Financial Econometrics

Lecture 3

Fu Ouyang

March 24, 2018

Outline

Autoregressive Processes

Stationary ARMA Processes

Autoregressive Processes: Definition

For a time series $\{X_t\}$, it is more intuitive to predict X_t using its history,

$$X_t = c + b_1 X_{t-1} + \cdots b_p X_{t-p} + \epsilon_t$$

where $\epsilon_t \sim WN(0, \sigma^2)$ and $c, b_1, \ldots, b_p$ are unknown parameters. We refer to this model as an *autoregressive* (AR) *process* of order p, $X_t \sim \text{AR}(p)$.

Suppose $X_t \sim \text{AR}(p)$ is stationary. Then,

$$\mu \equiv E(X_t) = c + \mu(b_1 + \cdots + b_p) \Rightarrow \mu = \frac{c}{1 - (b_1 + \cdots + b_p)}$$

and so

$$X_t - \mu = b_1(X_{t-1} - \mu) + \cdots + b_p(X_{t-p} - \mu) + \epsilon_t$$

In what follows, we assume X_t is "centralized", i.e., $E(X_t) = 0$ and $c = 0$.

AR(1) Processes

Example: AR(1) Model

$$X_t = bX_{t-1} + \epsilon_t$$

Assuming that $Cov(\epsilon_t, X_{t-k}) = 0, \forall k \geq 1$, $E(X_t^2) = b^2 E(X_{t-1}^2) + \sigma^2$.

"$\Rightarrow$": For weakly stationary AR(1), $E(X_t^2) = E(X_{t-1}^2)$, which then implies $|b| < 1$ and $E(X_t^2) = \sigma^2/(1-b^2)$. Therefore, $|b| < 1$ is a necessary condition for the stationarity of AR(1).

"$\Leftarrow$": By recursive substitution,

$$\begin{aligned} X_t &= bX_{t-1} + \epsilon_t \\ &= b^2 X_{t-2} + \epsilon_t + b\epsilon_{t-1} \\ &= \epsilon_t + b\epsilon_{t-1} + \cdots + b^k \epsilon_{t-k} + b^{k+1} X_{t-k-1} \end{aligned}$$

If $|b| < 1$, $X_t = \sum_{j=0}^{\infty} b^j \epsilon_{t-j}$ (in a mean squared error sense). To see this...

AR(1) Processes

Example: AR(1) Model

$$
\begin{aligned}
E[(X_t - \sum_{j=0}^{\infty} b^j \epsilon_{t-j})^2] &= \lim_{k\to\infty} E[(X_t - \sum_{j=0}^{k} b^j \epsilon_{t-j})^2] \\
&= \lim_{k\to\infty} |b|^{2(k+1)} E(X_{t-k-1}^2)
\end{aligned}
$$

Hence, if $|b| < 1$, $|b|^{2(k+1)} \to 0$ as $k \to \infty$, and $E[(X_t - \sum_{j=0}^{\infty} b^j \epsilon_{t-j})^2] = 0$.

AR(1) process is effectively a MA(∞) process and so weakly stationary (?).

To sum up, an AR(1) process is weakly stationary if and only if $|b| < 1$.

Backshift Operator

The recursive substitution can be compactly represented by the *backshift operator* B, i.e., for $k = \pm 1, \pm 2, \ldots$

$$B^k X_t = X_{t-k}$$

Then an AR(1) model can be written as $(1 - bB)X_t = \epsilon_t$. Recall the infinite series expansion of $(1 - bx)^{-1}$, we have

$$(1 - bx)^{-1} = \sum_{j=0}^{\infty} b^j x^j$$

as $(1 - bx)(1 + bx + b^2x^2 + \cdots) = 1$. An analogous definition of $(1 - bB)^{-1}$ gives the MA(∞) representation of the AR(1) process

$$X_t = (1 - bB)^{-1}\epsilon_t = \sum_{j=0}^{\infty} b^j \epsilon_{t-j}$$

Backshift Operator

The backshift operator B is useful in handling general AR(p) process:

$$X_t = b_1 X_{t-1} + \cdots + b_p X_{t-p} + \epsilon_t$$

which can be written as $b(B)X_t = \epsilon_t$ with $b(x) \equiv 1 - b_1 x - \cdots - b_p x^p$.

