Università di Pavia

Stationary ARMA Processes

Eduardo Rossi

$$Y_t = \mu + \epsilon_t + \theta\epsilon_{t-1} \quad t = 1, \ldots, T$$

$$\begin{aligned} \epsilon_t &\sim WN(0, \sigma^2) \\ E(\epsilon_t) &= 0 \\ E(\epsilon_t^2) &= \sigma^2 \\ E(\epsilon_t \epsilon_{t-j}) &= 0 \quad j \neq 0 \end{aligned}$$

$$\begin{aligned} E(Y_t) &= E(\mu + \epsilon_t + \theta\epsilon_{t-1}) \\ E(Y_t) &= \mu + E(\epsilon_t) + \theta E(\epsilon_{t-1}) \\ E(Y_t) &= \mu \end{aligned}$$

$$
\begin{aligned}
E(Y_t-\mu)^2 &= E(\epsilon_t+\theta\epsilon_{t-1})^2 \\
&= E(\epsilon_t^2+\theta^2\epsilon_{t-1}^2+2\theta\epsilon_{t-1}\epsilon_t) \\
&= \sigma^2+\theta^2\sigma^2+0 \\
&= (1+\theta^2)\sigma^2
\end{aligned}
$$

First Autocovariance

$$
\begin{aligned}
\gamma(1) &= E(Y_t-\mu)(Y_{t-1}-\mu) = E(\epsilon+\theta\epsilon_{t-1})(\epsilon_{t-1}+\theta\epsilon_{t-2}) \\
&= E(\epsilon_t\epsilon_{t-1}+\theta\epsilon_{t-1}^2+\theta\epsilon_t\epsilon_{t-2}+\theta^2\epsilon_{t-1}\epsilon_{t-2}) \\
&= 0+\theta\sigma^2+0+0=\theta\sigma^2
\end{aligned}
$$

Higher Autocovariances are all zero

$$
\gamma(j)=E[(Y_t-\mu)(Y_{t-j}-\mu)]=0
$$

MA(1) is covariance stationary regardless the value of θ.

$$\sum_{j=0}^{\infty} |\gamma(j)| = (1+\theta^2)\sigma^2 + |\theta\sigma^2| < \infty$$

If ϵ_t is a Gaussian White Noise, then MA(1) is ergodic for all moments.

Autocorrelation function

$$\rho(j) \equiv \frac{\gamma(j)}{\gamma(0)}$$

$$|\rho(j)| \leq 1$$

$$\begin{aligned} \rho(1) &= \frac{\theta\sigma^2}{(1+\theta^2)\sigma^2} = \frac{\theta}{1+\theta^2} \\ \rho(j) &= 0 \quad j > 0 \end{aligned}$$

The largest possible value for $\rho(1)$ is 0.5. This occurs if $\theta = 1$. The smallest is -0.5, $\theta = -1$. For $-0.5 < \rho(1) < 0.5$ there are two different values of θ that could produce that autocorrelation.

$\frac{\theta}{1+\theta^2}$ is unchanged if θ is replaced by $1/\theta$.

MA(1):

$$Y_t = \mu + \epsilon_t + \theta_1 \epsilon_{t-1}$$

$$Y_t = \mu + (1 + \theta_1 L)\epsilon_t$$

Autocorrelation function:

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

Replacing θ by $1/\theta$ and assuming that the (unobserved) shock process has a variance of $\theta^2\sigma^2$ instead of σ^2 yields a process with the same autocovariance structure as the original process.

