SPECIFICATION OF ARIMA MODELS BY THE BOX-JENKINS METHOD

Jean-Marie Dufour

May 2005

1 Basic steps

$$\begin{aligned}\varphi_p(B)(1-B)^d X_t &= \mu_0 + \theta_q(B)\,u_t \\ \varphi_p(B) &= 1-\varphi_1 B - \ldots - \varphi_p B^p \\ \theta_q(B) &= 1-\theta_1 B - \ldots - \varphi_q\, B^q\end{aligned}$$

(1) Specification (identification)

(a) Transformation of X_t
- Logarithm or power transformation
- Differencing (d)

(b) Values of p and q

(2) Estimation

(3) Validation (diagnostic checking)

$$(3) \rightarrow (1) \rightarrow (2) \rightarrow (3) \rightarrow (1) \ldots$$

up to a satisfactory model

2 Transformations

Objective : Obtain a series which looks stationary in mean and variance

(a) Variance stabilizing transformations

- Log or not

$$\begin{aligned} X_t^* &= X_t \\ &= \log(X_t) \end{aligned}$$

- Box-Cox transformations

$$\begin{aligned} X_t^* &= (X_t + m)^\lambda \quad , \quad \text{if } \lambda \neq 0 \\ &= \log(X_t + m) \ , \quad \text{if } \lambda = 0 \end{aligned}$$

or

$$X_t^* = \frac{(X_t + m)^\lambda - 1}{\lambda}$$

(b) Mean stabilizing transformations

$$\widetilde{X}_t = (1 - B)^d X_t^*$$

3 Identification of p and q

2 basic instruments

(1) Sample autocorrelations determine q
for $MA(q)$ model

(2) Sample partial autocorrelations determine p
for $AR(p)$ model

3.1 Identification of q for a $MA(q)$

For a $MA(q)$,

$$\rho_k = 0,\ \text{for } k > q\ .$$

If $k > q$, the asymptotic variance of r_ρ is

$$V(r_k) = \frac{1}{T}\left\{1 + 2\sum_{j=1}^{q} \rho_j^2\right\}\ .$$

If X_t follows a $MA(q)$,

$$\sqrt{T}\, r_k \underset{T\to\infty}{\longrightarrow} N\left[0, \bar{\sigma}_k^2\right]$$

$$\bar{\sigma}_k^2 = 1 + 2\sum_{j=1}^{q} \rho_j^2$$

σ_k can be consistently estimated by

$$\hat{\sigma}_k^2 = 1 + 2\sum_{j=1}^{q} r_j^2\ ,$$

hence

$$\sqrt{T}\frac{r_k}{\hat{\sigma}_k} = \frac{r_k}{\hat{\sigma}(r_k)} \underset{T\to\infty}{\longrightarrow} N(0,1)\ .$$

For $k > q$,

$$\hat{\sigma}(r_k) = \frac{1}{\sqrt{T}}\hat{\sigma}_k\ .$$

r any $k > q$,

$$\left|\frac{r_k}{\hat{\sigma}\left(r_k\right)}\right| > c\left(\alpha/2\right)$$

$$P\left[N\left(0,1\right) > c\left(\alpha/2\right)\right] = \frac{\alpha}{2}$$

is an indication that we do not have a $MA(q)$ process. For $j > q$ and $k > q$, r_j and r_k are asymptotically uncorrelated (independent since Gaussian).

To determine the order of a $MA(q)$, we look for a cut-off point in the autocorrelations :

$$\begin{aligned} r_k &\neq 0 \quad \text{for} \quad k \leq q\ , \\ r_k &\simeq 0 \quad \text{for} \quad k > q\ . \end{aligned}$$

For $AR\left(p\right)$ process

$$\rho_k = \sum_{j=1}^{p} \varphi_j \rho_{k-j}$$

i.e. an exponential decay of ρ_k with possibly oscillations.

3.2 Identification of p for an $AR(p)$

Consider the k equations system:

$\rho_j = a_{k1}\rho_{j-1} + a_{k2}\rho_{j-2} + \cdots + a_{kk}\rho_{j-k}\,, \quad j = 1, \ldots, k\,.$

a_{kk} is the partial autocorrelation at lag k.

For an $AR(p)$ process,

$$a_{kk} = 0, \quad \text{for } k > p\,.$$

a_{kk} can be consistently estimated on replacing ρ_j by r_j :

$r_j = \hat{a}_{k1} r_{j-1} + \hat{a}_{k2} r_{j-2} + \ldots + \hat{a}_{kk} r_{j-k}\,, \quad j = 1, \ldots, k\,.$

For an $AR(p)$ process

$$\sqrt{T}\hat{a}_{kk} \overset{a}{\sim} N[0,1]\,, \quad k > p\,.$$

we can test whether we have an $AR(p)$ by checking

$$\left|\sqrt{T}\hat{a}_{kk}\right| > c(\alpha/2)$$

$$\frac{\hat{a}_{kk}}{1/\sqrt{T}} \overset{a}{\sim} N[0,1]\ .$$

For a $MA(q)$ process, a_{kk} declines at an exponential rate.

For an $ARMA(p,q)$ with $p \geq 1, \quad q \geq 0$, both ρ_k and a_{kk} decline at exponential rates

Process type	Autocorrelations	Partial autocorrelations
$MA\,(q)$	$\rho_q \neq 0$ $\rho_k = 0,\ k > q$	Exponential decay oscillations possible $a_{kk} \underset{k\to\infty}{\longrightarrow} 0$
$AR\,(p)$	$\rho_k = \varphi_1 \rho_{k-1} + \ldots + \varphi_p \rho_{k-p}$ Exponential decay $\rho_k \underset{k\to\infty}{\longrightarrow} 0$	$a_{pp} \neq 0$ $a_{kk} = 0,\ k > p$
$ARMA\,(p,q)$	Irregular for $k = 1, \ldots, p$ $\rho_k = \varphi_1 \rho_{k-1} + \ldots + \varphi_p \rho_{k-p}$ $k > p$ $\rho_k \underset{k\to\infty}{\longrightarrow} 0$	$a_{kk} \underset{k\to\infty}{\longrightarrow} 0$