3.4.3 Moving average processes

Suppose that $\{Z_t\}$ is a purely random process with mean zero and variance σ_Z^2. Then a process $\{X_t\}$ is said to be a moving average process of order q (abbreviated to an MA(q) process) if

$$X_t = \beta_0 Z_t + \beta_1 Z_{t-1} + \cdots + \beta_q Z_{t-q} \tag{3.2}$$

where $\{\beta_i\}$ are constants. The Zs are usually scaled so that $\beta_0 = 1$.

We find immediately that

$$E(X_t) = 0$$

$$\text{Var}(X_t) = \sigma_Z^2 \sum_{i=0}^{q} \beta_i^2$$

since the Zs are independent. We also have

$$\begin{aligned}
\gamma(k) &= \text{Cov}(X_t, X_{t+k}) \\
&= \text{Cov}(\beta_0 Z_t + \cdots + \beta_q Z_{t-q}, \beta_0 Z_{t+k} + \cdots + \beta_q Z_{t+k-q}) \\
&= \begin{cases} 0 & k > q \\ \sigma_Z^2 \sum_{i=0}^{q-k} \beta_i \beta_{i+k} & k = 0, 1, \ldots, q \\ \gamma(-k) & k < 0 \end{cases}
\end{aligned}$$

since

$$\text{Cov}(Z_s, Z_t) = \begin{cases} \sigma_Z^2 & s = t \\ 0 & s \neq t \end{cases}$$

As $\gamma(k)$ does not depend on t, and the mean is constant, the process is second-order stationary for all values of the $\{\beta_i\}$. Furthermore, if the Zs are normally distributed, then so are the Xs, and we have a strictly stationary normal process.

The ac.f. of the MA(q) process is given by

$$\rho(k) = \begin{cases} 1 & k = 0 \\ \sum_{i=0}^{q-k} \beta_i \beta_{i+k} \Big/ \sum_{i=0}^{q} \beta_i^2 & k = 1, \ldots, q \\ 0 & k > q \\ \rho(-k) & k < 0 \end{cases}$$

Note that the ac.f. 'cuts off' at lag q, which is a special feature of MA processes.

In particular the MA(1) process with $\beta_0 = 1$ has an ac.f. given by

$$\rho(k) = \begin{cases} 1 & k=0 \\ \beta_1/(1+\beta_1^2) & k=\pm 1 \\ 0 & \text{otherwise} \end{cases}$$

No other restrictions on the $\{\beta_i\}$ are required for a (finite-order) MA process to be stationary, but it is generally desirable to impose restrictions on the $\{\beta_i\}$ to ensure that the process satisfies a condition called **invertibility** (e.g. Box and Jenkins, 1970, p. 50). This condition may be explained in the following way. Consider the following first-order MA processes:

$$\text{A} \qquad X_t = Z_t + \theta Z_{t-1}$$

$$\text{B} \qquad X_t = Z_t + \frac{1}{\theta} Z_{t-1}$$

It can easily be shown that these two different processes have exactly the same ac.f. (Are you surprised? Then check $\rho(k)$ for models A and B.) Thus we cannot identify an MA process uniquely from a given ac.f. Now, if we express models A and B by putting Z_t in terms of $X_t, X_{t-1}, \ldots$, we find by successive substitution that

$$\text{A} \qquad Z_t = X_t - \theta X_{t-1} + \theta^2 X_{t-2} - \cdots$$

$$\text{B} \qquad Z_t = X_t - \frac{1}{\theta} X_{t-1} + \frac{1}{\theta^2} X_{t-2} - \cdots$$

If $|\theta| < 1$, the series for A converges whereas that for B does not. Thus an estimation procedure which involves estimating the residuals (see Section 4.3.1) will lead naturally to model A. Thus if $|\theta| < 1$, model A is said to be invertible whereas model B is not. The imposition of the invertibility condition ensures that there is a unique MA process for a given ac.f.

