LECTURE 3

Linear Stochastic Models

Autcovariances of a Stationary Process

A temporal stochastic process is simply a sequence of random variables indexed by a time subscript. Such a process can be denoted by $x(t)$. The element of the sequence at the point $t = \tau$ is $x_\tau = x(\tau)$.

Let $\{x_{\tau+1}, x_{\tau+2}, \ldots, x_{\tau+n}\}$ denote n consecutive elements of the sequence. Then the process is said to be strictly stationary if the joint probability distribution of the elements does not depend on τ regardless of the size of n. This means that any two segments of the sequence of equal length have identical probability density functions. In consequence, the decision on where to place the time origin is arbitrary; and the argument τ can be omitted. Some further implications of stationarity are that

$$E(x_t) = \mu < \infty \quad \text{for all} \quad t \quad \text{and} \quad C(x_{\tau+t}, x_{\tau+s}) = \gamma_{|t-s|}. \tag{1}$$

The latter condition means that the covariance of any two elements depends only on their temporal separation $|t - s|$. Notice that, if the elements of the sequence are normally distributed, then the two conditions are sufficient to establish strict stationarity. On their own, they constitute the conditions of weak or 2nd-order stationarity.

The condition on the covariances implies that the dispersion matrix of the vector $[x_1, x_2, \ldots, x_n]$ is a bisymmetric Laurent matrix of the form

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

wherein the generic element in the (i, j)th position is $\gamma_{|i-j|} = C(x_i, x_j)$. Given that a sequence of observations of a time series represents only a segment of a single realisation of a stochastic process, one might imagine that there is little chance of making valid inferences about the parameters of the process.

However, provided that the process $x(t)$ is stationary and provided that the statistical dependencies between widely separated elements of the sequence are weak, it is possible to estimate consistently those parameters of the process which express the dependence of proximate elements of the sequence. If one is prepared to make sufficiently strong assumptions about the nature of the process, then a knowledge of such parameters may be all that is needed for a complete characterisation of the process.

Moving-Average Processes

The qth-order moving average process, or MA(q) process, is defined by the equation

$$y(t) = \mu_0\varepsilon(t) + \mu_1\varepsilon(t-1) + \cdots + \mu_q\varepsilon(t-q), \tag{3}$$

where $\varepsilon(t)$, which has $E\{\varepsilon(t)\} = 0$, is a white-noise process consisting of a sequence of independently and identically distributed random variables with zero expectations. The equation is normalised either by setting $\mu_0 = 1$ or by setting $V\{\varepsilon(t)\} = \sigma_\varepsilon^2 = 1$. The equation can be written in summary notation as $y(t) = \mu(L)\varepsilon(t)$, where $\mu(L) = \mu_0 + \mu_1 L + \cdots + \mu_q L^q$ is a polynomial in the lag operator.

A moving-average process is clearly stationary since any two elements y_t and y_s represent the same function of the vectors $[\varepsilon_t, \varepsilon_{t-1}, \ldots, \varepsilon_{t-q}]$ and $[\varepsilon_s, \varepsilon_{s-1}, \ldots, \varepsilon_{s-q}]$ which are identically distributed. In addition to the condition of stationarity, it is usually required that a moving-average process should be invertible such that it can be expressed in the form of $\mu^{-1}(L)y(t) = \varepsilon(t)$ where the LHS embodies a convergent sum of past values of $y(t)$. This is an infinite-order autoregressive representation of the process. The representation is available only if all the roots of the equation $\mu(z) = \mu_0 + \mu_1 z + \cdots + \mu_q z^q = 0$ lie outside the unit circle. This conclusion follows from our discussion of partial fractions.

