B.Sc./Grad. Dip.: Probability Models and Time Series
MAS programmes: Stochastic Models and Time Series

4 The properties of AR(1) and MA processes

4.1 Autocovariances and autocorrelations for an AR(1) process

We derive the autocovariance and autocorrelation functions for an AR(1) process $\{Y_t\}$, using two alternative methods. The first method is based upon the use of the model equation,

$$Y_t = \phi Y_{t-1} + \epsilon_t, \qquad t \in \mathbb{Z}. \tag{1}$$

Recall the infinite moving average expression, → due to $\sum_{i=0}^{n-1} \phi^i \varepsilon_{t-i}$ converging in mean square.

↳ $\sum_{i=0}^{n-1} \phi^i \varepsilon_{t-i} \xrightarrow{ms} Y_t$ as $n \to \infty$

$$Y_t = \sum_{i=0}^{\infty} \phi^i \epsilon_{t-i}, \tag{2}$$

shown by ↳ $E\left[\left(\sum_{i=0}^{n-1} \phi^i \varepsilon_{t-i} - Y_t\right)^2\right] = E[(\phi^n Y_{t-n})^2] \to 0$ as $n \to \infty$

derived in Section 3.7. The representation was shown to hold in the sense of mean square convergence for $|\phi| < 1$. However, as we will see later, the representation also holds in the sense of "almost sure" or "probability one" convergence: with such an interpretation, it is apparent that Y_t and ϵ_{t+i} are uncorrelated for $i \geq 1$. Therefore, it can be argued that Y_s and ϵ_t are uncorrelated for $s < t$, so that, since the process mean is zero,

i.e. when there is a shift between the white noise term & the process term

Recall that $E[Y_s]=0$ $E[\varepsilon_t]=0$ ∴ $Cov(Y_s, \varepsilon_t) = E[(Y_s - \mu_Y)(\varepsilon_t - \mu_\varepsilon)]$ → $\mu_Y = 0$, $\mu_\varepsilon = 0$

$$E[Y_s \epsilon_t] = 0 \qquad s < t.$$

$= E[Y_s \varepsilon_t] = 0$

since Y_s & ε_t uncorrelated for $t > s$

Squaring Equation (1) and taking expectations,

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

$$E[Y_t^2] = \phi^2 E[Y_{t-1}^2] + 2\phi E[Y_{t-1}\epsilon_t] + E[\epsilon_t^2],$$

0 $\qquad \sigma^2$

i.e.,

$E[Y_t]^2 = Var(Y_t) = Cov(Y_t, Y_t) = \gamma_0$

$$\gamma_0 = \phi^2 \gamma_0 + \sigma^2,$$

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

where $\sigma^2 = \text{var}(\epsilon_t)$. Hence,

$$\boxed{\gamma_0 = \frac{\sigma^2}{1 - \phi^2}} \tag{3}$$

→ $|\phi| < 1$

Multiplying Equation (1) by $Y_{t-\tau}$, where $\tau \geq 1$, and taking expectations, yields

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

$E[Y_t Y_{t-\tau}] = \phi[Y_{t-1} Y_{t-\tau}] + E[\varepsilon_t Y_{t-\tau}]$

$$\gamma_\tau = \phi \gamma_{\tau-1} \qquad \tau \geq 1. \tag{4}$$

$\gamma_{\tau-1} = \phi \gamma_{\tau-2}$

$\gamma_\tau = \phi^2 \gamma_{\tau-2}$

Equation (4) together with the initial condition of Equation (3) has the solution

$Cov(Y_t, Y_{t-\tau}) = \gamma_\tau$

$Cov(Y_{t-1}, Y_{t-\tau}) = \gamma$

$$\gamma_\tau = \frac{\sigma^2}{1 - \phi^2}\, \phi^\tau \qquad \tau \geq 0. \tag{5}$$

$\gamma_\tau = \phi^\tau \gamma_0$ ∴ $\rho_\tau = \phi^\tau \rho_0 = \phi^\tau$

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

Recalling that $\rho_\tau = \gamma_\tau / \gamma_0$, we obtain

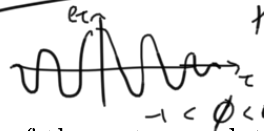
$$\rho_\tau = \phi^\tau \qquad \tau \geq 0. \tag{6}$$

remember, it's the difference between them:

and know $\gamma_1 = \frac{\sigma^2}{1-\phi^2}$

∴ $\rho_\tau = \dfrac{\cancel{\left(\frac{\sigma^2}{1-\phi^2}\right)} \phi^\tau}{\cancel{\left(\frac{\sigma^2}{1-\phi^2}\right)}}$

Autocorrelation function, $e_\tau = \phi^\tau$ ∴ if $\phi < 0$, then will oscillate, if $\phi > 0, < 1$, then will curve towards 0.



Note the geometric decline of the autocorrelation function. If $\phi < 0$ then the autocorrelation function oscillates and has negative correlation at lag 1.