The invertibility condition for the general-order MA process is best expressed by using the backward shift operator, denoted by B, which is defined by

$$B^j X_t = X_{t-j} \qquad \text{for all } j$$

Then equation (3.2) may be written as

$$\begin{aligned} X_t &= (\beta_0 + \beta_1 B + \cdots + \beta_q B^q) Z_t \\ &= \theta(B) Z_t \end{aligned}$$

where $\theta(B)$ is a polynomial of order q in B. An MA process of order q is invertible if the roots of the equation (regarding B as a complex variable and not an operator)

$$\theta(B)=\beta_0+\beta_1 B+\cdots+\beta_q B^q=0$$

all lie outside the unit circle (Box and Jenkins, 1970, p. 50). For example, in the first-order case we have $\theta(B)=1+\theta B$, which has root $B=-1/\theta$. Thus the root is outside the unit circle provided that $|\theta|<1$.

MA processes have been used in many areas, particularly econometrics. For example economic indicators are affected by a variety of 'random' events such as strikes, government decisions, shortages of key materials and so on. Such events will not only have an immediate effect but may also affect economic indicators to a lesser extent in several subsequent periods, and so it is at least plausible that an MA process may be appropriate.

Note that an arbitrary constant, μ say, may be added to the right-hand side of equation (3.2) to give a process with mean μ. This does not affect the ac.f. and has been omitted for simplicity.

3.4.4 Autoregressive processes

Suppose that $\{Z_t\}$ is a purely random process with mean zero and variance σ_Z^2. Then a process $\{X_t\}$ is said to be an autoregressive process of order p if

$$X_t=\alpha_1 X_{t-1}+\cdots+\alpha_p X_{t-p}+Z_t \tag{3.3}$$

This is rather like a multiple regression model, but X_t is regressed not on independent variables but on past values of X_t; hence the prefix 'auto'. An autoregressive process of order p will be abbreviated to an AR(p) process.

(a) *First-order process*

For simplicity, we begin by examining the first-order case, where $p=1$. Then

$$X_t=\alpha X_{t-1}+Z_t \tag{3.4}$$

The AR(1) process is sometimes called the Markov process, after the Russian A. A. Markov. By successive substitution in (3.4) we may write

$$\begin{aligned} X_t &= \alpha(\alpha X_{t-2}+Z_{t-1})+Z_t \\ &= \alpha^2(\alpha X_{t-3}+Z_{t-2})+\alpha Z_{t-1}+Z_t \end{aligned}$$

and eventually we find that X_t may be expressed as an infinite-order MA process in the form (provided $-1<\alpha<+1$)

$$X_t=Z_t+\alpha Z_{t-1}+\alpha^2 Z_{t-2}+\cdots$$

This duality between AR and MA processes is useful for a variety of purposes. Rather than use successive substitution, it is simpler to use the backward shift operator B. Then equation (3.4) may be written

$$(1-\alpha B)X_t=Z_t$$

so that

$$X_t = Z_t/(1-\alpha B)$$
$$= (1+\alpha B+\alpha^2 B^2+\cdots)Z_t$$
$$= Z_t+\alpha Z_{t-1}+\alpha^2 Z_{t-2}+\cdots$$

When expressed in this form it is clear that

$$E(X_t)=0$$
$$\text{Var}(X_t)=\sigma_Z^2(1+\alpha^2+\alpha^4+\cdots)$$

Thus the variance is finite provided that $|\alpha|<1$, in which case

$$\text{Var}(X_t)=\sigma_X^2=\sigma_Z^2/(1-\alpha^2)$$

The acv.f. is given by

$$\gamma(k)=E[X_t X_{t+k}]$$
$$=E[(\Sigma\alpha^i Z_{t-i})(\Sigma\alpha^j Z_{t+k-j})]$$
$$=\sigma_Z^2 \sum_{i=0}^{\infty} \alpha^i\alpha^{k+i} \qquad \text{for } k\geqslant 0$$

which converges for $|\alpha|<1$ to

$$\gamma(k)=\alpha^k\sigma_Z^2/(1-\alpha^2)$$
$$=\alpha^k\sigma_X^2$$

For $k<0$, we find $\gamma(k)=\gamma(-k)$. Since $\gamma(k)$ does not depend on t, an AR process of order 1 is second-order stationary provided that $|\alpha|<1$. The ac.f. is given by