The invertibility of $(1 + \theta_1 z)$ depends on the roots of

$$1 + \theta_1 z = 0$$

invertibility requires $|\theta| < 1$; if $|\theta| \geq 1$ the infinite sequence

$$(1 - \theta L + \theta^2 L^2 - \theta^3 L^3 + \ldots)$$

would not be well defined. For a MA(q) there are 2^q representations of the process having the same correlogram. Identification problem. To overcome this problem we impose the invertibility condition. AR(∞) representation:

$$\theta(L)^{-1} Y_t = \theta(1)^{-1}\mu + \epsilon_t$$

Moving average of order q (MA(q))

$$Y_t = \mu + \theta(L)\epsilon_t \quad t = 1, \ldots, T$$

$$\epsilon_t \sim WN(0, \sigma^2)$$

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

$$E(Y_t) = \mu$$

$$\begin{aligned} E[(Y_t - \mu)^2] &= E[(\theta(L)\epsilon_t)^2] \\ &= E(\epsilon_t^2 + \theta_1^2 \epsilon_{t-1}^2 + \ldots + \theta_q^2 \epsilon_{t-q}^2) \\ &= (1 + \theta_1^2 + \ldots + \theta_q^2)\sigma^2 \end{aligned}$$

because

$$E(\theta_i \theta_j \epsilon_{t-i}\epsilon_{t-j}) = 0 \quad \forall i \neq j \quad i,j = 0,\ldots,q \quad \theta_0 = 1$$

The autocovariance function

$$\begin{aligned}
\gamma(j) &= E[(\theta(L)\epsilon_t)(\theta(L)\epsilon_{t-j})] \\
&= E[(\epsilon_t + \ldots + \theta_j\epsilon_{t-j} + \ldots + \theta_q\epsilon_{t-q})(\epsilon_{t-j} + \theta_1\epsilon_{t-j-1} + \ldots + \theta_q\epsilon_{t-q-j})] \\
&= E(\theta_j\epsilon^2_{t-j} + \theta_1\theta_{j+1}\epsilon^2_{t-j-1} + \ldots + \theta_q\theta_{q-j}\epsilon^2_{t-q}) \\
&= (\theta_j + \theta_1\theta_{j+1} + \ldots + \theta_q\theta_{q-j})\sigma^2 \quad j = 1,\ldots,q
\end{aligned}$$

$$\gamma(j) = 0 \quad j > q$$

Example: MA(2)

$$Y_t = \mu + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2}$$

$$\gamma(0) = (1 + \theta_1^2 + \theta_2^2)\sigma^2$$

$$\begin{aligned} \gamma(1) &= E[(\epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2})(\epsilon_{t-1} + \theta_1 \epsilon_{t-2} + \theta_2 \epsilon_{t-3})] \\ &= \theta_1 E(\epsilon_{t-1}^2) + \theta_1 \theta_2 E(\epsilon_{t-2}^2) \\ &= (\theta_1 + \theta_1 \theta_2)\sigma^2 \end{aligned}$$

$$\gamma(j) = 0 \quad j = 3, 4, \ldots$$

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

$$\epsilon_t \sim WN(0, \sigma^2)$$

The infinite sequence generates a c.s. process provided that square summability holds

$$\sum_{j=0}^{\infty} \psi_j^2 < \infty$$

a slightly stronger condition is the absolute summability

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

$$E(Y_t) = \mu$$

$$\begin{aligned}
\gamma(0) &= E(Y_t - \mu)^2 \\
&= \lim_{T\to\infty} E(\psi_0 \epsilon_t + \psi \epsilon_{t-1} + \ldots + \psi_T \epsilon_{t-T}) \\
&= \lim_{T\to\infty} (\psi_0^2 + \psi_1^2 + \ldots + \psi_T^2)\sigma^2
\end{aligned}$$

$$\begin{aligned}
\gamma(j) &= E[(Y_t - \mu)(Y_{t-j} - \mu)] \\
&= \sigma^2(\psi_j \psi_0 + \psi_{j+1}\psi_1 + \ldots)
\end{aligned}$$

An MA(∞) with absolutely summable coefficients has absolutely summable covariances

$$\sum_{j=0}^{\infty} |\gamma(j)| < \infty$$

An MA(∞) absolutely summable is ergodic for the mean.

If $\epsilon_t \sim i.i.d.N(0, \sigma^2)$ then the process is ergodic for all moments.

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

$$\epsilon_t \sim WN(0, \sigma^2)$$

$$Y_t = c \sum_{j=0}^{T-1} \phi^j + \phi^T Y_{t-T} + \sum_{j=0}^{T-1} \phi^j \epsilon_{t-j}$$

Y_t is c.s. if $|\phi| < 1$.

