Moving-Average Processes

The MA($q$) process, is defined by

$$\begin{align*}
y(t) &= \mu_0 \varepsilon(t) + \mu_1 \varepsilon(t-1) + \cdots + \mu_q \varepsilon(t-q) \\
&= \mu(L) \varepsilon(t),
\end{align*}$$

where $\mu(L) = \mu_0 + \mu_1 L + \cdots + \mu_q L^q$ and where $\varepsilon(t)$ is white noise.

An MA model should be invertible such that $\mu^{-1}(L)y(t) = \varepsilon(t)$. This AR($\infty$) representation is available if and only if all the roots of $\mu(z) = 0$ lie outside the unit circle.

**Example.** Consider the MA(1) process

$$y(t) = \varepsilon(t) - \theta \varepsilon(t-1) = (1 - \theta L) \varepsilon(t).$$

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

$$\varepsilon(t) = (1 - \theta L)^{-1} y(t)$$

$$= \left\{ y(t) + \theta y(t-1) + \theta^2 y(t-2) + \cdots \right\}.$$

Imagine that $|\theta| > 1$ instead. Then we have to write

$$y(t + 1) = \varepsilon(t + 1) - \theta \varepsilon(t)$$

$$= -\theta(1 - L^{-1}/\theta) \varepsilon(t),$$

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

$$\varepsilon(t) = -\theta^{-1}(1 - L^{-1}/\theta)^{-1} y(t + 1)$$

$$= -\theta^{-1}\left\{ y(t + 1)/\theta + y(t + 2)/\theta^2 + y(t - 3)/\theta^3 + \cdots \right\}.$$

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

Consider

\[ \gamma_\tau = E(y_t y_{t-\tau}) = E\left\{ \sum_i \mu_i \varepsilon_{t-i} \sum_j \mu_j \varepsilon_{t-\tau-j} \right\} = \sum_i \sum_j \mu_i \mu_j E(\varepsilon_{t-i} \varepsilon_{t-\tau-j}). \]

(5.8)

Since \( \varepsilon(t) \) is white noise, it follows that

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

(5.9)

Therefore

\[ \gamma_\tau = \sigma_\varepsilon^2 \sum_j \mu_j \mu_{j+\tau}. \]

(5.10)

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

\[ \gamma_0 = \sigma_\varepsilon^2 (\mu_0^2 + \mu_1^2 + \cdots + \mu_q^2), \]

\[ \gamma_1 = \sigma_\varepsilon^2 (\mu_0 \mu_1 + \mu_1 \mu_2 + \cdots + \mu_{q-1} \mu_q), \]

\vdots

\[ \gamma_q = \sigma_\varepsilon^2 \mu_0 \mu_q. \]

(5.11)

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

**Example.** The MA(1) process \( y(t) = \varepsilon(t) - \theta \varepsilon(t-1) \) has

\[ \gamma_0 = \sigma_\varepsilon^2 (1 + \theta^2), \]

\[ \gamma_1 = -\sigma_\varepsilon^2 \theta, \]

\[ \gamma_\tau = 0 \quad \text{if} \quad \tau > 1. \]

(5.12)

Thus the dispersion matrix of \( y = [y_1, y_2, \ldots, y_T]' \) is

\[ D(y) = \sigma_\varepsilon^2 \begin{bmatrix} 1 + \theta^2 & -\theta & 0 & \cdots & 0 \\ -\theta & 1 + \theta^2 & -\theta & \cdots & 0 \\ 0 & -\theta & 1 + \theta^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 + \theta^2 \end{bmatrix}. \]

(5.13)
Autocovariance Generating Function

This is denoted by

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

To find the autocovariance generating function of the MA(\(q\)) process, consider

\[
\begin{align*}
\mu(z)\mu(z^{-1}) &= \sum_{i} \mu_{i}z^{i} \sum_{j} \mu_{j}z^{-j} \\
&= \sum_{i} \sum_{j} \mu_{i}\mu_{j}z^{i-j} \\
&= \sum_{\tau} \left( \sum_{j} \mu_{i}\mu_{j+\tau} \right) z^{\tau}, \quad \tau = i-j.
\end{align*}
\]

