Arima Model - 5

강의 : 김성범 교수님

정리 : 김선중

December 13, 2022

이것은 ARIMA 모델 개요 - Part 5 강의에 대한 노트이다. 이전과는 다르게 영어로 작성해보았다.

Contents

아리마 모델 (ARIMA Model) – Part 5

ARMA

Backward shift operator ← function

$B \cdot X_t = X_{t-1}$

$B^2 \cdot X_t = B \cdot (B \cdot X_t) = B \cdot X_{t-1} = X_{t-2}$

$B^3 \cdot X_t = X_{t-3} \quad \cdots \quad B^m \cdot X_t = X_{t-m}$

AR(1) $\quad X_t = \phi X_{t-1} + a_t$

$$= \phi B \cdot X_t + a_t$$

$$(1-\phi B) X_t = a_t$$

$$X_t = \frac{a_t}{1-\phi B}$$

$$= (1 + \phi B + \phi^2 B^2 + \phi^3 B^3 + \cdots) a_t$$

$$= a_t + \phi B a_t + \phi^2 B^2 a_t + \phi^3 B^3 a_t + \cdots$$

$$= a_t + \phi a_{t-1} + \phi^2 a_{t-2} + \phi^3 a_{t-3} + \cdots$$

AR(2) $\quad X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + a_t$

$$= \phi_1 B X_t + \phi_2 B^2 X_t + a_t$$

$$(1 - \phi_1 B - \phi_2 B^2) X_t = a_t$$

$$X_t = \frac{1}{1-\phi_1 B - \phi_2 B^2} \cdot a_t$$

$$= \frac{1}{(1-\alpha_1 B)(1-\alpha_2 B)} \cdot a_t$$

$$= \frac{1}{1-\alpha_1 B} \cdot \frac{1}{1-\alpha_2 B} \cdot a_t$$

$$= (1+\alpha_1 B + \alpha_1^2 B^2 + \cdots)(1 + \alpha_2 B + \alpha_2^2 B^2 + \cdots) a_t$$

$$= a_t + (\alpha_1 + \alpha_2) B a_t + (\alpha_1^2 + \alpha_2^2 + \alpha_1\alpha_2) B^2 a_t + \cdots$$

$$= a_t + (\alpha_1+\alpha_2) a_{t-1} + (\alpha_1^2 + \alpha_2^2 + \alpha_1\alpha_2) a_{t-2} + \cdots$$

* Forward shift operator

$F \cdot X_t = X_{t+1}$

$F^2 \cdot X_t = X_{t+2} \quad \cdots \quad F^K \cdot X_t = X_{t+K}$

ARMA(1,1)

$X_t = \phi X_{t-1} + a_t + \theta a_{t-1} \longrightarrow X_t - \phi B X_t = a_t + \theta a_{t-1}$

$(1-\phi B) X_t = a_t + \theta a_{t-1}$

$X_t = (1-\phi B)^{-1} (a_t + \theta a_{t-1})$

$= (1 + \phi B + \phi^2 B^2 + \cdots)(a_t + \theta a_{t-1})$

$= a_t + \theta a_{t-1} + \phi a_{t-1} + \phi a_{t-2} + \cdots$

$= a_t + (\phi+\theta) a_{t-1} + (\phi^2 + \phi\theta) a_{t-2} + \cdots$

$\boxed{= a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \cdots}$ (psi)

$X_t \cdot X_{t-h} = \phi X_{t-1} \cdot X_{t-h} + a_t \cdot X_{t-h} + \theta \cdot a_{t-1} \cdot X_{t-h}$

$E(X_t \cdot X_{t-h}) = \phi E(X_{t-1} \cdot X_{t-h}) + E(a_t \cdot X_{t-h}) + \theta E(a_{t-1} \cdot X_{t-h})$

$Cov(X_t, X_{t-h}) = \phi Cov(X_{t-1}, X_{t-h}) + E(a_t \cdot X_{t-h}) + \theta E(a_{t-1} X_{t-h})$

$h=0, \quad \gamma_X(0) = \phi \gamma_X(1) + \underbrace{E(a_t \cdot X_t)}_{①} + \theta \underbrace{E(a_{t-1}, X_t)}_{②}$

① $E(a_t \cdot X_t)$

$= E(a_t \cdot (a_t + \psi_1 a_{t-1} + \psi a_{t-2} + \cdots))$

$= E(a_t \cdot a_t) + \psi_1 E(a_t \cdot a_{t-1}) + \cdots$

$= Cov(a_t, a_t)$

$= \sigma_a^2$.

