UNIVERSITÀ DEGLI STUDI DI MILANO
Dipartimento di Economia, Management e Metodi Quantitativi

Academic Year 2019-2020

Time Series Econometics

Fabrizio Iacone

Chapter 4, ARMA Models

Topics: White noise, MA(1) model, MA(q) model, MA(∞) model, AR(1) model, AR(2) model, AR(p) model, ARMA(1,1) model, ARMA(p,q) model.
Sum of ARMA processes.

We said we are interested in the ψ_j of the representation

$$Y_t = \mu + \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$$

for the impulse response analysis and for forecasting (assuming ε_{t-j} is observable, of course).
However, in general we don't know the ψ_j, and we can't hope to estimate an infinite number of parameters, so we have to propose models that are a function of a very little number of parameters.

The first and most simple model is the

White Noise

$\{\varepsilon_t\}_{t=-\infty}^{\infty}$ is white noise if

$$E(\varepsilon_t) = 0 \ \forall t$$

$$E(\varepsilon_t^2) = \sigma^2 \ \forall t$$

$$E(\varepsilon_t \varepsilon_\tau) = 0 \ \forall t, \tau \text{ such that } \tau \neq t$$

So, if $Y_t = \varepsilon_t$,

$$\psi_j = 0 \ \forall j \neq 0$$

$$\gamma_j = 0 \ \forall j \neq 0$$

$$\alpha_j^{(j)} = 0 \ \forall j \neq 0$$

i.e. the process has no memory.

If ε_t is w.n.$(0, \sigma^2)$, and $Y_t = \varepsilon_t$,

- Y_t may be independent, but needs not be;

- Y_t may be strictly stationary, but needs not be;

- Y_t is covariance stationary;

If ε_t is w.n.$(0, \sigma^2)$, and
$Y_t = \mu + \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$,

- Y_t is stationary if $\sum_{j=0}^{\infty} \psi_j^2 < \infty$
- Y_t is stationary and ergodic (for the mean) if $\sum_{j=0}^{\infty} |\psi_j| < \infty$

MA(1)

Let ε_t w.n.$(0, \sigma^2)$, then

$$Y_t = \mu + \varepsilon_t + \theta\varepsilon_{t-1}$$

is MA(1).

We can check stationarity noticing that $\psi_0 = 1$, $\psi_1 = \theta$, so $\sum_{j=0}^{\infty} \psi_j^2 = 1 + \theta^2 < \infty$.

Otherwise, we can check that the first two moments do not depend on time.

Mean:

$$\begin{aligned} E(Y_t) &= E(\mu + \varepsilon_t + \theta\varepsilon_{t-1}) \\ &= E(\mu) + E(\varepsilon_t) + E(\theta\varepsilon_{t-1}) \\ &= \mu \quad + 0 \quad + 0 = \mu \end{aligned}$$

Autocovariances:

$$\begin{aligned}
\gamma_0 &= E\left[(Y_t - \mu)^2\right] \\
&= E\left[(\varepsilon_t + \theta\varepsilon_{t-1})^2\right] \\
&= E(\varepsilon_t^2 + \theta^2\varepsilon_{t-1}^2 + 2\theta\varepsilon_t\varepsilon_{t-1}) \\
&= E(\varepsilon_t^2) + \theta^2 E(\varepsilon_{t-1}^2) + 2\theta E(\varepsilon_t\varepsilon_{t-1}) \\
&= \sigma^2 \quad + \theta^2\sigma^2 \quad + 0 = (1+\theta^2)\sigma^2
\end{aligned}$$

$$\begin{aligned}
\gamma_1 &= E[(Y_t - \mu)(Y_{t-1} - \mu)] \\
&= E[(\varepsilon_t + \theta\varepsilon_{t-1})(\varepsilon_{t-1} + \theta\varepsilon_{t-2})] \\
&= E(\varepsilon_t\varepsilon_{t-1} + \theta\varepsilon_{t-1}^2 + \theta\varepsilon_t\varepsilon_{t-2} + \theta^2\varepsilon_{t-1}\varepsilon_{t-2}) \\
&= 0 \quad + \theta\sigma^2 \quad + 0 \quad + 0 = \theta\sigma^2
\end{aligned}$$

$$\gamma_{j\geq 2} = 0$$

★ So, if we want to check for stationarity by checking the moments, we verify

$$E(Y_t) = \mu \text{ for any } t$$

$$Cov(Y_t, Y_{t+j}) = \gamma_j \text{ for any } t$$

and, in particular,

$$\gamma_0 = (1+\theta^2)\sigma^2, \gamma_1 = \theta\sigma^2, \gamma_j = 0 \text{ for } j \geq 2.$$

Autocorrelations:

$$\rho_1 = \frac{\theta\sigma^2}{(1+\theta^2)\sigma^2} = \frac{\theta}{1+\theta^2}$$

$$\rho_{j\geq 2} = 0$$

Partial autocorrelations: using the definition it is possible to compute

$$\alpha_j^{(j)} = -\frac{(-\theta)^j}{(1+\theta^2+\ldots+\theta^{2j})} = -\frac{(-\theta)^j}{\sum_{i=0}^{j}\theta^{2i}}$$

★ Note: the same autocorrelation structure is generated by two values of θ. Consider

$$\theta = \theta_1 \text{ and } \theta = \theta_2 = 1/\theta_1:$$

$$\text{when } \theta = \theta_1,\ \rho_1|_{\theta=\theta_1} = \frac{\theta_1}{1+\theta_1^2};$$

$$\text{when } \theta = \theta_2,\ \rho_1|_{\theta=\theta_2} = \frac{\theta_2}{1+\theta_2^2}\ :$$

Replacing $\theta_2 = 1/\theta_1$ in $\rho_1|_{\theta=\theta_2}$,

$$\rho_1|_{\theta=\theta_2} = \frac{1/\theta_1}{1+1/\theta_1^2} = \frac{\theta_1^2}{\theta_1^2}\frac{1/\theta_1}{1+1/\theta_1^2}$$

$$= \frac{\theta_1}{\theta_1^2+1} = \rho_1|_{\theta=\theta_1}$$

i.e. θ_1 and θ_2 ($\theta_2 = 1/\theta_1$) generate two equally valid representations of the same process.

