Table of content I

Table of content II

Linear Difference Equations

Time series models can be represented or approximated by a linear difference equation. Consider the situation where a realization at time t, y_t, is a linear function of the last p realizations of y and a random disturbance term, denoted by ϵ_t.

$$y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \cdots + \alpha_p y_{t-p} + \epsilon_t. \tag{10}$$

$\Rightarrow$ $AR(p)$-Process

The Lag Operator

The *lag operator* (also called backward shift operator), denoted by L, is an operator that shifts the time index backward by one unit. Applying it to a variable at time t, we obtain the value of the variable at time $t-1$, i.e.,

$$Ly_t = y_{t-1}.$$

Applying the lag operator twice amount to lagging the variable twice, i.e., $L^2 y_t = L(Ly_t) = Ly_{t-1} = y_{t-2}$.

The Lag Operator

More formally, the lag operator transforms one time series, say $\{x_t\}_{t=-\infty}^{\infty}$, into another series, say $\{y_t\}_{t=-\infty}^{\infty}$, where $x_t = y_{t-1}$. Raising L to a negative power, we obtain a *delay* (or *lead*) *operator*, i.e.,

$$L^{-k} y_t = y_{t+k}.$$

The Lag Operator

The following statements hold for the lag operator L

$$
\begin{aligned}
Lc &= c \text{ for a constant c} && (11)\\
(L^j + L^i)y_t &= L^j y_t + L^i y_t \text{ (distributive law)} && (12)\\
L^i(L^j y_t) &= L^i y_{t-j} \text{ (associative law)} && (13)\\
aLy_t &= L(ay_t) = ay_{t-1} && (14)
\end{aligned}
$$

The Difference Operator

The *difference operator* Δ is used to express the difference between values of time series at different times. With Δy_t we denote the first difference of y_t, i.e.,

$$\Delta y_t = y_t - y_{t-1}.$$

It follows that

$$\begin{aligned} \Delta^2 y_t &= \Delta(\Delta y_t) = \Delta(y_t - y_{t-1}) \\ &= (y_t - y_{t-1}) - (y_{t-1} - y_{t-2}) = y_t - 2y_{t-1} + y_{t-2} \end{aligned}$$

etc. The difference operator can expressed in terms of the lag operator by $\Delta = 1 - L$. Hence, $\Delta^2 = (1-L)^2 = 1 - 2L + L^2$ and, in general, $\Delta^n = (1-L)^n$.

Transforming the Expression of Time Series Models

The lag operator enables us to express higher–order difference equations more compactly in form of polynomials in lag operator L. For example, the difference equation

$$y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \alpha_3 y_{t-3} + c$$

can be written as

$$y_t = \alpha_1 L y_t + \alpha_2 L^2 y_t + \alpha_3 L^3 y_t + c,$$

$$(1 - \alpha_1 L - \alpha_2 L^2 - \alpha_3 L^3) y_t = c$$

or, in short,

$$\boldsymbol{\alpha}(L) y_t = c.$$

The Characteristic Equation

Replacing in polynomial $a(L)$ lag operator L by variable λ, we obtain the *characteristic equation* associated with difference equation (10):

$$\boldsymbol{\alpha}(\lambda) = 0. \tag{15}$$

A value of λ which satisfies characteristic equation (15) is called a *root* of polynomial $\alpha(\lambda)$.
$\Rightarrow$ Will be important in later applications.

Solving Difference Equations

Expression (15) represents the so-called *coefficient form* of a characteristic equation, i.e.,

$$1 - \alpha_1 \lambda - \cdots - \alpha_p \lambda^p = 0.$$

An alternative is the *root form* given by

$$(\lambda_1 - \lambda)(\lambda_2 - \lambda) \cdots (\lambda_p - \lambda) = \prod_{i=1}^{p} (\lambda_i - \lambda) = 0.$$

Solving Difference Equations: An Example

Consider the difference equation

$$y_t = \frac{3}{2} y_{t-1} - \frac{1}{2} y_{t-2} + \epsilon_t.$$

The characteristic equation in coefficient form is given by

$$1 - \frac{3}{2}\lambda + \frac{1}{2}\lambda^2 = 0$$

or

$$2 - 3\lambda + 1\lambda^2 = 0,$$

which can be written in root form as

$$(1 - \lambda)(2 - \lambda) = 0.$$

Here, $\lambda_1 = 1$ and $\lambda_2 = 2$ represent the set of possible solutions for λ satisfying the characteristic equation $1 - \frac{3}{2}\lambda + \frac{1}{2}\lambda^2 = 0$.

Solving Difference Equations: An Example

Calculate the characteristic roots of the following difference equations

$$y_t = y_{t-1} - y_{t-2} + \epsilon_t \quad (16)$$

$$y_t = -y_{t-1} + y_{t-2} + \epsilon_t \quad (17)$$

$$y_t = 0.125 y_{t-3} + \epsilon_t \quad (18)$$

Autoregressive (AR) Processes

An *autoregressive process* of order p, or briefly an AR(p) process, is a process where realization y_t is a weighted sum of past p realizations, i.e., $y_{t-1}, y_{t-2}, \ldots, y_{t-p}$, plus an additive, contemporaneous disturbance term, denoted by ϵ_t.
The process can be represented by the p-th order difference equation

$$y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \ldots + \alpha_p y_{t-p} + \epsilon_t. \tag{19}$$

Autoregressive (AR) Processes

$$y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \ldots + \alpha_p y_{t-p} + \epsilon_t. \tag{20}$$

We assume that ϵ_t, $t = 0, \pm 1, \pm 2 \ldots$, is a zero-mean, independently and identically distributed (iid) sequence with

$$\mathrm{E}(\epsilon_t) = 0, \qquad \mathrm{E}(\epsilon_s \epsilon_t) = \begin{cases} \sigma^2, & \text{if } s = t, \\ 0, & \text{if } s \neq t, \end{cases} \tag{21}$$

for all t and s. Sequence (21) is called a zero–mean *white–noise process*, or simply *white noise.*

Autoregressive (AR) Processes

Using the lag operator L, the AR(p) process (19) can be expressed more compactly as

$$(1 - \alpha_1 L - \alpha_2 L^2 - \ldots - \alpha_p L^p) y_t = \epsilon_t$$

or

$$\boldsymbol{\alpha}(L) y_t = \epsilon_t, \tag{22}$$

where the autoregressive polynomial $\alpha(L)$ is defined by
$\boldsymbol{\alpha}(L) = 1 - \alpha_1 L - \alpha_2 L^2 - \ldots - \alpha_p L^p$.

