linear model

Lecture 3. Autoregressive Models

AR model

1. AR(1) Model

Definition 1.1. The *autoregressive* (AR) model of order p is defined by

we know $x_{t-1}, \cdots, x_{t-p}$

$$x_t = \phi_0 + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \phi_p x_{t-p} + \varepsilon_t,$$

$E[X_t | X_{t-1}, x_{t-2}, \cdots, x_{t-p}]$

$=$ —— //.

where $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$.

- The expectation of x_t conditioning on $x_{t-1}, \ldots, x_{t-p}$ is

$$E[x_t | x_{t-1}, \ldots, x_{t-p}] = \phi_0 + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \phi_p x_{t-p},$$

meaning that given the past data, the first p lagged variables x_{t-i} jointly determine the conditional expectation of x_t.

$E[x_t] = \phi_0 + \phi_1 E[x_{t-1}]$

$\boxed{|\phi_1| < 1}$

Theorem 1.2. *Consider an AR(1) model of the form*

$E[\varepsilon_t] = 0$

$E[\varepsilon_t^2] = \sigma_\varepsilon^2$

$$x_t = \phi_0 + \phi_1 x_{t-1} + \varepsilon_t, \tag{1.1}$$

AR model의 시계열이 stationarity가 있으려면.

where $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$*. Then, stationarity holds if and only if* $|\phi_1| < 1$*.*

Proof. Assume that x_t is stationary. Then, one obtains

$$\mu = \phi_0 + \phi_1 \mu$$

or

$$\mu = \frac{\phi_0}{1 - \phi_1},$$

where $E[x_t] = \mu$. Using $\phi_0 = (1 - \phi_1)\mu$, one uses repeated substitution to write (1.1) as

$x_t = \phi_0 + \phi_1 x_{t-1} + \varepsilon_t$

$= \mu - \mu\phi_1 + \phi_1 x_{t-1} + \varepsilon_t$

$x_t - \mu = \phi_1(x_{t-1} - \mu) + \varepsilon_t$

$x_{t-1} - \mu = \phi_1(x_{t-2} - \mu) + \varepsilon_{t-1}$

$$\begin{aligned} x_t - \mu &= \phi_1(x_{t-1} - \mu) + \varepsilon_t \\ &= \phi_1(\phi_1(x_{t-2} - \mu) + \varepsilon_{t-1}) + \varepsilon_t \\ &= \varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_1^2 \varepsilon_{t-2} + \cdots. \end{aligned} \tag{1.2}$$

From (1.2), one sees that an AR(1) model is expressed as a linear function of ε_{t-i} for $i \geq 0$. Since $\{\varepsilon_t\}$ is a white noise process, it shows

μ

$$\begin{aligned} Cov[x_{t-1}, \varepsilon_t] &= E[(x_{t-1} - E[x_{t-1}])(\varepsilon_t - E[\varepsilon_t])] \\ &= E[(x_{t-1} - \mu)\varepsilon_t] \\ &= E[(\varepsilon_{t-1} + \phi_1 \varepsilon_{t-2} + \phi_2^2 \varepsilon_{t-3} + \cdots)\varepsilon_t] \\ &= 0. \end{aligned}$$

$cov[\varepsilon_{t-1}, \varepsilon_{t-k}]$

$= E[(\varepsilon_{t-1} - E[\varepsilon_{t-1}])(\varepsilon_{t-k} - E[\varepsilon_{t-k}])] = E[\varepsilon_{t-1}\varepsilon_{t-k}] = 0$

Taking the square and the expectation of

$$x_t - \mu = \phi_1(x_{t-1} - \mu) + \varepsilon_t,$$

$E[\varepsilon_t^2] = E[(\varepsilon_t - 0)^2]$
$= E[(\varepsilon_t - E[\varepsilon_t])^2]$
$= Var[\varepsilon_t] = \sigma_\varepsilon^2$

one obtains

$$E[(x_t - \mu)^2] = \phi_1^2 E[(x_{t-1} - \mu)^2] + E[\varepsilon_t^2] + 2\phi_1 E[(x_{t-1} - \mu)\varepsilon_t]$$

$Cov[x_{t-1}, \varepsilon_t]$

or

$$Var[x_t] = \phi_1^2 Var[x_{t-1}] + \sigma_\varepsilon^2.$$

Since $Var[x_t] = Var[x_{t-1}]$ under the stationarity assumption, one obtains

$$Var[x_t] = \frac{\sigma_\varepsilon^2}{1 - \phi_1^2},$$

provided that $\phi_1^2 < 1$ which ensures that $Var[x_t]$ is bounded and nonnegative. Consequently, the stationarity of an AR(1) model implies that $|\phi_1| < 1$. ■

Assume that $|\phi_1| < 1$. Based on (1.2), one shows that

$$E[x_t] = \mu$$

and

$\phi_1^k \varepsilon_{t-k}$

$$\begin{aligned}
Cov[x_t, x_{t-k}] &= E[(\varepsilon_t + \phi_1\varepsilon_{t-1} + \phi_1^2\varepsilon_{t-2} + \cdots + \phi_1^{k+1}\varepsilon_{t-k-1} + \phi_1^{k+2}\varepsilon_{t-k-2} + \phi_1^{k+3}\varepsilon_{t-k-3} + \cdots) \\
&\quad \times (\varepsilon_{t-k} + \phi_1\varepsilon_{t-k-1} + \phi_1^2\varepsilon_{t-k-2} + \phi_1^3\varepsilon_{t-k-3} + \cdots)] \\
&= \phi_1^{k+1}\phi_1 E[\varepsilon_{t-k-1}^2] + \phi_1^{k+2}\phi_1^2 E[\varepsilon_{t-k-2}^2] + \phi_1^{k+3}\phi_1^3 E[\varepsilon_{t-k-3}^2] + \cdots \\
&= \sigma_\varepsilon^2\left(\phi_1^{k+2} + \phi_1^{k+4} + \phi_1^{k+6} + \cdots\right) \\
&= \frac{\sigma_\varepsilon^2 \phi_1^{k\cancel{+2}}}{1 - \phi_1^2},
\end{aligned}$$

$\phi_1^k E[\varepsilon_{t-k}^2]$

ϕ_1^k

both of which are finite and time invariant. Therefore, an AR(1) model is stationary if $|\phi_1| < 1$. In summary, the necessary and sufficient condition for an AR(1) model is $|\phi_1| < 1$. □