A slight variant of this method for obtaining the expression for ρ_τ is to divide through in Equation (4) by γ_0 to obtain the recurrence relation

$$\rho_\tau = \phi \rho_{\tau-1} \qquad \tau \geq 1. \tag{7}$$

Equation (7) together with the initial condition $\rho_0 = 1$ has the solution obtained previously as Equation (6). Using the symmetry property that $\gamma_{-\tau} = \gamma_\tau$ and $\rho_{-\tau} = \rho_\tau$, we may, if we wish, extend the range of the values of τ in the solutions for the autocovariance function $\{\gamma_\tau\}$ and the autocorrelation function $\{\rho_\tau\}$. Thus, for example, we may write

$$\boxed{\rho_\tau = \phi^{|\tau|}} \qquad \tau \in \mathbb{Z}.$$

An alternative approach to finding the expressions for $\{\gamma_\tau\}$ and $\{\rho_\tau\}$ is based upon use of the infinite moving average expression of Equation (2). For $\tau \geq 0$,

$Y_t = \sum_{i=0}^{\infty} \phi^i \varepsilon_{t-i} \quad Y_{t-\tau} = \sum_{j=0}^{\infty} \phi^j \varepsilon_{t-\tau-j}$

$$
\begin{aligned}
Cov(Y_t, Y_{t-\tau}) = \gamma_\tau &= E[Y_t Y_{t-\tau}] \\
&= E\left[\sum_{i=0}^{\infty} \phi^i \epsilon_{t-i} \sum_{j=0}^{\infty} \phi^j \epsilon_{t-\tau-j}\right] \\
&= E\left[\sum_{i=0}^{\infty} \phi^i \epsilon_{t-i} \sum_{j=\tau}^{\infty} \phi^{j-\tau} \epsilon_{t-j}\right] \\
&= \sum_{i=0}^{\infty} \sum_{j=\tau}^{\infty} \phi^{i+j-\tau} E[\epsilon_{t-i}\epsilon_{t-j}] \\
&= \sum_{j=\tau}^{\infty} \phi^{2j-\tau} \sigma^2 \\
&= \frac{\sigma^2}{1-\phi^2} \phi^\tau.
\end{aligned}
$$

change summation start index at $j = \tau$, and then change the $j \to j - \tau$.

$\sum_{i=0}^{\infty} x_i \sum_{j=0}^{\infty} y_j = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} x_i y_j$

→ linear property of expectation operator.

$E\left[\sum_{i=0}^{\infty} X_i\right] = E\left[\lim_{n\to\infty} \sum_{i=0}^{n} X_i\right] = \lim_{n\to\infty} E\left[\sum_{i=0}^{n} X_i\right] = \lim_{n\to\infty} n E[X_i]$

$Cov(\varepsilon_t, \varepsilon_{t-\tau}) = 0$ when $\tau \neq 0$, since white noise is uncorrelated by definition.

↳ $E[\varepsilon_{t-i}\varepsilon_{t-j}]$ only non-zero when $i = j$

Thus we have again the result of Equation (5). Note that in the above derivation we have used the properties of the autocovariance function of a white noise process as given in Equations (1) and (2) of Section 3.3.

$\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \phi^{i+j-\tau} E[\varepsilon_{t-i}\varepsilon_{t-j}] = \sum_{j=0}^{\infty} \phi^{2j-\tau} \underbrace{E[\varepsilon_{t-j}, \varepsilon_{t-j}]}_{Cov(\varepsilon_{t-j}, \varepsilon_{t-j})}$

all zero except when $i = j$

↳ $Var(\varepsilon_{t-j}) = \sigma^2$

4.2 The lag operator

It is often convenient to use the *lag operator* L to characterize models and to carry out mathematical manipulations. (An alternative terminology is *backward shift operator* with corresponding notation B.) The operator is defined by

$$\boxed{LY_t = Y_{t-1} \qquad t \in \mathbb{Z}.}$$

Applying L operator to a random variable at t returns same r.v. at $t-1$.

Defining L^j to be the 'j-fold' composition of L, then we note that

$$L^j Y_t = Y_{t-j}.$$

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

Applying L j-times means go back j steps.

↳ $L^j Y_t = Y_{t-j}$

THE ORDER OF HOW L is applied to Y_t is IMPORTANT! $LY_t = Y_{t-1}$

$LY_t = Y_{t-1}$

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

$Y_t = \phi L Y_t + \varepsilon_t$

$Y_t - \phi L Y_t = \varepsilon_t$

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

Equation (1) for the AR(1) model may be written as

$$(1 - \phi L)Y_t = \epsilon_t.$$

$Y_t L \neq Y_{t-1}$

↓ meaningless!