The mean of a stationary AR(1) process

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \epsilon_t$$

Taking Expectations (E) we get

$$\begin{aligned} E(y_t) &= \alpha_0 + \alpha_1 E(y_{t-1}) + E(\epsilon_t) \\ E(y_t) &= \alpha_0 + \alpha_1 E(y_t) \\ E(y_t) &= \mu = \frac{\alpha_0}{1-\alpha_1} \end{aligned}$$

The mean of a stationary AR(p) process

We the same technique one can obtain the mean of an AR(2) process

$$E(y_t) = \mu = \frac{\alpha_0}{1 - \alpha_1 - \alpha_2}$$

and an AR(p) process

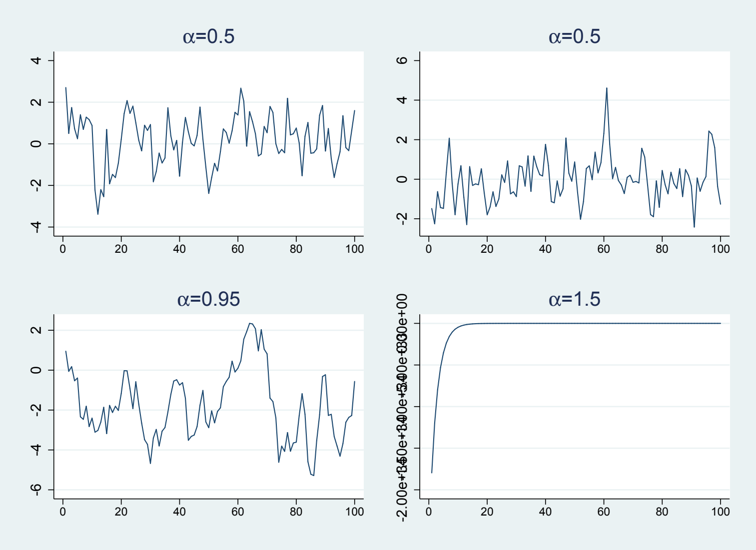
$$E(y_t) = \mu = \frac{\alpha_0}{1 - \alpha_1 - \alpha_2 - \ldots - \alpha_p}$$

Examples

Calculate the mean of the following AR processes

$$y_t = 0.5y_{t-1} + \epsilon_t \tag{23}$$
$$y_t = 0.5 + 0.5y_{t-1} + \epsilon_t \tag{24}$$
$$y_t = 0.5 - 0.5y_{t-1} + \epsilon_t \tag{25}$$
$$y_t = 0.5 + 0.5y_{t-1} + 0.5y_{t-2} + \epsilon_t \tag{26}$$

AR Examples


α=0.5
4
2
0
-2
-4
0 20 40 60 80 100
α=0.5
6
4
2
0
-2
0 20 40 60 80 100
α=0.95
2
0
-2
-4
-6
0 20 40 60 80 100
α=1.5
0 20 40 60 80 100

Moving Average (MA) Processes

A *moving average process* of order q, denoted by MA(q), is the weighted sum of the preceding q lagged disturbances plus a contemporaneous disturbance term, i.e.,

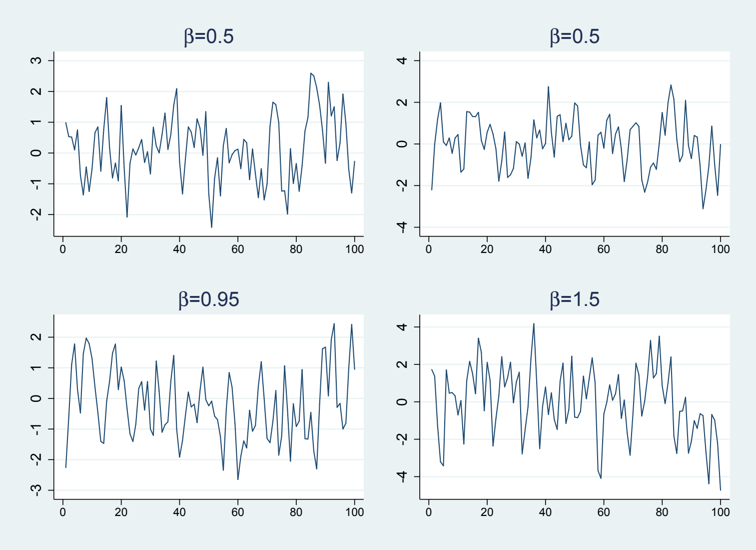
$$y_t = \beta_0 + \beta_1 \epsilon_{t-1} + \ldots + \beta_q \epsilon_{t-q} + \epsilon_t \tag{27}$$

or

$$y_t = \boldsymbol{\beta}(L)\epsilon_t. \tag{28}$$

Here $\boldsymbol{\beta}(L) = \beta_0 + \beta_1 L + \beta_2 L^2 + \ldots + \beta_q L^q$ denotes a moving average polynomial of degree q, and ϵ_t is again a zero-mean white noise process.

MA Examples



The mean of a stationary MA(q) process

$$y_t = \beta_0 + \beta_1 \epsilon_{t-1} + \ldots + \beta_q \epsilon_{t-q} + \epsilon_t$$

Taking expectations we get

$$\mathrm{E}(y_t) = \mu = \beta_0$$

because

$$\mathrm{E}(\epsilon_t) = \mathrm{E}(\epsilon_{t-1}) = \ldots = \mathrm{E}(\epsilon_{t-q}) = 0$$

Relationship between AR and MA

Consider the AR(1) process

$$y_t = \alpha_1 y_{t-1} + \epsilon_t$$

Repeated substitution yields

$$\begin{aligned} y_t &= \alpha_1(\alpha_1 y_{t-2} + \epsilon_{t-1}) + \epsilon_t \\ &= \alpha_1^2 y_{t-2} + \alpha_1 \epsilon_{t-1} + \epsilon_t \\ &= \alpha_1^2(\alpha_1 y_{t-3} + \epsilon_{t-1}) + \alpha_1 \epsilon_{t-1} + \epsilon_t \\ &= \ldots \\ &= \sum_{j=1}^{\infty} \alpha_1^j \epsilon_{t-j} + \epsilon_t \end{aligned}$$

i.e., each stationary AR(1) process can be represented as an MA(∞) process.

The mean of a stationary AR(q) process

Whiteboard

Alternative derivation of the mean of an stationary AR(1) process

$$y_t = c + \alpha y_{t-1} + \epsilon_t \tag{29}$$

with $|a| < 1$.

Relationship between AR and MA

For a general stationary AR(p) process

$$\begin{aligned} y_t &= \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \ldots + \alpha_p y_{t-p} + \epsilon_t \\ \boldsymbol{\alpha}(L) y_t &= \epsilon_t \end{aligned}$$

we have

$$y_t = \boldsymbol{\alpha}(L)^{-1} \epsilon_t = \boldsymbol{\phi}(L) \epsilon_t = \sum_{j=1}^{\infty} \phi_j \epsilon_{t-j} \tag{30}$$

where $\phi(L)$ is an operator satisfying $\alpha(L)\phi(L) = 1$.

Autoregressive Moving Average (ARMA) Processes

The AR and MA processes just discussed can be regarded as special cases of a mixed *autoregressive moving average process*, in short, an ARMA(p, q) process. It is written as

$$y_t = \alpha_1 y_{t-1} + \ldots + \alpha_p y_{t-p} + \epsilon_t + \beta_1 \epsilon_{t-1} + \ldots + \beta_q \epsilon_{t-q} \tag{31}$$

or

$$\boldsymbol{\alpha}(L) y_t = \boldsymbol{\beta}(L) \epsilon_t. \tag{32}$$

Clearly, ARMA($p, 0$) and ARMA($0, q$) processes correspond to pure AR(p) and MA(q) processes, respectively.

