

# Time Series Analysis

Autoregressive, MA and ARMA processes

Andrés M. Alonso    Carolina García-Martos

Universidad Carlos III de Madrid

Universidad Politécnica de Madrid

June – July, 2012

## 4. Autoregressive, MA and ARMA processes

### 4.1 Autoregressive processes

#### Outline:

- Introduction
- The first-order autoregressive process,  $AR(1)$
- The  $AR(2)$  process
- The general autoregressive process  $AR(p)$
- The partial autocorrelation function

#### Recommended readings:

- ▶ Chapter 2 of Brockwell and Davis (1996).
- ▶ Chapter 3 of Hamilton (1994).
- ▶ Chapter 3 of Peña, Tiao and Tsay (2001).

# Introduction

- ▶ In this section we will begin our study of models for stationary processes which are useful in representing the dependency of the values of a time series on its past.
- ▶ The simplest family of these models are the autoregressive, which generalize the idea of regression to represent the linear dependence between a dependent variable  $y$  ( $z_t$ ) and an explanatory variable  $x$  ( $z_{t-1}$ ), using the relation:

$$z_t = c + bz_{t-1} + a_t$$

where  $c$  and  $b$  are constants to be determined and  $a_t$  are i.i.d  $\mathcal{N}(0, \sigma^2)$ . Above relation define the **first order autoregressive process**.

- ▶ This linear dependence can be generalized so that the present value of the series,  $z_t$ , depends not only on  $z_{t-1}$ , but also on the previous  $p$  lags,  $z_{t-2}, \dots, z_{t-p}$ . Thus, an **autoregressive process of order  $p$**  is obtained.

# The first-order autoregressive process, AR(1)

▷ We say that a series  $z_t$  follows a **first order autoregressive process**, or **AR(1)**, if it has been generated by:

$$z_t = c + \phi z_{t-1} + a_t \quad (33)$$

where  $c$  and  $-1 < \phi < 1$  are constants and  $a_t$  is a white noise process with variance  $\sigma^2$ . The variables  $a_t$ , which represent the new information that is added to the process at each instant, are known as **innovations**.

## Example 35

*We will consider  $z_t$  as the quantity of water at the end of the month in a reservoir. During the month,  $c + a_t$  amount of water comes into the reservoir, where  $c$  is the average quantity that enters and  $a_t$  is the innovation, a random variable of zero mean and constant variance that causes this quantity to vary from one period to the next.*

*If a fixed proportion of the initial amount is used up each month,  $(1 - \phi)z_{t-1}$ , and a proportion,  $\phi z_{t-1}$ , is maintained the quantity of water in the reservoir at the end of the month will follow process (33).*

# The first-order autoregressive process, AR(1)

► The condition  $-1 < \phi < 1$  is necessary for the process to be stationary. To prove this, let us assume that the process begins with  $z_0 = h$ , with  $h$  being any fixed value. The following value will be  $z_1 = c + \phi h + a_1$ , the next,  $z_2 = c + \phi z_1 + a_2 = c + \phi(c + \phi h + a_1) + a_2$  and, substituting successively, we can write:

$$\begin{aligned} z_1 &= c + \phi h + a_1 \\ z_2 &= c(1 + \phi) + \phi^2 h + \phi a_1 + a_2 \\ z_3 &= c(1 + \phi + \phi^2) + \phi^3 h + \phi^2 a_1 + \phi a_2 + a_3 \\ &\vdots \\ z_t &= c \sum_{i=0}^{t-1} \phi^i + \phi^t h + \sum_{i=0}^{t-1} \phi^i a_{t-i} \end{aligned}$$

If we calculate the expectation of  $z_t$ , as  $E[a_t] = 0$ ,

$$E[z_t] = c \sum_{i=0}^{t-1} \phi^i + \phi^t h.$$

For the process to be stationary it is a necessary condition that this function does not depend on  $t$ .

# The first-order autoregressive process, AR(1)

- ▶ The mean is constant if both summands are, which requires that on increasing  $t$  the first term converges to a constant and the second is canceled. Both conditions are verified if  $|\phi| < 1$ , because then  $\sum_{i=0}^{t-1} \phi^i$  is the sum of an geometric progression with ratio  $\phi$  and converges to  $c/(1-\phi)$ , and the term  $\phi^t$  converges to zero, thus the sum converges to the constant  $c/(1-\phi)$ .
- ▶ With this condition, after an initial transition period, when  $t \rightarrow \infty$ , all the variables  $z_t$  will have the same expectation,  $\mu = c/(1-\phi)$ , independent of the initial conditions.
- ▶ We also observe that in this process the innovation  $a_t$  is uncorrelated with the previous values of the process,  $z_{t-k}$  for positive  $k$  since  $z_{t-k}$  depends on the values of the innovations up to that time,  $a_1, \dots, a_{t-k}$ , but not on future values. Since the innovation is a white noise process, its future values are uncorrelated with past ones and, therefore, with previous values of the process,  $z_{t-k}$ .

# The first-order autoregressive process, AR(1)

- ▶ The AR(1) process can be written using the notation of the **lag operator**,  $B$ , defined by

$$Bz_t = z_{t-1}. \quad (34)$$

Letting  $\tilde{z}_t = z_t - \mu$  and since  $B\tilde{z}_t = \tilde{z}_{t-1}$  we have:

$$(1 - \phi B)\tilde{z}_t = a_t. \quad (35)$$

- ▶ This condition indicates that a series follows an AR(1) process if on applying the operator  $(1 - \phi B)$  a white noise process is obtained.
- ▶ The operator  $(1 - \phi B)$  can be interpreted as a filter that when applied to the series converts it into a series with no information, a white noise process.

# The first-order autoregressive process, AR(1)

- ▶ If we consider the operator as an equation, in  $B$  the coefficient  $\phi$  is called the factor of the equation.
- ▶ The stationarity condition is that this factor be less than the unit in absolute value.
- ▶ Alternatively, we can talk about the root of the equation of the operator, which is obtained by making the operator equal to zero and solving the equation with  $B$  as an unknown;

$$1 - \phi B = 0$$

which yields  $B = 1/\phi$ .

- ▶ The condition of stationarity is then that the root of the operator be greater than one in absolute value.



# The first-order autoregressive process, AR(1)

## Expectation

▷ Taking expectations in (33) assuming  $|\phi| < 1$ , such that  $E[z_t] = E[z_{t-1}] = \mu$ , we obtain

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

Then, the **expectation** (or **mean**) is

$$\mu = \frac{c}{1 - \phi} \quad (36)$$

Replacing  $c$  in (33) with  $\mu(1 - \phi)$ , the process can be written in deviations to the mean:

$$z_t - \mu = \phi(z_{t-1} - \mu) + a_t$$

and letting  $\tilde{z}_t = z_t - \mu$ ,

$$\tilde{z}_t = \phi\tilde{z}_{t-1} + a_t \quad (37)$$

which is the most often used equation of the AR(1).

# The first-order autoregressive process, AR(1)

## Variance

► The variance of the process is obtained by squaring the expression (37) and taking expectations, which gives us:

$$E(\tilde{z}_t^2) = \phi^2 E(\tilde{z}_{t-1}^2) + 2\phi E(\tilde{z}_{t-1}a_t) + E(a_t^2).$$

We let  $\sigma_z^2$  be the variance of the stationary process. The second term of this expression is zero, since as  $\tilde{z}_{t-1}$  and  $a_t$  are independent and both variables have null expectation. The third is the variance of the innovation,  $\sigma^2$ , and we conclude that:

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

from which we find that the **variance** of the process is:

$$\sigma_z^2 = \frac{\sigma^2}{1 - \phi^2}. \quad (38)$$

Note that in this equation the condition  $|\phi| < 1$  appears, so that  $\sigma_z^2$  is finite and positive.

# The first-order autoregressive process, AR(1)

► It is important to differentiate the marginal distribution of a variable from the conditional distribution of this variable in the previous value. The marginal distribution of each observation is the same, since the process is stationary: it has mean  $\mu$  and variance  $\sigma_z^2$ . Nevertheless, the conditional distribution of  $z_t$  if we know the previous value,  $z_{t-1}$ , has a conditional mean:

$$E(z_t|z_{t-1}) = c + \phi z_{t-1}$$

and variance  $\sigma^2$ , which according to (38), is always less than  $\sigma_z^2$ .

- If we know  $z_{t-1}$  it reduces the uncertainty in the estimation of  $z_t$ , and this reduction is greater when  $\phi^2$  is greater.
- If the AR parameter is close to one, the reduction of the variance obtained from knowledge of  $z_{t-1}$  can be very important.