The infinite moving average representation of Y_t may be more simply derived using the formalism (and associated algebra) of the lag operator.

simply divide through by $(1-\phi L)$

$$\begin{aligned} Y_t &= (1 - \phi L)^{-1}\epsilon_t \\ &= \sum_{i=0}^{\infty} \phi^i L^i \epsilon_t \\ &= \sum_{i=0}^{\infty} \phi^i \epsilon_{t-i}, \end{aligned}$$

$Y_t = \sum_{i=0}^{\infty} \phi^i \varepsilon_{t-i} = \sum_{i=0}^{\infty} \phi^i L^i \varepsilon_t$

$\boxed{L^i \varepsilon_t = \varepsilon_{t-i}}$

$(1-\phi L)^{-1} = \sum_{i=0}^{\infty} \phi^i L^i$

assuming that the sum converges.

The reason it's called a linear process is because it's a linear combination of white noise terms.

4.3 Linear processes

Recall that $Y_t = \sum_{i=0}^{\infty} \phi^i \varepsilon_{t-i}$

$|\phi| < 1$ is a constant

$\{\phi^i : i \in \mathbb{N}\}$ → this could be seen as a sequence of coefficients

Definition 4.3.1 (Linear Process)

A process $\{Y_t\}$ that has the representation

time series process

LINEAR PROCESS REPRESENTATION $\{Y_t\}$

$$Y_t = \sum_{i=0}^{\infty} \psi_i \epsilon_{t-i}, \tag{8}$$

where $\{\epsilon_t\}$ is a white noise process and $\{\psi_i\}$ is a sequence of coefficients such that $\sum_{j=0}^{\infty} |\psi_j| < \infty$ is referred to as a *linear process*.

The sum of all the coefficients is less than ∞

$Y_t = \lim_{n\to\infty} \sum_{i=0}^{n} \psi_i \varepsilon_{t-i} = \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i}$

Proposition 4.3.2 (Convergence and Stationarity of Linear Process)

For a linear process $\{Y_t\}$, with representation $Y_t = \sum_{i=0}^{\infty} \psi_i \epsilon_{t-i}$:

Infinite sum from $0 \to \infty$ of a coefficient and w.n term.

(a) $Y_t = \sum_{i=0}^{\infty} \psi_i \epsilon_{t-i}$ is almost surely bounded (i.e. bounded with probability one);

(b) the sequence of partial sums, $\sum_{i=0}^{n} \psi_i \epsilon_{t-i}$, converges to Y_t in mean square as $n \longrightarrow \infty$;

(c) $\{Y_t\}$ is (weakly) stationary, with mean zero and autocovariance function given by

$\gamma_\tau = Cov(Y_t, Y_{t-\tau}) = E[(Y_t-\mu)(Y_{t-\tau}-\mu)]$

$= E[Y_t Y_{t-\tau}]$

$$\begin{aligned} \gamma_\tau &= E[Y_t Y_{t-\tau}] = Cov(Y_t, Y_{t-\tau}) = E[(Y_t-\mu)(Y_{t-\tau}-\mu)] \\ &= E\left[\sum_{i=0}^{\infty} \psi_i \epsilon_{t-i} \sum_{j=0}^{\infty} \psi_j \epsilon_{t-\tau-j}\right] \\ &= E\left[\sum_{i=0}^{\infty} \psi_i \epsilon_{t-i} \sum_{j=\tau}^{\infty} \psi_{j-\tau} \epsilon_{t-j}\right] \\ &= \sum_{i=0}^{\infty}\sum_{j=\tau}^{\infty} \psi_i \psi_{j-\tau} \underbrace{E[\epsilon_{t-i}\epsilon_{t-j}]}_{=0 \ \forall\ i \neq j} \\ &= \sigma^2 \sum_{j=\tau}^{\infty} \psi_j \psi_{j-\tau} \qquad \tau \geq 0. \end{aligned} \tag{9}$$

$Y_t = \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i}$

reset the indexing of second sum to start at $j=\tau$, and then all your j's must be $j-\tau$.

$Y_t = \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i}$

$Y_{t-\tau} = \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-\tau-i}$

$\gamma_\tau = E\left[\sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i} \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-\tau-j}\right]$

$t-\tau-(j-\tau) = t-\tau+\tau-j = t-j$

$\gamma_\tau = E\left[\sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i} \sum_{j=\tau}^{\infty} \psi_{j-\tau} \varepsilon_{t-j}\right]$

$\gamma_\tau = E\left[\sum_{i=0}^{\infty}\sum_{j=\tau}^{\infty} \psi_i \varepsilon_{t-i} \psi_{j-\tau} \varepsilon_{t-j}\right]$

$= \lim_{n\to\infty} \sum_{i=0}^{n}\sum_{j=\tau}^{n} \psi_i \psi_{j-\tau} E[\varepsilon_{t-i}\varepsilon_{t-j}]$

will be 0 for all i, except when $i=j$

$\sum_{j=\tau}^{\infty} \psi_j \psi_{j-\tau} Cov(\varepsilon_{t-j}, \varepsilon_{t-j})$

$= \sum_{j=\tau}^{\infty} \psi_j \psi_{j-\tau} \sigma^2$

Recall: white noise terms are uncorrelated, unless they're the same.