Invertibility

Assume $\theta = \theta_1$, $|\theta_1| < 1$, and set $\mu = 0$, then rewrite $Y_t = \varepsilon_t + \theta\varepsilon_{t-1}$ as

$$\varepsilon_t = Y_t - \theta\varepsilon_{t-1}$$

and notice that $\varepsilon_{t-1} = Y_{t-1} - \theta\varepsilon_{t-2}$ so, replacing in the formula for ε_t,

$$\begin{aligned}\varepsilon_t &= Y_t - \theta(Y_{t-1} - \theta\varepsilon_{t-2}) \\ &= Y_t - \theta Y_{t-1} + \theta^2\varepsilon_{t-2}\end{aligned}$$

In the same way, $\varepsilon_{t-2} = Y_{t-2} - \theta\varepsilon_{t-3}$ so

$$\begin{aligned}\varepsilon_t &= Y_t - \theta Y_{t-1} + \theta^2(Y_{t-2} - \theta\varepsilon_{t-3}) \\ &= Y_t - \theta Y_{t-1} + \theta^2 Y_{t-2} - \theta^3\varepsilon_{t-3}\end{aligned}$$

Iterating n times,

$$\varepsilon_t = \sum_{j=0}^{n}(-\theta)^j Y_{t-j} + (-\theta)^{n+1}\varepsilon_{t-(n+1)}$$

and, for $n \to \infty$, since $|\theta| < 1$, then $(-\theta)^{n+1} \to 0$, so

$$\varepsilon_t = \sum_{j=0}^{\infty}(-\theta)^j Y_{t-j}, \text{ i.e. } Y_t = \sum_{j=1}^{\infty} -(-\theta)^j Y_{t-j} + \varepsilon_t$$

So for an invertible MA(1) process, we can compute ε_t provided that we know $Y_t, \ldots, Y_{-\infty}$ and θ.

An alternative way to obtain this representation:

Rewrite $\varepsilon_t = Y_t - \theta\varepsilon_{t-1}$ as

$$\varepsilon_t = Y_t - \theta L \varepsilon_t$$

using the lag operator. Then,

$$\varepsilon_t + \theta L \varepsilon_t = Y_t$$

$$(1 + \theta L)\varepsilon_t = Y_t$$

so, for $|\theta| < 1$, then $\varepsilon_t = (1 + \theta L)^{-1} Y_t$, i.e.

$$\varepsilon_t = \frac{1}{(1 + \theta L)} Y_t.$$

Since

$$\frac{1}{(1 + \theta L)} = \sum_{j=0}^{\infty} (-\theta)^j L^j,$$

then

$$\varepsilon_t = \sum_{j=0}^{\infty} (-\theta)^j Y_{t-j},$$

i.e.

$$Y_t = \sum_{j=1}^{\infty} -(-\theta)^j Y_{t-j} + \varepsilon_t$$

However, if $\theta = \theta_2$, $|\theta_2| \geq 1$, then $(-\theta)^{n+1} \nrightarrow 0$ as $n \to \infty$, so the representation is not invertible if $\theta = \theta_2$.

Finally, if $\theta_1 = 1$, then $1/\theta_1 = 1$, and in both the cases $|\theta| < 1$ is not met. So, for $\theta = 1$ no invertible representation is avalailable.

MA(q)

Let ε_t w.n.$(0, \sigma^2)$, then

$$Y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q}$$

is MA(q).

Again, it is easy to verify that this MA(q) is stationary, as

$$\sum_{j=0}^{\infty} \psi_j^2 = 1 + \theta_1^2 + \ldots + \theta_q^2 < \infty.$$

Mean:

$$E(Y_t) = E(\mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q}) = \mu$$

Autocovariances:

$$\begin{aligned} \gamma_0 &= E\left[(\varepsilon_t + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q})^2 \right] \\ &= (\sigma^2 + \theta_1^2 \sigma^2 + \ldots + \theta_q^2 \sigma^2) \\ &= \left(1 + \theta_1^2 + \ldots + \theta_q^2\right) \sigma^2 \end{aligned}$$

(using $E(\varepsilon_{t-j} \varepsilon_{t-k}) = 0$ for $k \neq j$).

$$
\begin{aligned}
\gamma_{j\le q} &= E[(\varepsilon_t + \theta_1\varepsilon_{t-1} + \ldots + \theta_q\varepsilon_{t-q}) \\
&\quad \times (\varepsilon_{t-j} + \theta_1\varepsilon_{t-1-j} + \ldots + \theta_q\varepsilon_{t-q-j})]
\end{aligned}
$$

$$
= E(\theta_j\varepsilon_{t-j}^2 + \theta_{j+1}\theta_1\varepsilon_{t-j-1}^2 + \theta_{j+2}\theta_2\varepsilon_{t-j-2}^2 + \ldots + \theta_q\theta_{q-j}\varepsilon_{t-q}^2
$$

$$
= \left(\theta_j + \theta_{j+1}\theta_1 + \theta_{j+2}\theta_2 + \ldots + \theta_{q-j}\theta_q\right)\sigma^2
$$

$$
\gamma_{j>q} = 0
$$

Autocorrelations:
the autocorrelations drop to 0 after q lags

Impulse Response Function:
the impulse response are $\psi_j = \theta_j, j \le q$, and drop to 0 after q lags.

Invertibility

Set $\mu = 0$; recall

$$Y_t = (1 + \theta_1 L + \ldots + \theta_q L^q)\varepsilon_t$$

and factor

$$(1 + \theta_1 L + \ldots + \theta_q L^q)$$
$$= (1 - \lambda_1 L)(1 - \lambda_2 L)\ldots(1 - \lambda_q L)$$

in the MA(1) we asked that $|\lambda_1| < 1$: in the same way here we have to ask that $|\lambda_1| < 1$, $|\lambda_2| < 1$,...,$|\lambda_q| < 1$.

This is sometimes stated as asking that the roots of the equation in z

$$(1 + \theta_1 z + \ldots + \theta_q z^q) = 0$$

lie outside the unit circle.

If the MA(q) process is invertible, we can write

$$\varepsilon_t = (1 + \theta_1 L + \ldots + \theta_q L^q)^{-1} Y_t$$

and then derive π_0, π_1, π_2, ... such that

$$\varepsilon_t = \sum_{j=0}^{\infty} -\pi_j Y_{t-j}, \text{ i.e. } Y_t = \sum_{j=1}^{\infty} \pi_j Y_{t-j} + \varepsilon_t$$

MA(∞)

Let ε_t w.n.$(0, \sigma^2)$, then

$$Y_t = \mu + \varepsilon_t + \psi_1 \varepsilon_{t-1} + \ldots = \mu + \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$$

is MA(∞).
Under the additional assumption that

$$\sum_{j=0}^{\infty} |\psi_j| < \infty,$$

we can derive the moments replacing θ_j by ψ_j in a MA(q) and taking the limit for $q \to \infty$.

