aligned}$$

since $\operatorname{Corr}(Z_2, X_1) = 0$. Moreover $P_{\overline{\text{sp}}\{X_2, \ldots, X_k\}} X_{k+1} = .9X_k$ by Problem 2.12 and $P_{\overline{\text{sp}}\{X_2, \ldots, X_k\}} X_1 = .9X_2$ since $(X_1, X_2, \ldots, X_k)'$ has the same covariance matrix as $(X_{k+1}, X_k, \ldots, X_2)'$. Hence for $k \geq 2$,

$$\begin{aligned}\alpha(k) &= \operatorname{Corr}(X_{k+1} - .9X_k, X_1 - .9X_2) \\ &= \operatorname{Corr}(Z_{k+1}, X_1 - .9X_2) \\ &= 0.\end{aligned}$$

A realization of 100 observations $\{X_t, t = 1, \ldots, 100\}$ was displayed in Figure 3.2. Scatter diagrams of (X_{t-1}, X_t) and (X_{t-2}, X_t) are shown in Figures 3.4 and 3.5 respectively. The sample correlation $\hat{\rho}(1) = \sum_{t=1}^{99}(X_t - \bar{X})(X_{t+1} - \bar{X})/[\sum_{t=1}^{100}(X_t - \bar{X})^2]$ for Figure 3.4 is .814 (as compared with the corresponding

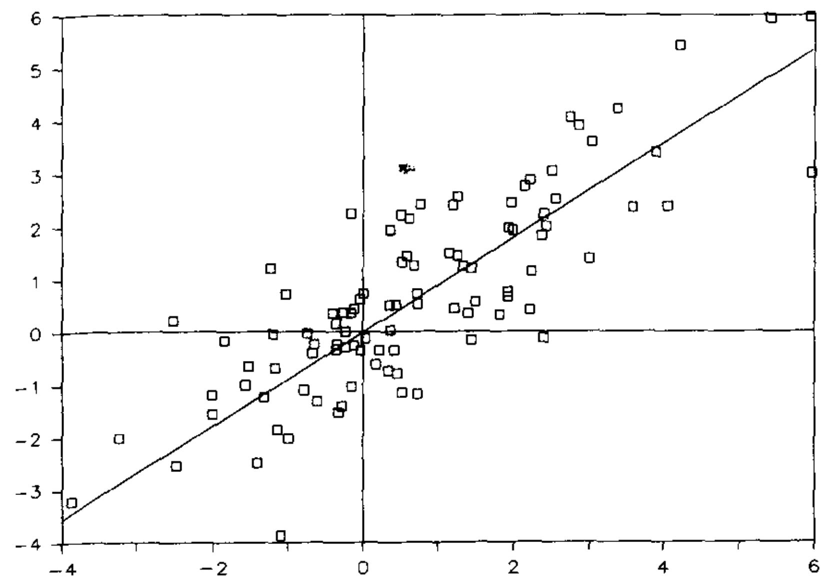


Figure 3.4. Scatter plot of the points (x_{t-1}, x_t) for the data of Figure 3.2, showing the line $x_t = .9x_{t-1}$.

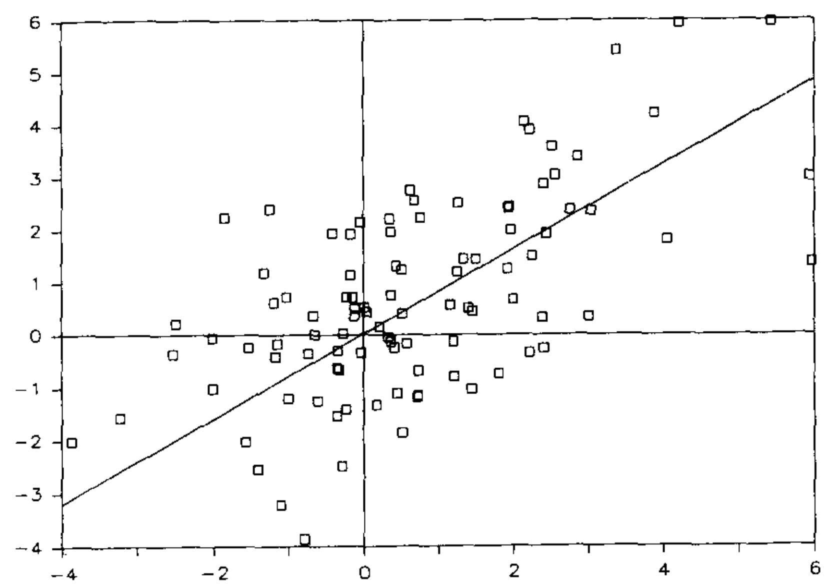


Figure 3.5. Scatter plot of the points (x_{t-2}, x_t) for the data of Figure 3.2, showing the line $x_t = .81x_{t-2}$.

theoretical correlation $\rho(1) = .9$). Likewise the sample correlation $\hat{\rho}(2) = \sum_{t=1}^{98}(X_t - \bar{X})(X_{t+2} - \bar{X})/[\sum_{t=1}^{100}(X_t - \bar{X})^2]$ for Figure 3.5 is .605 as compared with the theoretical correlation $\rho(2) = .81$. In Figure 3.6 we have plotted the points $(X_{t-2} - .9X_{t-1}, X_t - .9X_{t-1})$. It is apparent from the graph that the sample correlation between these variables is very small as expected from the fact that the theoretical partial autocorrelation at lag 2, i.e. $\alpha(2)$, is zero. One could say that the correlation between X_{t-2} and X_t is entirely eliminated when we remove the information in both variables explained by X_{t-1}.

EXAMPLE 3.4.2 (An MA(1) Process). For the moving average process,

$$X_t = Z_t + \theta Z_{t-1}, \qquad |\theta| < 1, \{Z_t\} \sim \mathrm{WN}(0, \sigma^2),$$

we have

$$\alpha(1) = \rho(1) = \theta/(1 + \theta^2).$$

A simple calculation yields $P_{\overline{sp}\{X_2\}}X_3 = [\theta/(1 + \theta^2)]X_2 = P_{\overline{sp}\{X_2\}}X_1$, whence

$$\alpha(2) = \mathrm{Corr}(X_3 - \theta(1 + \theta^2)^{-1}X_2, X_1 - \theta(1 + \theta^2)^{-1}X_2)$$
$$= -\theta^2/(1 + \theta^2 + \theta^4).$$

More lengthy calculations (Problem 3.21) give

$$\alpha(k) = -\frac{(-\theta)^k(1 - \theta^2)}{1 - \theta^{2(k+1)}}.$$

One hundred observations $\{X_t, t = 1, \ldots, 100\}$ of the process with $\theta = -.8$ and $\rho(1) = -.488$ were displayed in Figure 3.1. The scatter diagram of the points $(X_{t-2} + .488X_{t-1}, X_t + .488X_{t-1})$ is plotted in Figure 3.7 and the sample correlation of the two variables is found to be $-.297$, as compared with the theoretical correlation $\alpha(2) = -(.8)^2/(1 + .8^2 + .8^4) = -.312$.

EXAMPLE 3.4.3 (An AR(p) Process). For the causal AR process

$$X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = Z_t, \qquad \{Z_t\} \sim \mathrm{WN}(0, \sigma^2),$$

we have for $k > p$,

$$P_{\overline{sp}\{X_2,\ldots,X_k\}}X_{k+1} = \sum_{j=1}^{p} \phi_j X_{k+1-j}, \tag{3.4.1}$$

since if $Y \in \overline{sp}\{X_2, \ldots, X_k\}$ then by causality $Y \in \overline{sp}\{Z_j, j \le k\}$ and

$$\left\langle X_{k+1} - \sum_{j=1}^{p} \phi_j X_{k+1-j}, Y \right\rangle = \langle Z_{k+1}, Y\rangle = 0.$$

For $k > p$ we conclude from (3.4.1) that

$$\alpha(k) = \mathrm{Corr}\left(X_{k+1} - \sum_{j=1}^{p} \phi_j X_{k+1-j}, X_1 - P_{\overline{sp}\{X_2,\ldots,X_k\}}X_1\right)$$
$$= \mathrm{Corr}(Z_{k+1}, X_1 - P_{\overline{sp}\{X_2,\ldots,X_k\}}X_1)$$
$$= 0.$$

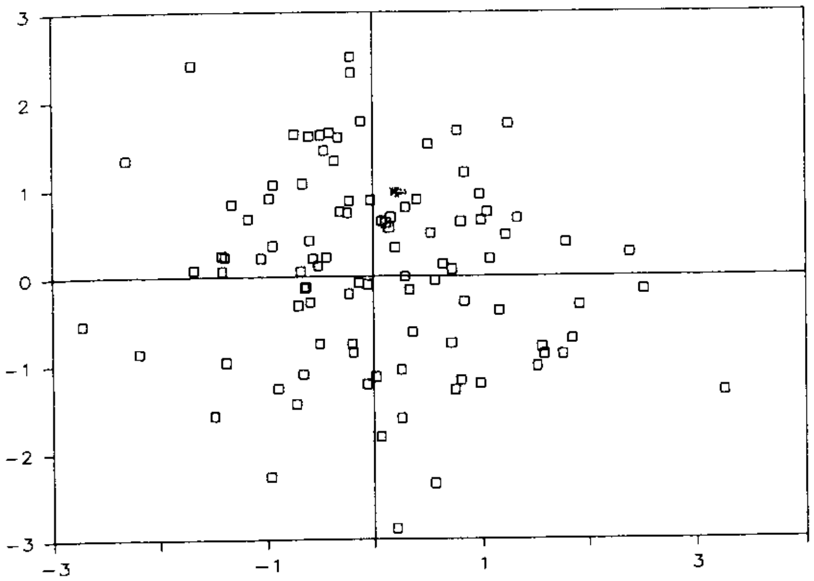


Figure 3.6. Scatter plot of the points $(x_{t-2} - .9x_{t-1}, x_t - .9x_{t-1})$ for the data of Figure 3.2.