Remark 1.3. For a stationary AR(1) model (i.e., $|\phi_1| < 1$), the kth order autocovariance is

$$\begin{aligned}
\gamma_k &= Cov[x_t, x_{t-k}] \\
&= Cov[\phi_0 + \phi_1 x_{t-1} + \varepsilon_t, x_{t-k}] \\
&= \phi_1\gamma_{k-1},
\end{aligned}$$

$x_t = \phi_0 + \phi_1 x_{t-1} + \varepsilon_t$

$\Downarrow$

$\frac{\gamma_k}{\gamma_0} = \phi_1 \frac{\gamma_{k-1}}{\gamma_0}$

$\Rightarrow \rho_k = \phi_1 \rho_{k-1}$

since

$$\begin{aligned}
Cov[\varepsilon_t, x_{t-k}] &= E[\varepsilon_t(x_{t-k} - \mu)] \\
&= E[\varepsilon_t(\varepsilon_{t-k} + \phi_1\varepsilon_{t-k-1} + \phi_2^2\varepsilon_{t-k-2} + \cdots)] \\
&= 0.
\end{aligned}$$

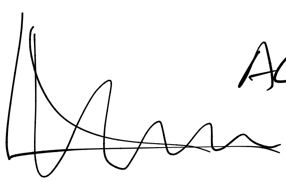
$\rho_0 = \frac{\gamma_0}{\gamma_0} = 1$

$\rho_1 = \phi_1$

$\rho_2 = \phi_1\rho_1 = \phi_1^2$

$\rho_3 = \phi\rho_2 = \phi_1(\phi_1^2) = \phi_1^3$

$\Rightarrow \rho_k = \phi_1^k$ $|\phi_1| < 1$



ACF line.

AR(1)

$$x_t = \phi_0 + \phi_1 x_{t-1} + \varepsilon_t$$

x_t is stationary $\Rightarrow |\phi_1| < 1$

$$\mu = \phi_0 + \phi_1 \mu$$

$$\mu = \frac{\phi_0}{1-\phi_1} \qquad \phi_0 = \mu(1-\phi_1)$$

$$x_t = \mu(1-\phi_1) + \phi_1 x_{t-1} + \varepsilon_t$$

$$x_t - \mu = \phi_1 (x_{t-1} - \mu) + \varepsilon_t$$

$$(x_{t-1} - \mu = \phi_1 (x_{t-2} - \mu) + \varepsilon_{t-1}).$$

$$\begin{aligned} x_t - \mu &= \phi_1(\phi_1(x_{t-2} - \mu) + \varepsilon_{t-1}) + \varepsilon_t \\ &= \phi_1(\phi_1(\phi_1(x_{t-3} - \mu) + \varepsilon_{t-2}) + \varepsilon_{t-1}) + \varepsilon_t \\ &= \phi_1(\phi_1(\phi_1(\phi_1(x_{t-4} - \mu) + \varepsilon_{t-3}) + \varepsilon_{t-2}) + \varepsilon_{t-1}) + \varepsilon_t \\ &= \cdots \\ &= \varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_1^2 \varepsilon_{t-2} + \phi_1^3 \varepsilon_{t-3} + \cdots \end{aligned}$$

$$\begin{aligned} x_t &= \phi_0 + \phi_1 x_{t-1} + \varepsilon_t \\ &= \underline{\mu + \varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_1^2 \varepsilon_{t-2}} + \cdots \end{aligned} \Rightarrow \text{linear.}$$

$$\begin{aligned} \mathrm{Cov}[x_{t-1}, \varepsilon_t] &= E[(x_{t-1} - \mu)\varepsilon_t] \\ &= E[(\varepsilon_{t-1} + \phi_1 \varepsilon_{t-2} + \phi_1^2 \varepsilon_{t-3} + \cdots)\varepsilon_t] \\ &= E[\varepsilon_t \cdot \varepsilon_{t-1} + \phi_1 \varepsilon_t \varepsilon_{t-2} + \phi_1^2 \varepsilon_t \varepsilon_{t-3} + \cdots] \\ &= E[\varepsilon_t \varepsilon_{t-1}] + \phi_1 E[\varepsilon_t \varepsilon_{t-2}] + \phi_1^2 E[\varepsilon_t \varepsilon_{t-3}] + \cdots \\ &= 0. \end{aligned}$$

$(\because E[\varepsilon_t \varepsilon_{t-1}] = \mathrm{Cov}[\varepsilon_t, \varepsilon_{t-1}] = 0)$

$\because \varepsilon_t \sim WN(0, \sigma_\varepsilon^2). \Rightarrow \mathrm{Cov}[x_t, x_{t+k}] = 0)$

$$x_t - \mu = \phi_1 (x_{t-1} - \mu) + \varepsilon_t$$

$$
\begin{aligned}
&E[(x_t - \mu)^2] \\
&= E[\phi_1^2 (x_{t-1} - \mu)^2 + 2\phi_1 (x_{t-1} - \mu)\varepsilon_t + \varepsilon_t^2] \\
&= \phi_1^2 E[(x_{t-1} - \mu)^2] + 2\phi_1 E[(x_{t-1} - \mu)\varepsilon_t] + E[\varepsilon_t^2] \\
&= \phi_1^2 E[(x_{t-1} - \mu)^2] + 2\phi_1 Cov[x_{t-1}, \varepsilon_t] + \sigma_\varepsilon^2 \\
&= \phi_1^2 E[(x_{t-1} - \mu)^2] + \sigma_\varepsilon^2
\end{aligned}
$$

$$\therefore Var[x_t] = \phi_1^2 Var[x_{t-1}] + \sigma_\varepsilon^2$$

In stationary assumtion,

$Var[x_t] = Var[x_{t-1}]$ (항상 설정).

$$\therefore Var[x_t] = \frac{\sigma_\varepsilon^2}{1 - \phi_1^2} > 0.$$

$$\therefore \phi_1^2 < 1, \quad |\phi_1| < 1$$

$|\phi_1| < 1 \Rightarrow x_t$ stationary.

