ARMA processes

Bojan Basrak
University of Zagreb

$\overline{\underline{IV}}$ ARMA PROCESSES

$\boxed{\text{DEF}}$ weakly stationary process $(X_t)_{t\in\mathbb{Z}}$ is called ARMA(p,q) process if it satisfies ARMA difference equations

$$X_t - \varphi_1 X_{t-1} - \dots - \varphi_p X_{t-p} = Z_t + \vartheta_1 Z_{t-1} + \dots + \vartheta_q Z_{t-q}, \quad t \in \mathbb{Z}$$

for some real numbers $\varphi_1, \dots, \varphi_p, \vartheta_1, \dots, \vartheta_q$ &

$Z_t \sim WN(0, \sigma^2)$

For simplicity we typically set

$$Z_t \sim \mathrm{IID}(0,\sigma^2)$$

ARMA equation can be written using the backward shift operator B,

$$B^d X_t = X_{t-d} \qquad d \geq 0,\ t \in \mathbb{Z}$$

& polynomials

$$\varphi(z) = 1 - \varphi_1 z - \cdots - \varphi_p z^p$$

$$\vartheta(z) = 1 + \vartheta_1 z + \cdots + \vartheta_q z^q$$

EXAMPLE 1 (MA(2) process)

$$X_t = Z_t + \vartheta_1 Z_{t-1} + \vartheta_2 Z_{t-2}, \quad t \in \mathbb{Z}$$

X_t is cleary stationary,

$$\varphi(z) = 1, \quad \vartheta(z) = 1 + \vartheta_1 z + \vartheta_2 z^2$$

EXAMPLE 2 (AR(p) process)

$$X_t = \varphi_1 X_{t-1} + \cdots + \varphi_p X_{t-p} + Z_t, \quad t \in \mathbb{Z}$$

Here

$$\varphi(z) = 1 + \varphi_1 z + \cdots + \varphi_p z^p, \quad \vartheta(z) = 1.$$

But, it's not clear apriori that such a stationary process exist.

We know for $p = 1$ &

$\varphi = \varphi_1 \in (-1,1)$ stationary solution is

$$X_t = \sum_0^\infty \varphi^j Z_{t-j}$$

ChI, Lemma 1

$$\Rightarrow \; E X_t = 0, \quad \rho_X(h) = \varphi^{|h|}$$

For $|\varphi| > 1$ stationary solution is not causal

$$X_t = \sum_{j=1}^{\infty} -\varphi^{-j} Z_{t+j}$$

In both cases stationary AR(p) process exists & is linear.

In general if ARMA(p,q) process has representation

$$X_t = \sum_{0}^{\infty} \psi_j Z_{t-j}$$

for (ψ_j) s.t. $\sum_{0}^{\infty} |\psi_j| < \infty$ we say that (X_t) is causal (cf Ch. I)

Note that for two real sequences (a_n), (b_n) corresponding linear filter can be composed if $\sum |a_j|, \sum |b_j| < \infty$

Then power series

$$\xi(z) = a(z)\, b(z) = \sum \xi_j z^j$$

converges absolutely for $|z| \leq 1$

$$\Rightarrow \quad a(B)\, b(B) X_t = \xi(B) X_t$$

(from analysis)

THEOREM 1

Suppose polynomials φ & ϑ have no common zeros in $\mathbb{C}$, then

i) if $\varphi(z) \neq 0$ for $z \in \mathbb{C}$, $|z| = 1$ then there exists a linear process (X_t) satisfying corresponding ARMA equations

ii) process (X_t) in i) is causal if and only if

$$\varphi(z) \neq 0 \quad \text{for} \quad |z| \leq 1, \ z \in \mathbb{C}$$

In both i) & ii) coefficients in linear representation of (X_t) are determined by

$$\psi(z) = \sum_{j=0}^{\infty} \psi_j z^j = \frac{\vartheta(z)}{\varphi(z)}, \quad |z| \leq 1.$$

Recall if

$$\psi(z) = \sum_{j=0}^{\infty} \psi_j z^j, \quad z \in \mathbb{C}, \quad \& \text{ all } |z| \leq \varepsilon > 0$$

then the coefficients ψ_j are uniquely determined

Recall that if

$$Z_t \sim WN(0, \sigma^2)$$

then

$$X_t = \sum_{0}^{\infty} \psi_j Z_{t-j}$$

is weakly stationary (Ch I, Lemma 1) &

$$\gamma_X(h) = \sigma^2 \sum_{j=0}^{\infty} \psi_j \cdot \psi_{j+|h|}$$

$$\rho_X(h) = \frac{\sum_{0}^{\infty} \psi_j \psi_{j+|h|}}{\sum_{0}^{\infty} \psi_j^2}$$