$$\rho(k)=\alpha^k \qquad k=0, 1, 2, \ldots$$

To get an even function defined for all integer k we can write

$$\rho(k)=\alpha^{|k|} \qquad k=0, \pm 1, \pm 2, \ldots$$

The ac.f. may also be obtained more simply by assuming *a priori* that the process is stationary, in which case $E(X_t)$ must be zero. Multiply through equation (3.4) by X_{t-k} (**not** X_{t+k}!) and take expectations. Then we find, for $k>0$, that

$$\gamma(-k)=\alpha\gamma(-k+1)$$

assuming that $E(Z_t X_{t-k})=0$ for $k>0$. Since $\gamma(k)$ is an even function, we must also have

$$\gamma(k)=\alpha\gamma(k-1) \qquad \text{for } k>0$$

Now $\gamma(0)=\sigma_X^2$, and so $\gamma(k)=\alpha^k\sigma_X^2$ for $k\geqslant 0$. Thus $\rho(k)=\alpha^k$ for $k\geqslant 0$. Now since

$|\rho(k)| \leqslant 1$, we must have $|\alpha| \leqslant 1$. But if $|\alpha| = 1$, then $|\rho(k)| = 1$ for all k, which is a degenerate case. Thus $|\alpha| < 1$ for a proper stationary process.

The above method of obtaining the ac.f. is often used, even though it involves 'cheating' a little by making an initial assumption of stationarity.

Three examples of the ac.f. of a first-order AR process are shown in Figure 3.1 for $\alpha = 0.8$, 0.3, -0.8. Note how quickly the ac.f. decays when $\alpha = 0.3$, and note how the ac.f. alternates when α is negative.

(b) *General-order case*

As in the first-order case, we can express an AR process of finite order as an MA process of infinite order. This may be done by successive substitution, or by using the backward shift operator. Then equation (3.3) may be written as

$$(1 - \alpha_1 B - \cdots - \alpha_p B^p) X_t = Z_t$$

or

$$\begin{aligned} X_t &= Z_t / (1 - \alpha_1 B - \cdots - \alpha_p B^p) \\ &= f(B) Z_t \end{aligned}$$

