Chapter 11 Identification of Box-Jenkins Models

In Chapter 10 we discussed the variety of Box-Jenkins *ARIMA* models. In this Chapter 11 we discuss the methods by which we determine the appropriate *ARIMA* model for the given time series data. We are thus at Stage 2 in the Forecasting Process — the Identification Stage of *ARIMA* modeling.

Every stationary time series has an *ACF*, an *autocorrelation function.* The *ACF* plays the analagous role as the scatterdiagram of Chapters 3, 4 and 5. The *ACF* suggests the kind of *ARIMA* model appropriate to the time series data.

We shall discuss the *ACF* of a Time Series Process in a general, non-mathematical framework. For those readers interested in the mathematics of the *ACF*, the proofs and derivations are found in Appendix B.

Chapter 11

An economist is an expert who will know tomorrow why the things he predicted yesterday didn't happen today.

– Evan Esar

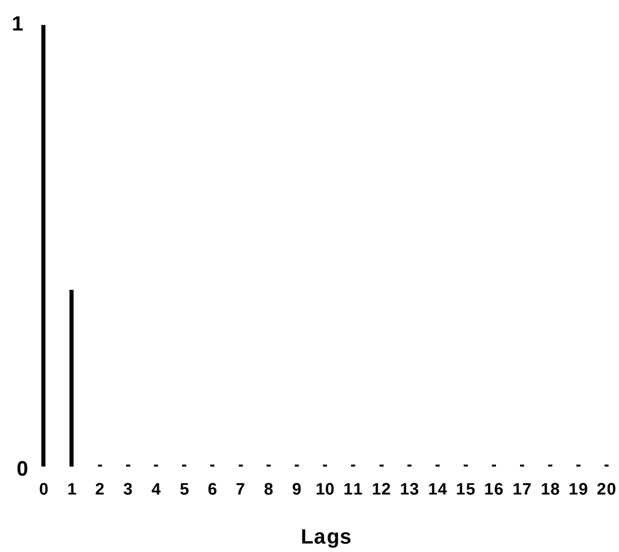
The Autocorrelation Functions of *MA* (Moving Average) processes

The *ACF* of an *MA(1)* process

In general terms an *MA(1)* process is written

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

The characteristic correlogram of the *MA(1)* process is a significant spike at lag 1 and zero autocorrelation at all other lags.



In other words, the series of autocorrelations

$$\rho_0 = 1, \quad \rho_1 = \text{significant}, \quad \rho_2 = 0, \quad \rho_3 = 0, \quad \rho_4 = 0, \ \ldots$$

is indicative of an *MA(1)* process.

Consequently, when we observe an *ACF* of an unknown process, and it has only one spike at lag 1, and zeros elsewhere, we then tentatively identify the time series as being generated by an *MA(1)* process.

There is a formula relating the *MA(1)* process and the "size" of the significant spike at lag 1, the autocorrelation at lag 1.

Autocorrelation at lag 1 of an MA(1) process

$$\rho_1 = \frac{-\theta_1}{(1+\theta_1^2)} \qquad \textbf{11.1}$$

Example 1

Consider an *MA(1)* process generated by

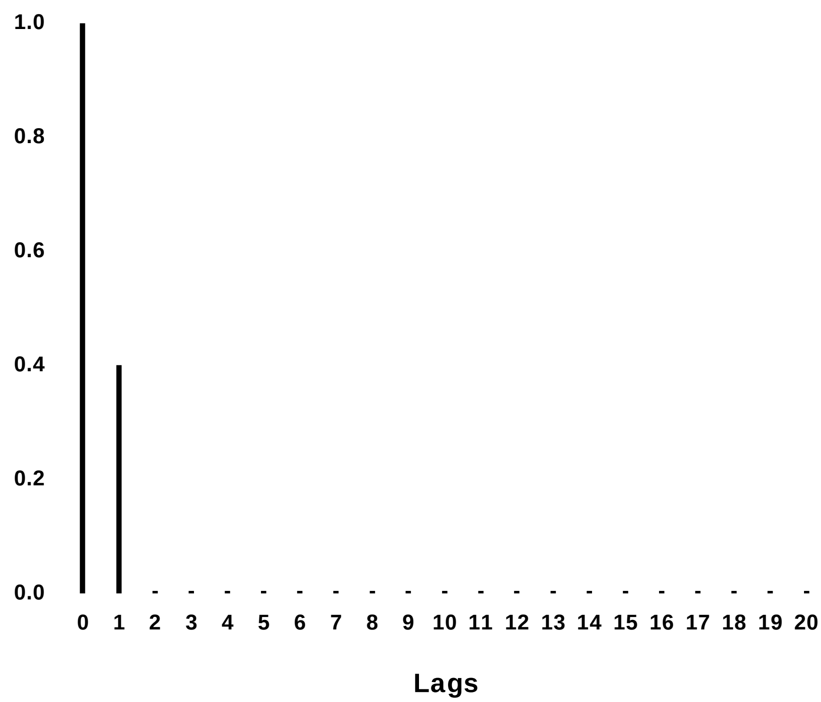
$$Y_t = 7 + .5\epsilon_{t-1} + \epsilon_t, \qquad \mu = 7 \qquad \theta_1 = -.5$$

$$\rho_1 = \frac{-\theta_1}{(1+\theta_1^2)} = \frac{-(-.5)}{\left(1+(-.5)^2\right)} = \frac{.5}{1.25} = .40$$

Table 11.1

ACF

Lags	0	1	2	3	4	5	$\cdots$
Autocorrelation	1	.40	0	0	0	0	$\cdots$



With the above example, the *MA(1)* model has about 40% "memory" of what occured 1 period back and then zero memory beyond 1 period.

Example 2

Consider an *MA(1)* process generated by

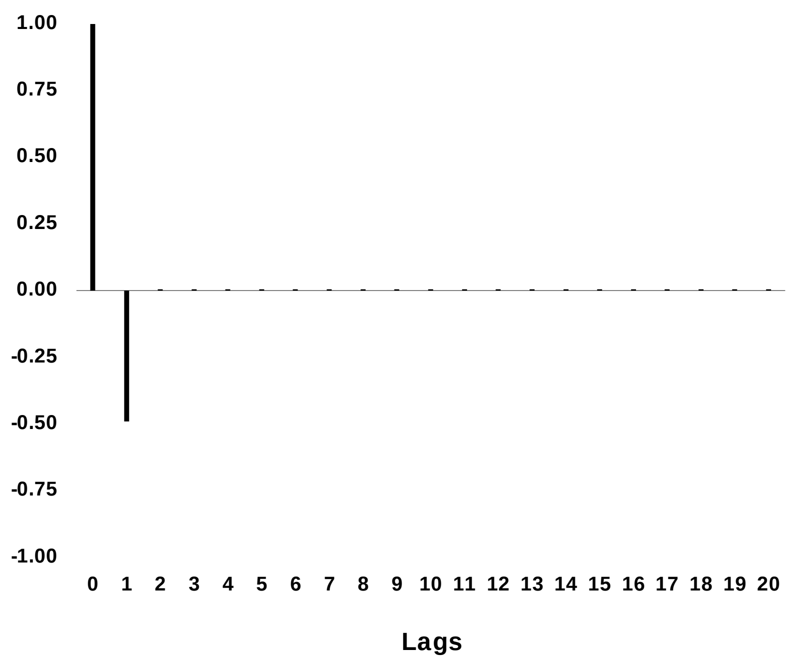
$$Y_t \;=\; 95 - .8\epsilon_{t-1} + \epsilon_t, \qquad \mu \;=\; 95 \qquad \theta_1 \;=\; .8$$

$$\rho_1 \;=\; \frac{-\theta_1}{(1+\theta_1^2)} \;=\; \frac{-(.8)}{(1+.8^2)} \;=\; \frac{-.8}{1.64} \;=\; -.49$$

Table 11.2

ACF

Lags	0	1	2	3	4	5	$\cdots$
Autocorrelations	1	-.49	0	0	0	0	$\cdots$



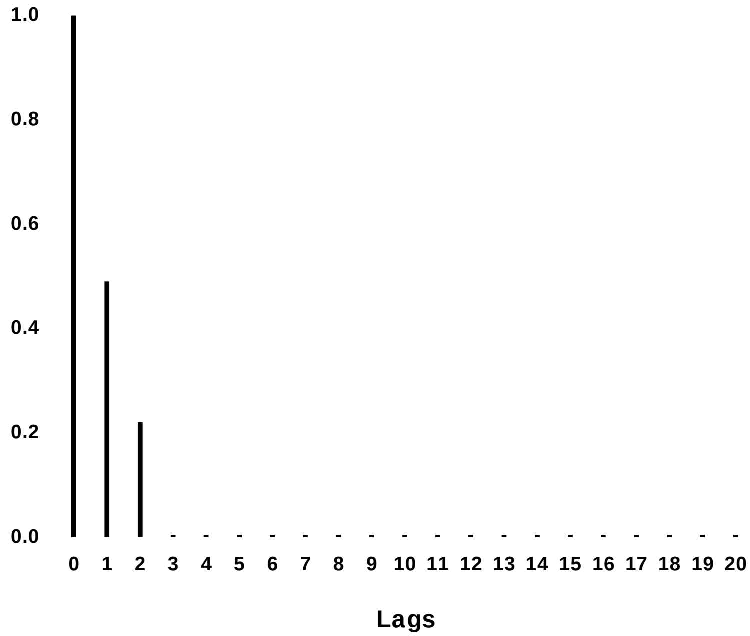
The Autocorrelation Functions of *MA* processes continued

The *ACF* of an *MA(2)* process

In general terms an *MA(2)* process is written

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

The characteristic correlogram of the *MA(2)* process is a significant spike at lag 1, a significant spike at lag 2 and zero autocorrelation at all other lags.



In other words, the series of autocorrelations

$$\rho_0 = 1, \quad \rho_1 = \text{significant}, \quad \rho_2 = \text{significant}, \quad \rho_3 = 0, \quad \rho_4 = 0, \ \ldots$$

is indicative of an *MA(2)* process.

Consequently, when we observe an *ACF* of an unknown process, and it has two significant spikes, at lags 1 and 2, and zeros elsewhere, we then tentatively identify the time series as being generated by an *MA(1)* process.

There are a formulas relating the *MA(2)* process and the "size" of the significant spike at lags 1 and 2, the autocorrelations at lags 1 and 2.

Autocorrelation at lag 1 of an MA(2) process

$$\rho_1 = \frac{-\theta_1 + \theta_1\theta_2}{\left(1 + \theta_1^2 + \theta_2^2\right)} \qquad \textbf{11.2}$$

Autocorrelation at lag 2 of an MA(2) process

$$\rho_2 = \frac{-\theta_2}{\left(1 + \theta_1^2 + \theta_2^2\right)} \qquad \textbf{11.3}$$

Example 3

Consider an MA(2) process with $\mu = 15$, $\theta_1 = .5$ and $\theta_2 = -.3$.

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

$$Y_t = 15 - .5\epsilon_{t-1} - (\text{-}.3)\epsilon_{t-2} + \epsilon_t$$

Then,

$$\rho_0 = 1$$

$$\rho_1 = \frac{-\theta_1 + \theta_1\theta_2}{(1+\theta_1^2+\theta_2^2)} = \frac{-.5 + (.5)(-.3)}{(1+(.5)^2+(-.3)^2)} = \frac{-.65}{1.34} = -.49$$

$$\rho_2 = \frac{-\theta_2}{(1+\theta_1^2+\theta_2^2)} = \frac{-(-.3)}{(1+(.5)^2+(-.3)^2)} = \frac{.3}{1.34} = .22$$

$$\rho_3 = 0$$

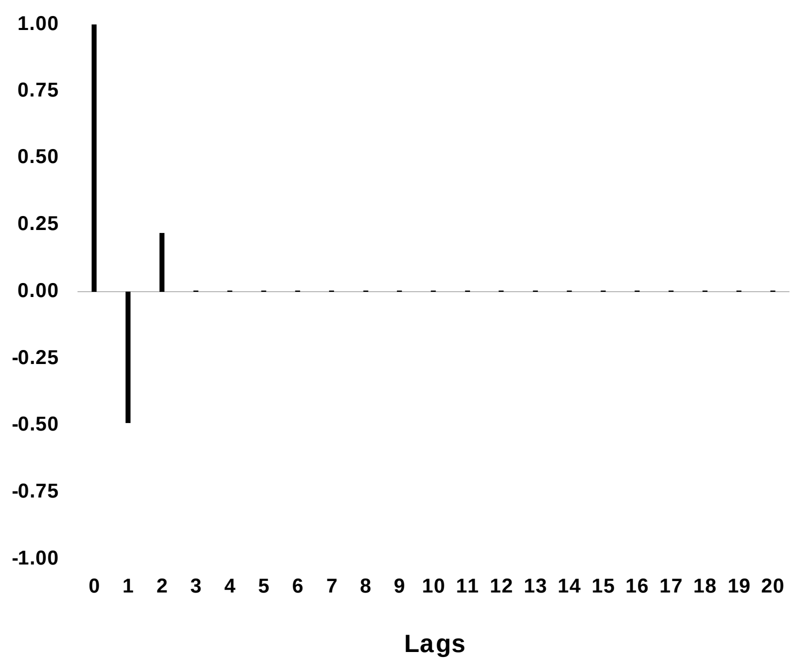
$$\rho_4 = 0$$

$$\vdots$$

Table 11.3

ACF

Lags	0	1	2	3	4	5	$\cdots$
Autocorrelations	1	-.49	.22	0	0	0	$\cdots$



An *MA(2)* process will have two predominant spikes and then zeros for the rest of the lags. Note: Depending on the signs and magnitudes of the parameters there will be different *ACF*'s and correlograms.

The *ACF* of a general *MA(q)* process

We hope the pattern is clear now; that an *MA(3)* process will have three significant spikes followed by zeros, and *MA(4)* process will have four spikes, then zeros, and so on. The *MA(q)* process will have spikes at lags 1, 2, 3, ... , q and then zeros beyond.

We generalize sequence of ρ_k's for the *MA(q)* process to:

ρ_k for an MA(q) process

$$\rho_k = \frac{-\theta_k + \theta_1\theta_{k+1} + \ldots + \theta_{q-k}\theta_q}{1 + \theta_1^2 + \theta_2^2 + \ldots + \theta_q^2}, \qquad \text{for } k = 1, 2, \ldots q \qquad 11.4$$

$$\rho_k = 0, \qquad \text{for } k > q$$

The *ACF* of the *MA(q)* process cuts off after lag q. The "memory" of such a time series process extends only q periods, beyond q periods the observations are uncorrelated.

