Lecture 2: ARMA(p,q) models (part 3)

Florian Pelgrin

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

Sept. 2011 - Jan. 2012

Motivation

- Characterize the main properties of ARMA(p,q) models.
- Estimation of ARMA(p,q) models

Road map

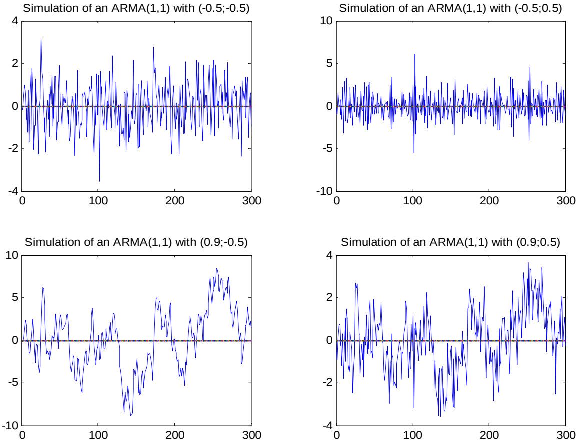
1. ARMA(1,1)
1.1. Definition and conditions

Definition

A stochastic process $(X_t)_{t\in\mathbb{Z}}$ is said to be a mixture autoregressive moving average model of order 1, ARMA(1,1), if it satisfies the following equation :

$$\begin{aligned} X_t &= \mu + \phi X_{t-1} + \epsilon_t + \theta\epsilon_{t-1} \ \forall t \\ \Phi(L)X_t &= \mu + \Theta(L)\epsilon_t \end{aligned}$$

where $\theta \neq 0$, $\theta \neq 0$, μ is a constant term, $(\epsilon_t)_{t\in\mathbb{Z}}$ is a weak white noise process with expectation zero and variance σ_ϵ^2 ($\epsilon_t \sim WN(0,\sigma_\epsilon^2)$), $\Phi(L) = 1-\phi L$ and $\Theta(L) = 1+\theta L$.


Simulation of an ARMA(1,1) with (-0.5;-0.5)
Simulation of an ARMA(1,1) with (-0.5;0.5)
Simulation of an ARMA(1,1) with (0.9;-0.5)
Simulation of an ARMA(1,1) with (0.9;0.5)
4
2
0
-2
-4
10
5
-5
-10
0
100
200
300

- The properties of an ARMA(1,1) process are a mixture of those of an AR(1) and MA(1) processes :
  - The (stability) stationarity condition is the one of an AR(1) process (or ARMA(1,0) process) :

$$|\phi| < 1.$$

  - The invertibility condition is the one of a MA(1) process (or ARMA(0,1) process) :

$$|\theta| < 1.$$

  - The representation of an ARMA(1,1) process is fundamental or causal if :

$$|\phi| < 1 \text{ and } |\theta| < 1.$$

  - The representation of an ARMA(1,1) process is said to be minimal and causal if :

$$|\phi| < 1, |\theta| < 1 \text{ and } \phi \neq \theta.$$

- If (X_t) is stable and thus weakly stationary, then (X_t) has an infinite moving average representation (MA(∞)) :

$$\begin{aligned} X_t &= \frac{\mu}{1-\phi} + (1-\phi L)^{-1}(1+\theta L)\epsilon_t \\ &= \frac{\mu}{1-\phi} + \sum_{k=0}^{\infty} a_k \epsilon_{t-k} \end{aligned}$$

where :

$$\begin{aligned} a_0 &= 1 \\ a_k &= \phi^k + \theta\phi^{k-1}. \end{aligned}$$

- If (X_t) is invertible, then (X_t) has an infinite autoregressive representation (AR(∞)) :

$$(1-\theta^* L)^{-1}(1-\phi L)X_t \quad = \quad \frac{\mu}{1-\theta^*} + \epsilon_t$$

i.e.

$$X_t = \quad = \quad \frac{\mu}{1-\theta^*} + \sum_{k=1}^{\infty} b_k X_{t-k} + \epsilon_t$$

where $\theta^* = -\theta$, and :

$$b_k = -\theta^{*k} - \theta^{*k-1}\phi.$$

1.2. Moments

Definition

Let (X_t) denote a stationary stochastic process that has a fundamental ARMA(1,1) representation, $X_t = \mu + \phi X_{t-1} + \epsilon_t + \theta \epsilon_{t-1}$. Then :

$$\begin{aligned}
\mathbb{E}[X_t] &= \frac{\mu}{1-\phi} \equiv m \\
\gamma_X(0) &\equiv \mathbb{V}(X_t) = \frac{1+2\phi\theta+\theta^2}{1-\phi^2}\sigma_\epsilon^2 \\
\gamma_X(1) &\equiv \mathbb{C}ov[X_t, X_{t-1}] = \frac{(\phi+\theta)(1+\phi\theta)}{1-\phi^2}\sigma_\epsilon^2 \\
\gamma_X(h) &= \phi\gamma_X(h-1) \text{ for } |h| > 1.
\end{aligned}$$

<u>Proof</u> : See Appendix 1.