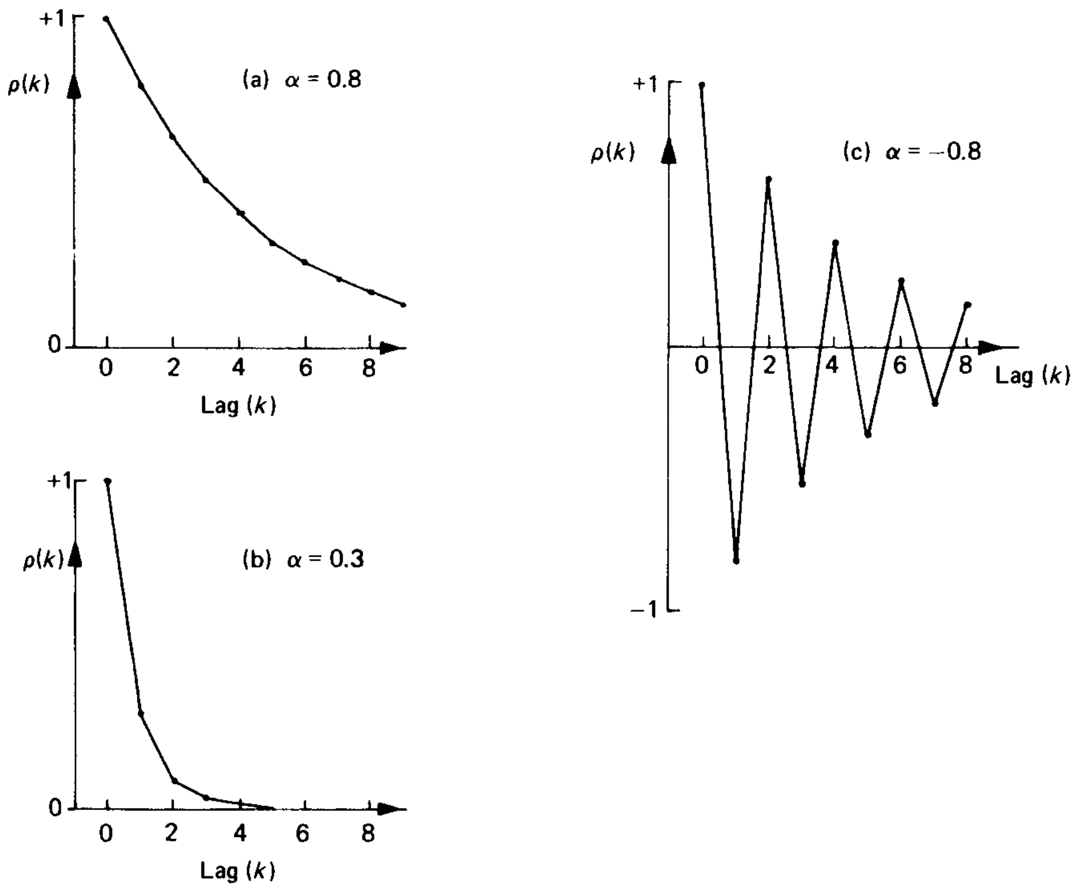

Figure 3.1 Three examples of the autocorrelation function of a first-order autoregressive process with, (a) $\alpha = 0.8$; (b) $\alpha = 0.3$; (c) $\alpha = -0.8$.

where

$$f(B) = (1 - \alpha_1 B - \cdots - \alpha_p B^p)^{-1}$$
$$= (1 + \beta_1 B + \beta_2 B^2 + \cdots)$$

The relationship between the αs and the βs may then be found. Having expressed X_t as an MA process, it follows that $E(X_t) = 0$. The variance is finite provided that $\Sigma \beta_i^2$ converges, and this is a necessary condition for stationarity. The acv.f. is given by

$$\gamma(k) = \sigma_Z^2 \sum_{i=0}^{\infty} \beta_i \beta_{i+k} \qquad \text{where } \beta_0 = 1$$

A sufficient condition for this to converge, and hence for stationarity, is that $\Sigma|\beta_i|$ converges.

We can in principle find the ac.f. of the general-order AR process using the above procedure, but the $\{\beta_i\}$ may be algebraically hard to find. The alternative simpler way is to **assume** the process is stationary, multiply through equation (3.3) by X_{t-k}, take expectations, and divide by σ_X^2, assuming that the variance of X_t is finite. Then, using the fact that $\rho(k) = \rho(-k)$ for all k, we find

$$\rho(k) = \alpha_1 \rho(k-1) + \cdots + \alpha_p \rho(k-p) \qquad \text{for all } k > 0$$

This set of equations is called the Yule-Walker equations after G. U. Yule and Sir Gilbert Walker. It is a set of difference equations and has the general solution

$$\rho(k) = A_1 \pi_1^{|k|} + \cdots + A_p \pi_p^{|k|}$$

where $\{\pi_i\}$ are the roots of the so-called auxiliary equation

$$y^p - \alpha_1 y^{p-1} \cdots - \alpha_p = 0$$

The constants $\{A_i\}$ are chosen to satisfy the initial conditions depending on $\rho(0) = 1$, which means that $\Sigma A_i = 1$. The first $(p-1)$ Yule-Walker equations provide $(p-1)$ further restrictions on the $\{A_i\}$ using $\rho(0) = 1$ and $\rho(k) = \rho(-k)$.

From the general form of $\rho(k)$, it is clear that $\rho(k)$ tends to zero as k increases provided that $|\pi_i| < 1$ for all i, and this is a necessary and sufficient condition for the process to be stationary.