$x_t = \phi_0 + \phi_1 x_{t-1} + \varepsilon_t$

we can also express AR(1) like this

$x_t - \alpha = \phi_1 (x_{t-1} - \alpha) + \varepsilon_t$

where $\phi_0 = \alpha(1-\phi_1)$, $\alpha = \frac{\phi_0}{1-\phi_1}$

let $y_t = x_t - \alpha$.

$$
\begin{aligned}
y_t &= \phi_1 y_{t-1} + \varepsilon_t \\
&= \phi_1 (\phi_1 y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\
&= \phi_1 (\phi_1 (\phi_1 y_{t-3} + \varepsilon_{t-2}) + \varepsilon_{t-1}) + \varepsilon_t \\
&= \cdots \\
&= \varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_1^2 \varepsilon_{t-2} + \cdots \quad \text{linear.}
\end{aligned}
$$

$E[y_t] = E[x_t - \alpha] = E[x_t] - \alpha = 0.$

$\therefore E[x_t] = \alpha$: time invariant.

$$
\begin{aligned}
Cov[x_t, x_{t-k}] &= E[(x_t - \alpha)(x_{t-k} - \alpha)] \\
&= E[(\varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_1^2 \varepsilon_{t-2} + \cdots)(\varepsilon_{t-k} + \phi_1 \varepsilon_{t-k-1} + \cdots)]. \\
&= E[\phi_1^k \varepsilon_{t-k} \cdot \varepsilon_{t-k} + \phi_1^{k+1} \varepsilon_{t-k-1} \phi_1 \varepsilon_{t-k+1} + \cdots] \\
&= \phi_1^k \sigma_\varepsilon^2 + \phi_1^{k+2} \sigma_\varepsilon^2 + \cdots \\
&= \sigma_\varepsilon^2 (\phi_1^k + \phi_1^{k+2} + \phi_1^{k+4} + \cdots) \\
&= \frac{\sigma_\varepsilon^2 \phi_1^k}{1-\phi_1^2} \quad : \text{time invariant.}
\end{aligned}
$$

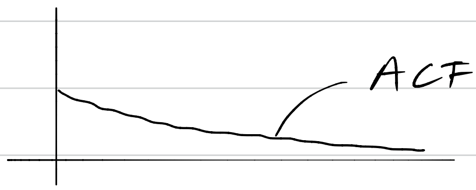
$\therefore x_t$ is stationary. $\blacksquare$

x_t is stationary $\Leftrightarrow |\phi_1| < 1$.

stationary AR(1) model.

$$\gamma_k = Cov[x_t, x_{t-k}]$$
$$= Cov[\phi_0 + \phi_1 x_{t-1} + \varepsilon_t,\ x_{t-k}]$$
$$= Cov[\phi_0, x_{t-k}] + Cov[\phi_1 x_{t-1}, x_{t-k}] + Cov[\varepsilon_t, x_{t-k}]$$
$$= \phi_1 Cov[x_{t-1}, x_{t-k}]$$
$$= \phi_1 Cov[x_t, x_{t-(k-1)}]$$
$$= \phi_1 \gamma_{k-1}$$

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\phi_1 \gamma_{k-1}}{\gamma_0} = \phi_1 \rho_{k-1} \quad (ACF)$$
$$\rho_k = \phi_1 \rho_{k-1}$$
$$\rho_0 = 1 \quad \rho_1 = \phi_1 \quad \rho_2 = \phi_1^2 \quad \rho_3 = \phi_1^3$$
$$\therefore \rho_k = \phi_1^k$$



pf of $Cov[x_{t-k}, \varepsilon_t] = 0$.
$$Cov[x_{t-k}, \varepsilon_t] = E[(x_{t-k} - \mu)\varepsilon_t]$$
$$=$$

$$x_t = \phi_0 + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \varepsilon_t$$
$$\mu = \phi_0 + \mu(\phi_1 + \phi_2)$$
$$= \frac{\phi_0}{1 - \phi_1 - \phi_2}$$
$$\phi_0 = \mu(1 - \phi_1 - \phi_2).$$
$$x_t - \mu = \phi_1(x_{t-1} - \mu) + \phi_2(x_{t-2} - \mu) + \varepsilon_t$$
$$Cov[x_{t-k}, \varepsilon_t] = E[(x_{t-k} - \mu)\varepsilon_t]$$
$$= E[(\phi_1(x_{t-k-1} - \mu) + \phi_2(x_{t-k-2} - \mu) + \varepsilon_{t-k})\varepsilon_t]$$

ACF of stationary AR(2).

$$x_t = \phi_0 + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \varepsilon_t$$

$$\begin{aligned}
\gamma_k &= \mathrm{cov}[x_t, x_{t-k}] \\
&= \mathrm{cov}[\phi_0 + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \varepsilon_t, x_{t-k}] \\
&= \mathrm{cov}[\phi_0, x_{t-k}] + \mathrm{cov}[\phi_1 x_{t-1}, x_{t-k}] \\
&\quad + \mathrm{cov}[\phi_2 x_{t-2}, x_{t-k}] + \mathrm{cov}[\varepsilon_t, x_{t-k}] \\
&= \phi_1 \mathrm{cov}[x_{t-1}, x_{t-k}] + \phi_2 \mathrm{cov}[x_{t-2}, x_{t-k}] \\
&= \phi_1 \mathrm{cov}[x_t, x_{t-(k-1)}] + \phi_2 \mathrm{cov}[x_t, x_{t-(k-2)}] \\
&= \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2}
\end{aligned}$$

$$\begin{aligned}
\gamma_k &= \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} \\
\gamma_0 &= \mathrm{Var}[x_t] = \sigma^2 \\
\gamma_1 &= \phi_1 \gamma_0 + \phi_2 \gamma_{-1} = \phi_1 \sigma^2 + \phi_2 \gamma_1 \\
&= \frac{\phi_1 \sigma^2}{1-\phi_2}
\end{aligned}$$