From (10) it follows that

\[ \gamma(z) = \sigma^2 \mu(z)\mu(z^{-1}). \]
The AR(p) process, is defined by
\[ \alpha_0 y(t) + \alpha_1 y(t - 1) + \cdots + \alpha_p y(t - p) = \varepsilon(t). \]
This can be written 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 \( \alpha(z) = 0 \) must lie outside the unit circle. In that case the AR process can be written as an MA(\( \infty \)) process:
\[ y(t) = \alpha^{-1}(L)\varepsilon(t). \]

The autovariance generating function for the AR(p) process is
\[ (5.23) \quad \gamma(z) = \frac{\sigma^2_\varepsilon}{\alpha(z)\alpha(z^{-1})}. \]

**Example.** Consider the AR(1) process defined by
\[ (5.18) \quad \varepsilon(t) = y(t) - \phi y(t - 1) = (1 - \phi L)y(t). \]
Provided that \(|\phi| < 1\), this can be represented in MA form as
\[ (5.19) \quad y(t) = (1 - \phi L)^{-1} \varepsilon(t) = \{\varepsilon(t) + \phi\varepsilon(t - 1) + \phi^2\varepsilon(t - 2) + \cdots\}. \]
The autocovariances of the AR(1) process can be obtained via the formula (10) for the autocovariances of an MA process. Thus
\[ (5.20) \quad \gamma_\tau = E(y_n y_{n-\tau}) = E\left\{ \sum_i \phi^i \varepsilon_{t-i} \sum_j \phi^j \varepsilon_{t-\tau-j} \right\} = \sum_i \sum_j \phi^i \phi^j E(\varepsilon_{t-i}\varepsilon_{t-\tau-j}); \]
and it follows from (9) that
\[ (5.21) \quad \gamma_\tau = \sigma^2_\varepsilon \sum_j \phi^j \phi^{j+\tau} \]
\[ = \frac{\sigma^2_\varepsilon \phi^{\tau}}{1 - \phi^2}. \]
The dispersion matrix of \( y = [y_1, y_2, \ldots, y_T]' \) is
\[ (5.22) \quad D(y) = \frac{\sigma^2_\varepsilon}{1 - \phi^2} \begin{bmatrix} 1 & \phi & \phi^2 & \cdots & \phi^{T-1} \\ \phi & 1 & \phi & \cdots & \phi^{T-2} \\ \phi^2 & \phi & 1 & \cdots & \phi^{T-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \phi^{T-1} & \phi^{T-2} & \phi^{T-3} & \cdots & 1 \end{bmatrix}. \]
The Yule-Walker Equations

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

\[
\sum_i \alpha_i E(y_{t-i}y_{t-\tau}) = E(\varepsilon_t y_{t-\tau}).
\]  

Given that \( \alpha_0 = 1 \), it follows that

\[
E(\varepsilon_t y_{t-\tau}) = \begin{cases} 
\sigma^2, & \text{if } \tau = 0; \\
0, & \text{if } \tau > 0.
\end{cases}
\]

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

\[
\sum_i \alpha_i \gamma_{\tau-i} = \begin{cases} 
\sigma^2, & \text{if } \tau = 0; \\
0, & \text{if } \tau > 0.
\end{cases}
\]

The second of these is a homogeneous difference equation which enables us to generate the sequence \( \gamma_0, \gamma_1, \ldots, \gamma_p \) 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:

\[
\begin{bmatrix}
\gamma_0 & \gamma_1 & \gamma_2 & \cdots & \gamma_p \\
\gamma_1 & \gamma_0 & \gamma_1 & \cdots & \gamma_{p-1} \\
\gamma_2 & \gamma_1 & \gamma_0 & \cdots & \gamma_{p-2} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\gamma_p & \gamma_{p-1} & \gamma_{p-2} & \cdots & \gamma_0
\end{bmatrix}
\begin{bmatrix}
1 \\
\alpha_1 \\
\alpha_2 \\
\vdots \\
\alpha_p
\end{bmatrix}
= \begin{bmatrix}
\sigma^2 \\
0 \\
0 \\
\vdots \\
0
\end{bmatrix}.
\]

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 \( \alpha_1, \ldots, \alpha_p, \sigma^2 \) or vice versa.

