Notes on ARMA Models

James L. Powell
Department of Economics
University of California, Berkeley

ARMA Processes

Autoregressive and moving average processes can be combined to obtain a very flexible class of univariate processes (proposed by Box and Jenkins), known as *ARMA* processes.

ARMA(p,q) Process: The time series y_t is an *ARMA(p,q)* process, written $y_t \sim ARMA(p,q)$, if

$$\begin{aligned} y_t &= \alpha + \beta_1 y_{t-1} + ... + \beta_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + ... + \theta_q \varepsilon_{t-q} \\ &= \alpha + \beta(L) y_{t-1} + \theta(L)\varepsilon_t, \end{aligned}$$

or

$$\phi(L) y_t = \alpha + \theta(L)\varepsilon_t$$

where $\varepsilon_t \sim WN(\sigma^2)$ and $Cov(\varepsilon_t, y_{t-s}) = 0$ if $s \geq 1$.

The requirements for stationarity of this process are the same as for stationarity of the corresponding $AR(p)$ process; that is, any roots of $\phi(z) = 0$ are outside the unit circle ($|z| > 1$), in which case

$$\begin{aligned} \mu &= E[y_t] \\ &= \frac{\alpha}{\phi(1)}, \\ y_t &= \mu + \psi(L)\varepsilon_t, \\ \psi(L) &\equiv \frac{\theta(L)}{\phi(L)}. \end{aligned}$$

Similarly, the moving average polynomial $\theta(L)$ is invertible if $\theta(z) = 0$ implies $|z| > 1$, so that

$$\begin{aligned} \tau(L) y_t &= \varepsilon_t, \\ \tau(L) &\equiv \frac{\phi(L)}{\theta(L)} \\ &= \frac{1}{\psi(L)}. \end{aligned}$$

Common Factors and Identification

In a sense, ARMA processes are *too* flexible, in the sense that low-order processes (i.e., those with p and q small) are nested in higher-order processes with certain parameter restrictions. In general, if $y_t \sim ARMA(p,q)$, then it can always be rewritten as an $ARMA(p+r, q+r)$ process for arbitrary positive integer r by suitable "generalized differencing". For example, suppose $y_t = \varepsilon_t \sim WN\left(\sigma^2\right)$, so that $y_t \sim ARMA(0,0)$. Then for any ρ with $|\rho| < 1$,

$$y_t - \rho y_{t-1} = \varepsilon_t - \rho\varepsilon_{t-1},$$

or

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

so $y_t \sim ARMA(1,1)$ with a particular restriction on the parameters (i.e., the sum of the first-order autoregressive and moving average coefficients is zero). For this example this redundancy is easy to find, but for more complicated $ARMA$ processes the restrictions on the parameters may be difficult to find in the population, and even harder to detect in estimation.

In general, writing a mean-zero $ARMA(p,q)$ process as

$$\phi(L)y_t = \theta(L)\varepsilon_t, \qquad \varepsilon_t \sim WN(\sigma^2),$$

the process can be rewritten as a lower-order process if the lag polynomials $\phi(L)$ and $\theta(L)$ have common roots (more precisely, if their associated polynomial equations $\tilde{\phi}(z) = 0$ and $\tilde{\theta}(z) = 0$ have common roots). If there are r roots in common, then the polynomials can be factored as

$$\begin{aligned} \phi(L) &= \alpha(L)\phi^*(L), \\ \theta(L) &= \alpha(L)\theta^*(L), \end{aligned}$$

where ϕ^* and θ^* are polynomials of order $p-r$ and $q-r$, respectively, so that $y_t \sim ARMA(p-r, q-r)$,

$$\phi^*(L)y_t = \theta^*(L)\varepsilon_t, \qquad \varepsilon_t \sim WN(\sigma^2).$$

In practice, with estimates $\hat{\phi}(L)$ and $\hat{\theta}(L)$ of the lag polynomials $\phi(L)$ and $\theta(L)$, determining if there are any common roots (and their number) can be problematic. Box and Jenkins' proposed solution to this *common factors* problem, which they called their "principle of parsimony", is simple enough – just pick p and q to be small enough to do the job (of forecasting, etc.). To implement this general idea, however, they proposed a methodology for model selection which they termed *time series identification* procedures. In econometric applications, the tradition has been to consider only *purely autoregressive* processes, i.e., assume that $y_t \sim AR(p)$ for some value of p (chosen in practice by a suitable model selection procedure). Purely autoregressive processes, while typically requiring a higher number of parameters to approximate complicated dynamic patterns, do not suffer from the common factor problem, since a redundant generalized difference in the autoregressive component is accompanied by an error term which is not white noise (i.e., $q = r > 0$). Furthermore, as will be discussed later, purely autoregressive processes are simpler to estimate, requiring only linear (not nonlinear) LS estimation.

Sums of ARMA Processes

Given the underlying linear structure of $ARMA$ processes, it is probably not surprising that sums of uncorrelated $ARMA$ processes are themselves $ARMA$ processes, though the orders of the combined process might not be obvious *a priori*. Suppose x_t is $ARMA(p_x, q_x)$ and z_t is $ARMA(p_z, q_z)$, and uncorrelated at all leads and lags:

$$\begin{aligned} \phi(L)x_t &= \theta(L)\varepsilon_t, \qquad \varepsilon_t \sim WN(\sigma^2), \\ \alpha(L)z_t &= \delta(L)\eta_t, \qquad \eta_t \sim WN(\nu^2), \end{aligned}$$

with

$$Cov(\varepsilon_t, \eta_s) = 0, \text{ all } t, s.$$

Writing

$$\begin{aligned} x_t &= \frac{\theta(L)}{\phi(L)}\varepsilon_t, \\ z_t &= \frac{\delta(L)}{\alpha(L)}\eta_t, \end{aligned}$$

the sum y_t of these two processes is

$$\begin{aligned} y_t &= x_t + z_t \\ &= \frac{\theta(L)}{\phi(L)}\varepsilon_t + \frac{\delta(L)}{\alpha(L)}\eta_t, \end{aligned}$$

and premultiplying y_t by the product of the two AR polynomials yields

$$\begin{aligned} \phi^*(L)y_t &\equiv \phi(L)\alpha(L)y_t \\ &= \alpha(L)\theta(L)\varepsilon_t + \phi(L)\delta(L)\eta_t \\ &\equiv u_t. \end{aligned}$$

The combined error term u_t is a moving-average process, with order no larger than the larger of $p_z + q_x$ (the order of the polynomial in front of ε_t) and $p_x + q_z$ (the order of the polynomial multiplying ε_t. Furthermore, if the AR polynomials ϕ and α have r common roots, then an r^{th}-order polynomial can be factored out of both sides of the equation for $y_t = x_t + z_t$, which is thus an $ARMA(p_y, q_y)$ process, with

$$
\begin{aligned}
p_y &= p_x + p_z - r, \\
q_y &= \max\{p_z + q_x - r, p_x + q_z - r\}.
\end{aligned}
$$

So, just as common roots in the AR and MA polynomials imply a lower order of an $ARMA$ process, common roots in the AR polynomials of two uncorrelated $ARMA$ processes yield lower orders for the $ARMA$ process describing their sum.

Other topics:

Markov Chains and AR Processes

Method of Undetermined Coefficients

Autocovariance Generating Function