$$\begin{aligned}
\rho_0 &= \frac{\gamma_0}{\gamma_0} = 1 \\
\rho_1 &= \frac{\gamma_1}{\gamma_0} = \frac{\gamma_1}{\sigma^2} = \frac{\phi_1}{1-\phi_2} \\
\rho_k &= \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} \\
\rho_1 &= \phi_1 \rho_0 + \phi_2 \rho_{-1} = \phi_1 + \phi_2 \rho_1 = \frac{\phi_1}{1-\phi_2} \quad \text{(check).} \\
\rho_2 &= \phi_1 \rho_1 + \phi_2 \rho_0 = \frac{\phi_1^2}{1-\phi_2} + \phi_2 \\
\rho_3 &= \phi_1 \rho_2 + \phi_2 \rho_1 = \phi_1\left(\frac{\phi_1^2}{1-\phi_2} + \phi_2\right) + \phi_2\left(\frac{\phi_1}{1-\phi_2}\right) \\
\rho_4 &= \phi_1 \rho_3 + \phi_2 \rho_2 = \phi_1\left(\phi_1\left(\frac{\phi_1^2}{1-\phi_2} + \phi_2\right) + \phi_2\left(\frac{\phi_1}{1-\phi_2}\right)\right) \\
&\quad + \phi_2\left(\frac{\phi_1^2}{1-\phi_2} + \phi_2\right) \\
\rho_2 &= \frac{1}{1-\phi_2}\left(\phi_1^2 + \phi_2(1-\phi_2)\right) \\
\rho_3 &= \frac{1}{1-\phi_2}\left(\phi_1(\phi_1^2 + \phi_2(1-\phi_2)) + \phi_1\phi_2\right) \\
&= \frac{1}{1-\phi_2}\left(\phi_1(\phi_1^2 + \phi_2(2-\phi_2))\right)
\end{aligned}$$

$$\rho_3 = \frac{1}{1-\phi_2}\left(\phi_1(\phi_1^2 + \phi_2(2-\phi_2))\right)$$

$$\rho_4 = \frac{1}{1-\phi_2}\left(\phi_1^2(\phi_1^2 + \phi_2(2-\phi_2))\right) + \frac{1}{1-\phi_2}\left(\phi_2(\cancel{\phi_1^2} + \phi_2(1-\phi_2))\right)$$

$$= \frac{1}{1-\phi_2}\left(\phi_1^2(\phi_1^2 + \phi_2(3-\phi_2)) + \phi_2^2(1-\phi_2)\right)$$

$$\rho_5 = \frac{1}{1-\phi_2}\Big[\phi_1^3(\phi_1^2 + \phi_2(3-\phi_2)) + \phi_1\phi_2^2(1-\phi_2))$$
$$+ \phi_1\phi_2(\phi_1^2 + \phi_2(2-\phi_2))\Big]$$

$$= \frac{1}{1-\phi_2}\Big[\phi_1^3(\phi_1^2 + \phi_2(4-\phi_2)) + \phi_1\phi_2^2(1-\phi_2))$$
$$+ \phi_1\phi_2^2(2-\phi_2)$$

$$= \frac{1}{1-\phi_2}\phi_1\left[\phi_1^2(\phi_1^2 + \phi_2(4-\phi_2)) + \phi_2^2(3-2\phi_2)\right]$$

$$\rho_6 = \frac{1}{1-\phi_2}\Big[\phi_1^4(\phi_1^2 + \phi_2(4-\phi_2)) + \phi_1^2\phi_2^2(3-2\phi_2)$$
$$+ \phi_1^2\phi_2(\phi_1^2 + \phi_2(3-\phi_2)) + \phi_2^3(1-\phi_2)\Big]$$

$$= \frac{1}{1-\phi_2}\Big[\phi_1^4(\phi_1^2 + \phi_2(5-\phi_2)) + \phi_1^2\phi_2^2(3-2\phi_2)$$
$$+ \phi_1^2\phi_2^2(3-\phi_2)) + \phi_2^3(1-\phi_2)\Big]$$

$$= \frac{1}{1-\phi_2}\Big[\phi_1^4(\phi_1^2 + \phi_2(5-\phi_2)) + \phi_1^2\phi_2^2(6-3\phi_2)$$
$$+ \phi_2^3(1-\phi_2)\Big]$$

$$\rho_k = \left(1 + \frac{1+\phi_2}{1-\phi_2}k\right)\left(\frac{\phi_1}{2}\right)^k$$

Consequently, the ACF of x_t satisfies $\rho_k = \phi_1 \rho_{k-1}$ for all $k > 0$. Since $\rho_0 = 1$, it further shows $\rho_k = \phi_1^k$. Therefore, the ACF of a stationary AR(1) series decays exponentially with rate ϕ_1 and starting value $\rho_0 = 1$.

2. AR(p) Model

Definition 2.1. The *lag operator* L is defined such that for any time series $\{x_t\}$,

$$Lx_t = x_{t-1}.$$

- The properties are (a) $L^2 x_t = L \cdot Lx_t = Lx_{t-1} = x_{t-2}$, (b) $L^j x_t = x_{t-j}$, (c) $L^0 = 1$, (d) $L^{-1} x_t = x_{t+1}$, and (e) $L \cdot a = a$ for any constant a.
- The operator $\Delta = 1 - L$ creates the first difference of a time series; i.e., $\Delta x_t = (1-L)x_t = x_t - x_{t-1}$.

Consider an AR model of order p of the form

$x_t - Lx_t = x_t - x_{t-1}$

$\parallel$

$(1-L)x_t$

Δ

$$x_t = \phi_0 + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \phi_p x_{t-p} + \varepsilon_t,$$

where $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$.

$\mu = \phi_0 + \mu(\phi_1 + \cdots + \phi_p)$

- Assume that x_t is stationary. Then, it shows

$\phi_0 = \mu(1 - \phi_1 - \cdots - \phi_p)$

$$E[x_t] = \boxed{\mu = \frac{\phi_0}{1 - \phi_1 - \cdots - \phi_p}},$$

$x_t = \mu(1 - \phi_1 - \cdots - \phi_p) + \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + \varepsilon_t$

provided $\phi_1 + \cdots + \phi_p \neq 1$. Using $\phi_0 = (1 - \phi_1 - \cdots - \phi_p)\mu$, one writes the AR($p$) model as

$$(x_t - \mu) = \phi_1(x_{t-1} - \mu) + \cdots + \phi_p(x_{t-p} - \mu) + \varepsilon_t. \tag{2.1}$$

