ARMA models

Definition 3.1 An autoregressive model of order p ($AR(p)$), is of the form

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + W_t, \qquad (1)$$

where X_t is stationary, $W_t \sim wn(0, \sigma_w^2)$, and $\phi_1, \phi_2, \dots, \phi_p$ are constants ($\phi_p \neq 0$). If $E(X_t) = \mu \neq 0$, we can replace X_t in (1) by $X_t - \mu$ to get

$$X_t = \alpha + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + W_t, \qquad (2)$$

where $\alpha = \mu(1 - \phi_1 - \dots - \phi_p)$.

Using the backshift operator, model (1) can be written as

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) X_t = \phi(B) X_t = W_t$$

$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$ is called the autoregressive operator.

Consider the AR(1) model $X_t = \phi X_{t-1} + W_t$

$$\begin{aligned} X_t &= \phi(\phi X_{t-2} + W_{t-1}) + W_t \\ &= \phi^2 X_{t-2} + \phi W_{t-1} + W_t \\ &= \dots \\ &= \phi^k X_{t-k} + \phi^{k-1} W_{t-(k-1)} + \dots + \phi W_{t-1} + W_t \end{aligned}$$

If $|\phi| < 1$ and $\sup_t Var(X_t) < \infty$, then $\phi^k X_{t-k} \to 0$ and hence

$$X_t = \sum_{j=0}^{\infty} \phi^j W_{t-j}$$

Since $\sum_{j=0}^{\infty} |\phi^j| = \frac{1}{1-|\phi|} < \infty$, so X_t is a linear process (1.31)

and by (1.32)

$$\begin{aligned} \gamma(h) &= Cov(X_{t+h}, X_t) \\ &= \sigma_w^2 \sum_{j=0}^{\infty} \phi^{j+h} \phi^j = \sigma_w^2 \phi^h \sum_{j=0}^{\infty} \phi^{2j} \\ &= \frac{\sigma_w^2 \phi^h}{1-\phi^2} \end{aligned}$$

and $\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \phi^h, \ h \geq 0$

If we can further assume W_t's are iid (e.g. $W_t \sim N(0, \sigma_w^2)$), then Theorem A.5, A.6, A.7 can then be applied to estimate the distribution of $\bar{X}$, $\hat{\gamma}(h)$ and $\hat{\rho}(h)$ for AR(1) models.

For example, from Theorem A.5, for $X_t = \mu_x + \sum_{j=-\infty}^{\infty} \psi_j w_{t-j}$,

$$\bar{X} \xrightarrow{d} N\left(\mu_x, \frac{V}{n}\right), \text{ where } V = \sigma_w^2 \left(\sum_{j=-\infty}^{\infty} \psi_j\right)^2$$

For general AR(p) model $X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + w_t$, we can write it as

$$\vec{Y}_t = \begin{pmatrix} X_t \\ X_{t-1} \\ \vdots \\ X_{t-p+1} \end{pmatrix} = \begin{pmatrix} \phi_1 & \phi_2 & \cdots & \phi_p \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & 0 \cdots & 0 \\ & & \ddots & \vdots \\ & & 1 & 0 \end{pmatrix} \begin{pmatrix} X_{t-1} \\ \vdots \\ \vdots \\ X_{t-p} \end{pmatrix} + \begin{pmatrix} w_t \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \Phi \vec{Y}_{t-1} + \vec{w}_t$$

$$= \Phi(\Phi \vec{Y}_{t-2} + \vec{w}_{t-1}) + \vec{w}_t$$

$$= \Phi^2 \vec{Y}_{t-2} + \Phi \vec{w}_{t-1} + \vec{w}_t$$

$$= \dots$$

$$= \Phi^k \vec{Y}_{t-k} + \sum_{j=0}^{k-1} \Phi^j \vec{w}_{t-j}$$

Let $\lambda_1, \dots, \lambda_p$ be the eigenvalues of Φ and $q_1, \dots, q_p$ be the corresponding eigenvectors. Suppose q_i, $i = 1, \dots, p$, are linearly independent, then we have the eigenvalue decomposition for $\Phi = Q \Lambda Q^{-1}$ where $Q = (q_1, \dots, q_p)$ and $\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_p)$, then $\Phi^j = Q \Lambda^j Q^{-1}$ and hence if $\max_j |\lambda_j| < 1$, $\Phi^k \vec{Y}_{t-k} \longrightarrow 0$ and we have $\vec{Y}_t = \sum_{j=0}^{\infty} \Phi^j \vec{w}_{t-j}$ or $X_t = \sum_{j=0}^{\infty} (\Phi^j)_{11} w_{t-j}$

Back to the AR(1) model, $X_t = \phi X_{t-1} + w_t$, if $|\phi| > 1$, then ϕ^k does not tend to 0. In such case, X_t is still stationary. (Note that if $\phi = 1$, $X_t = X_{t-1} + w_t$ is a random walk, which is non-stationary.

For $|\phi| > 1$, consider $X_{t+1} = \phi X_t + w_{t+1}$

$$\Rightarrow X_t = \phi^{-1} X_{t+1} - \phi^{-1} w_{t+1}$$

$$= \phi^{-1}(\phi^{-1} X_{t+2} - \phi^{-1} w_{t+2}) - \phi^{-1} w_{t+1}$$

$$= \dots$$

$$= \phi^{-k} X_{t+k} - \sum_{j=1}^{k-1} \phi^{-j} w_{t+j}$$

As $|\phi^{-1}| < 1$, we have $\phi^{-k} X_{t+k} \to 0$ and $X_t = -\sum_{j=1}^{\infty} \phi^{-j} w_{t+j}$. Since $\sum_{j=1}^{\infty} |\phi^{-j}| < \infty$, X_t is a linear process, which is stationary.