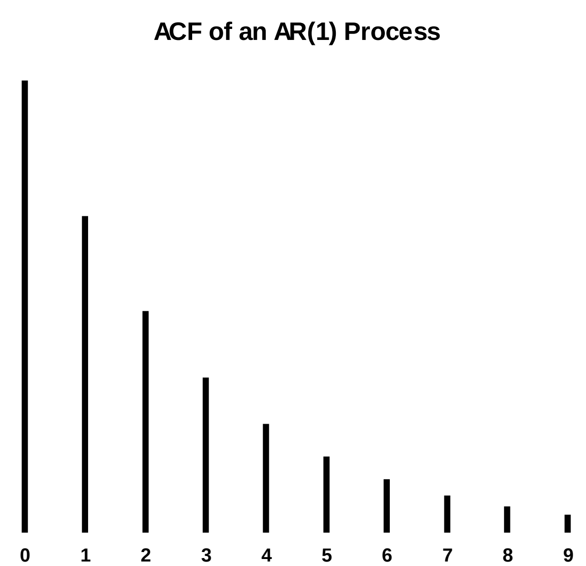
The Autocorrelation Functions of *AR* processes

The *ACF* of an *AR(1)* process

The general *AR(1)* process is written

$$Y_t = \delta + \phi_1 Y_{t-1} + \epsilon_t \qquad 11.5$$

The characteristic correlogram of an *AR(1)* process is a series of significant spikes at lags 1, 2, 3, ... exponentially decaying to zero.



As an example, the series of autocorrelations

$$\rho_0 = 1, \qquad \rho_1 = .70, \qquad \rho_2 = .49, \qquad \rho_3 = .343, \qquad \rho_4 = .2401, \qquad \ldots$$

is indicative of an *AR(1)* process.

Consequently, when we observe an *ACF* of an unknown process, and it has an exponentially decaying spikes, we then tentatively identify the time series as being generated by an *AR(1)* process.

The formulas for the autocorrelations of an *AR(1)* process are quite direct.

The Autocorrelation at lag 1 of an AR*(1) process*

$$\boldsymbol{\rho_1 = \phi_1} \qquad \textbf{11.6}$$

The Autocorrelation at lag 2 for an AR*(1) process*

$$\boldsymbol{\rho_2 = \phi_1^2} \qquad \textbf{11.7}$$

The Autocorrelation at lag 3 for an *AR*(1) *process*

$$\boldsymbol{\rho_3 = \phi_1^3} \qquad \textbf{11.8}$$

$$\vdots$$

In general,

The Autocorrelation at lag *k* for an *AR*(1) process

$$\boldsymbol{\rho_k = \phi_1^k} \qquad \textbf{11.9}$$

Example 4

Consider an *AR(1)* with $\phi_1 = .7$ and $\delta = 63$.

$$Y_t = \delta + \phi_1 Y_{t-1} + \epsilon_t$$

$$Y_t = 63 + .7Y_{t-1} + \epsilon_t$$

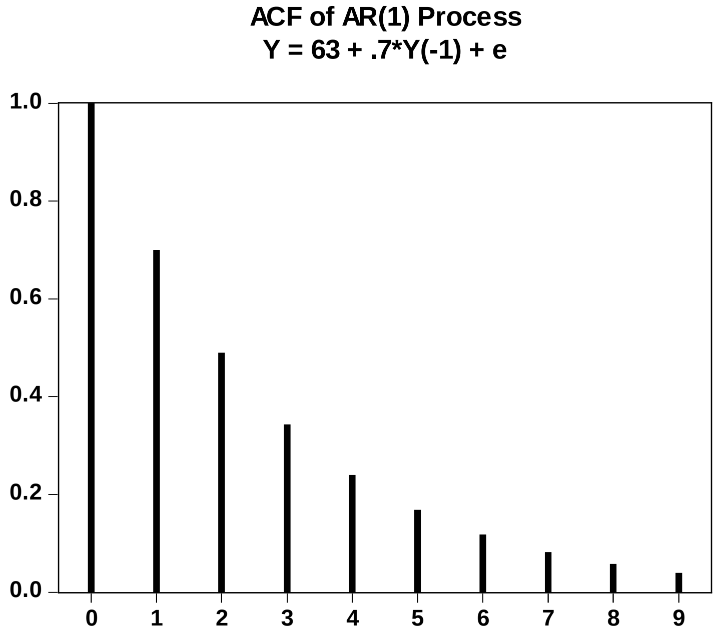
Then,

$$\begin{aligned} \rho_0 &= 1 \\ \rho_1 &= .7 = .7 \\ \rho_2 &= (.7)^2 = .49 \\ \rho_3 &= (.7)^3 = .343 \\ \rho_4 &= (.7)^4 = .2401 \\ &\vdots \end{aligned}$$

Table 11.4

ACF

Lags	0	1	2	3	4	5	⋯
Autocorrelations	1	.70	.49	.343	.2401	.168	⋯



Example 5

Consider an *AR(1)* with $\phi_1 = -.7$ and $\delta = 53$.

$$Y_t = \delta + \phi_1 Y_{t-1} + \epsilon_t$$

$$Y_t = 53 - .7 Y_{t-1} + \epsilon_t$$

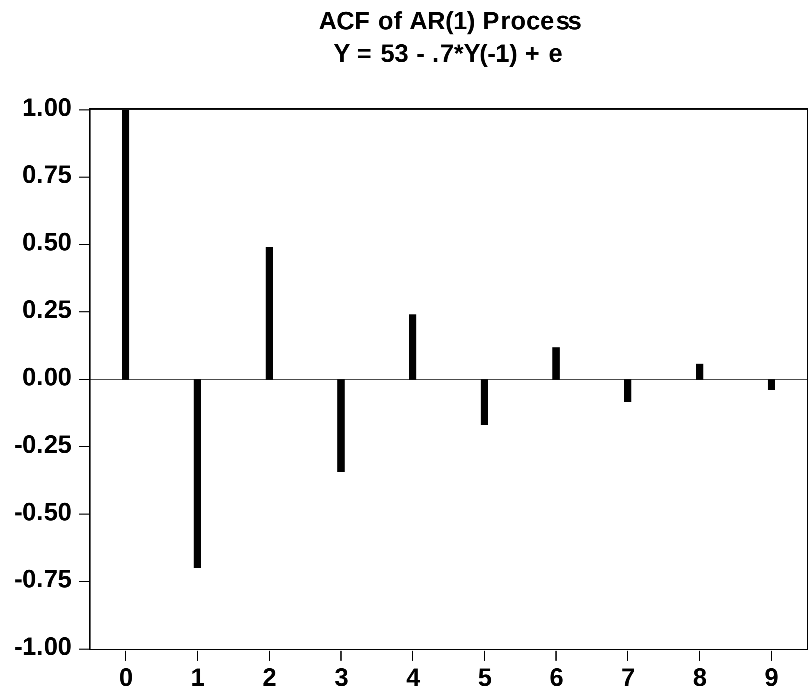
Then,

$$\begin{aligned} \rho_0 &= 1 \\ \rho_1 &= (-0.7)^1 = -0.7 \\ \rho_2 &= (-0.7)^2 = +.49 \\ \rho_3 &= (-0.7)^3 = -.343 \\ \rho_4 &= (-0.7)^4 = +.2401 \\ &\vdots \end{aligned}$$

Table 11.5

ACF

Lags	0	1	2	3	4	5	$\cdots$
Autocorrelations	1	-.7	.49	-.343	.2401	-.168	$\cdots$



Both correlograms have spikes that "decay exponentially." In the first example, $\phi_1 = .7$ is positive so the spikes are all positive. In the second example, $\phi_1 = -.7$ is negative so the spikes are alternately positive and negative. Unlike the *MA* models, the correlogram of the *AR(1)* never cuts off to zero, it dampens to zero.

The Autocorrelation Functions of *AR* processes continued

The *ACF* of an *AR*(2) process

The general *AR*(2) process is written

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

The characteristic correlogram of an *AR*(2) process is similar to the correlogram of an *AR*(1) process. The correlogram is a series of spikes at lags 1, 2, 3, $\ldots$ exponentially decaying to zero.

The formulas for the autocorrelations of an *AR*(2) process are usually written in a recursive manner.

The Autocorrelation at lag 1 of an *AR*(2) *process*

$$\boldsymbol{\rho_1 \ = \ \frac{\phi_1}{1-\phi_2}} \qquad \textbf{11.10}$$

The Autocorrelation at lag 2 for an *AR*(2) *process*

$$\boldsymbol{\rho_2 \ = \ \phi_1 \rho_1 \ + \ \phi_2 \rho_0} \qquad \textbf{11.11}$$

The Autocorrelation at lag 3 for an *AR*(2) *process*

$$\boldsymbol{\rho_3 \ = \ \phi_1 \rho_2 \ + \ \phi_2 \rho_1} \qquad \textbf{11.12}$$

The Autocorrelation at lag 4 for an *AR*(2) *process*

$$\boldsymbol{\rho_4 \ = \ \phi_1 \rho_3 \ + \ \phi_2 \rho_2} \qquad \textbf{11.13}$$

$$\vdots$$

Each new autocorrelation requires the autocorrelation of the previous lag and the autocorrelation of the lag before that.

The Autocorrelation at lag k for an *AR*(2) *process*

$$\boldsymbol{\rho_k \ = \ \phi_1 \rho_{k-1} \ + \ \phi_2 \rho_{k-2}} \qquad \textbf{11.14}$$

Example 6

Consider an *AR*(2) with $\delta \ = \ 17$, $\phi_1 \ = \ .5$, $\phi_2 \ = \ .3$

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

$$Y_t \ = \ 17 \ + \ .5Y_{t-1} \ + \ .3Y_{t-2} \ + \ \epsilon_t$$

$$\rho_0 \ = \ 1$$

$$\rho_1 \ = \ \frac{\phi_1}{1-\phi_2} \ = \ \frac{.5}{.7} \ = \ .7143$$

$$\rho_2 \ = \ \phi_1 \rho_1 \ + \ \phi_2 \rho_0 \ = \ (.5)(.7143) \ + \ (.3)(1) \ = \ .6571$$

$$\rho_3 = \phi_1\rho_2 + \phi_2\rho_1 = (.5)(.6571) + .3(.7143) = .5429$$

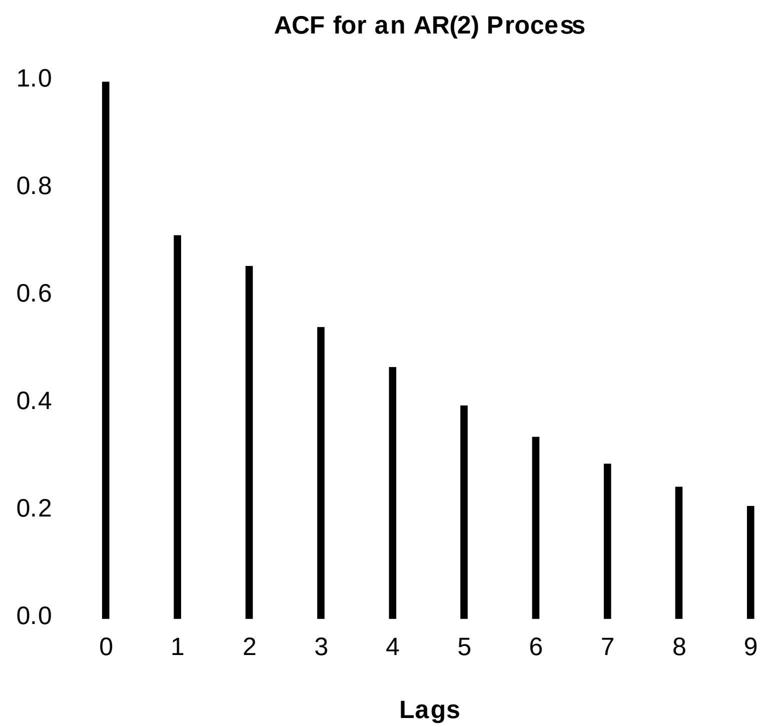
$$\rho_4 = \phi_1\rho_3 + \phi_2\rho_2 = (.5)(.5429) + .3(.6571) = .4686$$

$$\vdots$$

Table 11.6

ACF

Lags	0	1	2	3	4	5	$\cdots$
Autocorrelations	1	.7143	.6571	.5429	.4686	.3972	$\cdots$



Example 7

Consider an *AR*(2) with $\delta = 111$, $\phi_1 = -.45$, $\phi_2 = .40$

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

$$Y_t = 111 - .45Y_{t-1} + .40Y_{t-2} + \epsilon_t$$

$$\rho_0 = \qquad = 1$$

$$\rho_1 = \frac{\phi_1}{1-\phi_2} = \frac{-.45}{.60} = \qquad = -.750$$

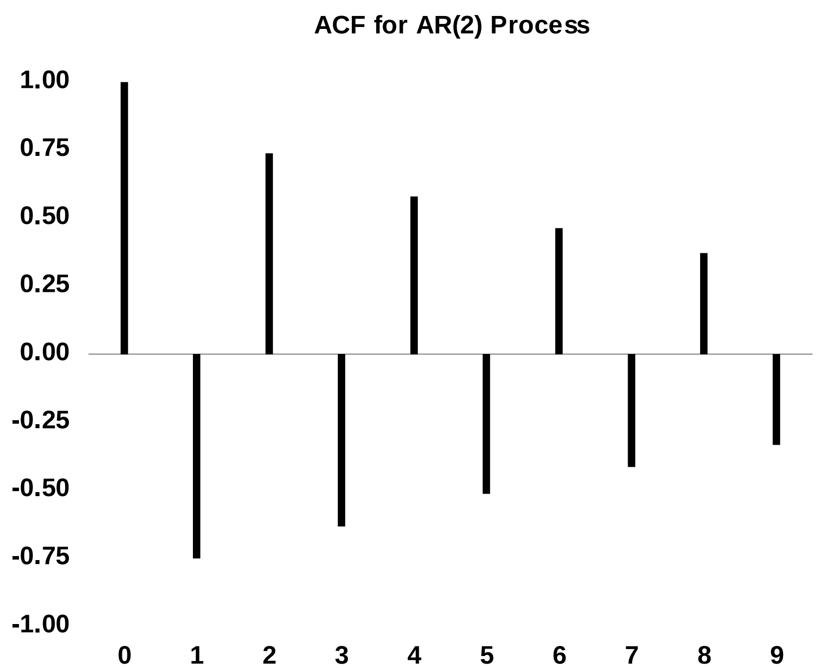
$$\rho_2 = \phi_1\rho_1 + \phi_2\rho_0 = (-.45)(-.750) + (.40)(1) = .738$$

$$\rho_3 = \phi_1\rho_2 + \phi_2\rho_1 = (-.45)(\ .738) + (.40)(-.750) = -.632$$

$$\rho_4 = \phi_1\rho_3 + \phi_2\rho_2 = (-.45)(-.632) + (.40)(.738) = .579$$

Table 11.7

ACF *Lags*	0	1	2	3	4	5	⋯
Autocorrelations	1	-.750	.738	-.632	.579	-.513	⋯



The correlogram of an *AR*(2) usually has two decaying sets of spikes and decays much more slowly than an *AR*(1).