The mean of a stationary ARMA(p, q) process

For

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \ldots + \alpha_p y_{t-p} + \beta_1 \epsilon_{t-1} + \ldots + \beta_q \epsilon_{t-q} + \epsilon_t \qquad (33)$$

we get

$$\mathrm{E}(y_t) = \mu = \frac{\alpha_0}{1 - \alpha_1 - \alpha_2 - \ldots - \alpha_p}$$

applying the previous arguments.

Examples

Calculate the mean of the following ARMA processes

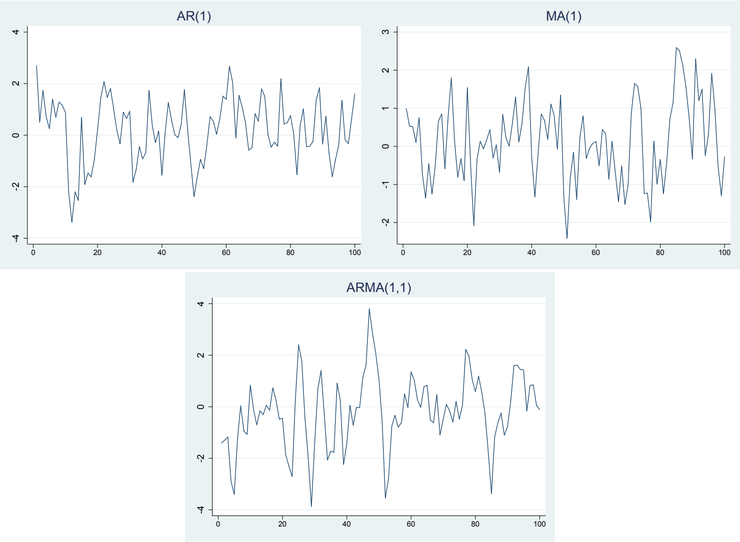
$$y_t = 0.5\epsilon_{t-1} + \epsilon_t \tag{34}$$

$$y_t = 1500\epsilon_{t-1} + 0.5 + 0.75y_{t-1} + \epsilon_t - 0.8\epsilon_{t-2} \tag{35}$$

$$y_t = 0.5 - 0.5y_{t-1} + 2\epsilon_{t-1} + 0.8\epsilon_{t-2} + \epsilon_t \tag{36}$$

$$y_t = y_{t-1} + 0.5\epsilon_{t-1} + \epsilon_t \tag{37}$$

ARMA Examples



ARMA Processes With Exogenous Variables (ARMAX Processes)

ARMA processes that also include current and/or lagged, exogenously determined variables are called *ARMAX processes*. Denoting the exogenous variable by y_t, an ARMAX process has the form

$$\alpha(L)y_t = \beta(L)\epsilon_t + \gamma(L)x_t. \tag{38}$$

Example: ARX-models for Forecasting

$$\alpha(L)y_t = \gamma(L)x_t + \epsilon_t \tag{39}$$

$$y_t = \alpha + \sum_{i=1}^{p} \beta_i y_{t-i} + \sum_{j=1}^{q} \gamma_j x_{t-j} + \epsilon_t \tag{40}$$

For example: Forecasting German Industrial Production with its own lagged values plus an exogenous indicator (e.g. the Ifo Business Climate)

$\Rightarrow$ Section about prediction

Integrated ARMA (ARIMA) Processes

Very often we observe that the mean and/or variance of economic time series increase over time. In this case, we say the series are nonstationary. However, a series of the *changes* from one period to the next, i.e., the first differences, may have a mean and/or variance that do not change over time.
$\Rightarrow$ Model the differenced series

Integrated ARMA (ARIMA) Processes

An ARMA model for the d-th difference of a series rather than the original series is called an *autoregressive integrated moving average process*, or an ARIMA (p, d, q), process and written as

$$\alpha(L)\Delta^d y_t = \beta(L)\epsilon_t. \tag{41}$$

Further Aspects

Seasonal ARMA Processes

$$\alpha_s(L^s)(1 - L^s)^D y_t = \beta_s(L^s)\epsilon_t, \tag{42}$$

ARMA Processes with deterministic Components:
Adding a constant

$$\boldsymbol{\alpha}(L)y_t = c + \boldsymbol{\beta}(L)\epsilon_t. \tag{43}$$

Or a linear Trend

$$\boldsymbol{\alpha}(L)y_t = c_0 + c_1 t + \boldsymbol{\beta}(L)\epsilon_t.$$

The Concept of Stationarity

Stationarity is a property that guarantees that the essential properties of a time series remain constant over time. An important concept of stationarity is that of *weak stationarity*.

Time series $\{y_t\}_{t=-\infty}^{\infty}$ is said to be weakly stationary if:

(1) the mean of y_t is constant over time, i.e., $\mathrm{E}(y_t) = \mu$, $|\mu| < \infty$;

(2) the variance of y_t is constant over time, i.e., $\mathrm{Var}(y_t) = \gamma_0 < \infty$;

(3) the covariance of y_t and y_{t-k} does not vary over time, but may depend on the lag k, i.e., $\mathrm{Cov}(y_t, y_{t-k}) = \gamma_k$, $|\gamma_k| < \infty$.

$\Rightarrow$ A process is called *strongly (strictly) stationary* if the joint distribution of $(y_1, ...y_k)$ is identical to that of $(y_{1+t}, ...y_{k+t})$.

Stationarity of AR(p) processes

An AR(p) is stationary if the absolute values of all the roots of the characteristic equation

$$\alpha_0 - \alpha_1 \lambda - \cdots - \alpha_p \lambda^p = 0.$$

are greater than 1 (with $\alpha_0 = 1$).

- This is in practice difficult to realize.
- What about forth order characteristic equations?
- Alternative: Employ the *Schur Criterion*

Stationarity of AR(p) processes: The Schur Criterion

If the determinants

$$A_1 = \begin{vmatrix} \alpha_0 & \alpha_p \\ \alpha_p & \alpha_0 \end{vmatrix}, A_2 = \begin{vmatrix} \alpha_0 & 0 & \alpha_p & \alpha_{p-1} \\ \alpha_1 & \alpha_0 & 0 & \alpha_p \\ \alpha_p & 0 & \alpha_0 & \alpha_1 \\ \alpha_{p-1} & \alpha_p & 0 & \alpha_0 \end{vmatrix} \dots$$

Stationarity of AR(p) processes: The Schur Criterion

and

$$A_p = \begin{vmatrix} \alpha_0 & 0 & \dots & 0 & \alpha_p & \alpha_{p-1} & \dots & \alpha_1 \\ \alpha_1 & \alpha_0 & \dots & 0 & 0 & \alpha_p & \dots & \alpha_2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \alpha_{p-1} & \alpha_{p-2} & \dots & \alpha_0 & 0 & 0 & \dots & \alpha_p \\ \alpha_p & 0 & \dots & 0 & \alpha_0 & \alpha_1 & \dots & \alpha_{p-1} \\ \alpha_{p-1} & \alpha_{p-1} & \dots & 0 & 0 & \alpha_0 & \dots & \alpha_{p-2} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \\ \alpha_1 & \alpha_2 & \dots & \alpha_p & 0 & 0 & \dots & \alpha_0 \end{vmatrix}.$$

are all positive, then an AR(p) process is stationary.