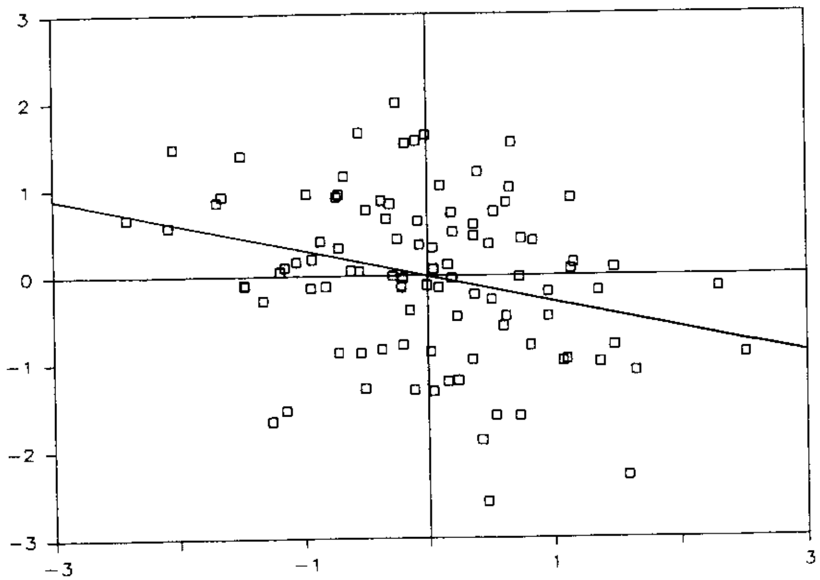


Figure 3.7. Scatter plot of the points $(x_{t-2} + .488x_{t-1}, x_t + .488x_{t-1})$ for the data of Figure 3.1, showing the line $y = -.312x$.

For $k \le p$ the values of $\alpha(k)$ can easily be computed from the equivalent Definition 3.4.2 below, after first determining $\rho(j) = \gamma(j)/\gamma(0)$ as described in Section 3.3.

In contrast with the partial autocorrelation function of an AR(p) process, that of an MA(q) process does not vanish for large lags. It is however bounded in absolute value by a geometrically decreasing function.

An Equivalent Definition of the Partial Autocorrelation Function

Let $\{X_t\}$ be a zero-mean stationary process with autocovariance function $\gamma(\cdot)$ such that $\gamma(h) \to 0$ as $h \to \infty$, and suppose that ϕ_{kj}, $j = 1, \ldots, k$; $k = 1, 2, \ldots$, are the coefficients in the representation

$$P_{\overline{\text{sp}}\{X_1,\ldots,X_k\}} X_{k+1} = \sum_{j=1}^{k} \phi_{kj} X_{k+1-j}.$$

Then from the equations

$$\langle X_{k+1} - P_{\overline{\text{sp}}\{X_1,\ldots,X_k\}} X_{k+1}, X_j \rangle = 0, \qquad j = k, \ldots, 1,$$

we obtain

$$\begin{bmatrix} \rho(0) & \rho(1) & \rho(2) & \cdots & \rho(k-1) \\ \rho(1) & \rho(0) & \rho(1) & \cdots & \rho(k-2) \\ \vdots & & & & \vdots \\ \rho(k-1) & \rho(k-2) & \rho(k-3) & \cdots & \rho(0) \end{bmatrix} \begin{bmatrix} \phi_{k1} \\ \phi_{k2} \\ \vdots \\ \phi_{kk} \end{bmatrix} = \begin{bmatrix} \rho(1) \\ \rho(2) \\ \vdots \\ \rho(k) \end{bmatrix}, \qquad k \ge 1. \tag{3.4.2}$$

Definition 3.4.2. The partial autocorrelation $\alpha(k)$ of $\{X_t\}$ at lag k is

$$\alpha(k) = \phi_{kk}, \qquad k \ge 1,$$

where ϕ_{kk} is uniquely determined by (3.4.2).

The equivalence of Definitions 3.4.1 and 3.4.2 will be established in Chapter 5, Corollary 5.2.1. The *sample* partial autocorrelation function is defined similarly.

Definition 3.4.3. The sample partial autocorrelation $\hat{\alpha}(k)$ at lag k of $\{x_1, \ldots, x_n\}$ is defined, provided $x_i \ne x_j$ for some i and j, by

$$\hat{\alpha}(k) = \hat{\phi}_{kk}, \qquad 1 \le k < n,$$

where $\hat{\phi}_{kk}$ is uniquely determined by (3.4.2) with each $\rho(j)$ replaced by the corresponding sample autocorrelation $\hat{\rho}(j)$.

§3.5 The Autocovariance Generating Function

If $\{X_t\}$ is a stationary process with autocovariance function $\gamma(\cdot)$, then its autocovariance generating function is defined by

$$G(z) = \sum_{k=-\infty}^{\infty} \gamma(k) z^k, \tag{3.5.1}$$

provided the series converges for all z in some annulus $r^{-1} < |z| < r$ with $r > 1$. Frequently the generating function is easy to calculate, in which case the autocovariance at lag k may be determined by identifying the coefficient of either z^k or z^{-k}. Clearly $\{X_t\}$ is white noise if and only if the autocovariance generating function $G(z)$ is constant for all z. If

$$X_t = \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j}, \qquad \{Z_t\} \sim \text{WN}(0, \sigma^2), \tag{3.5.2}$$

and there exists $r > 1$ such that

$$\sum_{j=-\infty}^{\infty} |\psi_j| z^j < \infty, \qquad r^{-1} < |z| < r, \tag{3.5.3}$$

the generating function $G(\cdot)$ takes a very simple form. It is easy to see that

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

and hence that

$$\begin{aligned} G(z) &= \sigma^2 \sum_{k=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \psi_j \psi_{j+|k|} z^k \\ &= \sigma^2 \left[\sum_{j=-\infty}^{\infty} \psi_j^2 + \sum_{k=1}^{\infty} \sum_{j=-\infty}^{\infty} \psi_j \psi_{j+k} (z^k + z^{-k}) \right] \\ &= \sigma^2 \left(\sum_{j=-\infty}^{\infty} \psi_j z^j \right) \left(\sum_{k=-\infty}^{\infty} \psi_k z^{-k} \right). \end{aligned}$$

Defining

$$\psi(z) = \sum_{j=-\infty}^{\infty} \psi_j z^j, \qquad r^{-1} < |z| < r,$$

we can write this result more neatly in the form

$$G(z) = \sigma^2 \psi(z) \psi(z^{-1}), \qquad r^{-1} < |z| < r. \tag{3.5.4}$$

EXAMPLE 3.5.1 (The Autocovariance Generating Function of an ARMA(p, q) Process). By Theorem 3.1.3 and (3.1.21), any ARMA process $\phi(B)X_t = \theta(B)Z_t$ for which $\phi(z) \ne 0$ when $|z| = 1$ can be written in the form (3.5.2) with

$$\psi(z) = \theta(z)/\phi(z), \qquad r^{-1} < |z| < r$$

for some $r > 1$. Hence from (3.5.4)

$$G(z) = \sigma^2 \frac{\theta(z)\theta(z^{-1})}{\phi(z)\phi(z^{-1})}, \qquad r^{-1} < |z| < r. \tag{3.5.5}$$

In particular for the MA(2) process

$$X_t = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2},$$

we have

$$\begin{aligned} G(z) &= \sigma^2(1 + \theta_1 z + \theta_2 z^2)(1 + \theta_1 z^{-1} + \theta_2 z^{-2}) \\ &= \sigma^2[(1 + \theta_1^2 + \theta_2^2) + (\theta_1 + \theta_1\theta_2)(z + z^{-1}) + \theta_2(z^2 + z^{-2})], \end{aligned}$$

from which we immediately find that

$$\begin{aligned} \gamma(0) &= \sigma^2(1 + \theta_1^2 + \theta_2^2), \\ \gamma(\pm 1) &= \sigma^2\theta_1(1 + \theta_2), \\ \gamma(\pm 2) &= \sigma^2\theta_2 \end{aligned}$$

and

$$\gamma(k) = 0 \quad \text{for } |k| > 2.$$

Example 3.5.2. Let $\{X_t\}$ be the non-invertible MA(1) process

$$X_t = Z_t - 2Z_{t-1}, \qquad \{Z_t\} \sim \text{WN}(0, \sigma^2).$$

The process defined by

$$\begin{aligned} Z_t^* &:= (1 - .5B)^{-1}(1 - 2B)Z_t \\ &= (1 - .5B)^{-1}X_t = \sum_{j=0}^{\infty} (.5)^j X_{t-j}, \end{aligned}$$

has autocovariance generating function,

$$\begin{aligned} G(z) &= \frac{(1 - 2z)(1 - 2z^{-1})}{(1 - .5z)(1 - .5z^{-1})}\sigma^2 \\ &= \frac{4(1 - 2z)(1 - 2z^{-1})}{(1 - 2z)(1 - 2z^{-1})}\sigma^2 \\ &= 4\sigma^2. \end{aligned}$$

It follows that $\{Z_t^*\} \sim \text{WN}(0, 4\sigma^2)$ and hence that $\{X_t\}$ has the invertible representation,

$$X_t = Z_t^* - .5Z_{t-1}^*.$$

A corresponding result for ARMA processes is contained in the following proposition.