As an example, let us consider the first-order moving-average process which is defined by

$$y(t) = \varepsilon(t) - \theta\varepsilon(t-1) = (1 - \theta L)\varepsilon(t). \tag{4}$$

Provided that $|\theta| < 1$, this can be written in autoregressive form as

$$\begin{aligned} \varepsilon(t) &= (1 - \theta L)^{-1} y(t) \\ &= \big\{y(t) + \theta y(t-1) + \theta^2 y(t-2) + \cdots\big\}. \end{aligned} \tag{5}$$

Imagine that $|\theta| > 1$ instead. Then, to obtain a convergent series, we have to write

$$\begin{aligned} y(t+1) &= \varepsilon(t+1) - \theta\varepsilon(t) \\ &= -\theta(1 - L^{-1}/\theta)\varepsilon(t), \end{aligned} \tag{6}$$

where $L^{-1}\varepsilon(t) = \varepsilon(t+1)$. This gives

$$\begin{aligned}\varepsilon(t) &= -\theta^{-1}(1 - L^{-1}/\theta)^{-1}y(t+1) \\ &= -\theta^{-1}\big\{y(t+1)/\theta + y(t+2)/\theta^2 + y(t-3)/\theta^3 + \cdots\big\}.\end{aligned} \tag{7}$$

Normally, an expression such as this, which embodies future values of $y(t)$, would have no reasonable meaning.

It is straightforward to generate the sequence of autocovariances from a knowledge of the parameters of the moving-average process and of the variance of the white-noise process. Consider

$$\begin{aligned}\gamma_\tau &= E(y_t y_{t-\tau}) \\ &= E\Big\{\sum_i \mu_i\varepsilon_{t-i}\sum_j \mu_j\varepsilon_{t-\tau-j}\Big\} \\ &= \sum_i\sum_j \mu_i\mu_j E(\varepsilon_{t-i}\varepsilon_{t-\tau-j}).\end{aligned} \tag{8}$$

Since $\varepsilon(t)$ is a sequence of independently and identically distributed random variables with zero expectations, it follows that

$$E(\varepsilon_{t-i}\varepsilon_{t-\tau-j}) = \begin{cases} 0, & \text{if } i \neq \tau + j; \\ \sigma_\varepsilon^2, & \text{if } i = \tau + j. \end{cases} \tag{9}$$

Therefore

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

Now let $\tau = 0, 1, \ldots, q$. This gives

$$\begin{aligned}\gamma_0 &= \sigma_\varepsilon^2(\mu_0^2 + \mu_1^2 + \cdots + \mu_q^2), \\ \gamma_1 &= \sigma_\varepsilon^2(\mu_0\mu_1 + \mu_1\mu_2 + \cdots + \mu_{q-1}\mu_q), \\ &\vdots \\ \gamma_q &= \sigma_\varepsilon^2\mu_0\mu_q.\end{aligned} \tag{11}$$

Also, $\gamma_\tau = 0$ for all $\tau > q$.

The first-order moving-average process $y(t) = \varepsilon(t) - \theta\varepsilon(t-1)$ has the following autocovariances:

$$\begin{aligned}\gamma_0 &= \sigma_\varepsilon^2(1 + \theta^2), \\ \gamma_1 &= -\sigma_\varepsilon^2\theta, \\ \gamma_\tau &= 0 \quad \text{if} \quad \tau > 1.\end{aligned} \tag{12}$$

Thus, for a vector $y = [y_1, y_2, \ldots, y_T]'$ of T consecutive elements from a first-order moving-average process, the dispersion matrix is

$$D(y) = \sigma_\varepsilon^2 \begin{bmatrix} 1+\theta^2 & -\theta & 0 & \ldots & 0 \\ -\theta & 1+\theta^2 & -\theta & \ldots & 0 \\ 0 & -\theta & 1+\theta^2 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 1+\theta^2 \end{bmatrix}. \tag{13}$$

In general, the dispersion matrix of a qth-order moving-average process has q subdiagonal and q supradiagonal bands of nonzero elements and zero elements elsewhere.

It is also helpful to define an autocovariance generating function which is a power series whose coefficients are the autocovariances γ_τ for successive values of τ. This is denoted by

$$\gamma(z) = \sum_\tau \gamma_\tau z^\tau; \quad \text{with} \quad \tau = \{0, \pm 1, \pm 2, \ldots\} \quad \text{and} \quad \gamma_\tau = \gamma_{-\tau}. \tag{14}$$

The generating function is also called the z-transform of the autocovariance function.