Depending on the magnitude and sign of the parameters the *AR*(2) ACF will take on different patterns.

The *ACF* of a general *AR*(p) process

Autocorrelation at lag k, ρ_k, of an $AR(p)$ process

$$\boldsymbol{\rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-2} + \ldots + \phi_p\rho_{k-p}} \qquad \textbf{11.15}$$

$$\begin{array}{ll}
p = 1 & \rho_k = \phi_1^k \\
p = 2 & \rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-2} \\
p = 3 & \rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-2} + \phi_3\rho_{k-3} \\
\vdots & \quad\vdots \\
p = p & \rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-2} + \phi_3\rho_{k-3} + \ldots + \phi_p\rho_{k-p} \\
\vdots & \quad\vdots
\end{array}$$

Partial Autocorrelation

Since all *AR* processes produce *ACF*'s which dampen out, it can be difficult to distinguish among *AR* models of different orders. However, by using a ***partial autocorrelation function, the PACF***, we may distinguish among *AR* models.

When we discuss autocorrelation at lag 2, ρ_2, we are considering the correlation between observations Y_t and Y_{t-2}; observations two periods apart.

$$\rho_2 = CORR(Y_t, Y_{t-2})$$

However, the reason that Y_t and Y_{t-2} are correlated may be due to their correlation with Y_{t-1}, the observation between them.

I.e. Y_t and Y_{t-2} are correlated because Y_t is correlated with Y_{t-1} and Y_{t-1} is correlated with Y_{t-2}.

To adjust for intermediary correlations, we construct the ***partial autocorrelations***, and the ***partial autocorrelation function, PACF***.

While autocorrelations at lag *k* are denoted ρ_k, partial autocorrelations at lag *k* are denoted ρ_{kk}.

Note that the partial autocorrelation at lag 1 is identical with the autocorrelation at lag 1 since there is no intermediary correlations. We shall list some of the formulas for the partial autocorrelations.

AR(1) process

Partial Autocorrelations

$$\begin{array}{ll}
\text{Lag 1} & \rho_{11} = \rho_1 = \phi_1 \\
\text{Lag 2} & \rho_{22} = 0 \\
\text{Lag 3} & \rho_{33} = 0 \\
\vdots & \quad\vdots \\
\text{Lag k} & \rho_{kk} = 0
\end{array}$$

$\rho_{kk} = 0$ for *k* = 2, 3, . . .

***AR*(2)** process

Partial Autocorrelations

$$\begin{array}{ll}
\text{Lag 1} & \rho_{11} = \rho_1 \\
\text{Lag 2} & \rho_{22} = \frac{\rho_2 - \rho_1^2}{2} \\
\text{Lag 3} & \rho_{33} = 0 \\
\vdots & \quad\vdots \\
\text{Lag } k & \rho_{kk} = 0
\end{array}$$

$\rho_{kk} = 0$ for *k* = 3, 4, . . .

In general, for an *AR*(p) process

$$\rho_{kk} = 0 \text{ for } k = p + 1, p + 2, \ldots$$

The partial autocorrelations are zero for lags that are larger than the order of the process. This fact allows us to better identify an *AR* processes.

Hence, an *AR*(p) process is identified by

1) An *ACF* which is an infinited damped series,
2) A *PACF* that is zero for lags larger than *p*.

In other words, an *AR*(p) will have a decaying *ACF* and a *PACF* which looks like an *MA*(p), as illustrated below by Example 8.

Example 8

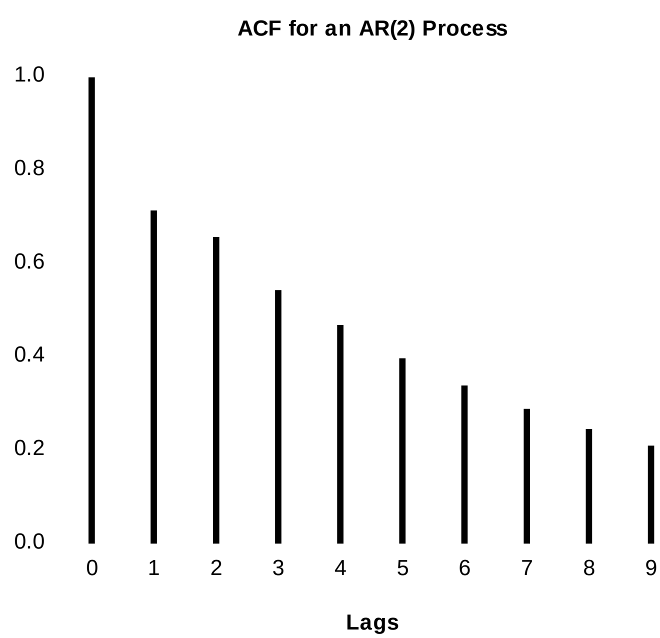
Consider an *AR*(2) with $\delta = 17$, $\phi_1 = .5$, $\phi_2 = .3$

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

$$Y_t = 17 + .5Y_{t-1} + .3Y_{t-2} + \epsilon_t$$

Table 11.8

ACF							
Lags	0	1	2	3	4	5	$\cdots$
Autocorrelations	1	.7143	.6571	.5429	.4686	.3972	$\cdots$



Partial Autocorrelations

Lag 1 $\rho_{11} = \rho_1 = .7143$

Lag 2 $\qquad \rho_{22} = \frac{\rho_2 - \rho_1^2}{2} = \frac{.6571 - (.7143)^2}{2} = .073$

Lag 3 $\qquad \rho_{33} = 0$

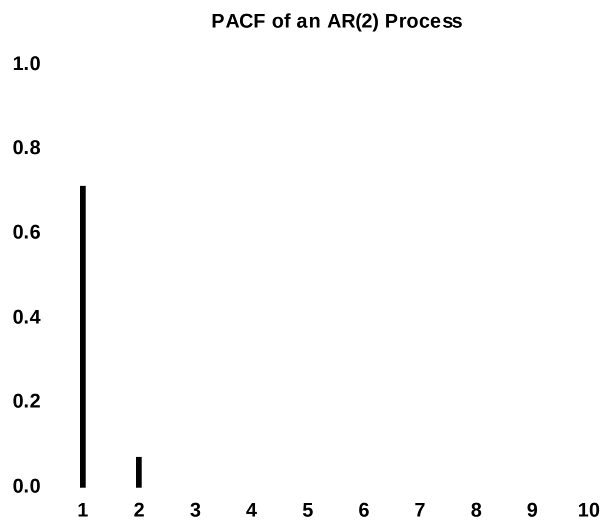
$\vdots \qquad \vdots$

Lag k $\qquad \rho_{kk} = 0$

$\vdots \qquad \vdots$

Table 11.9

PACF						
Lags	1	2	3	4	5	$\cdots$
Partial Autocorrelations	.7143	.073	0	0	0	$\cdots$



The *PACF* for *MA*(q) models

It is not difficult to show that the *PACF* for an *MA* process is the converse of the *AR* process. While the *ACF* of an *MA*(q) model cuts off after *q* lags, the *PACF* of an *MA*(q) is infinite in extent and decays.

Example 9

Consider the *ACF* and the *PACF* of the following *MA*(1) process, with $\mu = 7$ and $\theta_1 = $ -,5.

$$Y_t = 7 + .5\epsilon_{t-1} + \epsilon_t$$

Table 11.10

ACF							
Lags	0	1	2	3	4	5	$\cdots$
Autocorrelation	1	.40	0	0	0	0	$\cdots$

Table 11.11

PACF						
Lags	1	2	3	4	5	$\cdots$
Partial Autocorrelation	.40	-.191	.094	-.047	.0204	$\cdots$

Mixed Models

The *ACF* and the *PACF* of *ARMA*(p,q) processes and the Multplicity of Models

The *ACF* and *PACF* of *ARMA*(p,q) processes are combinations of the *AR* and *MA* *ACF*'s and *PACF*'s. Obviously, the higher the order of *p* and *q* the more difficult it is identify the precise *ARMA*(p,q) model. However, it has been proven that a stochastic time series can be modelled by more than one *ARMA*(p,q) model. Thus, if we are not certain if our time series is an *ARMA*(1,2) or an *ARMA*(2,2), it is not necessarily critical, since each may perform equally well for forecasting purposes.

Related to this notion of a multiplicity of models for a single time series is the ***principle of parsimony***. A parsimonious model is one that uses the fewest number of parameters. We mention the concept of parsimony here in that when attempting to identify a model, the preferred choice is the model with the fewest parameters. So, for example, if we are uncertain whether a given process is an *ARMA*(1,1) or an *ARMA*(1,2), we choose the *ARMA*(1,1), first, for more thorough analysis. An *ARMA*(1,1) has three parameters, whereas an *ARMA*(1,2) has four.

Example 10

Consider an *ARMA*(1, 1) process, with $\delta = 23$, $\phi_1 = .5$, and $\theta_1 = .3$.
In backshift notation

$$(1 - \phi_1 B)Y_t = \delta + (1 - \theta_1 B)\epsilon_t$$

In expanded notation

$$Y_t = \delta + \phi_1 Y_{t-1} - \theta_1 \epsilon_{t-1} + \epsilon_t$$
$$Y_t = 23 + .5Y_{t-1} - .3\epsilon_{t-1} + \epsilon_t$$

The *ACF* and *PACF* of this process are listed below, as well as the correlograms.

Table 11.12

ACF

Lags	0	1	2	3	4	5	$\cdots$
Autocorrelation	1	.215	.103	.053	.027	.013	$\cdots$

Table 11.13

PACF

Lags	1	2	3	4	5	$\cdots$
Partial Autocorrelation	.215	.041	.058	-.014	.029	$\cdots$

The *ACF* of the *ARMA*(1,1) is similar to that of an *AR*(1), having an exponential dampening. However, the dampening does not begin at lag 0, but rather at lag 1. At ρ_1 we have a spike, then the exponential dampening follows. The *PACF* consists of a single initial value ρ_{11}. From there it behaves like the *PACF* of an *MA*(1) process that is dominated by an exponential dampening.

Nonstationary Processes

All of the discussion in this chapter has assumed that the processes we are identifying are stationary. However, as we are aware, most business and economic time series are non-stationary, usually having trend or changing variance. As we have discussed in Chapter 10, differencing is the primary method of removing trend.

Recall that first differencing is

$$Z_t = Y_t - Y_{t-1}$$

With Z_t now the stationary series, we consider the *ACF* and the *PACF* of the series Z_t. We conclude that Z_t is an *ARMA*(1,1), which in turn means that Y_t is an *ARIMA*(1,1,1) process.

Z_t being an *ARMA*(1,1) means

$$(1 - \phi_1 B)Z_t = \delta + (1 - \theta_1 B)\epsilon_t$$

$$Z_t = \delta + \phi_1 Z_{t-1} - \theta_1 \epsilon_{t-1} + \epsilon_t$$

This is equivalent to Y_t being an *ARIMA*(1,1,1).

$$(1 - \phi_1 B)(1 - B)Y_t = \delta + (1 - \theta_1 B)\epsilon_t$$

$$Y_t = \delta + (1 + \phi_1)Y_{t-1} - Y_{t-2} - \theta_1 \epsilon_{t-1} + \epsilon_t$$

The *ACF* and *PACF* of Time Series containing Seasonality

Many business and economic time series have seasonality, either monthly, quarterly, or yearly seasonality. To remove seasonality from a time series, so as to make it stationary, we usually difference by the degree of the seasonality.

If the data are quarterly, then we difference by lag 4. For example, if Y_t is a quarterly time series, then

$$Z_t = Y_t - Y_{t-4}$$

$$Z_t = (1 - B^4)Y_t$$

If Y_t is a monthly time series, then

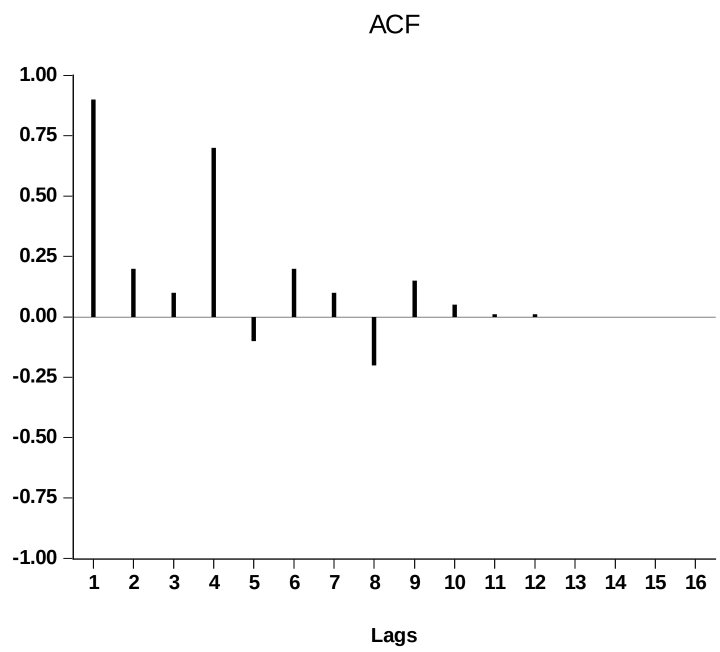
$$Z_t = Y_t - Y_{t-12}$$

$$Z_t = (1 - B^{12})Y_t$$

The seasonal component of a time series may, in itself, be an *ARMA* process. In other words, there may be a *seasonal ARMA* process in conjunction with the underlying *regular ARMA* process.

Example 11

Consider a quarterly time series which is a *seasonal MA*(1) in addition to being a *regular MA*(1). The *ACF* of such a process would appear as



A significant spike at lag 1 identified the *regular MA*(1), an another significant spike at lag 4 identified the *seasonal MA*(1).

In this example we need to *regular difference at lag 1* to remove the trend, and we need to *seasonal difference at lag 4* to remove the quarterly seasonality.