Recall that a linear process $X_t = \mu + \sum_{j=-\infty}^{\infty} \psi_j w_{t-j}$ is called causal if $\psi_j = 0$ for $j<0$, i.e. $X_t = \mu + \sum_{j=0}^{\infty} \psi_j w_{t-j}$. Therefore, for $|\phi|>1$, $X_t = -\sum_{j=1}^{\infty} \phi^{-j} w_{t+j} = \sum_{j=-\infty}^{-1} (-\phi^{j}) w_{t-j}$ is stationary, but not causal. 3

Non-causal linear process is not preferred because it means X_t depend on the errors $(w_{t+j}, j>0)$ in the future. However, for the case of AR(1) with $|\phi|>1$, we can find an equivalent causal process. Suppose $\{X_1, X_2, \dots, X_n\}$ is the observed time series. Define y_t such that $y_1 = X_n$, $y_2 = X_{n-1}, \dots, y_n = X_1$, then $X_t = \phi X_{t-1} + w_t$

$$\Rightarrow \quad y_{s-1} = \phi y_s + w_{s-1}, \quad s = n+2-t$$

$$\Rightarrow \quad y_s = \phi^{-1} y_{s-1} + v_s$$

where $v_s = -\phi^{-1} w_{s-1} \sim wn(0, \phi^{-2}\sigma_w^2)$. Now $|\phi^{-1}|<1$ and so

$$y_s = \sum_{j=0}^{\infty} \phi^{-j} v_{s-j}$$ which is a causal linear process

Since $\{X_t\}, t=1,\dots,n$ and $\{y_s\}, s=1,\dots,n$ are the same set of values, the analysis results for $\{y_s\}$ are also applied for $\{X_t\}$, e.g. $\bar{y}$, $\gamma_y(h)$ et

---

Problem 3.3 (a) For $X_t = \phi X_{t-1} + w_t$, $|\phi|>1$, $w_t \sim iid\ N(0, \sigma_w^2)$

Show that $E(X_t) = 0$ and $\gamma_x(h) = \sigma_w^2 \phi^{-2} \phi^{-h} / (1-\phi^{-2})$ for $h \geq 0$

$\because X_t = \sum_{j=-\infty}^{-1} (-\phi^{j}) w_{t-j}$, $\therefore E(X_t) = 0$ and

by (1.32) with $\psi_j = -\phi^j$ if $j<0$, $\psi_j = 0$ if $j \geq 0$

$$\gamma_x(h) = \sigma_w^2 \sum_{j=-\infty}^{\infty} \psi_{j+h}\psi_j = \sigma_w^2 \sum_{j=-\infty}^{-1} \psi_{j+h}\psi_j = \sigma_w^2 \sum_{j=-\infty}^{-1} \psi_j \psi_{j-h} = \sigma_w^2 \sum_{j=-\infty}^{-1} \phi^j \phi^{j-h} = \sigma_w^2 \phi^{-h} \sum_{j=1}^{\infty} (\phi^{-1})$$

$$= \frac{\sigma_w^2 \phi^{-2} \phi^{-h}}{1-\phi^2}$$

(b) For $y_t = \phi^{-1} y_{t-1} + v_t$, $v_t \sim iid\ N(0, \sigma_w^2 \phi^{-2})$

$y_t = \sum_{j=0}^{\infty} \phi^{-j} v_{t-j}$, $\therefore E y_t = 0$ and

$$\gamma_y(h) = Var(v_t) \sum_{j=0}^{\infty} \phi^{-(j+h)} \phi^{-j}$$

$$= \sigma_w^2 \phi^{-2} \phi^{-h} \frac{1}{1-\phi^{-2}}$$

Importance! However, in practical sense, AR(1) model with $|\phi|>1$ is NOT stationary, but explosive. Recall the argument of stationarity of $X_t = \phi X_{t-1} + W_t$ with $|\phi|>1$, we assume $W_1, \dots, W_n$ are generated at the same time so that generating X_t from W_1 to W_n using $X_t = \phi X_{t-1} + W_t$ and generating Y_s from W_n to W_1 using $Y_s = \phi^{-1} Y_{s-1} + V_s$, $V_s = -\phi^{-1} W_{s-1}$, $s = n+2-t$, are the same. However, in practice, W_t's are generated in one direction only, which is $t = 1, 2, 3, \dots$ Therefore, AR(1) with $|\phi|>1$ is not stationary in practical sense.

For AR(p) model, consider $\vec{Y}_t = \Phi \vec{Y}_{t-1} + \vec{W}_t = \Phi^k \vec{Y}_{t-k} + \sum_{j=0}^{k-1} \Phi^j \vec{W}_{t-j}$

If $\max_j |\lambda_j| < 1$, then we have seen that $X_t = \sum_{j=0}^{\infty} (\Phi^j)_{11} W_{t-j}$, which is causal stationary.