Mean:

$$E(Y_t) = E(Y_t = \mu + \varepsilon_t + \psi_1 \varepsilon_{t-1} + \ldots) = \mu$$

Autocovariances:

$$\gamma_0 = \sum_{k=0}^{\infty} \psi_k^2 \sigma^2$$

$$\gamma_j = \sum_{k=0}^{\infty} \psi_k \psi_{k+j} \sigma^2$$

AR(1)

Let ε_t w.n.$(0, \sigma^2)$, then

$$Y_t = c + \phi Y_{t-1} + \varepsilon_t$$

is AR(1).
Assume further that

$$|\phi| < 1.$$

Since $Y_{t-1} = c + \phi Y_{t-2} + \varepsilon_{t-1}$,

$$\begin{aligned} Y_t &= c + \phi(c + \phi Y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\ &= (1+\phi)c + \phi^2 Y_{t-2} + \phi\varepsilon_{t-1} + \varepsilon_t \end{aligned}$$

Next, replace $Y_{t-2} = c + \phi Y_{t-3} + \varepsilon_{t-2}$

$$\begin{aligned} Y_t &= (1+\phi)c + \phi^2(c + \phi Y_{t-3} + \varepsilon_{t-2}) \\ &\quad + \phi\varepsilon_{t-1} + \varepsilon_t \\ &= (1+\phi+\phi^2)c + \phi^3 Y_{t-3} \\ &\quad + \phi^2\varepsilon_{t-2} + \phi\varepsilon_{t-1} + \varepsilon_t \end{aligned}$$

Iterating n times,

$$Y_t = \sum_{j=0}^{n} \phi^j c + \phi^{n+1} Y_{t-n-1} + \sum_{j=0}^{n} \phi^j \varepsilon_{t-j};$$

as $n \to \infty$, since $|\phi| < 1$, then $\phi^{n+1} \to 0$ and $\sum_{j=0}^{n} \phi^j \to \frac{1}{1-\phi}$, so

$$Y_t = \frac{1}{1-\phi} c + \sum_{j=0}^{\infty} \phi^j \varepsilon_{t-j}$$

So an AR(1) with $|\phi| < 1$ may be written as a MA(∞): notice that the condition $\sum_{j=0}^{\infty} |\psi_j| < \infty$ is met, because $\psi_j = \phi^j$, so $\sum_{j=0}^{\infty} |\psi_j| = \sum_{j=0}^{\infty} |\phi|^j = \frac{1}{1-|\phi|}$ (then it also follows that the process is stationary and ergodic for the mean).

This can also be obtained rewriting Y_t as

$$Y_t = c + \phi L Y_t + \varepsilon_t$$

using the lag operator, and then

$$(1 - \phi L)Y_t = c + \varepsilon_t$$

Since $|\phi| < 1$,

$$Y_t = (1 - \phi L)^{-1} c + (1 - \phi L)^{-1} \varepsilon_t$$

and since

$$\frac{1}{(1 - \phi L)} = \sum_{j=0}^{\infty} \phi^j L^j,$$

so

$$Y_t = \sum_{j=0}^{\infty} \phi^j c + \sum_{j=0}^{\infty} \phi^j \varepsilon_{t-j}$$

The representation

$$Y_t = \frac{1}{1 - \phi} c + \sum_{j=0}^{\infty} \phi^j \varepsilon_{t-j}$$

then follows.

Mean:

$$E(Y_t) = \frac{1}{1-\phi}c + 0 + 0 + 0 + \ldots = \frac{1}{1-\phi}c$$

(so set $\mu = \frac{1}{1-\phi}c$);

Autocovariances:

using the formula for the MA(∞) process,

$$\gamma_0 = \sum_{k=0}^{\infty} \psi_k^2 \sigma^2 = \sum_{k=0}^{\infty} \phi^{2k} \sigma^2 = \frac{1}{1-\phi^2}\sigma^2$$

$$\gamma_j = \sum_{k=0}^{\infty} \psi_k \psi_{k+j} \sigma^2 = \sum_{k=0}^{\infty} \phi^k \phi^{k+j} \sigma^2$$

$$= \sum_{k=0}^{\infty} \phi^{2k} \phi^j \sigma^2 = \frac{\phi^j}{1-\phi^2}\sigma^2$$

Autocorrelations

$$\rho_j = \frac{\gamma_j}{\gamma_0} = \phi^j$$

Partial autocorrelations

$$\alpha_1^{(1)} = \phi$$

$$\alpha_{j\geq 2}^{(j)} = 0$$

Impulse Response Function

$$\psi_j = \phi^j$$

Upon knowing that the process is stationary, we could derive the mean and autocovariances using that property:
Mean:

$$E(Y_t) = E(c + \phi Y_{t-1} + \varepsilon_t)$$
$$= c + \phi E(Y_{t-1}) + E(\varepsilon_t)$$

using stationarity, $E(Y_t) = \mu$, $E(Y_{t-1}) = \mu$, so

$$\mu = c + \phi\mu$$

and then

$$\mu = \frac{c}{1-\phi}$$

Autocovariances:
Replacing $c = (1-\phi)\mu$, rewrite Y_t as

$$Y_t = \mu - \phi\mu + \phi Y_{t-1} + \varepsilon_t$$

$$Y_t - \mu = \phi(Y_{t-1} - \mu) + \varepsilon_t$$

$$\gamma_0 = E(Y_t - \mu)^2 = E(\phi(Y_{t-1} - \mu) + \varepsilon_t)^2$$
$$= \phi^2 E(Y_{t-1} - \mu)^2 + E(\varepsilon_t^2)$$
$$+ 2\phi E((Y_{t-1} - \mu)\varepsilon_t)$$
$$= \phi^2\gamma_0 + \sigma^2$$

solving for γ_0,

$$\gamma_0 = \frac{\sigma^2}{1-\phi^2}.$$

$$\begin{aligned}
\gamma_{j\geq 1} &= E[(Y_t-\mu)(Y_{t-j}-\mu)] \\
&= E[(\phi(Y_{t-1}-\mu)+\varepsilon_t)(Y_{t-j}-\mu)] \\
&= \phi E[(Y_{t-1}-\mu)(Y_{t-j}-\mu)] \\
&\quad + E(\varepsilon_t(Y_{t-j}-\mu)) \\
&= \phi\gamma_{j-1}
\end{aligned}$$

so

$$\gamma_{j\geq 1} = \frac{\phi^j}{1-\phi^2}\sigma^2.$$

AR(p)

Let ε_t w.n.$(0, \sigma^2)$, then

$$Y_t = c + \phi_1 Y_{t-1} + \ldots + \phi_p Y_{t-p} + \varepsilon_t$$

is AR(p).

How can we check for stationarity? Factoring

$$(1 - \phi_1 L - \ldots - \phi_p L^p) = (1 - \lambda_1 L)\ldots(1 - \lambda_p L)$$

stationarity follows if $|\lambda_j| < 1$ for all j.

Another way to state this condition is check that the solutions of the equation in z

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

are all outside the unit circle.