First Order Difference Equation

$$(1 - \phi L)y_t = w_t$$

If $|\phi| < 1$ then backward solution:

$$y_t = w_t + \phi w_{t-1} + \ldots$$

If $|\phi| > 1$ forward solution based on

$$(1 - \phi L)^{-1} = \frac{-\phi^{-1}L^{-1}}{1 - \phi^{-1}L^{-1}} = -\phi^{-1}L^{-1}[1 + \phi^{-1}L^{-1} + \phi^{-2}L^{-2} + \ldots]$$

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

when it is applied to a bounded sequence $\{w_t\}_{t=-\infty}^{\infty}$ the result is another bounded sequence. Applying $(1 - \phi L)^{-1}$ we are implicitly imposing a boundedness assumption.

Premultiplying by

$$[1 + \phi^{-1}L^{-1} + \ldots + \phi^{-(T-1)}L^{-(T-1)}][-\phi^{-1}L^{-1}]$$

the limit of this operator exists and is $(1-\phi L)^{-1}$ when $|\phi| > 1$

$$(1-\phi L)^{-1} = [-\phi^{-1}L^{-1}][1 + \phi^{-1}L^{-1} + \ldots]$$

Applying this operator amounts to solving the difference equation forward.

For a AR(1) process with $|\phi| > 1$:

$$Y_t = (1 - \phi L)^{-1}\epsilon_t = [-\phi^{-1}L^{-1}][1 + \phi^{-1}L^{-1} + \ldots]\epsilon_t$$

$$Y_t = [-\phi^{-1}L^{-1}][\epsilon_t + \phi^{-1}\epsilon_{t+1} + \ldots] = -\sum_{j=1}^{\infty}\phi^{-j}\epsilon_{t+j}$$

this is the unique stationary solution. This is regarded as unnatural since Y_t is correlated with $\{\epsilon_s, s > t\}$ a property not shared by the solution obtained when $|\phi| < 1$.

It is customary when modelling stationary time series to restrict attention to AR(1) processes with $|\phi| < 1$ for which Y_t has the representation in terms of $\{\epsilon_s, s \le t\}$.

If $|\phi| = 1$ there is no stationary solution.

When Y_t is c.s. we can write:

$$\begin{aligned} Y_t &= (c+\epsilon_t) + \phi(c+\epsilon_{t-1}) + \phi^2(c+\epsilon_{t-2}) + \dots \\ &= c + c\phi + c\phi^2 + \dots + \epsilon_t + \phi\epsilon_{t-1} + \dots \\ &= \frac{c}{1-\phi} + \underbrace{\epsilon_t + \phi\epsilon_{t-1} + \phi^2\epsilon_{t-2}^2 + \dots}_{\text{MA}(\infty)_{\psi_j=\phi^j}} \end{aligned}$$

when $|\phi| < 1$

$$\sum_{j=0}^{\infty} |\psi_j| = \sum_{j=0}^{\infty} |\phi|^j = \frac{1}{1-|\phi|}$$

this ensures that the MA(∞) representation exists.

The AR(1) process is ergodic for the mean.

$$\mu \equiv E(Y_t) = \frac{c}{1-\phi}$$

$$\begin{aligned}
\gamma(0) &= E[(Y_t - \mu)^2] \\
&= E[(\epsilon_t + \phi\epsilon_{t-1} + \phi\epsilon_{t-2} + \ldots)^2] \\
&= (1 + \phi^2 + \phi^4 + \ldots)\sigma^2 \\
&= \sigma^2/(1-\phi^2)
\end{aligned}$$

$$\begin{aligned}
\gamma(j) &= E[(Y_t - \mu)(Y_{t-j} - \mu)] \\
&= E[(\epsilon_t + \phi\epsilon_{t-1} + \phi^2\epsilon_{t-2} + \ldots)(\epsilon_{t-j} + \phi\epsilon_{t-j-1} + \phi^2\epsilon_{t-j-2} + \ldots)] \\
&= (\phi^j + \phi^{j+2} + \phi^{j+4} + \ldots)\sigma^2 \\
&= \phi^j(1 + \phi^2 + \phi^4 + \ldots)\sigma^2 \\
&= \sigma^2(\phi^j/(1-\phi^2))
\end{aligned}$$