The *regular differencing at lag 1* and the *regular MA*(1) we denote by

ARIMA(0,1,1)

However, we also wish to include the *seasonal difference at lag 4* and the *seasonal MA*(1). We extend the above notation to

$$\underset{\text{regular factor}\uparrow}{ARIMA(0,1,1)} \times \underset{\text{seasonal}\uparrow\text{factor}}{(0,1,1)_4}$$

In backshift notation,

$$ARIMA(0,1,1) \times (0,1,1)_4$$

$$(1 - \mathrm{B})(1 - \mathrm{B}^4)\mathrm{Y}_t = (1 - \theta_1 \mathrm{B})(1 - \Theta_1 \mathrm{B}^4)\epsilon_t$$

We generalize this notation to

Seasonal ARIMA models

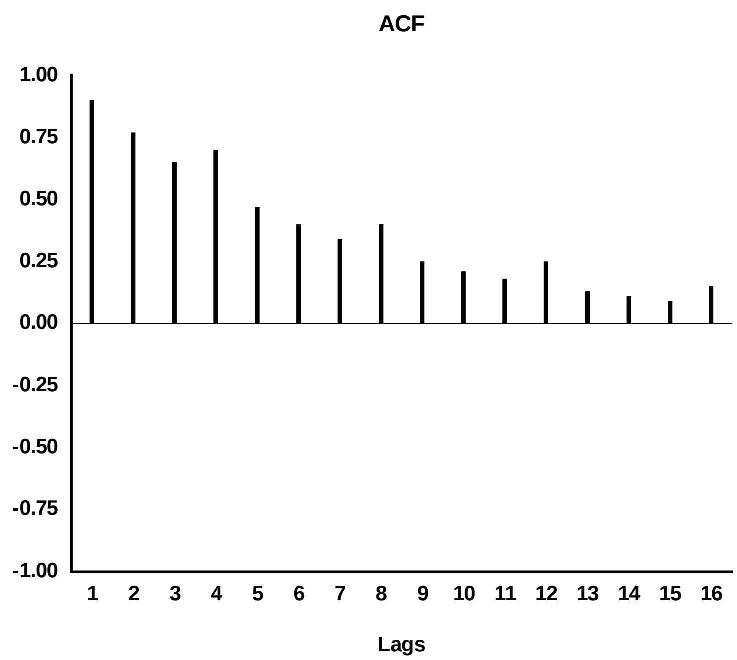
$$\boldsymbol{ARIMA(p,d,q) \times (P,D,Q)_s} \qquad \textbf{11.16}$$

where $(P,D,Q,)_s$ now denotes the *seasonal P, D,* and *Q* orders of *seasonal P, AR, seasonal D, differencing,* and *seasonal Q, MA* factors.

We shall use uppercase Greek letters will be used to denote the seasonal parameters.

Example 12

Consider the *ACF* of the time series in which it has regular first differencing and quarterly first differencing.



The *ACF* identifies a *regular AR*(1) process and a *seasonal AR*(1) process, so the suggested *seasonal ARIMA* model is

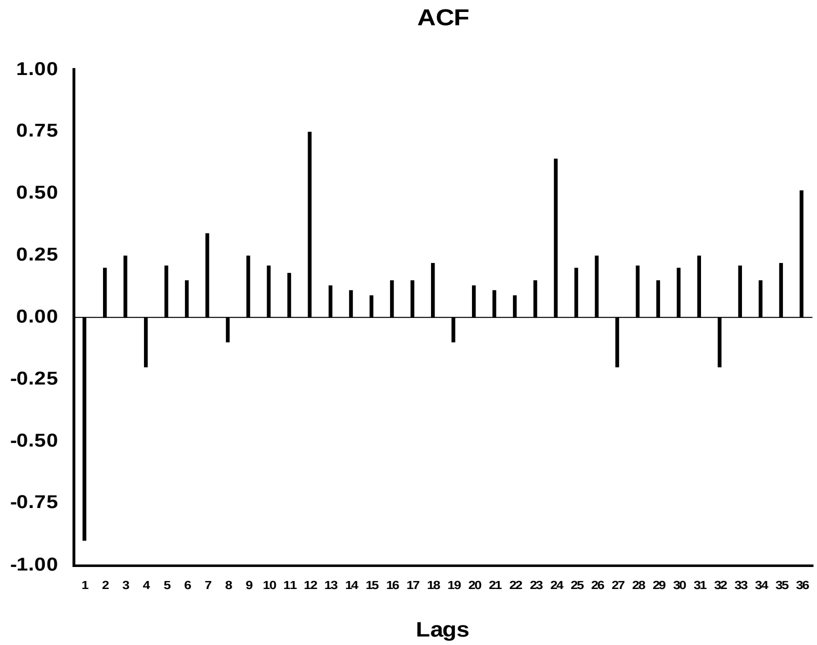
$$ARIMA(1,1,0) \times (1,1,0)_4$$

In backshift notation

$$(1 - \phi_1 B)(1 - B)(1 - \Phi_1 B^4)(1 - B^4)Y_t = \epsilon_t$$

Example 13

Suppose, after *regular first differencing* and *monthly seasonal first differencing,* a time series process had the following *ACF*.



The *ACF* reveals a *regular MA*(1) process and a *seasonal AR*(1) process. The suggested *ARIMA* model is

$$ARIMA(0,1,1) \times (1,1,0)_{12}$$

In backshift notation

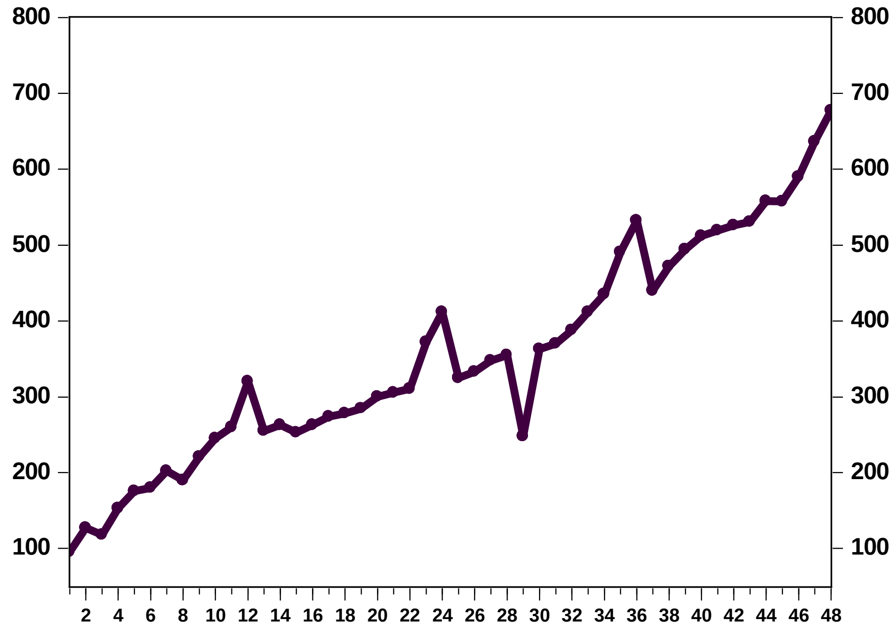
$$(1 - \Phi_1 B^{12})(1 - B)(1 - B^{12})Y_t = (1 - \theta_1 B)\epsilon_t$$

Let us consider the orginal Sales data of Chapter 6.

Table 11.14

Four Years of Time Series Data

Period	Year 1	Period	Year 2	Period	Year 3	Period	Year 4
1	95.5	13	255.0	25	325.0	37	440.0
2	127.0	14	263.0	26	333.0	38	472.0
3	118.0	15	253.0	27	348.0	39	494.5
4	153.0	16	263.0	28	355.0	40	512.0
5	175.5	17	274.0	29	248.0	41	519.0
6	180.0	18	278.0	30	363.0	42	526.0
7	202.0	19	284.5	31	370.0	43	530.5
8	190.0	20	300.0	32	388.0	44	558.0
9	220.5	21	305.5	33	412.0	45	557.5
10	245.0	22	310.5	34	435.0	46	590.0
11	260.0	23	372.0	35	490.5	47	636.4
12	320.0	24	412.0	36	532.5	48	677.4



This data clearly has trend and seasonality. To remove the trend we use first differencing at lag 1 and to remove seasonality we use seasonal differencing at lag 12. The original series is Y_t, the transformed series is

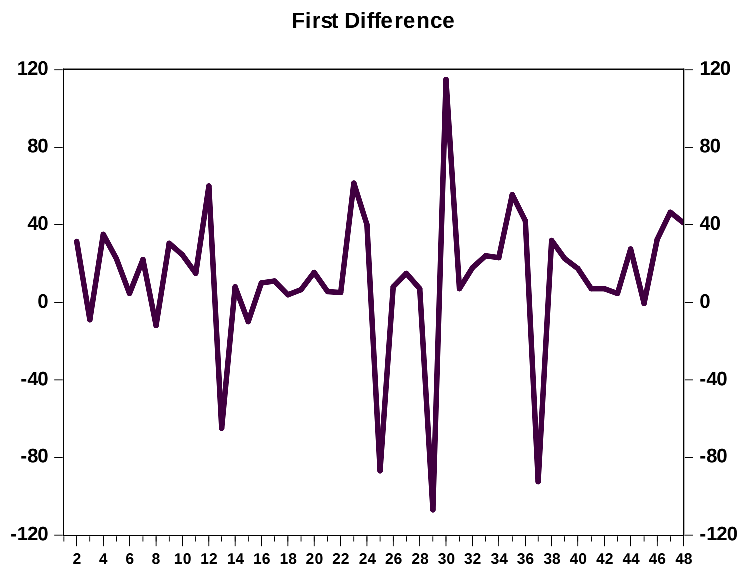
$$Z_t = (1 - B)(1 - B^{12})Y_t$$

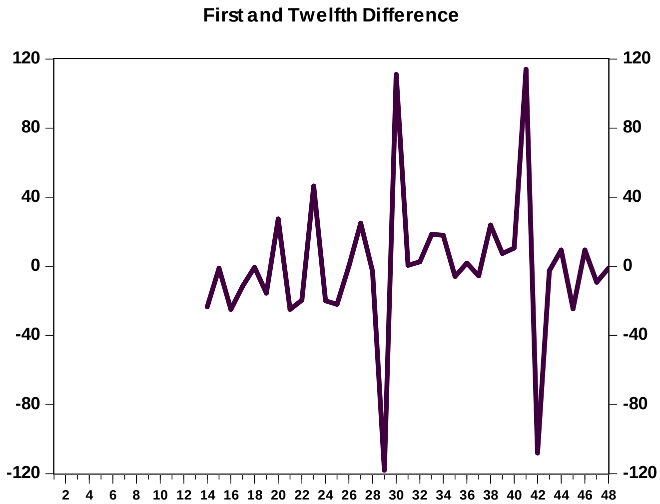
Table 11.15

Period	Y_t	*First Difference*	Z_t *Twelfth Difference of the First Difference*
1	95.5	*	*
2	127.0	31.5	*
3	118.0	-9.0	*
4	153.0	35.0	*
5	175.5	22.5	*
6	180.0	4.5	*
7	202.0	22.0	*
8	190.0	-12.0	*
9	220.5	30.5	*
10	245.0	24.5	*
11	260.0	15.0	*
12	320.0	60.0	*
13	255.0	-65.0	*
14	263.0	8.0	-23.5
15	253.0	-10.0	-1.0
16	263.0	10.0	-25.0
⋮	⋮	⋮	⋮

Notice that Z_t has 13 fewer observations than the original series. The regular difference lag 1 has one few observation than the original series, and the seasonal difference lag 12 has 12 fewer observations than the first difference lag 1, so a total of 13 less observations than the original series.

Figure 11-21 is a plot of the first difference of the data, the seasonality of the data is still evident. Figure 11-22 is a plot of the twelth difference of the first difference. It is a stationary series.





Note that the order of differencing is unimportant:

Either $$Z_t = (1 - B)(1 - B^{12})Y_t$$
or, $$Z_t = (1 - B^{12})(1 - B)Y_t$$

will achieve the same result, Z_t.

We now consider the *ACF* and the *PACF* of the series Z_t

Sample: 1 48
Included observations: 35

	AC	PAC	Q-Stat	Prob
1	-0.423	-0.423	6.8063	0.009
2	-0.095	-0.333	7.1611	0.028
3	0.204	0.012	8.8382	0.032
4	-0.067	0.029	9.0272	0.060
5	0.009	0.070	9.0307	0.108
6	-0.141	-0.182	9.9233	0.128
7	0.161	0.011	11.116	0.134
8	0.079	0.167	11.413	0.179
9	-0.186	0.020	13.136	0.157
10	0.097	0.003	13.624	0.191
11	0.172	0.201	15.223	0.173
12	-0.416	-0.323	24.973	0.015
13	0.220	-0.059	27.828	0.010
14	0.020	-0.032	27.852	0.015
15	-0.060	0.044	28.084	0.021
16	-0.026	-0.070	28.129	0.031

Recall that spikes greater in absolute value than $\frac{2}{\sqrt{n}}$ are usually significant. Since we have $n - 13 = 48 - 13 = 35$ observations

$$\frac{2}{\sqrt{35}} = .338$$

We have significant spikes are lag 1 and lag 12. (Notice that we requested the *ACF* to compute to lag 25, so we could check for a significant spike at lag 24.)

This *ACF* identifies a *regular MA*(1) and a *seasonal MA*(1) as the likely *ARIMA* model.

ARIMA(0,1,1) $\times$ $(0,1,1)_{12}$

Thus we request of the software for such an estimated model of the original series, Y_t. The estimated model from the computer output estimates the *regular MA*(1) parameter at $\hat{\theta}_1 = .556$ and the *seasonal SMA*(1) parameter estimate at $\hat{\Theta}_1 = .674$.

Hence,

ARIMA(0,1,1)X$(0,1,1)_{12}$

In backshift notation:

$$(1 - B)(1 - B^{12})Y_t = (1 - \hat{\theta}_1 B)(1 - \hat{\Theta}_1 B^{12})\epsilon_t$$

$$(1 - B)(1 - B^{12})Y_t = (1 - .556B)(1 - .674B^{12})\epsilon_t$$

In expanded notation:

$$Y_t = Y_{t-1} + Y_{t-12} - Y_{t-13} - .556\epsilon_{t-1} - .674\epsilon_{t-12} + .375\epsilon_{t-13} + \epsilon_t$$

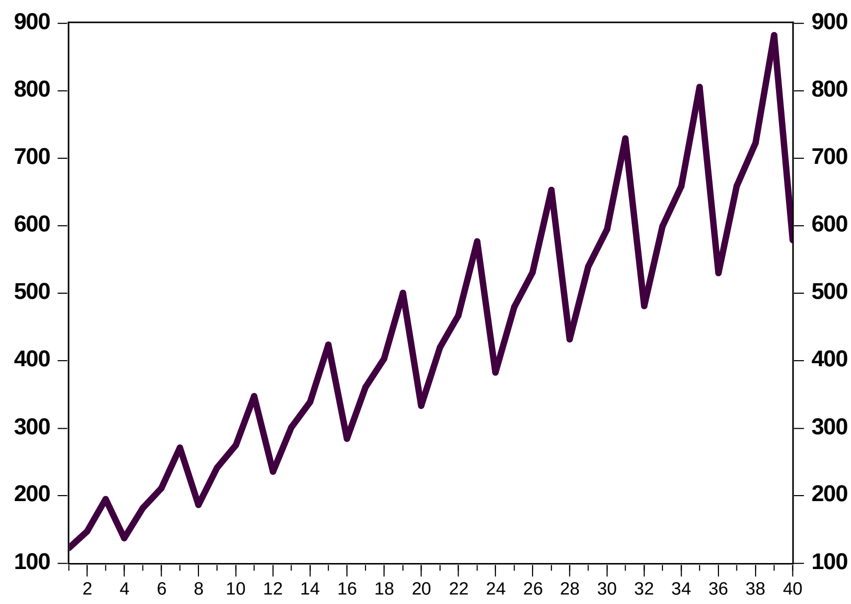
Voila!