$= \sigma^2 \sum_{j=\tau}^{\infty} \psi_j \psi_{j-\tau}.$

$\gamma_\tau = \mathrm{Cov}(Y_t, Y_{t-\tau}) = \sigma^2 \sum_{j=\tau}^{\infty} \psi_j \psi_{j-\tau}$

When $\tau = 0$, $\gamma_0 = \mathrm{Cov}(Y_t, Y_t) = \mathrm{Var}(Y_t) = \boxed{\sigma^2 \sum_{j=0}^{\infty} \psi_j^2}$

But we know for Y_t to be a linear process: $Y_t = \sum_{i=0}^{\infty} \psi_i \epsilon_{t-i}$ and that $\sum_{i=0}^{\infty} |\psi_i| < \infty$

In particular,

$$\boxed{\text{WHEN } \tau = 0} \quad \mathrm{var}(Y_t) = \gamma_0 = \sigma^2 \sum_{j=0}^{\infty} \psi_j^2.$$

So we know that $\mathrm{Var}(Y_t)$ is finite

But since $\sum_{j=0}^{\infty} |\psi_j| < \infty$, then $\sum_{j=0}^{\infty} \psi_j^2 < \infty$, and so $\mathrm{var}(Y_t)$ is finite. Finiteness of γ_τ for $\tau \neq 0$ follows from the calculation to be used in Examples 4 Qu. 3. □

Remarks 4.3.3

(i) A linear process $\{Y_t\}$, as defined above, is sometimes said to be *causal* or *a causal function of* $\{\epsilon_t\}$.

$Y_t = \sum_{i=0}^{\infty} \phi^i \epsilon_{t-i} = \sum_{i=0}^{\infty} \psi_i \epsilon_{t-i} \rightarrow$ Value of Y_t only depends on white noise terms from prev. periods.

(ii) We see from the infinite moving average representation of Equation (8) that the value Y_s of the process at time s depends only on $\{\epsilon_t : t \leq s\}$ and not on $\{\epsilon_t : t > s\}$.

Hence: Y_t and ϵ_t are uncorrelated if $t > s$.

(iii) As for the special case of the AR(1) process, Y_s and ϵ_t are uncorrelated for $s < t$, so that

$$\boxed{E[Y_s \epsilon_t] = 0, \qquad s < t,} \tag{10}$$

a fact that we shall make use of later.

Good to know how $E[Y_t, Y_{t-\tau}]$, $E[Y_t, \epsilon_t]$, and $E[\epsilon_t, \epsilon_{t-\tau}]$

Hence why so! The $\psi_i \to \phi^i$ varies as this is a special case of the linear process.

(iv) In the special case of $\{\psi_i = \phi^i : i \in \mathbb{N}\}$ we have the infinite moving average representation (2) of the AR(1) process, and thus $\sum_{j=0}^{\infty} |\psi_j| < \infty$ is satisfied if and only if $|\phi| < 1$. It is readily checked by substitution that the process $\{Y_t\}$ defined by $Y_t = \sum_{i=0}^{\infty} \phi^i \epsilon_{t-i}$ satisfies Equation (1). It follows that a viable (i.e. stationary) AR(1) process with autoregressive parameter ϕ exists if $|\phi| < 1$.

$\mathrm{Var}(Y_t) = \sigma^2 \sum_{j=0} \psi_j^2$

The ϕ^i comes in the infinite moving average case.

(v) The square summability of the coefficients i.e. $\sum_{j=0}^{\infty} \psi_j^2 < \infty$ is actually both necessary and sufficient for mean square convergence of the linear process representation.

When $\psi_i = \phi^i$, need $\sum_{i=0} \phi^i < \infty$

$Y_t = \sum_{i=0}^{\infty} \phi^i \epsilon_{t-i} = \sum_{i=0}^{\infty} \psi_i \epsilon_{t-i} \leftarrow$ linear process

special case

$\sum_{i=1}^{\infty} |\psi_i| < \infty$ and $\sum_{i=0}^{\infty} \psi_i^2 < \infty \rightarrow$ finite variance

4.4 The moving average model

New info represented by the white noise term.

Consider a market that every working day receives fresh information which affects the price of a certain commodity. Let Y_t denote the price change on day t. The immediate effect of the information received on day t upon Y_t is represented by ϵ_t, where $\{\epsilon_t\}$ is assumed to be a white noise process. But there is also a residual effect, such that Y_t is affected by the information received on the q previous days. A simple model represents $\{Y_t\}$ as a moving average process.

A *moving average process of order q*, an MA(q) process, with zero mean is a process $\{Y_t\}$ which satisfies the relation

MA(q) process

$$\boxed{Y_t = \epsilon_t + \sum_{i=1}^{q} \theta_i \epsilon_{t-i} \qquad t \in \mathbb{Z},} \tag{11}$$

So it includes the white noise term on t, as well as a bunch of other white noise terms pre-t.

where $\{\epsilon_t\}$ is a white noise process with mean zero and variance σ^2 and where $\theta_q \neq 0$.