The autocovariance generating function of the qth-order moving-average process can be found quite readily. Consider the convolution

$$\begin{aligned} \mu(z)\mu(z^{-1}) &= \sum_i \mu_i z^i \sum_j \mu_j z^{-j} \\ &= \sum_i \sum_j \mu_i \mu_j z^{i-j} \\ &= \sum_\tau \Big(\sum_j \mu_i \mu_{j+\tau} \Big) z^\tau, \quad \tau = i - j. \end{aligned} \tag{15}$$

By referring to the expression for the autocovariance of lag τ of a moving-average process given under (10), it can be seen that the autocovariance generating function is just

$$\gamma(z) = \sigma_\varepsilon^2 \mu(z)\mu(z^{-1}). \tag{16}$$

Autoregressive Processes

The pth-order autoregressive process, or AR(p) process, is defined by the equation

$$\alpha_0 y(t) + \alpha_1 y(t-1) + \cdots + \alpha_p y(t-p) = \varepsilon(t). \tag{17}$$

This equation is invariably normalised by setting $\alpha_0 = 1$, although it would be possible to set $\sigma_\varepsilon^2 = 1$ instead. The equation can be written in summary notation as $\alpha(L)y(t) = \varepsilon(t)$, where $\alpha(L) = \alpha_0 + \alpha_1 L + \cdots + \alpha_p L^p$. For the process to be stationary, the roots of the equation $\alpha(z) = \alpha_0 + \alpha_1 z + \cdots + \alpha_p z^p = 0$ must lie outside the unit circle. This condition enables us to write the autoregressive process as an infinite-order moving-average process in the form of $y(t) = \alpha^{-1}(L)\varepsilon(t)$.

As an example, let us consider the first-order autoregressive process which is defined by

$$\begin{aligned} \varepsilon(t) &= y(t) - \phi y(t-1) \\ &= (1 - \phi L)y(t). \end{aligned} \tag{18}$$

Provided that the process is stationary with $|\phi| < 1$, it can be represented in moving-average form as

$$\begin{aligned} y(t) &= (1 - \phi L)^{-1}\varepsilon(t) \\ &= \big\{\varepsilon(t) + \phi\varepsilon(t-1) + \phi^2\varepsilon(t-2) + \cdots\big\}. \end{aligned} \tag{19}$$

The autocovariances of the process can be found by using the formula of (10) which is applicable to moving-average process of finite or infinite order. Thus

$$\begin{aligned} \gamma_\tau &= E(y_t y_{t-\tau}) \\ &= E\Big\{\sum_i \phi^i \varepsilon_{t-i} \sum_j \phi^j \varepsilon_{t-\tau-j}\Big\} \\ &= \sum_i \sum_j \phi^i \phi^j E(\varepsilon_{t-i}\varepsilon_{t-\tau-j}); \end{aligned} \tag{20}$$

and the result under (9) indicates that

$$\begin{aligned} \gamma_\tau &= \sigma_\varepsilon^2 \sum_j \phi^j \phi^{j+\tau} \\ &= \frac{\sigma_\varepsilon^2 \phi^\tau}{1 - \phi^2}. \end{aligned} \tag{21}$$

For a vector $y = [y_1, y_2, \ldots, y_T]'$ of T consecutive elements from a first-order autoregressive process, the dispersion matrix has the form

$$D(y) = \frac{\sigma_\varepsilon^2}{1 - \phi^2} \begin{bmatrix} 1 & \phi & \phi^2 & \cdots & \phi^{T-1} \\ \phi & 1 & \phi & \cdots & \phi^{T-2} \\ \phi^2 & \phi & 1 & \cdots & \phi^{T-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \phi^{T-1} & \phi^{T-2} & \phi^{T-3} & \cdots & 1 \end{bmatrix}. \tag{22}$$

To find the autocovariance generating function for the general pth-order autoregressive process, we may consider again the function $\alpha(z) = \sum_i \alpha_i z^i$. Since an autoregressive process may be treated as an infinite-order moving-average process, it follows that

$$\gamma(z) = \frac{\sigma_\varepsilon^2}{\alpha(z)\alpha(z^{-1})}. \tag{23}$$

For an alternative way of finding the autocovariances of the pth-order process, consider multiplying $\sum_i \alpha_i y_{t-i} = \varepsilon_t$ by $y_{t-\tau}$ and taking expectations to give

$$\sum_i \alpha_i E(y_{t-i}y_{t-\tau}) = E(\varepsilon_t y_{t-\tau}). \tag{24}$$

Taking account of the normalisation $\alpha_0 = 1$, we find that

$$E(\varepsilon_t y_{t-\tau}) = \begin{cases} \sigma_\varepsilon^2, & \text{if } \tau = 0; \\ 0, & \text{if } \tau > 0. \end{cases} \tag{25}$$

