Financial Econometrics
Lecture 1: Autoregressive and Moving Average Processes

Richard G. Pierse

1 Introduction

Autoregressive and moving average processes are two alternative models for stationary time series variables that are widely used in financial econometrics. Both imply some form of persistence so that shocks to a variable continue to have an effect after the period in which they occurred although stationarity ensures that this effect eventually dies away. These two alternative processes also form the basis for the *Box and Jenkins ARIMA* model that is the topic of next week's lecture.

2 Stationarity

The concept of stationarity of a random variable is very important in the analysis of financial time series. We make use of two definitions: *weak stationarity* and *strict stationarity*.

2.1 Weak stationarity

A variable y_t is said to be *weakly stationary* if both its *mean* and *variance* are constant over time and its *autocovariances* γ_s defined by

$$\gamma_s = \mathrm{cov}(y_t y_{t-s}) = \mathrm{E}(y_t - \mathrm{E}(y_t))(y_{t-s} - \mathrm{E}(y_{t-s}))$$

depend only on the distance between observations s and not on t. Weak stationarity implies time independence of the first two moments of the distribution only.

2.2 Strict stationarity

A variable is said to be *strictly stationary* if its joint distribution does not depend on t.

$$f(y_t, y_{t-1}, y_{t-2}, \cdots, y_{t-s}) = f(y_{t'}, y_{t'-1}, y_{t'-2}, \cdots, y_{t'-s}), \qquad \forall t, t', s.$$

Strict stationarity demands time independence of the complete distribution of y_t and is thus stronger than weak stationarity. However, it is possible for a variable to be strictly stationary without being weakly stationary if its first two moments do not exist.

3 Autoregressive processes

3.1 The first order autoregressive process

Let y_t be defined by the process

$$y_t = \phi y_{t-1} + \varepsilon_t \tag{3.1}$$

where ε_t is an independently identically distributed random variable of innovations with

$$\varepsilon_t \sim iid(0, \sigma^2).$$

For the process to be stationary we require that the parameter ϕ satisfies the restriction that

$$|\phi| < 1.$$

This process is known as a first order autoregressive or *AR(1)* process. By repeated substitution we can write y_t in terms of the innovations ε_t

$$y_t = \varepsilon_t + \phi\varepsilon_{t-1} + \phi^2\varepsilon_{t-2} + \phi^3\varepsilon_{t-3} + \cdots$$

The variable y_t has the properties that

$$\mathrm{E}(y_t) = \mathrm{E}(\varepsilon_t) + \phi\,\mathrm{E}(\varepsilon_{t-1}) + \phi^2\,\mathrm{E}(\varepsilon_{t-2}) + \phi^3\,\mathrm{E}(\varepsilon_{t-3}) + \cdots = 0$$

$$\begin{aligned} \mathrm{var}(y_t) &= \mathrm{E}(\varepsilon_t^2) + \phi^2\,\mathrm{E}(\varepsilon_{t-1}^2) + \phi^4\,\mathrm{E}(\varepsilon_{t-2}^2) + \phi^6\,\mathrm{E}(\varepsilon_{t-3}^2) + \cdots \\ &= (1 + \phi^2 + \phi^4 + \phi^6 + \cdots)\sigma^2 \\ &= \sigma^2/(1-\phi^2) \end{aligned}$$

and

$$\begin{aligned} \mathrm{cov}(y_t y_{t-s}) &= \phi^s\,\mathrm{var}(y_{t-s}) \\ &= \phi^s\sigma^2/(1-\phi^2). \end{aligned}$$

It can be seen that y_t has a constant mean of zero, a constant variance and autocovariances that depend only on the distance s between observations. Thus y_t is indeed a *weakly stationary* variable.

The autocorrelation function (ACF) of any stationary process y_t is defined by

$$\rho_s = \frac{\text{cov}(y_t, y_{t-s})}{\text{var}(y_t)} \,. \tag{3.2}$$

For the *AR(1)* process

$$\rho_s = \phi^s$$

and the ACF coefficients die away (in absolute value) since $|\phi^s| < |\phi^{s-1}|$.

The partial autocorrelation function ($PACF$) of y_t is defined by

$$p_s = \frac{\text{cov}(y_t, y_{t-s}|y_{t-1}, y_{t-2}, \cdots, y_{t-s-1})}{\text{var}(y_t)}. \tag{3.3}$$

It measures the autocorrelation between y_t and y_{t-s} taking into account the effect of all intermediate lags y_{t-r}, $r < s$. It is based on conditional variances and covariances and is identical to the coefficient p_s in the equation

$$y_t = p_1 y_{t-1} + p_2 y_{t-2} + \cdots + p_s y_{t-s} + \varepsilon_t.$$

In the *AR(1)* case, clearly $p_1 = \phi$ and $p_s = 0$, $s > 1$. It can be seen that the $PACF$ coefficients cut off abruptly after $s = 1$.