For convenience, we may define $\theta_0 = 1$ and rewrite Equation (11) as

or start at $i=0$ for the present day.

$$Y_t = \sum_{i=0}^{q} \theta_i \epsilon_{t-i} \qquad t \in \mathbb{Z}. \tag{12}$$

LINEAR process

$Y_t = \sum_{i=0}^{\infty} \psi_i \epsilon_{t-i} = \sum_{i=0}^{\infty} \phi^i \epsilon_{t-i}$

So, AR(1) $\Rightarrow$

AR(p), the p is the number of previous Y values you regress on

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

MA(q) is a moving average process, where q is the number of θ's in the sum

$\boxed{MA(q) = \epsilon_t + \sum_{i=1}^{q} \theta_i \epsilon_{t-i}}$

o MA(q) → FINITE MOVING AVERAGE

o AR(1) → can be represented by an INFINITE MOVING AVERAGE

→ MA(FINITE q)

- Note that we are dealing here with a *finite* moving average, having a finite number of moving average parameters, as against the *infinite* moving average representation that we described in the discussion of the AR(1) process.

↳ $Y_t = \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i} = \sum_{i=0}^{\infty} \phi^i \varepsilon_{t-i}$

the weights sum to 1.

- Historically, "moving averages" were introduced rather differently, with the coefficients θ_i defined in such a way that $\sum \theta_i = 1$. Each Y_t value is then a weighted average of the ϵ_t values. The average "moves" as t moves through successive values.

↳ $Y_t = \sum_{i=0}^{q} \theta_i \varepsilon_{t-i}$ → θ_i is the weight for lag $\tau = i$

An MA(q) process is a special case of a linear process as defined in Section 4.3, with

AR(1) : $Y_t = \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i} = \sum_{i=0}^{\infty} \phi^i \varepsilon_{t-i}$

MA(q) : $Y_t = \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i} = \sum_{i=0}^{q} \theta_i \varepsilon_{t-i}$

coefficients $\psi_i = 0$ for $i > q$.

$$\boxed{\psi_i = \begin{cases} \theta_i & \text{for } 0 \le i \le q \\ 0 & \text{for } i > q \end{cases}.}$$

→ remember linear process is an infinite sum.

$Y_t = \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i}$

Because all but a finite number of the coefficients ψ_i are zero, an MA(q) process necessarily satisfies the summability conditions on the coefficients in the definition of a linear process. Thus, for all moving average parameter values, $\theta_1, \theta_2, \ldots, \theta_q$, Equation (11) or (12) defines a stationary process with mean zero.

Applying the result of Equation (9) to the present case, for $\tau \ge 0$,

Recall: $\gamma_\tau = Cov(Y_t, Y_{t-\tau}) = E[(Y_t - \mu)(Y_{t-\tau} - \mu)]$

$= E[Y_t Y_{t-\tau}] = E[\sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i} \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-\tau-j}]$

$= \sum_{i=0} \sum_{j=\tau} \psi_i \psi_{j-\tau} E[\varepsilon_{t-j} \varepsilon_{t-j}]$

$Y_t = \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i}$

$= \sum_{j=\tau}^{\infty} \psi_j \psi_{j-\tau} Cov(\varepsilon_{t-j}, \varepsilon_{t-j})$

$\boxed{= \sigma^2 \sum_{j=\tau}^{\infty} \psi_j \psi_{j-\tau}}$

$$\gamma_\tau = \sigma^2 \sum_{j=\tau}^{q} \theta_j \theta_{j-\tau}.$$

MA(q) $\gamma_\tau = Var(Y_t) = \sigma^2 \sum_{j=\tau} \psi_j \psi_{j-\tau}$

$\boxed{\gamma_\tau = Var(Y_t) = \sigma^2 \sum_{j=\tau}^{q} \theta_j \theta_{j-\tau}}$

↳ know this will be $< \infty$, and therefore you will have finite variance.

↓ rejigging indices

Thus

$$\gamma_\tau = \begin{cases} \sigma^2 \sum_{i=0}^{q-\tau} \theta_i \theta_{i+\tau} & \text{for } 0 \le \tau \le q \\ 0 & \text{for } \tau > q \end{cases}.$$

In particular,

$$\text{var}(Y_t) = \gamma_0 = \sigma^2 \sum_{i=0}^{q} \theta_i^2.$$

→ $Var(Y_t) = \gamma_0 = Cov(Y_t, Y_t)$

$= \sigma^2 \sum_{j=0}^{q} \theta_j^2$

The autocorrelation function is given by

$\rho_\tau = \frac{\gamma_\tau}{\gamma_0} = \frac{\sigma^2 \sum_{i=0}^{q-\tau} \theta_i \theta_{i+\tau}}{\sum_{i=0}^{q} \theta_i^2}$

$$\rho_\tau = \begin{cases} \sum_{i=0}^{q-\tau} \theta_i \theta_{i+\tau} / \sum_{i=0}^{q} \theta_i^2 & \text{for } 0 \le \tau \le q \\ 0 & \text{for } \tau > q \end{cases}.$$

→ Auto correlation function → shown via correlogram → has a distinct end at $\tau = q$.