Proposition 3.5.1. *Let $\{X_t\}$ be the ARMA(p, q) process satisfying the equations*

$$\phi(B)X_t = \theta(B)Z_t, \qquad \{Z_t\} \sim \text{WN}(0, \sigma^2),$$

where $\phi(z) \neq 0$ and $\theta(z) \neq 0$ for all $z \in \mathbb{C}$ such that $|z| = 1$. Then there exist polynomials, $\tilde{\phi}(z)$ and $\tilde{\theta}(z)$, nonzero for $|z| \le 1$, of degree p and q respectively, and a white noise sequence $\{Z_t^*\}$ such that $\{X_t\}$ satisfies the causal invertible equations

$$\tilde{\phi}(B)X_t = \tilde{\theta}(B)Z_t^*.$$

Proof. Define

$$\begin{aligned} \tilde{\phi}(z) &= \phi(z) \prod_{r<j\le p} \frac{(1 - a_j z)}{(1 - a_j^{-1}z)}, \\ \tilde{\theta}(z) &= \theta(z) \prod_{s<j\le q} \frac{(1 - b_j z)}{(1 - b_j^{-1}z)}, \end{aligned}$$

where $a_{r+1}, \ldots, a_p$ and $b_{s+1}, \ldots, b_q$ are the zeroes of $\phi(z)$ and $\theta(z)$ which lie inside the unit circle. Since $\tilde{\phi}(z) \neq 0$ and $\tilde{\theta}(z) \neq 0$ for all $|z| \le 1$, it suffices to show that the process defined by

$$\begin{aligned} Z_t^* &= \frac{\tilde{\phi}(B)}{\tilde{\theta}(B)} X_t \\ &= \left(\prod_{r<j\le p} \frac{1 - a_j B}{1 - a_j^{-1}B}\right)\left(\prod_{s<k\le q} \frac{1 - b_k^{-1}B}{1 - b_k B}\right) Z_t \end{aligned}$$

is white noise. Using the same calculation as in Example 3.5.2, we find that the autocovariance generating function for $\{Z_t^*\}$ is given by

$$G(z) = \sigma^2 \left(\prod_{r<j\le p} |a_j|^2\right)\left(\prod_{s<k\le q} |b_k|^{-2}\right).$$

Since $G(z)$ is constant, we conclude that $\{Z_t^*\}$ is white noise as asserted. □

§3.6* Homogeneous Linear Difference Equations with Constant Coefficients

In this section we consider the solution $\{h_t\}$ of the k^{th} order linear difference equation

$$h_t + \alpha_1 h_{t-1} + \cdots + \alpha_k h_{t-k} = 0, \qquad t \in T, \tag{3.6.1}$$

where $\alpha_1, \ldots, \alpha_k$ are real constants with $\alpha_k \neq 0$ and T is a subinterval of the integers which without loss of generality we can assume to be $[k, \infty)$, $(-\infty, \infty)$ or $[k, k + r]$, $r > 0$. Introducing the backward shift operator B defined by equation (3.1.8), we can write (3.6.1) in the more compact form

$$\alpha(B)h_t = 0, \qquad t \in T, \tag{3.6.2}$$

where $\alpha(B) = 1 + \alpha_1 B + \cdots + \alpha_k B^k$.

Definition 3.6.1. A set of $m \le k$ solutions, $\{h_t^{(1)}, \ldots, h_t^{(m)}\}$, of (3.6.2) will be called linearly independent if from

$$c_1 h_t^{(1)} + c_2 h_t^{(2)} + \cdots + c_m h_t^{(m)} = 0 \quad \text{for all } t = 0, 1, \ldots, k-1,$$

it follows that $c_1 = c_2 = \cdots = c_m = 0$.

We note that if $\{h_t^1\}$ and $\{h_t^2\}$ are any two solutions of (3.6.2) then $\{c_1 h_t^1 + c_2 h_t^2\}$ is also a solution. Moreover for any specified values of $h_0, h_1, \ldots, h_{k-1}$, henceforth referred to as *initial conditions*, all the remaining values h_t, $t \notin [0, k-1]$, are uniquely determined by one or other of the recursion relations

$$h_t = -\alpha_1 h_{t-1} - \cdots - \alpha_k h_{t-k}, \qquad t = k, k+1, \ldots, \tag{3.6.3}$$

and

$$\alpha_k h_t = -h_{t+k} - \alpha_1 h_{t+k-1} - \cdots - \alpha_{k-1} h_{t+1}, \qquad t = -1, -2, \ldots. \tag{3.6.4}$$

Thus if we can find k linearly independent solutions $\{h_t^{(1)}, \ldots, h_t^{(k)}\}$ of (3.6.2) then by linear independence there will be exactly one set of coefficients $c_1, \ldots, c_k$ such that the solution

$$h_t = c_1 h_t^{(1)} + \cdots + c_k h_t^{(k)}, \tag{3.6.5}$$

has prescribed initial values $h_0, h_1, \ldots, h_{k-1}$. Since these values uniquely determine the entire sequence $\{h_t\}$ we conclude that (3.6.5) is the unique solution of (3.6.2) satisfying the initial conditions. The remainder of this section is therefore devoted to finding a set of k linearly independent solutions of (3.6.2).

Theorem 3.6.1. *If $h_t = (a_0 + a_1 t + \cdots + a_j t^j)m^t$ where $a_0, \ldots, a_j, m$ are (possibly complex-valued) constants, then there are constants $b_0, \ldots, b_{j-1}$ such that*

$$(1 - mB)h_t = (b_0 + b_1 t + \cdots + b_{j-1} t^{j-1})m^t.$$

PROOF.

$$\begin{aligned}(1 - mB)h_t &= (a_0 + a_1 t + \cdots + a_k t^j)m^t - m(a_0 + a_1(t-1) + \cdots \\ &\quad + a_k(t-1)^j)m^{t-1} \\ &= m^t\left[\sum_{r=0}^{j} a_r(t^r - (t-1)^r)\right]\end{aligned}$$

and $\sum_{r=0}^{j} a_r(t^r - (t-1)^r)$ is clearly a polynomial of degree $j - 1$. □

Corollary 3.6.1. *The functions $h_t^{(j)} = t^j \xi^{-t}$, $j = 0, 1, \ldots, k-1$ are k linearly independent solutions of the difference equation*

$$(1 - \xi^{-1}B)^k h_t = 0. \tag{3.6.6}$$

PROOF. Repeated application of the operator $(1 - \xi^{-1}B)$ to $h_t^{(j)}$ in conjunction with Theorem 3.6.1 establishes that $h_t^{(j)}$ satisfies (3.6.6). If

$$(c_0 + c_1 t + \cdots + c_{k-1} t^{k-1})\xi^{-t} = 0 \quad \text{for } t = 0, 1, \ldots, k-1,$$

then the polynomial $\sum_{j=0}^{k-1} c_j t^j$, which is of degree less than k, has k zeroes. This is only possible if $c_0 = c_1 = \cdots = c_{k-1} = 0$. □

Solution of the General Equation of Order k

For the general equation (3.6.2), the difference operator $\alpha(B)$ can be written as

$$\alpha(B) = \prod_{i=1}^{j} (1 - \xi_i^{-1}B)^{r_i}$$

where ξ_i, $i = 1, \ldots, j$ are the distinct zeroes of $\alpha(z)$ and r_i is the multiplicity of ξ_i. It follows from Corollary 3.6.1 that $t^n \xi_i^{-t}$, $n = 0, 1, \ldots, r_i - 1$; $i = 1, \ldots, j$, are k solutions of the difference equation (3.6.2) since

$$\alpha(B)t^n \xi_i^{-t} = \prod_{s \ne i} (1 - \xi_s^{-1}B)^{r_s}(1 - \xi_i^{-1}B)^{r_i} t^n \xi_i^{-t} = 0.$$

It is shown below in Theorem 3.6.2 and Corollary 3.6.2 that these solutions are indeed linearly independent and hence that the general solution of (3.6.2) is

$$h_t = \sum_{i=1}^{j} \sum_{n=0}^{r_i - 1} c_{in} t^n \xi_i^{-t}. \tag{3.6.7}$$

In order for this general solution to be real, the coefficients corresponding to a pair of complex conjugate roots must themselves be complex conjugates. More specifically if $(\xi_i, \bar{\xi}_i)$ is a pair of complex conjugate zeroes of $\alpha(z)$ and $\xi_i = d\exp(i\theta_i)$, then the corresponding terms in (3.6.7) are

$$\sum_{n=0}^{r_i-1} c_{in} t^n \xi_i^{-t} + \sum_{n=0}^{r_i-1} \bar{c}_{in} t^n \bar{\xi}_i^{-t},$$

which can be rewritten as

$$\sum_{n=0}^{r_i-1} 2[\mathrm{Re}(c_{in})\cos(\theta_i t) + \mathrm{Im}(c_{in})\sin(\theta_i t)]t^n d^{-t},$$

or equivalently as

$$\sum_{n=0}^{r_i-1} a_{in} t^n d^{-t} \cos(\theta_i t + b_{in}),$$

with appropriately chosen constants a_{in} and b_{in}.

EXAMPLE 3.6.1. Suppose h_t satisfies the first order linear difference equation $(1 - \xi^{-1}B)h_t = 0$. Then the general solution is given by $h_t = C\xi^{-t} = h_0\xi^{-t}$. Observe that if $|\xi| > 1$, then h_t decays at an exponential rate as $t \to \infty$.

EXAMPLE 3.6.2. Consider the second order difference equation $(1 + \alpha_1 B + \alpha_2 B^2)h_t = 0$. Since $1 + \alpha_1 B + \alpha_2 B^2 = (1 - \xi_1^{-1}B)(1 - \xi_2^{-1}B)$, the character of the general solution will depend on ξ_1 and ξ_2.