If $\max_j |\lambda_j| > 1$, the trick for AR(1) may not be applied. Note that

$\Phi = Q \Lambda Q^{-1} \Rightarrow \Phi^{-1} = Q \Lambda^{-1} Q^{-1}$

$$\vec{Y}_{t+1} = \Phi \vec{Y}_t + \vec{W}_{t+1} \Rightarrow \vec{Y}_t = \Phi^{-1} \vec{Y}_{t+1} - \Phi^{-1} \vec{W}_{t+1}$$
$$= \Phi^{-k} \vec{Y}_{t+k} - \sum_{j=1}^{k-1} \Phi^{-j} \vec{W}_{t+j}$$

Unlike AR(1) that $|\phi|>1 \Rightarrow |\phi^{-1}|<1$, we don't have $\max_j |\lambda_j^{-1}| < 1$ in general for $\max_j |\lambda_j| > 1$. To get the linear process for $X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + W_t$, we can consider the autoregressive operator

$$\phi(B) = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)$$

To illustrate the idea, we start from an simple example

$$X_t = \tfrac{1}{4} X_{t-2} + W_t \Rightarrow \phi(B) = (1 - \tfrac{1}{4}B^2) = (1 - \tfrac{1}{2}B)(1 + \tfrac{1}{2}B)$$

$$\therefore \quad \phi(B) X_t = W_t \Leftrightarrow (1 - \tfrac{1}{2}B)(1 + \tfrac{1}{2}B) X_t = W_t$$

let $U_t = (1 + \tfrac{1}{2}B) X_t = X_t + \tfrac{1}{2} X_{t-1}$, then $(1 - \tfrac{1}{2}B) U_t = W_t$

$$\Leftrightarrow \quad U_t = \tfrac{1}{2} U_{t-1} + W_t = \sum_{j=0}^{\infty} (\tfrac{1}{2})^j W_{t-j}$$

Now $(1+\frac{1}{2}B)X_t = U_t$

$$\Rightarrow \quad X_t = -\frac{1}{2}X_{t-1} + U_t = \sum_{j=0}^{\infty}\left(-\frac{1}{2}\right)^j U_{t-j}$$

$$= \sum_{j=0}^{\infty}\left(-\frac{1}{2}\right)^j\left(\sum_{\ell=0}^{\infty}\left(\frac{1}{2}\right)^\ell W_{t-j-\ell}\right)$$

$$= \sum_{m=0}^{\infty}\left(\sum_{\substack{j,\ell\geq 0\\ j+\ell=m}}\left(-\frac{1}{2}\right)^j\left(\frac{1}{2}\right)^\ell\right) W_{t-m}$$

In general, we can decompose the polynomial $\phi(z) = (1-\phi_1 z - \phi_2 z^2 - \dots \phi_p z^p)$

$$= \prod_{i=1}^{p}(1 - r_i^{-1} z)$$

where r_i, $i=1,\dots,p$, are the roots of $\phi(z)$. We can see that AR(p) model is causal stationary if and only if $|r_i|>1$ for all i.

Inverse operator $\phi^{-1}(B)$

A general technique for finding the coefficients of the linear process is that of matching coefficients. Take AR(1) $X_t = \phi X_{t-1} + W_t$, or $\phi(B)X_t = W_t$ with $\phi(B) = 1-\phi B$, as an example.

Set $X_t = \sum_{j=0}^{\infty}\psi_j W_{t-j} = \psi(B)W_t$ (Here suppose $|\phi|<1$)

then $\phi(B)X_t = W_t \Rightarrow \phi(B)\psi(B)W_t = W_t$

$$\Rightarrow (1-\phi B)(1+\psi_1 B + \psi_2 B^2 + \dots + \psi_j B^j + \dots) = 1$$

$$\Rightarrow 1 + (\psi_1 - \phi)B + (\psi_2 - \psi_1\phi)B^2 + \dots + (\psi_j - \psi_{j-1}\phi)B^j + \dots = 1$$

$$\Rightarrow \psi_1 = \phi, \quad \psi_2 = \psi_1\phi, \dots, \psi_j = \psi_{j-1}\phi, \dots$$

$$\Rightarrow \psi_j = \phi^j, \quad j=1,2,\dots$$

The fact that $\phi(B)\psi(B) = 1$ makes $\psi(B)$ look like the inverse of $\phi(B)$. $\phi(B)X_t = W_t \Rightarrow X_t = \phi^{-1}(B)W_t$

Actually, as we have seen so far, $\phi(B)$ performs similar to a polynomial. If we consider the polynomial $\phi(z) = 1-\phi z$, then we know that

$$\phi^{-1}(z) = \frac{1}{1-\phi z} = 1 + \phi z + \phi^2 z^2 + \dots + \phi^j z^j + \dots, \quad |z|\leq 1$$

While we may treat the backshift operator B, as a complex number z, they are different

Definition 3.3 The moving average model of order q, $MA(q)$, is $\quad X_t = W_t + \theta_1 W_{t-1} + \theta_2 W_{t-2} + \dots + \theta_q W_{t-q}$,
where $W_t \sim wn(0, \sigma_w^2)$ and $\theta_1, \dots, \theta_q \neq 0$ are parameters.

Define the moving average operator $\theta(B) = 1 + \theta_1 B + \dots + \theta_q B^q$, then $MA(q)$ model is $\quad X_t = \theta(B) W_t$

By definition, $MA(q)$ model is causal stationary for any values of $\theta_1, \dots, \theta_q$.

Example 3.5 Consider $MA(1)$ model $X_t = W_t + \theta W_{t-1}$, we can directly compute $\quad E(X_t) = 0 \quad \gamma(h) = \begin{cases} (1+\theta^2)\sigma_w^2 & h=0 \\ \theta\sigma_w^2 & h=1 \\ 0 & h>1 \end{cases}$

or we can consider $\quad X_t = \sum_{j=-\infty}^{\infty} \psi_j W_{t-j} \quad$ with $\psi_0 = 1 \quad \psi_1 = \theta$
$\psi_j = 0$ otherwise

then by (1.32)

$$\gamma(h) = \sigma_w^2 \sum_{j=-\infty}^{\infty} \psi_{j+h}\psi_j = \sigma_w^2 (\psi_h \psi_0 + \psi_{h+1}\psi_1)$$
$$= \sigma_w^2 (\psi_h + \theta \psi_{h+1})$$

Example 3.6 Note that we get the same $\gamma(h)$ (and hence $\rho(h)$) for $(\theta, \sigma_w^2) = (5, 1)$ and $(\frac{1}{5}, 25)$

We can check that $\quad X_t = W_t + \frac{1}{5} W_{t-1}, \quad W_t \sim iid\ N(0, 25)$
and $\quad Y_t = V_t + 5 V_{t-1}, \quad V_t \sim iid\ N(0, 1)$

are the same for all finite distributions. Since we can only observe X_t or Y_t, but not W_t or V_t, so we cannot distinguish between the models.