3.2 Higher order processes

We can define the *AR(p)* process

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t \tag{3.4}$$

where as before

$$\varepsilon_t \sim iid(0, \sigma^2).$$

Introducing the *lag operator*, L, defined by

$$L^k x_t = x_{t-k} \quad , \quad L^0 x_t = 1$$

we can rewrite the model as

$$y_t - \phi_1 L y_t - \phi_2 L^2 y_t - \cdots - \phi_p L^p y_t = \varepsilon_t$$

or

$$(1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p) y_t = \phi(L) y_t = \varepsilon_t$$

where $\phi(L)$ is a polynomial function in the lag operator. This polynomial can be factorised as the product of its roots

$$\phi(L) = \prod_{j=1}^{p}(1 - \alpha_j L) = (1 - \alpha_1 L)(1 - \alpha_2 L)\cdots(1 - \alpha_p L) \tag{3.5}$$

where the roots are

$$\frac{1}{\alpha_1}, \frac{1}{\alpha_2}, \cdots, \frac{1}{\alpha_p}.$$

These roots can either be *real* numbers or *complex* numbers of the form $a + bi$ where a and b are real numbers and where i is the imaginary number defined by $i = \sqrt{-1}$. Where roots are complex, they must appear in *complex conjugate pairs* of the form

$$\alpha_j = a_j + b_j i \qquad \text{and} \qquad \alpha_{j+1} = a_j - b_j i$$

so that their sum

$$\alpha_j + \alpha_{j+1} = (a_j + b_j i) + (a_j - b_j i) = 2a_j$$

and their product

$$\alpha_j \alpha_{j+1} = (a_j + b_j i)(a_j - b_j i) = a_j^2 - b_j^2 i^2 = a_j^2 + b_j^2$$

are both *real* numbers. Real roots correspond to damped exponentials whereas complex conjugate pairs of roots correspond to damped sine waves (cycles).

For the *AR(p)* process to be stationary we require that

$$\|\alpha_j\| < 1, \forall j$$

where $\|\alpha_j\|$ is the *norm* of α_j defined for the *real* case as $\|\alpha_j\| = |\alpha_j|$ and for the *complex* case as

$$\|\alpha_j\| = \sqrt{(a_j + b_j i)(a_j - b_j i)} = \sqrt{a_j^2 + b_j^2}.$$

The condition for stationarity is sometimes stated as the condition that all the roots of the *AR(p)* process lie *outside* the unit circle.

The *AR(p)* model can be written in terms of the innovations ε_t as

$$y_t = \phi(L)^{-1}\varepsilon_t$$

where

$$\phi(L)^{-1} = \prod_{j=1}^{p}(1 - \alpha_j L)^{-1}$$

is a polynomial in the lag operator. This polynomial will only exist when the *AR(p)* process satisfies the conditions for stationarity. In general it will be of infinite order so that y_t depends on the whole past history of ε_t.

The *AR(p)* process has the properties that

$$\mathrm{E}(y_t) = \mathrm{E}(\phi(L)^{-1}\varepsilon_t) = 0,$$

$$\begin{aligned} \mathrm{var}(y_t) &\equiv \gamma_0 = E[y_t(\phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t)] \\ &= \phi_1\gamma_1 + \phi_2\gamma_2 + \cdots + \phi_p\gamma_p + \sigma^2 \end{aligned}$$

$$\begin{aligned} \mathrm{cov}(y_t y_{t-s}) &\equiv \gamma_s = E[y_{t-s}(\phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t)] \\ &= \phi_1\gamma_{s-1} + \phi_2\gamma_{s-2} + \cdots + \phi_p\gamma_{s-p}. \end{aligned}$$

The coefficients of the *ACF*, ρ_s the autocovariances, will die out as s increases but will never disappear completely. The *PACF* partial autocovariances p_s are non-zero up to p_p but then cut off completely.

4 Moving Average processes

Let y_t be defined by the process

$$y_t = \varepsilon_t + \theta\varepsilon_{t-1} \tag{4.1}$$

where ε_t is an independently identically distributed random variable with

$$\varepsilon_t \sim iid(0, \sigma^2).$$

This process is known as a first order moving average or *MA(1)* process.