Given stationarity,

Mean:

$$\begin{aligned} E(Y_t) &= E(c + \phi_1 Y_{t-1} + \ldots + \phi_p Y_{t-p} + \varepsilon_t) \\ &= c + \phi_1 E(Y_{t-1}) + \ldots + \phi_p E(Y_{t-p}) + E(\varepsilon_t) \end{aligned}$$

$$\mu = c + \phi_1 \mu + \ldots + \phi_p \mu$$

$$\mu = \frac{c}{1 - \phi_1 - \ldots - \phi_p}$$

Autocovariances

$$\begin{aligned}
\gamma_0 &= E(Y_t - \mu)^2 \\
&= E[(\phi_1(Y_{t-1} - \mu) + \ldots + \phi_p(Y_{t-p} - \mu) + \varepsilon_t) \\
&\quad \times (Y_t - \mu)] \\
&= E[\phi_1(Y_{t-1} - \mu)(Y_t - \mu) + \ldots \\
&\quad + \phi_p(Y_{t-p} - \mu)(Y_t - \mu) + \varepsilon_t(Y_t - \mu)] \\
&= \phi_1\gamma_1 + \ldots. + \phi_p\gamma_p + \sigma^2
\end{aligned}$$

$$\begin{aligned}
\gamma_{j\geq 1} &= E[(Y_t - \mu)(Y_{t-j} - \mu)] \\
&= E[(\phi_1(Y_{t-1} - \mu) + \ldots + \phi_p(Y_{t-p} - \mu) + \varepsilon_t) \\
&\quad \times (Y_{t-j} - \mu)] \\
&= E[\phi_1(Y_{t-1} - \mu)(Y_{t-j} - \mu) + \ldots \\
&\quad + \phi_p(Y_{t-p} - \mu)(Y_{t-j} - \mu) + \varepsilon_t(Y_{t-j} - \mu)] \\
&= \phi_1\gamma_{j-1} + \ldots. + \phi_p\gamma_{j-p}
\end{aligned}$$

This is a linear system in $\gamma_j, j = 0, \ldots, p$.

Autocorrelations
(Yule Walker equations)

$$\rho_{j\geq 1} = \phi_1 \rho_{j-1} + \ldots + \phi_p \rho_{j-p}$$

Partial autocorrelations AR(p)

$$\alpha_1^{(1)} = \rho_1$$

$$\alpha_j^{(j)} \neq 0 \ (\text{for } 1 < j < p)$$

$$\alpha_p^{(p)} = \phi_p$$

$$\alpha_{j>p}^{(j)} = 0$$

For example, AR(2),

$$\gamma_0 = \phi_1\gamma_1 + \phi_2\gamma_2 + \sigma^2$$
$$\gamma_1 = \phi_1\gamma_0 + \phi_2\gamma_{-1}$$
$$\gamma_2 = \phi_1\gamma_1 + \phi_2\gamma_0$$

and notice that $\gamma_1 = \gamma_{-1}$, so replacing γ_1 and γ_2,

$$\gamma_1 = \frac{\phi_1}{1-\phi_2}\gamma_0, \gamma_2 = \left(\frac{\phi_1^2}{1-\phi_2} + \phi_2\right)\gamma_0$$

$$\gamma_0 = \frac{(1-\phi_2)}{(1+\phi_2)\left[(1-\phi_2)^2 - \phi_1^2\right]}\sigma^2$$

and

$$\rho_1 = \phi_1 + \phi_2\rho_1$$
$$\rho_2 = \phi_1\rho_1 + \phi_2$$

so

$$\rho_1 = \frac{\phi_1}{1-\phi_2}$$
$$\rho_2 = \frac{\phi_1^2 + \phi_2 - \phi_2^2}{1-\phi_2}.$$

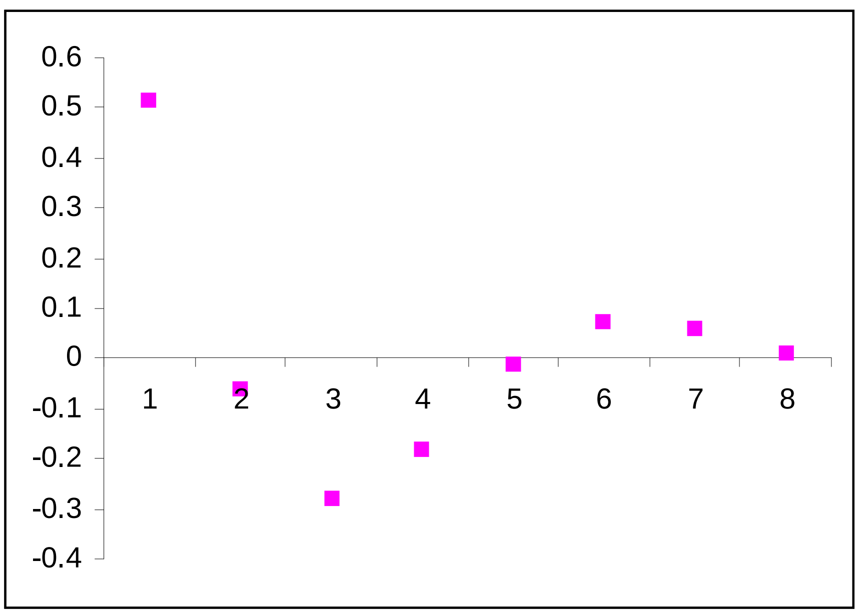
★ AR(2) If the roots of $1 - \phi_1 z - \phi_2 z^2 = 0$ are complex, then the autocorrelations show a cyclical dynamics. This is very important because both economics and natural phenomena often display cyclical dynamics.

Example: $\phi_1 = 0.75$, $\phi_2 = -0.45$.

Solutions of $1 - 0.75z + 0.45z^2 = 0$ are $z_{1,2} = 0.83333 \pm 1.236i$.

Note that $|z_i| = \sqrt{0.83333^2 + 1.236^2} = 1.4907$ so this is process is stationary.

Autocorrelation function:



$$\phi_1 = 0.75, \ \phi_2 = -0.45$$

★ AR(p). If some roots of $1 - \phi_1 z - \ldots - \phi_p z^p = 0$ are complex, then the autocorrelations show cyclical dynamics.