Note that, for an MA(q) process, there is a cut-off point in the autocorrelation function – all autocorrelations beyond the q-th are zero. By comparison, for the AR(1) process and, as we shall see, for more general autoregressive processes, all autocorrelations are generally non-zero but die away geometrically as a function of the lag. This fact may be borne in mind when we are examining the sample autocorrelation function of an observed time series and considering what model to fit to the data.

$\gamma_\tau = \sigma^2 \sum_{i=\tau}^{q} \theta_i \theta_{i-\tau}$

$\gamma_\tau = \begin{cases} \sigma^2 \sum_{i=0}^{q-\tau} \theta_i \theta_{i+\tau} & \text{if } 0 \le q \le \tau \\ 0 & \text{if } q - \tau < 0 \end{cases}$

↳ $\tau > q$

Want to turn this to sum from $\boxed{i = 0 \to i = q - \tau}$

↳ means the indices need to go from $i \to i + \tau$

↳ $\gamma_\tau = \sigma^2 \sum_{i=0}^{q-\tau} \theta_{i+\tau} \theta_{i+\tau-\tau} = \boxed{\sigma^2 \sum_{i=0}^{q-\tau} \theta_i \theta_{i+\tau} = Cov(Y_t, Y_{t-\tau}) = \gamma_\tau}$

$\rightarrow$ MA(q) $\Rightarrow Y_t = \sum_{i=0}^{q} \theta_i \varepsilon_{i-t}$ when q=1 $\rightarrow$ MA(1)

$Y_t = \theta_0 \varepsilon_t + \theta_1 \varepsilon_{t-1}$, where $\theta_0 = 1$

$\boxed{Y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1}}$

4.5 The first order moving average process

In the special case of the MA(1) process $\{Y_t\}$, which satisfies the equation

$\theta_0 = 1$, $\theta_1 \to \theta$

$$Y_t = \epsilon_t + \theta\epsilon_{t-1} \qquad t \in \mathbb{Z}, \tag{13}$$

the autocorrelation function is given by

linear process: $\boxed{Y_t = \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i}}$

MA(q): $\boxed{Y_t = \sum_{i=0}^{q} \theta_i \varepsilon_{t-i}}$

$$\boxed{\begin{aligned}\gamma_\tau &= Cov(Y_t, Y_{t-\tau}) = E[Y_t, Y_{t-\tau}] \\ &= E\left[\sum_{i=0}^{\infty}\psi_i \varepsilon_{t-i} \sum_{j=\tau}^{\infty}\psi_{j-\tau}\varepsilon_{t-j}\right] \\ &= \sum_{i=0}^{\infty}\sum_{j=\tau}\psi_i\psi_{j-\tau} E[\varepsilon_{t-i}\varepsilon_{t-j}] \\ &= \sum_{j=\tau}^{\infty}\psi_j\psi_{j-\tau}\sigma^2 = \sigma^2\sum_{j=\tau}^{\infty}\psi_j\psi_{j-\tau}\end{aligned}}$$

$$\begin{aligned} \rho_0 &= 1 \\ \rho_1 &= \frac{\theta}{1+\theta^2} \\ \rho_\tau &= 0, \qquad \tau \geq 2. \end{aligned}$$

$\gamma_\tau = \sigma^2 \sum_{j=\tau}^{q} \theta_j \theta_{j-\tau} = \sigma^2 \sum_{j=0}^{q-\tau} \theta_{j+\tau}\theta_j$

$\boxed{\therefore \gamma_0 = \sigma^2 \sum_{j=0}^{q} \theta_j^2}$

Note that if $\theta > 0$ then the MA(1) process is smoother than a white noise process but that if $\theta < 0$ then the MA(1) process is more jagged than a white noise process.

for MA(1) process: q=1

$\gamma_0 = \sigma^2(\theta_0^2 + \theta_1^2) = \sigma^2(1+\theta_1^2)$

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

$\gamma_1 = \sigma^2 \sum_{j=0} \theta_{0+1}\theta_0 = \sigma^2\theta_1$

Using Equation (13) recursively, $\rightarrow$ MA(1) $\hookrightarrow Y_t = \varepsilon_t + \theta_1\varepsilon_{t-1}$

Knowing that MA(1), q=1, there will be a discontinuity at $\tau = q$. So, will get ρ_τ values for $0 \leq \tau \leq q$, and when $\tau > q$, $\rho_\tau = 0$.