# The first-order autoregressive process, AR(1)

## Autocovariance function

► Using (37), multiplying by  $z_{t-k}$  and taking expectations gives us  $\gamma_k$ , the covariance between observations separated by  $k$  periods, or the **autocovariance of order  $k$** :

$$\gamma_k = E[(z_{t-k} - \mu)(z_t - \mu)] = E[\tilde{z}_{t-k}(\phi\tilde{z}_{t-1} + a_t)]$$

and as  $E[\tilde{z}_{t-k}a_t] = 0$ , since the innovations are uncorrelated with the past values of the series, we have the following recursion:

$$\gamma_k = \phi\gamma_{k-1} \quad k = 1, 2, \dots \quad (39)$$

where  $\gamma_0 = \sigma_z^2$ .

► This equation shows that since  $|\phi| < 1$  the dependence between observations decreases when the lag increases.

► In particular, using (38):

$$\gamma_1 = \frac{\phi\sigma^2}{1 - \phi^2} \quad (40)$$

# The first-order autoregressive process, AR(1)

## Autocorrelation function, ACF

► Autocorrelations contain the same information as the autocovariances, with the advantage of not depending on the units of measurement. From here on we will use the term simple autocorrelation function (ACF) to denote the autocorrelation function of the process in order to differentiate it from other functions linked to the autocorrelation that are defined at the end of this section.

► Let  $\rho_k$  be the **autocorrelation of order  $k$** , defined by:  $\rho_k = \gamma_k / \gamma_0$ , using (39), we have:

$$\rho_k = \phi \gamma_{k-1} / \gamma_0 = \phi \rho_{k-1}.$$

Since, according to (38) and (40),  $\rho_1 = \phi$ , we conclude that:

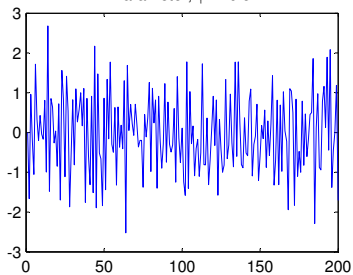
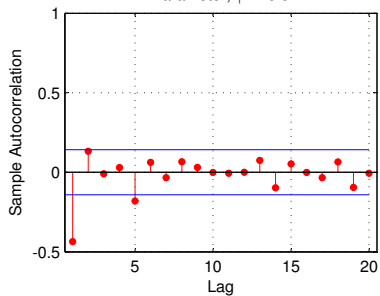
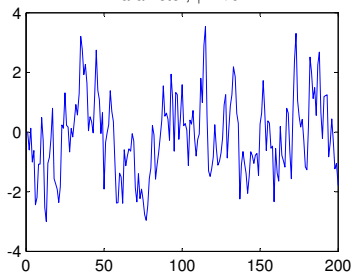
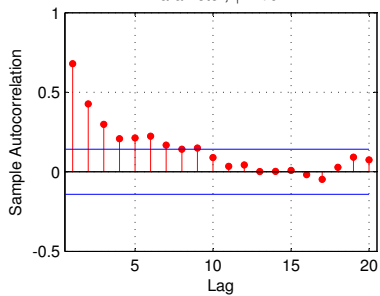
$$\rho_k = \phi^k \tag{41}$$

and when  $k$  is large,  $\rho_k$  goes to zero at a rate that depends on  $\phi$ .

# The first-order autoregressive process, AR(1)

## Autocorrelation function, ACF

- ▶ The expression (41) shows that the autocorrelation function of an AR(1) process is equal to the powers of the AR parameter of the process and decreases geometrically to zero.
- ▶ If the parameter is positive the linear dependence of the present on past values is always positive, whereas if the parameter is negative this dependence is positive for even lags and negative for odd ones.
- ▶ When the parameter is positive the value at  $t$  is similar to the value at  $t - 1$ , due to the positive dependence, thus the graph of the series evolves smoothly. Whereas, when the parameter is negative the value at  $t$  is, in general, the opposite sign of that at  $t - 1$ , thus the graph shows many changes of signs.

Parameter,  $\phi = -0.5$ Parameter,  $\phi = -0.5$ Parameter,  $\phi = +0.7$ Parameter,  $\phi = +0.7$ 

# Representation of an AR(1) process as a sum of innovations

► The AR(1) process can be expressed as a function of the past values of the innovations. This representation is useful because it reveals certain properties of the process. Using  $\tilde{z}_{t-1}$  in the expression (37) as a function of  $\tilde{z}_{t-2}$ , we have

$$\tilde{z}_t = \phi(\phi\tilde{z}_{t-2} + a_{t-1}) + a_t = a_t + \phi a_{t-1} + \phi^2 \tilde{z}_{t-2}.$$

If we now replace  $\tilde{z}_{t-2}$  with its expression as a function of  $\tilde{z}_{t-3}$ , we obtain

$$\tilde{z}_t = a_t + \phi a_{t-1} + \phi^2 a_{t-2} + \phi^3 \tilde{z}_{t-2}$$

and repeatedly applying this substitution gives us:

$$\tilde{z}_t = a_t + \phi a_{t-1} + \phi^2 a_{t-2} + \dots + \phi^{t-1} a_1 + \phi^t \tilde{z}_1$$

► If we assume  $t$  to be large, since  $\phi^t$  will be close to zero we can represent the series as a function of all the past innovations, with weights that decrease geometrically.



# Representation of an AR(1) process as a sum of innovations

- ▶ Other possibility is to assume that the series starts in the infinite past:

$$\tilde{z}_t = \sum_{j=0}^{\infty} \phi^j a_{t-j}$$

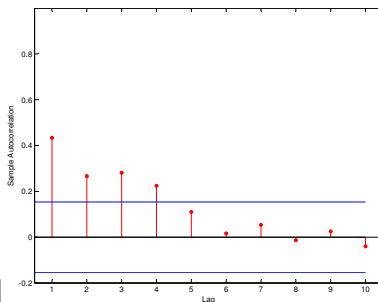
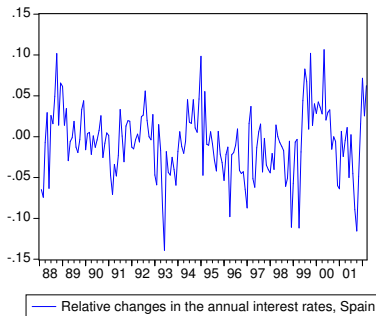
and this representation is denoted as the infinite order moving average,  $MA(\infty)$ , of the process.

- ▶ Observe that the coefficients of the innovations are precisely the coefficients of the simple autocorrelation function.
- ▶ The expression  $MA(\infty)$  can also be obtained directly by multiplying the equation (35) by the operator  $(1 - \phi B)^{-1} = 1 + \phi B + \phi^2 B^2 + \dots$ , thus obtaining:

$$\tilde{z}_t = (1 - \phi B)^{-1} a_t = a_t + \phi a_{t-1} + \phi^2 a_{t-2} + \dots$$

## Example 36

The figures show the monthly series of relative changes in the annual interest rate, defined by  $z_t = \log(y_t/y_{t-1})$  and the ACF. The AC coefficients decrease with the lag: the first is of order .4, the second close to  $.4^2 = .16$ , the third is a similar value and the rest are small and not significant.



Datafile `interestrates.xls`

# The AR(2) process

► The dependency between present and past values which an AR(1) establishes can be generalized allowing  $z_t$  to be linearly dependent not only on  $z_{t-1}$  but also on  $z_{t-2}$ . Thus the second order autoregressive, or AR(2) is obtained:

$$z_t = c + \phi_1 z_{t-1} + \phi_2 z_{t-2} + a_t \quad (42)$$

where  $c$ ,  $\phi_1$  and  $\phi_2$  are now constants and  $a_t$  is a white noise process with variance  $\sigma^2$ .

► We are going to find the conditions that must verify the parameters for the process to be stationary. Taking expectations in (42) and imposing that the mean be constant, results in:

$$\mu = c + \phi_1 \mu + \phi_2 \mu$$

which implies

$$\mu = \frac{c}{1 - \phi_1 - \phi_2}, \quad (43)$$

and the condition for the process to have a finite mean is that  $1 - \phi_1 - \phi_2 \neq 0$ .

# The AR(2) process

► Replacing  $c$  with  $\mu(1 - \phi_1 - \phi_2)$  and letting  $\tilde{z}_t = z_t - \mu$  be the process of deviations to the mean, the AR(2) process is:

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + a_t. \quad (44)$$

► In order to study the properties of the process it is advisable to use the operator notations. Introducing the lag operator,  $B$ , the equation of this process is:

$$(1 - \phi_1 B - \phi_2 B^2) \tilde{z}_t = a_t. \quad (45)$$

► The operator  $(1 - \phi_1 B - \phi_2 B^2)$  can always be expressed as  $(1 - G_1 B)(1 - G_2 B)$ , where  $G_1^{-1}$  and  $G_2^{-1}$  are the roots of the equation of the operator considering  $B$  as a variable and solving

$$1 - \phi_1 B - \phi_2 B^2 = 0. \quad (46)$$

# The AR(2) process

- ▶ The equation (46) is called **the characteristic equation** of the operator.
- ▶  $G_1$  and  $G_2$  are also said to be **factors** of the **characteristic polynomial** of the process. These roots can be real or complex conjugates.
- ▶ It can be proved that the condition of stationarity is that  $|G_i| < 1$ ,  $i = 1, 2$ .
- ▶ This condition is analogous to that studied for the AR(1).
- ▶ Note that this result is consistent with the condition found for the mean to be finite. If the equation

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

has a unit root it is verified that  $1 - \phi_1 - \phi_2 = 0$  and the process is not stationary, since it does not have a finite mean.