For convenience, we will choose the one with an infinite AR representation

Consider $\quad W_t = -\theta W_{t-1} + X_t$

$$= \sum_{j=0}^{\infty} (-\theta)^j X_{t-j} \quad \text{for } |\theta| < 1$$

This process is called an invertible process. We will choose the model with $\sigma_w^2 = 25$ and $\theta = \frac{1}{5}$ because it is invertible.

Definition 3.5 | A time series $\{X_t; t = 0, \pm 1, \pm 2, \dots\}$ is ARMA(p,q) if it is stationary and

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + W_t + \theta_1 W_{t-1} + \dots + \theta_q W_{t-q}$$

or $\phi(B) X_t = \theta(B) W_t$ $\quad (\phi_p \neq 0, \ \theta_q \neq 0)$

where $\phi(z) = 1 - \phi_1 z - \dots - \phi_p z^p$ is the AR polynomial and

$\theta(z) = 1 + \theta_1 z + \dots + \theta_q z^q$ is the MA polynomial

$W_t \sim wn(0, \sigma_w^2)$ and the parameters p and q are called the autoregressive and the moving average orders, respectively.

Example 3.7 | Parameter Redundancy

Consider $X_t = W_t$, if we apply the same operator $\eta(B) = 1 - 0.5B$ on both sides, we get

$$(1 - 0.5B) X_t = (1 - 0.5B) W_t$$

$$\Leftrightarrow \quad X_t = 0.5 X_{t-1} - 0.5 W_{t-1} + W_t$$

which is a ARMA(1,1)

To avoid redundancy, we require $\phi(z)$ and $\theta(z)$ have no common factors

$\Rightarrow$ If we find that there are roots of $\hat{\phi}(z)$ and $\hat{\theta}(z)$ that are close, we should consider ARMA models with lower orders p, q.

Property 3.1 | An ARMA(p,q) model $\phi(B) X_t = \theta(B) W_t$ is causal, i.e. X_t can be written as $X_t = \sum_{j=0}^{\infty} \psi_j W_{t-j} = \psi(B) W_t$ with $\sum_{j=0}^{\infty} |\psi_j| < \infty$, if and only if $\phi(z) \neq 0$ for $|z| \leq 1$, i.e. the roots of $\phi(z)$ lie outside the unit circle. (We set $\psi_0 = 1$)

Property 3.2 | An ARMA(p,q) model is invertible, i.e. W_t can be written as $W_t = \sum_{j=0}^{\infty} \pi_j X_{t-j} = \pi(B) X_t$ with $\sum_{j=0}^{\infty} |\pi_j| < \infty$, $(\pi_0 = 1)$ if and only if $\theta(z) \neq 0$ for $|z| \leq 1$

Example 3.8 | Consider $X_t = 0.4X_{t-1} + 0.45X_{t-2} + W_t + W_{t-1} + 0.25W_{t-2}$

$$\Leftrightarrow \quad (1-0.4B-0.45B^2)X_t = (1+B+0.25B^2)W_t$$

Note that $\phi(B) = 1-0.4B-0.45B^2 = (1+0.5B)(1-0.9B)$

$\theta(B) = (1+B+0.25B^2) = (1+0.5B)^2$

$$\therefore \quad (1+0.5B)(1-0.9B)X_t = (1+0.5B)^2 W_t$$

$$\Leftrightarrow \quad (1-0.9B)X_t = (1+0.5B)W_t$$

or

$$X_t = 0.9X_{t-1} + W_t + 0.5W_{t-1}$$

which is a ARMA(1,1) model.

The model is causal as $\phi(z) = 1-0.9z = 0 \Leftrightarrow z = \frac{10}{9} > 1$

and the model is invertible as $\theta(z) = 1+0.5z = 0 \Leftrightarrow z = -2$ $(\because |z|>1)$

To write $X_t = \sum_{j=0}^{\infty}\psi_j W_{t-j}$, consider

$$(1-0.9B)\left(\sum_{j=0}^{\infty}\psi_j B^j\right)W_t = (1+0.5B)W_t$$

$$\Rightarrow \quad (1-0.9B)(1+\psi_1 B+\psi_2 B^2+\dots+\psi_j B^j+\dots) = 1+0.5B$$

$$\Rightarrow \quad 1+(\psi_1-0.9)B+(\psi_2-0.9\psi_1)B^2+\dots+(\psi_j-0.9\psi_{j-1})B^j+\dots = 1+0.5B$$

Matching the coefficients,

$\psi_1 - 0.9 = 0.5 \Rightarrow \psi_1 = 1.4$ $\qquad \psi_j - 0.9\psi_{j-1} = 0 \Rightarrow \psi_j = 0.9\psi_{j-1}$