Case 1 ξ_1 and ξ_2 are real and distinct. In this case, $h_t = c_1\xi_1^{-t} + c_2\xi_2^{-t}$ where c_1 and c_2 are determined by the two initial conditions $c_1 + c_2 = h_0$ and $c_1\xi_1^{-1} + c_2\xi_2^{-1} = h_1$. These have a unique solution since $\xi_1 \neq \xi_2$.

Case 2 $\xi_1 = \xi_2$. Using (3.6.7) with $j = 1$ and $r_1 = 2$ we have $h_t = (c_0 + c_1 t)\xi_1^{-t}$.

Case 3 $\xi_1 = \bar{\xi}_2 = de^{i\theta}$, $0 < \theta < 2\pi$. The solution can be written either as $c\xi_1^{-t} + \bar{c}\bar{\xi}_1^{-t}$ or as the sinusoid $h_t = ad^{-t}\cos(\theta t + b)$.

Observe that if $|\xi_1| > 1$ and $|\xi_2| > 1$, then in each of the three cases, h_t approaches zero at a geometric rate as $t \to \infty$. In the third case, h_t is a damped sinusoid. More generally, if the roots of $\alpha(z)$ lie outside the unit circle, then the general solution is a sum of exponentially decaying functions and exponentially damped sinusoids.

We now return to the problem of establishing linear independence of the solutions $t^n\xi_i^{-t}$, $n = 0, 1, \ldots, r_i - 1$; $i = 1, \ldots, j$, of (3.6.2).

Theorem 3.6.2. *If*

$$\sum_{l=1}^{q}\sum_{j=0}^{p} c_{lj}t^j m_l^t = 0 \quad \text{for } t = 0, 1, 2, \ldots \tag{3.6.8}$$

where $m_1, m_2, \ldots, m_q$ are distinct numbers, then $c_{lj} = 0$ for $l = 1, 2, \ldots, q$; $j = 0, 1, \ldots, p$.

PROOF. Without loss of generality we can assume that $|m_1| \geq |m_2| \geq \cdots \geq |m_q| > 0$. It will be sufficient to show that (3.6.8) implies that

$$c_{1j} = 0, \qquad j = 0, \ldots, p \tag{3.6.9}$$

since if this is the case then equations (3.6.8) reduce to

$$\sum_{l=2}^{q}\sum_{j=0}^{p} c_{lj}t^j m_l^t = 0, \qquad t = 0, 1, 2, \ldots,$$

which in turn imply that $c_{2j} = 0$, $j = 0, \ldots, p$. Repetition of this argument shows then that $c_{lj} = 0, j = 0, \ldots, p; l = 1, \ldots, q$.

To prove that (3.6.8) implies (3.6.9) we need to consider two separate cases.

Case 1 $|m_1| > |m_2|$. Dividing each side of (3.6.8) by $t^p m_1^t$ and letting $t \to \infty$, we find that $c_{1p} = 0$. Setting $c_{1p} = 0$ in (3.6.8), dividing each side by $t^{p-1}m_1^t$ and letting $t \to \infty$, we then obtain $c_{2p} = 0$. Repeating the procedure with divisors $t^{p-2}m_1^t, t^{p-3}m_1^t, \ldots, m_1^t$ (in that order) we find that $c_{1j} = 0, j = 0, 1, \ldots, p$ as required.

Case 2 $|m_1| = |m_2| = \cdots = |m_s| > |m_{s+1}| > 0$, where $s \leq q$. In this case we can write $m_j = re^{i\theta_j}$ where $-\pi < \theta_j \leq \pi$ and $\theta_1, \ldots, \theta_s$ are all different. Dividing each side of (3.6.8) by $t^p r^t$ and letting $t \to \infty$ we find that

$$\sum_{l=1}^{s} c_{lp}e^{i\theta_l t} \to 0 \quad \text{as } t \to \infty. \tag{3.6.10}$$

We shall now show that this is impossible unless $c_{1p} = c_{2p} = \cdots = c_{sp} = 0$. Set $g_t = \sum_{l=1}^{s} c_{lp}e^{i\theta_l t}$ and let A_n, $n = 0, 1, 2, \ldots$, be the matrix

$$A_n = \begin{bmatrix} e^{i\theta_1 n} & e^{i\theta_2 n} & & e^{i\theta_s n} \\ e^{i\theta_1(n+1)} & e^{i\theta_2(n+1)} & & e^{i\theta_s(n+1)} \\ \vdots & \vdots & & \vdots \\ e^{i\theta_1(n+s-1)} & e^{i\theta_2(n+s-1)} & \cdots & e^{i\theta_s(n+s-1)} \end{bmatrix}. \tag{3.6.11}$$

Observe that $\det A_n = e^{i(\theta_1+\cdots+\theta_s)n}(\det A_0)$. The matrix A_0 is a Vandermonde matrix (Birkhoff and Mac Lane (1965)) and hence has a non-zero determinant. Applying Cramer's rule to the equation

$$A_n\begin{bmatrix} c_{1p} \\ \vdots \\ c_{sp} \end{bmatrix} = \begin{bmatrix} g_n \\ \vdots \\ g_{n+s-1} \end{bmatrix},$$

we have

$$c_{1p} = \frac{\det M}{\det A_n}, \tag{3.6.12}$$

where

$$M = \begin{bmatrix} g_n & e^{i\theta_2 n} & \cdots & e^{i\theta_s n} \\ \vdots & \vdots & & \vdots \\ g_{n+s-1} & e^{i\theta_2(n+s-1)} & \cdots & e^{i\theta_s(n+s-1)} \end{bmatrix}.$$

Since $g_n \to 0$ as $n \to \infty$, the numerator in (3.6.12) approaches zero while the denominator remains bounded away from zero because $|\det A_n| = |\det A_0| > 0$. Hence c_{1p} must be zero. The same argument applies to the other coefficients $c_{2p}, \ldots, c_{sp}$ showing that they are all necessarily zero as claimed.

We now divide (3.6.8) by $t^{p-1}r^t$ and repeat the preceding argument, letting $t \to \infty$ to deduce that

$$\sum_{l=1}^{s} c_{l,p-1}e^{i\theta_l t} \to 0 \quad \text{as } t \to \infty,$$

and hence that $c_{l,p-1} = 0, l = 1, \ldots, s$. We then divide by $t^{p-2}r^t, \ldots, r^t$ (in that order), repeating the argument at each stage to deduce that

$$c_{lj} = 0, \quad j = 0, 1, \ldots, p \quad \text{and} \quad l = 1, 2, \ldots, s.$$

This shows that (3.6.8) implies (3.6.9) in this case, thereby completing the proof of the theorem. □

Corollary 3.6.2. *The k solutions* $t^n \xi_i^{-t}$, $n = 0, 1, \ldots, r_i - 1$; $i = 1, \ldots, j$, *of the difference equation* (3.6.2) *are linearly independent.*

PROOF. We must show that each c_{in} is zero if $\sum_{i=1}^{j} \sum_{n=0}^{r_i-1} c_{in} t^n \xi_i^{-t} = 0$ for $t = 0, 1, \ldots, k-1$. Setting h_t equal to the double sum we have $\alpha(B)h_t = 0$ and $h_0 = h_1 = \cdots = h_{k-1} = 0$. But by the recursions (3.6.3) and (3.6.4), this necessarily implies that $h_t = 0$ for all t. Direct application of Theorem 3.6.2 with $p = \max\{r_1, \ldots, r_j\}$ completes the proof. □

Problems

3.1. Determine which of the following processes are causal and/or invertible:

(a) $X_t + .2X_{t-1} - .48X_{t-2} = Z_t$,

(b) $X_t + 1.9X_{t-1} + .88X_{t-2} = Z_t + .2Z_{t-1} + .7Z_{t-2}$,

(c) $X_t + .6X_{t-2} = Z_t + 1.2Z_{t-1}$,

(d) $X_t + 1.8X_{t-1} + .81X_{t-2} = Z_t$,

(e) $X_t + 1.6X_{t-1} = Z_t - .4Z_{t-1} + .04Z_{t-2}$.

3.2. Show that in order for an AR(2) process with autoregressive polynomial $\phi(z) = 1 - \phi_1 z - \phi_2 z^2$ to be causal, the parameters (ϕ_1, ϕ_2) must lie in the triangular region determined by the intersection of the three regions,

$$\phi_2 + \phi_1 < 1,$$
$$\phi_2 - \phi_1 < 1,$$
$$|\phi_2| < 1.$$

3.3. Let $\{X_t, t = 0, \pm 1, \ldots\}$ be the stationary solution of the non-causal AR(1) equations,

$$X_t = \phi X_{t-1} + Z_t, \qquad \{Z_t\} \sim \text{WN}(0, \sigma^2), \qquad |\phi| > 1.$$

Show that $\{X_t\}$ also satisfies the causal AR(1) equations,

$$X_t = \phi^{-1} X_{t-1} + \tilde{Z}_t, \quad \{\tilde{Z}_t\} \sim \text{WN}(0, \tilde{\sigma}^2),$$

for a suitably chosen white noise process $\{\tilde{Z}_t\}$. Determine $\tilde{\sigma}^2$.

3.4. Show that there is no stationary solution of the difference equations

$$X_t = \phi X_{t-1} + Z_t, \qquad \{Z_t\} \sim \text{WN}(0, \sigma^2),$$

if $\phi = \pm 1$.

3.5. Let $\{Y_t, t = 0, \pm 1, \ldots\}$ be a stationary time series. Show that there exists a stationary solution $\{X_t\}$ of the difference equations,

$$X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = Y_t + \theta_1 Y_{t-1} + \cdots + \theta_q Y_{t-q},$$

if $\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p \neq 0$ for $|z| = 1$. Furthermore, if $\phi(z) \neq 0$ for $|z| \le 1$ show that $\{X_t\}$ is a causal function of $\{Y_t\}$.