Example 2

Let us consider the Quarterly Sales data of Chapter 6; (Table 4. , page).

Table 11.16

1	122.15	11	347.43	21	420.18	31	729.29
2	147.04	12	235.41	22	466.81	32	480.67
3	194.69	13	300.97	23	576.54	33	599.00
4	137.30	14	338.90	24	382.56	34	658.67
5	181.75	15	423.80	25	479.79	35	805.66
6	210.99	16	284.46	26	530.76	36	529.72
7	271.06	17	360.58	27	652.92	37	658.61
8	186.36	18	402.86	28	431.62	38	722.62
9	241.36	19	500.17	29	539.40	39	882.03
10	274.95	20	333.51	30	594.72	40	578.77



This data clearly has trend and seasonality. To remove the trend we use first differencing at lag 1 and to remove seasonality we use seasonal differencing at lag 4. The original series is Y_t, the transformed series is

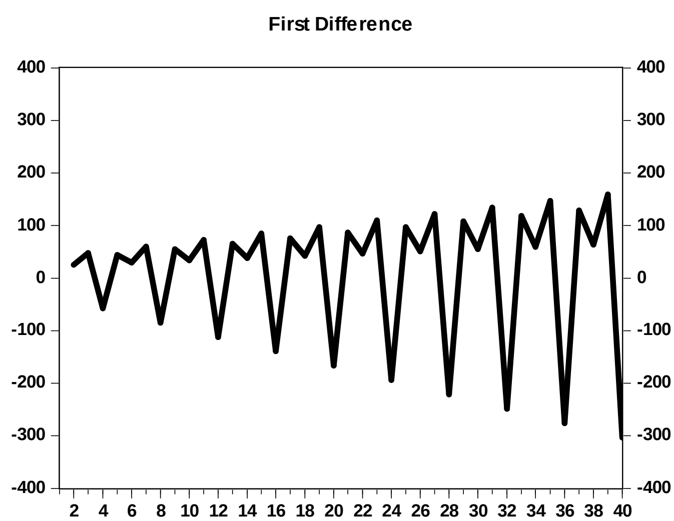
$$Z_t = (1 - B)(1 - B^4)Y_t$$

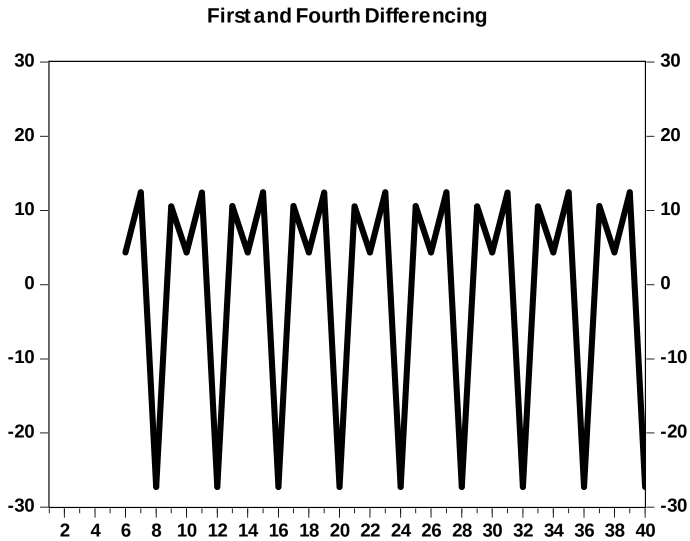
Table 11.17

Period	Y_t	*First Difference*	Z_t *Fourth Difference of the First Difference*
1	122.15	*	*
2	47.04	24.89	*
3	194.69	47.64	*
4	137.30	-57.38	*
5	181.75	44.45	*
6	210.99	29.24	4.35
7	271.06	60.06	12.42
8	186.36	-84.70	-27.32
9	241.36	55.00	10.56
10	274.95	33.59	4.35
11	347.43	72.48	12.42
12	235.41	-112.02	-27.32
⋮	⋮	⋮	⋮

Notice that Z_t has 5 fewer observations than the original series. The regular difference lag 1 has one few observation than the original series, and the seasonal difference lag 4 has 4 fewer observations than the first difference lag 1, so a total of 5 less observations than the original series.

Figure 11-26 is a plot of the first difference of the data, the seasonality of the data is still evident. Figure 11-27 is a plot of the fourth difference of the first difference. It is a stationary series.





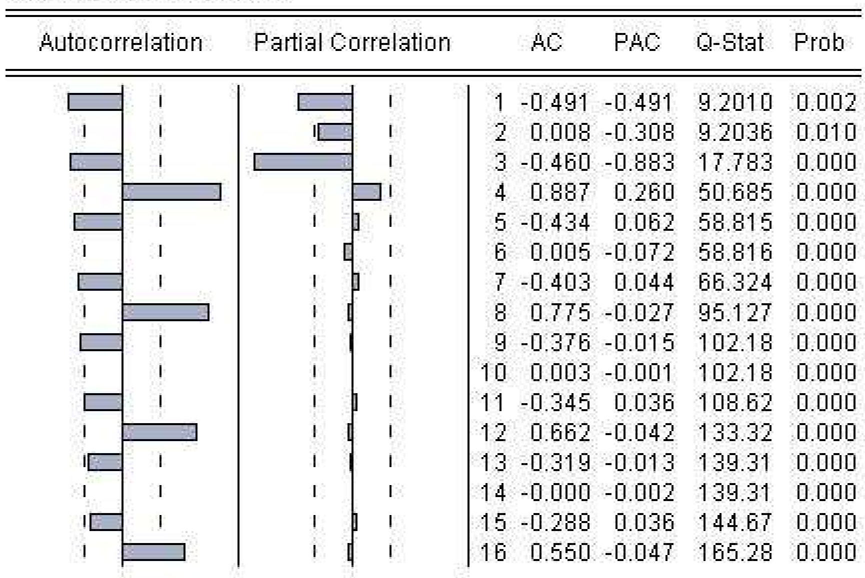
Note that the order of differencing is unimportant:

Either $Z_t = (1 - B)(1 - B^4)Y_t$

or, $Z_t = (1 - B^4)(1 - B)Y_t$

will achieve the same result, Z_t.

We now consider the *ACF* and the *PACF* of the series Z_t .



Sample: 1 40

Included observations: 35

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.491	-0.491	9.2010	0.002
		2	0.008	-0.308	9.2036	0.010
		3	-0.460	-0.883	17.783	0.000
		4	0.887	0.260	50.685	0.000
		5	-0.434	0.062	58.815	0.000
		6	0.005	-0.072	58.816	0.000
		7	-0.403	0.044	66.324	0.000
		8	0.775	-0.027	95.127	0.000
		9	-0.376	-0.015	102.18	0.000
		10	0.003	-0.001	102.18	0.000
		11	-0.345	0.036	108.62	0.000
		12	0.662	-0.042	133.32	0.000
		13	-0.319	-0.013	139.31	0.000
		14	-0.000	-0.002	139.31	0.000
		15	-0.288	0.036	144.67	0.000
		16	0.550	-0.047	165.28	0.000

Recall that spikes greater in absolute value than $\frac{2}{\sqrt{n}}$ are usually significant. Since we have $n - 13 = 40 - 5 = 35$ observations

$$\frac{2}{\sqrt{35}} = .338$$

We have an *ACF* with decaying spikes, indicating *AR* processes.

In this example, the key to the identification is with the *PACF*. The *PACF* has significant spikes at lags 1 through 4. The first two spikes suggest a regular *AR*(2) process, and the spike at lag 4 suggests a *seasonal MA*(1) process. The major spike at lag 3 is probably the result of the combination of the regular *AR*(2) and the seasonal *MA*(1).

A *regular AR*(2) and a *seasonal MA*(1) is written

$$ARIMA(2,1,0) \times (0,1,1)_4$$

Thus we request of the computer software for such an estimated model of the original series, Y_t

The estimated model from the computer output estimates the *regular AR*(2) parameters at $\hat{\phi}_1 = -0.717$, $\hat{\phi}_2 = -0.359$ and the *seasonal SMA*(1) parameter estimate at $\hat{\Theta}_1 = -0.873$.

Hence,

$$ARIMA(2,1,0)\mathrm{X}(0,1,1)_4$$

In backshift notation:

$$(1 - \hat{\phi}_1\mathrm{B} - \hat{\phi}_2\mathrm{B}^2)(1 - \mathrm{B})(1 - \mathrm{B}^4)\mathrm{Y}_t = (1 - \hat{\Theta}_1\mathrm{B}^4)\epsilon_t$$

$$(1 - (-0.717)B - (-0.359)B^2)(1 - B)(1 - B^4)Y_t = (1 - (-0.873)\mathrm{B}^4)\epsilon_t$$

In expanded notation:

$$Y_t = Y_{t-1} + .717Y_{t-2} + .359Y_{t-3} + Y_{t-4} - .283Y_{t-5} - .359Y_{t-6} - .359Y_{t-7} + .873\epsilon_{t-4} + \epsilon_t$$

Voila!

PROBLEMS AND QUESTIONS

Stage 2 Identification of the Appropriate Model

Determine the *ACF* of the following *MA(1)* processes.

11.1 $Y_t = 11 + .60\epsilon_{t-1} + \epsilon_t$

11.2 $Y_t = 102 - .65\epsilon_{t-1} + \epsilon_t$

Write down the equation and then determine the *ACF* of the following *MA(1)* processes.

11.3 *MA(1)* process with $\mu = 38$ and $\theta_1 = .34$

11.4 *MA(1)* process with $\mu = 253$ and $\theta_1 =$ -.76

Determine the *ACF* of the following *MA(2)* processes.

11.5 $Y_t = 41 + .72\epsilon_{t-1} + .36\epsilon_{t-2} + \epsilon_t$

11.6 $Y_t = 402 - .65\epsilon_{t-1} + .45\epsilon_{t-2} + \epsilon_t$

Write down the equation and then determine the *ACF* of the following *MA(2)* processes.

11.7 *MA(2)* process with $\mu = 45$, $\theta_1 = .64$, $\theta_2 = .38$

11.8 *MA(2)* process with $\mu = 28$, $\theta_1 =$ -.76, $\theta_2 = .46$

Determine the first 5 lags of the *ACF* of the following *AR(1)* processes.

11.9 $Y_t = 11 + .60\mathrm{Y}_{t-1} + \epsilon_t$

11.10 $Y_t = 102 - .65\mathrm{Y}_{t-1} + \epsilon_t$

Write down the equation and then determine the first 5 lags of the *ACF* of the following *AR(1)* processes.

11.11 *AR(1)* process with $\delta = 38$ and $\phi_1 = .34$

11.12 *AR(1)* process with $\delta = 253$ and $\phi_1 =$ -.76

Determine the first 5 lags of the *ACF* of the following *AR(2)* processes.

11.13 $Y_t = 41 + .72\mathrm{Y}_{t-1} + .36\mathrm{Y}_{t-2} + \epsilon_t$

11.14 $Y_t = 402 - .65\mathrm{Y}_{t-1} + .45\mathrm{Y}_{t-2} + \epsilon_t$

Write down the equation and then determine the first 5 lags of the *ACF* of the following *AR(2)* processes.

11.15 *AR(2)* process with $\delta = 45$, $\phi_1 = .64$, $\phi_2 = .38$

11.16 *AR(2)* process with $\delta = 28$, $\phi_1 =$ -.76, $\phi_2 = .46$

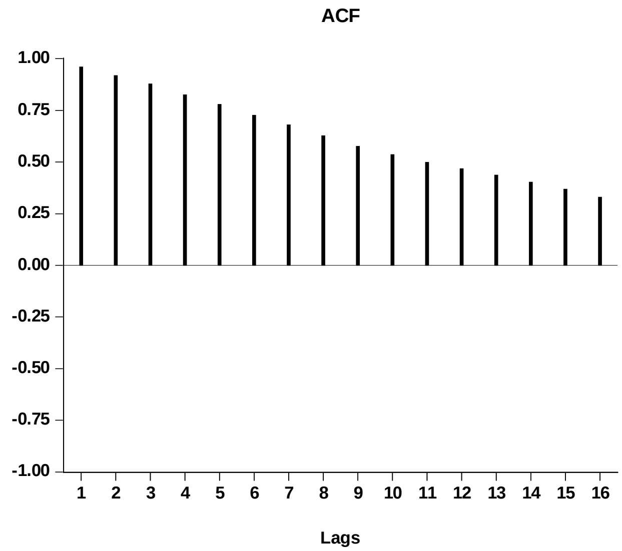
Determine the first 5 lags of the *ACF* of the following:

11.17 $Y_t = 41 + .72\epsilon_{t-1} + .36\epsilon_{t-2} - .26\epsilon_{t-3} + \epsilon_t$

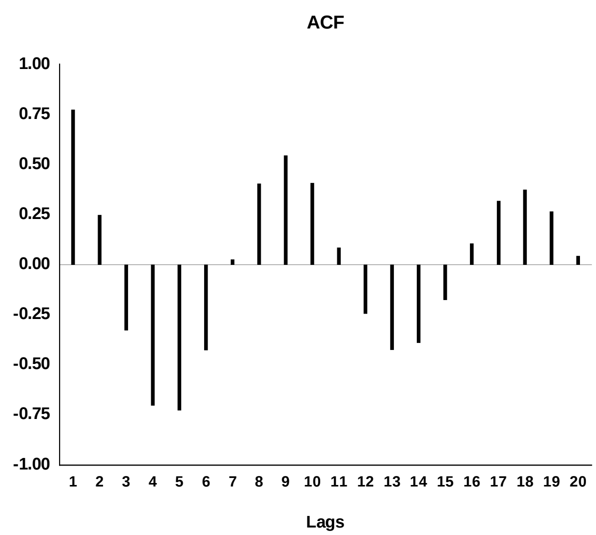
11.18 $Y_t = 402 + .55Y_{t-1} + .25Y_{t-2} + .20Y_{t-3} + \epsilon_t$