Definition

The autocorrelation function of an ARMA(1,1) process satisfies :

$$\rho_X(h) = \begin{cases} 1 \text{ if } h = 0 \\ \frac{(\phi+\theta)(1+\phi\theta)}{1+2\phi\theta+\theta^2} \text{ if } |h| = 1 \\ \phi\rho_X(h-1) \text{ if } |h| > 1. \end{cases}$$

- The autocorrelation function of an ARMA(1,1) process exhibits exponential decay towards zero : it does not cut off but gradually dies out as h increases.

- The autocorrelation function of an ARMA(1,1) process displays the shape of that of an AR(1) process for $|h| > 1$.

- Partial Autocorrelation :

The partial autocorrelation function of an ARMA(1,1) process will gradually die out (the same property as a moving average model).

Estimation

- Same techniques as before, especially those of MA models.
- Yule-Walker estimator : the extended Yule-Walker equations could be used in principe to estimate the AR coefficients but the MA coefficients need to be estimated by other means.
- In the presence of moving average components, the least squares estimator becomes nonlinear and the corresponding estimator is the conditional nonlinear least squares estimator (see estimation of MA(q) models). It has to be solved with numerical methods.
- Taking explicit distributional assumption for the error term, the conditional or exact maximum likelihood estimator can be computed (using also numerical or optimization methods).
- Other methods are also available : the Kalman filter, the generalized method of moments, etc.

Estimation of ARMA(p,q) models (true DGP: $\mu = 0$, $\varphi = 0.9$, and $\theta = 0.5$)

	Coefficient	Std. Error	t-Statistic	p-value.	Akaike info criterion	Schwarz criterion
			ARMA(1,1)			
C	0.4532	0.7558	0.5996	0.5490	**2.8841**	**2.9134**
AR(1)	0.9112	0.0184	49.4884	0.0000		
MA(1)	0.4639	0.0409	11.3368	0.0000		
			ARMA(2,2)			
C	0.3137	0.6942	0.4518	0.6516	2.8865	2.9288
AR(1)	1.4836	0.2655	5.5879	0.0000		
AR(2)	-0.5230	0.2422	-2.1598	0.0313		
MA(1)	-0.1233	0.2666	-0.4625	0.6439		
MA(2)	-0.2837	0.1245	-2.2784	0.0231		
			ARMA(2,1)			
	0.3941	0.7431	0.5303	0.5961	2.8857	2.9195
AR(1)	0.8581	0.0964	8.8992	0.0000		
AR(2)	0.0491	0.0937	0.5244	0.6002		
MA(1)	0.5047	0.0843	5.9848	0.0000		
			AR(2)			
C	0.4276	0.6628	0.6451	0.5192	2.9206	2.9459
AR(1)	1.2819	0.0419	30.5971	0.0000		
AR(2)	-0.3523	0.0416	-8.4649	0.0000		
			AR(1)			
C	0.2857	0.9394	0.3042	0.7611	3.0559	3.0728
AR(1)	0.9467	0.0135	70.1399	0.0000		
			MA(2)			
C	0.6545	0.2043	3.2042	0.0014	3.6527	3.6779
MA(1)	1.3803	0.0329	41.9665	0.0000		
MA(2)	0.6740	0.0324	20.8027	0.0000		
			MA(4)			
C	0.6259	0.2633	2.3768	0.0178	3.2039	3.2461
MA(1)	1.4334	0.0408	35.1707	0.0000		
MA(2)	1.2250	0.0657	18.6595	0.0000		
MA(3)	0.8552	0.0658	12.9989	0.0000		
MA(4)	0.4131	0.0407	10.1483	0.0000		

Note: C, AR(j), and MA(j) are respectively the estimate of the constant term, the jth autogressive term, and the jth moving average term.