Autocorrelation function

$$\rho(j) = \frac{\gamma(j)}{\gamma(0)} = \phi^j$$

geometric decay.

$$\gamma(j) = \phi\gamma(j-1)$$

solution:

$$\gamma(j) = \phi^j\gamma(0)$$

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \epsilon_t$$

$$\epsilon_t \sim WN(0, \sigma^2)$$

$$(1 - \phi_1 L - \phi_2 L^2) Y_t = c + \epsilon_t$$

The difference equation is stable provided that the roots of

$$1 - \phi_1 z - \phi_2 z^2 = 0$$

lie outside the unit circle.

$$\psi(L) = (1 - \phi_1 L - \phi_2 L^2)^{-1} = \psi_0 + \psi_1 L + \psi_2 L^2 + \dots$$

$$\mu = \frac{c}{1 - \phi_1 - \phi_2}$$

The autocovariances

$$Y_t = \mu(1 - \phi_1 - \phi_2) + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \epsilon_t$$

$$Y_t - \mu = \phi_1(Y_{t-1} - \mu) + \phi_2(Y_{t-2} - \mu) + \epsilon_t$$

multiplying both sides by $(Y_{t-j} - \mu)$ and taking expectations produces

$$\begin{aligned} E[(Y_t - \mu)(Y_{t-j} - \mu)] &= \phi_1 E[(Y_{t-1} - \mu)(Y_{t-j} - \mu)] \\ &\quad + \phi_2 E[(Y_{t-2} - \mu)(Y_{t-j} - \mu)] + E[\epsilon_t(Y_{t-j} - \mu)] \end{aligned}$$

$$\gamma(j) = \phi_1 \gamma(j-1) + \phi_2 \gamma(j-2) \quad j = 1, 2, \ldots$$

The autocovariances follow the same the second-order difference equation as does the process for Y_t. The autocorrelations

$$\rho(j) = \phi_1 \rho(j-1) + \phi_2 \rho(j-2) \quad j = 1, 2, \ldots$$

Setting $j = 1$

$$\rho(1) = \phi_1 + \phi_2 \rho(1)$$

$$\rho(1) = \frac{\phi_1}{1 - \phi_2}$$

For $j = 2$

$$\rho(2) = \phi_1 \rho(1) + \phi_2$$

The variance of a c.s. AR(2)

$$E[(Y_t - \mu)^2] = \phi_1 E[(Y_{t-1} - \mu)(Y_t - \mu)] + \phi_2 E[(Y_{t-2} - \mu)(Y_t - \mu)] + E[(\epsilon_t)(Y_t - \mu)]$$

$$
\begin{aligned}
E(\epsilon_t)(Y_t - \mu) &= E(\epsilon_t)[\phi_1(Y_{t-1} - \mu) + \phi_2(Y_{t-2} - \mu) + \epsilon_t] \\
&= \phi_1 \cdot 0 + \phi_2 \cdot 0 + \sigma^2
\end{aligned}
$$

$$\gamma(0) = \phi_1\gamma(1) + \phi_2\gamma(2) + \sigma^2$$

$$\gamma(0) = \phi_1\rho(1)\gamma(0) + \phi_2\rho(2)\gamma(0) + \sigma^2$$

Substituting $\rho(1)$ and $\rho(2)$

$$
\begin{aligned}
\gamma(0) &= \left[\frac{\phi_1^2}{1-\phi_2} + \phi_2(\phi_1\rho(1) + \phi_2)\right]\gamma(0) + \sigma^2 \\
&= \left[\frac{\phi_1^2}{1-\phi_2} + \frac{\phi_2\phi_1^2}{1-\phi_2} + \phi_2^2\right]\gamma(0) + \sigma^2
\end{aligned}
$$