Multiplying (2.1) by $x_{t-k} - \mu$ and using $E[(x_{t-k} - \mu)\varepsilon_t] = 0$ for $k > 0$, one obtains

$$(x_{t-k} - \mu)(x_t - \mu) = \phi_1(x_{t-k} - \mu)(x_{t-1} - \mu) + \cdots + \phi_{\cancel{2}\,p}(x_{t-k} - \mu)(x_{t-p} - \mu) + (x_{t-k} - \mu)\varepsilon_t,$$

which leads to

$$\gamma_k = \phi_1 \gamma_{k-1} + \cdots + \phi_p \gamma_{k-p}$$

for $k > 0$. Consequently, the ACF of x_t satisfies

$$\rho_k = \phi_1 \rho_{k-1} + \cdots + \phi_p \rho_{k-p}$$

or

$$(1 - \phi_1 L - \cdots - \phi_p L^p)\rho_k = 0.$$

Definition 2.2. A pth-order polynomial equation

$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p = 0$$

AR(p).

$$x_t = \phi_0 + \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + \varepsilon_t$$

$$(1 - \phi_1 L - \phi_2 L^2 - \cdots \phi_p L^p) x_t = \phi_0 + \varepsilon_t$$

$$\Rightarrow \phi(L) x_t = \phi_0 + \varepsilon_t$$

AR(2).

$$x_t = \phi_0 + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \varepsilon_t$$

$$\mu = \phi_0 + \phi_1 \mu + \phi_2 \mu$$

$$= \frac{\phi_0}{1 - \phi_1 - \phi_2}$$

$$\phi_0 = \mu(1 - \phi_1 - \phi_2).$$

$$x_t - \mu = \phi_1 (x_{t-1} - \mu) + \phi_2 (x_{t-2} - \mu) + \varepsilon_t$$

$$E[(x_t - \mu)^2] = \phi_1^2 E[(x_{t-1} - \mu)^2] + \phi_2^2 E[(x_{t-2} - \mu)^2] + E[\varepsilon_t^2]$$

$$Var[x_t] = \phi_1^2 Var[x_{t-1}] + \phi_2^2 Var[x_{t-2}] + \sigma_\varepsilon^2$$

$$Var[x_t] = \frac{\sigma_\varepsilon^2}{1 - \phi_1^2 - \phi_2^2} > 0$$

$$\therefore \phi_1^2 + \phi_2^2 < 1 \qquad \sqrt{\phi_1^2 + \phi_2^2} < 1$$

$$(1 - \phi_1 L - \phi_2 L^2) x_t = \phi_0 + \varepsilon_t$$

is referred to as the *characteristic equation* of the AR(p) model. Inverses of the solutions of the equation $\phi(z)$ are referred to as the *characteristic roots*.

Theorem 2.3. *The stationarity condition for an AR(p) model is that all characteristic roots are less than 1 in modulus, or equivalently, all the solutions of the characteristic equation are greater than 1 in modulus.*

Example 2.4. Consider the quarterly growth rate of US gross national product (GNP), seasonally adjusted, from 2Q 1947 to 1Q 2010.

modulus

$a+bi$

$\sqrt{a^2+b^2}$

```
> mydat <- read.table("data1.txt", header = T)
> head(mydat)
  Year Mon Dat VALUE
1 1947   1   1 238.1
...
6 1948   4   1 268.7
> gnp <- ts(log(mydat$VALUE), start = c(1947, 1), freq = 4)
> gnp.g <- ts(diff(gnp), start = c(1947, 2), freq = 4)
> (reg <- arima(gnp.g, order = c(3,0,0), method = "CSS"))
Call:
arima(x = gnp.g, order = c(3, 0, 0), method = "CSS")
Coefficients:
        ar1    ar2     ar3 intercept
     0.4380 0.2089 -0.1542    0.0162
s.e. 0.0615 0.0661  0.0616    0.0012
sigma^2 estimated as 9.441e-05: part log likelihood = 810.18
> eqn <- c(1, -reg$coef[1:3])
> (root <- polyroot(eqn))
[1] 1.632519+0.854625i -1.910083+0.000000i 1.632519-0.854625i
> Mod(root)
[1] 1.842688 1.910083 1.842688
```

$\rightarrow$ solution > 1. $\Rightarrow$ stationarity.

- All the solutions to the characteristic equation are greater than 1 in modulus, satisfying the stationarity of x_t.

3. Order Determination

In application, the order p of an AR model is unknown, so that one must specify it empirically. Consider the following AR models

$$
\begin{aligned}
x_t &= \phi_{0,1} + \phi_{1,1}x_{t-1} + \varepsilon_{t,1} \quad \text{AR(1)} \\
x_t &= \phi_{0,2} + \phi_{1,2}x_{t-1} + \phi_{2,2}x_{t-2} + \varepsilon_{t,2} \quad \text{AR(2)} \\
x_t &= \phi_{0,3} + \phi_{1,3}x_{t-1} + \phi_{2,3}x_{t-2} + \phi_{3,3}x_{t-3} + \varepsilon_{t,3} \quad \text{AR(3)} \\
&\vdots
\end{aligned}
$$

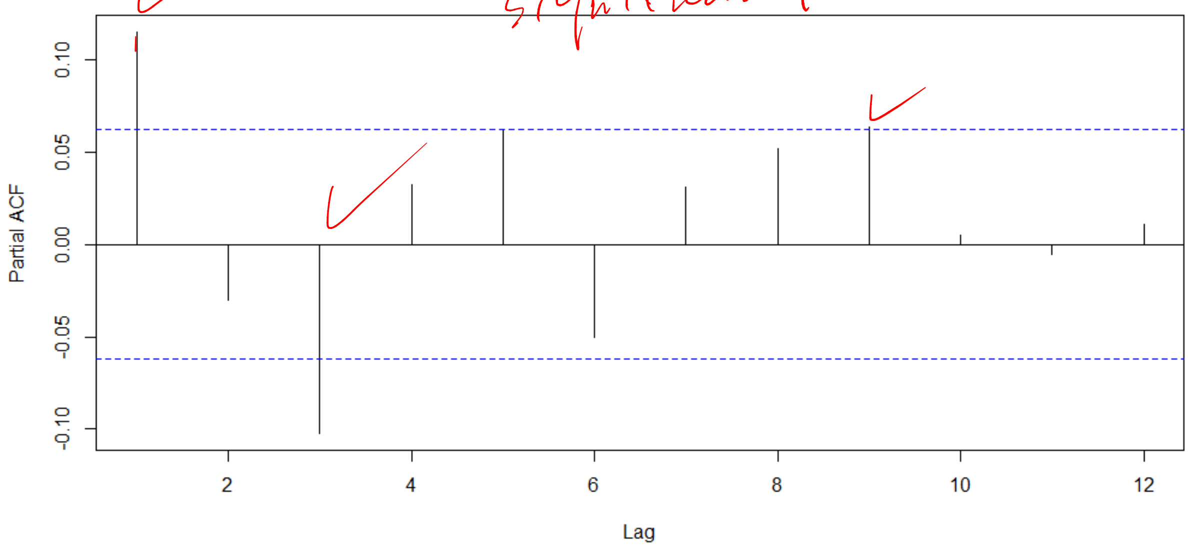
AR(2) 에서 $\phi_{2,2}$가 0 이면 AR(2)를 따르지 X.