2. ARMA(p,q)
2.1. Definition and conditions

Definition

A stochastic process $(X_t)_{t\in\mathbb{Z}}$ is said to be a mixture autoregressive moving average model of order p and q, ARMA(p,q), if it satisfies the following equation :

$$\begin{aligned} X_t &= \mu + \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q} \ \forall t \\ \Phi(L) X_t &= \mu + \Theta(L) \epsilon_t \end{aligned}$$

where $\theta_q \neq 0$, $\phi_p \neq 0$, μ is a constant term, $(\epsilon_t)_{t\in\mathbb{Z}}$ is a weak white noise process with expectation zero and variance σ_ϵ^2 ($\epsilon_t \sim WN(0, \sigma_\epsilon^2)$), $\Phi(L) = 1 - \phi_1 L - \cdots - \phi_p L^p$ and $\Theta(L) = 1 + \theta_1 L + \cdots + \theta_q L^q$.

Main idea of ARMA(p,q) models

- Approximate Wold form of stationary time series by parsimonious parametric models
  - AR and MA models can be cumbersome because one may need a high-order model with many parameters to adequately describe the data dynamics (see the effective Fed fund rate application)
  - By mixing AR and MA models into a more compact form, the number of parameters is kept small...

- The properties of an ARMA(p,q) process are a mixture of those of an AR(p) and MA(q) processes :
  - The (stability) stationarity conditions are those of an AR(p) process (or ARMA(p,0) process) :
$$z^p\Phi(z^{-1}) = 0 \equiv z^p - \phi_1 z^{p-1} - \cdots - \phi_p = 0 \Leftrightarrow |z_i| < 1.$$
for $i = 1, \cdots, p$.
  - The invertibility conditions are those of an MA(q) process (or ARMA(0,q) process) :
$$z^q\Theta(z^{-1}) = 0 \equiv z^q + \theta_1 z^{q-1} + \cdots + \theta_q = 0 \Leftrightarrow |\tilde{z}_i| < 1.$$
for $i = 1, \cdots, q$.
  - The representation of an ARMA(p,q) process is fundamental or causal if it is stable and invertible
  - The representation of an ARMA(1,1) process is said to be minimal and causal if it is stable, invertible and the characteristic polynomials $z^p\Phi(z^{-1})$ and $z^q\Theta(z^{-1})$ have no common roots.

Definition

The representation of a mixture autoregressive moving average process of order p and q defined by :

$$X_t = \mu + \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q},$$

is said to be a minimal causal (fundamental) representation—(ϵ_t) is the innovation process—if :

(i) All the roots of the characteristic equation associated to Φ $z^p - \phi_1 z^{p-1} - \cdots - \phi_p = 0$ are of modulus less than one, $|z_i| < 1$ for $i = 1, \cdots, p$;

(ii) All the roots of the characteristic equation associated to Θ $z^q + \theta_1 z^{q-1} + \cdots + \theta_q = 0$ are of modulus less than one, $|\tilde{z}_i| < 1$ for $i = 1, \cdots, q$;

(iii) The characteristic polynomials $z^p \Phi(z^{-1})$ and $z^q \Theta(z^{-1})$ have no common roots.

- If (X_t) is stable and thus weakly stationary, then (X_t) has an infinite moving average representation (MA(∞)) :

$$\begin{aligned} X_t &= \frac{\mu}{1-\sum_{k=1}^{p}\phi_k} + \Phi(L)^{-1}(1+\theta L)\epsilon_t \\ &= \frac{\mu}{1-\sum_{k=1}^{p}\phi_k} + \sum_{k=0}^{\infty} a_k \epsilon_{t-k} \end{aligned}$$

where :

$$a_0 = 1$$
$$\sum_{k=0}^{\infty} |a_k| < \infty$$

- If (X_t) is invertible, then (X_t) has an infinite autoregressive representation (AR(∞)) :

$$\Theta(L)^{-1}\Phi(L)X_t \quad = \quad \frac{\mu}{1-\sum_{k=1}^{q}\theta_k^*} + \epsilon_t$$

i.e.

$$X_t = \quad = \quad \frac{\mu}{1-\sum_{k=1}^{q}\theta_q^*} + \sum_{k=1}^{\infty} b_k X_{t-k} + \epsilon_t$$

where $\theta_k^* = -\theta_k$.