Therefore, on setting $E(y_{t-i}y_{t-\tau}) = \gamma_{\tau-i}$, equation (24) gives

$$\sum_i \alpha_i \gamma_{\tau-i} = \begin{cases} \sigma_\varepsilon^2, & \text{if } \tau = 0; \\ 0, & \text{if } \tau > 0. \end{cases} \tag{26}$$

The second of these is a homogeneous difference equation which enables us to generate the sequence $\{\gamma_p, \gamma_{p+1}, \ldots\}$ once p starting values $\gamma_0, \gamma_1, \ldots, \gamma_{p-1}$ are known. By letting $\tau = 0, 1, \ldots, p$ in (26), we generate a set of $p+1$ equations which can be arrayed in matrix form as follows:

$$\begin{bmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \cdots & \gamma_p \\ \gamma_1 & \gamma_0 & \gamma_1 & \cdots & \gamma_{p-1} \\ \gamma_2 & \gamma_1 & \gamma_0 & \cdots & \gamma_{p-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \gamma_p & \gamma_{p-1} & \gamma_{p-2} & \cdots & \gamma_0 \end{bmatrix} \begin{bmatrix} 1 \\ \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_p \end{bmatrix} = \begin{bmatrix} \sigma_\varepsilon^2 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{27}$$

These are called the Yule–Walker equations, and they can be used either for generating the values $\gamma_0, \gamma_1, \ldots, \gamma_p$ from the values $\alpha_1, \ldots, \alpha_p, \sigma_\varepsilon^2$ or vice versa.

For an example of the two uses of the Yule–Walker equations, let us consider the second-order autoregressive process. In that case, we have

$$\begin{aligned} \begin{bmatrix} \gamma_0 & \gamma_1 & \gamma_2 \\ \gamma_1 & \gamma_0 & \gamma_1 \\ \gamma_2 & \gamma_1 & \gamma_0 \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{bmatrix} &= \begin{bmatrix} \alpha_2 & \alpha_1 & \alpha_0 & 0 & 0 \\ 0 & \alpha_2 & \alpha_1 & \alpha_0 & 0 \\ 0 & 0 & \alpha_2 & \alpha_1 & \alpha_0 \end{bmatrix} \begin{bmatrix} \gamma_2 \\ \gamma_1 \\ \gamma_0 \\ \gamma_1 \\ \gamma_2 \end{bmatrix} \\ &= \begin{bmatrix} \alpha_0 & \alpha_1 & \alpha_2 \\ \alpha_1 & \alpha_0 + \alpha_2 & 0 \\ \alpha_2 & \alpha_1 & \alpha_0 \end{bmatrix} \begin{bmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \end{bmatrix} = \begin{bmatrix} \sigma_\varepsilon^2 \\ 0 \\ 0 \end{bmatrix}. \end{aligned} \tag{28}$$

Given $\alpha_0 = 1$ and the values for $\gamma_0, \gamma_1, \gamma_2$, we can find σ_ε^2 and α_1, α_2. Conversely, given $\alpha_0, \alpha_1, \alpha_2$ and σ_ε^2, we can find $\gamma_0, \gamma_1, \gamma_2$. It is worth recalling at this juncture that the normalisation $\sigma_\varepsilon^2 = 1$ might have been chosen instead of $\alpha_0 = 1$. This would have rendered the equations more easily intelligible. Notice also how the matrix following the first equality is folded across the axis which divides it vertically to give the matrix which follows the second equality. Pleasing effects of this sort often arise in time-series analysis.