An equivalent way of expressing the stationarity condition is to say that the roots of the equation

$$\phi(B) = 1 - \alpha_1 B - \cdots - \alpha_p B^p = 0 \tag{3.5}$$

must lie outside the unit circle (Box and Jenkins, 1970, Section 3.2).

Of particular interest is the AR(2) process, when π_1, π_2 are the roots of the quadratic equation

$$y^2 - \alpha_1 y - \alpha_2 = 0$$

Thus $|\pi_i| < 1$ if

$$\left| \frac{\alpha_1 \pm \sqrt{(\alpha_1^2 + 4\alpha_2)}}{2} \right| < 1$$

from which it can be shown (Exercise 3.6) that the stationarity region is the triangular region satisfying

$$\alpha_1 + \alpha_2 < 1$$
$$\alpha_1 - \alpha_2 > -1$$
$$\alpha_2 > -1$$

The roots are real if $\alpha_1^2 + 4\alpha_2 > 0$, in which case the ac.f. decreases exponentially with k, but the roots are complex if $\alpha_1^2 + 4\alpha_2 < 0$, in which case we find that the ac.f. is a damped cosine wave. (See Example 3.1 at the end of this section.)

When the roots are real, the constants A_1, A_2 are found as follows. Since $\rho(0) = 1$, we have

$$A_1 + A_2 = 1$$

From the first of the Yule-Walker equations, we have

$$\begin{aligned}\rho(1) &= \alpha_1 \rho(0) + \alpha_2 \rho(-1) \\ &= \alpha_1 + \alpha_2 \rho(1)\end{aligned}$$

Thus

$$\begin{aligned}\rho(1) &= \alpha_1/(1 - \alpha_2) \\ &= A_1 \pi_1 + A_2 \pi_2 \\ &= A_1 \pi_1 + (1 - A_1)\pi_2\end{aligned}$$

Hence we find

$$\begin{aligned}A_1 &= [\alpha_1/(1 - \alpha_2) - \pi_2]/(\pi_1 - \pi_2) \\ A_2 &= 1 - A_1\end{aligned}$$

AR processes have been applied to many situations in which it is reasonable to assume that the present value of a time series depends on the immediate past values together with a random error. For simplicity we have only considered processes with mean zero, but non-zero means may be dealt with by rewriting equation (3.3) in the form

$$X_t - \mu = \alpha_1(X_{t-1} - \mu) + \cdots + \alpha_p(X_{t-p} - \mu) + Z_t$$

This does not affect the ac.f.

Example 3.1 Consider the AR(2) process given by

$$X_t = X_{t-1} - \tfrac{1}{2}X_{t-2} + Z_t$$

Is this process stationary? If so, what is its ac.f.?

In order to answer the first question we find the roots of equation (3.5), which in this case is

$$\phi(B) = 1 - B + \tfrac{1}{2}B^2 = 0$$

The roots of this equation (regarding B as a variable) are $1 \pm \mathrm{i}$. As the modulus of both roots exceeds one, the roots are both outside the unit circle and so the process is stationary.

In order to find the ac.f. of the process, we use the first Yule-Walker equation to give

$$\begin{aligned}\rho(1) &= \rho(0) - \tfrac{1}{2}\rho(-1)\\ &= 1 - \tfrac{1}{2}\rho(1)\end{aligned}$$

giving $\rho(1) = 2/3$.

For $k \geqslant 2$, the Yule-Walker equations are

$$\rho(k) = \rho(k-1) - \tfrac{1}{2}\rho(k-2)$$

We could find $\rho(2)$, then $\rho(3)$, and so on by successive substitution, but it is easier to find the general solution by solving as a difference equation, which has the auxiliary equation

$$y^2 - y + \tfrac{1}{2} = 0$$

with roots $y = (1 \pm \mathrm{i})/2 = [\cos(\pi/4) \pm \mathrm{i}\sin(\pi/4)]/\sqrt{2} = \mathrm{e}^{\pm \mathrm{i}\pi/4}/\sqrt{2}$. Since $\alpha_1^2 + 4\alpha_2 = (1-2)$ is less than zero, the ac.f. is a damped cosine wave. Using $\rho(0) = 1$ and $\rho(1) = 2/3$, some messy trigonometry and algebra gives

$$\rho_k = \left(\frac{1}{\sqrt{2}}\right)^k \left(\cos\frac{\pi k}{4} + \frac{1}{3}\sin\frac{\pi k}{4}\right)$$

for $k = 0, 1, 2, \ldots$.