② $E(a_{t-1} (a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \cdots))$

$= E(a_{t-1} \cdot a_t) + \boxed{\psi_1 E(a_{t-1} \cdot a_{t-1})} + \psi_2 E(a_{t-1}, a_{t-2}) + \cdots)$

$= \psi_1 \sigma_a^2$

$$\boxed{\therefore \gamma_X(0) = \sigma_X^2 = \phi \gamma_X(1) + \sigma_a^2 + \theta \underbrace{\psi_1}_{=(\phi+\theta)} \sigma_a^2}$$

Autovariance function of ARMA(1,1) $h=0$

$h=1, \quad Cov(X_t, X_{t-1}) = \phi Cov(X_{t-1}, X_{t-1}) + E(a_t \cdot X_{t-1}) + \theta E(a_{t-1} \cdot X_{t-1})$

$\gamma_X(1) = \phi \gamma_X(0) + \theta \cdot \sigma_a^2 \rightarrow ①$

$\boxed{\gamma_X(0) = \phi \gamma_X(1) + \sigma_a^2 + \theta(\phi+\theta)\sigma_a^2} \rightarrow ②$

①, ②

$$\gamma_X(0) = \frac{1 + 2\theta\phi + \theta^2}{1-\phi^2} \cdot \sigma_a^2 \longrightarrow h=0$$

$$\gamma_X(1) = \frac{(1+\theta\phi)(\phi+\theta)}{1-\phi^2} \sigma_a^2 \longrightarrow h=1$$

$h \geq 2. \quad Cov(X_t, X_{t-h}) = \phi Cov(X_{t-1}, X_{t-h}) + E(a_t \cdot X_{t-h}) + \theta E(a_{t-1} \cdot X_{t-h})$

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

$\gamma_X(h) = \phi^{h-1} \gamma_X(1)$ leave

$$\boxed{= \phi^{h-1} \cdot \frac{(1+\theta\phi)(\phi+\theta)}{1-\phi^2} \cdot \sigma_a^2}$$

AR process $\quad$ AR(p)

(WN) $\rightarrow \boxed{\phi^{-1}(B)} \rightarrow$ AR(p)

$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + a_t$

$X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \cdots - \phi_p X_{t-p} = a_t$

$X_t - \phi_1 B X_t - \phi_2 B^2 X_t - \cdots - \phi_p B^p X_t = a_t$

$\underbrace{(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p)}_{\phi(B)} X_t = a_t$

$\phi(B) X_t = a_t$

$\boxed{X_t = \phi^{-1}(B) \cdot a_t}$ WN

⇒ AR process can be thought of as the output X_t from a linear filter with transfer function $\phi^{-1}(B)$ when the input is white noise a_t.

MA(q)

$X_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_q a_{t-q}$

$= a_t - \theta_1 B \cdot a_t - \theta_2 B^2 a_t - \cdots - \theta_q B^q \cdot a_t$

$= (1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q) a_t$

$= \theta(B) \cdot a_t$.

$a_t \rightarrow \boxed{\theta(B)} \rightarrow$ MA(q)

ARMA(p,q)

$X_t = \underbrace{\phi_1 X_{t-1} + \cdots + \phi_p X_{t-p}}_{AR} + \underbrace{a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q}}_{MA}$

$\phi(B) X_t = \theta(B) \cdot a_t$

$X_t = \boxed{\phi^{-1}(B) \cdot \theta(B)} a_t$

$a_t \rightarrow \boxed{} \rightarrow X_t$ ARMA(p,q)

1 Backward shift operator B

An operator B is called the *backward shift operator*. It is a function that maps a random variable which depends on time (i.e. X_t or a_t), backward by one step.

$$B(X_t) = X_{t-1} \tag{1}$$

We occasionally write BX_t instead of $B(X_t)$ to stand for the same thing. Denoting $B^m = B \circ B \circ \cdots \circ B$ by the m times composition of B, we have

$$\begin{aligned} BX_t &= X_{t-1} \\ B^2 X_t &= X_{t-2} \\ &\vdots \\ B^m X_t &= X_{t-m}. \end{aligned}$$

The professor tries to explain AR(1) model and AR(2) model in terms of this operator B, which is illustrated as follows.