The Partial Autocorrelation Function

Let $\alpha_{r(r)}$ be the coefficient associated with $y(t-r)$ in an autoregressive process of order r whose parameters correspond to the autocovariances $\gamma_0, \gamma_1, \ldots, \gamma_r$. Then the sequence $\{\alpha_{r(r)}; r = 1, 2, \ldots\}$ of such coefficients, whose index corresponds to models of increasing orders, constitutes the partial autocorrelation function. In effect, $\alpha_{r(r)}$ indicates the role in explaining the variance of $y(t)$ which is due to $y(t-r)$ when $y(t-1), \ldots, y(t-r+1)$ are also taken into account.

Much of the theoretical importance of the partial autocorrelation function is due to the fact that, when γ_0 is added, it represents an alternative way of conveying the information which is present in the sequence of autocorrelations. Its role in identifying the order of an autoregressive process is evident; for, if $\alpha_{r(r)} \neq 0$ and if $\alpha_{p(p)} = 0$ for all $p > r$, then it is clearly implied that the process has an order of r.

The sequence of partial autocorrelations may be computed efficiently via the recursive Durbin–Levinson Algorithm which uses the coefficients of the AR model of order r as the basis for calculating the coefficients of the model of order $r+1$.

To derive the algorithm, let us imagine that we already have the values $\alpha_{0(r)} = 1, \alpha_{1(r)}, \ldots, \alpha_{r(r)}$. Then, by extending the set of rth-order Yule–Walker equations to which these values correspond, we can derive the system

$$\begin{bmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_r & \gamma_{r+1} \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{r-1} & \gamma_r \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \gamma_r & \gamma_{r-1} & \cdots & \gamma_0 & \gamma_1 \\ \gamma_{r+1} & \gamma_r & \cdots & \gamma_1 & \gamma_0 \end{bmatrix} \begin{bmatrix} 1 \\ \alpha_{1(r)} \\ \vdots \\ \alpha_{r(r)} \\ 0 \end{bmatrix} = \begin{bmatrix} \sigma^2_{(r)} \\ 0 \\ \vdots \\ 0 \\ g \end{bmatrix}, \tag{29}$$

wherein

$$g = \sum_{j=0}^{r} \alpha_{j(r)} \gamma_{r+1-j} \qquad \text{with} \quad \alpha_{0(r)} = 1. \tag{30}$$

The system can also be written as

$$\begin{bmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_r & \gamma_{r+1} \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{r-1} & \gamma_r \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \gamma_r & \gamma_{r-1} & \cdots & \gamma_0 & \gamma_1 \\ \gamma_{r+1} & \gamma_r & \cdots & \gamma_1 & \gamma_0 \end{bmatrix} \begin{bmatrix} 0 \\ \alpha_{r(r)} \\ \vdots \\ \alpha_{1(r)} \\ 1 \end{bmatrix} = \begin{bmatrix} g \\ 0 \\ \vdots \\ 0 \\ \sigma^2_{(r)} \end{bmatrix}. \tag{31}$$

The two systems of equations (29) and (31) can be combined to give

$$\begin{bmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_r & \gamma_{r+1} \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{r-1} & \gamma_r \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \gamma_r & \gamma_{r-1} & \cdots & \gamma_0 & \gamma_1 \\ \gamma_{r+1} & \gamma_r & \cdots & \gamma_1 & \gamma_0 \end{bmatrix} \begin{bmatrix} 1 \\ \alpha_{1(r)} + c\alpha_{r(r)} \\ \vdots \\ \alpha_{r(r)} + c\alpha_{1(r)} \\ c \end{bmatrix} = \begin{bmatrix} \sigma^2_{(r)} + cg \\ 0 \\ \vdots \\ 0 \\ g + c\sigma^2_{(r)} \end{bmatrix}. \tag{32}$$

