Lecture 3: ARIMA(p,d,q) models

Florian Pelgrin

University of Lausanne, École des HEC
Department of mathematics (IMEA-Nice)

Motivation

- Characterize the main properties of ARIMA(p) models.

- Discuss asymptotic equivalence with ARMA(p,q) models
Road map

1. Introduction

2. ARIMA(p,d,q) model
   - Definition
   - Fundamental representations
   - Equivalence with ARMA(p,q) models
   - Autoregressive approximation
   - Moving average approximation

3. Application
1. Definition

- To some extent, ARIMA(p,d,q) models are a generalization of ARMA(p,q) models: the d-differenced process $\Delta^dX_t$ is (asymptotically) an ARMA(p,q) process:

- On the other hand, the statistical properties of the two models are different, especially in terms of forecasting.
Definition

A stochastic process \((X_t)_{t \geq -p-d}\) is said to be an \(ARIMA(p, d, q)\)—an integrated mixture autoregressive moving average model—if it satisfies the following equation:

\[
\Phi(L)(1 - L)^d X_t = \mu + \Theta(L) \epsilon_t \quad \forall t \geq 0
\]

where \(\epsilon_t\) is a (weak) white noise process with variance \(\sigma^2_\epsilon\), the lag polynomials are given by:

\[
\Phi(L) = 1 - \phi_1 L - \cdots - \phi_p L^p \quad \text{with} \quad \phi_p \neq 0
\]

\[
\Theta(L) = 1 + \theta_1 L + \cdots + \theta_q L^q \quad \text{with} \quad \theta_q \neq 0,
\]

and the initial conditions:

\[
Z_{-1} = \{X_{-1}, \cdots, X_{-p-d}, \epsilon_{-1}, \cdots, \epsilon_{-q}\}
\]

are such that:
Remarks:

1. The stochastic process $X_t$ also writes:

$$\Phi(L)\Delta^d X_t = \mu + \Theta(L)\epsilon_t \text{ or } \Phi(L)Y_t = \mu + \Theta(L)\epsilon_t$$

where $Y_t = \Delta^d X_t = (1 - L)^d X_t$.

2. Initial conditions are fundamental.

Example: Consider the two stochastic processes:

$$X_t = \rho X_{t-1} + \epsilon_t$$

$$Y_t \equiv X_t - X_{t-1} = \eta_t$$

where $\rho = 0.9$, $\epsilon_t$ and $\eta_t$ are (weak) white noises.
ARIMA(p,d,q) model

Definition

AR(1) model
Random walk (without drift)

Note: The black, blue, and red solid lines correspond respectively to $X_0 = 0$, $X_0 = 2$, and $X_0 = -2$.

**Figure:** Initial conditions and stationarity
1.2. Fundamental representations

**Definition (Fundamental representation)**

Let \((X_t)_{t \geq -p-d}\) denote the following ARIMA\((p,d,q)\) stochastic process:

\[
\Phi(L) \Delta^d X_t = \mu + \Theta(L) \epsilon_t \quad \forall t \geq 0
\]

where \(\Phi(L) = 1 - \phi_1 L - \cdots - \phi_p L^p\) and \(\Theta(L) = 1 + \theta_1 L + \cdots + \theta_q L^q\), \(\theta_q \neq 0, \phi_p \neq 0\), \(\mu\) is a constant term, \((\epsilon_t)\) is a weak white noise, and the initial conditions are uncorrelated with \(\epsilon_t\) \((t \geq 0)\). This representation is said to be causal or fundamental if and only if:

(i) All of the roots of the (inverse) characteristic equation associated to \(\Phi\) are of modulus less (larger) than one.