2.2. Moments of an ARMA(p,q)

- The properties of the moments of an ARMA(p,q) are also a mixture of those of an AR(1) and MA(1) processes.
- The mean is the same as the one of an AR(p) model (with a constant term) :

$$\mathbb{E}(X_t) = \frac{\mu}{1 - \sum_{k=1}^{p} \phi_k} \equiv m.$$

- Autocorrelation :
  - The autocorrelation function of an ARMA(p,q) process exhibits exponential decay towards zero : it does not cut off but gradually dies out as h increases (possibly damped oscillations.
  - The autocorrelation function of an ARMA(p,q) process displays the shape of that of an AR(p) process for $|h| > max(p, q+1)$.
- Partial Autocorrelation : The partial autocorrelation function of an ARMA(p,q) process will gradually die out (the same property as a MA(q) model).

2.3. Estimation

- Same techniques as in previous models...
  1. Conditional least squares method
  2. Maximum likelihood estimator (conditional or exact)
  3. Generalized method of moments
  4. Etc

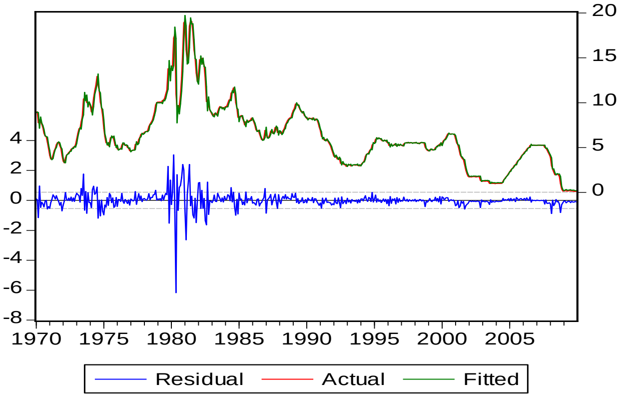
3. Application

- Effective Fed fund rate : 1970 :01-2010 :01 (monthly observations)
- An ARMA(1,2) captures better the dynamics of the effective Fed fund rate.

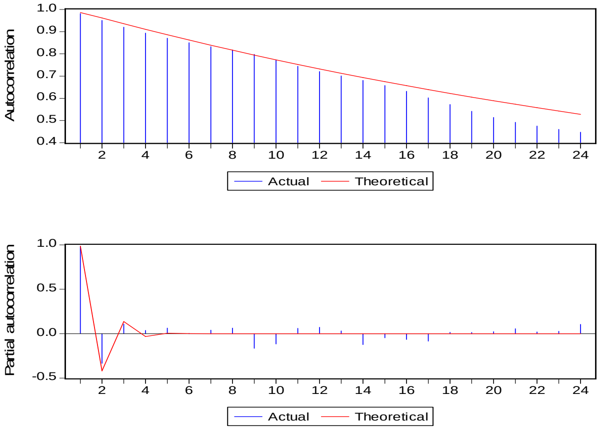
ML estimation of the effective Fed fund rate : ARMA(1,2)

Coefficients	Estimates	Std. Error	P-value
μ	6.225	1.263	0.000
ϕ	0.967	0.012	0.000
θ_1	0.488	0.048	0.000
θ_2	0.082	0.048	0.088

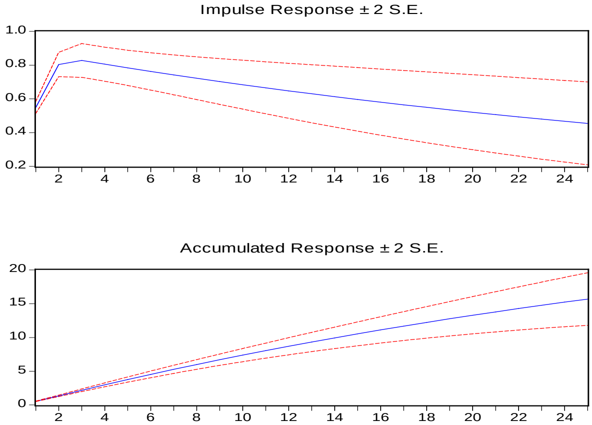
Effective Fed fund rate: ARMA(1,2) specification



Effective Fed fund rate: diagnostics of the ARMA(1,2) specification



Effective Fed fund rate: Impulse response function of the estimated ARMA(1,2) specification