$$
\begin{aligned}
\gamma(0) &= \left[1 - \frac{\phi_1^2}{1-\phi_2} - \frac{\phi_2\phi_1^2}{1-\phi_2} - \phi_2^2\right]^{-1} \sigma^2 \\
&= \left[\frac{1-\phi_2-\phi_1^2-\phi_2\phi_1^2-\phi_2^2(1-\phi_2)}{1-\phi_2}\right]^{-1} \sigma^2 \\
&= \frac{(1-\phi_2)\sigma^2}{1-\phi_2-\phi_1^2-\phi_2\phi_1^2-\phi_2^2(1-\phi_2)} \\
&= \frac{(1-\phi_2)\sigma^2}{1-\phi_2-\phi_1^2-\phi_2\phi_1^2-\phi_2^2(1-\phi_2)} \\
&= \frac{(1-\phi_2)\sigma^2}{(1+\phi_2)[(1-\phi_2)^2-\phi_1^2]}
\end{aligned}
$$

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + \epsilon_t$$

$$\epsilon_t \sim WN(0, \sigma^2)$$

provided that the roots of

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

all lie the unit circle.

Covariance-stationary representation:

$$\begin{aligned} Y_t &= \mu + \psi(L)\epsilon_t \\ &= \frac{c}{1 - \phi_1 - \phi_2 - \ldots - \phi_p} + \frac{1}{1 - \phi_1 L - \ldots - \phi_p L^p}\epsilon_t \end{aligned}$$

where

$$\psi(z) = (1 - \phi_1 z - \ldots - \phi_p z^p)^{-1} = \phi(z)^{-1}$$

and

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

The mean is

$$\mu = E(Y_t) = \frac{c}{1 - \phi_1 - \ldots - \phi_p}$$

$$Y_t - \mu = \phi_1(Y_{t-1} - \mu) + \phi_2(Y_{t-2} - \mu) + \ldots + \phi_p(Y_{t-p} - \mu) + \epsilon_t$$

Autocovariances are found by multiplying both sides by $(Y_{t-j} - \mu)$ and taking expectations

The autocovariance function

$$\gamma(j) = \begin{cases} \phi_1\gamma(j-1) + \phi_2\gamma(j-2) + \ldots + \phi_p\gamma(j-p) & j = 1, 2, \ldots \\ \phi_1\gamma(1) + \ldots + \phi_p\gamma_p + \sigma^2 & j = 0 \end{cases}$$

Dividing the autocovariance function by γ_0 we obtain the Yule-Walker equations:

$$\rho_j = \phi_1 \rho_{j-1} + \phi_2 \rho_{j-2} + \ldots + \phi_p \rho_{j-p} \quad j = 1, 2, \ldots$$

Thus the autocovariances and autocorrelations follow the same p-th order difference equation as does the process itself. For distinct roots, their solutions take the form

$$\gamma(j) = g_1 \lambda_1^j + g_2 \lambda_2^j + \ldots + g_p \lambda_p^j$$

where the eigenvalues $(\lambda_1, \ldots, \lambda_p)$ are the solutions to

$$\lambda^p - \phi_1 \lambda^{p-1} - \ldots - \phi_p = 0$$

$$Y_t = c + \phi_1 Y_{t-1} + \ldots + \phi_p Y_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \ldots + \theta_q \epsilon_{t-q}$$

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

$$\phi(L) Y_t = c + \theta(L) \epsilon_t$$

where

$$\begin{aligned} \phi(L) &= 1 - \phi_1 L - \ldots - \phi_p L^p \\ \theta(L) &= 1 + \theta_1 + \ldots + \theta_q L^q \end{aligned}$$

the stationarity depends on the roots of

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

If the roots are outside the unit circle then the inverse of $\phi(z)$ exists, then dividing by $\phi(L)$ both sides

$$Y_t = \mu + \psi(L)\epsilon_t$$

$$\psi(L) = \frac{\theta(L)}{\phi(L)}$$

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

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

$$\begin{aligned} c &= \mu(1 - \phi_1 - \ldots - \phi_p) \\ Y_t &= \mu(1 - \phi_1 - \ldots - \phi_p) + \phi_1 Y_{t-1} + \ldots + \phi_p Y_{t-p} + \epsilon_t + \ldots + \theta_q \epsilon_{t-q} \\ Y_t - \mu &= \phi_1 (Y_{t-1} - \mu) + \ldots + \phi_p (Y_{t-p} - \mu) + \epsilon_t + \theta_1 \epsilon_{t-1} + \ldots + \theta_q \epsilon_{t-q} \end{aligned}$$