# The AR(2) process

## Autocovariance function

▷ Squaring expression (44) and taking expectations, we find that the variance must satisfy:

$$\gamma_0 = \phi_1^2 \gamma_0 + \phi_2^2 \gamma_0 + 2\phi_1 \phi_2 \gamma_1 + \sigma^2. \quad (47)$$

▷ In order to calculate the autocovariance, multiplying the equation (44) by  $\tilde{z}_{t-k}$  and taking expectations, we obtain:

$$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2}, \quad k \geq 1 \quad (48)$$

▷ Specifying this equation for  $k = 1$ , since  $\gamma_{-1} = \gamma_1$ , we have

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

which provides  $\gamma_1 = \phi_1 \gamma_0 / (1 - \phi_2)$ . Using this expression in (47) results in the formula for the variance:

$$\sigma_z^2 = \gamma_0 = \frac{(1 - \phi_2) \sigma^2}{(1 + \phi_2)(1 - \phi_1 - \phi_2)(1 + \phi_1 - \phi_2)}. \quad (49)$$

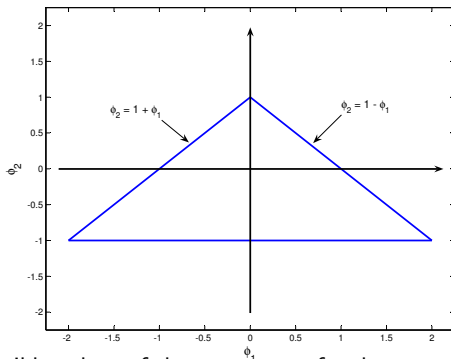
## Autocovariance function

► For the process to be stationary this variance must be positive, which will occur if the numerator and the denominator have the same sign. It can be proved that the values of the parameters that make AR(2) a stationary process are those

$$-1 < \phi_2 < 1$$

included in the region:  $\phi_1 + \phi_2 < 1$

$$\phi_2 - \phi_1 < 1$$



It represents the admissible values of the parameters for the process to be stationary.

# The AR(2) process

► In this process it is important again to differentiate the marginal and conditional properties. Assuming that the conditions of stationarity are verified, the marginal mean is given by

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

and the marginal variance is

$$\sigma_z^2 = \frac{(1 - \phi_2) \sigma^2}{(1 + \phi_2)(1 - \phi_1 - \phi_2)(1 + \phi_1 - \phi_2)}.$$

► Nevertheless, the conditional mean of  $z_t$  given the previous values is:

$$E(z_t | z_{t-1}, z_{t-2}) = c + \phi_1 z_{t-1} + \phi_2 z_{t-2}$$

and its variance will be  $\sigma^2$ , the variance of the innovations which will always be less than the marginal variance of the process  $\sigma_z^2$ .



# The AR(2) process

## Autocorrelation function

► Dividing by the variance in equation (48), we obtain the relationship between the autocorrelation coefficients:

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} \quad k \geq 1 \quad (50)$$

specifying (50) for  $k = 1$ , as in a stationary process  $\rho_1 = \rho_{-1}$ , we obtain:

$$\rho_1 = \frac{\phi_1}{1 - \phi_2} \quad (51)$$

and specifying (50) for  $k = 2$  and using (51):

$$\rho_2 = \frac{\phi_1^2}{1 - \phi_2} + \phi_2. \quad (52)$$

# The AR(2) process

## Autocorrelation function

▷ For  $k \geq 3$  the autocorrelation coefficients can be obtained recursively starting from the difference equation (50). It can be proved that the general solution to this equation is:

$$\rho_k = A_1 G_1^k + A_2 G_2^k \quad (53)$$

where  $G_1$  and  $G_2$  are the factors of the characteristic polynomial of the process and  $A_1$  and  $A_2$  are constants to be determined from the initial conditions  $\rho_0 = 1$ , (which implies  $A_1 + A_2 = 1$ ) and  $\rho_1 = \phi_1 / (1 - \phi_2)$ .

▷ According to (53) the coefficients  $\rho_k$  will be less than or equal to the unit if  $|G_1| < 1$  and  $|G_2| < 1$ , which are the conditions of stationarity of the process.

# The AR(2) process

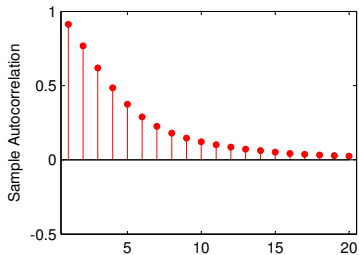
## Autocorrelation function

► If the factors  $G_1$  and  $G_2$  are complex of type  $a \pm bi$ , where  $i = \sqrt{-1}$ , then this condition is  $\sqrt{a^2 + b^2} < 1$ . We may find ourselves in the following cases:

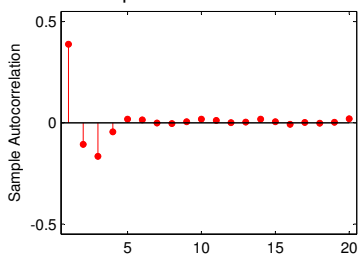
- 1 The two factors  $G_1$  and  $G_2$  are real. The decrease of (53) is the sum of the two exponentials and the shape of the autocorrelation function will depend on whether  $G_1$  and  $G_2$  have equal or opposite signs.
- 2 The two factors  $G_1$  and  $G_2$  are complex conjugates. In this case, the function  $\rho_k$  will decrease sinusoidally.

► The four types of possible autocorrelation functions for an AR(2) are shown in the next figure.

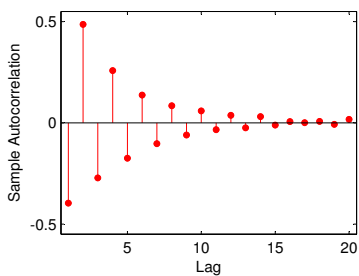
Real roots = 0.75 and 0.5



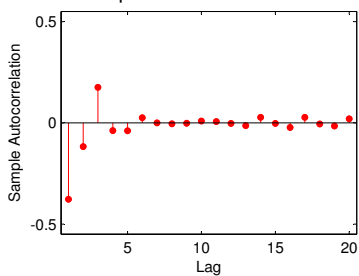
Complex roots =  $0.25 \pm 0.5i$



Real roots = -0.75 and 0.5



Complex roots =  $-0.25 \pm 0.5i$



# Representation of an AR(2) process as a sum of innovations

► The AR(2) process can be represented, as with an AR(1), as a linear combination of the innovations. Writing (45) as

$$(1 - G_1 B)(1 - G_2 B)\tilde{z}_t = a_t$$

and inverting these operators, we have

$$\tilde{z}_t = (1 + G_1 B + G_1^2 B^2 + \dots)(1 + G_2 B + G_2^2 B^2 + \dots)a_t \quad (54)$$

which leads to the MA( $\infty$ ) expression of the process:

$$\tilde{z}_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots \quad (55)$$

► We can obtain the coefficients  $\psi_i$  as a function of the roots equating powers of  $B$  in (54) and (55).

# Representation of an AR(2) process as a sum of innovations

► We can also obtain the coefficients  $\psi_i$  as a function of the coefficients  $\phi_1$  and  $\phi_2$ . Letting  $\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \dots$  since  $\psi(B) = (1 - \phi_1 B - \phi_2 B^2)^{-1}$ , we have

$$(1 - \phi_1 B - \phi_2 B^2)(1 + \psi_1 B + \psi_2 B^2 + \dots) = 1. \quad (56)$$

► Imposing the restriction that all the coefficients of the powers of  $B$  in (56) are null, the coefficient of  $B$  in this equation is  $\psi_1 - \phi_1$ , which implies  $\psi_1 = \phi_1$ . The coefficient of  $B^2$  is  $\psi_2 - \phi_1\psi_1 - \phi_2$ , which implies the equation:

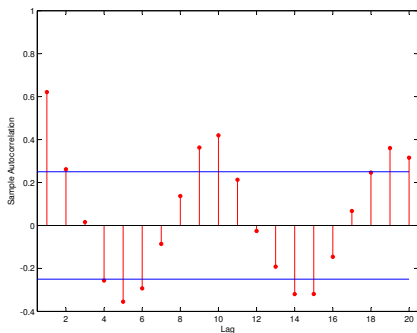
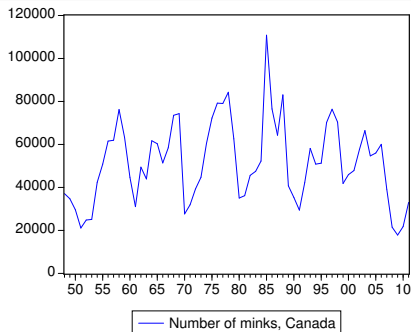
$$\psi_k = \phi_1\psi_{k-1} + \phi_2\psi_{k-2} \quad (57)$$

for  $k = 2$ , since  $\psi_0 = 1$ . The coefficients of  $B^k$  for  $k \geq 2$  verify the equation (57) which is similar to the one that must verify the autocorrelation coefficients.