Impulse Response Function:
in general we can compute the IRF inverting $\phi(L)Y_t = \varepsilon_t$

$$Y_t = \phi(L)^{-1}\varepsilon_t$$

(here we used stationarity) so

$$\phi(L)^{-1} = \psi(L)$$

i.e.

$$\begin{aligned}
1 &= \phi(L)\psi(L) \\
1 &= (1 - \phi_1 L - \ldots - \phi_p L^p) \\
&\quad \times (1 + \psi_1 L + \psi_2 L^2 + \psi_3 L^3 \ldots) \\
1 &= 1 - \phi_1 L + \psi_1 L - \phi_2 L^2 - \phi_1 \psi_1 L^2 + \psi_2 L^2 \\
&\quad - \phi_3 L^3 - \phi_2 \psi_1 L^3 - \phi_1 \psi_2 L^3 + \psi_3 L^3 + \ldots.
\end{aligned}$$

solve this for the various powers of L :

$$L^0 : \quad 1 = 1$$

$$L : \quad -\phi_1 + \psi_1 = 0, \text{ so } \psi_1 = \phi_1$$

$$L^2 : \quad -\phi_2 - \phi_1\psi_1 + \psi_2 = 0,$$

$$\text{so } \psi_2 = \phi_1\psi_1 + \phi_2$$

$$L^3 : \quad -\phi_3 - \phi_2\psi_1 - \phi_1\psi_2 + \psi_3 = 0,$$

$$\text{so } \psi_3 = \phi_3 + \phi_2\psi_1 + \phi_1\psi_2$$

In the AR(2) case, then,

$$\psi_1 = \phi_1,$$

$$\psi_{j\geq 2} = \psi_{j-1}\phi_1 + \phi_2\psi_{j-2}$$

so if for example $\phi_1 = 0.75$, $\phi_2 = -0.45$,

$$\psi_1 = 0.75$$

$$\psi_2 = 0.75 \times 0.75 + (-0.45) \times 1 = 0.1125$$

$$\psi_3 = 0.1125 \times 0.75 + (-0.45) \times 0.75 = -0.25313$$

$$\psi_4 = -0.25313 \times 0.75 + (-0.45) \times 0.1125 = -0.24047$$

ARMA(p, q)

Let ε_t w.n.$(0, \sigma^2)$, then

$$\begin{aligned} Y_t &= c + \phi_1 Y_{t-1} + \ldots + \phi_p Y_{t-p} \\ &\quad + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q} \end{aligned}$$

is ARMA(p, q).

Stationarity of the whole ARMA(p, q) depends on the autoregressive part only: we have to check if the roots of

$$1 - \phi_1 z - \ldots - \phi_p z^p = 0$$

are all outside the unit circle.

For invertibility, we require that the roots of

$$1 + \theta_1 z + \ldots + \theta_q z^q = 0$$

are outside the unit circle.

Using the lag operator, the ARMA(p, q) is

$$(1 - \phi_1 L - \ldots - \phi_p L^p) Y_t = (1 + \theta_1 L + \ldots + \theta_q L^q) \varepsilon_t$$

★ Using stationarity, we can rewrite the model as a MA(∞):

$$Y_t = (1 - \phi_1 L - \ldots - \phi_p L^p)^{-1} (1 + \theta_1 L + \ldots + \theta_q L^q) \varepsilon_t$$

$$= \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$$

★ Using invertibility, we can rewrite the model as a AR(∞)

$$\varepsilon_t = (1 + \theta_1 L + \ldots + \theta_q L^q)^{-1} (1 - \phi_1 L - \ldots - \phi_p L^p) Y_t$$

$$\varepsilon_t = \sum_{j=0}^{\infty} -\pi_j Y_{t-j}, \text{ i.e. } Y_t = \sum_{j=1}^{\infty} \pi_j Y_{t-j} + \varepsilon_t$$

Given stationarity,
Mean:

$$\begin{aligned} E(Y_t) &= E(c + \phi_1 Y_{t-1} + \ldots + \phi_p Y_{t-p} \\ &\quad + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q}) \\ &= c + \phi_1 \mu + \ldots + \phi_p \mu + 0 + \ldots + 0 \end{aligned}$$

$$\mu = c + \phi_1 \mu + \ldots + \phi_p \mu$$

$$\mu = \frac{c}{1 - \phi_1 - \ldots - \phi_p}$$

Autocovariances
The autocovariances are a combination between those of an AR(p) and a MA(q), so for $j > q$,

$$\gamma_j = \phi_1 \gamma_{j-1} + \ldots + \phi_p \gamma_{j-p}$$

For example, ARMA(1, 1),
$Y_t = c + \phi Y_{t-1} + \varepsilon_t + \theta\varepsilon_{t-1}$ ($|\phi| < 1$):
first notice that

$$\begin{aligned} & E[(Y_t - \mu)\varepsilon_t] \\ = & E[(\phi(Y_{t-1} - \mu) + \varepsilon_t + \theta\varepsilon_{t-1})\varepsilon_t] \\ = & \quad 0 \quad + \sigma^2 + 0 = \sigma^2 \end{aligned}$$

$$\begin{aligned} & E[(Y_t - \mu)\varepsilon_{t-1}] \\ = & E[(\phi(Y_{t-1} - \mu) + \varepsilon_t + \theta\varepsilon_{t-1})\varepsilon_{t-1}] \\ = & \quad \phi\sigma^2 \quad + 0 \quad + \theta\sigma^2 = (\phi + \theta)\sigma^2 \end{aligned}$$

so

$$\begin{aligned} \gamma_0 &= E[(\phi(Y_{t-1} - \mu) + \varepsilon_t + \theta\varepsilon_{t-1})(Y_t - \mu)] \\ &= \phi E[(Y_{t-1} - \mu)(Y_t - \mu)] \\ &\quad + E[\varepsilon_t(Y_t - \mu)] + \theta E[\varepsilon_{t-1}(Y_t - \mu)] \\ &= \phi\gamma_1 + \sigma^2 + \theta(\phi + \theta)\sigma^2 \\ &= \phi\gamma_1 + \sigma^2(1 + \phi\theta + \theta^2) \end{aligned}$$

$$\begin{aligned} \gamma_1 &= E[(Y_t - \mu)(Y_{t-1} - \mu)] \\ &= E[\phi(Y_{t-1} - \mu)(Y_{t-1} - \mu)] \\ &\quad + E[\varepsilon_t(Y_{t-1} - \mu)] + E[\theta\varepsilon_{t-1}(Y_{t-1} - \mu)] \\ &= \phi\gamma_0 + 0 + \theta\sigma^2 = \phi\gamma_0 + \theta\sigma^2 \end{aligned}$$

so

$$
\begin{aligned}
\gamma_0 &= \phi(\phi\gamma_0 + \theta\sigma^2) + \sigma^2(1 + \phi\theta + \theta^2) \\
&= \phi^2\gamma_0 + \sigma^2(\phi\theta + 1 + \phi\theta + \theta^2) \\
&= \frac{\sigma^2(1 + 2\phi\theta + \theta^2)}{1 - \phi^2} \\
&= \frac{\sigma^2(1 - \phi^2 + \phi^2 + 2\phi\theta + \theta^2)}{1 - \phi^2} \\
&= \sigma^2\left(1 + \frac{(\phi + \theta)^2}{1 - \phi^2}\right)
\end{aligned}
$$

$$
\gamma_1 = \sigma^2\left(\theta + \phi + \frac{(\theta + \phi)^2\phi}{1 - \phi^2}\right)
$$

and

$$
\gamma_{j\geq 2} = \phi\gamma_{j-1}
$$

The autocorrelations can be derived in the same way: for the generic ARMA(p, q), for $j > q$,