3.6. Suppose that $\{X_t\}$ is the ARMA process defined by

$$\phi(B)X_t = \theta(B)Z_t, \qquad \{Z_t\} \sim \text{WN}(0, \sigma^2),$$

where $\phi(\cdot)$ and $\theta(\cdot)$ have no common zeroes and $\phi(z) \neq 0$ for $|z| = 1$. If $\xi(\cdot)$ is any polynomial such that $\xi(z) \neq 0$ for $|z| = 1$, show that the difference equations,

$$\xi(B)\phi(B)Y_t = \xi(B)\theta(B)Z_t,$$

have the unique stationary solution, $\{Y_t\} = \{X_t\}$.

3.7. Suppose $\{X_t\}$ is an invertible ARMA(p, q) process satisfying (3.1.4) with

$$Z_t = \sum_{j=0}^{\infty} \pi_j X_{t-j}.$$

Show that the sequence $\{\pi_j\}$ is determined by the equations

$$\pi_j + \sum_{k=1}^{\min(q,j)} \theta_k \pi_{j-k} = -\phi_j, \qquad j = 0, 1, \ldots$$

where we define $\phi_0 = -1$ and $\theta_k = 0$ for $k > q$ and $\phi_j = 0$ for $j > p$.

3.8. The process $X_t = Z_t - Z_{t-1}$, $\{Z_t\} \sim \text{WN}(0, \sigma^2)$, is not invertible according to Definition 3.1.4. Show however that $Z_t \in \overline{\text{sp}}\{X_j, -\infty < j \le t\}$ by considering the mean square limit of the sequence $\sum_{j=0}^{n} (1 - j/n) X_{t-j}$ as $n \to \infty$.

3.9. Suppose $\{X_t\}$ is the two-sided moving average

$$X_t = \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j}, \qquad \{Z_t\} \sim \text{WN}(0, \sigma^2)$$

where $\sum_j |\psi_j| < \infty$. Show that $\sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty$ where $\gamma(\cdot)$ is the autocovariance function of $\{X_t\}$.

3.10. Let $\{Y_t\}$ be a stationary zero-mean time series. Define

$$X_t = (1 - .4B)Y_t = Y_t - .4Y_{t-1}$$

and

$$W_t = (1 - 2.5B)Y_t = Y_t - 2.5Y_{t-1}.$$

(a) Express the autocovariance functions of $\{X_t\}$ and $\{W_t\}$ in terms of the autocovariance function of $\{Y_t\}$.

(b) Show that $\{X_t\}$ and $\{W_t\}$ have the same autocorrelation functions.

(c) Show that the process $U_t = -\sum_{j=1}^{\infty} (.4)^j X_{t+j}$ satisfies the difference equations $U_t - 2.5U_{t-1} = X_t$.

3.11. Let $\{X_t\}$ be an ARMA process with $\phi(z) \neq 0$, $|z| = 1$, and autocovariance function $\gamma(\cdot)$. Show that there exist constants $C > 0$ and $s \in (0, 1)$ such that $|\gamma(h)| \le Cs^{|h|}$, $h = 0, \pm 1, \ldots$ and hence that $\sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty$.

3.12. For those processes in Problem 3.1 which are causal, compute and graph their autocorrelation and partial autocorrelation functions using PEST.

3.13. Find the coefficients $\psi_j, j = 0, 1, 2, \ldots$, in the representation

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}$$

of the ARMA(2, 1) process,

$$(1 - .5B + .04B^2)X_t = (1 + .25B)Z_t, \qquad \{Z_t\} \sim \mathrm{WN}(0, \sigma^2).$$

3.14. Find the autocovariances $\gamma(j), j = 0, 1, 2, \ldots,$ of the AR(3) process,

$$(1 - .5B)(1 - .4B)(1 - .1B)X_t = Z_t, \qquad \{Z_t\} \sim \mathrm{WN}(0, 1).$$

Check your answers for $j = 0, \ldots, 4$ with the aid of the program PEST.

3.15. Find the mean and autocovariance function of the ARMA(2, 1) process,

$$X_t = 2 + 1.3X_{t-1} - .4X_{t-2} + Z_t + Z_{t-1}, \qquad \{Z_t\} \sim \mathrm{WN}(0, \sigma^2).$$

Is the process causal and invertible?

3.16. Let $\{X_t\}$ be the ARMA(1, 1) process,

$$X_t - \phi X_{t-1} = Z_t + \theta Z_{t-1}, \qquad \{Z_t\} \sim \mathrm{WN}(0, \sigma^2),$$

where $|\phi| < 1$ and $|\theta| < 1$. Determine the coefficients $\{\psi_j\}$ in Theorem 3.1.1 and show that the autocorrelation function of $\{X_t\}$ is given by $\rho(1) = (1 + \phi\theta)(\phi + \theta)/(1 + \theta^2 + 2\phi\theta)$, $\rho(h) = \phi^{h-1}\rho(1)$ for $h \geq 1$.

3.17. For an MA(2) process find the largest possible values of $|\rho(1)|$ and $|\rho(2)|$.

3.18. Let $\{X_t\}$ be the moving average process

$$X_t = Z_t - 2Z_{t-1}, \qquad \{Z_t\} \sim \mathrm{IID}(0, 1).$$

(a) If $Z_t^* := (1 - .5B)^{-1}X_t$, show that

$$Z_t^* = X_t - P_{\mathcal{M}_{t-1}}X_t,$$

where $\mathcal{M}_{t-1} = \overline{\mathrm{sp}}\{X_s, -\infty < s < t\}$.

(b) Conclude from (a) that

$$X_t = Z_t^* + \theta Z_{t-1}^*, \qquad \{Z_t^*\} \sim \mathrm{WN}(0, \sigma^2).$$

Specify the values of θ and σ^2.

(c) Find the linear filter which relates $\{Z_t\}$ to $\{Z_t^*\}$, i.e. determine the coefficients $\{\alpha_j\}$ in the representation $Z_t^* = \sum_{j=-\infty}^{\infty} \alpha_j Z_{t-j}$.

(d) If $EZ_t^3 = c$, compute $E((Z_1^*)^2 Z_2^*)$. If $c \neq 0$, are Z_1^* and Z_2^* independent? If $Z_t \sim N(0, 1)$, are Z_1^* and Z_2^* independent?

3.19. Suppose that $\{X_t\}$ and $\{Y_t\}$ are two zero-mean stationary processes with the same autovariance function and that $\{Y_t\}$ is an ARMA(p, q) process. Show that $\{X_t\}$ must also be an ARMA(p, q) process. (Hint: If $\phi_1, \ldots, \phi_p$ are the AR coefficients for $\{Y_t\}$, show that $\{W_t := X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p}\}$ has an autocovariance function which is zero for lags $|h| > q$. Then apply Proposition 3.2.1 to $\{W_t\}$.)

3.20. (a) Calculate the autocovariance function $\gamma(\cdot)$ of the stationary time series

$$Y_t = \mu + Z_t + \theta_1 Z_{t-1} + \theta_{12} Z_{t-12}, \qquad \{Z_t\} \sim \mathrm{WN}(0, \sigma^2).$$

(b) Use program PEST to compute the sample mean and sample autocovariances $\hat{\gamma}(h), 0 \leq h \leq 20$, of $\{\nabla\nabla_{12}X_t\}$ where $\{X_t, t = 1, \ldots, 72\}$ is the accidental deaths series of Example 1.1.6.

(c) By equating $\hat{\gamma}(1)$, $\hat{\gamma}(11)$ and $\hat{\gamma}(12)$ from part(b) to $\gamma(1)$, $\gamma(11)$ and $\gamma(12)$ respectively from part(a), find a model of the form defined in (a) to represent $\{\nabla\nabla_{12}X_t\}$.

3.21. By matching the autocovariances and sample autocovariances at lags 0 and 1, fit a model of the form

$$X_t - \mu = \phi(X_{t-1} - \mu) + Z_t, \qquad \{Z_t\} \sim \mathrm{WN}(0, \sigma^2),$$

to the strikes data of Example 1.1.3. Use the fitted model to compute the best linear predictor of the number of strikes in 1981. Estimate the mean squared error of your predictor.

3.22. If $X_t = Z_t - \theta Z_{t-1}$, $\{Z_t\} \sim \mathrm{WN}(0, \sigma^2)$ and $|\theta| < 1$, show from the prediction equations that the best linear predictor of X_{n+1} in $\overline{\mathrm{sp}}\{X_1, \ldots, X_n\}$ is

$$\hat{X}_{n+1} = \sum_{j=1}^{n} \phi_j X_{n+1-j},$$

where $\phi_1, \ldots, \phi_n$ satisfy the difference equations,

$$-\theta\phi_{j-1} + (1 + \theta^2)\phi_j - \theta\phi_{j+1} = 0, \qquad 2 \leq j \leq n - 1,$$

with boundary conditions,

$$(1 + \theta^2)\phi_n - \theta\phi_{n-1} = 0$$

and

$$(1 + \theta^2)\phi_1 - \theta\phi_2 = -\theta.$$

3.23. Use Definition 3.4.2 and the results of Problem 3.22 to determine the partial autocorrelation function of a moving average of order 1.

3.24. Let $\{X_t\}$ be the stationary solution of $\phi(B)X_t = \theta(B)Z_t$, where $\{Z_t\} \sim \mathrm{WN}(0, \sigma^2)$, $\phi(z) \neq 0$ for all $z \in \mathbb{C}$ such that $|z| = 1$, and $\phi(\cdot)$ and $\theta(\cdot)$ have no common zeroes. If A is any zero-mean random variable in L^2 which is uncorrelated with $\{X_t\}$ and if $|z_0| = 1$, show that the process $\{X_t + Az_0^t\}$ is a complex-valued stationary process (see Definition 4.1.1) and that $\{X_t + Az_0^t\}$ and $\{X_t\}$ both satisfy the equations $(1 - z_0 B)\phi(B)X_t = (1 - z_0 B)\theta(B)Z_t$.