The variable y_t has the properties that

$$\mathrm{E}(y_t) = \mathrm{E}(\varepsilon_t) + \theta\,\mathrm{E}(\varepsilon_{t-1}) = 0,$$

$$\begin{aligned} \mathrm{var}(y_t) &= \mathrm{E}(\varepsilon_t + \theta\varepsilon_{t-1})^2 \\ &= \mathrm{E}(\varepsilon_t^2) + 2\theta\,\mathrm{E}(\varepsilon_t\varepsilon_{t-1}) + \theta^2\,\mathrm{E}(\varepsilon_{t-1}^2) \\ &= (1+\theta^2)\sigma^2, \end{aligned}$$

and

$$\begin{aligned} \mathrm{cov}(y_t y_{t-s}) &= \mathrm{E}(\varepsilon_t + \theta\varepsilon_{t-1})(\varepsilon_{t-s} + \theta\varepsilon_{t-s-1}) \\ &= \mathrm{E}(\varepsilon_t\varepsilon_{t-s}) + \theta\,\mathrm{E}(\varepsilon_t\varepsilon_{t-s-1}) + \theta\,\mathrm{E}(\varepsilon_{t-1}\varepsilon_{t-s}) + \theta^2\,\mathrm{E}(\varepsilon_{t-1}\varepsilon_{t-s-1}) \\ &= \theta\sigma^2 \quad , \quad s = 1 \qquad \text{and} \qquad = 0 \quad , \quad s > 1. \end{aligned}$$

It can be seen that y_t has a constant mean of zero, a constant variance and autocovariances that depend only on the distance s between observations. Thus y_t is a weakly stationary variable and this is true for all values of θ, including the unit moving average root cases $\theta = \pm 1$.

The autocovarances ρ_s cut off after the first order whereas it can be shown that the partial autocovariances p_s will die out slowly.

4.1 Identification

For the autocorrelation function of the *MA(1)* process we have $\rho_1 = \theta/(1+\theta^2)$ and $\rho_s = 0$ for all $s > 1$. Consider an alternative *MA(1)* process with *MA* coefficient $\theta^* = 1/\theta$. The autocorrelation function for this process has

$$\rho_1^* = \frac{\theta^*}{(1+\theta^{*2})} = \frac{1}{\theta}\frac{\theta^2}{(1+\theta^2)} = \frac{\theta}{(1+\theta^2)} = \rho_1$$

and $\rho_s^* = 0$ for all $s > 1$. Thus an *MA(1)* process with coefficient $1/\theta$ has exactly the same autocorrelation function as a process with coefficient θ and it is impossible to distinguish between the two processes from their autocorrelations. This is an *identification* problem and to resolve this we impose the identification restriction on the *MA(1)* model that

$$|\theta| \leq 1.$$

As before, note that the identification restriction does not exclude the unit root cases $\theta = \pm 1$.

4.2 Higher order processes

We can define the *MA(q)* process

$$y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q}$$

where as before

$$\varepsilon_t \sim iid(0, \sigma^2).$$

This process has the properties

$$\mathrm{E}(y_t) = \mathrm{E}(\varepsilon_t) + \theta_1 \mathrm{E}(\varepsilon_{t-1}) + \theta_2 \mathrm{E}(\varepsilon_{t-2}) + \cdots + \theta_q \mathrm{E}(\varepsilon_{t-q}) = 0,$$

$$\begin{aligned} \mathrm{var}(y_t) &= \mathrm{E}(\varepsilon_t^2) + \theta_1^2 \mathrm{E}(\varepsilon_{t-1}^2) + \theta_2^2 \mathrm{E}(\varepsilon_{t-2}^2) + \cdots + \theta_q^2 \mathrm{E}(\varepsilon_{t-q}^2) \\ &= (1 + \theta_1^2 + \theta_2^2 + \cdots + \theta_q^2)\sigma^2 \end{aligned}$$

and

$$
\begin{aligned}
\operatorname{cov}(y_t y_{t-s}) &= (\theta_s + \theta_1\theta_{s+1} + \cdots + \theta_{q-s}\theta_q)\sigma^2 \quad , \quad s \leq q \\
&= 0 \quad , \quad s > q.
\end{aligned}
$$

Autocovariances ρ_s are non-zero up to order q but cut off after this point. Partial autocovariances p_s will die away gradually.

References

[1] Box, G.E.P, G.M. Jenkins and G.C. Reinsel (1994), *Time Series Analysis: Forecasting and Control*, (3rd ed.), Prentice Hall, Englewood Cliffs, NJ.