$= 0.9^{j-1}\psi_1 = (1.4)0.9^{j-1}$

$$\therefore \quad X_t = W_t + 1.4\sum_{j=1}^{\infty}0.9^{j-1}W_{t-j}$$

We can also do the same for $W_t = \sum_{j=0}^{\infty}\pi_j X_{t-j}$ to get

$$X_t = 1.4\sum_{j=1}^{\infty}(-0.5)^{j-1}X_{t-j} + W_t$$

For general case, $\phi(B)X_t = \theta(B)W_t$, how to find ψ_j such that $X_t = \sum_{j=0}^{\infty}\psi_j W_{t-j}$ (given that the zeros of $\phi(z)$ are outside the unit circle)

By matching the coefficients, $\phi(B)\psi(B)W_t = \theta(B)W_t$

$$(1-\phi_1 B-\phi_2 B^2-\dots)(\psi_0+\psi_1 B+\psi_2 B^2+\dots)=(1+\theta_1 B+\theta_2 B^2+\dots)$$ 9

$$\Rightarrow \quad \psi_0+\psi_1 B+\psi_2 B^2+\psi_3 B^3+\psi_4 B^4+\dots$$
$$-\phi_1\psi_0 B-\phi_1\psi_1 B^2-\phi_1\psi_2 B^3-\dots \qquad = 1+\theta_1 B+\theta_2 B^2+\theta_3 B^3+\dots$$
$$-\phi_2\psi_0 B^2-\phi_2\psi_1 B^3-\dots$$

The first few values are $\psi_0=1$

$$\psi_1-\phi_1\psi_0=\theta_1$$
$$\psi_2-\phi_1\psi_1-\phi_2\psi_0=\theta_2$$
$$\psi_3-\phi_1\psi_2-\phi_2\psi_1-\phi_3\psi_0=\theta_3$$
$$\vdots$$

For ARMA(p, q) model $\phi_j=0$ for $j>p$ and $\theta_j=0$ for $j>q$

$\therefore$ For $j>q$ (or $j\geq q+1$), and $j\geq p$ (i.e. $j\geq \max(p,q+1)$)
$$\psi_j-\sum_{k=1}^{p}\phi_k\psi_{j-k}=0 \qquad (3.40)$$

Otherwise, if $0\leq j<\max(p,q+1)$, $\psi_j-\sum_{k=1}^{j}\phi_k\psi_{j-k}=\theta_j$ (3.41)

Given $\phi_1,\dots,\phi_p$ and $\theta_1,\dots,\theta_q$, how to solve $\psi_1,\psi_2,\dots$ from (3.40) and (3.41)?

While there are $\max(p,q+1)$ equations for $\psi_j-\sum_{k=1}^{j}\phi_k\psi_{j-k}=\theta_j$, and hence can be solved directly; the problem comes from $\psi_j-\sum_{k=1}^{p}\phi_k\psi_{j-k}=0$

Suppose we know $\psi_0,\psi_1,\dots\psi_{m-1}$, we want to find ψ_j for $j\geq m$

$m=\max(p,q+1)\geq p$

For $m=1$ ($p=1$), we know $\psi_0=c\ (=1)$, $\psi_j=\phi_1\psi_{j-1}$ for $j=1,2,\dots$
$$=\phi_1^j\psi_0$$

Let z_0 be the root of $\phi(z)=1-\phi_1 z \Rightarrow z_0=\phi_1^{-1}$

$$\therefore \psi_j=(z_0^{-1})^j c \qquad \text{for } j\geq 1$$

For $m=2$ ($p=2$), we have ψ_0 and ψ_1,

$$\psi_j - \phi_1 \psi_{j-1} - \phi_2 \psi_{j-2} = 0 \qquad \text{for } j=2,3,\dots$$

Let z_1 and z_2 be the roots of $\phi(z) = 1-\phi_1 z - \phi_2 z^2$

Then, if $z_1 \neq z_2$, we have $\psi_j = c_1 z_1^{-j} + c_2 z_2^{-j}$

Checking: $(c_1 z_1^{-j} + c_2 z_2^{-j}) - \phi_1 (c_1 z_1^{-(j-1)} + c_2 z_2^{-(j-1)}) - \phi_2 (c_1 z_1^{-(j-2)} + c_2 z_2^{-(j-2)})$

$$= c_1 z_1^{-j} (1-\phi_1 z_1 - \phi_2 z_1^2) + c_2 z_2^{-j} (1-\phi_1 z_2 - \phi_2 z_2^2) = 0$$

c_1 and c_2 can be determined by

$$\psi_1 = c_1 z_1^{-1} + c_2 z_2^{-1} \qquad \psi_0 = c_1 + c_2$$

---

if $z_1 = z_2 = z_0$, i.e. $1-\phi_1 z - \phi_2 z^2 = (1-z_0^{-1} z)^2 = 1 - 2z_0^{-1} z + z_0^{-2} z^2$

$$\Rightarrow \phi_1 = 2 z_0^{-1} \qquad \phi_2 = -z_0^{-2}$$

the solution is $\psi_j = z_0^{-j} (c_1 + c_2 j)$

Checking: $z_0^{-j}(c_1 + c_2 j) - \phi_1 z_0^{-(j-1)} (c_1 + c_2(j-1)) - \phi_2 z_0^{-(j-2)} (c_1 + c_2(j-2))$

$$= z_0^{-j} (c_1 + c_2 j) [1 - \phi_1 z_0 - \phi_2 z_0^2] + c_2 z_0^{-j+1} (\phi_1 + 2\phi_2 z_0)$$

$$= 0$$

Again, c_1 and c_2 can be determined by

$$\psi_1 = z_0^{-1} (c_1 + c_2) \qquad \psi_0 = c_1$$

In general, for $\psi_j - \phi_1 \psi_{j-1} - \dots - \phi_p \psi_{j-p} = 0$ for $j = p, p+1, \dots$