► We conclude that the shape of the coefficients  $\psi_i$  will be similar to that of the autocorrelation coefficients.

## Example 37

*The figures show the number of mink sighted yearly in an area of Canada and the ACF. The series shows a cyclical evolution that could be explained by an AR(2) with negative roots corresponding to the sinusoidal structure of the autocorrelation.*



## Example 38

Write the autocorrelation function of the  $AR(2)$  process

$$z_t = 1.2z_{t-1} - 0.32z_{t-2} + a_t$$

▷ The characteristic equation of that process is:

$$0.32X^2 - 1.2X + 1 = 0$$

whose solution is:

$$X = \frac{1.2 \pm \sqrt{1.2^2 - 4 \times 0.32}}{0.64} = \frac{1.2 \pm 0.4}{0.64}$$

▷ The solutions are  $G_1^{-1} = 2.5$  and  $G_2^{-1} = 1.25$  and the factors are  $G_1 = 0.4$  and  $G_2 = 0.8$ .



▷ The characteristic equation can be written:

$$0.32X^2 - 1.2X + 1 = (1 - 0.4X)(1 - 0.8X).$$

Therefore, the process is stationary with real roots and the autocorrelation coefficients verify:

$$\rho_k = A_1 0.4^k + A_2 0.8^k.$$

▷ To determine  $A_1$  and  $A_2$  we impose the initial conditions  $\rho_0 = 1, \rho_1 = 1.2 / (1.322) = 0.91$ . Then, for  $k = 0$ :

$$1 = A_1 + A_2$$

and for  $k = 1$ ,

$$0.91 = 0.4A_1 + 0.8A_2$$

solving these equations we obtain  $A_2 = 0.51/0.4$  and  $A_1 = -0.11/0.4$ .

▷ Therefore, the autocorrelation function is:

$$\rho_k = -\frac{0.11}{0.4}0.4^k + \frac{0.51}{0.4}0.8^k$$

which gives us the following table:

$k$	0	1	2	3	4	5	6	7	8
$\rho_k$	1	0.91	0.77	0.63	0.51	0.41	0.33	0.27	0.21

▷ To obtain the representation as a function of the innovations, writing

$$(1 - 0.4B)(1 - 0.8B)z_t = a_t$$

and inverting both operators:

$$z_t = (1 + 0.4B + .16B^2 + .06B^3 + \dots)(1 + 0.8B + .64B^2 + \dots)a_t$$

yields:

$$z_t = (1 + 1.2B + 1.12B^2 + \dots)a_t.$$

# The general autoregressive process, AR(p)

► We say that a stationary time series  $z_t$  follows an **autoregressive process of order  $p$**  if:

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \dots + \phi_p \tilde{z}_{t-p} + a_t \quad (58)$$

where  $\tilde{z}_t = z_t - \mu$ , with  $\mu$  being the mean of the stationary process  $z_t$  and  $a_t$  a white noise process.

► Utilizing the operator notation, the equation of an AR(p) is:

$$(1 - \phi_1 B - \dots - \phi_p B^p) \tilde{z}_t = a_t \quad (59)$$

and letting  $\phi_p(B) = 1 - \phi_1 B - \dots - \phi_p B^p$  be the polynomial of degree  $p$  in the lag operator, whose first term is the unit, we have:

$$\phi_p(B) \tilde{z}_t = a_t \quad (60)$$

which is the general expression of an autoregressive process.

# The general autoregressive process, AR(p)

- ▷ The **characteristic equation** of this process is defined by:

$$\phi_p(B) = 0 \quad (61)$$

considered as a function of  $B$ .

- ▷ This equation has  $p$  roots  $G_1^{-1}, \dots, G_p^{-1}$ , which are generally different, and we can write:

$$\phi_p(B) = \prod_{i=1}^p (1 - G_i B)$$

such that the coefficients  $G_i$  are the factors of the characteristic equation.

- ▷ It can be proved that the process is stationary if  $|G_i| < 1$ , for all  $i$ .

# The general autoregressive process, AR(p)

## Autocorrelation function

► Operating with (58), we find that the autocorrelation coefficients of an AR(p) verify the following difference equation:

$$\rho_k = \phi_1 \rho_{k-1} + \dots + \phi_p \rho_{k-p}, \quad k > 0.$$

► In the above sections we saw particular cases in this equation for  $p = 1$  and  $p = 2$ . We can conclude that the autocorrelation coefficients satisfy the same equation as the process:

$$\phi_p(B) \rho_k = 0 \quad k > 0. \quad (62)$$

► The general solution to this equation is:

$$\rho_k = \sum_{i=1}^p A_i G_i^k, \quad (63)$$

where the  $A_i$  are constants to be determined from the initial conditions and the  $G_i$  are the factors of the characteristic equation.

# The general autoregressive process, AR(p)

## Autocorrelation function

- ▶ For the process to be stationary the modulus of  $G_i$  must be less than one or, the roots of the characteristic equation (61) must be greater than one in modulus, which is the same.
- ▶ To prove this, we observe that the condition  $|\rho_k| < 1$  requires that there not be any  $G_i$  greater than the unit in (63), since in that case, when  $k$  increases the term  $G_i^k$  will increase without limit.
- ▶ Furthermore, we observe that for the process to be stationary there cannot be a root  $G_i$  equal to the unit, since then its component  $G_i^k$  would not decrease and the coefficients  $\rho_k$  would not tend to zero for any lag.
- ▶ Equation (63) shows that the autocorrelation function of an AR(p) process is a mixture of exponents, due to the terms with real roots, and sinusoids, due to the complex conjugates. As a result, their structure can be very complex.

# Yule-Walker equations

▷ Specifying the equation (62) for  $k = 1, \dots, p$ , a system of  $p$  equations is obtained that relate the first  $p$  autocorrelations with the parameters of the process. This is called the Yule-Walker system:

$$\begin{aligned}\rho_1 &= \phi_1 + \phi_2 \rho_1 + \dots + \phi_p \rho_{p-1} \\ \rho_2 &= \phi_1 \rho_1 + \phi_2 + \dots + \phi_p \rho_{p-2} \\ &\vdots \\ \rho_p &= \phi_1 \rho_{p-1} + \phi_2 \rho_{p-2} + \dots + \phi_p.\end{aligned}$$

▷ Defining:

$$\phi' = [\phi_1, \dots, \phi_p], \quad \rho' = [\rho_1, \dots, \rho_p], \quad \mathbf{R} = \begin{bmatrix} 1 & \rho_1 & \dots & \rho_{p-1} \\ \vdots & \vdots & & \vdots \\ \rho_{p-1} & \rho_{p-2} & \dots & 1 \end{bmatrix}$$

the above system is written as a matrix:

$$\rho = \mathbf{R}\phi \tag{64}$$

and the parameters can be determined using:  $\phi = \mathbf{R}^{-1}\rho$ .

# Yule-Walker equations - Example

## Example 39

*Obtain the parameters of an AR(3) process whose first autocorrelations are  $\rho_1 = 0.9$ ;  $\rho_2 = 0.8$ ;  $\rho_3 = 0.5$ . Is the process stationary?*

▷ The Yule-Walker equation system is:

$$\begin{bmatrix} 0.9 \\ 0.8 \\ 0.5 \end{bmatrix} = \begin{bmatrix} 1 & 0.9 & 0.8 \\ 0.9 & 1 & 0.9 \\ 0.8 & 0.9 & 1 \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \end{bmatrix}$$

whose solution is:

$$\begin{bmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \end{bmatrix} = \begin{bmatrix} 5.28 & -5 & 0.28 \\ -5 & 10 & -5 \\ 0.28 & -5 & 5.28 \end{bmatrix} \begin{bmatrix} 0.9 \\ 0.8 \\ 0.5 \end{bmatrix} = \begin{bmatrix} 0.89 \\ 1 \\ -1.11 \end{bmatrix}.$$



▷ As a result, the AR(3) process with these correlations is:

$$(1 - 0.89B - B^2 + 1.11B^3) z_t = a_t.$$

▷ To prove that the process is stationary we have to calculate the factors of the characteristic equation. The quickest way to do this is to obtain the solutions to the equation

$$X^3 - 0.89X^2 - X + 1.11 = 0$$

and check that they all have modulus less than the unit.