Determine the processes suggested by the following *ACF*'s and *PACF*'s

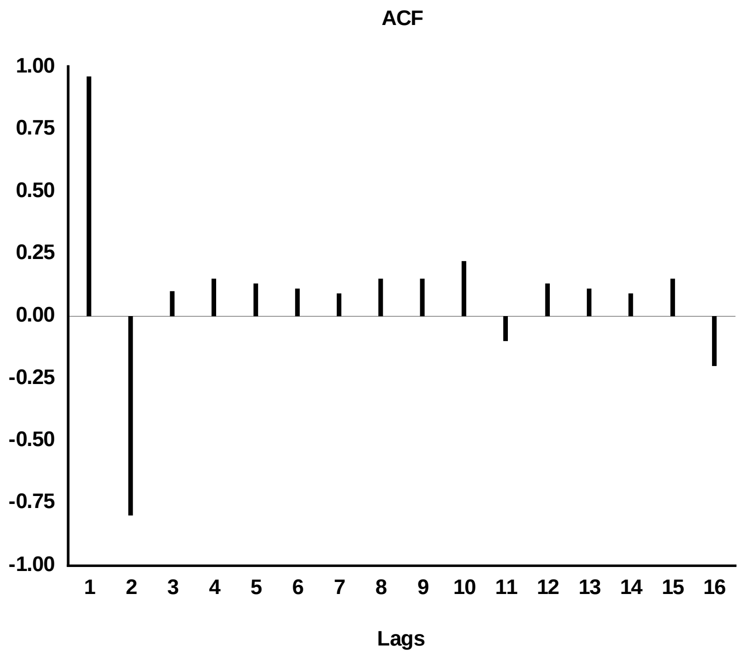
11.19



11.20



11.21


ACF
1.00
0.75
0.50
0.25
0.00
-0.25
-0.50
-0.75
-1.00
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
Lags

Appendix A

Formula Summary

***Autocorrelation at lag 1 of an MA*(1) process**

$$\rho_1 = \frac{-\theta_1}{(1 + \theta_1^2)}$$

***Autocorrelation at lag 1 of an MA*(2) process**

$$\rho_1 = \frac{-\theta_1 + \theta_1\theta_2}{\left(1 + \theta_1^2 + \theta_2^2\right)}$$

***Autocorrelation at lag 2 of an MA*(2) process**

$$\rho_2 = \frac{-\theta_2}{\left(1 + \theta_1^2 + \theta_2^2\right)}$$

***ρ_k for an MA*(q) process**

$$\rho_k = \frac{-\theta_k + \theta_1\theta_{k+1} + \ldots + \theta_{q-k}\theta_q}{1 + \theta_1^2 + \theta_2^2 + \ldots + \theta_q^2}, \qquad \text{for } k = 1, 2, \ldots q$$

$$\rho_k = 0, \qquad \text{for } k > q$$

The Autocorrelation at lag 1 of an AR*(1) *process

$$\rho_1 = \phi_1$$

The Autocorrelation at lag 2 for an AR*(1) *process

$$\rho_2 = \phi_1^2$$

The Autocorrelation at lag 3 for an AR*(1) *process

$$\rho_3 = \phi_1^3$$

$$\vdots$$

In general,

The Autocorrelation at lag k for an AR*(1) *process

$$\rho_k = \phi_1^k$$

The Autocorrelation at lag 1 of an AR*(2) *process

$$\rho_1 = \frac{\phi_1}{1 - \phi_2}$$

The Autocorrelation at lag 2 for an AR*(2) *process

$$\rho_2 = \phi_1\rho_1 + \phi_2\rho_0$$

The Autocorrelation at lag 3 for an AR*(2) *process

$$\rho_3 = \phi_1\rho_2 + \phi_2\rho_1$$

The Autocorrelation at lag 4 for an AR*(2) *process

$$\boxed{\rho_4 = \phi_1\rho_3 + \phi_2\rho_2}$$

$\vdots$

The Autocorrelation at lag k for an AR(2) process

$$\boxed{\rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-2}}$$

$\vdots$

Autocorrelation at lag k, ρ_k, of an AR(p) process

$$\boxed{\rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-2} + \ldots + \phi_p\rho_{k-p}}$$

$p = 1$	$\rho_k = \phi_1^k$
$p = 2$	$\rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-2}$
$p = 3$	$\rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-2} + \phi_3\rho_{k-3}$
$\vdots$	$\vdots$
$p = p$	$\rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-2} + \phi_3\rho_{k-3} + \ldots + \phi_p\rho_{k-p}$
$\vdots$	$\vdots$

Partial Autocorrelations

AR(1) process

Lag 1	$\rho_{11} = \rho_1 = \phi_1$
Lag 2	$\rho_{22} = 0$
Lag 3	$\rho_{33} = 0$
$\vdots$	$\vdots$
Lag k	$\rho_{kk} = 0$

$\rho_{kk} = 0$ for $k = 2, 3, \ldots$

AR(2) process

Lag 1	$\rho_{11} = \rho_1$
Lag 2	$\rho_{22} = \frac{\rho_2 - \rho_1^2}{2}$
Lag 3	$\rho_{33} = 0$
$\vdots$	$\vdots$
Lag k	$\rho_{kk} = 0$

$\rho_{kk} = 0$ for $k = 3, 4, \ldots$

In general, for an *AR*(p) process

$\rho_{kk} = 0$ for $k = p + 1, p + 2, \ldots$

Seasonal ARIMA models

$$\boxed{\textbf{\textit{ARIMA}(p,d,q)} \times \textbf{(P,D,Q)}_s}$$

where (P,D,Q,)$_s$ now denotes the *seasonal P, D,* and *Q* orders of *seasonal P, AR, seasonal D, differencing,* and *seasonal Q, MA* factors.

Appendix B

Proofs and Derivations

We begin this review with the very useful statistical device called the *expectation operator*. This operator and its application to the error term are important for our understanding of the *ARIMA* models.

The Expectation Operator

Recall from Chapter 2, that we denoted
$E(Y_t)$ as the *expected value of* Y_t, or the mean value of Y_t.

$$E(Y_t) = \mu_Y$$

Since, $E(Y_t) = \mu_Y = \frac{\sum Y_t}{N}$,

mneumonically, we can think of the expectation operator as

$$\boxed{\boldsymbol{E(\quad) = \frac{\sum(\quad)}{N}}}$$

We can use the expectation operator to define the variance since by definition the variance is given by

$$\sigma_Y^2 = \frac{\sum(Y_t - \mu_Y)^2}{N} \qquad \text{B.1}$$

So,

$$\sigma_Y^2 = E(Y_t - \mu_Y)^2 \qquad \text{B.2}$$

and since $E(Y_t) = \mu_Y$,

$$\sigma_{Y_t}^2 = E[Y_t - E(Y_t)]^2 \qquad \text{B.3}$$

The covariance is also written using the expectation operator

$$Cov(X_t, Y_t) = \frac{\sum(X_t - \mu_X)(Y_t - \mu_Y)}{N}$$

Using the expectation notation of equation (B.4) we have

$$Cov(X_t, Y_t) = E[(X_t - E(X_t))(Y_t - E(Y_t))] \qquad \text{B.4}$$

There are two facts about the expectation operator that we need, and we list here without proof.

Fact Number 1.

If X and Y are random variables, then

$$E(X + Y) = E(X) + E(Y) \qquad \text{B.5}$$

$$E(X - Y) = E(X) - E(Y) \qquad \text{B.6}$$

Fact Number 2.

If c is a constant, then

$$E(cX) = cE(X) \qquad \text{B.7}$$

$$E(c) = c \qquad \text{B.8}$$

Since much of the following discussion is involved with the random disturbance term, ϵ_t, we include it here with the discussion of the expectation operator.

The expectation notation of the mean of ϵ

$\boldsymbol{E(\epsilon_t) = 0}$	**B.9**

For the variance of ϵ_t we use the format of equation (B.3)

variance $\sigma_\epsilon^2 = E(\epsilon_t - E(\epsilon_t))^2$

And since $E(\epsilon_t) = 0$, the variance of ϵ_t reduces to

The expectation notation of the variance of ϵ

$\boldsymbol{\sigma_\epsilon^2 = E(\epsilon_t^2)}$	**B.10**

Recall that the ϵ_t's are assumed to be independent random errors so that
$Cov(\epsilon_t, \epsilon_{t-1}) = 0.$

$$Cov(\epsilon_t, \epsilon_{t-1}) = E[(\epsilon_t - E(\epsilon_t))(\epsilon_{t-1} - E(\epsilon_{t\text{-}1}))] = 0.$$

Again, since $E(\epsilon_t) = 0$ and $E(\epsilon_{t-1}) = 0$, we have

$$Cov(\epsilon_t, \epsilon_{t-1}) = E[(\epsilon_t - 0)(\epsilon_{t-1} - 0)] = 0.$$

$$= E[(\epsilon_t)(\epsilon_{t-1})] = 0.$$

$$= E[\epsilon_t \epsilon_{t-1}] = 0.$$

Notice, that in the case of the covariance between ϵ_t and ϵ_{t-1}, this is written as the expectation of their product, and it always equals zero.

The expectation notation of the covariance between different ϵ's

$\boldsymbol{Cov(\epsilon_t, \epsilon_{t-1}) = E[\epsilon_t \epsilon_{t-1}] = 0}$	**B.11**

Equations (B.9), (B.10), and (B.11) will be used repeatedly throughout this chapter, so we summarize them here.

Properties of ϵ_t

(1) mean zero:	$E(\epsilon_t) = 0$	B.9
(2) variance σ_ϵ^2:	$E(\epsilon_t^2) = \sigma_\epsilon^2$	B.10
(3) uncorrelated ϵ's:	$E(\epsilon_t \epsilon_{t-1}) = 0$	B.11

The *ACF* of an *MA*(1) process

In all of these processes we need the mean and the variance. So we begin with general mean and variance of an *MA*(1) process.

The general *MA*(1) model is written

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

The Mean of an *MA*(1) process.

Using the expectation operator we determine the mean of the process.

$$E(Y_t) \;=\; E(\mu_Y \;-\; \theta_1\epsilon_{t-1} \;+\; \epsilon_t) \qquad \text{B.12}$$

$$\;=\; E(\mu_Y) \;-\; E(\theta_1\epsilon_{t-1}) \;+\; E(\epsilon_t) \qquad \text{B.13}$$

We have applied Fact Number 1 regarding the expectation operator, equations B.5 and B.6, and since μ_Y and θ_1 are constants, we apply Fact Number 2 and re-write equation B.14 as

$$E(Y_t) \;=\; \mu_Y \;-\; \theta_1 E(\epsilon_{t-1}) \;+\; E(\epsilon_t) \qquad \text{B.14}$$

Lastly we apply the property of ϵ_t under expectation, i.e.

$E(\epsilon_t) \;=\; 0$ and $\qquad E(\epsilon_{t-1}) \;=\; 0.$

Thus,

$$E(Y_t) \;=\; \mu_Y$$

No surprise, we knew that the mean of an *MA*(1) process is μ_Y. And we have provided an easy demonstration of using the expectation operator.

Our interest is in determining the *ACF*, the autocorrelation at various lags, so that the mean of the *MA*(1) is unimportant. Indeed, without loss of generality we can assume that $\mu_Y \;=\; 0$. We can make this assumption, because by setting the mean equal to zero we are changing only the level of the process. The change of level has no effect on the autocorrelation at various lags.

By setting $\mu_Y \;=\; 0$, we have

$$E(Y_t) \;=\; 0.$$

and thus

$$MA(1) \qquad Y_t \;=\; -\;\theta_1\epsilon_{t-1} \;+\; \epsilon_t \qquad \text{B.15}$$

The Variance of an *MA*(1) process.

By definition of variance (equation B.3),

$$\sigma^2_{Y_t} \;=\; E[Y_t \;-\; E(Y_t)]^2$$

Since, $E(Y_t) \;=\; 0$, $\sigma^2_{Y_t} \;=\; E[Y_t]^2$ B.16

Now, because $Y_t \;=\; -\theta_1\epsilon_{t-1} \;+\; \epsilon_t$ (equation B.17) becomes

$$\sigma^2_{Y_t} \;=\; E[-\theta_1\epsilon_{t-1} \;+\; \epsilon_t]^2 \qquad \text{B.17}$$

We expand equation (B.18)[1]

$$\;=\; E[(-\theta_1\epsilon_{t-1} \;+\; \epsilon_t)(-\theta_1\epsilon_{t-1} \;+\; \epsilon_t)] \qquad \text{B.}\mathit{18}$$

$$\;=\; E[\theta_1^2\epsilon_{t-1}^2 \;-\; 2\theta_1\epsilon_{t-1}\epsilon_t \;+\; \epsilon_t^2] \qquad \text{B.19}$$

1

We then re-write equation B.20 using Facts 1 and 2 of the expectation operator.

$$= \theta_1^2 \mathrm{E}[\epsilon_{t-1}^2] - 2\theta_1 \mathrm{E}[\epsilon_{t-1}\epsilon_t] + \mathrm{E}[\epsilon_t^2] \qquad \text{B.20}$$

Now, $\mathrm{E}[\epsilon_{t-1}^2] = \sigma_\epsilon^2$, and $\mathrm{E}[\epsilon_t^2] = \sigma_\epsilon^2$, by equation B.10

and $\mathrm{E}[\epsilon_{t-1}\epsilon_t] = 0$, by equation B.11.

So, equation B.21 reduces to

$$\sigma_{Y_t}^2 = \theta_1^2 \sigma_\epsilon^2 + \sigma_\epsilon^2 \qquad \text{B.21}$$

$$\sigma_{Y_t}^2 = (1 + \theta_1^2)\sigma_\epsilon^2 \qquad \text{B.22}$$

The Variance of an MA(1) process

$$\boldsymbol{\sigma_{Y_t}^2 = (1 + \theta_1^2)\sigma_\epsilon^2} \qquad \textbf{B.23}$$

The variance is equivalent to γ_0, the autocovariance at lag 0.