Computation of this predictor using the time-domain methods of Section 5.1 is a considerably more difficult task.

Example 5.6.2 (Prediction of an ARMA Process). Consider the causal invertible ARMA process,

$$\phi(B)X_t = \theta(B)Z_t, \qquad \{Z_t\} \sim \mathrm{WN}(0,\sigma^2),$$

with spectral density

$$f(\lambda) = a(\lambda)\overline{a(\lambda)}, \tag{5.6.9}$$

where

$$a(\lambda) = (2\pi)^{-1/2}\sigma \sum_{k=0}^{\infty} \psi_k e^{-ik\lambda} \tag{5.6.10}$$

and $\sum_{k=0}^{\infty}\psi_k z^k = \theta(z)/\phi(z)$, $|z| \le 1$. Convergence of the series in (5.6.10) is uniform on $[-\pi,\pi]$ since, by the causality assumption, $\sum_k|\psi_k| < \infty$.

The function $g(\cdot) = P_{\overline{\mathrm{sp}}\{\exp(it\cdot), -\infty<t\le n\}} e^{i(n+h)\cdot}$ must satisfy

$$\int_{-\pi}^{\pi} (e^{i(n+h)\lambda} - g(\lambda))e^{-im\lambda}a(\lambda)\overline{a(\lambda)}\,d\lambda = 0, \qquad m \le n. \tag{5.6.11}$$

This equation implies that $(e^{i(n+h)\cdot} - g(\cdot))a(\cdot)\overline{a(\cdot)}$ is an element of the subspace $\mathcal{M}_+ = \overline{\mathrm{sp}}\{\exp(im\cdot),\ m > n\}$ of $L^2([-\pi,\pi], \mathcal{B}, d\lambda)$. Noting from (5.6.10) that $1/\overline{a(\cdot)} \in \overline{\mathrm{sp}}\{\exp(im\cdot),\ m \ge 0\}$, we deduce that the function $(e^{i(n+h)\cdot} - g(\cdot))a(\cdot)$ is also an element of $\mathcal{M}_+$. Let us now write

$$e^{i(n+h)\lambda}a(\lambda) = g(\lambda)a(\lambda) + (e^{i(n+h)\lambda} - g(\lambda))a(\lambda), \tag{5.6.12}$$

observing that $g(\cdot)a(\cdot)$ is orthogonal to $\mathcal{M}_+$ (in $L^2(d\lambda)$). But from (5.6.10),

$$e^{i(n+h)\lambda}a(\lambda) = (2\pi)^{-1/2}\sigma e^{in\lambda} \sum_{k=-h}^{\infty} \psi_{k+h}e^{-ik\lambda}, \tag{5.6.13}$$

and since the element $e^{i(n+h)\cdot}a(\cdot)$ of $L^2(d\lambda)$ has a unique representation as a sum of two components, one in $\mathcal{M}_+$ and one orthogonal to $\mathcal{M}_+$, we can immediately make the identification,

$$g(\lambda)a(\lambda) = (2\pi)^{-1/2}\sigma e^{in\lambda} \sum_{k=0}^{\infty} \psi_{k+h}e^{-ik\lambda}.$$

Using (5.6.10) again we obtain

$$g(\lambda) = e^{in\lambda}[\phi(e^{-i\lambda})/\theta(e^{-i\lambda})] \sum_{k=0}^{\infty} \psi_{k+h}e^{-ik\lambda},$$

i.e.

$$g(\lambda) = \sum_{j=0}^{\infty} \alpha_j e^{i(n-j)\lambda}, \tag{5.6.14}$$

where $\sum_{j=0}^{\infty}\alpha_j z^j = [\phi(z)/\theta(z)]\sum_{k=0}^{\infty}\psi_{k+h}z^k$, $|z| \le 1$. Applying the mapping I to each side of (5.6.14) and using (5.6.2) and (5.6.3), we conclude that

$$P_{\mathcal{M}_n}X_{n+h} = \sum_{j=0}^{\infty} \alpha_j X_{n-j}. \tag{5.6.15}$$

It is not difficult to check (Problem 5.17) that this result is equivalent to (5.5.4).

§5.7* The Wold Decomposition

In Example 5.6.1, the values $X_{n+j}, j \ge 1$, of the process $\{X_t, t \in \mathbb{Z}\}$ were perfectly predictable in terms of elements of $\mathcal{M}_n = \overline{\mathrm{sp}}\{X_t, -\infty < t \le n\}$. Such processes are called deterministic. Any zero-mean stationary process $\{X_t\}$ which is not deterministic can be expressed as a sum $X_t = U_t + V_t$ of an MA(∞) process $\{U_t\}$ and a deterministic process $\{V_t\}$ which is uncorrelated with $\{U_t\}$. In the statement and proof of this decomposition (Theorem 5.7.1) we shall use the notation σ^2 for the one-step mean squared error,

$$\sigma^2 = E|X_{n+1} - P_{\mathcal{M}_n}X_{n+1}|^2,$$

and $\mathcal{M}_{-\infty}$ for the closed linear subspace,

$$\mathcal{M}_{-\infty} = \bigcap_{n=-\infty}^{\infty} \mathcal{M}_n,$$

of the Hilbert space $\mathcal{M} = \overline{\mathrm{sp}}\{X_t, t \in \mathbb{Z}\}$. All subspaces and orthogonal complements should be interpreted as relative to $\mathcal{M}$. For orthogonal subspaces $\mathcal{S}_1$ and $\mathcal{S}_2$ we define $\mathcal{S}_1 \oplus \mathcal{S}_2 := \{x + y : x \in \mathcal{S}_1 \text{ and } y \in \mathcal{S}_2\}$.

Remark 1. The process $\{X_t\}$ is said to be deterministic if and only if $\sigma^2 = 0$, or equivalently if and only if $X_t \in \mathcal{M}_{-\infty}$ for each t (Problem 5.18).

Theorem 5.7.1 (The Wold Decomposition). *If $\sigma^2 > 0$ then X_t can be expressed as*

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j} + V_t, \tag{5.7.1}$$

where

(i) $\psi_0 = 1$ *and* $\sum_{j=0}^{\infty} \psi_j^2 < \infty$,

(ii) $\{Z_t\} \sim \mathrm{WN}(0,\sigma^2)$,

(iii) $Z_t \in \mathcal{M}_t$ *for each* $t \in \mathbb{Z}$,

(iv) $E(Z_t V_s) = 0$ *for all* $s, t \in \mathbb{Z}$,

(v) $V_t \in \mathcal{M}_{-\infty}$ *for each* $t \in \mathbb{Z}$,

and

(vi) $\{V_t\}$ *is deterministic.*

((v) *and* (vi) *are not the same since* $\mathcal{M}_{-\infty}$ *is defined in terms of* $\{X_t\}$, *not* $\{V_t\}$.) *The sequences* $\{\psi_j\}$, $\{Z_j\}$ *and* $\{V_j\}$ *are uniquely determined by* (5.7.1) *and the conditions* (i)–(vi).

PROOF. We first show that the sequences defined by

$$Z_t = X_t - P_{\mathcal{M}_{t-1}}X_t, \tag{5.7.2}$$

$$\psi_j = \langle X_t, Z_{t-j}\rangle/\sigma^2 \tag{5.7.3}$$

and

$$V_t = X_t - \sum_{j=0}^{\infty} \psi_j Z_{t-j}, \tag{5.7.4}$$

satisfy (5.7.1) and conditions (i)–(vi). The proof is then completed by establishing the uniqueness of the three sequences.

Clearly Z_t as defined by (5.7.2) is an element of $\mathcal{M}_t$ and is orthogonal to $\mathcal{M}_{t-1}$ by the definition of $P_{\mathcal{M}_{t-1}}X_t$. Hence

$$Z_t \in \mathcal{M}_{t-1}^{\perp} \subset \mathcal{M}_{t-2}^{\perp} \subset \cdots,$$

which shows that for $s < t$, $E(Z_s Z_t) = 0$. By Problem 5.19 this establishes (ii) and (iii). Now by Theorem 2.4.2(ii) we can write

$$P_{\overline{\text{sp}}\{Z_j, j\le t\}}X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}, \tag{5.7.5}$$

where ψ_j is defined by (5.7.3) and $\sum_{j=0}^{\infty}\psi_j^2 < \infty$. The coefficients ψ_j are independent of t by stationarity and

$$\psi_0 = \sigma^{-2}\langle X_t, X_t - P_{\mathcal{M}_{t-1}}X_t\rangle = \sigma^{-2}\|X_t - P_{\mathcal{M}_{t-1}}X_t\|^2 = 1.$$

Equations (5.7.4) and (5.7.5) and the definition of $P_{\overline{\text{sp}}\{Z_j, j\le t\}}X_t$ imply that

$$\langle V_t, Z_s\rangle = 0 \quad \text{for } s \le t.$$

On the other hand if $s > t$, $Z_s \in \mathcal{M}_{s-1}^{\perp} \subset \mathcal{M}_t^{\perp}$, and since $V_t \in \mathcal{M}_t$ we conclude that

$$\langle V_t, Z_s\rangle = 0 \quad \text{for } s > t,$$

establishing (iv.) To establish (v) and (vi) it will suffice (by Remark 1) to show that

$$\overline{\text{sp}}\{V_j, j \le t\} = \mathcal{M}_{-\infty} \quad \text{for every } t. \tag{5.7.6}$$

Since $V_t \in \mathcal{M}_t = \mathcal{M}_{t-1} \oplus \overline{\text{sp}}\{Z_t\}$ and since $\langle V_t, Z_t\rangle = 0$, we conclude that $V_t \in \mathcal{M}_{t-1} = \mathcal{M}_{t-2} \oplus \overline{\text{sp}}\{Z_{t-1}\}$. But since $\langle V_t, Z_{t-1}\rangle = 0$ it then follows that $V_t \in \mathcal{M}_{t-2}$. Continuing with this argument we see that $V_t \in \mathcal{M}_{t-j}$ for each $j \ge 0$, whence $V_t \in \bigcap_{j=0}^{\infty}\mathcal{M}_{t-j} = \mathcal{M}_{-\infty}$. Thus

$$\overline{\text{sp}}\{V_j, j \le t\} \subseteq \mathcal{M}_{-\infty} \quad \text{for every } t. \tag{5.7.7}$$

Now by (5.7.4), $\mathcal{M}_t = \overline{\text{sp}}\{Z_j, j \le t\} \oplus \overline{\text{sp}}\{V_j, j \le t\}$. If $Y \in \mathcal{M}_{-\infty}$ then $Y \in \mathcal{M}_{s-1}$ for every s, so that $\langle Y, Z_s\rangle = 0$ for every s, and consequently $Y \in \overline{\text{sp}}\{V_j, j \le t\}$.