Stationarity of an AR(1) Process

Consider the AR(1) process

$$y_t = \alpha_1 y_{t-1} + \epsilon_t$$

The characteristic equation is

$$1 - \alpha_1 \lambda = 0$$

We have

$$\begin{aligned} A_1 &= \begin{vmatrix} \alpha_0 & \alpha_p \\ \alpha_p & \alpha_0 \end{vmatrix} = \begin{vmatrix} 1 & -\alpha_1 \\ -\alpha_1 & 1 \end{vmatrix} \\ &= 1 - \alpha_1^2 > 0 \Longleftrightarrow \left|\alpha_1\right| < 1 \end{aligned}$$

Stationarity of AR(p) processes: **An Alternative Schur Criterion**

For the AR polynomial $a(L) = 1 - \alpha_1 L - \ldots - \alpha_p L^p$, the Schur criterion requires the construction two lower-triangular Toeplitz matrices, A_1 and A_2, whose first columns consist of the vectors $(1, -\alpha_1, -\alpha_2, \ldots, -\alpha_{p-1})'$ and $(-\alpha_p, -\alpha_{p-1}, \ldots, -\alpha_1)'$, respectively, i.e.,

$$A_1 = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ -\alpha_1 & 1 & & & 0 \\ -\alpha_2 & -\alpha_1 & \ddots & & \vdots \\ \vdots & & & & 0 \\ -\alpha_{p-1} & -\alpha_{p-2} & \cdots & -\alpha_1 & 1, \end{bmatrix}$$

Stationarity of AR(p) processes: **An Alternative Schur Criterion**

$$A_2 = \begin{bmatrix} -\alpha_p & 0 & \cdots & 0 & 0 \\ -\alpha_{p-1} & -\alpha_p & & & 0 \\ -\alpha_{p-2} & -\alpha_{p-1} & & & \vdots \\ \vdots & & \ddots & & 0 \\ -\alpha_1 & -\alpha_2 & \cdots & -\alpha_{p-1} & -\alpha_p \end{bmatrix}.$$

Then, the AR (p) process is covariance stationary if and only if the so-called *Schur matrix*, defined by

$$S_a = A_1 A_1' - A_2 A_2', \tag{44}$$

is positive definite.

Stationarity of AR(1) processes: **An Alternative Schur Criterion**

For

$$y_t = \alpha_1 y_{t-1} + \epsilon_t$$

we get $A_1 = [1]$ and $A_2 = [-\alpha_1]$

$$|S_a| = 1 \cdot 1' - (-\alpha_1) \cdot (-\alpha_1)' = 1 - \alpha_1^2 > 0 \Longleftrightarrow |\alpha_1| < 1$$

Stationarity of AR(2) processes: **An Alternative Schur Criterion**

For

$$y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \epsilon_t$$

we get

$$A_1 = \begin{bmatrix} 1 & 0 \\ -\alpha_1 & 1 \end{bmatrix}, A_2 = \begin{bmatrix} -\alpha_2 & 0 \\ -\alpha_1 & -\alpha_2 \end{bmatrix}$$

$$S_a = \begin{bmatrix} 1 - \alpha_2^2 & -\alpha_1 - \alpha_2\alpha_1 \\ -\alpha_1 - \alpha_2\alpha_1 & 1 - \alpha_2^2 \end{bmatrix}$$

Stationarity of an AR(2) Process

For an AR(2) process covariance stationarity requires that the AR coefficients satisfy

$$\begin{aligned} |\alpha_2| &< 1, \\ \alpha_2 + \alpha_1 &< 1, \\ \alpha_2 - \alpha_1 &< 1. \end{aligned} \tag{45}$$

Stationarity of MA(q) Processes

Pure MA processes are always stationary, because it has no autoregressive roots.

Stationarity of ARMA(p, q) Processes

The stationarity property of the mixed ARMA process

$$\alpha(L)y_t = \beta(L)\epsilon_t \tag{46}$$

does not dependent on the values of the MA parameters. Stationarity is a property that depends solely on the AR parameters.

Stationarity: Examples

	α_1	α_2	Stationary?
AR(1)	0.5		
AR(1)	-0.99		
AR(1)	1		
AR(1)	1.5		
AR(2)	0.5	0.4	
AR(2)	0.2	-0.9	
AR(2)	1.5	-0.5	

$\Rightarrow$ Same conclusions for ARMA models with q MA lags with arbitrary parameters (β_i).

Examples

Are the following process stationary? Employ the Schur-Criterion:

$$y_t = 0.5y_{t-1} + \epsilon_t \quad (47)$$

$$y_t = 0.5 + 0.5y_{t-1} + \epsilon_t \quad (48)$$

$$y_t = 0.5 - 0.5y_{t-1} + \epsilon_t \quad (49)$$

$$y_t = 0.5 + 0.5y_{t-1} + 0.5y_{t-2} + \epsilon_t \quad (50)$$

$$y_t = 0.5 + 0.5y_{t-1} + 0.5y_{t-2} - 0.8y_{t-3} + 0.5\epsilon_{t-1} + \epsilon_t \quad (51)$$

ARMA Models

Remember the definition of ARMA models

$$y_t = \alpha_1 y_{t-1} + \ldots + \alpha_p y_{t-p} + \epsilon_t + \beta_1 \epsilon_{t-1} + \ldots + \beta_q \epsilon_{t-q} \tag{52}$$

or

$$\boldsymbol{\alpha}(L) y_t = \boldsymbol{\beta}(L) \epsilon_t. \tag{53}$$

Causality

A linear process is **causal** if

$$\alpha(L) = \alpha_0 + \alpha_1 L + \alpha_1 L^2 ... \tag{54}$$

satisfies

$$\sum_{j=0}^{\infty} = |\alpha_j| < \infty \tag{55}$$

or equivalently: all roots of the *characteristic* equation are outside the unit circle

Invertibility

A linear process is **invertible** if

$$\beta(L) = \beta_0 + \beta_1 L + \beta_1 L^2 ... \tag{56}$$

satisfies

$$\sum_{j=0}^{\infty} = |\beta_j| < \infty \tag{57}$$

or equivalently: all roots of the *characteristic* equation are outside the unit circle

Examples

$$y_t = 1.5y_{t-1} + 0.2\epsilon_{t-1} + \epsilon_t$$

or equivalently

$$(1 - 1.5L)y_t = (1 + 0.2L)\epsilon_t$$

Characteristic equations

$(1 - 1.5\lambda) = 0 \Rightarrow \lambda = 2/3 \Rightarrow$ *in*side the unit circle, **not causal**

and

$(1 + 0.2\lambda) = 0 \Rightarrow \lambda = -5 \Rightarrow$ *out*side the unit circle, **invertible**

Examples

$$y_t = 0.25 y_{t-2} + 2\epsilon_{t-1} + \epsilon_t$$

or equivalently

$$(1 + 0.25L^2) y_t = (1 + 2L)\epsilon_t$$

Characteristic equations

$(1 + 0.25\lambda^2) = 0 \Rightarrow \lambda_{1,2} = \pm 2 \Rightarrow$ *out*side the unit circle, **causal**

and

$(1 + 2\lambda) = 0 \Rightarrow \lambda = -0.5 \Rightarrow$ *in*side the unit circle, **not invertible**

Autocovariance and Autocorrelation Functions

How to determine the order of an ARMA(p, q) process?