The Autocovariance at lag 0, γ_0, of an MA(1) process

$$\boldsymbol{\gamma_0 = (1 + \theta_1^2)\sigma_\epsilon^2} \qquad \textbf{B.24}$$

Autocorrelation at lag 1, γ_1, of an *MA*(1) process

By definition, $\gamma_1 = Cov(Y_t, Y_{t-1})$ B.25

$$\gamma_1 = E[(Y_t - E(Y_t))(Y_{t-1} - E(Y_{t-1}))] \qquad \text{B.26}$$

Since $E(Y_t) = E(Y_{t-1}) = 0$ equation B.25 reduces to[2]

$$\gamma_1 = E[Y_t Y_{t-1}] \qquad \text{B.27}$$

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

then $Y_{t-1} = -\theta_1\epsilon_{t-2} + \epsilon_{t-1}$

Substituting the expressions for Y_t and Y_{t-1} into equation B.26

$$\gamma_1 = E[(-\theta_1\epsilon_{t-1} + \epsilon_t)(-\theta_1\epsilon_{t-2} + \epsilon_{t-1})] \qquad \text{B.28}$$

Multiplying out the binomial and taking expectations

$$= \mathrm{E}[\theta_1^2\epsilon_{t-1}\epsilon_{t-2} - \theta_1\epsilon_t\epsilon_{t-2} - \theta_1\epsilon_{t-1}^2 + \epsilon_t\epsilon_{t-1}] \qquad \text{B.29}$$

$$= \theta_1^2\mathrm{E}[\epsilon_{t-1}\epsilon_{t-2}] - \theta_1\mathrm{E}[\epsilon_t\epsilon_{t-2}] - \theta_1\mathrm{E}[\epsilon_{t-1}^2] + \mathrm{E}[\epsilon_t\epsilon_{t-1}] \qquad \text{B.30}$$

Invoking property B.11, that $\mathrm{E}[\epsilon_{t-1}\epsilon_{t-2}] = \mathrm{E}[\epsilon_t\epsilon_{t-2}] = \mathrm{E}[\epsilon_t\epsilon_{t-1}] = 0$ and that $\mathrm{E}[\epsilon_{t-1}^2] = \sigma_\epsilon^2$, equation B.29 reduces to:

The Autocovariance at lag 1, γ_1, for an MA(1) process

$$\boldsymbol{\gamma_1 = -\theta_1\sigma_\epsilon^2} \qquad \textbf{B.31}$$

We now have

$$\gamma_0 = (1 + \theta_1^2)\sigma_\epsilon^2$$

2

and $\gamma_1 = \text{-}\theta_1 \sigma_\epsilon^2$

Thus, by defintion of *autocorrelation at lag 1*, ρ_1

$$\rho_1 = \frac{\textit{autocovariance at lag 1}}{\textit{variance}}$$

$$\rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{\text{-}\theta_1 \sigma_\epsilon^2}{(1+\theta_1^2)\sigma_\epsilon^2} = \frac{\text{-}\theta_1}{(1+\theta_1^2)}$$

autocorrelation at lag 1, ρ_1, for an MA(1) process

$$\boxed{\rho_1 = \frac{\text{-}\theta_1}{(1+\theta_1^2)}}$$

Autocorrelation at lag 2, γ_2, for an *MA*(1) process

By definition the autocovariance at lag 2 is

$$\gamma_2 = Cov(Y_t, Y_{t-2}) \qquad \textit{B.32}$$

$$\gamma_2 = E[(Y_t - E(Y_t))(Y_{t-2} - E(Y_{t-2}))] \qquad \text{B.33}$$

Since $E(Y_t) = E(Y_{t-2}) = 0$ equation B.32 reduces to

$$\gamma_2 = E[Y_t Y_{t-2}] \qquad \text{B.34}$$

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

then $Y_{t-2} = -\theta_1 \epsilon_{t-3} + \epsilon_{t-2}$

Substituting the expressions for Y_t and Y_{t-2} into equation B.33

$$\gamma_2 = E[(-\theta_1 \epsilon_{t-1} + \epsilon_t)(-\theta_1 \epsilon_{t-3} + \epsilon_{t-2})] \qquad \text{B.35}$$

Multiplying out the binomial and taking expectations

$$= E[\theta_1^2 \epsilon_{t-1}\epsilon_{t-3} - \theta_1 \epsilon_t \epsilon_{t-3} - \theta_1 \epsilon_{t-1}\epsilon_{t-2} + \epsilon_t \epsilon_{t-2}] \qquad \textit{B.36}$$

$$\gamma_2 = \theta_1^2 E[\epsilon_{t-1}\epsilon_{t-3}] - \theta_1 E[\epsilon_t \epsilon_{t-3}] - \theta_1 E[\epsilon_{t-1}\epsilon_{t-2}] + E[\epsilon_t \epsilon_{t-2}] \qquad \textit{B.37}$$

Invoking property B.11, that $\text{E}[\epsilon_{t-1}\epsilon_{t-3}] = \text{E}[\epsilon_t \epsilon_{t-3}] = \text{E}[\epsilon_{t-1}\epsilon_{t-2}] = \text{E}[\epsilon_t \epsilon_{t-1}] = 0$ means that equation B.36 goes to zero under expectation. Hence,

The Autocovariance at lag 2, γ_2, for an MA(1) process

$$\boxed{\gamma_2 = 0 \qquad \textbf{B.38}}$$

We now have

$$\gamma_0 = (1 + \theta_1^2)\sigma_\epsilon^2$$

and $\gamma_2 = 0$

Thus, by defintion of *autocorrelation at lag 2*, ρ_2

$$\rho_2 = \frac{\textit{autocovariance at lag 2}}{\textit{variance}}$$

$$\rho_2 = \frac{\gamma_2}{\gamma_0} = \frac{0}{(1+\theta_1^2)\sigma_\epsilon^2} = 0$$

The autocorrelation at lag 2, ρ_2, for an MA(1) process

$$\boldsymbol{\rho_2 = 0}$$

It is not difficult to see that

$$\rho_3 = 0$$
$$\rho_4 = 0$$
$$\vdots$$

An *MA*(1) process has a very short "memory"

$$\rho_0 = 1$$
$$\rho_1 = \frac{-\theta_1}{(1+\theta_1^2)}$$
$$\rho_2 = 0$$
$$\rho_3 = 0$$
$$\vdots$$

Examples of *MA*(1) processes and their autocorrelation functions

Example B.1

Suppose $\theta_1 = .5$ in an *MA*(1) process.

$$Y_t = -\theta_1\epsilon_{t-1} + \epsilon_t$$
$$Y_t = -.5\epsilon_{t-1} + \epsilon_t$$

Then,

$$\rho_0 = 1$$
$$\rho_1 = \frac{-\theta_1}{(1+\theta_1^2)} = \frac{-.5}{1+(.5)^2} = \frac{-.5}{1.25} = -.4$$
$$\rho_2 = 0$$
$$\rho_3 = 0$$
$$\vdots$$

Table 11.18

ACF

Lags	*0*	*1*	*2*	*3*	*4*	*5*	$\cdots$
Autocorrelation	*1*	*.40*	*0*	*0*	*0*	*0*	$\cdots$

Figure 11-1

With the above example, the *MA*(1) model has about 40% "memory" of what occured 1 period back and then zero memory beyond 1 period.

Example B.2

Suppose $\theta_1 = -.7$ in an *MA*(1) process.

$$Y_t = -\theta_1\epsilon_{t-1} + \epsilon_t$$
$$Y_t = -(-.7)\epsilon_{t-1} + \epsilon_t$$
$$Y_t = +.7\epsilon_{t-1} + \epsilon_t$$

Then,

$$\rho_0 = 1$$

$$\rho_1 = \frac{-\theta_1}{(1+\theta_1^2)} = \frac{+.7}{1+(-.7)^2} = \frac{+.7}{1.49} = +.47$$

$$\rho_2 = 0$$

$$\rho_3 = 0$$

$$\vdots$$

Table 11.19

ACF

Lags	*0*	*1*	*2*	*3*	*4*	*5*	$\cdots$
Autocorrelations	*1*	*.47*	*0*	*0*	*0*	*0*	$\cdots$

The characteristic graph of an *MA*(1) process is a spike at lag 1 and zeros elsewhere. With the above example, the *MA*(1) model has about 47% "memory" of what occured 1 period back and then zero memory beyond 1 period.

Consequently, when we observe an ACF of an unknown process, and it has only one spike at lag 1, and zeros elsewhere, we can then tentatively identify the time series as being generated by an *MA*(1) process.

The general *MA*(2) model is written

$$Y_t = \mu_Y - \theta_1\epsilon_{t-1} - \theta_2\epsilon_{t-2} + \epsilon_t \qquad \text{B.39}$$

It is not difficult to show that for an *MA*(2) process the mean is again μ_Y

$$E(Y_t) = \mu_Y$$

So again, without loss of generality, we set $\mu_Y = 0$.

$$MA(2) \qquad Y_t = -\theta_1\epsilon_{t-1} - \theta_2\epsilon_{t-2} + \epsilon_t \qquad \text{B.40}$$

The Variance of an *MA*(2) process.

By definition of variance (equation B.3),

$$\sigma^2_{Y_t} = E[Y_t - E(Y_t)]^2$$

Since, $E(Y_t) = 0$, $\sigma^2_{Y_t} = E[Y_t]^2$ B.41

Now, because $Y_t = -\theta_1\epsilon_{t-1} - \theta_2\epsilon_{t-2} + \epsilon_t$ (equation B.40) becomes

$$\sigma^2_{Y_t} = E[-\theta_1\epsilon_{t-1} - \theta_2\epsilon_{t-2} + \epsilon_t]^2 \qquad \text{B.42}$$

$$= \mathrm{E}[\theta_1^2\epsilon_{t-1}^2 + \theta_2^2\epsilon_{t-2}^2 + 2\theta_1\theta_2\epsilon_{t-1}\epsilon_{t-2} - 2\theta_1\epsilon_{t-1}\epsilon_t - 2\theta_2\epsilon_{t-2}\epsilon_t + \epsilon_t^2] \qquad \text{B.43}$$

We then re-write equation B.20 using Facts 1 and 2 of the expectation operator.

$$= \theta_1^2\mathrm{E}[\epsilon_{t-1}^2] + \theta_2^2\mathrm{E}[\epsilon_{t-2}^2] + 2\theta_1\theta_2\mathrm{E}[\epsilon_{t-1}\epsilon_{t-2}] - 2\theta_1\mathrm{E}[\epsilon_{t-1}\epsilon_t] - 2\theta_2\mathrm{E}[\epsilon_{t-2}\epsilon_t] + \mathrm{E}[\epsilon_t^2] \qquad \text{B.44}$$

Now, the terms having errors for different periods go to zero under expectation, and the squared errror terms become σ^2_ϵ under expectation, so that equation B.43 reduces to:

$$= \theta_1^2\sigma^2_\epsilon + \theta_2^2\sigma^2_\epsilon + 2\theta_1\theta_2 \cdot 0 - 2\theta_1 \cdot 0 - 2\theta_2 \cdot 0 + \sigma^2_\epsilon \qquad \text{B.45}$$

So,

$$\sigma^2_{Y_t} = \theta_1^2\sigma^2_\epsilon + \theta_2^2\sigma^2_\epsilon + \sigma^2_\epsilon \qquad \text{B.46}$$

$$\sigma^2_{Y_t} = (1 + \theta_1^2 + \theta_2^2)\sigma^2_\epsilon \qquad \text{B.47}$$

The Variance of an MA(2) process

$\boldsymbol{\sigma^2_{Y_t} \ = \ (1 \ + \ \theta_1^2 \ + \ \theta_2^2)\sigma_\epsilon^2}$	**B.48**

The variance is equivalent to γ_0, the autocovariance at lag 0.

The Autocovariance at lag 0, γ_0, of an MA(2) process

$\boldsymbol{\gamma_0 \ = \ (1 \ + \ \theta_1^2 \ + \theta_2^2)\sigma_\epsilon^2}$	**B.49**

Autocorrelation at lag 1, ρ_1, for an *MA*(2) process

This first requires the autocovariance at lag 1.

By definition,
$$\gamma_1 \ = \ Cov(Y_t, Y_{t-1}) \quad \text{B.24}$$
$$\gamma_1 \ = \ E[(Y_t \ - \ E(Y_t))(Y_{t-1} \ - \ E(Y_{t-1}))] \quad \text{B.25}$$

Since $E(Y_t) \ = \ E(Y_{t-1}) \ = \ 0$ equation B.25 reduces to
$$\gamma_1 \ = \ E[Y_t Y_{t-1}] \quad \text{B.26}$$

Since the MA(2) is: $Y_t \ = \ - \ \theta_1\epsilon_{t-1} \ - \ \theta_2\epsilon_{t-2} + \ \epsilon_t$

then $Y_{t-1} \ = \ - \ \theta_1\epsilon_{t-2} \ - \ \theta_2\epsilon_{t-3} \ + \ \epsilon_{t-1}$

Substituting the expressions for Y_t and Y_{t-1} into equation B.26
$$\gamma_1 \ = \ E[(- \ \theta_1\epsilon_{t-1} \ - \ \theta_2\epsilon_{t-2} + \ \epsilon_t)(- \ \theta_1\epsilon_{t-2} \ - \ \theta_2\epsilon_{t-3} + \ \epsilon_{t-1})]$$

Multiplying out the binomial and taking expectations
$$= \ E(\theta_1^2\epsilon_{t-1}\epsilon_{t-2} \ + \ \theta_1\theta_2\epsilon_{t-1}\epsilon_{t-3} - \ \theta_1\epsilon_{t-1}^2 \ + \ \theta_1\theta_2\epsilon_{t-2}^2 + \ \theta_2^2\epsilon_{t-2}\epsilon_{t-3}$$
$$- \ \theta_2\epsilon_{t-1}\epsilon_{t-2} \ - \ \theta_1\epsilon_t\epsilon_{t-2} + \ \theta_2\epsilon_t\epsilon_{t-2} \ + \ \epsilon_t\epsilon_{t-1})$$

$$\gamma_1 \ = \ (\text{-}\theta_1 \ + \ \theta_1\theta_2)\sigma_\epsilon^2$$
$$= \ (\text{-}\theta_1 \ + \ \theta_1\theta_2)\sigma_\epsilon^2$$

The Autocovariance at lag 1, γ_1, for an MA(2) process

$\boldsymbol{\gamma_1 \ = \ (\text{-}\theta_1 \ + \ \theta_1\theta_2)\sigma_\epsilon^2}$	**B.50**

We now have for an MA(2)
$$\gamma_0 \ = \ (1 \ + \ \theta_1^2 \ + \ \theta_2^2)\sigma_\epsilon^2$$

and
$$\gamma_1 \ = \ (\text{-}\theta_1 \ + \ \theta_1\theta_2)\sigma_\epsilon^2$$

Thus, by defintion of *autocorrelation at lag 1*, ρ_1

$$\rho_1 \ = \ \frac{\textit{autocovariance at lag 1}}{\textit{variance}}$$

$$\rho_1 \ = \ \frac{\gamma_1}{\gamma_0} \ = \ \frac{(\text{-}\theta_1 + \theta_1\theta_2)\sigma_\epsilon^2}{(1 + \theta_1^2)\sigma_\epsilon^2} \ = \ \frac{\text{-}\theta_1 + \theta_1\theta_2}{(1 + \theta_1^2 + \theta_2^2)}$$

The Autocorrelation at lag 1, ρ_1, for an MA(2) process

$$\boldsymbol{\rho_1 = \frac{-\theta_1 + \theta_1\theta_2}{(1 + \theta_1^2 + \theta_2^2)}} \qquad \textbf{B.51}$$

Autocorrelation at lag 2, ρ_2, for an *MA*(2) process

Using similar expectation operations we derive that

The Autocovariance at lag 2, γ_2, for an MA(2) process

$$\boldsymbol{\gamma_2 = -\theta_2\sigma_\epsilon^2} \qquad \textbf{B.52}$$

and then

$$\rho_2 = \frac{\gamma_2}{\gamma_0} = \frac{-\theta_2\sigma_\epsilon^2}{(1 + \theta_1^2 + \theta_2^2)\sigma_\epsilon^2} = \frac{-\theta_2}{(1 + \theta_1^2 + \theta_2^2)}$$

The Autocorrelation at lag 2, ρ_2, for an MA(2) process

$$\boldsymbol{\rho_2 = \frac{-\theta_2}{(1 + \theta_1^2 + \theta_2^2)}} \qquad \textbf{B.53}$$

Autocorrelation at lag 3, ρ_3, for an MA(2) process

It is not difficult to show that the autocovariance at lag 3, γ_3, is zero. And since $\gamma_3 = 0$, then $\rho_3 = 0$.