Theorem 1 characterizes ARMA equations which have a stationary solution since it holds that

<u>THEOREM 2</u>

Suppose φ has a zero on the sphere $|z|=1$ which is not a zero of ϑ, then corresponding ARMA equations have no stationary solution

EXAMPLE 3 (ARMA (2,1) process)

Suppose

$$(1 - B + \tfrac{1}{4}B^2)X_t = (1+B)Z_t$$

$\Downarrow$

$$\varphi(z) = 1 - z + \tfrac{1}{4}z^2, \quad \vartheta(z) = 1 + z$$

$$= (1 - \tfrac{z}{2})^2$$

$\Downarrow$

By tm1 there is a stationary linear & causal solution of this ARMA equation

$$X_t = \psi(B) Z_t$$

$$\psi(z) = \frac{\vartheta(z)}{\varphi(z)} \quad \Rightarrow \quad \text{coefficients of } \psi \text{ can be calculated}$$

e.g. by comparing coefficients of

$$\psi(z)\varphi(z) = \vartheta(z)$$

$$\sum_0^\infty \psi_j z^j \left(1 - z + \frac{1}{4} z^2\right) = 1 + z$$

$\Downarrow$

$j = 0$: $\quad \psi_0 \cdot 1 = 1$

$j = 1$: $\quad \psi_1 - \psi_0 = 1 \quad \Rightarrow \psi_1 = 2$

$j \geq 2$: $\quad \psi_j - \psi_{j-1} + \frac{1}{4}\psi_{j-2} = 0 \quad \leftarrow$ difference equation

General solution can be found as

$$\psi_n = (\alpha + n\beta)\, 2^{-n} \qquad n \geq 0$$

Using: $\psi_0 = 1, \; \psi_1 = 2$

$$\Rightarrow \quad \alpha = 1, \; \beta = 3 \quad \text{i.e.}$$

$$\psi_n = (1 + 3n) \cdot 2^{-n} \qquad n = 0, 1, 2, \ldots$$

EXE1 > Show that ARMA(2,1) equations

$$\left(1-\tfrac{1}{2}B\right)X_t = \left(1+\tfrac{1}{2}B\right)\left(1+0.7B\right)Z_t$$

have causal weakly stationary solution.

Find coefficients in the linear representation

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}$$

$\boxed{\text{DEF}}$ ARMA(p,q) process (X_t) is <u>invertible</u> if for some sequence $(\pi_j)_{j\in\mathbb{N}_0}$, $\sum|\pi_j|<\infty$

$$Z_t = \sum_{j=0}^{\infty} \pi_j X_{t-j}$$

THEOREM 3

Suppose (X_t) is weakly stationary ARMA process, s.t. φ & ϑ have no common zeros. Then (X_t) is invertible

$\iff$ $\vartheta(z) \neq 0$ for $|z| \leq 1$, $z \in \mathbb{C}$

EXAMPLE 4

$X_t = Z_t + \frac{1}{2} Z_{t-2}$ is invertible

$X_t = Z_t - 1.01\, Z_{t-1}$ is not invertible

FUNCTIONS γ, ρ & α FOR ARMA PROCESSES

Assume (X_t) is causal ARMA process

s.t. $\varphi(B) X_t = \vartheta(B) Z_t \quad t \in \mathbb{Z}$

$\Rightarrow \quad X_t = \sum_0^\infty \psi_j Z_{t-j} \quad , \quad \sum_0^\infty |\psi_j| < \infty$

L1, Ch1

$\Rightarrow \quad E X_t = 0$

& $\quad \gamma_X(h) = E(X_t \cdot X_{t+h}) = \sigma^2 \sum_0^\infty \psi_j \cdot \psi_{j+|h|}$

EXAMPLE 5

a) MA(q) process : $X_t = Z_t + \vartheta_1 Z_{t-1} + \dots + \vartheta_q Z_{t-q}$

we showed

$$\gamma_X(h) = \begin{cases} \sigma^2 \sum_0^{q-|h|} \vartheta_j \vartheta_{j+|h|} , & |h| < q \\ 0 , & \text{otherwise} \end{cases}$$

b) ARMA (1,1) process

$$X_t - \varphi X_{t-1} = Z_t + \vartheta Z_{t-1}$$

Suppose $|\varphi| < 1$ (p 12-13)

b)

$$\gamma_X(0) = \sigma^2 \left(1 + \frac{(\vartheta+\varphi)^2}{1-\varphi^2}\right)$$

$$\gamma_X(1) = \sigma^2 \left(\vartheta + \varphi + \frac{(\vartheta+\varphi)^2 \cdot \varphi}{1-\varphi^2}\right), \quad \gamma_X(h) = \varphi^{h-1} \gamma_X(1)$$
$h \geq 2$

c) AR(2) process

$$X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} = Z_t$$

In c)

$$\gamma(h) = \frac{\sigma^2 \cdot r^4}{r^2-1} \frac{\sin(h\vartheta + \psi)}{(r^4 - 2r^2\cos 2\vartheta + 1)\sin\vartheta} \cdot r^{-h}$$

if ϕ has zeros $re^{\pm i\vartheta}$, $\psi = \arctan\left(\frac{r^2+1}{r^2-1}\tan\vartheta\right)$