Let $\alpha_1^{-1}, ..., \alpha_p^{-1}$ be roots of $b(x) = 0$, i.e.,

$$b(x) = \prod_{j=1}^{p} (1 - \alpha_j x)$$

There is a sequence $\{a_k\}$ with each a_k determined by $\alpha_1, ..., \alpha_p$ such that

$$b(x)^{-1} = \prod_{j=1}^{p} (1 - \alpha_j x)^{-1} = \prod_{j=1}^{p} \left(\sum_{l=0}^{\infty} \alpha_j^l x^l \right) = 1 + \sum_{k=1}^{\infty} a_k x^k$$

The MA(∞) Representation of AR(p) Processes

From the derivation above, we know for $k = 1, 2, ...,$

$$|a_k| = O\left(\max_{1\leq j\leq p} |\alpha_j|^k\right)$$

If $|\alpha_j| < 1$ for all $1 \leq j \leq p$, then the AR(p) process can be written as

$$X_t = \epsilon_t + \sum_{k=1}^{\infty} a_k \epsilon_{t-k}$$

with $\sum_{k=1}^{\infty} a_k^2 < \infty$ (since $p < \infty$), i.e., $X_t \sim$ MA(∞) and stationary.

The MA(∞) representation for a stationary AR(p) process $\{X_t\}$ indicates that it is a *causal process*, i.e., X_t only depends on $\{\epsilon_t, \epsilon_{t-1}, ...\}$, and is uncorrelated with any future innovations.

The MA(∞) Representation of AR(p) Processes

Recall that for MA(∞) process

$$\gamma(0) = \sigma^2 \sum_{j=0}^{\infty} a_j^2, \gamma(k) = \sigma^2 \sum_{j=0}^{\infty} a_j a_{j+|k|}$$

and so for the AR(p) process, $\rho(k) = O(\max_{1 \le j \le p} |\alpha_j|^k) \to 0$ as $k \to \infty$, which means it only suitable for modeling *short memory* data.

AR(p) Model

1. An AR(p) process is stationary if the p roots of the characteristic equation $1 - b_1 x - \cdots - b_p x^p = 0$ are outside the unit cycle.
2. The ACF of a stationary AR(p) process decays at an exponential rate, i.e., $\rho(k) = O(\alpha^k)$ for some $\alpha \in (0, 1)$.

Yule-Walker Equation

Using the MA(∞) representation to compute $\gamma(k)$ and $\rho(k)$ of the AR(p) model below is cumbersome

$$X_t = b_1 X_{t-1} + \cdots + b_p X_{t-p} + \epsilon_t$$

An alternatively way is to use the *Yule-Walker equation*: For $k \geq 1$, we have

$$\gamma(k) = b_1 \gamma(k-1) + \cdots + b_p \gamma(k-p)$$

which by $\gamma(-k) = \gamma(k)$ yields

$$\begin{pmatrix} \gamma(1) \\ \gamma(2) \\ \vdots \\ \gamma(p) \end{pmatrix} = \begin{pmatrix} \gamma(0) & \gamma(1) & \cdots & \gamma(p-1) \\ \gamma(1) & \gamma(0) & \cdots & \gamma(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma(p-1) & \gamma(p-2) & \cdots & \gamma(0) \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_p \end{pmatrix}$$

ACVF and ACF

For $k = 0$, we have

$$\gamma(0) = b_1\gamma(1) + \cdots + b_p\gamma(p) + \sigma^2$$

Putting all equations together, we have

$$\begin{pmatrix} \gamma(0) - \sigma^2 \\ \gamma(1) \\ \gamma(2) \\ \vdots \\ \gamma(p) \end{pmatrix} = \begin{pmatrix} \gamma(1) & \gamma(2) & \cdots & \gamma(p) \\ \gamma(0) & \gamma(1) & \cdots & \gamma(p-1) \\ \gamma(1) & \gamma(0) & \cdots & \gamma(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma(p-1) & \gamma(p-2) & \cdots & \gamma(0) \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_p \end{pmatrix}$$

i.e., $p + 1$ linear equations to solve $p + 1$ unknowns $\gamma(0), \gamma(1), \ldots, \gamma(p)$.

$\forall k > p$, $\gamma(k)$ can be obtained recursively using the Yule-Walker equation.

Partial Autocorrelation Function

The *partial autocorrelation function* (PACF) at lag k, denoted by $\pi(k)$, is the conditional correlation between X_1 and X_{1+k} given all the intermediate variables $X_2, ..., X_k$. More concretely, let

$$(b_{k1}, ..., b_{kk}) \equiv \arg \min_{\beta_1, ..., \beta_k} E\left[(X_{1+k} - \beta_1 X_k - \cdots - \beta_k X_1)^2\right]$$

and then $\pi(k) \equiv b_{kk}$.