▷ The roots of this equation are  $-1.0550$ ,  $0.9725 + 0.3260i$  and  $0.9725 - 0.3260i$ .

▷ The modulus of the real factor is greater than the unit, thus we conclude that there is no an AR(3) stationary process that has these three autocorrelation coefficients.

# Representation of an AR(p) process as a sum of innovations

- ▶ To obtain the coefficients of the representation MA( $\infty$ ) form we use:

$$(1 - \phi_1 B - \dots - \phi_p B^p)(1 + \psi_1 B + \psi_2 B^2 + \dots) = 1$$

and the coefficients  $\psi_i$  are obtained by setting the powers of  $B$  equal to zero.

- ▶ It is proved that they must verify the equation

$$\psi_k = \phi_1 \psi_{k-1} + \dots + \phi_p \psi_{k-p}$$

which is analogous to that which verifies that autocorrelation coefficients of the process.

- ▶ As mentioned earlier, the autocorrelation coefficients,  $\rho_k$ , and the coefficients of the structure MA( $\infty$ ) are not identical: although both sequences satisfy the same difference equation and take the form  $\sum A_i G_i^k$ , the constants  $A_i$  depend on the initial conditions and will be different in both sequences.

# The partial autocorrelation function

- ▶ Determining the order of an autoregressive process from its autocorrelation function is difficult. To resolve this problem the partial autocorrelation function is introduced.
- ▶ If we compare an AR(1) with an AR(2) we see that although in both processes each observation is related to the previous ones, the type of relationship between observations separated by more than one lag is different in both processes:
  - In the AR(1) the effect of  $z_{t-2}$  on  $z_t$  is always through  $z_{t-1}$ , and given  $z_{t-1}$ , the value of  $z_{t-2}$  is irrelevant for predicting  $z_t$ .
  - Nevertheless, in an AR(2) in addition to the effect of  $z_{t-2}$  which is transmitted to  $z_t$  through  $z_{t-1}$ , there exists a direct effect on  $z_{t-2}$  on  $z_t$ .
- ▶ In general, an AR( $p$ ) has *direct* effects on observations separated by 1, 2, ...,  $p$  lags and the *direct* effects of the observations separated by more than  $p$  lags are null.

# The partial autocorrelation function

▶ The **partial autocorrelation coefficient of order  $k$** , denoted by  $\rho_k^p$ , is defined as the correlation coefficient between observations separated by  $k$  periods, when we eliminate the linear dependence due to intermediate values.

- ① We eliminate from  $\tilde{z}_t$ , the effect of  $\tilde{z}_{t-1}, \dots, \tilde{z}_{t-k+1}$  using the regression:

$$\tilde{z}_t = \beta_1 \tilde{z}_{t-1} + \dots + \beta_{k-1} \tilde{z}_{t-k+1} + u_t,$$

where the variable  $u_t$  contains the part of  $\tilde{z}_t$  not common to  $\tilde{z}_{t-1}, \dots, \tilde{z}_{t-k+1}$ .

- ② We eliminate the effect of  $\tilde{z}_{t-1}, \dots, \tilde{z}_{t-k+1}$  from  $\tilde{z}_{t-k}$  using the regression:

$$\tilde{z}_{t-k} = \gamma_1 \tilde{z}_{t-1} + \dots + \gamma_{k-1} \tilde{z}_{t-k+1} + v_t,$$

where, again,  $v_t$  contains the part of  $\tilde{z}_{t-k}$  not common to the intermediate observations.

- ③ We calculate the simple correlation coefficient between  $u_t$  and  $v_t$  which, by definition, is the **partial autocorrelation coefficient of order  $k$** .

# The partial autocorrelation function

► This definition is analogous to that of the partial correlation coefficient in regression. It can be proved that the three above steps are equivalent to fitting the multiple regression:

$$\tilde{z}_t = \alpha_{k1}\tilde{z}_{t-1} + \dots + \alpha_{kk}\tilde{z}_{t-k} + \eta_t$$

and thus  $\rho_k^p = \alpha_{kk}$ .

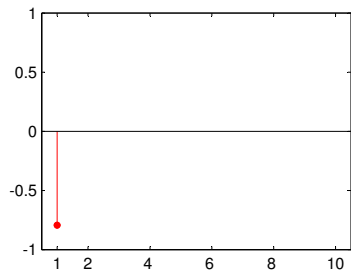
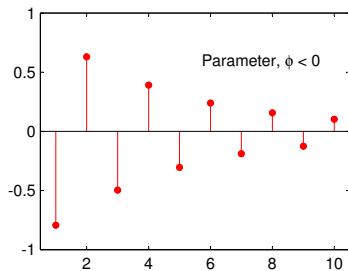
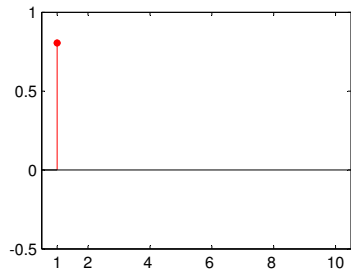
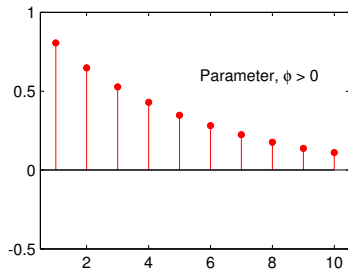
► The partial autocorrelation coefficient of order  $k$  is the coefficient  $\alpha_{kk}$  of the variable  $z_{t-k}$  after fitting an  $AR(k)$  to the data of the series. Therefore, if we fit the family of regressions:

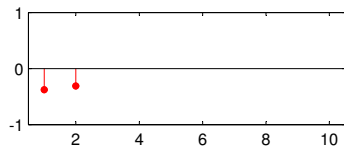
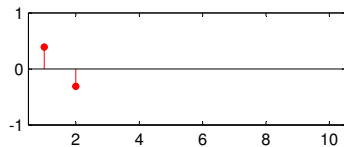
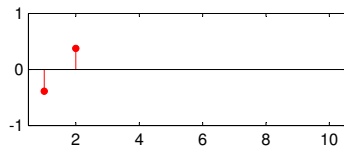
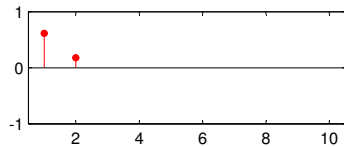
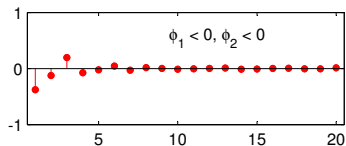
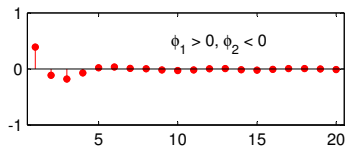
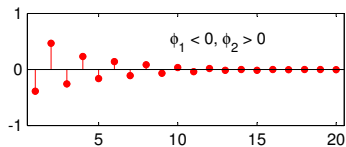
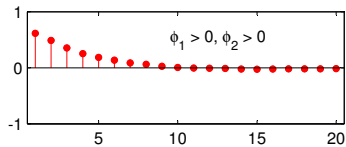
$$\begin{aligned}\tilde{z}_t &= \alpha_{11}\tilde{z}_{t-1} + \eta_{1t} \\ \tilde{z}_t &= \alpha_{21}\tilde{z}_{t-1} + \alpha_{22}\tilde{z}_{t-2} + \eta_{2t} \\ &\vdots \\ \tilde{z}_t &= \alpha_{k1}\tilde{z}_{t-1} + \dots + \alpha_{kk}\tilde{z}_{t-k} + \eta_{kt}\end{aligned}$$

the sequence of coefficients  $\alpha_{jj}$  provides the partial autocorrelation function.

# The partial autocorrelation function

- ▶ From this definition it is clear that an  $AR(p)$  process will have the first  $p$  nonzero partial autocorrelation coefficients and, therefore, in the **partial autocorrelation function (PACF)** the number of nonzero coefficients indicates the order of the AR process.
- ▶ This property will be a key element in identifying the order of an autoregressive process.
- ▶ Furthermore, the partial correlation coefficient of order  $p$  always coincides with the parameter  $\phi_p$ .
- ▶ The Durbin-Levinson algorithm is an efficient method for estimating the partial correlation coefficients.