$$\rho_j = \phi_1 \rho_{j-1} + \ldots + \phi_p \rho_{j-p}$$

In particular, for the ARMA(1, 1),

$$\rho_1 = \frac{(\theta + \phi)(1 + \theta\phi)}{1 + \theta^2 + 2\phi\theta}$$

$$\rho_{j\geq 2} = \phi \rho_{j-1}$$

Impulse response function

Given stationarity, inverting $\phi(L)Y_t = \theta(L)\varepsilon_t$

$$Y_t = \phi(L)^{-1}\theta(L)\varepsilon_t$$

$$\phi(L)^{-1}\theta(L) = \psi(L)$$

$$\theta(L) = \phi(L)\psi(L)$$

$$\begin{aligned} &(1 + \theta_1 L + \ldots + \theta_q L^q) \\ &= (1 - \phi_1 L - \ldots\ldots - \phi_p L^p) \\ &\quad \times (1 + \psi_1 L + \psi_2 L^2 + \psi_3 L^3 + \ldots) \end{aligned}$$

$$1 + \theta_1 L + \theta_2 L^2 + \theta_3 L^3 + \ldots + \theta_q L^q$$

$$= 1 - \phi_1 L + \psi_1 L - \phi_2 L^2 - \phi_1 \psi_1 L^2 + \psi_2 L^2$$

$$- \phi_3 L^3 - \phi_2 \psi_1 L^3 - \phi_1 \psi_2 L^3 + \psi_3 L^3 \ldots$$

solve this for the various powers of L :

$$L^0 : \quad 1 = 1$$

$$L : \quad -\phi_1 + \psi_1 = \theta_1, \text{ so } \psi_1 = \theta_1 + \phi_1$$

$$L^2 : \quad -\phi_2 - \phi_1 \psi_1 + \psi_2 = \theta_2,$$

$$\text{so } \psi_2 = \theta_2 + \phi_1 \psi_1 + \phi_2$$

$$L^3 : \quad -\phi_3 - \phi_2 \psi_1 - \phi_1 \psi_2 + \psi_3 = \theta_3,$$

$$\text{so } \psi_3 = \theta_3 + \phi_3 + \phi_2 \psi_1 + \phi_1 \psi_2$$

In the ARMA$(1,1)$ case, then,

$$\psi_1 = \theta + \phi,$$

$$\psi_{j \geq 2} = \psi_{j-1} \phi \text{ i.e. } \psi_{j \geq 2} = (\theta + \phi) \phi^{j-1}$$

The ARMA(1,1) could also be decomposed in impulse responses by looking at

$$Y_t = \phi Y_{t-1} + \xi_t \text{ where } \xi_t = \varepsilon_t + \theta \varepsilon_{t-1}$$

(and $\mu = 0$ to keep notation short). Then,

$$\begin{aligned}
Y_t &= \sum_{j=0}^{\infty} \phi^j \xi_{t-j} = \sum_{j=0}^{\infty} \phi^j (\varepsilon_{t-j} + \theta \varepsilon_{t-j-1}) \\
&= \sum_{j=0}^{\infty} \phi^j \varepsilon_{t-j} + \theta \sum_{j=0}^{\infty} \phi^j \varepsilon_{t-j-1} \\
&= \sum_{j=0}^{\infty} \phi^j \varepsilon_{t-j} + \theta \sum_{l=1}^{\infty} \phi^{l-1} \varepsilon_{t-l} \\
&= \varepsilon_t + \sum_{j=1}^{\infty} \phi \phi^{j-1} \varepsilon_{t-j} + \theta \sum_{j=1}^{\infty} \phi^{j-1} \varepsilon_{t-j} \\
&= \varepsilon_t + (\theta + \phi) \sum_{j=1}^{\infty} \phi^{j-1} \varepsilon_{t-j}
\end{aligned}$$

Common Factors

in ARMA modelling, it may be that the same factor appears both in $\phi(L)$ and of $\theta(L)$: in this case, the ARMA(p, q) process cannot be distinguished, on the basis of the autocorrelation structure (or from the weights in the MA(∞) representation), from an ARMA$(p-1, q-1)$ process.

In this case, it is sometimes also said that the model ARMA(p, q) is overparametrised.

The ARMA(p, q) model may be simplified (and indeed it may desirable to do so, especially if the parameters $\phi_1, \ldots, \phi_p$ and $\theta_1, \ldots, \theta_q$ have to be estimated).

Example:

$$Y_t = 1.2Y_{t-1} - 0.35Y_{t-2} + \varepsilon_t - 0.7\varepsilon_{t-1}$$

is

$$\begin{aligned} Y_t - 1.2Y_{t-1} + 0.35Y_{t-2} &= \varepsilon_t - 0.7\varepsilon_{t-1} \\ (1 - 1.2L + 0.35L^2)Y_t &= (1 - 0.7L)\varepsilon_t \\ (1 - 0.7L)(1 - 0.5L)Y_t &= (1 - 0.7L)\varepsilon_t \end{aligned}$$

so, simplifying $(1 - 0.7L)$, the process has the same autocorrelation structure (and the same weights in the MA(∞) representation) of

$$(1 - 0.5L)Y_t = \varepsilon_t$$

i.e.

$$Y_t = 0.5Y_{t-1} + \varepsilon_t$$

A final comment on stationary and invertible ARMA.

We already saw that for a stationary ARMA(p, q), it is also possible to give a MA(∞) representation; in the same way, it is also possible to give an AR(∞) representation (indeed, this is a proper definition of "invertibility"). All these models have the same autocovariances / autocorrelation structures, and are therefore indistinguishable.

We can choose the representation that is more convenient for our purpose: for example, we may like the MA(∞) if we are interested in the impulse rensponse function, the AR(∞) if we want to compute ε_t given observations on $\{Y_t\}_{-\infty}^{\infty}$ (and assuming we know the parameters), or we may prefer the ARMA(p, q) if we are interested in estimating the parameters.

Stationary and ergodic ARMA.

Let

$$Y_t = \phi_1 Y_{t-1} + \ldots + \phi_p Y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q}$$

for any t and assume also that $\{Y_t\}$ is stationary and $\{\varepsilon_t\}$ is independently and identically distributed with $E(\varepsilon_t) = 0$ and $E(\varepsilon_t^2) = \sigma^2$ (i.i.d.$(0, \sigma^2)$).

Then Y_t stationary and ergodic.

Sum of ARMA processes

Sometimes processes are obtained as sums of other processes, for example we may be looking at the dynamics of an aggregate process composed of individual processes.

Example:

$$Y_t = X_t + v_t$$

where

$$X_t = u_t + \delta u_{t-1}$$

and u_t is w.n.$(0, \sigma_u^2)$, v_t is w.n.$(0, \sigma_v^2)$, $E(u_t v_\tau) = 0$ for all t, τ.