1.1 AR(1) model described by B

Recall that AR(1) is called the *autoregressive model* of order 1, which predict X_t by means of X_{t-1} and a_t.

$$X_t = \phi X_{t-1} + a_t. \tag{2}$$

Here, a_t is a white noise so that the expectation of a_t is 0 ;

$$\mathbb{E}[a_t] = 0, \tag{3}$$

the variance of a_t is independent of t ;

$$\mathbb{V}[a_t] = {\sigma_a}^2, \tag{4}$$

and a_t's in distinct timsteps are independent one another ;

$$\mathrm{Cov}(a_t, a_{t+h}) = 0 \quad \text{for} \quad h \neq 0. \tag{5}$$

Recall that $\mathbb{V}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$ for any random variable X. By (3) and (4),

$$\mathbb{E}[{a_t}^2] = \mathbb{V}[a_t] + (\mathbb{E}[a_t])^2 = {\sigma_a}^2 + 0^2 = {\sigma_a}^2. \tag{6}$$

Moreover, by (3) and (5),

$$\mathbb{E}[a_t a_s] = \mathbb{E}[(a_t - \mathbb{E}[a_t])(a_s - \mathbb{E}[a_s])] = \mathrm{Cov}(a_t, a_s) = 0 \tag{7}$$

if $t \neq s$. Summarizing the above two equations ((6) and (7)) yields

$$\mathbb{E}[a_t a_s] = \begin{cases} {\sigma_a}^2 & (t = s) \\ 0 & (t \neq s) \end{cases} \tag{8}$$

We can rewrite the equation (2) as

$$X_t = \phi B X_t + a_t$$
$$(1-\phi B)X_t = a_t$$

where $1-\phi B$ can be thought of as an oprator, which maps X_t to $X_t - \phi B X_t$. Here, 1 should be viewed as a identity operator. But, I think the professor regard $1-\phi B$ as a '1 minus a real number ϕB. Moreover, he assumes that $|\phi B| < 1$ in that

$$\begin{aligned} X_t &= \frac{1}{1-\phi B} a_t \\ &= (1 + \phi B + \phi^2 B^2 + \cdots) a_t \\ &= a_t + \phi a_{t-1} + \phi^2 a_{t-2} + \cdots . \end{aligned} \tag{9}$$

where he use the formula for geometric series.[1] As a result, X_t is expressed as a (infinite) linear combination of a_{t-h} for $h = 0, 1, 2, \cdots$.

1.2 AR(2) model described by B

Recall that AR(2) is an autoregressive model of order 2, which predict X_t by means of X_{t-1}, X_{t-2} and a_t.

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + a_t \tag{10}$$

Using the backward shift operator B, we get

$$X_t = \phi_1 B X_t + \phi_2 B^2 X_t + a_t$$
$$(1 - \phi_1 B - \phi_2 B^2) X_t = a_t$$

Again, regarding B as a real number, $1 - \phi_1 B - \phi_2 B^2$ can be thought of as a quadratic polynomial of B. So there exists numbers[2] α_1 and α_2 such that $(\alpha_1 + \alpha_2 = \phi_1,\ \alpha_1\alpha_2 = -\phi_2)$

$$1 - \phi_1 B - \phi_2 B^2 = (1-\alpha_1 B)(1-\alpha_2 B).$$

Thus,

$$\begin{aligned} X_t &= \frac{1}{1-\phi_1 B - \phi_2 B^2} a_t \\ &= \frac{1}{(1-\alpha_1 B)(1-\alpha_2 B)} a_t \\ &= (1 + \alpha_1 B + {\alpha_1}^2 B^2 + \cdots)(1 + \alpha_2 B + {\alpha_2}^2 B^2 + \cdots) a_t \\ &= a_t + (\alpha_1 + \alpha_2) a_{t-1} + ({\alpha_1}^2 + \alpha_1\alpha_2 + {\alpha_2}^2) a_{t-2} + ({\alpha_1}^3 + {\alpha_1}^2\alpha_2 + \alpha_1{\alpha_2}^2 + {\alpha_2}^3) a_{t-3} + \cdots . \end{aligned}$$

[1] An alternative way to equate $(1-\phi B)^{-1}$ and $1 + \phi B + \phi^2 B^2 + \cdots$ is illustrated in Shumway's book. See appendix.