To define partial autocorrelation function $\alpha(h)$ we have used

$$\alpha(0) = 1, \quad \alpha(h) = \varphi_h^{(h)} \qquad h \geq 1$$

where

$$\Gamma_h \begin{pmatrix} \varphi_1^{(h)} \\ \vdots \\ \varphi_h^{(h)} \end{pmatrix} = \begin{pmatrix} \gamma(1) \\ \vdots \\ \gamma(h) \end{pmatrix}$$

So if Γ_h is regular matrix $\varphi_h^{(h)}$ can be directly calculated.

We have seen that for AR(1) process

$$X_t = \varphi X_{t-1} + Z_t, \quad |\varphi| < 1$$

$$\alpha_X(h) = \begin{cases} \varphi, & h = 1 \\ 0 & h \geq 2 \end{cases}$$

EXAMPLE 6 (AR(p) process)

$$\Rightarrow \quad \alpha_X(p) = \varphi_p$$

$$\alpha_X(h) = 0 \quad \text{for } h > p$$

EXAMPLE 7 (MA(1) process)

$$X_t = Z_t + \vartheta Z_{t-1}$$

One can show

$$\alpha_X(h) = \varphi_h^{(h)} = \cdots = \frac{-(-\vartheta)^h}{(1+\vartheta^2+\cdots+\vartheta^{2h})}$$

which decays exponentially

$\boxed{\text{DUALITY}}$ OF ρ & α

for AR & MA processes !

By plotting sample versions of functions α & ρ i.e. $\hat{\alpha}$ & $\hat{\rho}$ we can get an idea about suitability of AR(p) & MA(q) processes for modeling of a given data set.

ESTIMATION FOR ARMA PROCESSES

Suppose that for given observations

$$X_1, X_2, \ldots, X_n$$

we want to estimate a certain ARMA model. To select the model we can use graphs of $\hat{\rho}$ & $\hat{\alpha}$ or so called <u>information</u> <u>criteria</u>

(to be mentioned later)

YULE - WALKER ESTIMATORS

Suppose we have decided we want to fit AR(p) model

$$X_t - \varphi_1 X_{t-1} - \cdots - \varphi_p X_{t-p} = Z_t \qquad (*)$$

How do we estimate $\varphi_1, \ldots, \varphi_p$?

Idea multiply $(*)$ with $X_t, \ldots, X_{t-p}$

& take expectation to obtain :

e.g. $EX_t^2 - \varphi_1 EX_tX_{t-1} - \cdots - \varphi_p EX_tX_{t-p} = EX_tZ_t$

show $EX_tZ_t = \sigma^2$ (Le 1, Ch 1) to get

$$\gamma(0) - \varphi_1\gamma(1) - \cdots - \varphi_p\gamma(p) = \sigma^2$$

$$\gamma(1) - \varphi_1\gamma(0) - \cdots - \varphi_p\gamma(p-1) = 0$$

$$\vdots$$

$$\gamma(p) - \varphi_1\gamma(p-1) - \cdots - \varphi_p\gamma(0) = 0$$

That is

$$\Gamma_p \begin{pmatrix} \varphi_1 \\ \vdots \\ \varphi_p \end{pmatrix} = \begin{pmatrix} \gamma(1) \\ \vdots \\ \gamma(p) \end{pmatrix}$$

&

$$\gamma(0) - (\varphi_1 \ldots \varphi_p)\begin{pmatrix} \gamma(1) \\ \vdots \\ \gamma(p) \end{pmatrix} = \sigma^2$$

These equations motivate so called <u>Yule-Walker equations</u>

$$\hat{\Gamma}_p \begin{pmatrix} \hat{\phi}_1 \\ \vdots \\ \hat{\phi}_p \end{pmatrix} = \begin{pmatrix} \hat{\gamma}(1) \\ \vdots \\ \hat{\gamma}(p) \end{pmatrix}$$

$$\hat{\sigma}^2 = \hat{\gamma}(0) - (\hat{\phi}_1 \cdots \hat{\phi}_p) \begin{pmatrix} \hat{\gamma}(1) \\ \vdots \\ \hat{\gamma}(p) \end{pmatrix}$$

One can show that if

$$\hat{\gamma}(0) > 0 \quad \Rightarrow \quad \hat{\Gamma}_p \text{ is nonsingular}$$

& Y-W equations have unique solution

<u>REMARK</u> In practice these equations are solved using some efficient algorithms (eg. Durbin – Levinson / Innovations).