(ii) All of the roots of the (inverse) characteristic equation associated to \(\Theta\) are of modulus less (larger) than one.
Definition (Fundamental minimal representation)

Let \((X_t)_{t \geq -p-d}\) denote the following ARIMA\((p,d,q)\) stochastic process:

\[
\Phi(L) \Delta^d X_t = \mu + \Theta(L) \epsilon_t \quad \forall t \geq 0
\]

where \(\Phi(L) = 1 - \phi_1 L - \cdots - \phi_p L^p\) and \(\Theta(L) = 1 + \theta_1 L + \cdots + \theta_q L^q\), \(\theta_q \neq 0, \phi_p \neq 0\), \(\mu\) is a constant term, \((\epsilon_t)\) is a weak white noise, and the initial conditions are uncorrelated with \(\epsilon_t (t \geq 0)\). This representation is said to be causal or fundamental if and only if:

(i) All of the roots of the (inverse) characteristic equation associated to \(\Phi\) are of modulus less (larger) than one.

(ii) All of the roots of the (inverse) characteristic equation associated to \(\Theta\) are of modulus less (larger) than one.

(iii) The (inverse) characteristic polynomials have no common roots.
1.3. Equivalence with ARMA\((p,q)\) models

**Proposition**

Let \((X_t)_{t \geq -p-d}\) denote a minimal and causal ARIMA\((p,d,q)\) stochastic process:

\[
\Phi(L)(1 - L)^d X_t = \mu + \Theta(L)\varepsilon_t.
\]

The stochastic process defined by:

\[
Y_t = \Delta^d X_t = (1 - L)^d X_t
\]

is asymptotically equivalent to an ARMA\((p,q)\) process.
### Definition

The autoregressive approximation (and not the AR(∞) representation) of a causal and minimal ARIMA(p,d,q) stochastic process is given by:

\[ A_t(L)X_t = \mu_0 + \epsilon_t + h(t)'Z_{-1}. \]

where \( A_t(L) = \sum_{j=0}^{t} a_j L^j \), \( a_0 = 1 \), and the \( a_j \) terms are the coefficients of the division (by increasing powers) of \( \Phi(u) \) by \( \Theta(u) \), \( \mu_0 \) is a constant term, and \( h(t) \in \mathbb{R}^{p+d+q} \) (with \( \lim_{t \to +\infty} h(t) = 0 \)).
1.5. Moving average approximation

Definition

The moving average approximation (and not the MA(∞) representation) of a causal and minimal ARIMA(p,d,q) stochastic process is given by:

\[ X_t = \mu_1 + B_t(L)\epsilon_t + \tilde{h}(t)'Z_{t-1}. \]

where \( B_t(L) = \sum_{j=0}^{t} b_j L^j \), \( b_0 = 1 \), and the \( b_j \) terms are the coefficients of the division (by increasing powers) of \( \Theta(u) \) by \( \Phi(u) \), \( \mu_1 \) is a constant term, and \( \tilde{h}(t) \in \mathbb{R}^{p+d+q} \) (with \( \lim_{t \to +\infty} \tilde{h}(t) = 0 \)).
Example

Starting from:

$$(1 - \phi L)(1 - L)X_t = \epsilon_t - \theta \epsilon_{t-1} \ \forall t \geq 0$$

with the initial conditions $\epsilon_{-1}$, $X_{-1}$ and $X_{-2}$;

One gets:

$$\Delta X_t = \frac{1 - \theta L}{1 - \phi L} \tilde{\epsilon}_t - \phi^t \theta \epsilon_{-1} + \phi^{t+1}(X_{-1} - X_{-2})$$

$$= \frac{1 - \theta L}{1 - \phi L} \tilde{\epsilon}_t + h'_t Z_{-1}.$$
Visual inspection of the autocorrelation function may indicate that the (log) US dividend is not stationary. The autocorrelogram of the first-differenced variable decreases quickly to zero—the first-difference variable may be stationary.

Unit root tests tends to favor the assumption of non-stationarity (see later on).