[2] real numbers if $D = b^2 - 4ac = (-\phi_1)^2 - 4 \cdot (-\phi_2) \cdot 1 = {\phi_1}^2 + 4\phi_2 > 0$. but in general, they are complex numbers

1.3 Forward shift operator F

In contrast to the backward shift operator B, the *forward shift operator* F shift a random variable forward.

$$\begin{aligned} FX_t &= X_{t+1} \\ F^2X_t &= X_{t+2} \\ &\vdots \\ F^kX_t &= X_{t+k} \end{aligned} \tag{11}$$

for any positive integer k.

2 ARMA(1,1) model

ARMA(1,1) model is a combination of AR(1) model and MA(1) model. We are assuming that X_t is obtained from X_{t-1}, a_t and a_{t-1}.

$$X_t = \phi X_{t-1} + a_t + \theta a_{t-1}. \tag{12}$$

Here, we compute the covariance $\gamma(h) = \mathrm{Cov}(X_t, X_{t-h})$ of this model for $h = 0, 1, 2, \cdots$. To this aim, multiply (12) by X_{t-h} on both sides

$$X_tX_{t-h} = \phi X_{t-1}X_{t-h} + a_tX_{t-h} + \theta a_{t-1}X_{t-h}$$

and take the expectation on both sides to get[3];

$$\begin{aligned} \gamma(h) &= \mathbb{E}[X_tX_{t-h}] \\ &= \phi\mathbb{E}[X_{t-1}X_{t-h}] + \mathbb{E}[a_tX_{t-h}] + \theta\mathbb{E}[a_{t-1}X_{t-h}] \\ &= \phi\mathrm{Cov}(X_{t-1}, X_{t-h}) + \mathbb{E}[a_tX_{t-h}] + \theta\mathbb{E}[a_{t-1}X_{t-h}] \end{aligned} \tag{13}$$

2.1 autocovariance for $h = 0$: $\gamma(0)$

Suppose $h = 0$. Before computation, we can convert the equation (12) using the operator B ;

$$\begin{aligned} (1 - \phi B)X_t &= a_t + \theta a_{t-1} \\ X_t &= \frac{1}{1 - \phi B}(a_t + \theta a_{t-1}) \\ &= (1 + \phi B + \phi^2B^2 + \cdots)(a_t + \theta a_{t-1}) \\ &= a_t + \theta a_{t-1} + \phi a_{t-1} + \phi\theta a_{t-2} + \phi^2a_{t-2} + \phi^2\theta a_{t-3} + \cdots \\ &= a_t + (\theta + \phi)a_{t-1} + \phi(\theta + \phi)a_{t-2} + \cdots \\ &= a_t + \psi_1a_{t-1} + \psi_2a_{t-2} + \cdots \end{aligned} \tag{14}$$

[3]여기에서 $\mathbb{E}[X_t] = \mathbb{E}[X_{t-h}] = 0$임을 가정하고 있는 것일까? Covariance의 정의는 $\mathrm{Cov}(X, Y) = \mathbb{E}[(X - [X])(Y - [Y])]$인데 마치 $\mathrm{Cov}(X, Y) = \mathbb{E}[XY]$인 것처럼 쓰고 있다. Shumway의 책에서는 x_t의 평균이 0임을 가정하고 있다. x_t의 평균이 0이 아닐 경우 평균 μ를 뺀 $x_t - \mu$를 고려하라고 되어 있다. (p. 86)