What are the properties of Y_t?

$$E(Y_t) = 0 \text{ for all } t$$

$$\gamma_0 = E(X_t + v_t)^2 = E(X_t^2) + E(v_t^2) + 2E(X_t v_t)$$
$$= (1 + \delta^2)\sigma_u^2 + \sigma_v^2$$

$$\gamma_1 = E[(X_t + v_t)(X_{t-1} + v_{t-1})]$$
$$= E(X_t X_{t-1}) + E(v_t X_{t-1})$$
$$+ E(X_t v_{t-1}) + E(v_t v_{t-1})$$
$$= \delta\sigma_u^2$$

$$\gamma_{j\geq 2} = 0$$

So Y_t is MA(1), i.e., we can represent it as

$$Y_t = \varepsilon_t + \theta\varepsilon_{t-1}$$

where ε_t is wn$(0, \sigma^2)$.

Check:

Given δ, σ_u^2 and σ_v^2, we want to characterise θ and σ^2. From γ_1 and γ_2 compute

$$\rho_1 = \frac{\delta\sigma_u^2}{(1+\delta^2)\sigma_u^2 + \sigma_v^2}.$$

Since in a MA(1)

$$\frac{\theta}{1+\theta^2} = \rho_1$$

we can derive θ solving

$$\frac{\theta}{1+\theta^2} = \frac{\delta\sigma_u^2}{(1+\delta^2)\sigma_u^2 + \sigma_v^2}$$

and then we can derive σ^2, for example from $\gamma_1 = \theta\sigma^2$, so

$$\sigma^2 = \frac{\delta}{\theta}\sigma_u^2.$$

Notice that ε_t is not $u_t + v_t$.

In general, consider

$$Y_t = X_t + W_t$$

where X_t and W_t are (zero mean) stationary processes such that X_t and W_τ are not correlated at any t, τ, then

$$E(Y_t Y_{t-j}) = E(X_t X_{t-j}) + E(W_t W_{t-j})$$

$$\text{i.e. } \gamma_j^Y = \gamma_j^X + \gamma_j^W$$

Sum of two MA processes

If X_t is MA(q_1) and W_t is MA(q_2), then Y_t is MA(max[q_1, q_2])

Sum of two AR(1) processes

$$Y_t = X_t + W_t \text{ where}$$

$$(1 - \pi L)X_t = u_t, \, (1 - \rho L)W_t = v_t \quad (\pi \neq \rho)$$

then

$$(1 - \rho L)(1 - \pi L)X_t = (1 - \rho L)u_t$$

$$(1 - \rho L)(1 - \pi L)W_t = (1 - \pi L)v_t$$

$$(1 - \rho L)(1 - \pi L)(X_t + W_t)$$

$$= (1 - \rho L)u_t + (1 - \pi L)v_t$$

so Y_t is ARMA$(2, 1)$.

Note: (If $\rho = \pi$, Y_t is AR(1))

Check:

$$(1-\rho L)u_t + (1-\pi L)v_t$$

is the sum of two MA(1), so this is also an MA(1)

$$\varepsilon_t + \theta\varepsilon_{t-1}$$

with

$$\frac{\theta}{1+\theta^2} = \frac{\rho\sigma_u^2 + \pi\sigma_v^2}{(1+\rho^2)\sigma_u^2 + (1+\pi^2)\sigma_v^2}$$

and

$$\sigma^2 = \frac{\rho\sigma_u^2 + \pi\sigma_v^2}{\theta}.$$

Finally, recalling $Y_t = (X_t + W_t)$,

$$\begin{aligned}
&(1-(\rho+\pi)L + \rho\pi L^2)(X_t + W_t) \\
&= (1-(\rho+\pi)L + \rho\pi L^2)Y_t \\
&= (1-\phi_1 L - \phi_2 L^2)Y_t
\end{aligned}$$

simply setting

$$\phi_1 = (\rho+\pi), \phi_2 = -\rho\pi$$

Sum of two ARMA processes

If X_t is ARMA(p_1, q_1),
W_t is ARMA(p_2, q_2),
then Y_t is ARMA(p, q) with

$$p \leq p_1 + p_2$$

and

$$q \leq \max(p_1 + q_2, p_2 + q_1)$$

Signal extraction.

Sometimes the process that we observe is a sum because it is the sum of the process that we are interested in, and of a disturbance. Suppose we are interested in X_t, but we can only observe Y_t

$$Y_t = X_t + v_t$$

where v_t is a disturbance.

For example, X_t may be the "core inflation" and "v_t" is a disturbance.

Does averaging reveal a signal?

One common practise is to measure the inflation taking the average over some months: for example, with monthly data, this is done taking the inflation rate over the last year.

Suppose v_t is wn$(0, \sigma^2)$, and consider

$$\frac{1}{k}\sum_{j=0}^{k-1} v_{t-j}$$

(as in average of quarterly or monthly data to a yearly basis): this is now a MA(k).

Therefore, averaging induced dependendence where there was none.

In the same way, $\frac{1}{k}\sum_{j=0}^{k-1} X_{t-j}$ also increases the dependence of X_t.

So, the "new" process will in general have more dependence (when we look at monthly or quarterly series), but this has been introduced artificially, and it may have nothing to do with the core inflation.

Consider again the example of a MA(1)+wn process. As we observe Y_t, we can estimate θ and σ^2. However (without an identification assumption), we cannot estimate δ, σ_u^2 and σ_v^2 . In other words, Y_t contains less information than X_t and v_t.

Forecasting with ARMA models

Recall that the best linear forecast of Y_{t+1} using Y_t $, \dots, Y_{t-m+1}$

$$\widehat{Y}_{t+1|t,\dots,t-m+1} = \alpha_1^{(m)} Y_t + \alpha_2^{(m)} Y_{t-1} + \dots + \alpha_m^{(m)} Y_{t-m+1}$$

is obtained setting

$$\alpha = \begin{pmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{m-2} & \gamma_{m-1} \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{m-3} & \gamma_{m-2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \gamma_{m-2} & \gamma_{m-3} & \cdots & \gamma_0 & \gamma_1 \\ \gamma_{m-1} & \gamma_{m-2} & \cdots & \gamma_1 & \gamma_0 \end{pmatrix}^{-1} \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \cdots \\ \gamma_{m-1} \\ \gamma_m \end{pmatrix}$$

This may be heavy to compute, as, when m is large, it requires the inversion of an $m \times m$ matrix.

If we know that $\{Y_t\}_{t=-\infty}^{\infty}$ is a stationary invertible ARMA process, we can use this information to simplify this forecast.

★ EXAMPLE: AR(p).

If $\{Y_t\}_{t=-\infty}^{\infty}$ is AR(p) and $m \geq p$,

$$\widehat{Y}_{t+1|t,\ldots,t-m+1} = \phi_1 Y_t + \phi_2 Y_{t-1} + \ldots + \phi_p Y_{t-p+1}$$

ie, using the parameters of the AR(p), which means that we do not need to invert an $m \times m$ matrix.