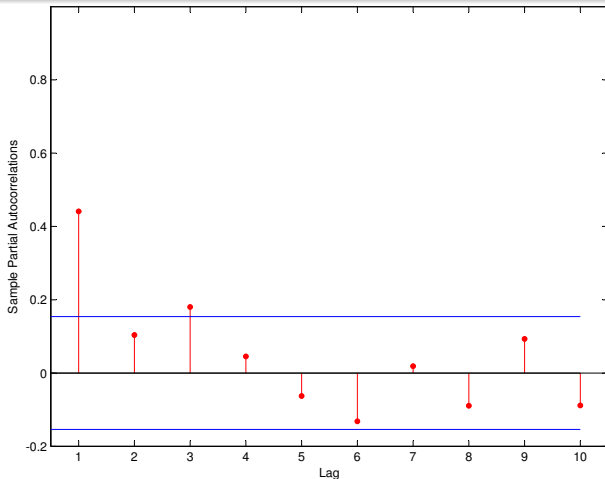






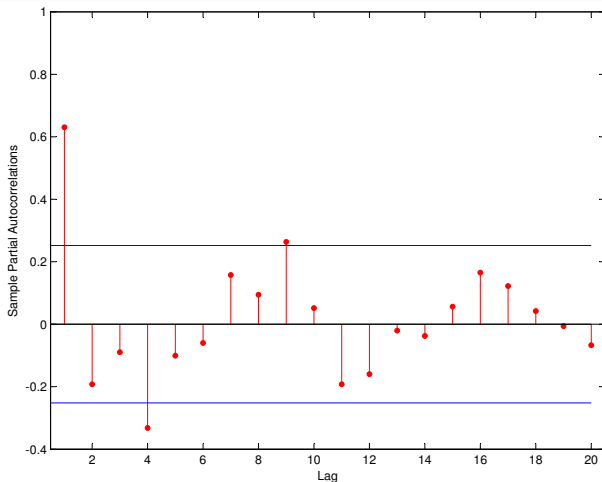
## Example 40

*The figure shows the partial autocorrelation function for the interest rates series from example 36. We conclude that the variations in interest rates follow an AR(1) process, since there is only one significant coefficient.*



## Example 41

*The figure shows the partial autocorrelation function for the data on *mink* from example 37. This series presents significant partial autocorrelation coefficients up to the fourth lag, suggesting that the model is an  $AR(4)$ .*



# Time Series Analysis

Moving average and ARMA processes

Andrés M. Alonso    Carolina García-Martos

Universidad Carlos III de Madrid

Universidad Politécnica de Madrid

June – July, 2012

# 4. Autoregressive, MA and ARMA processes

## 4.2 Moving average and ARMA processes

### Outline:

- Introduction
- The first-order moving average process,  $MA(1)$
- The  $MA(q)$  process
- The  $MA(\infty)$  process and Wold decomposition
- The  $ARMA(1,1)$  process
- The  $ARMA(p,q)$  processes
- The ARMA processes and the sum of stationary processes

### Recommended readings:

- ▷ Chapters 2 and 3 of Brockwell and Davis (1996).
- ▷ Chapter 3 of Hamilton (1994).
- ▷ Chapter 3 of Peña, Tiao and Tsay (2001).

# Introduction

- ▶ The autoregressive processes have, in general, infinite non-zero autocorrelation coefficients that decay with the lag. The AR processes have a relatively “long” memory, since the current value of a series is correlated with all previous ones, although with decreasing coefficients.
- ▶ This property means that we can write an AR process as a linear function of all its innovations, with weights that tend to zero with the lag. The AR processes cannot represent short memory series, where the current value of the series is only correlated with a small number of previous values.
- ▶ A family of processes that have this “very short memory” property are the moving average, or MA processes. The MA processes are a function of a finite, and generally small, number of its past innovations.
- ▶ Later, we will combine the properties of the AR and MA processes to define the ARMA processes, which give us a very broad and flexible family of stationary stochastic processes useful in representing many time series.

# The first order moving average, MA(1)

- ▷ A **first order moving average**, MA(1), is defined by a linear combination of the last two innovations, according to the equation:

$$\tilde{z}_t = a_t - \theta a_{t-1} \quad (65)$$

where  $\tilde{z}_t = z_t - \mu$ , with  $\mu$  being the mean of the process and  $a_t$  a white noise process with variance  $\sigma^2$ .

- ▷ The MA(1) process can be written with the operator notation:

$$\tilde{z}_t = (1 - \theta B) a_t. \quad (66)$$

- ▷ This process is the sum of the two stationary processes,  $a_t$  and  $-\theta a_{t-1}$  and, therefore, will always be stationary for any value of the parameter, unlike the AR processes.

# The first order moving average, MA(1)

- ▶ In these processes we will assume that  $|\theta| < 1$ , so that the past innovation has less weight than the present. Then, we say that the process is **invertible** and has the property whereby the effect of past values of the series decreases with time.
- ▶ To justify this property, we substitute  $a_{t-1}$  in (65) as a function of  $z_{t-1}$ :

$$\tilde{z}_t = a_t - \theta(\tilde{z}_{t-1} + \theta a_{t-2}) = -\theta \tilde{z}_{t-1} - \theta^2 a_{t-2} + a_t$$

and repeating this operation for  $a_{t-2}$ :

$$\tilde{z}_t = -\theta \tilde{z}_{t-1} - \theta^2(\tilde{z}_{t-2} + \theta a_{t-3}) + a_t = -\theta \tilde{z}_{t-1} - \theta^2 \tilde{z}_{t-2} - \theta^3 a_{t-3} + a_t$$

using successive substitutions of  $a_{t-3}$ ,  $a_{t-4}$ ..., etc., we obtain:

$$\tilde{z}_t = -\sum_{i=1}^{t-1} \theta^i \tilde{z}_{t-1} - \theta^t a_0 + a_t \quad (67)$$

# The first order moving average, MA(1)

- ▶ Notice that when  $|\theta| < 1$ , the effect of  $\tilde{z}_{t-k}$  tends to zero with  $k$  and the process is called **invertible**.
- ▶ If  $|\theta| \geq 1$  it produces the paradoxical situation in which the effect of past observations increases with the distance. From here on, we assume that the process is invertible.
- ▶ Thus, since  $|\theta| < 1$ , there exists an inverse operator  $(1 - \theta B)^{-1}$  and we can write equation (66) as:

$$(1 + \theta B + \theta^2 B^2 + \dots) \tilde{z}_t = a_t \quad (68)$$

that implies:

$$\tilde{z}_t = - \sum_{i=1}^{\infty} \theta^i \tilde{z}_{t-1} + a_t$$

which is equivalent to (67) assuming that the process begins in the infinite past. This equation represents the MA(1) process with  $|\theta| < 1$  as an AR( $\infty$ ) with coefficients that decay in a geometric progression.



# The first order moving average, MA(1)

## Expectation and variance

- ▶ The expectation can be derived from relation (65) which implies that  $E[\tilde{z}_t] = 0$ , so

$$E[z_t] = \mu.$$

- ▶ The variance of the process is calculated from (65). Squaring and taking expectations, we obtain:

$$E(\tilde{z}_t^2) = E(a_t^2) + \theta^2 E(a_{t-1}^2) - 2\theta E(a_t a_{t-1})$$

since  $E(a_t a_{t-1}) = 0$ ,  $a_t$  is a white noise process and  $E(a_t^2) = E(a_{t-1}^2) = \sigma^2$ , then we have that:

$$\sigma_z^2 = \sigma^2 (1 + \theta^2). \quad (69)$$

- ▶ This equation tells us that the marginal variance of the process,  $\sigma_z^2$ , is always greater than the variance of the innovations,  $\sigma^2$ , and this difference increases with  $\theta^2$ .

# The first order moving average, MA(1)

## Simple and partial autocorrelation function

▶ The first order autocovariance is calculated by multiplying equation (65) by  $\tilde{z}_{t-1}$  and taking expectations:

$$\gamma_1 = E(\tilde{z}_t \tilde{z}_{t-1}) = E(a_t \tilde{z}_{t-1}) - \theta E(a_{t-1} \tilde{z}_{t-1}).$$

▶ In this expression the first term  $E(a_t \tilde{z}_{t-1})$  is zero, since  $\tilde{z}_{t-1}$  depends on  $a_{t-1}$ , and  $a_{t-2}$ , but not on future innovations, such as  $a_t$ .

▶ To calculate the second term, replacing  $\tilde{z}_{t-1}$  with its expression according to (65), gives us

$$E(a_{t-1} \tilde{z}_{t-1}) = E(a_{t-1}(a_{t-1} - \theta a_{t-2})) = \sigma^2$$

from which we obtain:

$$\gamma_1 = -\theta \sigma^2. \quad (70)$$

# The first order moving average, MA(1)

## Simple and partial autocorrelation function

► The second order autocovariance is calculated in the same way:

$$\gamma_2 = E(\tilde{z}_t \tilde{z}_{t-2}) = E(a_t \tilde{z}_{t-2}) - \theta E(a_{t-1} \tilde{z}_{t-2}) = 0$$

since the series is uncorrelated with its future innovations the two terms are null. The same result is obtained for covariances of orders higher than two.

► In conclusion:

$$\gamma_j = 0, \quad j > 1. \quad (71)$$

Dividing the autocovariances (70) and (71) by expression (69) of the variance of the process, we find that the autocorrelation coefficients of an MA(1) process verify:

$$\rho_1 = \frac{-\theta}{1 + \theta^2}, \quad \rho_k = 0 \quad k > 1, \quad (72)$$

and the (ACF) will only have one value different from zero in the first lag.