**Example.** Consider the second-order autoregressive process. We have

\[
\begin{bmatrix}
\gamma_0 & \gamma_1 & \gamma_2 \\
\gamma_1 & \gamma_0 & \gamma_1 \\
\gamma_2 & \gamma_1 & \gamma_0
\end{bmatrix}
\begin{bmatrix}
\alpha_0 \\
\alpha_1 \\
\alpha_2
\end{bmatrix}
= \begin{bmatrix}
\alpha_2 & \alpha_1 & \alpha_0 \\
0 & \alpha_2 & \alpha_1 \\
0 & 0 & \alpha_2
\end{bmatrix}
\begin{bmatrix}
\gamma_0 \\
\gamma_1 \\
\gamma_2
\end{bmatrix}
= \begin{bmatrix}
\sigma^2 \\
0 \\
0
\end{bmatrix}.
\]

Given \( \alpha_0 = 1 \) and the values for \( \gamma_0, \gamma_1, \gamma_2 \), we can find \( \sigma^2 \) and \( \alpha_1, \alpha_2 \). Conversely, given \( \alpha_0, \alpha_1, \alpha_2 \) and \( \sigma^2 \), we can find \( \gamma_0, \gamma_1, \gamma_2 \).
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.

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 $r + 1$.

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

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

wherein

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

The system can also be written as

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

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

$$\begin{bmatrix}
\gamma_0 & \gamma_1 & \cdots & \gamma_r & \gamma_{r+1} \\
\gamma_1 & \gamma_0 & \cdots & \gamma_{r-1} & \gamma_r \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
\gamma_r & \gamma_{r-1} & \cdots & \gamma_0 & \gamma_1 \\
\gamma_{r+1} & \gamma_r & \cdots & \gamma_1 & \gamma_0
\end{bmatrix}
\begin{bmatrix}
1 \\
\alpha_{1(r)} + c\alpha_{r(r)} \\
\vdots \\
\alpha_{r(r)} + c\alpha_{1(r)} \\
\sigma^2_{(r)} + cg
\end{bmatrix}
= 
\begin{bmatrix}
\sigma^2_{(r)} + cg \\
0 \\
\vdots \\
0 \\
g + c\sigma^2_{(r)}
\end{bmatrix}.$$
LECTURE 4 : ARMA PROCESSES

If we take the coefficient of the combination to be

\[(5.33) \quad c = -\frac{g}{\sigma^2_{(r)}}.\]

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

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

\[(5.34) \quad \begin{bmatrix} \alpha_{1(r+1)} \\ \vdots \\ \alpha_{r(r+1)} \end{bmatrix} = \begin{bmatrix} \alpha_{1(r)} \\ \vdots \\ \alpha_{r(r)} \end{bmatrix} + \alpha_{r+1(r+1)} \begin{bmatrix} \alpha_{r(r)} \\ \vdots \\ \alpha_{1(r)} \end{bmatrix}
\]

\[
\sigma^2_{(r+1)} = \sigma^2_{(r)} \left\{ 1 - (\alpha_{r+1(r+1)})^2 \right\}.
\]

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

\[(5.35) \quad \alpha_{1(1)} = -\gamma_1/\gamma_0 \quad \text{and} \quad \sigma^2_{(1)} = \gamma_0 \{1 - (\alpha_{1(1)})^2\}.\]
**Autoregressive Moving Average Processes**

The ARMA\((p,q)\) process, is defined by

\[
\alpha_0 y(t) + \alpha_1 y(t - 1) + \cdots + \alpha_p y(t - p) = \mu_0 \varepsilon(t) + \mu_1 \varepsilon(t - 1) + \cdots + \mu_q \varepsilon(t - q).
\]

This can also be written as \(\alpha(L)y(t) = \mu(L)\varepsilon(t)\). If the roots of \(\alpha(z) = 0\) lie outside the unit circle, then the process has an MA\((1)\) form: \(y(t) = \mu^{-1}(L)\alpha(L)y(t) = \varepsilon(t)\).