3.4.5 Mixed ARMA models

A useful class of models for time series is formed by combining MA and AR processes. A mixed autoregressive/moving-average process containing p AR terms and q MA terms is said to be an ARMA process of order (p, q). It is given by

$$\begin{aligned}X_t = \alpha_1 X_{t-1} + \cdots + \alpha_p X_{t-p} + Z_t + \beta_1 Z_{t-1}\\ + \cdots + \beta_q Z_{t-q}\end{aligned} \qquad (3.6)$$

Using the backward shift operator B, equation (3.6) may be written in the form

$$\phi(B)X_t=\theta(B)Z_t \tag{3.6a}$$

where $\phi(B)$, $\theta(B)$ are polynomials of order p, q respectively, such that

$$\phi(B)=1-\alpha_1 B-\cdots-\alpha_p B^p$$

and

$$\theta(B)=1+\beta_1 B+\cdots+\beta_q B^q$$

As for an AR process, the values of $\{\alpha_i\}$ which make the process stationary are such that the roots of

$$\phi(B)=0$$

lie outside the unit circle. As for an MA process, the values of $\{\beta_i\}$ which make the process invertible are such that the roots of

$$\theta(B)=0$$

lie outside the unit circle.

It is straightforward in principle, though algebraically rather tedious, to calculate the ac.f. of an ARMA process, but this will not be discussed here. (See Exercise 3.11; and see Box and Jenkins, 1970, Section 3.4).

The importance of ARMA processes lies in the fact that a stationary time series may often be described by an ARMA model involving fewer parameters than a pure MA or AR process by itself.

It is sometimes helpful to express an ARMA model as a pure MA process in the form

$$X_t=\psi(B)Z_t \tag{3.6b}$$

where $\psi(B)=\Sigma\psi_i B^i$ is the MA operator which may be of infinite order. The ψ weights, $\{\psi_i\}$, can be useful in calculating forecasts (see Chapter 5) and in assessing the properties of a model (e.g. see Exercise 3.11). By comparison with equation (3.6a), we see that $\psi(B)=\theta(B)/\phi(B)$. Alternatively, it can be helpful to express an ARMA model as a pure AR process in the form

$$\pi(B)X_t=Z_t \tag{3.6c}$$

where $\pi(B)=\phi(B)/\theta(B)$. By convention we write $\pi(B)=1-\sum_{i\geqslant 1}\pi_i B^i$, since the natural way to write an AR model is in the form

$$X_t=\sum_{i=1}^{\infty}\pi_i X_{t-i}+Z_t$$

By comparing (3.6b) and (3.6c), we see that

$$\pi(B)\psi(B)=1$$

The ψ weights or π weights may be obtained directly by division or by equating powers of B in an equation such as

$$\psi(B)\phi(B)=\theta(B)$$

Example 3.2 Find the ψ weights and π weights for the ARMA(1, 1) process given by

$$X_t=0.5X_{t-1}+Z_t-0.3Z_{t-1}$$

Here $\phi(B)=(1-0.5B)$ and $\theta(B)=(1-0.3B)$, so the process is stationary and invertible. Then

$$\begin{aligned}\psi(B)=\theta(B)/\phi(B)&=(1-0.3B)(1-0.5B)^{-1}\\&=(1-0.3B)(1+0.5B+0.5^2B^2+\cdots)\\&=1+0.2B+0.1B^2+0.005B^3+\cdots\end{aligned}$$

Hence

$$\psi_i=0.2\times0.5^{i-1}\qquad\text{for } i=1,2,\ldots$$

Similarly we find

$$\pi_i=0.2\times0.3^{i-1}\qquad\text{for } i=1,2,\ldots$$

Note that both the ψ weights and π weights die away quickly, and this also indicates a stationary, invertible process.