For a stationary AR(p) process, the PACF cuts off at p, i.e., $\pi(k) = 0$ for all $k > p$. To see this, recall that

$$E(X_{1+k} | X_1, ..., X_k) = \arg \min_{g \in \mathcal{G}} E[(X_{1+k} - g(X_1, ..., X_k))^2]$$

and by definition of AR(p),

Partial Autocorrelation Function

$$E(X_{1+k}|X_1, \dots, X_k) = b_1 X_k + \cdots + b_p X_{k-p+1} + 0 \cdot X_{k-p} + \cdots + 0 \cdot X_1$$

Hence $b_{k1} = b_1, \dots, b_{kp} = b_p, b_{k,p+1} = 0, \dots, b_{kk} = 0$.

The sample PACF at lag k, denoted by $\widehat{\pi}(k)$, is the sample analogue of $\pi(k)$, i.e.,

$$(\widehat{b}_{k1}, \dots, \widehat{b}_{kk}) = \arg \min_{\beta_1, \dots, \beta_k} \sum_{t=1+k}^{T} (X_t - \beta_1 X_{t-1} - \cdots - \beta_k X_{t-k})^2$$

and $\widehat{\pi}(k) \equiv \widehat{b}_{kk}$, one can obtain $\widehat{\pi}(k)$ by running a least square estimation.

The sample PACF of a stationary AR(p) process does *not* necessarily cuts off at p. The sample PACF will be used for model selection (next lecture).

Outline

ARMA Processes: Definition

A general *autoregressive and moving average* (ARMA) *model* with the order (p, q) is the combination of a AR(p) and a MA(q) process:

$$X_t = b_1 X_{t-1} + \cdots b_p X_{t-p} + \epsilon_t + a_1 \epsilon_{t-1} + \cdots + a_q \epsilon_{t-q}$$

where $\epsilon_t \sim WN(0, \sigma^2)$ and $(b_1, ..., b_p, a_1, ..., a_q)$ are unknown parameters.

Let

$$\begin{aligned} a(x) &= 1 + a_1 x + \cdots + a_q x^q \\ b(x) &= 1 - b_1 x - \cdots - b_p x^p \end{aligned}$$

Then an ARMA(p,q) model can be compactly represented as

$$b(B) X_t = a(B) \epsilon_t$$

provided that $a(x) = 0$ and $b(x) = 0$ do not have common roots.

ARMA Processes: Properties

Stationarity of ARMA(p,q)

When the p roots of $b(x) = 0$ are all outside of the unit cycle, the ARMA(p,q) process $\{X_t\}$ is stationary and has an MA(∞) representation

$$X_t = b(B)^{-1}a(B)\epsilon_t \sim \mathrm{MA}(\infty)$$

Similar to stationary AR(p) processes, a stationary ARMA(p,q) process

1. $\{X_t\}$ is a causal process, i.e., X_t only depends on $\{\epsilon_t, \epsilon_{t-1}, ...\}$.
2. $\{X_t\}$ has short memory, i.e., $\rho(k) = O(\theta^k)$ as $k \to \infty$ for some $|\theta| < 1$.

Yule-Walker Equation

For all $k > q$ (why?),

$$\gamma(k) = b_1\gamma(k-1) + \cdots + b_p\gamma(k-p)$$

Invertibility: Definition

If an MA(q) process

$$X_t = a(B)\epsilon_t$$

where $a(x) = 1 + a_1 x + \cdots + a_q x^q$ can be written as an AR(∞) process, then it is *invertible*, i.e., the innovations $\epsilon_t, \epsilon_{t-1}, \ldots$ can be recovered from the observed $X_t, X_{t-1}, \ldots$.

With similar derivation as for the MA(∞) representation of an AR(p) process, we can show that $X_t \sim$ AR(∞) (again in a mean-squared error sense) if the q roots of $a(x)$ are outside the unit cycle.

In practice, the above invertibility condition is imposed to the MA(q) process for the identification of $(a_1, \ldots, a_q)$ in terms of its ACF.

The following example shows the necessity of doing so.

Invertibility: Example

Consider two MA(1) models with $|a| < 1$,

$$X_t = \epsilon_t + a\epsilon_{t-1}, \epsilon_t \sim WN(0, \sigma^2)$$
$$Y_t = e_t + a^{-1}e_{t-1}, e_t \sim WN(0, a^2\sigma^2)$$

It is easy to show that $\{X_t\}$ and $\{Y_t\}$ share the same ACF, and so they are *not* distinguishable in terms of the ACF. However, $\{Y_t\}$ is not invertible. In fact, by recursive substitution, we have

$$\begin{aligned} Y_t &= a^{-1}e_{t-1} - \sum_{j=1}^{k} (-a)^j \, Y_{t+j} + (-a)^k \, e_{t+k} \\ &\overset{m.s.}{\rightarrow} a^{-1}e_{t-1} - \sum_{j=1}^{\infty} (-a)^j \, Y_{t+j} \end{aligned}$$

as $k \to \infty$, so $\{Y_t\}$ is "invertible in the future", not useful for forecasting.