But this means that

$$\mathcal{M}_{-\infty} \subseteq \overline{\text{sp}}\{V_j, j \le t\} \quad \text{for every } t, \tag{5.7.8}$$

which completes the proof of (5.7.6) and hence of (v) and (vi).

To establish uniqueness we observe from (5.7.1) that if $\{Z_t\}$ and $\{V_t\}$ are any sequences satisfying (5.7.1) and having the properties (i)–(vi), then $\mathcal{M}_{t-1} \subseteq \overline{\text{sp}}\{Z_j, j \le t-1\} \oplus \overline{\text{sp}}\{V_j, j \le t-1\}$ from which it follows, using (ii) and (iv), that Z_t is orthogonal to $\mathcal{M}_{t-1}$. Projecting each side of (5.7.1) onto $\mathcal{M}_{t-1}$ and subtracting the resulting equation from (5.7.1), we then find that the process $\{Z_t\}$ must satisfy (5.7.2). By taking inner products of each side of (5.7.1) with Z_{t-j} we see that ψ_j must also satisfy (5.7.3). Finally, if (5.7.1) is to hold, it is obviously necessary that V_t must be defined as in (5.7.4). □

In the course of the preceding proof we have established a number of results which are worth collecting together as a corollary.

Corollary 5.7.1

(a) $\overline{\text{sp}}\{V_j, j \le t\} = \mathcal{M}_{-\infty}$ *for every t.*
(b) $\mathcal{M}_t = \overline{\text{sp}}\{Z_j, j \le t\} \oplus \mathcal{M}_{-\infty}$.
(c) $\mathcal{M}_{-\infty}^{\perp} = \overline{\text{sp}}\{Z_j, j \in \mathbb{Z}\}$.
(d) $\overline{\text{sp}}\{U_j, j \le t\} = \overline{\text{sp}}\{Z_j, j \le t\}$, *where* $U_t = \sum_{j=0}^{\infty}\psi_j Z_{t-j}$.

PROOF.

(a) This is a restatement of (5.7.6).
(b) Use part (a) together with the relation, $\mathcal{M}_t = \overline{\text{sp}}\{Z_j, j \le t\} \oplus \overline{\text{sp}}\{V_j, j \le t\}$.
(c) Observe that $\mathcal{M} = \overline{\text{sp}}\{X_t, t \in \mathbb{Z}\} = \overline{\text{sp}}\{Z_t, t \in \mathbb{Z}\} \oplus \mathcal{M}_{-\infty}$.
(d) This follows from the fact that $\mathcal{M}_t = \overline{\text{sp}}\{U_j, j \le t\} \oplus \mathcal{M}_{-\infty}$. □

In view of part (b) of the corollary it is now possible to interpret the representation (5.7.1) as the decomposition of the subspace $\mathcal{M}_t$ into two orthogonal subspaces $\overline{\text{sp}}\{Z_j, j \le t\}$ and $\mathcal{M}_{-\infty}$.

A stationary process is said to be *purely non-deterministic* if and only if $\mathcal{M}_{-\infty} = \{0\}$. In this case the Wold decomposition has no deterministic component, and the process can be represented as an MA(∞), $X_t = \sum_{j=0}^{\infty}\psi_j Z_{t-j}$. Many of the time series dealt with in this book (e.g. ARMA processes) are purely non-deterministic.

Observe that the h-step predictor for the process (5.7.1) is

$$P_{\mathcal{M}_t}X_{t+h} = \sum_{j=h}^{\infty}\psi_j Z_{t+h-j} + V_{t+h},$$

since $Z_j \perp \mathcal{M}_t$ for all $j < t$, and $V_{t+h} \in \mathcal{M}_t$. The corresponding mean squared error is

$$\|X_{t+h} - P_{\mathcal{M}_t}X_{t+h}\|^2 = \text{Var}\left(\sum_{j=0}^{h-1}\psi_j Z_{t+h-j}\right) = \sigma^2\sum_{j=0}^{h-1}\psi_j^2,$$

which should be compared with the result (5.5.5). For a purely non-deterministic process it is clear that the h-step prediction mean squared error converges as $h \to \infty$ to the variance of the process. In general we have from part (d) of Corollary 5.7.1,

$$P_{\overline{sp}\{U_j, j \le t\}} U_{t+h} = P_{\overline{sp}\{Z_j, j \le t\}} U_{t+h} = \sum_{j=h}^{\infty} \psi_j Z_{t+h-j},$$

which shows that the h-step prediction error for the $\{U_t\}$ sequence coincides with that of the $\{X_t\}$ process. This is not unexpected since the purely deterministic component does not contribute to the prediction error.

EXAMPLE 5.7.1. Consider the stationary process $X_t = Z_t + Y$, where $\{Z_t\} \sim \mathrm{WN}(0, \sigma^2)$, $\{Z_t\}$ is uncorrelated with the random variable Y and Y has mean zero and variance σ^2. Since

$$\frac{1}{n}\sum_{j=0}^{n-1} X_{t-j} = \frac{1}{n}\sum_{j=0}^{n-1} Z_{t-j} + Y \xrightarrow{\text{m.s.}} Y,$$

it follows that $Y \in \mathscr{M}_t$ for every t. Also $Z_t \perp \mathscr{M}_s$ for $s < t$ so $Z_t \perp \mathscr{M}_{-\infty}$. Hence $Y = P_{\mathscr{M}_{-\infty}} X_t$ is the deterministic component of the Wold decomposition and $Z_t = X_t - Y$ is the purely non-deterministic component.

For a stationary process $\{X_t\}$ satisfying the hypotheses of Theorem 5.7.1, the spectral distribution function F_X is the sum of two spectral distribution functions F_U and F_V corresponding to the two components $U_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}$ and V_t appearing in (5.7.1) (see Problem 4.7). From Chapter 4, F_U is absolutely continuous with respect to Lebesgue measure and has the spectral density

$$f_U(\lambda) = |\psi(e^{-i\lambda})|^2 \sigma^2/(2\pi), \qquad \text{where } \psi(e^{-i\lambda}) = \sum_{j=0}^{\infty} \psi_j e^{-ij\lambda}. \tag{5.7.9}$$

On the other hand, the spectral distribution F_V has no absolutely continuous component (see Doob (1953)). Consequently the Wold decomposition of a stationary process is analogous to the Lebesgue decomposition of the spectral measure into its absolutely continuous and singular parts. We state this as a theorem.

Theorem 5.7.2. *If* $\sigma^2 > 0$, *then*

$$F_X = F_U + F_V$$

where F_U *and* F_V *are respectively the absolutely continuous and singular components in the Lebesgue decomposition of* F_X. *The density function associated with* F_U *is defined by* (5.7.9).

The requirement $\sigma^2 > 0$ is critical in the above theorem. In other words it is possible for a deterministic process to have an absolutely continuous spectral distribution function. This is illustrated by Example 5.6.1. In the next section, a formula for σ^2 will be given in terms of the derivative of F_X which

is valid even in the case $\sigma^2 = 0$. This immediately yields a necessary and sufficient criterion for a stationary process to be deterministic.

§5.8* Kolmogorov's Formula

Let $\{X_t\}$ be a real-valued zero-mean stationary process with spectral distribution function F_X and let f denote the derivative of F_X (defined everywhere on $[-\pi, \pi]$ except possibly on a set of Lebesgue measure zero). We shall assume, to simplify the proof of the following theorem, that f is continuous on $[-\pi, \pi]$ and is bounded away from zero. Since $\{X_t\}$ is real, we must have $f(\lambda) = f(-\lambda)$, $0 \le \lambda \le \pi$. For a general proof, see Hannan (1970) or Ash and Gardner (1975).