Let $z_1, z_2, \dots, z_r$ be the roots of $\phi(z) = 1 - \phi_1 z - \dots - \phi_p z^p$ with multiplicity $m_1, m_2, \dots, m_r$ respectively such that $m_1 + m_2 + \dots + m_r = p$

The solution is $\psi_j = z_1^{-j} P_1(j) + z_2^{-j} P_2(j) + \dots + z_r^{-j} P_r(j)$

where $P_i(j)$ for $i=1,2,\dots,r$ is a polynomial in j of degree $m_i - 1$

Example 3.12 $X_t = 0.9X_{t-1} + 0.5w_{t-1} + w_t$

$m = \max(p, q+1) = 2$, $p=1$, $\phi(z) = 1 - 0.9z$ $\quad \theta(z) = 1 + 0.5z$

$\psi_0 = 1 \quad \psi_1 - \phi_1\psi_0 = \theta_1 \Rightarrow \psi_1 = 0.9(1) + 0.5 = 1.4$

$\psi_j - \phi_1\psi_{j-1} = 0$ for $j = 2, 3, \ldots$ (Note that the initial value is ψ_1)

The root for $\phi(z)$ is 0.9^{-1}. Hence $\psi_j = 0.9^j P_1(j)$

Since $m_1 = 1$, $P_1(j) = c$, from $\psi_1 = 0.9c \Rightarrow c = 1.4(0.9)^{-1}$

$\therefore \psi_j = 1.4(0.9)^{j-1}$ for $j \geq 1$ as we saw in Example 3.8

Example 3.10 Consider an AR(2) model $X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + w_t$

Suppose X_t can be written as $X_t = \sum_{j=0}^{\infty} \psi_j w_{t-j}$ (i.e. causal)

then consider $E(X_t X_{t-h}) = \phi_1 E(X_{t-1}X_{t-h}) + \phi_2 E(X_{t-2}X_{t-h}) + E(w_t X_{t-h})$

for $h > 0$ $\Rightarrow \gamma(h) = \phi_1\gamma(h-1) + \phi_2\gamma(h-2)$

Dividing by $\gamma(0)$ $\Rightarrow \rho(h) - \phi_1\rho(h-1) - \phi_2\rho(h-2) = 0$, $h = 1, 2, \ldots$

Put $h = 1$, $\Rightarrow \rho(1) - \phi_1\rho(0) - \phi_2\rho(-1) = 0$

$\Rightarrow \rho(-1) = \rho(1) = \phi_1/(1-\phi_2)$ (Note that $\rho(0) = 1$)

Let z_1 and z_2 be the roots of $\phi(z) = 1 - \phi_1 z - \phi_2 z^2$. $|z_1| > 1$, $|z_2| > 1$

Case (i), if z_1 and z_2 are real and distinct,

then $\rho(h) = c_1 z_1^{-h} + c_2 z_2^{-h} \longrightarrow 0$ exponentially fast as $h \to \infty$

Case (ii), if $z_1 = z_2 (= z_0)$, then

$\rho(h) = z_0^{-h}(c_1 + c_2 h) \longrightarrow 0$ exponentially fast as $h \to \infty$

Case (iii), $z_1 = a + ib$ (complex) and hence $z_2 = a - ib$

$\Rightarrow z_1 = \sqrt{a^2+b^2}\left(\frac{a}{\sqrt{a^2+b^2}} + i\frac{b}{\sqrt{a^2+b^2}}\right)$ $\qquad z_2 = \sqrt{a^2+b^2}\left(\frac{a}{\sqrt{a^2+b^2}} - i\frac{b}{\sqrt{a^2+b^2}}\right)$

$= |z_1|(\cos\theta + i\sin\theta)$ $\qquad = |z_1|(\cos\theta - i\sin\theta)$

$= |z_1| e^{i\theta}$ $\qquad = |z_1| e^{-i\theta}$

Since $z_1 \neq z_2$, we have $\rho(h) = c_1 z_1^{-h} + c_2 z_2^{-h}$

$$= c_1 |z_1|^{-h} e^{ih\theta} + c_2 |z_1|^{-h} e^{-ih\theta}$$

Let $c_1 = \frac{a}{2} e^{ib}$, then $c_1 |z_1|^{-h} e^{ih\theta} = \frac{a}{2} |z_1|^{-h} e^{i(h\theta + b)}$

$$= \frac{a}{2} |z_1|^{-h} [\cos(h\theta + b) + i \sin(h\theta + b)]$$

Since $\rho(h)$ is a real number, the imaginary terms must be removed, and hence $c_2 |z_1|^{-h} e^{-ih\theta}$ must be equal to $\frac{a}{2} |z_1|^{-h} e^{-i(h\theta + b)}$

$$= \frac{a}{2} |z_1|^{-h} [\cos(h\theta + b) - i \sin(h\theta + b)]$$

$\therefore \rho(h) = a |z_1|^{-h} \cos(h\theta + b) \to 0$ exponentially as $h \to \infty$

$\rho(h)$ will look periodic due to the $\cos(h\theta + b)$ term. For example, for $\theta = \frac{2\pi}{12}$ in Example 3.11, $\cos((12m + h)\theta + b)$

$$= \cos\left(12m\left(\frac{2\pi}{12}\right) + h\theta + b\right)$$

$$= \cos(2m\pi + h\theta + b) = \cos(h\theta + b)$$

for all integer m. Therefore, we see a cycle for every 12 points.