The chosen specification is then an ARIMA($p,1,q$) model:

$$\Phi(L)(1 - L) d_t = \Theta(L) \epsilon_t$$

where $\epsilon_t$ is a white noise process, and:

$$\Phi(L) = 1 - \phi_1 L - \cdots - \phi_p L^p$$

$$\Theta(L) = 1 + \theta_1 L + \cdots + \theta_q L^q.$$
Figure 34: (Partial) autocorrelogram of (log) US stock market dividends

<table>
<thead>
<tr>
<th>Level</th>
<th>First-difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Autocorrelation</td>
<td>Partial Correlation</td>
</tr>
<tr>
<td>1</td>
<td>0.989</td>
</tr>
<tr>
<td>2</td>
<td>0.998</td>
</tr>
<tr>
<td>3</td>
<td>0.979</td>
</tr>
<tr>
<td>4</td>
<td>0.973</td>
</tr>
<tr>
<td>5</td>
<td>0.905</td>
</tr>
<tr>
<td>6</td>
<td>0.959</td>
</tr>
<tr>
<td>7</td>
<td>0.955</td>
</tr>
<tr>
<td>8</td>
<td>0.942</td>
</tr>
<tr>
<td>9</td>
<td>0.939</td>
</tr>
<tr>
<td>10</td>
<td>0.927</td>
</tr>
<tr>
<td>11</td>
<td>0.918</td>
</tr>
<tr>
<td>12</td>
<td>0.910</td>
</tr>
<tr>
<td>13</td>
<td>0.902</td>
</tr>
<tr>
<td>14</td>
<td>0.893</td>
</tr>
<tr>
<td>15</td>
<td>0.894</td>
</tr>
<tr>
<td>16</td>
<td>0.876</td>
</tr>
<tr>
<td>17</td>
<td>0.867</td>
</tr>
<tr>
<td>18</td>
<td>0.859</td>
</tr>
<tr>
<td>19</td>
<td>0.848</td>
</tr>
<tr>
<td>20</td>
<td>0.841</td>
</tr>
<tr>
<td>21</td>
<td>0.832</td>
</tr>
<tr>
<td>22</td>
<td>0.834</td>
</tr>
<tr>
<td>23</td>
<td>0.815</td>
</tr>
<tr>
<td>24</td>
<td>0.807</td>
</tr>
<tr>
<td>25</td>
<td>0.798</td>
</tr>
<tr>
<td>26</td>
<td>0.789</td>
</tr>
<tr>
<td>27</td>
<td>0.781</td>
</tr>
<tr>
<td>28</td>
<td>0.772</td>
</tr>
<tr>
<td>29</td>
<td>0.763</td>
</tr>
<tr>
<td>30</td>
<td>0.755</td>
</tr>
<tr>
<td>31</td>
<td>0.746</td>
</tr>
<tr>
<td>32</td>
<td>0.737</td>
</tr>
<tr>
<td>33</td>
<td>0.729</td>
</tr>
<tr>
<td>34</td>
<td>0.720</td>
</tr>
<tr>
<td>35</td>
<td>0.711</td>
</tr>
<tr>
<td>36</td>
<td>0.702</td>
</tr>
</tbody>
</table>
### Estimation of an ARIMA(2,1,1) model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t-stat.</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c$</td>
<td>0.0007</td>
<td>0.0003</td>
<td>1.8777</td>
<td>0.0611</td>
</tr>
<tr>
<td>$\phi_1$</td>
<td>0.7342</td>
<td>0.1096</td>
<td>6.6982</td>
<td>0.0000</td>
</tr>
<tr>
<td>$\phi_2$</td>
<td>0.1206</td>
<td>0.0619</td>
<td>1.9477</td>
<td>0.0521</td>
</tr>
<tr>
<td>$\theta_1$</td>
<td>-0.6442</td>
<td>0.1038</td>
<td>-6.2008</td>
<td>0.0000</td>
</tr>
</tbody>
</table>
Figure 35: Adjusted error terms