Our cell phones solve one such equation every 10 ms during the call (cf. Dutoit)

EXAMPLE 8 (Y-W estimators in AR(1) case)

$$X_t = \varphi X_{t-1} + Z_t \quad , \quad |\varphi| < 1$$

$$\hat{\sigma}^2 = \hat{\gamma}(0) - \hat{\varphi}\hat{\gamma}(1) \quad \& \quad 0 = \hat{\gamma}(1) - \hat{\varphi}\hat{\gamma}(0)$$

$$\Rightarrow \quad \hat{\varphi} = \frac{\hat{\gamma}(1)}{\hat{\gamma}(0)} = \hat{\rho}(1) \quad \& \quad \hat{\sigma}^2 = \hat{\gamma}(0)\left(1 - \hat{\rho}^2(1)\right)$$

EXE 2) Show that estimators of σ^2 & φ are consistent in this case if $(Z_t) \sim \mathrm{IID}(0, \sigma^2)$.

<u>REMARK</u> By construction Y-W estimators belong to class of moment estimators

One can show (see Brockwell & Davis) that for AR(p) process with iid noise (Z_t)

$$\sqrt{n}\left[\begin{pmatrix}\hat{\varphi}_1\\ \vdots\\ \hat{\varphi}_p\end{pmatrix}-\begin{pmatrix}\varphi_1\\ \vdots\\ \varphi_p\end{pmatrix}\right]\xrightarrow{d} N(0,\sigma^2\Gamma_p^{-1})$$

which can be used to obtain confidence intervals for φ_i; moreover

$\hat{\sigma}^2 \xrightarrow{P} \sigma^2$.

GAUSSIAN MAXIMUM LIKELIHOOD

Yule Walker is suited for AR(p) model, but there are methods which can estimate parameters of general ARMA(p,q) process.

One of them is derived under assumptions:

(X_t) is Gaussian, invertible & causal

$(Z_t) \sim \text{IID}, N(0,\sigma^2)$

For a given vectors of observations

$$\vec{X}_n = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$$

& parameters $\vec{\varphi} = \begin{pmatrix} \varphi_1 \\ \vdots \\ \varphi_p \end{pmatrix}$, $\vec{\vartheta} = \begin{pmatrix} \vartheta_1 \\ \vdots \\ \vartheta_q \end{pmatrix}$

we maximize likelihood w.r.t. $(\vec{\varphi}, \vec{\vartheta}, \sigma^2) = (\beta, \sigma^2)$

$$L(\beta, \sigma^2) = L(\Gamma_n(\beta, \sigma^2)) =$$

$$(2\pi)^{-n/2} \cdot (\det \Gamma_n)^{-1/2} \cdot \exp\left\{ -\frac{1}{2} \vec{X}_n^{\tau} \Gamma_n^{-1} \vec{X}_n \right\}$$

In practice MLE is found by numerical optimization, such an estmotor $\hat{\beta}_n$ of β is called Gaussian MLE.

It works even when data are not Gaussian (cf. Brockwell & Davis) reasonably well.

One can show that for causal, invertible, Gaussian time series (X_t) with iid noise

$$\sqrt{n}\left[\begin{pmatrix}\hat{\varphi}_1\\ \vdots\\ \hat{\vartheta}_q\end{pmatrix}-\begin{pmatrix}\varphi_1\\ \vdots\\ \vartheta_q\end{pmatrix}\right]=\sqrt{n}\left(\hat{\beta}_n-\beta_0\right)\xrightarrow{d}N\left(0,W(\beta_0)\right)$$

MODEL SELECTION

There are several information criteria that can be used to select the "correct order" of the ARMA model. They all attempt to penalize overfitting.

Thus if RSS_k denotes the residual sum of squares under the model with k parameters, we define

$$\hat{\sigma}^2 = \frac{RSS_k}{n}$$

The earliest methods of order selection are due to Akaike (1969, '73, '74). They chose the model which minimizes apropriate information criteria

1. Akaike information criterion (AIC)

$$AIC = \ln \hat{\sigma}_k^2 + \underbrace{\frac{n+2k}{n}}_{\text{penalty}}$$

Other penalties are possible

2. Corrected AIC (AICC)

$$AICC = \ln \hat{\sigma}_k^2 + \frac{n+k}{n-k-2}$$

3. Bayesian (Schwarz's) inf. criterion (BIC)

$$BIC = \ln \hat{\sigma}_k^2 + \frac{k \cdot \ln n}{n}$$

see (Shumway & Stoffer or Brockwell & Dains)