If we take the coefficient of the combination to be

$$c = -\frac{g}{\sigma^2_{(r)}}, \tag{33}$$

then the final element in the vector on the RHS becomes zero and the system becomes the set of Yule–Walker equations of order $r + 1$. The solution of the equations, from the last element $\alpha_{r+1(r+1)} = c$ through to the variance term $\sigma^2_{(r+1)}$ is given by

$$\alpha_{r+1(r+1)} = \frac{1}{\sigma^2_{(r)}} \left\{ \sum_{j=0}^{r} \alpha_{j(r)} \gamma_{r+1-j} \right\}$$

$$\begin{bmatrix} \alpha_{1(r+1)} \\ \vdots \\ \alpha_{r(r+1)} \end{bmatrix} = \begin{bmatrix} \alpha_{1(r)} \\ \vdots \\ \alpha_{r(r)} \end{bmatrix} + \alpha_{r+1(r+1)} \begin{bmatrix} \alpha_{r(r)} \\ \vdots \\ \alpha_{1(r)} \end{bmatrix} \tag{34}$$

$$\sigma^2_{(r+1)} = \sigma^2_{(r)} \{1 - (\alpha_{r+1(r+1)})^2\}.$$

Thus the solution of the Yule–Walker system of order $r + 1$ is easily derived from the solution of the system of order r, and there is scope for devising a recursive procedure. The starting values for the recursion are

$$\alpha_{1(1)} = -\gamma_1/\gamma_0 \qquad \text{and} \qquad \sigma^2_{(1)} = \gamma_0 \{1 - (\alpha_{1(1)})^2\}. \tag{35}$$

Autoregressive Moving Average Processes

The autoregressive moving-average process of orders p and q, which is referred to as the ARMA(p, q) process, is defined by the equation

$$\begin{aligned}\alpha_0 y(t) + \alpha_1 y(t-1) + \cdots + \alpha_p y(t-p) \\ = \mu_0 \varepsilon(t) + \mu_1 \varepsilon(t-1) + \cdots + \mu_q \varepsilon(t-q).\end{aligned} \tag{36}$$

The equation is normalised by setting $\alpha_0 = 1$ and by setting either $\mu_0 = 1$ or $\sigma_\varepsilon^2 = 1$. A more summary expression for the equation is $\alpha(L)y(t) = \mu(L)\varepsilon(t)$. Provided that the roots of the equation $\alpha(z) = 0$ lie outside the unit circle, the process can be represented by the equation $y(t) = \alpha^{-1}(L)\mu(L)\varepsilon(t)$ which corresponds to an infinite-order moving-average process. Conversely, provided the roots of the equation $\mu(z) = 0$ lie outside the unit circle, the process can be represented by the equation $\mu^{-1}(L)\alpha(L)y(t) = \varepsilon(t)$ which corresponds to an infinite-order autoregressive process.

By considering the moving-average form of the process, and by noting the form of the autocovariance generating function for such a process which is given by equation (16), it can be seen that the autocovariance generating function for the autoregressive moving-average process is

$$\gamma(z) = \sigma_\varepsilon^2 \frac{\mu(z)\mu(z^{-1})}{\alpha(z)\alpha(z^{-1})}. \tag{37}$$

This generating function, which is of some theoretical interest, does not provide a practical means of finding the autocovariances. To find these, let us consider multiplying the equation $\sum_i \alpha_i y_{t-i} = \sum_i \mu_i \varepsilon_{t-i}$ by $y_{t-\tau}$ and taking expectations. This gives

$$\sum_i \alpha_i \gamma_{\tau-i} = \sum_i \mu_i \delta_{i-\tau}, \tag{38}$$

where $\gamma_{\tau-i} = E(y_{t-\tau}y_{t-i})$ and $\delta_{i-\tau} = E(y_{t-\tau}\varepsilon_{t-i})$. Since ε_{t-i} is uncorrelated with $y_{t-\tau}$ whenever it is subsequent to the latter, it follows that $\delta_{i-\tau} = 0$ if $\tau > i$. Since the index i in the RHS of the equation (38) runs from 0 to q, it follows that

$$\sum_i \alpha_i \gamma_{i-\tau} = 0 \quad \text{if} \quad \tau > q. \tag{39}$$

Given the $q+1$ nonzero values $\delta_0, \delta_1, \ldots, \delta_q$, and p initial values $\gamma_0, \gamma_1, \ldots, \gamma_{p-1}$ for the autocovariances, the equations can be solved recursively to obtain the subsequent values $\{\gamma_p, \gamma_{p+1}, \ldots\}$.

