Capital University of Economics and Business

ISEM

Financial Econometrics

ARMA(p, q) Model

Instructor:
CHEUNG Ying Lun

Term:
Fall Semester, 2019

November 4, 2019

ARMA(p, q) process

We say X_t is an ARMA(p, q) process if it takes the form

$$X_t = a_0 + a_1 X_{t-1} + \cdots + a_p X_{t-p} + \varepsilon_t - b_1 \varepsilon_{t-1} - \cdots - b_q \varepsilon_{t-q}, \qquad \varepsilon_t \overset{\text{iid}}{\sim} \left(0, \sigma^2\right). \tag{1}$$

Moments of a weakly stationary ARMA(1,1) process

Suppose that X_t is weakly stationary ARAM(1,1) process

$$X_t = a_0 + a_1 X_{t-1} + \varepsilon_t - b_1 \varepsilon_{t-1}, \qquad \varepsilon_t \overset{\text{iid}}{\sim} \left(0, \sigma^2\right). \tag{2}$$

Then the mean of X_t can be obtained simply by taking the expectation of Eq.(2):

$$\mu_X = a_0 + a_1 \mu_X \implies \mu_X = \frac{a_0}{1 - a_1}.$$

Next, taking the variance of Eq.(2),

$$\begin{aligned} \operatorname{var}(X_t) &= a_1^2 \operatorname{var}(X_{t-1}) + \operatorname{var}(\varepsilon_t) + b_1^2 \operatorname{var}(\varepsilon_{t-1}) + 2 \operatorname{cov}(a_1 X_{t-1}, -b_1 \varepsilon_{t-1}) \\ &= a_1^2 \operatorname{var}(X_t) + \sigma^2 + b_1^2 \sigma^2 - 2 a_1 b_1 \sigma^2 \\ &= \frac{(1 - 2a_1 b_1 + b_1^2)\sigma^2}{1 - a_1^2} \end{aligned}$$

Subtracting the mean from both sides of Eq.(1),

$$X_t - \mu_X = a_0 + a_1 (X_{t-1} - \mu_X) + \varepsilon_t - b_1 \varepsilon_{t-1} \tag{3}$$

Now multiply $X_{t-h} - \mu_X$ to the both sides of Eq.(3) and take expectation, we obtain the Yule-Walker equations. For $h = 1$,

$$\gamma_X(1) = a_1 \gamma_X(0) - b_1 \mathbb{E}\left[(X_{t-1} - \mu_X)\varepsilon_{t-1}\right] = a_1 \gamma_X(0) - b_1 \sigma^2.$$

For $h > 1$,

$$\gamma_X(h) = a_1 \gamma_X(h-1)$$

which is the same as in an AR(1) process.

Estimation

In an AR(p) model, we can estimate the coefficients by linear regression. However, in an ARMA(p, q) model, linear regression will result in a biased estimator since the regressors X_{t-j} are correlated with the error terms ε_{t-j}. Therefore, here we estimate the model by the *maximum-likelihood estimation*.

MLE

Consider a set of random variables $\{X_1, \ldots, X_T\}$ with joint density function $f_{X_1,\ldots,X_T}(x_1, \ldots, x_T; \boldsymbol{\theta})$, where $\boldsymbol{\theta}$ collects all the parameters in the density function. Suppose now that we observe the realization of the set of random variables $\{x_1, \ldots, x_T\}$. The idea of MLE is as follow:

- We want to choose a set of parameters $\hat{\boldsymbol{\theta}}$ such that the probability of observing $\{x_1, \ldots, x_T\}$ is maximized.

Therefore, we set the likelihood function as

$$\mathcal{L}(\boldsymbol{\theta}; x_1, \ldots, x_T) = f_{X_1,\ldots,X_T}(x_1, \ldots, x_T; \boldsymbol{\theta}).$$

The maximum-likelihood estimator is then

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} \mathcal{L}(\boldsymbol{\theta}; x_1, \ldots, x_T).$$

In particular, assume $X_t \overset{\text{iid}}{\sim} \mathcal{N}\left(\mu, \sigma^2\right)$, then the density function of X_t is

$$f_{X_t}(x_t; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_t - \mu)^2}{2\sigma^2}\right)$$

The joint density is given by

$$\begin{aligned} f_{X_1,\ldots,X_T}(x_1, \ldots, x_T; \mu, \sigma^2) &= \prod_{t=1}^{T} f_{X_t}(x_t; \mu, \sigma^2) \\ &= (2\pi\sigma^2)^{-T/2} \exp\left(-\sum_{t=1}^{T} \frac{(x_t - \mu)^2}{2\sigma^2}\right) \end{aligned}$$

In this case, $\boldsymbol{\theta} = (\mu, \sigma^2)$ contains only two parameters. Let $\mathcal{L}(\mu, \sigma^2; x_1, \ldots, x_T) = f_{X_1,\ldots,X_T}(x_1, \ldots, x_T; \mu, \sigma$ the maximum-likelihood estimator of μ and σ^2 are

$$(\hat{\mu}, \hat{\sigma}^2) = \arg\max_{\mu,\sigma^2} \mathcal{L}(\mu, \sigma^2; x_1, \ldots, x_T).$$