- Useful tools are the
  - *sample autocovariance function* (SACovF)
  - and its scaled counterpart *sample autocorrelation function* (SACF)

Deriving the ACovF and ACF for an AR(1) Process

Derive the Autocovariance Function for an AR(1) process.

$$y_t = \alpha y_{t-1} + \epsilon_t, \tag{58}$$

where ϵ_t is the usual white–noise process with $E(\epsilon_t^2) = \sigma^2$.

Deriving the ACovF and ACF for an AR(1) Process

Consider the stationary AR(1) process

$$y_t = \alpha y_{t-1} + \epsilon_t, \tag{59}$$

where ϵ_t is the usual white–noise process with $\mathrm{E}(\epsilon_t^2) = \sigma^2$.
To obtain the variance $\gamma_0 = \mathrm{E}(y_t^2)$, multiply both sides of (58) by y_t,

$$y_t^2 = \alpha y_t y_{t-1} + y_t \epsilon_t,$$

and take expectations, i.e.,

$$\mathrm{E}(y_t^2) = \alpha \mathrm{E}(y_t y_{t-1}) + \mathrm{E}(y_t \epsilon_t)$$

or

$$\gamma_0 = \alpha \gamma_1 + \mathrm{E}(y_t \epsilon_t).$$

Deriving the ACovF and ACF for an AR(1) Process

Thus, to specify γ_0, we have to determine γ_1 and $\mathrm{E}(y_t\epsilon_t)$. To obtain the latter quantity, substitute the RHS of (58) for y_t,

$$\begin{aligned}\mathrm{E}(y_t\epsilon_t) &= \mathrm{E}[(\alpha y_{t-1}+\epsilon_t)\epsilon_t] \\ &= \alpha\mathrm{E}(y_{t-1}\epsilon_t)+\mathrm{E}(\epsilon_t^2).\end{aligned}$$

Since y_{t-1} is independent of the future disturbances ϵ_{t+i}, $i=0,1,\ldots$, $\mathrm{E}(y_{t-1}\epsilon_t)=0$ and $\mathrm{E}(\epsilon_t^2)=\sigma^2$,

$$\mathrm{E}(\epsilon_t y_t)=\sigma^2.$$

Therefore,

$$\gamma_0=\alpha\gamma_1+\sigma^2. \tag{60}$$

Deriving the ACovF and ACF for an AR(1) Process

To determine $\gamma_1 = E(y_t y_{t-1})$, we basically repeat the above procedure. Multiplying (58) by y_{t-1} and taking expectations on both sides gives

$$\mathrm{E}(y_t y_{t-1}) = \alpha \mathrm{E}(y_{t-1}^2) + \mathrm{E}(y_{t-1}\epsilon_t).$$

Using $\mathrm{E}(y_{t-1}\epsilon_t) = 0$ and the fact that stationarity implies that $\mathrm{E}(y_{t-1}^2) = \mathrm{E}(y_t^2) = \gamma_0$, we have

$$\gamma_1 = \alpha\gamma_0. \tag{61}$$

Deriving the ACovF and ACF for an AR(1) Process

Substituting (61) into (60) and solving for γ_0 gives the expression for the theoretical variance of an AR(1) process, which we derived in the previous section,

$$\gamma_0 = \frac{\sigma^2}{1-\alpha^2}. \tag{62}$$

It follows from (61) that

$$\gamma_1 = \alpha \frac{\sigma^2}{1-\alpha^2}. \tag{63}$$

Deriving the ACovF and ACF for an AR(1) Process

In fact, since

$$E(y_t y_{t-k}) = \alpha E(y_{t-1} y_{t-k}) + E(\epsilon_t y_{t-k}), \qquad k = 1, 2, \ldots,$$

and $E(\epsilon_t y_{t-k}) = 0$, for $k = 1, 2, \ldots$, first and higher-order autocovariances are derived recursively by

$$\gamma_k = \alpha \gamma_{k-1}, \qquad k = 1, 2, \ldots \, . \tag{64}$$

It is obvious that the recursive relationship (64) holds also for the autocorrelation function, $\rho_k = \gamma_k / \gamma_0$, of the AR(1) process, i.e., $\rho_k = \alpha \rho_{k-1}$, for $k = 1, 2, \ldots$.

Deriving the ACov and ACF for an ARMA(1,1) Process

Consider the stationary, zero-mean ARMA(1,1) process

$$y_t = \alpha y_{t-1} + \epsilon_t + \beta\epsilon_{t-1}, \tag{65}$$

where ϵ_t is again an white–noise process with variance σ^2.

Deriving the ACov and ACF for an ARMA(1,1) Process

As in the previous example, multiplying (65) by y_t and taking expectations yields

$$\gamma_0 = \alpha\gamma_1 + \mathrm{E}[y_t(\epsilon_t + \beta\epsilon_{t-1})]. \tag{66}$$

To determine $\mathrm{E}[y_t(\epsilon_t + \beta\epsilon_{t-1})]$, replace y_t by the right hand side of (65), i.e.,

$$\begin{aligned}
\mathrm{E}[y_t(\epsilon_t + \beta\epsilon_{t-1})] &= \mathrm{E}[(\alpha y_{t-1} + \epsilon_t + \beta\epsilon_{t-1})(\epsilon_t + \beta\epsilon_{t-1})] \\
&= \mathrm{E}(\alpha y_{t-1}\epsilon_t + \epsilon_t^2 + \beta\epsilon_{t-1}\epsilon_t + \alpha\beta y_{t-1}\epsilon_{t-1} \\
&\quad + \beta\epsilon_t\epsilon_{t-1} + \beta^2\epsilon_{t-1}^2) \\
&= \sigma^2 + \alpha\beta\sigma^2 + \beta^2\sigma^2.
\end{aligned}$$

Deriving the ACov and ACF for an ARMA(1,1) Process

Taking the expectation operator inside the parentheses and noting the fact that $\mathrm{E}(y_{t-1}\epsilon_t) = \mathrm{E}(\epsilon_{t-1}\epsilon_t) = 0$ and $\mathrm{E}(y_{t-1}\epsilon_{t-1}) = \sigma^2$, we have

$$\mathrm{E}[y_t(\epsilon_t + \beta\epsilon_{t-1})] = (1 + \alpha\beta + \beta^2)\sigma^2. \tag{67}$$