The variance

$$E[(Y_t - \mu)^2] = \phi_1 E[(Y_{t-1} - \mu)(Y_t - \mu)] + \ldots + \phi_p E[(Y_{t-p} - \mu)(Y_t - \mu)]$$
$$+ E[\epsilon_t (Y_t - \mu)] + \theta_1 E[\epsilon_{t-1}(Y_t - \mu)] + \ldots + \theta_q E[\epsilon_{t-q}(Y_t - \mu)]$$

$$E[(Y_t - \mu)^2] = \phi_1 [\sigma^2 (\psi_1 \psi_0 + \psi_2 \psi_1 + \ldots)] + \ldots +$$
$$\phi_p [\sigma^2 (\psi_p \psi_0 + \psi_{p+1} \psi_1 + \ldots)]$$
$$+ E[\psi_0 \epsilon_t^2] + \theta_1 E[\psi_1 \epsilon_{t-1}^2] + \ldots + \theta_q E[\psi_q \epsilon_{t-q}^2]$$

$$
\begin{aligned}
&E[(Y_t - \mu)^2] = \phi_1[\sigma^2(\psi_1\psi_0 + \psi_2\psi_1 + \ldots)] + \ldots + \\
&\phi_p[\sigma^2(\psi_p\psi_0 + \psi_{p+1}\psi_1 + \ldots)] \\
&+\psi_0\sigma^2 + \theta_1\psi_1\sigma^2 + \ldots + \theta_q\psi_q\sigma^2
\end{aligned}
$$

An ARMA(p,q) process will have more complicated autocovariances for lags 1 through q than would the corresponding AR(p) process

$$
\begin{aligned}
\gamma(j) &= E[(Y_t - \mu)(Y_{t-j} - \mu)] \\
&= \phi_1 E[(Y_{t-1} - \mu)(Y_{t-j} - \mu)] + \ldots + \phi_p E[(Y_{t-p} - \mu)(Y_{t-j} - \mu)] \\
&\quad + E[\epsilon_t(Y_{t-j} - \mu)] + \theta_1 E[\epsilon_{t-1}(Y_{t-j} - \mu)] + \ldots + \theta_q E[\epsilon_{t-q}(Y_{t-j} - \mu)]
\end{aligned}
$$

For $j > q$ autocovariances are given by

$$E[(Y_t - \mu)(Y_{t-j} - \mu)] =$$
$$\phi_1 E[(Y_{t-1} - \mu)(Y_{t-j} - \mu)] + \ldots + \phi_p E[(Y_{t-p} - \mu)(Y_{t-j} - \mu)]$$
$$+E[\epsilon_t(Y_{t-j} - \mu)] + \theta_1 E[\epsilon_{t-1}(Y_{t-j} - \mu)] + \ldots + \theta_q E[\epsilon_{t-q}(Y_{t-j} - \mu)]$$

$$\begin{aligned} E[\epsilon_{t-q}(Y_{t-j} - \mu)] &= E[\epsilon_{t-q}(\psi(L)\epsilon_{t-j})] \\ &= E[\epsilon_{t-q}(\psi_0\epsilon_{t-j} + \psi_1\epsilon_{t-j-1} + \ldots)] \\ &= \psi_0 E[\epsilon_{t-q}\epsilon_{t-j}] + \psi_1 E[\epsilon_{t-q}\epsilon_{t-j-1}] + \ldots \\ &= 0 \end{aligned}$$

then

$$E[(Y_t - \mu)(Y_{t-j} - \mu)] =$$
$$\phi_1 E[(Y_{t-1} - \mu)(Y_{t-j} - \mu)] + \ldots + \phi_p E[(Y_{t-p} - \mu)(Y_{t-j} - \mu)]$$

$$\gamma(j) = \phi_1 \gamma(j-1) + \phi_2 \gamma(j-2) + \ldots + \phi_p \gamma(j-p) \quad j = q+1, q+2, \ldots$$

Thus after q lags the autocovariance function follow the p-th order difference equation governed by the autoregressive parameters.