The estimate of $\hat{\phi}_{j,j}$ of the jth equation is called the jth *sample partial autocorrelation function* (PACF) of x_t. For an AR(p) model, the pth sample PACF should not be 0, but $\hat{\phi}_{j,j}$ should be close to 0 for all $j > p$; that is, the sample PACF cuts off at lag p.

Example 3.1. Consider the monthly simple returns of CRSP value-weighted index from January 1926 to December 2008.

```
> pacf(vwrtn, lag = 12, main = "")
```

significant.



- The PACF identifies an AR(3) model since the sample PACF $\hat{\phi}_{j,j}$ is significantly different from 0 at the 5% level for $j = 3$ but is not for $j > 3$.

4. Parameter Estimation

Consider an AR(p) model of the form

$$x_t = \phi_0 + \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + \varepsilon_t, \tag{4.1}$$

where $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$. The *conditional least-squares* (LS) method, which starts with the $(p+1)$th observation, is used to estimate the parameters. Conditioning on the first p observations, one takes (4.1) as a multiple linear regression wherein x_t is a dependent variable and $x_{t-1}, \ldots, x_{t-p}$ are a set of independent variables.

AR(p)

- With the estimate of ϕ_i, denoted by $\hat{\phi}_i$, the fitted model is

$$\hat{x}_t = \hat{\phi}_0 + \hat{\phi}_1 x_{t-1} + \cdots + \hat{\phi}_p x_{t-p}$$

  and the residual is

$$\hat{\varepsilon}_t = x_t - \hat{x}_t.$$

conditional.

The LS estimate of σ_ε^2 is

$$\hat{\sigma}_\varepsilon^2 = \frac{\sum_{t=p+1}^T \hat{\varepsilon}_t^2}{T-2p-1}.$$

Remark 4.1. In R, an AR(p) model is in the form

$$(x_t - \mu) = \phi_1(x_{t-1} - \mu) + \cdots + \phi_p(x_{t-p} - \mu) + \varepsilon_t,$$

where $\mu = \phi_0(1 - \phi_1 - \cdots - \phi_p)^{-1}$ is referred to as the `intercept`.

Example 4.2. Consider the monthly simple returns of CRSP value-weighted index from January 1926 to December 2008.

```
> mydat <- read.table("data2.txt", header = T)
> head(mydat)
      date    ibmrtn    vwrtn    ewrtn    sprtn
1 19260130 -0.010381 0.000724 0.023174 0.022472
...
6 19260630  0.068493 0.056888 0.050487 0.043184
> vwrtn <- mydat$vwrtn
> (reg <- arima(vwrtn, order = c(3, 0, 0), method = "CSS"))
Call:
arima(x = vwrtn, order = c(3, 0, 0), method = "CSS")
Coefficients:
        ar1     ar2     ar3 intercept
     0.1148 -0.0188 -0.1043    0.0091
s.e. 0.0315  0.0317  0.0317    0.0017
sigma^2 estimated as 0.002877: part log likelihood = 1500.5
> (phi0 <- as.numeric((1 - coef(reg)[1] - coef(reg)[2] - coef(reg)[3])
     *coef(reg)[4]))
[1] 0.009129547
```

AR(3) 5

Conditional Least square method

intercept = μ

- The fitted model is

$$(x_t - 0.0091) = 0.1148(x_{t-1} - 0.0091) - 0.0188(x_{t-2} - 0.0091) - 0.1043(x_{t-3} - 0.0091) + \hat{\varepsilon}_t.$$

Using $\phi_0 = \mu(1 - \phi_1 - \phi_2 - \phi_3)$, an alternative representation is

$$x_t = 0.0092 + 0.1148x_{t-1} - 0.0188x_{t-2} - 0.1043x_{t-3} + \hat{\varepsilon}_t.$$

ϕ_0 ϕ_1 ϕ_2 ϕ_3 white mean process

5. Model Checking

If the fitted model is adequate, the residual series $\{\hat{\varepsilon}_t\}$ should behave as a white noise. The Ljung-Box test is used to check the closeness of the residuals to a white noise.

H_0 - white noise process

H_1 ≠ "

Theorem 5.1. *Under the null hypothesis that ε_t is a white noise in an AR(p) model, it shows*

$$Q(m) \approx \chi^2_{(m-g)}, \qquad g = p.$$

where g is the number of AR coefficients used in the model.

Example 5.2. Consider the monthly simple returns of CRSP value-weighted index from January 1926 to December 2008.

```
> reg <- arima(vwrtn, order = c(3, 0, 0), method = "CSS")
> e <- reg$residuals
> mytest <- Box.test(e, lag = 12, type = "Ljung")
> (pvalue <- 1 - pchisq(as.numeric(mytest$statistic), 9))
[1] 0.05507046
```

Q(m). m=12.

$m-g$ // $12-3$

cannot reject. (∵ > 0.05).

- The null hypothesis of no residual autocorrelation in the first 12 lags is "barely" not rejected at the 5% level.

```
> tstat <- reg$coef/sqrt(diag(reg$var.coef))
> print(tstat, digits = 3)
  ar1    ar2    ar3 intercept
3.646 -0.591 -3.288     5.370
```

|ar1| > 1.64

|ar2| < 1.64

→ insignificant.