Thus, for an *MA*(2) process,

$$\begin{aligned}
\gamma_0 &= (1 + \theta_1^2 + \theta_2^2)\sigma_\epsilon^2 \\
\gamma_1 &= (-\theta_1 + \theta_1\theta_2)\sigma_\epsilon^2 \\
\gamma_2 &= -\theta_2\sigma_\epsilon^2 \\
\gamma_3 &= 0 \\
\gamma_4 &= 0 \\
&\vdots
\end{aligned}$$

So,

$$\begin{aligned}
\rho_0 &= 1 \\
\rho_1 &= \frac{-\theta_1 + \theta_1\theta_2}{(1 + \theta_1^2 + \theta_2^2)} \\
\rho_2 &= \frac{-\theta_2}{(1 + \theta_1^2 + \theta_2^2)} \\
\rho_3 &= 0 \\
\rho_4 &= 0 \\
&\vdots
\end{aligned}$$

The Autocovariance and Autocorrelation of an *AR*(1) process

The general *AR*(1) process is written

$$Y_t = \delta + \phi_1 Y_{t-1} + \epsilon_t$$

Recall that δ is <u>not</u> the mean of the process; it is a constant, like the β_0 in a regression model. In other words, $E(Y_t) \neq \delta$ in an AR(1) process.

The mean of an *AR*(1) process

By condition (1) of stationarity,

$$E(Y_t) = E(Y_{t-1})$$

So, if we take expectations of equation B.50,

$$E(Y_t) = E(\delta + \phi_1 Y_{t-1} + \epsilon_t) \qquad \text{B.}\mathit{54}$$

$$E(Y_t) = \delta + \phi_1 E(Y_{t-1}) + E(\epsilon_t) \qquad \text{B.}\mathit{55}$$

$$E(Y_t) = \delta + \phi_1 E(Y_t) + \mathit{0} \qquad \text{B.56}$$

Solving for $E(\mathrm{Y}_t)$,

$$\mathrm{E}(\mathrm{Y}_t) - \phi_1 \mathrm{E}(\mathrm{Y}_t) = \delta$$

$$\mathrm{E}(\mathrm{Y}_t) = \frac{\delta}{1 - \phi_1} \qquad \text{B.57}$$

The mean of an AR(1) process

$$\boxed{\boldsymbol{E(Y_t) = \frac{\delta}{1 - \phi_1}} \qquad \textbf{B.58}}$$

The variance of an *AR*(1) process, γ_0

Without loss of generality, we set $\delta = 0$, which, in turn, means that

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

and $\quad E(Y_t) = \mathit{0}.$

By definition of variance (equation B.3),

$$\sigma^2_{Y_t} = E[Y_t - E(Y_t)]^2$$

Since,

$$E(Y_t) = 0, \quad \sigma^2_{Y_t} = E[Y_t]^2 \qquad \text{B.59}$$

Now, because $Y_t = \phi_1 Y_{t-1} + \epsilon_t$ equation B.57 becomes

$$\sigma^2_{Y_t} = E[\phi_1 Y_{t-1} + \epsilon_t]^2 \qquad \text{B.60}$$

$$= E[\phi_1^2 Y_{t-1}^2 + \mathit{2}\phi_1 Y_{t-1}\epsilon_t + \epsilon_t^2]$$

$$\sigma^2_{Y_t} = \phi_1^2 E[Y_{t-1}^2] + \mathit{2}\phi_1 E[Y_{t-1}\epsilon_t] + E[\epsilon_t^2] \qquad \text{B.61}$$

Equation B.59 may be simplied.

We are assuming that the process is stationary, so $E[Y_{t-1}^2] = \sigma^2_{Y_t}$ in equation B.59.

Y_{t-1} and ϵ_t are uncorrelated, so $E[Y_{t-1}\epsilon_t] = \mathit{0}$ in equation B.59.

And by property 2 of the ϵ's, $E[\epsilon_t^2] = \sigma^2_\epsilon$ in equation B.59.

Thus, equation B.59 simplies to

$$\sigma^2_{Y_t} = \phi_1^2 \sigma^2_{Y_t} + \mathit{2}\phi_1 \cdot \mathit{0} + \sigma^2_\epsilon$$

Solving for $\sigma^2_{Y_t}$

***The variance of an AR*(1) process**

$$\boxed{\boldsymbol{\gamma_0 = \sigma^2_{Y_t} = \frac{\sigma^2_\epsilon}{1 - \phi_1^2}} \qquad \textbf{B.62}}$$

Determining γ_1 and ρ_1 of an *AR*(1) process

By definition, $\gamma_1 = COV(Y_t, Y_{t-1})$ B.63

$$= E[(Y_t - E(Y_t))(Y_{t-1} - E(Y_{t-1}))]$$

$$= E[(Y_t)(Y_{t-1})]$$

$$= E[(\phi_1 Y_{t-1} + \epsilon_t)(Y_{t-1})]$$

$$= \phi_1 E(Y_{t-1}^2) + E(\epsilon_t Y_{t-1})$$

$E(Y_{t-1}^2) = \gamma_0$, and because ϵ_t and Y_{t-1} are uncorrelated $E(\epsilon_t Y_{t-1}) = 0$

$$\gamma_1 = \phi_1 \gamma_0$$

The Autocovariance at lag 1, γ_1, for an AR(1) process

$$\boldsymbol{\gamma_1 = \phi_1 \gamma_0} \qquad \textbf{6.64}$$

However, we are more concerned with the ***autocorrelation at lag 1***. By definition,

$$\rho_1 = \frac{\gamma_1}{\gamma_0}$$

using equation (B.41) we have

$$\rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{\phi_1 \gamma_0}{\gamma_0}$$

$$\rho_1 = \phi_1$$

The Autocorrelation at lag 1, ρ_1, for an AR(1) process

$$\boldsymbol{\rho_1 = \phi_1} \qquad \textbf{B.65}$$

Determining γ_2 and ρ_2 of an AR(1) process

By definition, $\gamma_2 = COV(Y_t, Y_{t-2})$ *B.66*

$$= E[(Y_t - E(Y_t))(Y_{t-2} - E(Y_{t-2}))]$$

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

$$= \phi_1 E(Y_{t-1} Y_{t-2}) + E(\epsilon_t Y_{t-2})$$

$E(Y_{t-1} Y_{t-2}) = \gamma_1$, and because ϵ_t and Y_{t-2} are uncorrelated $E(\epsilon_t Y_{t-2}) = 0$

$$\gamma_2 = \phi_1 \gamma_1$$

The Autocovariance at lag 2, γ_2, for an AR(1) process

$$\boldsymbol{\gamma_2 = \phi_1 \gamma_1} \qquad \textbf{6.67}$$

However, we are more concerned with the ***autocorrelation at lag 2***. By definition,

$$\rho_2 = \frac{\gamma_2}{\gamma_0}$$

using equation (B.41) we have

$$\rho_2 = \frac{\gamma_2}{\gamma_0} = \frac{\phi_1 \gamma_1}{\gamma_0} = \frac{\phi_1(\phi_1 \gamma_0)}{\gamma_0}$$

$$\rho_2 = \phi_1^2$$

The Autocorrelation at lag 1, ρ_2, for an AR(1) process

$$\boldsymbol{\rho_2 = \phi_1^2} \qquad \textbf{B.68}$$

Equivalently, it is not difficult to see that the

The Autocovariance at lag k, γ_k, for an AR(1) process

$$\boldsymbol{\gamma_k = \phi_1 \gamma_{k-1}} \qquad \textbf{B.69}$$

The Autocorrelation at lag k, ρ_k, for an AR(1) process

$$\boldsymbol{\rho_k = \phi_1^k} \qquad \textbf{B.70}$$

Example B.3

Suppose we have an *AR(1)* with $\phi_1 = .7$ and $\delta = 0$.

$$Y_t = \delta + \phi_1 Y_{t-1} + \epsilon_t$$
$$Y_t = .7Y_{t-1} + \epsilon_t$$

Then,

$$\begin{aligned} \rho_0 &= 1 \\ \rho_1 &= .7 = .7 \\ \rho_2 &= (.7)^2 = .49 \\ \rho_3 &= (.7)^3 = .343 \\ \rho_4 &= (.7)^4 = .2401 \\ &\vdots \end{aligned}$$

Table 11.20

ACF

Lags	*0*	*1*	*2*	*3*	*4*	$\cdots$
Autocorrelations	*1*	*.7*	*.49*	*.343*	*.2401*	$\cdots$

Figure 11-2

Example B.4

Suppose we have an *AR(1)* with $\phi_1 = -.3$ and $\delta = 29$.

$$Y_t = \delta + \phi_1 Y_{t-1} + \epsilon_t$$
$$Y_t = -.3Y_{t-1} + \epsilon_t$$

Then,

$$\begin{aligned} \rho_0 &= 1 \\ \rho_1 &= -.3 = -.3 \\ \rho_2 &= (-.3)^2 = .09 \\ \rho_3 &= (-.3)^3 = -.027 \\ \rho_4 &= (-.3)^4 = .0081 \end{aligned}$$

$\vdots$

Table 11.21

ACF

Lags	*0*	*1*	*2*	*3*	*4*	$\cdots$
Autocorrelations	*1*	*-0.3*	*.09*	*-0.027*	*.0081*	$\cdots$

Both correlograms have spikes that "decay exponentially." In the first example, $\phi_1 = .7$ is positive so the spikes are all positive. In the second example, $\phi_1 = -.3$ is negative so the spikes are alternately positive and negative. Unlike the *MA* models, the correlogram of the *AR(1)* never cuts off to zero, it decays to zero.

Figure 11-3

The Autocovariance and Autocorrelation of an *AR(2)* process

The general *AR(2)* process is written

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

Remember that δ is not the mean of the process; it is a constant, like the β_0 in a regression model. In other words,

$E(Y_t) \neq \delta$ in an *AR(2)* process.

The mean of an *AR(2)* process

It is not difficult to show that the mean of the process is

The mean of an AR(2) process

$$E(Y_t) = \frac{\delta}{1 - \phi_1 - \phi_2} \qquad \text{B.71}$$

Notice that this is a generalization of the mean of an *AR(1)* process.

The mean of an AR(1) process

$$E(Y_t) = \frac{\delta}{1 - \phi_1}$$

Consequently, again, without loss of generality we let $\delta = 0$, so that $E(Y_t) = 0$.

The Variance of an *AR(2)* process, γ_0

The variance of an *AR(2)* is a generalization of the variance of an *AR(1)* process.

The variance of an AR(2) process

$$\sigma^2_{Y_t} = \frac{\sigma^2_\epsilon}{1 - \phi_1^2 - \phi_2^2} \qquad \text{B.72}$$

Determining γ_1 and ρ_1 of an *AR(2)* process

By definition,
$$\gamma_1 = COV(Y_t, Y_{t-1}) \quad \text{B.73}$$
$$= E[(Y_t)(Y_{t-1})]$$
$$= E[(\phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \epsilon_t)(Y_{t-1})]$$
$$= \phi_1 E(Y_{t-1}^2) + \phi_2 E(Y_{t-1}Y_{t-2}) + E(\epsilon_t Y_{t-1})$$

$E(Y_{t-1}^2) = \gamma_0$

$E(Y_{t-1}Y_{t-2}) = \gamma_1$

And because ϵ_t and Y_{t-1} are uncorrelated $E(\epsilon_t Y_{t-1}) = 0$

equation B.69 reduces to

$$\gamma_1 = \phi_1\gamma_0 + \phi_2\gamma_1$$

Solving for γ_1, $\gamma_1 = \frac{\phi_1\gamma_0}{1-\phi_2}$

The Autocovariance at lag 1, γ_1, for an AR(2) process

$$\boxed{\gamma_1 = \frac{\phi_1\gamma_0}{1-\phi_2}}$$

By the definition of ρ_1,
$$\rho_1 = \frac{\gamma_1}{\gamma_0}$$

The Autocorrelation at lag 1, γ_1, for an AR(2) process

$$\boxed{\rho_1 = \frac{\phi_1}{1-\phi_2}}$$

Determining γ_2 and ρ_2 of an *AR(2)* process

Using the same process as before we find that
$$\gamma_2 = \phi_1\gamma_1 + \phi_2\gamma_0$$

Then, using the usual definition of ρ_2

$$\rho_2 = \frac{\gamma_2}{\gamma_0}$$

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

Determining higher over values of γ and ρ of an AR(2) process

Notice that γ_3 has the pattern
$$\gamma_3 = \phi_1\gamma_2 + \phi_2\gamma_1$$
So that,
$$\rho_3 = \phi_1\rho_2 + \phi_2\rho_1$$

Thus, in general at lag *k*

Autocovariance at lag k, γ_k, of an AR(2) process

$$\boxed{\gamma_k = \phi_1\gamma_{k-1} + \phi_2\gamma_{k-2}} \quad \textbf{B.74}$$

Autocorrelation at lag k, ρ_k, of an AR(2) process

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} \tag{B.75}$$

These last two equations are called the ***Yule-Walker equations***.

The mean of an AR(p) process

$$E(Y_t) = \frac{\delta}{1 - \phi_1 - \phi_2 - \phi_3 - \ldots - \phi_p} \tag{B.76}$$

The variance of an AR(p) process

$$\gamma_0 = \frac{\sigma_\epsilon^2}{1 - \phi_1^2 - \phi_2^2 - \ldots - \phi_p^2} \tag{B.77}$$

Autocovariance at lag k, γ_k, of an AR(p) process

$$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \ldots + \phi_p \gamma_{k-p} \tag{B.78}$$

Autocorrelation at lag k, ρ_k, of an AR(p) process

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \ldots + \phi_p \rho_{k-p} \tag{B.79}$$

The first p equations are the generalizations of the Yule-Walker equations.

$$\begin{array}{ll}
k = 1 & \rho_1 = \phi_1 \\
k = 2 & \rho_2 = \phi_1 \rho_1 + \phi_2 \\
k = 3 & \rho_3 = \phi_1 \rho_2 + \phi_2 \rho_1 + \phi_3 \\
 & \vdots
\end{array}$$