$$\begin{aligned} Y_t &= \epsilon_t + \theta(Y_{t-1} - \theta\epsilon_{t-2}) \\ &= \epsilon_t + \theta Y_{t-1} - \theta^2(Y_{t-2} - \theta\epsilon_{t-3}) \\ &= \dots \\ &= \epsilon_t - \sum_{k=1}^{n}(-\theta)^k Y_{t-k} - (-\theta)^{n+1}\epsilon_{t-n-1}. \end{aligned} \tag{14}$$

$\hookrightarrow$ recursive usage of MA(1)

Here we would like to take the limit as $n \to \infty$ in some appropriate sense for Equation (14) to obtain the *infinite order autoregressive representation* of $\{Y_t\}$,

MA(1) $\rightarrow Y_t = \varepsilon_t + \theta\varepsilon_{t-1}$

AR(1) $\rightarrow Y_t = \sum_{i=0}^{\infty} \phi^i \varepsilon_{t-i}$

$$Y_t = -\sum_{k=1}^{\infty}(-\theta)^k Y_{t-k} + \epsilon_t. \tag{15}$$

$\rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{\sigma^2\theta_1}{\sigma^2(1+\theta_1^2)}$

$\boxed{\rho_1 = \frac{\theta_1}{1+\theta_1^2}}$

Let us consider again the concept of convergence in mean square.

MA(q) $= \sum_{i=0}^{q}\theta_i\varepsilon_{t-i}$

$\downarrow$

finite moving average

$$\begin{aligned} E\left[\left(-\sum_{k=1}^{n}(-\theta)^k Y_{t-k} + \epsilon_t - Y_t\right)^2\right] &= E\left[\theta^{2(n+1)}\epsilon_{t-n-1}^2\right] \\ &= \theta^{2(n+1)}\sigma^2, \end{aligned}$$

which tends to zero as $n \to \infty$ if and only if $|\theta| < 1$.

As a matter of fact, it can be shown that equation (15) is an exact representation (with probability one), in which case (as we will see later) the MA(1) process is an *invertible* process: this is the case if and only if $|\theta| < 1$.

Using the lag operator, we may rewrite Equation (13) as

MA(1) $\rightarrow Y_t = \varepsilon_t + \theta\varepsilon_{t-1}$

$Y_{t-1} = \varepsilon_{t-1} + \theta\varepsilon_{t-2}$

$Y_{t-2} = \varepsilon_{t-2} + \theta\varepsilon_{t-3}$

$$Y_t = (1 + \theta L)\epsilon_t.$$

$\varepsilon_{t-1} = Y_{t-1} - \theta\varepsilon_{t-2}$

$\varepsilon_{t-2} = Y_{t-2} - \theta\varepsilon_{t-3}$

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

MA(1)

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

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

dividing through by $(1+\theta L)$

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

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

↳ $\varepsilon_t = \theta_0 Y_t + \sum_{k=1}^{\infty}(-\theta)^k Y_{t-k}$

$= \sum_{k=0}^{\infty}(-\theta)^k Y_{t-k}$

$= \sum_{k=0}^{\infty}(-\theta)^k L^k Y_t$

Formally inverting this relationship, we may write

Recall: $(1+\theta L)^{-1}$ can be expressed as an infinite sum

$$\epsilon_t = (1+\theta L)^{-1}Y_t = \boxed{\sum_{k=0}^{\infty}(-\theta)^k L^k Y_t} = Y_t + \sum_{k=1}^{\infty}(-\theta)^k Y_{t-k},$$

which is equivalent to Equation (15). As shown above, this procedure is justifiable if and only if $|\theta| < 1$.

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

Note that the MA(1) processes with parameters θ and θ^{-1}, respectively, have the same autocorrelation function, since

$$\boxed{\frac{\theta^{-1}}{1+\theta^{-2}} = \frac{\theta}{1+\theta^2}.}$$

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

Hence, given the autocorrelation function of an MA(1) process, it is impossible to identify uniquely the parameter value. But if we impose the condition that the process must be invertible then there is a unique parameter value corresponding to the given autocorrelation function.

The MA(1) process with $|\theta| = 1$ is exceptional in that it is not invertible but may be uniquely identified from the autocorrelation function.

↳ MA(1) processes should be invertible, with edge-case of $|\theta| = 1$.

4.6 Invertibility

Using the lag operator, we may write the general MA(q) model of Equation (11) in the form

General MA(q) linear: $X_t = \sum_{i=0}^{\infty}\psi_i \varepsilon_{t-i}$

$$Y_t = \theta(L)\epsilon_t, \qquad (16)$$

$\theta(L)$ here is a linear fn.
↳ called the MA characteristic polynomial.

where $\theta(z)$ is the *MA characteristic polynomial*, defined by

MA(q): $Y_t = \sum_{i=0}^{q}\theta_i \varepsilon_{t-i}$

→ for MA(q)
→ q+1 terms ↳ as sum from zero to q.

$$\theta(z) = 1 + \theta_1 z + \theta_2 z^2 + \ldots + \theta_q z^q. \qquad (17)$$

$\theta(L) = 1 + \theta_1 L + \theta_2 L^2 + \theta_3 L^3 + \ldots + \theta_q L^q$

Thus $\theta(z)$ is the generating function of the moving average coefficients. The corresponding *MA characteristic equation* is

$$\boxed{1 + \theta_1 z + \theta_2 z^2 + \ldots + \theta_q z^q = 0.} \qquad (18)$$

⇒ set characteristic polynomial to zero.