To find the requisite values $\delta_0, \delta_1, \ldots, \delta_q$, consider multiplying the equation $\sum_i \alpha_i y_{t-i} = \sum_i \mu_i \varepsilon_{t-i}$ by $\varepsilon_{t-\tau}$ and taking expectations. This gives

$$\sum_i \alpha_i \delta_{\tau-i} = \mu_\tau \sigma_\varepsilon^2, \tag{40}$$

where $\delta_{\tau-i} = E(y_{t-i}\varepsilon_{t-\tau})$. The equation may be rewritten as

$$\delta_\tau = \frac{1}{\alpha_0}\Big(\mu_\tau \sigma_\varepsilon^2 - \sum_{i=1} \delta_{\tau-i}\Big), \tag{41}$$

and, by setting $\tau = 0, 1, \ldots, q$, we can generate recursively the required values $\delta_0, \delta_1, \ldots, \delta_q$.

Example. Consider the ARMA(2, 2) model which gives the equation

$$\alpha_0 y_t + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} = \mu_0 \varepsilon_t + \mu_1 \varepsilon_{t-1} + \mu_2 \varepsilon_{t-2}. \tag{42}$$

Multiplying by y_t, y_{t-1} and y_{t-2} and taking expectations gives

$$\begin{bmatrix} \gamma_0 & \gamma_1 & \gamma_2 \\ \gamma_1 & \gamma_0 & \gamma_1 \\ \gamma_2 & \gamma_1 & \gamma_0 \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} \delta_0 & \delta_1 & \delta_2 \\ 0 & \delta_0 & \delta_1 \\ 0 & 0 & \delta_0 \end{bmatrix} \begin{bmatrix} \mu_0 \\ \mu_1 \\ \mu_2 \end{bmatrix}. \tag{43}$$

Multiplying by ε_t, ε_{t-1} and ε_{t-2} and taking expectations gives

$$\begin{bmatrix} \delta_0 & 0 & 0 \\ \delta_1 & \delta_0 & 0 \\ \delta_2 & \delta_1 & \delta_0 \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} \sigma_\varepsilon^2 & 0 & 0 \\ 0 & \sigma_\varepsilon^2 & 0 \\ 0 & 0 & \sigma_\varepsilon^2 \end{bmatrix} \begin{bmatrix} \mu_0 \\ \mu_1 \\ \mu_2 \end{bmatrix}. \tag{44}$$

When the latter equations are written as

$$\begin{bmatrix} \alpha_0 & 0 & 0 \\ \alpha_1 & \alpha_0 & 0 \\ \alpha_2 & \alpha_1 & \alpha_0 \end{bmatrix} \begin{bmatrix} \delta_0 \\ \delta_1 \\ \delta_2 \end{bmatrix} = \sigma_\varepsilon^2 \begin{bmatrix} \mu_0 \\ \mu_1 \\ \mu_2 \end{bmatrix}, \tag{45}$$

they can be solved recursively for δ_0, δ_1 and δ_2 on the assumption that that the values of α_0, α_1, α_2 and σ_ε^2 are known. Notice that, when we adopt the normalisation $\alpha_0 = \mu_0 = 1$, we get $\delta_0 = \sigma_\varepsilon^2$. When the equations (43) are rewritten as

$$\begin{bmatrix} \alpha_0 & \alpha_1 & \alpha_2 \\ \alpha_1 & \alpha_0 + \alpha_2 & 0 \\ \alpha_2 & \alpha_1 & \alpha_0 \end{bmatrix} \begin{bmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \end{bmatrix} = \begin{bmatrix} \mu_0 & \mu_1 & \mu_2 \\ \mu_1 & \mu_2 & 0 \\ \mu_2 & 0 & 0 \end{bmatrix} \begin{bmatrix} \delta_0 \\ \delta_1 \\ \delta_2 \end{bmatrix}, \tag{46}$$

they can be solved for γ_0, γ_1 and γ_2. Thus the starting values are obtained which enable the equation

$$\alpha_0 \gamma_\tau + \alpha_1 \gamma_{\tau-1} + \alpha_2 \gamma_{\tau-2} = 0; \quad \tau > 2 \tag{47}$$

to be solved recursively to generate the succeeding values $\{\gamma_3, \gamma_4, \ldots\}$ of the autocovariances.