4. Appendix

1. Moments of an ARMA(1,1).

1. Moments of an ARMA(1,1)

- The properties of the moments of an ARMA(1,1) are a mixture of those of an AR(1) and MA(1) processes.
- The mean is the same as the one of an AR(1) model (with a constant term) :

$$\begin{aligned} \mathbb{E}(X_t) &= \mathbb{E}\left(\mu + \phi X_{t-1} + \epsilon_t + \theta \epsilon_{t-1}\right) \\ &= \mu + \phi \mathbb{E}(X_{t-1}) + \mathbb{E}(\epsilon_t) + \theta \mathbb{E}(\epsilon_{t-1}) \\ &= \mu + \phi \mathbb{E}(X_t) \end{aligned}$$

since $\mathbb{E}(X_t) = \mathbb{E}(X_{t-j})$ for all j (stationarity property) and $\mathbb{E}(\epsilon_{t-j}) = 0$ for all j(white noise). Therefore,

$$\mathbb{E}(X_t) = \frac{\mu}{1-\phi} \equiv m.$$

- Autocovariances
  Trick : Proceed in the same way and note that the Yule-Walker equations are the same as those of an AR(1) for $|h| > 1$.
- For $h = 0$:

$$\begin{aligned}
\gamma_X(0) &= \mathbb{E}\left[(X_t - m)(X_t - m)\right] \\
&= \mathbb{E}\left[(\phi(X_{t-1} - m) + \epsilon_t + \theta\epsilon_{t-1})(X_t - m)\right] \\
&= \phi\mathbb{E}\left[(X_t - m)(X_{t-1} - m)\right] + \mathbb{E}\left[\epsilon_t(X_t - m)\right] + \theta\,\underbrace{\mathbb{E}\left[\epsilon_{t-1}(X_t - m)\right]}_{\neq 0} \\
&= \underbrace{\phi\gamma_X(1) + \sigma_\epsilon^2}_{\text{AR(1) part}} + \theta\mathbb{E}\left[(\phi(X_{t-1} - m) + \epsilon_t + \theta\epsilon_{t-1})\,\epsilon_{t-1}\right] \\
&= \phi\gamma_X(1) + \sigma_\epsilon^2 + \theta\phi\mathbb{E}\left[(X_{t-1} - m)\epsilon_{t-1}\right] + \theta\mathbb{E}\left[\epsilon_t\epsilon_{t-1}\right] + \theta^2\mathbb{E}\left[\epsilon_{t-1}^2\right] \\
&= \phi\gamma_X(1) + \sigma_\epsilon^2\left(1 + \theta(\phi + \theta)\right).
\end{aligned}$$

- For $h = 1$:

$$
\begin{aligned}
\gamma_X(1) &= \mathbb{E}\left[(X_t - m)(X_{t-1} - m)\right] \\
&= \mathbb{E}\left[\left(\phi(X_{t-1} - m) + \epsilon_t + \theta\epsilon_{t-1}\right)(X_{t-1} - m)\right] \\
&= \phi\mathbb{E}\left[(X_{t-1} - m)(X_{t-1} - m)\right] + \mathbb{E}\left[\epsilon_t(X_{t-1} - m)\right] \\
&\quad +\theta\mathbb{E}\left[\epsilon_{t-1}(X_{t-1} - m)\right] \\
&= \underbrace{\phi\gamma_X(0)}_{\text{AR(1) part}} + \theta\sigma_\epsilon^2
\end{aligned}
$$

- Solving for $\gamma_X(0)$ and $\gamma_X(1)$:

$$
\begin{aligned}
\gamma_X(0) &\equiv \mathbb{V}(X_t) = \frac{1 + 2\phi\theta + \theta^2}{1 - \phi^2}\sigma_\epsilon^2 \\
\gamma_X(1) &\equiv \mathbb{C}ov(X_t, X_{t-1}) = \frac{(\phi + \theta)(1 + \phi\theta)}{1 - \phi^2}\sigma_\epsilon^2.
\end{aligned}
$$

- For $|h| > 1$:

$$\begin{aligned}
\gamma_X(h) &= \mathbb{E}\left[(X_t - m)(X_{t-h} - m)\right] \\
&= \mathbb{E}\left[(\phi(X_{t-1} - m) + \epsilon_t + \theta\epsilon_{t-1})\,(X_{t-h} - m)\right] \\
&= \phi\mathbb{E}\left[(X_{t-1} - m)(X_{t-h} - m)\right] + \mathbb{E}\left[\epsilon_t(X_{t-h} - m)\right] \\
&\quad +\theta\mathbb{E}\left[\epsilon_{t-1}(X_{t-h} - m)\right] \\
&= \underbrace{\phi\gamma_X(h-1)}_{\text{AR(1) part}}
\end{aligned}$$

since $\mathbb{E}\left[\epsilon_{t-j}(X_{t-h} - m)\right] = 0$ for all $h > j$—(ϵ_t) is the innovation process.

The expression of the autocovariance of order h displays the same difference or recurrence equation as in an AR(1) model—only the initial value $\gamma_X(1)$ changes !