For general ARMA(p, q) model, $\phi(B) X_t = \theta(B) w_t$ such that $X_t = \sum_{j=0}^{\infty} \psi_j w_{t-j}$. We can compute $\gamma(h) = \mathrm{Cov}(X_{t+h}, X_t) = \sigma_w^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+h}$ and $\rho(h) = \frac{\gamma(h)}{\gamma(0)}$. To get an explicit expression for $\rho(h)$, consider

$$\gamma(h) = \mathrm{cov}(X_{t+h}, X_t) = \mathrm{Cov}\left(\sum_{j=1}^{p} \phi_j X_{t+h-j} + \sum_{j=0}^{q} \theta_j w_{t+h-j}, X_t\right)$$

$$= \sum_{j=1}^{p} \phi_j \gamma(h-j) + \sum_{j=0}^{q} \theta_j \mathrm{Cov}(w_{t+h-j}, X_t)$$

Note that $\mathrm{Cov}\left(w_{t+h-j}, \sum_{k=0}^{\infty} \psi_k w_{t-k}\right) = \mathrm{Cov}(w_{t+h-j}, \psi_{j-h} w_{t-(j-h)}) = \sigma_w^2 \psi_{j-h}$

$\therefore \gamma(h) = \sum_{j=1}^{p} \phi_j \gamma(h-j) + \sigma_w^2 \sum_{j=h}^{q} \theta_j \psi_{j-h}$ (Note that $\psi_{j-h} = 0$ if $j < h$)

$\therefore \gamma(h) - \phi_1 \gamma(h-1) - \dots - \phi_p \gamma(h-p) = 0$ if $h \geq \max(p, q+1)$

and $\gamma(h) - \sum_{j=1}^{p} \phi_j \gamma(h-j) = \sigma_w^2 \sum_{j=h}^{q} \theta_j \psi_{j-h}$ if $h < \max(p, q+1)$

to ensure at least p initial values.

Example 3.14 $x_t = \phi x_{t-1} + \theta w_{t-1} + w_t$, $|\phi|<1$

$$\gamma(h) = \sum_{j=1}^{p} \phi_j \gamma(h-j) + \sigma_w^2 \sum_{j=h}^{q} \theta_j \psi_{j-h} \quad (3.48)$$

$$\Rightarrow \gamma(h) = \phi\gamma(h-1) + \sigma_w^2 \theta \quad \text{if } h=1 \quad (\psi_0 = 1)$$

$$\gamma(h) - \phi\gamma(h-1) = 0 \quad \text{for } h=2,3,\ldots$$

$$\Rightarrow \gamma(h) = \phi\gamma(h-1) = \phi^{h-1}\gamma(1)$$

Put $h=1$ in (3.48), $\gamma(1) = \phi\gamma(0) + \sigma_w^2\theta$

$h=0$ in (3.48), $\gamma(0) = \phi\gamma(1) + \sigma_w^2(\theta_0\psi_0 + \theta\psi_1)$

Recall that $\psi_j - \sum_{k=1}^{j} \phi_k \psi_{j-k} = \theta_j$ for $j < \max(p, q+1)$

$$\therefore \psi_1 = \phi_1\psi_0 + \theta_1 = \phi + \theta \Rightarrow \gamma(0) = \gamma(1) + \sigma_w^2(1+\theta\phi+\theta^2)$$

$$\Rightarrow \gamma(0) = \phi^2\gamma(0) + \sigma_w^2\theta\phi + \sigma_w^2(1+\theta\phi+\theta^2)$$

$$\Rightarrow \gamma(0) = \sigma_w^2 \frac{1+2\theta\phi+\theta^2}{1-\phi^2}$$

$$\Rightarrow \gamma(1) = \sigma_w^2 \frac{(1+\theta\phi)(\phi+\theta)}{1-\phi^2}$$

$$\Rightarrow \gamma(h) = \phi^{h-1}\gamma(1) = \sigma_w^2 \frac{(1+\theta\phi)(\phi+\theta)}{1-\phi^2}\phi^{h-1}$$

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \frac{(1+\theta\phi)(\phi+\theta)}{1+2\theta\phi+\theta^2}\phi^{h-1}$$

Partial autocorrelation function (PACF)

For MA(q) models, the ACF, $\gamma(h)$, will be zero for lags greater than q. Therefore, we can choose appropriate q from the ACF plot.

However, it is not true for AR(p) models.

For example, let say we want to choose between AR(1) $x_t = \phi_1 x_{t-1} + w_t$ and AR(2) $x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + w_t$ model. Even if $\phi_2 = 0$,

$$\gamma_x(2) = Cov(x_t, x_{t-2}) = Cov(\phi_1 x_{t-1} + w_t, x_{t-2})$$

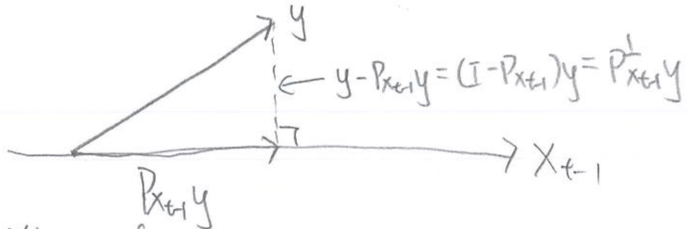
$$= Cov(\phi_1^2 x_{t-2} + \phi w_{t-1} + w_t, \sum_{j=0}^{\infty} \psi_j w_{t-2-j}) = \phi_1^2 \gamma_x(0) \neq 0$$

The correlation between X_t and X_{t-2} come from their correlation with X_{t-1}. If AR(1) is true, then the correlation between X_t and X_{t-2} would be zero after the effect of X_{t-1} is "removed."