3.4.6 Integrated ARIMA models

In practice most time series are non-stationary. In order to fit a stationary model, such as those discussed in Sections 3.4.3–3.4.5, it is necessary to remove non-stationary sources of variation. If the observed time series is non-stationary in the mean then we can difference the series, as suggested in Section 2.5.3, and this approach is widely used in econometrics. If X_t is replaced by $\nabla^d X_t$ in equation (3.6) then we have a model capable of describing certain types of non-stationary series. Such a model is called an 'integrated' model because the stationary model which is fitted to the differenced data has to be summed or 'integrated' to provide a model for the non-stationary data. Writing

$$W_t=\nabla^d X_t=(1-B)^d X_t$$

the general autoregressive integrated moving average process (abbreviated ARIMA process) is of the form

$$W_t=\alpha_1 W_{t-1}+\cdots+\alpha_p W_{t-p}+Z_t+\cdots+\beta_q Z_{t-q} \tag{3.7}$$

By analogy with equation (3.6a), we may write equation (3.7) in the form

$$\phi(B)W_t = \theta(B)Z_t \tag{3.7a}$$

or

$$\phi(B)(1-B)^d X_t = \theta(B)Z_t \tag{3.7b}$$

Thus we have an ARMA(p, q) model for W_t, while the model in equation (3.7b), describing the dth differences of X_t, is said to be an ARIMA process of order (p, d, q). The model for X_t is clearly non-stationary, as the AR operator $\phi(B)(1-B)^d$ has d roots on the unit circle. In practice the value of d is often taken to be one. Note that the random walk can be regarded as an ARIMA(0, 1, 0) process.

ARIMA models can be generalized to include seasonal terms, as discussed in Section 4.6.

†3.4.7 The general linear process

The MA process, of possibly infinite order, is given by

$$X_t = \sum_{i=0}^{\infty} \psi_i Z_{t-i} \tag{3.8}$$

in the notation of equation (3.6b), where $\{Z_t\}$ denotes a purely random process. A sufficient condition for the sum to converge and for the process to be stationary is that $\sum_{i=0}^{\infty} \psi_i^2 < \infty$. The process described by equation (3.8) is sometimes called a **general linear process**. Stationary AR, MA, and ARMA processes are special cases of this model and the duality between AR and MA processes is easily demonstrated using equations (3.6b) and (3.6c).

†3.4.8 Continuous process

So far, we have only considered stochastic processes in discrete time, because these are the main type of process the statistician uses in practice. Continuous-time processes have been used in some applications, notably in the study of control theory by electrical engineers. Here we shall only indicate some of the problems connected with their use.

By analogy with a discrete-time purely random process, we might expect to define a continuous-time purely random process as having an ac.f. given by

$$\rho(\tau) = \begin{cases} 1 & \tau = 0 \\ 0 & \tau \neq 0 \end{cases}$$

However, this is a discontinuous function, and it can be shown that such a process would have an infinite variance and hence be a physically unrealizable

†These sections should be omitted at first reading.

phenomenon. Nevertheless, some processes which arise in practice do appear to have the properties of continuous-time white noise even when sampled at quite small discrete intervals. We may approximate continuous-time white noise by considering a purely random process in discrete time at intervals Δt, and letting $\Delta t \to 0$, or by considering a process in continuous time with ac.f. $\rho(\tau) = e^{-\lambda|\tau|}$ and letting $\lambda \to \infty$ so that the ac.f. decays very quickly.