- The coefficient ϕ_2 is not significantly different from 0 at the 5% level. If some of the estimated AR coefficients are not significantly different from 0, then one may remove those insignificant parameters.

```
> (reg1 <- arima(vwrtn, order = c(3, 0, 0), fixed = c(NA, 0, NA, NA),
    method = "CSS"))
Call:
arima(x = vwrtn, order = c(3, 0, 0), fixed = c(NA, 0, NA, NA), method = "CSS")
Coefficients:
        ar1 ar2     ar3 intercept
     0.1126   0 -0.1064    0.0091
s.e. 0.0313   0  0.0315    0.0017
sigma^2 estimated as 0.002878: part log likelihood = 1500.33
> tstat1 <- reg1$coef[-2]/sqrt(diag(reg1$var.coef))
> print(tstat1, digits = 3)
 ar1   ar3 intercept
3.63 -3.37      5.22
> e1 <- reg1$residuals
> mytest1 <- Box.test(e1, lag = 12, type = "Ljung")
> (pvalue <- 1 - pchisq(as.numeric(mytest1$statistic), 10))
[1] 0.07238978
```

ar2 = 0 [illegible].

12 − 2

(∵ ignore ar2).

> 0.05

- The null hypothesis of no residual autocorrelation in the first 12 lags is not rejected at the 5% level. Therefore, the fitted model is

$$(x_t - 5.22) = 3.63(x_{t-1} - 5.22) - 3.37(x_{t-3} - 5.22) + \hat{\varepsilon}_t.$$

ar2 coeff ignore.

6. Forecasting

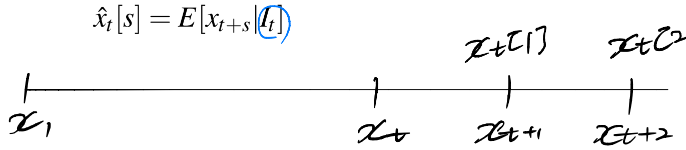
Suppose that one knows $\phi_1, \ldots, \phi_p$ and has a full history up to the time t, $\{x_t, x_{t-1}, x_{t-2}, \ldots\}$. The *s-step ahead forecast* of x_t, denoted by $\hat{x}_t[s]$, is the conditional expectation of x_{t+s} given the information set available at time t, denoted by I_t: i.e.,

$\{x_t, x_{t-1}, x_{t-2}, \cdots\}^c = \phi_1, \cdots, \phi_p$

$\phi_1, \cdots, \phi_p$

$$\hat{x}_t[s] = E[x_{t+s}|I_t]$$

for $s \geq 1$.



- For $x_t \sim \mathrm{AR}(p)$, the 1- step ahead forecast is

$\hat{x}_t[1] = E[X_{t+1}|I_t]$

$x_t = \phi_0 + \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + \varepsilon_t$

$$\begin{aligned}
\hat{x}_t[1] &= E[x_{t+1}|I_t] \\
&= E[\phi_0 + \phi_1 x_t + \cdots + \phi_p x_{t+1-p} + \varepsilon_{t+1}|I_t] \\
&= \phi_0 + \phi_1 x_t + \cdots + \phi_p x_{t+1-p}
\end{aligned}$$

and the associated forecast error is

$$\begin{aligned}
e_t[1] &= x_{t+1} - \hat{x}_t[1] \\
&= \varepsilon_{t+1}.
\end{aligned}$$

Consequently, the variance of $e_t[1]$ is $Var[e_t[1]] = \sigma_\varepsilon^2$. If ε_t is normally distributed, a 95% 1-step ahead interval forecast is $\hat{x}_t[1] \pm 1.96\sigma_\varepsilon$.

- For $x_t \sim \mathrm{AR}(p)$, the 2- step ahead forecast is

$E[\phi_1 x_{t+1} + \phi_2 x_t | I_t] = \phi_1 E[x_{t+1}|I_t] + \phi_2 x_t$

$x_{t+1} \notin I_t$

$$\begin{aligned}
\hat{x}_t[2] &= E[x_{t+2}|I_t] \\
&= E[\phi_0 + \phi_1 x_{t+1} + \cdots + \phi_p x_{t+2-p} + \varepsilon_{t+2}|I_t] \\
&= \phi_0 + \phi_1 E[x_{t+1}|I_t] + \phi_2 x_t + \cdots + \phi_p x_{t+2-p} \\
&= \phi_0 + \phi_1 \hat{x}_t[1] + \phi_2 x_t + \cdots + \phi_p x_{t+2-p}
\end{aligned}$$

and the associated forecast error is

$\hat{x}_t[2]$ require $\hat{x}_t[1]$. $\therefore$ first compute this.

$$\begin{aligned}
e_t[2] &= x_{t+2} - \hat{x}_t[2] \\
&= \phi_1(x_{t+1} - \hat{x}_t[1]) + \varepsilon_{t+2} \\
&= \varepsilon_{t+2} + \phi_1 \varepsilon_{t+1}.
\end{aligned}$$

So, the variance of the forecast error is $Var[e_t[2]] = (1 + \phi_1^2)\sigma_\varepsilon^2$.

$x_t[1] = E[X_{t+1} \mid I_t]$

$= \phi_0 + \underline{\phi_1 x_t} + \cdots + \phi_p x_{t+1-p} + 0$ $\qquad E[\varepsilon_{t+1} \mid I_t] = 0$

$E[\phi_1 x_t \mid I_t] = \phi_1 E[x_t \mid I_t] = \phi_1 x_t$

$x_t = \phi_0 + \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + \varepsilon_t$ $\qquad$ x_t가 I_t 에 포함되어 있음.

$x_{t+1} = \phi_0 + \underline{\phi_1 x_t} + \cdots + \phi_p x_{t+1-p} + \varepsilon_{t+1}$

$e_t[1] = x_{t+1} - x_t[1] = \varepsilon_{t+1}$ $\qquad$ cf) $x_t \sim AR(p)$

$Var[e_t[1]] = Var[\varepsilon_{t+1}] = \sigma_\varepsilon^2$ $\qquad$ $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$

$e_t[1] \sim N(0, \sigma_\varepsilon^2)$ $\qquad$ Assume normality

$x_{t+1} - x_t[1] \sim N(0, \sigma_\varepsilon^2)$ $\qquad$ $\varepsilon_t \sim$ Normal