Definition 4.6.1 (Invertibility)
A stationary process $\{Y_t\}$ is said to be invertible if it can be expressed in the form

will often require $\{Y_t\}$ to be both stationary + invertible

$$\boxed{Y_t = \sum_{k=1}^{\infty}\pi_k Y_{t-k} + \epsilon_t,} \qquad (19)$$

MA(q) model can either be expressed as:
$Y_t = \sum_{i=0}^{q}\theta_i \varepsilon_{t-i}$ $\equiv$ $Y_t = \varepsilon_t + \sum_{i=1}^{q}\theta_i \varepsilon_{t-i}$

such that $\sum_{j=1}^{\infty}|\pi_j| < \infty$.

For an MA(q) process, the following result can be established.

Theorem 4.6.2 (Invertibility condition for $MA(q)$)
A necessary and sufficient condition for the moving average parameters $\theta_1, \theta_2, \ldots, \theta_q$ of the MA(q) model to specify an invertible moving average process is that all the roots of the characteristic equation, Equation (18), lie strictly outside the unit circle in the complex plane (i.e., all the roots are greater than one in modulus/absolute value). □

Suspect this is a key skill to learn:- test for invertibility of a time series process.

Lingo: ϕ: autoregressive parameters
θ: moving average parameters

Recall characteristic polynomial: $\theta(z) = 1 + \theta_1 z + \theta_2 z^2 + \dots + \theta_q z^q$
and characteristic equation : $\theta(z) = 1 + \theta_1 z + \theta_2 z^2 + \dots + \theta_q z^q = 0$

We see from the expression (19) that the value at time t of an invertible process can be expressed as a linear combination of the values of the process at previous time points, with the addition of the white noise term at time t. This makes it plausible that it should be possible to forecast future values of such a process in a straightforward manner, given knowledge of its history.

Generally, for purposes of estimation and forecasting, the invertibility condition is imposed on stationary processes. Given the autocorrelation function of a MA process, there is generally a multiplicity of parameter values that will yield the given autocorrelations, but there is only one corresponding set of parameters which specifies an *invertible* MA process.

Assuming that a given MA(q) process is invertible, we may rewrite (19) as

$Y_t = \sum_{k=1}^{\infty} \pi_k Y_{t-k} + \varepsilon_t$

$$\boxed{\epsilon_t = \pi(L) Y_t,} \tag{20}$$

where

$\varepsilon_t = Y_t - \sum_{i=1}^{\infty} \pi_i Y_{t-i}$

$$\boxed{\pi(z) = 1 - \sum_{k=1}^{\infty} \pi_k z^k.}$$

We may invert (16) to obtain

$Y_t = \sum_{i=0}^{q} \theta_i \varepsilon_{t-i}$ ↳ $Y_t = \theta(L)\varepsilon_t$

$$\boxed{\epsilon_t = \theta(L)^{-1} Y_t.} \tag{21}$$

MA(q) in terms of characteristic polynomial

It follows from equations (20) and (21) that we may write

$$\boxed{\pi(z) = \theta(z)^{-1}}$$

or, equivalently,

$$\boxed{\theta(z)\pi(z) = 1.}$$

In the special case of the MA(1) process, $\theta(z) = 1 + \theta z$, $\pi(z) = (1 + \theta z)^{-1}$, and the invertibility condition reduces to $|\theta| < 1$.

MA(1) → $q = 1$.

Recall: $\theta(z) = 1 + \theta_1 z + \theta_2 z^2 + \dots + \theta_q z^q$

$\boxed{\theta(L) = 1 + \theta L}$ ↳ $\theta(L)^{-1} = \pi(L) = (1 + \theta L)^{-1}$

and $\pi(L) = 1 - \sum_{k=1}^{\infty} \pi_k L^k$

Since $\boxed{\theta(z)\pi(z) = 1.}$

What does $\pi(z)$ mean?
↳ Falls directly out of the invertibility criterion

$Y_t = \sum_{k=1}^{\infty} \pi_k Y_{t-k} + \varepsilon_t$

in terms of ε_t ↳ $\varepsilon_t = Y_t - \sum_{k=1}^{\infty} \pi_k Y_{t-k}$

in terms of lag ↳ $\varepsilon_t = \pi(L) X_t$ lag operator, applied i times

where $\pi(z) = 1 - \sum_{i=1}^{\infty} \pi_i z^i$
↳ will be the lag