As an example of the difficulties involved with continuous-time processes, we briefly consider a first-order continuous AR process. A first-order discrete AR process may be written in terms of X_t, ∇X_t and Z_t. As differencing in discrete time corresponds to differentiation in continuous time, a natural way of trying to define a continuous first-order AR process is by

$$\frac{dX(t)}{dt} + aX(t) = Z(t) \tag{3.9}$$

where a is a constant, and $Z(t)$ denotes continuous white noise. In the theory of Brownian motion, this is called Langevin's equation. However, as $Z(t)$ does not physically exist, it is more legitimate to write equation (3.9) in a form involving infinitesimal small changes as

$$dX(t) + aX(t)\, dt = dU(t) \tag{3.10}$$

where $\{U(t)\}$ is a process with orthogonal increments such that the random variables $[U(t_2) - U(t_1)]$ and $[U(t_4) - U(t_3)]$ are uncorrelated for any two non-overlapping intervals (t_1, t_2) and (t_3, t_4). It can then be shown that the process $X(t)$ defined in equation (3.10) has ac.f.

$$\rho(\tau) = e^{-a|\tau|}$$

which is similar to the ac.f. of a first-order discrete AR process in that both decay exponentially. However, the rigorous study of continuous processes, such as that in equation (3.9), requires considerable mathematical machinery, including a knowledge of stochastic integration, and we will not pursue it here. The reader is referred for example to Yaglom (1962) and Cox and Miller (1968, Section 7.4).

†3.5 THE WOLD DECOMPOSITION THEOREM

This section gives a brief introduction to a famous result, called the Wold decomposition theorem, which is of mainly theoretical interest. The treatment in this section is a shortened version of that given by Cox and Miller (1968). The Wold decomposition theorem says that any discrete stationary process can be expressed as the sum of two uncorrelated processes, one purely deterministic and one purely indeterministic. The terms 'deterministic' and

†This section should be omitted at first reading.

'indeterministic' are defined as follows. We can regress X_t on $(X_{t-q}, X_{t-q-1}, \ldots)$ and denote the residual variance from the resulting linear regression model by τ_q^2. As $\tau_q^2 \leqslant \mathrm{Var}(X_t)$, it is clear that, as q increases, τ_q^2 is a non-decreasing bounded sequence and therefore tends to a limit as $q \rightarrow \infty$. If $\lim_{q\rightarrow\infty}\tau_q^2 = \mathrm{Var}(X_t)$ then linear regression on the remote past is useless for prediction purposes, and we say that $\{X_t\}$ is **purely indeterministic**. But if $\lim_{q\rightarrow\infty}\tau_q^2$ is zero then the process can be forecast exactly, and we say that $\{X_t\}$ is purely **deterministic**.

All the stationary processes we have considered in this chapter, such as AR and MA processes, are purely indeterministic. The best-known examples of purely deterministic processes are sinusoidal processes (see Exercise 3.14), such as

$$X_t = g\cos(\omega t + \theta) \tag{3.11}$$

where g is a constant, ω is a constant in $(0, \pi)$ called the frequency of the process, and θ is a random variable called the phase which is uniformly distributed on $(0, 2\pi)$ but which is fixed for a single realization. Note that we must include the term θ so that

$$E(X_t) = 0 \qquad \text{for all } t$$

otherwise (3.11) would not define a stationary process. As θ is fixed for a single realization, once enough values of X_t have been observed to evaluate θ, all subsequent values of X_t are completely determined. It is then obvious that (3.11) defines a purely deterministic process.

The Wold decomposition theorem also says that the purely indeterministic component can be written as the linear sum of an 'innovation' process $\{Z_t\}$, which is a sequence of uncorrelated random variables. A special class of processes of particular interest arise when the Zs are independent and not merely uncorrelated, as we then have a general linear process (Section 3.4.7). On the other hand when processes are generated in a non-linear way the Wold decomposition is usually of little interest.

The concept of a purely indeterministic process is a useful one, and most of the stationary stochastic processes which are considered in the rest of this book are of this type.

EXERCISES

In all the following questions $\{Z_t\}$ is a discrete, purely random process, such that $E(Z_t)=0$, $\mathrm{Var}(Z_t)=\sigma_Z^2$, $\mathrm{Cov}(Z_t, Z_{t+k})=0$, $k \neq 0$.

Exercise 3.14 is harder than the others and may be omitted.

3.1 Find the ac.f. of the second-order MA process given by

$$X_t = Z_t + 0.7Z_{t-1} - 0.2Z_{t-2}$$