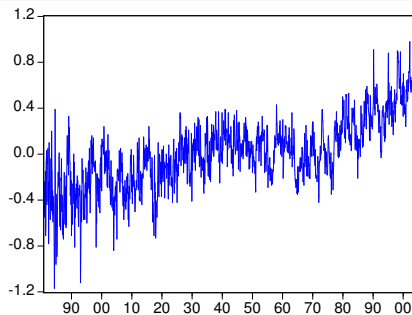
# The first order moving average, MA(1)

## Simple and partial autocorrelation function

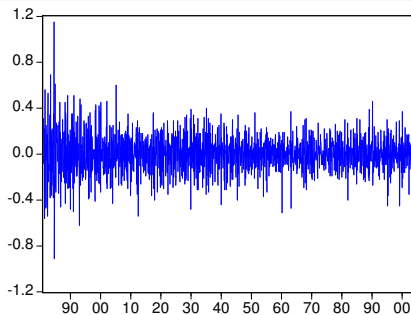
- ▶ This result proves that the autocorrelation function (*ACF*) of an MA(1) process has the same properties as the partial autocorrelation function (*PACF*) of an AR(1) process: there is a first coefficient different from zero and the rest are null.
- ▶ This duality between the AR(1) and the MA(1) is also seen in the partial autocorrelation function, *PACF*.
- ▶ According to (68), when we write an MA(1) process in autoregressive form  $z_{t-k}$  has a direct effect on  $z_t$  of magnitude  $\theta^k$ , no matter what  $k$  is.
- ▶ Therefore, the *PACF* have all non-null coefficients and they decay geometrically with  $k$ .
- ▶ This is the structure of the *ACF* in an AR(1) and, hence, we conclude that the *PACF* of an MA(1) has the same structure as the *ACF* of an AR(1).

## Example 42

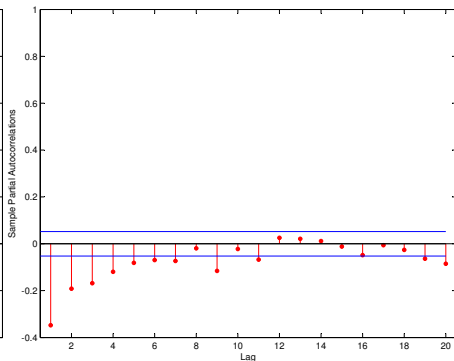
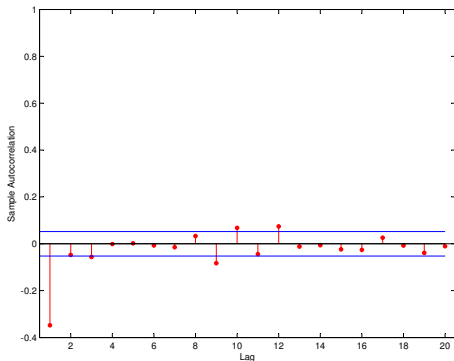
The left figure show monthly data from the years 1881 - 2002 and represent the deviation between the average temperature of a month and the mean of that month calculated by averaging the temperatures in the 25 years between 1951 and 1975. The right figure show  $z_t = y_t - y_{t-1}$ , which represents the variations in the Earth's mean temperature from one month to the next.



— Earth temperature (deviation to monthly mean)



— Earth temperature (monthly variations)



- ▷ In the autocorrelation function a single coefficient different from zero is observed, and in the PACF a geometric decay is observed.
- ▷ Both graphs suggest an MA(1) model for the series of differences between consecutive months,  $z_t$ .

# The MA( $q$ ) process

▶ Generalizing on the idea of an MA(1), we can write processes whose current value depends not only on the last innovation but on the last  $q$  innovations. Thus the **MA( $q$ ) process** is obtained, with general representation:

$$\tilde{z}_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}.$$

▶ Introducing the operator notation:

$$\tilde{z}_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) a_t \quad (73)$$

it can be written more compactly as:

$$\tilde{z}_t = \theta_q(B) a_t. \quad (74)$$

▶ An MA( $q$ ) is always stationary, as it is a sum of stationary processes. We say that the process is **invertible** if the roots of the operator  $\theta_q(B) = 0$  are, in modulus, greater than the unit.

# The MA( $q$ ) process

► The properties of this process are obtained with the same method used for the MA(1). Multiplying (73) by  $\tilde{z}_{t-k}$  for  $k \geq 0$  and taking expectations, the autocovariances are obtained:

$$\gamma_0 = (1 + \theta_1^2 + \dots + \theta_q^2) \sigma^2 \quad (75)$$

$$\gamma_k = (-\theta_k + \theta_1\theta_{k+1} + \dots + \theta_{q-k}\theta_q) \sigma^2 \quad k = 1, \dots, q, \quad (76)$$

$$\gamma_k = 0 \quad k > q, \quad (77)$$

showing that an MA( $q$ ) process has exactly the first  $q$  coefficients of the autocovariance function different from zero.

► Dividing the covariances by  $\gamma_0$  and utilizing a more compact notation, the autocorrelation function is:

$$\rho_k = \frac{\sum_{i=0}^{i=q} \theta_i \theta_{k+i}}{\sum_{i=0}^{i=q} \theta_i^2}, \quad k = 1, \dots, q \quad (78)$$

$$\rho_k = 0, \quad k > q,$$

where  $\theta_0 = -1$ , and  $\theta_k = 0$  for  $k \geq q + 1$ .



# The MA( $q$ ) process

► To compute the partial autocorrelation function of an MA( $q$ ) we express the process as an AR( $\infty$ ):

$$\theta_q^{-1}(B) \tilde{z}_t = a_t,$$

and letting  $\theta_q^{-1}(B) = \pi(B)$ , where:

$$\pi(B) = 1 - \pi_1 B - \dots - \pi_k B^k - \dots$$

and the coefficients of  $\pi(B)$  are obtained imposing  $\pi(B)\theta_q(B) = 1$ . We say that the process is invertible if all the roots of  $\theta_q(B) = 0$  lie outside the unit circle. Then the series  $\pi(B)$  is convergent.

► For invertible MA processes, setting the powers of  $B$  to zero, we find that the coefficients  $\pi_i$  verify the following equation:

$$\pi_k = \theta_1 \pi_{k-1} + \dots + \theta_q \pi_{k-q}$$

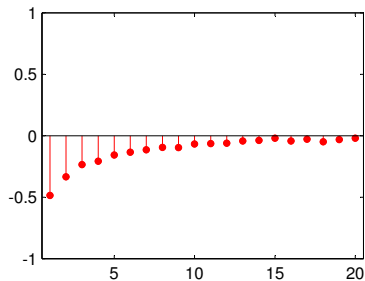
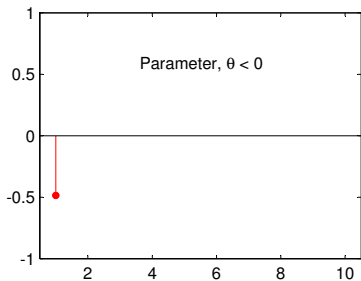
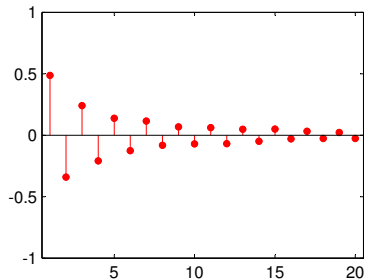
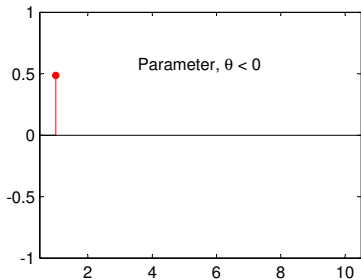
where  $\pi_0 = -1$  and  $\pi_j = 0$  for  $j < 0$ .