Multiplying (65) by y_{t-1} and taking expectations gives

$$\begin{aligned} \gamma_1 &= \mathrm{E}[\alpha y_{t-1}^2 + y_{t-1}(\epsilon_t + \beta\epsilon_{t-1})] \\ &= \alpha\gamma_0 + \beta\sigma^2. \end{aligned}$$

Deriving the ACov and ACF for an ARMA(1,1) Process

Combining (66)–(68) and solving for γ_0 gives us the formula for the variance of an ARMA(1,1) process

$$\gamma_0 = \frac{1 + 2\alpha\beta + \beta^2}{1 - \alpha^2}\sigma^2. \tag{68}$$

For the first order autocovariance we obtain from (65) and (66)

$$\begin{aligned} \gamma_1 &= \left(\frac{\alpha(1 + 2\alpha\beta + \beta^2)}{1 - \alpha^2} + \beta^2 \right) \sigma^2 \\ &= \frac{(1 + \alpha\beta)(\alpha + \beta)}{1 - \alpha^2}\sigma^2. \end{aligned} \tag{69}$$

Deriving the ACov and ACF for an ARMA(1,1) Process

Higher–order autocovariances can be computed recursively by

$$\gamma_k = \alpha\gamma_{k-1}, \qquad k = 2, 3, \dots \ . \tag{70}$$

Excursion: The AcovF for a general ARMA(p, q) process

Let y_t be generated by the stationary ARMA (p, q) process

$$\boldsymbol{\alpha}(L)y_t = \boldsymbol{\beta}(L)\epsilon_t, \tag{71}$$

where ϵ_t is the usual white–noise process with $\mathrm{E}(\epsilon_t) = 0$ and $\mathrm{E}(\epsilon_t^2) = \sigma^2$; and $\boldsymbol{\alpha}(L)$ and $\boldsymbol{\beta}(L)$ are polynomials defined by $a(L) = 1 - \alpha_1 L - \ldots - \alpha_r L^r$ and $\boldsymbol{\beta}(L) = \beta_0 + \beta_1 L + \ldots + \beta_r L^r$, with $r = \max(p, q)$ and $\alpha_i = 0$ for $i = p+1, p+2, \ldots, r$, if $r > p$ or $\beta_i = 0$ for $i = q+1, q+2, \ldots, r$, if $r > q$.

Excursion: The AcovF for a general ARMA(p, q) process

From the definition of the autocovariance, $\gamma_k = \mathrm{E}(y_t y_{t-k})$, it follows that

$$\begin{aligned} \gamma_k &= \alpha_1\gamma_{k-1} + \alpha_2\gamma_{k-2} + \ldots + \alpha_r\gamma_{k-r} \qquad (72) \\ &\quad + \mathrm{E}(\beta_0\epsilon_t y_{t-k} + \beta_1\epsilon_{t-1} y_{t-k} + \ldots + \beta_r\epsilon_{t-r} y_{t-k}), \qquad k = 0, 1, \ldots, r. \end{aligned}$$

Replacing y_{t-k} by its moving average representation,
$y_{t-k} = \beta(L)/\alpha(L)\epsilon_{t-k} = c(L)\epsilon_{t-k}$, where $c(L) = c_0 + c_1 L + c_2 L^2 \ldots$, we obtain

$$\mathrm{E}(\epsilon_{t-i} y_{t-k}) = \begin{cases} c_{i-k}\sigma^2, & \text{if } i = k, k+1, \ldots, r, \\ 0, & \textit{otherwise}. \end{cases}$$

Excursion: The AcovF for a general ARMA(p, q) process

Defining $\boldsymbol{\gamma} = (\gamma_0, \gamma_1, \ldots, \gamma_r)'$, $\boldsymbol{c} = (c_0, c_1, \ldots, c_r)'$ and using the fact that $\gamma_{k-i} = \gamma_{i-k}$, expression (72) can be rewritten in matrix terms as

$$\boldsymbol{\gamma} = M_\alpha \boldsymbol{\gamma} + N_\beta \boldsymbol{c} \sigma^2. \tag{73}$$

The $(r+1) \times (r+1)$ matrix M_α is the sum of two matrices, $M_\alpha = T_\alpha + H_\alpha$, with T_α denoting the lower-triangular Toeplitz matrix

$$T_\alpha = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ \alpha_1 & 0 & & & 0 \\ \alpha_2 & \alpha_1 & \ddots & & \vdots \\ \vdots & & & & 0 \\ \alpha_r & \alpha_{r-1} & \cdots & \alpha_1 & 0 \end{bmatrix},$$

Excursion: The AcovF for a general ARMA(p, q) process

and H_α is “almost" a Hankel matrix and given by

$$H_\alpha = \begin{bmatrix} 0 & \alpha_1 & \alpha_2 & \cdots & \alpha_{r-1} & \alpha_r \\ 0 & \alpha_2 & \alpha_3 & \cdots & \alpha_r & 0 \\ \vdots & \vdots & & & & \vdots \\ 0 & \alpha_{r-1} & \alpha_r & & & 0 \\ 0 & \alpha_r & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

Excursion: The AcovF for a general ARMA(p, q) process

Note that matrix H_α is not exactly Hankel due to the zeros in the first column. Finally, the Hankel matrix N_β is defined by

$$N_\beta = \begin{bmatrix} \beta_0 & \beta_1 & \cdots & \beta_{r-1} & \beta_r \\ \beta_1 & \beta_2 & \cdots & \beta_r & 0 \\ \vdots & & & & \vdots \\ \beta_{r-1} & \beta_r & & & 0 \\ \beta_r & 0 & \cdots & 0 & 0 \end{bmatrix}.$$

Excursion: The AcovF for a general ARMA(p, q) process

The initial autocovariances can be computed by

$$\gamma = (I - M_\alpha)^{-1} N_\beta c \sigma^2. \tag{74}$$

Since $c = (I - T_\alpha)^{-1}\beta$, a closed-form expression, relating the autocovariances of an ARMA process to its parameters α_i, β_i, and σ^2 is given by

$$\gamma = (I - M_\alpha)^{-1} N_\beta (I - T_\alpha)^{-1} \beta \sigma^2. \tag{75}$$

Excursion: The AcovF for a general ARMA(p, q) process

Note that $(I - T_\alpha)^{-1}$ always exists, since $\mid I - T_\alpha \mid = 1$, and that

$$N_\beta(I - T_\alpha)^{-1} = [(I - T_\alpha)^{-1}]' N_\beta,$$

since N_β is Hankel with zeros below the main counterdiagonal and $(I - T_\alpha)^{-1}$ is a lower-triangular Toeplitz matrix. Hence, (75) can finally be rewritten as

$$\gamma = [(I - T_\alpha')(I - M_\alpha)]^{-1} N_\beta \boldsymbol{\beta} \sigma^2. \tag{76}$$

Excursion: The AcovF for a general ARMA(p, q) process

Note that for $p < q = r$ only $p + 1$ equations have to be solved simultaneously. The corresponding system of equations is obtained by eliminating the last $p - q$ rows in (73); and higher–order autocovariances can be derived recursively by

$$\gamma_k = \begin{cases} \sum_{i=1}^{p} \alpha_i \gamma_{k-i} + \sigma^2 \sum_{j=k}^{q} \beta_j c_{j-k}, & \text{if } k = p+1, p+2, \ldots, q, \\ \sum_{i=1}^{p} \alpha_i \gamma_{k-i}, & \text{if } k = q+1, q+2, \ldots . \end{cases} \tag{77}$$

Excursion: The AcovF for a general ARMA(p, q) process

For pure autoregressive processes expression (76) reduces to

$$\gamma = [(I - T'_\alpha)(I - M_\alpha)]^{-1} s, \tag{78}$$

where the $(r+1) \times 1$ vector s is defined by $s = \sigma^2(\beta_0, 0, \ldots, 0)^T$. Thus, vector γ is given by the first column of $[(I - T'_\alpha)(I - M_\alpha)]^{-1}$ multiplied by $\sigma^2\beta_0$.