Theorem 5.8.1 (Kolmogorov's Formula). *The one-step mean square prediction error of the stationary process* $\{X_t\}$ *is*

$$\sigma^2 = 2\pi \exp\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi} \ln f(\lambda)\, d\lambda\right\}. \tag{5.8.1}$$

PROOF. Using a Taylor series expansion of $\ln(1 - z)$ for $|z| < 1$ and the identity $\int_{-\pi}^{\pi} e^{ik\lambda}\, d\lambda = 0$, $k \ne 0$, we have for $|a| < 1$,

$$\begin{aligned}\int_{-\pi}^{\pi} \ln|1 - ae^{-i\lambda}|^2\, d\lambda &= \int_{-\pi}^{\pi} \ln(1 - ae^{-i\lambda})(1 - \bar{a}e^{i\lambda})\, d\lambda \\ &= \int_{-\pi}^{\pi}\left(\sum_{j=1}^{\infty} \frac{a^j e^{-ij\lambda}}{j} + \sum_{k=1}^{\infty} \frac{\bar{a}^k e^{ik\lambda}}{k}\right) d\lambda \\ &= 0. \end{aligned}\tag{5.8.2}$$

If $\{X_t\}$ is an AR(p) process satisfying $\phi(B)X_t = Z_t$, where $\{Z_t\} \sim \mathrm{WN}(0, \sigma^2)$ and $\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p \ne 0$ for $|z| \le 1$, then $\{X_t\}$ has spectral density,

$$g(\lambda) = \frac{\sigma^2}{2\pi}|1 - \phi_1 e^{-i\lambda} - \cdots - \phi_p e^{-ip\lambda}|^{-2} = \frac{\sigma^2}{2\pi}\prod_{j=1}^{p} |1 - a_j e^{-ij\lambda}|^{-2},$$

where $|a_j| < 1$, $j = 1, \ldots, p$. Hence

$$\int_{-\pi}^{\pi} \ln g(\lambda)\, d\lambda = \int_{-\pi}^{\pi} \ln \frac{\sigma^2}{2\pi}\, d\lambda - \sum_{j=1}^{p} \int_{-\pi}^{\pi} \ln|1 - a_j e^{-i\lambda}|^2\, d\lambda = 2\pi \ln \frac{\sigma^2}{2\pi},$$

establishing Kolmogorov's formula for causal AR processes.

Under the assumptions made on f, it is clear that $\min_{-\pi \le \lambda \le \pi} f(\lambda) > 0$. Moreover, it is easily shown from Corollary 4.4.2 that for any $\varepsilon \in (0, \min f(\lambda))$, there exist causal AR processes with spectral densities $g_\varepsilon^{(1)}$ and $g_\varepsilon^{(2)}$ such that

$$f(\lambda) - \varepsilon \le g_\varepsilon^{(1)}(\lambda) \le f(\lambda) \le g_\varepsilon^{(2)}(\lambda) \le f(\lambda) + \varepsilon. \tag{5.8.3}$$

Now define

$$\sigma_n^2(f) = E[(X_t - P_{\overline{sp}\{X_{t-1},\dots,X_{t-n}\}} X_t)^2]$$

$$= \min_{c_1,\dots,c_n} E(X_t - c_1 X_{t-1} - \cdots - c_n X_{t-n})^2$$

$$= \min_{c_1,\dots,c_n} \int_{-\pi}^{\pi} |1 - c_1 e^{-i\lambda} - \cdots - c_n e^{-in\lambda}|^2 f(\lambda)\, d\lambda.$$

By (5.8.3) and the definition of $\sigma_n^2(\cdot)$,

$$\sigma_n^2(g_\varepsilon^{(1)}) \le \sigma_n^2(f) \le \sigma_n^2(g_\varepsilon^{(2)}).$$

Since, by Problem 2.18, $\sigma_n^2(f) \to \sigma^2(f) := E[(X_t - P_{\overline{sp}\{X_s, -\infty < s < t\}} X_t)^2]$,

$$\sigma^2(g_\varepsilon^{(1)}) \le \sigma^2(f) \le \sigma^2(g_\varepsilon^{(2)}). \tag{5.8.4}$$

However we have already established that

$$\sigma^2(g_\varepsilon^{(i)}) = 2\pi \exp\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi} \ln g_\varepsilon^{(i)}(\lambda)\, d\lambda\right\}, \qquad i = 1, 2.$$

If follows therefore from (5.8.4) that $\sigma^2(f)$ must equal the common limit, as $\varepsilon \to 0$, of $\sigma^2(g_\varepsilon^{(1)})$ and $\sigma^2(g_\varepsilon^{(2)})$, i.e.

$$\sigma^2(f) = 2\pi \exp\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi} \ln f(\lambda)\, d\lambda\right\}. \qquad \square$$

Remark 1. Notice that $-\infty \le \int_{-\pi}^{\pi} \ln f(\lambda)\, d\lambda < \infty$ since $\ln f(\lambda) \le f(\lambda)$. If $\int_{-\pi}^{\pi} \ln f(\lambda)\, d\lambda = -\infty$, the theorem is still true with $\sigma^2 = 0$. Thus

$$\sigma^2 > 0 \quad \text{if and only if} \quad \int_{-\pi}^{\pi} \ln f(\lambda)\, d\lambda > -\infty,$$

and in this case $f(\lambda) > 0$ almost everywhere.

Remark 2. Equation (5.8.1) was first derived by Szegö in the absolutely continous case and was later extended by Kolmogorov to the general case. In the literature however it is usually referred to as Kolmogorov's formula.

EXAMPLE 5.8.1. For the process defined in Example 5.7.1, $F_U(d\lambda) = \sigma^2\, d\lambda/2\pi$ and $F_V(d\lambda) = \sigma^2 \delta_0(d\lambda)$ where δ_0 is the unit mass at the origin. Not surprisingly, the one-step mean square prediction error is therefore

$$2\pi \exp\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi} \ln\left(\frac{\sigma^2}{2\pi}\right) d\lambda\right\} = \sigma^2.$$

Problems

5.1. Let $\{X_t\}$ be a stationary process with mean μ. Show that

$$P_{\overline{sp}\{1, X_1, \dots, X_n\}} X_{n+h} = \mu + P_{\overline{sp}\{Y_1, \dots, Y_n\}} Y_{n+h},$$

where $\{Y_t\} = \{X_t - \mu\}$.

5.2. Suppose that $\{\mathscr{H}_n, n = 1, 2, \dots\}$ is a sequence of subspaces of a Hilbert space $\mathscr{H}$ with the property that $\mathscr{H}_n \subseteq \mathscr{H}_{n+1}$, $n = 1, 2, \dots$. Let $\mathscr{H}_\infty$ be the smallest closed subspace of $\mathscr{H}$ containing $\bigcup \mathscr{H}_n$, and let X be an element of $\mathscr{H}$. If $P_n X$ and $P_\infty X$ are the projections of X onto $\mathscr{H}_n$ and $\mathscr{H}_\infty$ respectively, show that
(a) $P_1 X, (P_2 - P_1)X, (P_3 - P_2)X, \dots$, are orthogonal,
(b) $\sum_{j=1}^{\infty} \|(P_{j+1} - P_j)X\|^2 < \infty$,
and
(c) $P_n X \to P_\infty X$.

5.3. Show that the converse of Proposition 5.1.1 is not true by constructing a stationary process $\{X_t\}$ such that Γ_n is non-singular for all n and $\gamma(h) \not\to 0$ as $h \to \infty$.

5.4. Suppose that $\{X_t\}$ is a stationary process with mean zero and spectral density

$$f_X(\lambda) = (\pi - |\lambda|)/\pi^2, \qquad -\pi \le \lambda \le \pi.$$

Find the coefficients $\{\theta_{ij}, j = 1, \dots, i; i = 1, \dots, 5\}$ and the mean squared errors $\{v_i, i = 0, \dots, 5\}$.

5.5. Let $\{X_t\}$ be the MA(1) process of Example 5.2.1. If $|\theta| < 1$, show that as $n \to \infty$,
(a) $\|X_n - \hat{X}_n - Z_n\| \to 0$,
(b) $v_n \to \sigma^2$,
and
(c) $\theta_{n1} \to \theta$. (Note that $\theta = E(X_{n+1} Z_n)\sigma^{-2}$ and $\theta_{n1} = v_{n-1}^{-1} E(X_{n+1}(X_n - \hat{X}_n))$.)

5.6. Let $\{X_t\}$ be the invertible MA(q) process

$$X_t = Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q}, \qquad \{Z_t\} \sim \text{WN}(0, \sigma^2).$$

Show that as $n \to \infty$,
(a) $\|X_n - \hat{X}_n - Z_n\| \to 0$,
(b) $v_n \to \sigma^2$,
and that
(c) there exist constants $K > 0$ and $c \in (0, 1)$ such that $|\theta_{nj} - \theta_j| \le K c^n$ for all n.

5.7. Verify equations (5.3.20) and (5.3.21).

5.8. The values .644, −.442, −.919, −1.573, .852, −.907, .686, −.753, −.954, .576, are simulated values of $X_1, \dots, X_{10}$ where $\{X_t\}$ is the ARMA(2, 1) process,

$$X_t - .1X_{t-1} - .12X_{t-2} = Z_t - .7Z_{t-1}, \qquad \{Z_t\} \sim \text{WN}(0, 1).$$

(a) Compute the forecasts $P_{10}X_{11}$, $P_{10}X_{12}$ and $P_{10}X_{13}$ and the corresponding mean squared errors.
(b) Assuming that $Z_t \sim N(0, 1)$, construct 95% prediction bounds for X_{11}, X_{12} and X_{13}.
(c) Using the method of Problem 5.15, compute $\tilde{X}_{11}^T$, $\tilde{X}_{12}^T$ and $\tilde{X}_{13}^T$ and compare these values with those obtained in (a).
[The simulated values of X_{11}, X_{12} and X_{13} were in fact .074, 1.097 and −.187 respectively.]

5.9. Repeat parts (a)–(c) of Problem 5.8 for the simulated values −1.222, 1.707, .049, 1.903, −3.341, 3.041, −1.012, −.779, 1.837, −3.693 of $X_1, \dots, X_{10}$, where $\{X_t\}$ is the MA(2) process

$$X_t = Z_t - 1.1Z_{t-1} + .28Z_{t-2}, \qquad \{Z_t\} \sim \text{WN}(0, 1).$$