$\Rightarrow \underline{x_{t+1} \sim N(x_t[1], \sigma_\varepsilon^2)}$ //

95% C.I : $x_t[1] \pm 1.96\, \sigma_\varepsilon$

$e_t[2] = x_{t+2} - \hat{x}_t[2]$

$= \varepsilon_{t+2} + \phi_1 \varepsilon_{t+2}$

$Var[e_t[2]] = Var(\varepsilon_{t+2} + \phi_1 \varepsilon_{t+1})$

$= Var(\varepsilon_{t+2}) + \phi_1^2 \cdot Var(\varepsilon_{t+1})$

$= \sigma_\varepsilon^2 + \phi_1^2 \sigma_\varepsilon^2$

$= (1 + \phi_1^2) \sigma_\varepsilon^2$

$$\hat{x}_t[1] = E[x_{t+1} \mid I_t]$$
$$= E[\phi_0 + \phi_1 x_t + \phi_2 x_{t-1} + \cdots + \phi_p x_{t+1-p} + \varepsilon_{t+1} \mid I_t]$$
$$= \phi_0 + \phi_1 x_t + \phi_2 x_{t-1} + \cdots + \phi_p x_{t+1-p}$$

$$e_t[1] = x_{t+1} - \hat{x}_t[1]$$
$$= \varepsilon_{t+1}$$

$$Var[e_t[1]] = Var[\varepsilon_{t+1}] = \sigma_\varepsilon^2$$

95% $\quad \hat{x}_t[1] \pm 1.96\,\sigma_\varepsilon$

$$\hat{x}_t[2] = E[x_{t+2} \mid I_t]$$
$$= E[\phi_0 + \phi_1 x_{t+1} + \phi_2 x_t + \phi_3 x_{t-1} + \cdots + \phi_p x_{t+2-p} + \varepsilon_{t+2} \mid I_t]$$
$$= \phi_0 + \phi_1 E[x_{t+1} \mid I_t] + \phi_2 x_t + \phi_3 x_{t-1} + \cdots + \phi_p x_{t+2-p}$$
$$= \phi_0 + \phi_1 \hat{x}_t[1] + \phi_2 x_t + \cdots + \phi_p x_{t+2-p}$$

$$e_t[2] = x_{t+2} - \hat{x}_t[2]$$
$$= \phi_1 x_{t+1} - \phi_1 \hat{x}_t[1] + \varepsilon_{t+2}.$$
$$= \phi_1 (x_{t+1} - \hat{x}_t[1]) + \varepsilon_{t+2}$$
$$= \phi_1 e_t[1] + \varepsilon_{t+2}$$
$$= \phi_1 \varepsilon_{t+1} + \varepsilon_{t+2}$$

$$Var[e_t[2]] = Var[\phi_1 \varepsilon_{t+1} + \varepsilon_{t+2}]$$
$$= \phi_1^2 Var[\varepsilon_{t+1}] + 2\phi\, cov[\varepsilon_{t+1}, \varepsilon_{t+2}] + Var[\varepsilon_{t+2}]$$
$$= \phi_1^2 \sigma_\varepsilon^2 + \sigma_\varepsilon^2$$
$$= \sigma_\varepsilon^2 (1 + \phi_1^2)$$

95% $\quad \hat{x}_t[2] \pm 1.96\,\sigma_\varepsilon^2 (1 + \phi_1^2).$

- One sees that $Var[e_t[2]] \geq Var[e_t[1]]$, which means that as the forecast horizon s increases the uncertainty in forecast also increases. The s-step ahead forecast can be computed recursively in a similar manner.

Example 6.1. Consider the monthly simple returns of CRSP value-weighted index from January 1926 to December 2008.

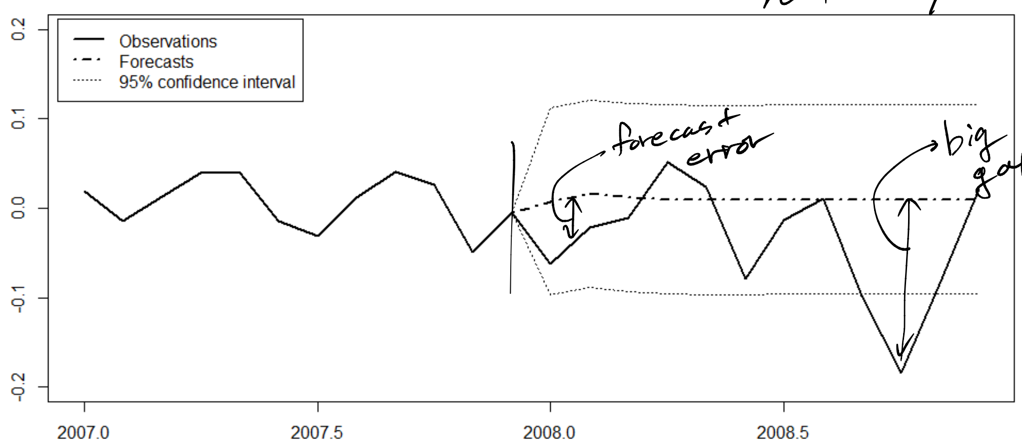
```
> vwrtn <- ts(mydat$vwrtn, start = c(1926, 1), freq = 12)
> vwrtn1 <- window(vwrtn, end = c(2007, 12))
> (reg <- arima(vwrtn1, order = c(3, 0, 0), method = "CSS"))
Call:
arima(x = vwrtn1, order = c(3, 0, 0), method = "CSS")
Coefficients:
        ar1      ar2      ar3  intercept
     0.1024  -0.0201  -0.1090     0.0096
s.e. 0.0317   0.0318   0.0317     0.0017
sigma^2 estimated as 0.002842: part log likelihood = 1488.54
> vwrtn2 <- window(vwrtn, start = c(2007, 1))
> plot(vwrtn2, lwd = 2, xlab = "", ylab = "", ylim = c(-0.2, 0.2))
> x.ahead <- predict(reg, 12)
> pred <- x.ahead$pred
> lines(ts(c(vwrtn2[12], pred), start = c(2007, 12), freq = 12), lty = 4,
    lwd = 2)
> upper <- x.ahead$pred + 1.96*x.ahead$se
> lines(ts(c(vwrtn2[12], upper), start = c(2007, 12), freq = 12), lty = 3)
> lower <- x.ahead$pred - 1.96*x.ahead$se
> lines(ts(c(vwrtn2[12], lower), start = c(2007, 12), freq = 12), lty = 3)
> legend("topleft", c("Observations", "Forecasts", "95% confidence interval"),
    lty = c(1, 4, 3), lwd = c(2, 2, 1), inset = 0.01)
```

vwrtn

vwrtn1 vwrtn2

using this forecast vwrtn2 and compare with real. How precise

not very precise.



forecast error

big gap.