Excursion: The AcovF for a general ARMA(p, q) process

In the case of a pure MA process, (76) simplifies to

$$\gamma = N_\beta \boldsymbol{\beta} \sigma^2, \tag{79}$$

or

$$\gamma_k = \begin{cases} \sigma^2 \sum_{i=k}^{q} \beta_i \beta_{i-k}, & \text{if } k = 0, 1, \ldots, q, \\ 0, & \text{if } k > q. \end{cases} \tag{80}$$

The AcovF of an ARMA(1,1) reconsidered

Consider again the ARMA(1,1) process $y_t = \alpha_1 y_{t-1} + \epsilon_t + \beta_1 \epsilon_{t-1}$ from Example 3.4.2. To compute $\gamma = (\gamma_0, \gamma_1)'$, we now apply formula (76). Matrices T_α, H_α, N_β and vector $\boldsymbol{\beta}$ become:

$$T_\alpha = \begin{bmatrix} 0 & 0 \\ \alpha_1 & 0 \end{bmatrix}, \quad H_\alpha = \begin{bmatrix} 0 & \alpha_1 \\ 0 & 0 \end{bmatrix}, \quad N_\beta = \begin{bmatrix} 1 & \beta_1 \\ \beta_1 & 0 \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} 1 \\ \beta_1 \end{bmatrix}.$$

The AcovF of an ARMA(1,1) reconsidered

Simple matrix manipulations produce the desired result:

$$
\begin{aligned}
\gamma &= [(I - T'_\alpha)(I - M_\alpha)]^{-1} N_\beta \beta \sigma^2 \\
&= \begin{bmatrix} 1+\alpha_1^2 & -2\alpha_1 \\ -\alpha_1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & \beta_1 \\ \beta_1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ \beta_1 \end{bmatrix} \sigma^2 \\
&= \frac{1}{1-\alpha_1^2} \begin{bmatrix} 1 & 2\alpha_1 \\ \alpha_1 & 1+\alpha_1^2 \end{bmatrix} \begin{bmatrix} 1+\beta_1^2 \\ \beta_1 \end{bmatrix} \sigma^2 \\
&= \frac{\sigma^2}{1-\alpha^2} \begin{bmatrix} 1+\beta_1^2+2\alpha_1\beta_1 \\ \alpha_1(1+\beta_1^2)+\beta_1(1+\alpha_1^2) \end{bmatrix},
\end{aligned}
$$

which coincides with results (68) and (69) in the previous example.

An Example

Derive γ_0 and γ_1 using the stated procedure for the following process

$$y_t = 0.5y_{t-1} + \epsilon_t \tag{81}$$

with $\epsilon_t \sim N(0, 1)$.

The Yule-Walker Equations

Consider the AR(p) process

$$y_t = \alpha_1 y_{t-1} + \ldots + \alpha_p y_{t-p} + \epsilon_t$$

Multiplying both sides with y_{t-j} and taking expectations yields

$$\mathrm{E}(y_t y_{t-j}) = \alpha_1 \mathrm{E}(y_{t-1} y_{t-j}) + \ldots + \alpha_p \mathrm{E}(y_{t-p} y_{t-j})$$

which gives rise to the following equation system

$$\begin{aligned}
\gamma_1 &= \alpha_1\gamma_0 + \alpha_2\gamma_1 + \ldots + \alpha_p\gamma_{p-1} \\
\gamma_2 &= \alpha_1\gamma_1 + \alpha_2\gamma_0 + \ldots + \alpha_p\gamma_{p-2} \\
& \ldots \\
\gamma_p &= \alpha_1\gamma_{p-1} + \alpha_2\gamma_{p-2} + \ldots + \alpha_p\gamma_0
\end{aligned}$$

The Yule-Walker Equations

Or in matrix notation

$$\gamma = \boldsymbol{\alpha}\Gamma$$

with

$$\Gamma = \begin{bmatrix} \gamma_0 & \gamma_1 & \cdots \gamma_{p-1} \\ \gamma_1 & \gamma_0 & \cdots \gamma_{p-2} \\ \vdots & \ddots & \vdots \\ \gamma_{p-1} & \gamma_{p-2} & \cdots \gamma_0 \end{bmatrix}$$

We obtain a similar structure for the autocorrelation function by dividing by γ_0.

Partial Autocorrelation Function

The *partial autocorrelation function* (PACF) represents an additional tool for portraying the properties of an ARMA process. The definition of a *partial correlation coefficient* eludes to the difference between the PACF and the ACF. The ACF $\rho_k, k = 0, \pm 1, \pm 2, \ldots$, represents the *unconditional correlation* between y_t and y_{t-k}. By *unconditional correlation* we mean the correlation between y_t and y_{t-k} without taking the influence of the intervening variables $y_{t-1}, y_{t-2}, \ldots, y_{t-k+1}$ into account.

Partial Autocorrelation Function

The PACF, denoted by $\alpha_{kk}, k = 1, 2, \ldots$, reflects the net association between y_t and y_{t-k} over and above the association of y_t and y_{t-k} which is due to their common relationship with the intervening variables $y_{t-1}, y_{t-2}, \ldots, y_{t-k+1}$.

The PACF for an AR(1)

Consider the stationary AR(1) process

$$y_t = \alpha_1 y_{t-1} + \epsilon_t$$

Given that y_t and y_{t-2} are both correlated with y_{t-1}, we would like to know whether or not there is an additional association between y_t and y_{t-2} which goes beyond their common association with y_{t-1}.

The PACF for an AR(1)

Let ρ_{12}=Corr(y_t, y_{t-1}), ρ_{13}=Corr(y_t, y_{t-2}) and ρ_{23}=Corr(y_{t-1}, y_{t-2}). The partial correlation between y_t and y_{t-2} conditional on y_{t-1}, denoted by $\rho_{13,2}$, is

$$\rho_{13,2} = \frac{\rho_{13} - \rho_{12}\rho_{23}}{\sqrt{(1-\rho_{13}^2)(1-\rho_{23}^2)}}.$$

The PACF for an AR(1)

Considering an AR(1) process, we know that $\rho_{12} = \rho_{23} = \alpha_1$ and $\rho_{13} = \rho_2 = \alpha_1^2$. Hence, the partial autocorrelation between y_t and y_{t-2}, $\rho_{13,2}$, is zero. Denoting the partial autocorrelation between y_t and y_{t-k} by α_{kk}, it can be easily verified that for any AR(1) process $\alpha_{kk} = 0$, for $k = 2, 3, \ldots$. Since there are no intervening variables between y_t and y_{t-1}, the first-order partial autocorrelation coefficient is equivalent to the first order autocorrelation coefficient, i.e., $\alpha_{11} = \rho_1$. In particular for an AR(1) process we have $\alpha_{11} = \alpha_1$.