To do so, we write

$$X_t = P_{X_{t-1}} X_t + P^{\perp}_{X_{t-1}} X_t \quad \text{and} \quad X_{t-2} = P_{X_{t-1}} X_{t-2} + P^{\perp}_{X_{t-1}} X_{t-2}$$

where $P_{X_{t-1}}$ is the projection onto $\text{span}\{X_{t-1}\} = \{\beta X_{t-1} : \beta \in \mathbb{R}\}$ and $P^{\perp}_{X_{t-1}} = I - P_{X_{t-1}}$ is the projection onto the space orthogonal to $\text{span}\{X_{t-1}\}$



By the definition of projection $P_{X_{t-1}} y = \hat{\beta} X_{t-1}$ such that $\hat{\beta} = \arg\min_{\beta} \|y - \beta X_{t-1}\|^2$ which is the OLS estimate

Then the correlation between $P^{\perp}_{X_{t-1}} X_t$ and $P^{\perp}_{X_{t-1}} X_{t-2}$ represent the correlation between X_t and X_{t-2} with the effect of X_{t-1} is "removed".

If $\text{Corr}(P^{\perp}_{X_{t-1}} X_t, P^{\perp}_{X_{t-1}} X_{t-2}) \simeq 0$, we choose AR(1) model. If not, we may further ask whether we should use AR(2) $X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + W_t$ or AR(3) $X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \phi_3 X_{t-3} + W_t$.

Again $X_{t-3} = \sum_{j=0}^{\infty} \psi_j W_{t-3-j}$ correlated with X_t through X_{t-1} and X_{t-2}

We then compute $P_{\{X_{t-1}, X_{t-2}\}}$, which is the projection onto $\text{span}\{X_{t-1}, X_{t-2}\} = \{\beta_1 X_{t-1} + \beta_2 X_{t-2}\}$. Let $P_{\{t-1:t-2\}} = P_{\{t-2:t-1\}} = P_{\{X_{t-1}, X_{t-2}\}}$, then

$P_{\{t-1:t-2\}} X_t = \hat{\beta}_1 X_{t-1} + \hat{\beta}_2 X_{t-2}$, the regression of X_t on $\{X_{t-1}, X_{t-2}\}$

$P_{\{t-1:t-2\}} X_{t-3} = \tilde{\beta}_1 X_{t-1} + \tilde{\beta}_2 X_{t-2}$, the regression of X_{t-3} on $\{X_{t-1}, X_{t-2}\}$

And we check if $\text{Corr}(P^{\perp}_{\{t-1:t-2\}} X_t, P^{\perp}_{\{t-1:t-2\}} X_{t-3})$

$$= \text{Corr}(X_t - P_{\{t-1:t-2\}} X_t, X_{t-3} - P_{\{t-1:t-2\}} X_{t-3})$$

is close to 0 or not.

Definition 3.9 The partial autocorrelation function (PACF)

For stationary X_t, define $\phi_{11} = \mathrm{Corr}(X_{t+1}, X_t) = \rho(1)$

$$\phi_{hh} = \mathrm{Corr}(X_{t+h} - \hat{X}_{t+h},\ X_t - \hat{X}_t),\quad h \geq 2$$

where $\hat{X}_{t+h} = P_{\{t+1:t+h-1\}} X_{t+h}$, $\quad \hat{X}_t = P_{\{t+1:t+h-1\}} X_t$

Table 3.1

	AR(p)	MA(q)	ARMA(p, q)
ACF ($\rho(h)$)	Tails off	Cuts off after lag q	Tails off
PACF (ϕ_{hh})	Cuts off after lag p	Tails off	Tails off

Example 3.15 $X_t = \phi X_{t-1} + W_t$, $\quad |\phi| < 1$

To compute ϕ_{22}, we first compute $\hat{X}_{t+2} = P_{\{t+1\}} X_{t+2} = \hat{\beta} X_{t+1}$

where
$$\begin{aligned}\hat{\beta} = \arg\min_\beta \| X_{t+2} - \beta X_{t+1} \|^2 &= \arg\min_\beta E(X_{t+2} - \beta X_{t+1})^2 \\ &= \arg\min_\beta (\gamma(0) - 2\beta\gamma(1) + \beta^2\gamma(0)) \\ &= \frac{\gamma(1)}{\gamma(0)} = \rho(1) = \phi \ (= \phi_{11} \text{ by definition})\end{aligned}$$

Similarly, $\hat{X}_t = P_{\{t+1\}} X_{t+1} = \tilde{\beta} X_{t+1}$

$$\begin{aligned}\tilde{\beta} = \arg\min_\beta E(X_t - \beta X_{t+1})^2 &= \arg\min_\beta (\gamma(0) - 2\beta\gamma(1) + \beta^2\gamma(0)) \\ &= \hat{\beta} = \phi\end{aligned}$$

$$\begin{aligned}\therefore \phi_{22} &= \mathrm{Corr}(X_{t+2} - \hat{X}_{t+2},\ X_t - \hat{X}_t) \\ &= \mathrm{Corr}(X_{t+2} - \phi X_{t+1},\ X_t - \phi X_{t+1}) = \mathrm{Corr}\left(W_{t+2}, \sum_{j=0}^\infty \psi_j W_{t-j} - \phi \sum_{j=0}^\infty \psi_j W_{t+1-j}\right) \\ &= 0\end{aligned}$$

For ARMA(p, q) $\phi(B)X_t = \theta(B)W_t$, it can be writeen as a MA(∞) $X_t = \phi^{-1}(B)\theta(B)W_t$ or a AR(∞) $\theta^{-1}(B)\phi(B)X_t = W_t$ and hence both ACF, $\rho(h)$, and PACF, ϕ_{hh}, do not cut off at finite lags.