The autocovariance generating function for the ARMA process is

\[
\gamma(z) = \sigma_\varepsilon^2 \frac{\mu(z)\mu(z^{-1})}{\alpha(z)\alpha(z^{-1})}.
\]

To find the autocovariances in practice, 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

\[
\sum_i \alpha_i \gamma_{t-i} = \sum_i \mu_i \delta_{t-i},
\]

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

\[
\sum_i \alpha_i \gamma_{i-\tau} = 0 \quad \text{if} \quad \tau > q.
\]

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

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

\[
\sum_i \alpha_i \delta_{t-i} = \mu_\tau \sigma_\varepsilon^2,
\]

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

\[
\delta_\tau = \frac{1}{\alpha_0} \left( \mu_\tau \sigma_\varepsilon^2 - \sum_{i=1}^{\tau} \delta_{t-i} \right),
\]

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

\( \alpha_0 y_t + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} = \mu_0 \epsilon_t + \mu_1 \epsilon_{t-1} + \mu_2 \epsilon_{t-2}. \)

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

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

Multiplying by \( \epsilon_t, \epsilon_{t-1} \) and \( \epsilon_{t-2} \) and taking expectations gives

\[
\begin{bmatrix}
\delta_0 & 0 & 0 \\
\delta_1 & 0 & 0 \\
\delta_2 & \delta_0 & \delta_0
\end{bmatrix}
\begin{bmatrix}
\alpha_0 \\
\alpha_1 \\
\alpha_2
\end{bmatrix}
= \begin{bmatrix}
\sigma_\epsilon^2 & 0 & 0 \\
0 & \sigma_\epsilon^2 & 0 \\
0 & 0 & \sigma_\epsilon^2
\end{bmatrix}
\begin{bmatrix}
\mu_0 \\
\mu_1 \\
\mu_2
\end{bmatrix}.
\]

When the latter equations are written as

\[
\begin{bmatrix}
\alpha_0 & 0 & 0 \\
\alpha_1 & \alpha_0 & 0 \\
\alpha_2 & \alpha_1 & \alpha_0
\end{bmatrix}
\begin{bmatrix}
\delta_0 \\
\delta_1 \\
\delta_2
\end{bmatrix}
= \sigma_\epsilon^2
\begin{bmatrix}
\mu_0 \\
\mu_1 \\
\mu_2
\end{bmatrix},
\]

they can be solved recursively for \( \delta_0, \delta_1 \) and \( \delta_2 \) on the assumption that that the values of \( \alpha_0, \alpha_1, \alpha_2 \) and \( \sigma_\epsilon^2 \) are known. Notice that, when we adopt the normalisation \( \alpha_0 = \mu_0 = 1 \), we get \( \delta_0 = \sigma_\epsilon^2 \). When the equations (43) are rewritten as

\[
\begin{bmatrix}
\alpha_0 & \alpha_1 & \alpha_2 \\
\alpha_1 & \alpha_0 & 0 \\
\alpha_2 & \alpha_1 & \alpha_0
\end{bmatrix}
\begin{bmatrix}
\gamma_0 \\
\gamma_1 \\
\gamma_2
\end{bmatrix}
= \begin{bmatrix}
\mu_0 & \mu_1 & \mu_2 \\
\mu_1 & \mu_2 & 0 \\
\mu_2 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
\delta_0 \\
\delta_1 \\
\delta_2
\end{bmatrix},
\]

they can be solved for \( \gamma_0, \gamma_1 \) and \( \gamma_2 \). Thus the starting values are obtained which enable the equation

\( \alpha_0 \gamma_\tau + \alpha_1 \gamma_{\tau-1} + \alpha_2 \gamma_{\tau-2} = 0; \quad \tau > 2 \)

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