The PACF for a general AR process

Another way of interpreting the PACF is to view it as the sequence of the k-th autoregressive coefficients in a k-th order autoregression. Letting $\alpha_{k\ell}$ denote the ℓ-th autoregressive coefficient of an AR(k) process, the **Yule–Walker equations**

$$\rho_\ell = \alpha_{k1}\rho_{\ell-1} + \cdots + \alpha_{k(k-1)}\rho_{\ell-k+1} + \alpha_{kk}\rho_{\ell-k}, \qquad \ell = 1, 2, \ldots, k, \quad (82)$$

The PACF for a general AR process

$$\rho_\ell = \alpha_{k1}\rho_{\ell-1} + \cdots + \alpha_{k(k-1)}\rho_{\ell-k+1} + \alpha_{kk}\rho_{\ell-k}, \qquad \ell = 1, 2, \ldots, k, \quad (83)$$

give rise to the system of linear equations

$$\begin{bmatrix} 1 & \rho_1 & \cdots & \rho_{k-1} \\ \rho_1 & 1 & & \rho_{k-2} \\ \rho_2 & \rho_1 & & \rho_{k-3} \\ \vdots & & & \vdots \\ \rho_{k-2} & & & \rho_1 \\ \rho_{k-1} & \rho_{k-2} & \cdots & 1 \end{bmatrix} \begin{bmatrix} \alpha_{k1} \\ \alpha_{k2} \\ \alpha_{k3} \\ \vdots \\ \alpha_{k(k-1)} \\ \alpha_{kk} \end{bmatrix} = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \vdots \\ \rho_{k-1} \\ \rho_k \end{bmatrix}$$

or, in short,

$$P_k \boldsymbol{\alpha}_k = \underline{\rho}_k, \qquad k = 1, 2, \ldots \ . \quad (84)$$

The PACF for a general AR process

Using Cramér's rule, to successively solve (84) for α_{kk}, $k = 1, 2 \ldots$, we have

$$\alpha_{kk} = \frac{\mid P_k^* \mid}{\mid P_k \mid}, \qquad k = 1, 2, \ldots, \tag{85}$$

where matrix P_k^* is obtained by replacing the last column of matrix P_k by vector $\underline{\rho}_k = (\rho_1, \rho_2, \ldots, \rho_k)'$, i.e.,

$$P_k^* = \left[\begin{array}{cccc|c} 1 & \rho_1 & \cdots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & & \rho_{k-3} & \rho_2 \\ \rho_2 & \rho_1 & & \rho_{k-4} & \rho_3 \\ \vdots & & & \vdots & \vdots \\ \rho_{k-2} & & & 1 & \rho_{k-1} \\ \rho_{k-1} & \rho_{k-2} & \cdots & \rho_1 & \rho_k \end{array}\right]$$

The PACF for a general AR process

Applying (85), the first three terms of the PACF are given by

$$
\begin{aligned}
\alpha_{11} &= \frac{|\rho_1|}{|1|} = \rho_1, \\
\alpha_{22} &= \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2},
\end{aligned}
$$

The PACF for a general AR process

$$\alpha_{33} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}} = \frac{\rho_3 + \rho_1 \rho_2 (\rho_2 - 2) - \rho_1^2 (\rho_3 - \rho_1)}{(1 - \rho_2) - (1 - \rho_2 - 2\rho_1^2)}.$$

The PACF for a general AR process

From the Yule–Walker equations it is evident that $\mid P_k^* \mid = 0$ for an AR process whose order is less than k, since the last column of matrix P_k^* can always be obtained from a linear combination of the first $k-1$ (or less) columns of P_k^*. Hence, the theoretical PACF of an AR(p) will generally be different from zero for the first p terms and exactly zero for terms of higher order. This property allows us to identify the order of a pure AR process from its PACF.

The PACF for a MA(1) process

Consider the MA(1) process $y_t = \epsilon_t + \beta_1 \epsilon_{t-1}$. Its ACF is given by

$$\rho_k = \begin{cases} \frac{\beta_1}{1+\beta_1}, & \text{if k=1}, \\ 0, & \text{if k=2,3,} \ldots . \end{cases}$$

Applying (85), the first 4 terms of the PACF are:

$$\begin{aligned} \alpha_{11} &= \rho_1, \ \alpha_{22} = -\frac{\rho_1^2}{1-\rho_1^2}, \\ \alpha_{33} &= \frac{\rho_1^3}{1-2\rho_1^2}, \ \alpha_{44} = -\frac{\rho_1^4}{1-3\rho_1^2+\rho_1^4}. \end{aligned} \tag{86}$$

The PACF for a MA(1) process

In fact, the general expression for the PACF of an MA(1) process in terms of the MA coefficient β_1 is

$$\alpha_{kk} = -\frac{(-\beta_1)^k(1-\beta_1^2)}{1-\beta_1^{2(k+1)}}.$$

$\Rightarrow$ PACF gradually dies out, in contrast to an AR process
$\Rightarrow$ this allows us to identify processes by looking at its corresponding ACF and PACF

Characteristics of specific processes

Identification Functions:

1. **autocorrelation function (ACF)**, ρ_k,
2. **partial autocorrelation function (PACF)**, α_{kk},

Characteristics of AR processes

- ACF*:* The Yule–Walker equations

$$\rho_k = \alpha_1 \rho_{k-1} + \alpha_2 \rho_{k-2} + \ldots + \alpha_p \rho_{k-p}, \qquad k = 1, 2, \ldots$$

  imply that the ACF of a stationary AR process is generally different from zero but gradually dies out as k approaches infinity.
- PACF: The first p terms are generally different from zero; higher-order terms are identically zero.

Characteristics of MA Processes

- ACF: We know that the ACF of an MA(q) process is given by

$$\gamma_k = \begin{cases} \sigma^2 \sum_{i=k}^{q} \beta_i \beta_{i-k}, & \text{if } k = 0, 1, \ldots, q \\ 0, & \text{if } k > q, \end{cases}$$

which implies that the ACF is generally different from zero up to lag q and equal to zero thereafter.

- PACF: The PACF is computed successively by

$$\alpha_{kk} = \frac{\mid P_k^* \mid}{\mid P_k \mid}, \qquad k = 1, 2, \ldots,$$

with matrices P_k^* and P_k defined in the Section before. Example 3.6.2 demonstrated the pattern of the PACF of an MA(1) process.

ACF and PACF

	Model		
	AR(p)	MA(q)	ARMA(p, q)
ACF	tails off	cuts off after q	tails off
PACF	cuts off after p	tails off	tails off

Table: Patterns for Identifying ARMA Processes