SPECIFICATION OF ARIMA MODELS
BY THE BOX-JENKINS METHOD

Jean-Marie Dufour

May 2005
1 Basic steps

\[ \varphi_p(B)(1 - B)^d X_t = \mu_0 + \theta_q(B) u_t \]
\[ \varphi_p(B) = 1 - \varphi_1 B - \ldots - \varphi_p B^p \]
\[ \theta_q(B) = 1 - \theta_1 B - \ldots - \varphi_q B^q \]

(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
  \[ X_t^* = X_t \]
  \[ = \log(X_t) \]

- Box-Cox transformations
  \[ X_t^* = (X_t + m)^\lambda, \quad \text{if } \lambda \neq 0 \]
  \[ = \log(X_t + m), \quad \text{if } \lambda = 0 \]
  or
  \[ X_t^* = \frac{(X_t + m)^\lambda - 1}{\lambda} \]

(b) Mean stabilizing transformations
  \[ \tilde{X}_t = (1 - B)^d X_t^* \]
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 $r_\rho$ is 
\[ V(r_k) = \frac{1}{T} \left\{ 1 + 2 \sum_{j=1}^{q} \rho_j^2 \right\}. \]

If $X_t$ follows a $MA(q)$, 
\[ \sqrt{T} r_k \xrightarrow{T \to \infty} N \left[ 0, \sigma_k^2 \right] \]
\[ \hat{\sigma}_k^2 = 1 + 2 \sum_{j=1}^{q} \rho_j^2 \]

$\sigma_k$ can be consistently estimated by 
\[ \hat{\sigma}_k^2 = 1 + 2 \sum_{j=1}^{q} r_j^2, \]

hence 
\[ \sqrt{T} \frac{r_k}{\hat{\sigma}_k} = \frac{r_k}{\hat{\sigma}(r_k)} \xrightarrow{T \to \infty} N(0, 1). \]

For $k > q$, 
\[ \hat{\sigma}(r_k) = \frac{1}{\sqrt{T}} \hat{\sigma}_k. \]
r any \( k > q \),
\[
\left| \frac{r_k}{\hat{\sigma}(r_k)} \right| > c (\alpha/2)
\]

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

\[
r_k \not= 0 \quad \text{for} \quad k \leq q,
\]

\[
r_k \simeq 0 \quad \text{for} \quad k > q.
\]

For \( AR(p) \) process

\[
\rho_k = \sum_{j=1}^{p} \varphi_j \rho_{k-j}
\]
i.e. an exponential decay of \( \rho_k \) 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 $\rho_j$ by $r_j$:

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

For an $AR(p)$ process

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

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

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

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

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 $\rho_k$ and $a_{kk}$ decline at exponential rates.
<table>
<thead>
<tr>
<th>Process type</th>
<th>Partial autocorrelations</th>
<th>Autocorrelations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Autocorrelations</td>
<td>$0 \overset{d}{\overset{\infty}{\longleftarrow}} \tau d \rho$</td>
<td>$0 \overset{d}{\overset{\infty}{\leftarrow}} \tau d$</td>
</tr>
<tr>
<td>AR</td>
<td>$\rho_k = \phi_1 \rho_{k-1} + \ldots + \phi_p \rho_{k-p}$</td>
<td>$\rho_k = \phi_1 \rho_{k-1} + \ldots + \phi_p \rho_{k-p}$</td>
</tr>
<tr>
<td>ARMA</td>
<td>Irregular for $k = 1, \ldots, p$</td>
<td>$\rho_k \overset{d}{\rightarrow} 0$</td>
</tr>
<tr>
<td>$d &lt; \tau, 0 = \frac{\tau \rho_{d}}{\phi_1}$</td>
<td>$d &lt; \tau, 0 = \frac{\tau \rho_{d}}{\phi_1}$</td>
<td></td>
</tr>
<tr>
<td>$0 \not= d \phi_1$</td>
<td>$0 \not= d \phi_1$</td>
<td></td>
</tr>
<tr>
<td>Exponential decay</td>
<td>Exponential decay</td>
<td></td>
</tr>
<tr>
<td>$d &lt; \tau, 0 = \frac{\tau \rho_{d}}{\phi_1}$</td>
<td>$d &lt; \tau, 0 = \frac{\tau \rho_{d}}{\phi_1}$</td>
<td></td>
</tr>
<tr>
<td>$0 \not= d \phi_1$</td>
<td>$0 \not= d \phi_1$</td>
<td></td>
</tr>
<tr>
<td>Exponential decay</td>
<td>Exponential decay</td>
<td></td>
</tr>
<tr>
<td>$b &lt; \tau, 0 = \frac{\tau b \rho_{d}}{\phi_1}$</td>
<td>$b &lt; \tau, 0 = \frac{\tau b \rho_{d}}{\phi_1}$</td>
<td></td>
</tr>
<tr>
<td>$0 \not= b \phi_1$</td>
<td>$0 \not= b \phi_1$</td>
<td></td>
</tr>
</tbody>
</table>