ARIMA Models

Richard G. Pierse

1 Introduction

Time Series Analysis looks at the properties of time series from a purely statistical point of view. No attempt is made to relate variables using *a priori* economic theory (*c.f. econometrics*).

Why use time series methods?

1) **Short term forecasting**. Time series models may well forecast better than econometric models over the short term. One would expect that they would do less well in the longer term.

2) **Projecting exogenous variables**. even in an econometric model, exogenous variables need to be projected for forecasting. Time series models may be used to *complete* the model.

Time Series methods can be either univariate or multivariate. This lecture deals with the Box and Jenkins methodology of *ARIMA* models as a univariate analysis, although it can just as easily be used to jointly model several time series..

2 *AR(I)MA* Processes

Any *stationary* linear process x_t can be written

$$x_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} \tag{1}$$

where ε_t is a sequence of uncorrelated random variables with mean 0 and constant variance σ^2. This is called the *infinite moving average* representation of x_t, or the *Wold* representation. Note the condition that

$$\sum_{j=0}^{\infty} \psi_j^2 < \infty\,.$$

Introducing the *lag operator*, L, defined by

$$L^k w_t = w_{t-k}$$

equation (1) can be rewritten as

$$x_t = (1 + \psi_1 L + \psi_2 L^2 + \cdots)\varepsilon_t = \psi(L)\varepsilon_t$$

where without loss of generality we impose the normalisation restriction that $\psi_0 = 1$.

The polynomial $\psi(L)$ can be factorised as

$$\psi(L) = \prod_{j=0}^{\infty}(1 + \beta_j L) = (1 + \beta_1 L)(1 + \beta_2 L) \ \cdots = 0$$

with roots given by

$$-\frac{1}{\beta_1}, -\frac{1}{\beta_2}, \quad \text{etc.}$$

which, *for identifiability*, must satisfy the condition that

$$\|\beta_i\| \leq 1 \quad , \quad \forall i \, .$$

If *all* the roots satisfy the stronger condition that $\|\beta_i\| < 1$, then the process is said to be *invertible* and x_t can be written in the *autoregressive* representation

$$\psi(L)^{-1} x_t = \varepsilon_t \, .$$

More generally, if *some* of the roots satisfy $\|\beta_i\| < 1$, then it can be factorised into two polynomials

$$x_t = \psi(L)\varepsilon_t = \phi(L)^{-1}\theta(L)\varepsilon_t$$

or

$$\phi(L) x_t = \theta(L)\varepsilon_t \, . \tag{2}$$

(2) is a *mixed ARMA(p,q)* model where

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

and

$$\theta(L) = 1 + \theta_1 L + \theta_2 L^2 + \cdots + \theta_q L^q \, .$$

Finally, suppose that there is a variable y_t which is *integrated* of order d, ($y_t \sim I(d)$). Then, by definition,

$$\Delta^d y_t = z_t$$

is stationary, and can be expressed as an *ARMA(p,q)* process. Therefore, it follows that

$$\phi(L)\Delta^d y_t = \theta(L)\varepsilon_t \tag{3}$$

Such a process is said to be an integrated *ARMA* process or an *ARIMA(p,d,q)* process.

2.1 Mixed Processes

2.1.1 Advantages of mixed processes

Why is it necessary to consider mixed processes? Box and Jenkins (1976) stress *parsimony*. They argue that an *ARMA(p,q)* model with small values of p and q will do as well as a high order $AR(p^*)$ or $MA(q^*)$ process.

Allowing an MA component may give evidence of *over-differencing*. Suppose that $x_t = \varepsilon_t$, then $\Delta x_t = \Delta\varepsilon_t = \varepsilon_t + \theta\varepsilon_{t-1}$, where $\theta = -1$. If we find an estimated parameter $\widehat{\theta}$ close to -1, then this is evidence for over-differencing. This could not be picked up in a *pure* AR model.

2.1.2 Problems with mixed processes

One problem with estimating mixed processes is that of *common factors*. Suppose the 'true' model is *ARMA(p,q)* but the investigator mistakenly fits *ARMA(p+1,q+1)*. If the true model is given by $\phi(L)x_t = \theta(L)\varepsilon_t$, then the estimated model can be written

$$(1 - \alpha_{p+1}L)\phi(L)x_t = (1 + \beta_{q+1}L)\theta(L)\varepsilon_t\,.$$

This model is *not identified* since $\alpha_{p+1} = -\beta_{q+1} = \gamma$, reduces the model to *ARMA(p,q)* for *any* value of the root γ. This means that a *general-to-simple modelling strategy* will not work with mixed $ARMA$ models.

3 Identifying the Order of the *ARMA* Process

3.1 The Autocorrelation Function

The autocorrelation function of a stationary variable x_t is the correlation between x_t and x_{t-j}, considered as a function of j.

$$\rho_j = \frac{Cov(x_t, x_{t-j})}{Var(x_t)}\,.$$

The estimated sample autocorrelation function or *correlogram* is given by

$$\widehat{\rho}_j = \frac{\sum_{t=j+1}^{T}(x_t - \overline{x})(x_{t-j} - \overline{x})}{\sum_{t=1}^{T}(x_t - \overline{x})^2}\,.$$

In large samples the correlogram should mirror the shape of the theoretical autocorrelation function. This may help distinguish between pure AR and pure MA processes.

4 Diagnostic Checking

Having estimated an *ARIMA* model of chosen order, the residuals should be checked for evidence of autocorrelation.

4.1 Box-Pierce or Q statistic

This statistic, proposed by Box and Pierce (1970), is a *portmanteau* statistic for testing autocorrelation. The form of the statistic is given by

$$Q = T \sum_{j=1}^{n} \widehat{\rho}_j^2 \sim_a \chi_n^2$$

where $\widehat{\rho}_j^2$ is the *squared* sample autocorrelation coefficient of jth order.

4.2 Box-Ljung (or Modified Box-Pierce) statistic

Ljung and Box (1978) suggested a correction to the Box-Pierce statistic that has better properties in small samples. The modified statistic is given by

$$Q^* = T(T+2) \sum_{j=1}^{n} (T-j)^{-1} \widehat{\rho}_j^2 \sim_a \chi_n^2$$

4.3 Lagrange Multiplier tests

The Box-Pierce statistic can be shown to be a *Lagrange Multiplier* test of the hypothesis

$$H_0 : AR(0) \quad \text{or} \quad MA(0)$$

against the general alternative

$$H_1 : AR(n) \quad \text{or} \quad MA(n)\,.$$

It is well known that this test has little power against specific alternatives within the general class. Alternatively, we can construct LM tests against a more *specific* alternative hypothesis. These are likely to have higher power.

5 Further reading

Hamilton (1994) is a good textbook treatment. Harvey (1993) treats *ARIMA* models from the state-space perspective. Box and Jenkins (1976) is a surprisingly easy and good read.

References

[1] Box, G.E.P, and G.M. Jenkins, (1976), *Time Series Analysis: Forecasting and Control*, (2nd ed.), Holden-Day, San Fransisco.

[2] Box, G.E.P. and D.A. Pierce, (1970), 'Distribution of residual autocorrelations in autoregressive integrated moving average time series models', *Journal of the American Statistical Association*, 65, 1509–1526.

[3] Hamilton, J.D. (1994), *Time Series Analysis*, Princeton University Press, Princeton, NJ.

[4] Harvey, A.C. (1993), *Time Series Models*, (2nd. ed.), Philip Allan, New York.

[5] Ljung, G.M. and G.E.P. Box (1978), 'On a measure of lack of fit in time series models', *Biometrika*, 66, 67–72.