There is a potential for redundant parameterization with ARMA processes. Consider a simple white noise process

$$Y_t = \epsilon_t$$

Suppose both sides are multiplyied by $(1 - \rho L)$:

$$(1 - \rho L)Y_t = (1 - \rho L)\epsilon_t$$

Both are valid representations, thus the latter might be described as an ARMA(1,1) process, with $\phi_1 = \rho$ and $\theta_1 = -\rho$. Since any value of ρ describes the data equally well, we will get into trouble trying to estimate the parameter ρ by maximum likelihood.

A related overparameterization can arise with an ARMA(p,q) model. Consider the factorization of the lag polynomial operators:

$$(1-\lambda_1 L)(1-\lambda_2 L)\ldots(1-\lambda_p L)(Y_t-\mu) = (1-\eta_1 L)\ldots(1-\eta_q L)\epsilon_t$$

Assume that $|\lambda_i| < 1$ for all i, so that the process is c.s.. If $\phi(L)$ and $\theta(L)$ have any roots in common, $\lambda_i = \eta_j$ for some i and j, then both sides can be divided by $(1-\lambda_i L)$

$$\prod_{k=1,k\neq i}^{p}(1-\lambda_k L)(Y_t-\mu) = \prod_{k=1,k\neq j}^{q}(1-\eta_k L)\epsilon_t$$

$$(1-\phi_1^* L - \ldots - \phi_{p-1}^* L^{p-1})(Y_t - \mu) = (1+\theta_1^* L + \ldots + \theta_{q-1}^* L^{q-1})\epsilon_t$$

where

$$(1-\phi_1^* L - \ldots - \phi_{p-1}^* L^{p-1}) \equiv \prod_{k=1, k\neq i}^{p} (1-\lambda_k L)$$

$$(1+\theta_1^* L + \ldots + \theta_{q-1}^* L^{q-1}) \equiv \prod_{k=1, k\neq j}^{q} (1-\eta_k L)$$

The stationary process ARMA(p,q) process is clearly identical to the stationary ARMA(p-1,q-1) process.

$$Y_t = c + \phi_1 Y_{t-1} + \epsilon_t + \theta_1 \epsilon_{t-1}$$

$$(1 - \phi_1 L) Y_t = c + (1 + \theta_1 L)\epsilon_t$$

$$|\phi_1| < 1$$

$$\psi(L) = \frac{1 + \theta_1 L}{1 - \phi_1 L} = (1 + \theta_1 L)(1 + \phi_1 L + \phi_1^2 L^2 + \ldots)$$

$$\psi_0 + \psi_1 L + \psi_2 L^2 + \ldots = (1 + \theta_1 L)(1 + \phi_1 L + \phi_1^2 L^2 + \ldots)$$

$$\begin{aligned}\psi_0 &= 1\\ \psi_1 &= \theta_1 + \psi_1\\ \psi_2 &= \phi_1^2 + \phi_1\theta_1\\ \ldots &= \ldots\end{aligned}$$

$$\gamma(0) = \phi_1 E[(Y_{t-1} - \mu)(Y_t - \mu)] + E[\epsilon_t(Y_t - \mu)] + \theta_1 E[\epsilon_{t-1}(Y_t - \mu)]$$

$$\gamma(0) = \phi_1[\sigma^2(\psi_1\psi_0 + \psi_2\psi_1 + \ldots)] + \sigma^2 + \theta_1\psi_1\sigma^2$$

$$\gamma(1) = \phi_1\gamma(0) + \theta_1\sigma^2$$

$$\gamma(2) = \phi_1\gamma(1)$$