MLE of an ARMA(p, q) process

Consider the ARAM(p, q) process

$$X_t = a_0 + a_1 X_{t-1} + \cdots + a_p X_{t-p} + \varepsilon_t - b_1 \varepsilon_{t-1} - \cdots - b_q \varepsilon_{t-q}, \qquad \varepsilon_t \overset{\text{iid}}{\sim} \mathcal{N}\left(0, \sigma^2\right).$$

Then the joint density of ε_t is given by

$$f_{\varepsilon_1,\varepsilon_T}(\varepsilon_1, \ldots, \varepsilon_T; \boldsymbol{\theta}) = (2\pi\sigma^2)^{-T/2} \exp\left(-\sum_{t=1}^{T} \frac{\varepsilon_t^2}{2\sigma^2}\right)$$

However, we do not observe ε_t. Rearranging terms in the model,

$$\varepsilon_t = X_t - a_0 - a_1 X_{t-1} - \cdots - a_p X_{t-p} + b_1 \varepsilon_{t-1} + \cdots + b_q \varepsilon_{t-q}.$$

Since we only observe $\{X_1, \ldots, X_T\}$, we start at $t = p + 1$ and assume that $\varepsilon_{p-j} = 0$ for all $j \geq 0$.

$$\begin{aligned}
\varepsilon_{p+1} &= X_{p+1} - a_0 - a_1 X_p - \cdots - a_p X_1 \\
\varepsilon_{p+2} &= X_{p+2} - a_0 - a_1 X_{p+1} - \cdots - a_p X_2 + b_1 \varepsilon_{p+1} \\
&\vdots \\
\varepsilon_{p+1+q} &= X_{p+1+q} - a_0 - a_1 X_{p+q} - \cdots - a_p X_{q+1} + b_1 \varepsilon_{p+q} + \cdots + b_q \varepsilon_{p+1} \\
&\vdots \\
\varepsilon_T &= X_T - a_0 - a_1 X_{T-1} - \cdots - a_p X_{T-p} + b_1 \varepsilon_{T-1} + \cdots + b_q \varepsilon_{T-q}
\end{aligned}$$

Therefore, given the parameters $\{a_0, \ldots, a_p\}$ and $\{b_1, \ldots, b_q\}$, together with the observations $\{X_1, \ldots, X_T\}$, we can obtain the values of $\{\varepsilon_{p+1}, \ldots, \varepsilon_T\}$. We can then substitute these values to the joint density function. Finally, we can find the MLE of the parameters as

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} \mathcal{L}(\mu, \sigma^2; x_1, \ldots, x_T)$$

where $\mathcal{L}(\mu, \sigma^2; x_1, \ldots, x_T) = f_{\varepsilon_{p+1}, \varepsilon_T}(\varepsilon_{p+1}, \ldots, \varepsilon_T; \boldsymbol{\theta})$ and $\boldsymbol{\theta} = (a_0, \ldots, a_p, b_1, \ldots, b_q, \sigma^2)$.

Representation of an ARMA process

Using the lag polynomial, the ARMA(p, q) process can be written as

$$(1 - a_1 L - \cdots - a_p L^p)(X_t - \mu_X) = (1 - b_1 L - \cdots - b_q L^q)\varepsilon_t.$$

Factoring the polynomials, we have

$$(1 - \phi_1 L) \ldots (1 - \phi_p L)(X_t - \mu_X) = (1 - \theta_1 L) \ldots (1 - \theta_q L)\varepsilon_t$$

where we assume that $\phi_i \neq \theta_j$ for any i, j. If $|\phi_j| < 1$ for all j, then

$$\begin{aligned}
X_t &= \mu_X + (1 - \phi_1 L)^{-1} \ldots (1 - \phi_p L)^{-1} (1 - \theta_1 L) \ldots (1 - \theta_q L)\varepsilon_t \\
&= \mu_X + \left(\sum_{j=0}^{\infty} \phi_1^j L^j \right) \ldots \left(\sum_{j=0}^{\infty} \phi_p^j L^j \right) (1 - \theta_1 L) \ldots (1 - \theta_q L)\varepsilon_t
\end{aligned}$$

This is the MA(∞) representation of an ARMA(p, q) process. In this case, we say X_t is *casual*. X_t is also weakly stationary. Similarly, if $|\theta_j| < 1$ for all j, then

$$\begin{aligned}
\varepsilon_t &= (1 - \theta_1 L)^{-1} \ldots (1 - \theta_q L)^{-1} (1 - \phi_1 L) \ldots (1 - \phi_p L)(X_t - \mu_X) \\
&= \left(\sum_{j=0}^{\infty} \theta_1^j L^j \right) \ldots \left(\sum_{j=0}^{\infty} \theta_q^j L^j \right) (1 - \phi_1 L) \ldots (1 - \phi_p L)(X_t - \mu_X)
\end{aligned}$$

This is the AR(∞) representation of an ARMA(p, q) process. In this case, we say X_t is *invertible*.