By (14), the second term and the third term of (13) are evalauted as

$$
\begin{aligned}
\mathbb{E}[a_t X_t] &= \mathbb{E}[a_t(a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \cdots)] \\
&= \mathbb{E}[{a_t}^2] + \psi_1 \mathbb{E}[a_t a_{t-1}] + \psi_2 \mathbb{E}[a_t a_{t-2}] + \cdots \\
&= {\sigma_a}^2 + 0 + 0 + \cdots = {\sigma_a}^2,
\end{aligned}
$$

and

$$
\begin{aligned}
\theta \mathbb{E}[a_{t-1} X_t] &= \theta \mathbb{E}[a_{t-1}(a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \cdots)] \\
&= \theta \left(\mathbb{E}[a_{t-1} a_t] + \psi_1 \mathbb{E}[{a_{t-1}}^2] + \psi_2 \mathbb{E}[a_{t-1} a_{t-2}] + \cdots \right) \\
&= \theta \left(0 + \psi_1 {\sigma_a}^2 + 0 + \cdots \right) \\
&= \theta \psi_1 {\sigma_a}^2,
\end{aligned}
$$

where we used (8) in the computations. Therefore, the autocovariance function of ARMA(1,1) when $h = 0$ is

$$
\begin{aligned}
\gamma(0) &= \phi \mathrm{Cov}(X_{t-1}, X_t) + {\sigma_a}^2 + \theta \psi_1 {\sigma_a}^2 \\
&= \phi \gamma(1) + (1 + \theta^2 + \theta\phi) {\sigma_a}^2,
\end{aligned}
\tag{15}
$$

which is also the variance $\mathbb{V}(X_t)$ of the ARMA(1,1) model.

2.2 autocovariance for $h = 1$: $\gamma(1)$

Suppose $h = 1$. The equation (12) applied to $t - 1$ yields

$$
\begin{aligned}
X_{t-1} &= \phi X_{t-2} + a_{t-1} + \theta a_{t-2} \\
X_{t-1} &= \phi B X_{t-1} + a_{t-1} + \theta a_{t-2} \\
(1 - \phi B) X_{t-1} &= a_{t-1} + \theta a_{t-2} \\
X_{t-1} &= (1 + \phi B + \phi^2 B^2 + \cdots)(a_{t-1} + \theta a_{t-2}) \\
&= a_{t-1} + \psi_1 a_{t-2} + \psi_2 a_{t-3} + \cdots
\end{aligned}
$$

for the same ψ_{t-h} $(h = 1, 2, \cdots)$. Plugging $h = 1$ and the above equation into (13),

$$
\begin{aligned}
\gamma(1) &= \phi \mathrm{Cov}(X_{t-1}, X_{t-1}) + \mathbb{E}[a_t X_{t-1}] + \theta \mathbb{E}[a_{t-1} X_{t-1}] \\
&= \phi \gamma(0) + \mathbb{E}[a_t(a_{t-1} + \psi_1 a_{t-2} + \psi_2 a_{t-3} + \cdots)] + \theta \mathbb{E}[a_{t-1}(a_{t-1} + \psi_1 a_{t-2} + \psi_2 a_{t-3} + \cdots)] \\
&= \phi \gamma(0) + 0 + \theta \times {\sigma_a}^2 \\
&= \phi \gamma(0) + \theta {\sigma_a}^2
\end{aligned}
\tag{16}
$$

by repeated uses of (8).

As a result, we get a system of linear equations (15) and (16) with unknowns $\gamma(0)$ and $\gamma(1)$. Solving the system by substitution yields

$$
\begin{aligned}
\gamma(0) &= \phi \gamma(1) + (1 + \theta^2 + \theta\phi) {\sigma_a}^2 \\
&= \phi^2 \gamma(0) + (1 + 2\theta\phi + \theta^2) {\sigma_a}^2
\end{aligned}
$$

$$\gamma(0) = \frac{1 + 2\theta\phi + \theta^2}{1 - \phi^2}{\sigma_a}^2 \tag{17}$$

and

$$\begin{aligned}\gamma(1) &= \phi\gamma(0) + \theta{\sigma_a}^2 \\ &= \phi^2\gamma(1) + (\phi + \phi\theta^2 + \phi^2\theta + \theta){\sigma_a}^2\end{aligned}$$

$$\gamma(1) = \frac{\phi + \theta\phi^2 + \theta^2\phi + \theta}{1 - \phi^2}{\sigma_a}^2 = \frac{(\phi + \theta)(1 + \phi\theta)}{1 - \phi^2}{\sigma_a}^2 \tag{18}$$