★ EXAMPLE: MA(1), approximation to the optimal forecast.

Let $\widehat{Y}_{t+1|\varepsilon_t}$ be the forecast of Y_{t+1} if ε_t was observable. Then

$$\widehat{Y}_{t+1|\varepsilon_t} = \theta \varepsilon_t.$$

If the process is invertible,

$$\varepsilon_t = \frac{1}{1+\theta L} Y_t = \sum_{j=0}^{\infty} (-\theta)^j Y_{t-j}$$

so, if we had the infinite past of Y_t, then ε_t would be observable, i.e.,

$$\widehat{Y}_{t+1|t,t-1,\ldots,-\infty} = \theta \sum_{j=0}^{\infty} (-\theta)^j Y_{t-j}$$

Of course, we do not have an infinite number of past values for Y_t. However, if we assume that

$$\varepsilon_0 = 0$$

then we can compute

$$\hat{\varepsilon}_1 = Y_1, \hat{\varepsilon}_2 = Y_2 - \theta Y_1, \ldots$$

(the notation $\hat{\varepsilon}_t$ means that this was computed using the assumption $\varepsilon_0 = 0$). Iterating, we can compute $\hat{\varepsilon}_t$ using Y_1, .., Y_t, so

$$\hat{Y}_{t+1|t,\ldots,1,\varepsilon_0=0} = \hat{Y}_{t+1|\hat{\varepsilon}_t} = \theta\hat{\varepsilon}_t$$

This is an approximation to the optimal forecast (because it depends on $\varepsilon_0 = 0$, which is not usually true), but one that is worth considering, because it means that we do not need to invert a $t \times t$ matrix.

What is the error that we make if we assume $\varepsilon_0 = 0$ when it is not? Setting $\varepsilon_0 = 0$ is equivalent to setting $Y_{-1} = Y_{-2} = \ldots = 0$. As these have weights $(-\theta)^{t+1}$, $(-\theta)^{t+2}$ ... in $\varepsilon_t = \sum_{j=0}^{\infty}(-\theta)^j Y_{t-j}$, then $(-\theta)^{t+1} \to 0$ very fast as t is large and $|\theta| < 1$, and the approximation error is therefore little.

★ EXAMPLE: ARMA(p, q), approximation to the optimal forecast.

The approximation to the optimal forecast for an ARMA(p, q) may be computed in the same way.

★ In fact, in practice the parameters ϕ_1, ..., ϕ_p, θ_1, ..., θ_q, are not known, and must be estimated. Many estimation algorithms also generate the series $\widehat{\varepsilon}_1$, ..., $\widehat{\varepsilon}_T$ as part of the computation.

Forecasting with ARMA models

Example 1. AR(p)

Compute the best linear forecast, $\widehat{Y}_{t+1|t,\ldots}$, assuming

$$Y_t = 2 + 0.2Y_{t-1} - 0.4Y_{t-2} + \varepsilon_t \text{ where } \varepsilon_t \text{ is white noise}$$

$$Y_t = 5,\ Y_{t-1} = -2,\ Y_{t-2} = -1,\ Y_{t-3} = 1$$

$$\widehat{Y}_{t+1|t,t-1,t-2,t-3} = 2 + 0.2 \times 5 - 0.4 \times (-2) = 3.8$$

Notice that only up to $p = 2$ observations are used for the forecast, $\widehat{Y}_{t+1|t,t-1,t-2,t-3,\ldots} = \widehat{Y}_{t+1|t,t-1}$.

Example 2. MA(q) part 1

Compute an approximation to the best linear forecast of Y_{t+1}, assuming

$$Y_t = 1 + \varepsilon_t + 0.5\varepsilon_{t-1} \text{ where } \varepsilon_t \text{ is white noise}$$

$$\widehat{\varepsilon}_t = 0.77930.$$

Then

$$\widehat{Y}_{t+1|\widehat{\varepsilon}_t} = \mu + \theta\widehat{\varepsilon}_t = 1 + 0.5 \times 0.779\,30 = 1.389\,65$$

Example 3. MA(q) part 2

Compute an approximation to the best linear forecast of Y_{t+1}, assuming

$$Y_t = 1 + \varepsilon_t + 0.5\varepsilon_{t-1} \text{ where } \varepsilon_t \text{ is white noise}$$

and

t	10	9	8	7	6	5	4	3	2	1
y_t	2.0	1.5	0.5	0.0	1.5	1.0	0.5	0.0	1.0	2.0

Setting $\hat{\varepsilon}_0 = 0$, and focussing on $t = 10$, then

$$\hat{\varepsilon}_1 = (Y_1 - \mu) = 1$$

$$\hat{\varepsilon}_2 = (Y_2 - \mu) - \theta(Y_1 - \mu) = 0 - 0.5 \times 1 = -0.5, \ldots$$

$$\hat{\varepsilon}_t = \sum_{j=0}^{t} (-\theta)^j (Y_{t-j} - \mu) = 0.77930,$$

$$\hat{Y}_{t+1|t,t-1,\ldots,\hat{\varepsilon}_0=0} = \mu + \theta\hat{\varepsilon}_t = 1 + 0.5 \times 0.77930 = 1.38965$$

Note that the "exact", best linear forecast, $\hat{Y}_{t+1|t,t-1,\ldots,1}$, is

$$\hat{Y}_{t+1|t,t-1,\ldots,1}$$

$$= \mu + \alpha_1^{(t)}(Y_t - \mu) + \alpha_2^{(t)}(Y_{t-1} - \mu) + \ldots + \alpha_t^{(t)}(Y_1 - \mu)$$

$$= 1 + 0.5 \times 1 - 0.25 \times 0.5 \ldots \ldots - 0.00074 \times 1$$

$$= 1.38984$$

Three ways to check stationarity, a summary

- Check moments

$$E(Y_t) = \mu \text{ (i.e., constant) for any } t$$

$$E[(Y_t - \mu)(Y_{t+j} - \mu)] = \gamma_j \text{ (i.e., constant) for any } t$$

(it may change with j)

- Check MA representation

A sufficient condition is checking if we can write

$$Y_t = \mu + \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}, \text{ where } \sum_{j=0}^{\infty} \psi_j^2 < \infty$$

and ε_t is white noise

- Check roots of AR polynomial

for any ARMA

$$Y_t = c + \phi_1 Y_{t-1} + \ldots + \phi_p Y_{t-p} + \varepsilon_t + \ldots + \theta_q \varepsilon_{t-q}$$

where ε_t is white noise, sufficient condition is that the roots $z_1, \ldots, z_p$ of

$$1 - \phi_1 z - \ldots - \phi_p z^p = 0$$

are all outside of the unit circle.