2.3 autocovariance for $h \geq 2$: $\gamma(h)$

By the similar reasoning as in 2.1 and 2.2, we have

$$\begin{aligned}X_{t-h} &= (1 + \phi B + \phi^2 B^2 + \cdots)(a_{t-h} + \theta a_{t-h-1}) \\ &= a_{t-h} + \psi_1 a_{t-h-1} + \psi_2 a_{t-h-2} + \cdots\end{aligned}$$

Substituting the above equation to (13) yields

$$\begin{aligned}\gamma(h) &= \phi\mathrm{Cov}(X_{t-1}, X_{t-h}) + \mathbb{E}[a_t X_{t-h}] + \theta\mathbb{E}[a_{t-1}X_{t-h}] \\ &= \phi\gamma(h-1) + \mathbb{E}[a_t(a_{t-h} + \psi_1 a_{t-h-1} + \psi_2 a_{t-h-2} + \cdots)] + \theta\mathbb{E}[a_{t-1}(a_{t-h} + \psi_1 a_{t-h-1} + \psi_2 a_{t-h-2} + \cdots)] \\ &= \phi\gamma(h-1).\end{aligned} \tag{19}$$

Here, all the terms of the form $\mathbb{E}[a_t a_{t-l}]$ or $\mathbb{E}[a_{t-1}a_{t-l}]$ has vanished since $h \geq 2$. Using induction on the above reccurence formula (19), we have $\gamma(h) = \phi^{h-1}\gamma(1)$ for $h \geq 2$. Combined with (17) and (18), we get

$$\gamma(h) = \begin{cases} \frac{1+2\theta\phi+\theta^2}{1-\phi^2}{\sigma_a}^2 & (h = 0) \\ \frac{(\phi+\theta)(1+\phi\theta)\phi^{h-1}}{1-\phi^2}{\sigma_a}^2 & (h \geq 1) \end{cases} \tag{20}$$

Since $\gamma(-h) = \gamma(h)$, equation (20) is the explicit formula of $\gamma(h)$ for all integers h.

3 AR, MA and ARMA described by operators

The lecture ends by giving a summary for AR model, MA model and ARMA model. The moral is that those models can be expressed in terms of *polynomials* of B.

3.1 AR model

Consider the AR(p) model, characterized by the equation

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + a_t \tag{21}$$

The above equation can be converted into

$$(1 - \phi B - \phi^2 B^2 - \cdots - \phi^p B^p)X_t = a_t \tag{22}$$

We denote

$$\phi(B) = 1 - \phi B - \phi^2 B^2 - \cdots - \phi^p B^p, \tag{23}$$

so that $\phi(B)$ is an operator (a polynomial of operator B) acting on a random variable. The equation (22) can be expressed as

$$\phi(B) X_t = a_t \tag{24}$$

If the inverse operator $\phi^{-1}(B)$ of $\phi(B)$ exists[4], the above equation goes

$$X_t = \phi^{-1}(B) a_t.$$

Here is a comment by the professor, summarizing the above equation :

> AR process can be thought of as making the output X_t from a linear filter with transfer function $\phi^{-1}(B)$ when the input is a white noise a_t.

3.2 MA model

The MA(q) model is governed by the equation[5]

$$X_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_q a_{t-q}. \tag{25}$$

We have

$$X_t = (1 - \theta B - \theta^2 B^2 - \cdots - \theta^q B^q) a_t \tag{26}$$

$$\theta(B) = 1 - \theta B - \theta^2 B^2 - \cdots - \theta^q B^q \tag{27}$$

$$X_t = \theta(B) a_t \tag{28}$$

3.3 ARMA model

The ARMA(p, q) model has the equation

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_q a_{t-q}. \tag{29}$$

By the similar procedures that we did earlier,

$$\phi(B) X_t = \theta(B) a_t \tag{30}$$

holds for the same $\phi(B)$ and $\theta(B)$ defined in (23) and (27), repsectively. Taking the inverse $\phi^{-1}(B)$ of $\phi(B)$ on both sides yields

$$X_t = \phi^{-1}(B) \theta(B) a_t. \tag{31}$$

[4] actually, it exists if $|\phi| < 1$ whose explicit form is written in (9). The way of geometric series doesn't seem to be rigorous. Shumway explain an alternative way in section 3.2.

[5] Note that the professor wrote minus sign instead of plus sign. The choice of signs are irrelavant of the context. Confer to the second footnote at page 90, in Shumway's book.