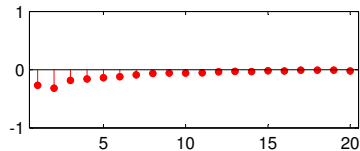
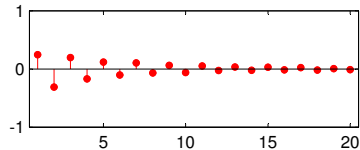
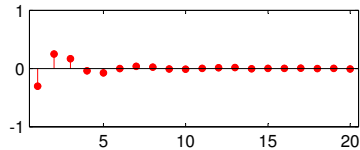
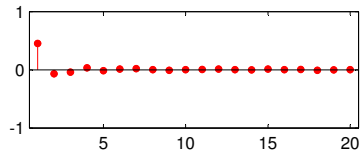
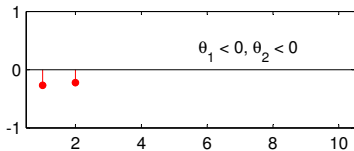
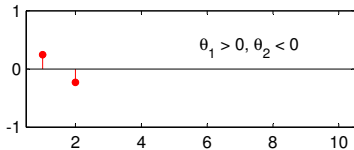
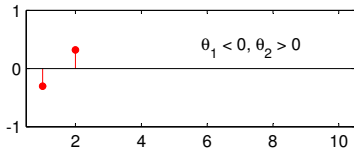
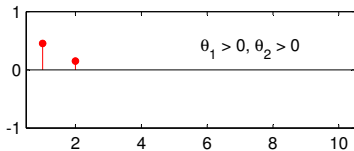
# The MA(q) process

- ▷ The solution to this difference equation is of the form  $\sum A_i G_i^k$ , where now the  $G_i^{-1}$  are the roots of the moving average operator. Having obtained the coefficients  $\pi_i$  of the representation  $AR(\infty)$ , we can write the MA process as:

$$\tilde{z}_t = \sum_{i=1}^{\infty} \pi_i \tilde{z}_{t-i} + a_t.$$

- ▷ From this expression we conclude that the *PACF* of an MA is non-null for all lags, since a direct effect of  $\tilde{z}_{t-i}$  on  $\tilde{z}_t$  exists for all  $i$ . The *PACF* of an MA process thus has the same structure as the *ACF* of an AR process of the same order.
- ▷ We conclude that a duality exists between the AR and MA processes such that the *PACF* of an  $MA(q)$  has the structure of the *ACF* of an  $AR(q)$  and the *ACF* of an  $MA(q)$  has the structure of the *PACF* of an  $AR(q)$ .





# The $MA(\infty)$ process and Wold decomposition

- ▶ The autoregressive and moving average processes are specific cases of a general representation of stationary processes obtained by Wold (1938).
- ▶ Wold proved that any weakly stationary stochastic process,  $z_t$ , with finite mean,  $\mu$ , that does not contain deterministic components, can be written as a linear function of uncorrelated random variables,  $a_t$ , as:

$$z_t = \mu + a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots = \mu + \sum_{i=0}^{\infty} \psi_i a_{t-i} \quad (\psi_0 = 1) \quad (79)$$

where  $E(z_t) = \mu$ , and  $E[a_t] = 0$ ;  $Var(a_t) = \sigma^2$ ;  $E[a_t a_{t-k}] = 0$ ,  $k > 1$ .

- ▶ Letting  $\tilde{z}_t = z_t - \mu$ , and using the lag operator, we can write:

$$\tilde{z}_t = \psi(B)a_t, \quad (80)$$

with  $\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \dots$  being an indefinite polynomial in the lag operator  $B$ .

# The $MA(\infty)$ process and Wold decomposition

- ▶ We denote (80) as the **general linear representation** of a non-deterministic stationary process.
- ▶ This representation is important because it guarantees that any stationary process admits a linear representation.
- ▶ In general, the variables  $a_t$  make up a white noise process, that is, they are uncorrelated with zero mean and constant variance.
- ▶ In certain specific cases the process can be written as a function of normal independent variables  $\{a_t\}$ . Thus the variable  $\tilde{z}_t$  will have a normal distribution and the weak coincides with strict stationarity.
- ▶ The series  $\tilde{z}_t$ , can be considered as the result of passing a process of impulses  $\{a_t\}$  of uncorrelated variables through a linear filter  $\psi(B)$  that determines the weight of each "impulse" in the response.

# The $MA(\infty)$ process and Wold decomposition

- ▶ The properties of the process are obtained as in the case of an MA model. The variance of  $z_t$  in (79) is:

$$Var(z_t) = \gamma_0 = \sigma^2 \sum_{i=0}^{\infty} \psi_i^2 \quad (81)$$

and for the process to have finite variance the series  $\{\psi_i^2\}$  must be convergent.

- ▶ We observe that if the coefficients  $\psi_i$  are zero after lag  $q$  the general model is reduced to an  $MA(q)$  and formula (81) coincides with (76).

- ▶ The covariances are obtained with

$$\gamma_k = E(\tilde{z}_t \tilde{z}_{t-k}) = \sigma^2 \sum_{i=0}^{\infty} \psi_i \psi_{i+k},$$

which for  $k = 0$  provide, as a particular case, formula (81) for the variance.

- ▶ Furthermore, if the coefficients  $\psi_i$  are zero after lag  $q$  on, this expression provides the autocovariances of an  $MA(q)$  expression.

# The $MA(\infty)$ process and Wold decomposition

- ▶ The autocorrelation coefficients are given by:

$$\rho_k = \frac{\sum_{i=0}^{\infty} \psi_i \psi_{i+k}}{\sum_{i=0}^{\infty} \psi_i^2}, \quad (82)$$

which generalizes the expression (78) of the autocorrelations of an  $MA(q)$ .

- ▶ A consequence of (79) is that any stationary process also admits an autoregressive representation, which can be of infinite order. This representation is the inverse of that of Wold, and we write

$$\tilde{z}_t = \pi_1 \tilde{z}_{t-1} + \pi_2 \tilde{z}_{t-2} + \dots + a_t,$$

which in operator notation is reduced to

$$\pi(B) \tilde{z}_t = a_t.$$

- ▶ The  $AR(\infty)$  representation is the dual representation of the  $MA(\infty)$  and it is shown that:  $\pi(B)\psi(B) = 1$  such that by setting the powers of  $B$  to zero we can obtain the coefficients of one representation from those of another.



# The AR and MA processes and the general process

► It is straightforward to prove that an MA process is a particular case of the Wold representation, as are the AR processes.

► For example, the AR(1) process

$$(1 - \phi B) \tilde{z}_t = a_t \quad (83)$$

can be written, multiplying by the inverse operator  $(1 - \phi B)^{-1}$

$$\tilde{z}_t = (1 + \phi B + \phi^2 B^2 + \dots) a_t$$

which represents the AR(1) process as a particular case of the MA( $\infty$ ) form of the general linear process, with coefficients  $\psi_i$  that decay in geometric progression.

► The condition of stationarity and finite variance, convergent series of coefficients  $\psi_i^2$ , is equivalent now to  $|\phi| < 1$ .

# The AR and MA processes and the general process

- ▶ For higher order AR process to obtain the coefficients of the  $MA(\infty)$  representation we impose the condition that the product of the AR and  $MA(\infty)$  operators must be the unit.
- ▶ For example, for an  $AR(2)$  the condition is:

$$(1 - \phi_1 B - \phi_2 B^2) (1 + \psi_1 B + \psi_2 B^2 + \dots) = 1$$

and imposing the cancelation of powers of  $B$  we obtain the coefficients:

$$\psi_1 = \phi_1$$

$$\psi_2 = \phi_1 \psi_1 + \phi_2$$

$$\psi_i = \phi_1 \psi_{i-1} + \phi_2 \psi_{i-2}, \quad i \geq 2$$

where  $\psi_0 = 1$ .

# The AR and MA processes and the general process

▶ Analogously, for an  $AR(p)$  the coefficients  $\psi_i$  of the general representation are calculated by:

$$(1 - \phi_1 B - \dots - \phi_p B^p)(1 + \psi_1 B + \psi_2 B^2 + \dots) = 1$$

and for  $i \geq p$  they must verify the condition:

$$\psi_i = \phi_1 \psi_{i-1} + \dots + \phi_p \psi_{i-p}, \quad i \geq p.$$

▶ The condition of stationarity implies that the roots of the characteristic equation of the  $AR(p)$  process,  $\phi_p(B) = 0$ , must lie outside the unit circle.

# The AR and MA processes and the general process

▷ Writing the operator  $\phi_p(B)$  as:

$$\phi_p(B) = \prod_{i=1}^p (1 - G_i B)$$

where  $G_i^{-1}$  are the roots of  $\phi_p(B) = 0$ , it is shown that, expanding in partial fractions:

$$\phi_p^{-1}(B) = \sum \frac{k_i}{(1 - G_i B)}$$

will be convergent if  $|G_i| < 1$ .

▷ Summarizing, the AR processes can be considered as particular cases of the general linear process characterized by the fact that: (1) all the  $\psi_i$  are different from zero; (2) there are restrictions on the  $\psi_i$ , that depend on the order of the process.

▷ In general they verify the sequence  $\psi_i = \phi_1 \psi_{i-1} + \dots + \phi_p \psi_{i-p}$ , with initial conditions that depend on the order of the process.

# The ARMA(1,1) process

▷ One conclusion from the above section is that the AR and MA processes approximate a general linear  $MA(\infty)$  process from a complementary point of view:

- The AR admit an  $MA(\infty)$  structure, but they impose restrictions on the decay patterns of the coefficients  $\psi_i$ .
- The MA require a number of finite terms, however, they do not impose restrictions on the coefficients.
- From the point of view of the autocorrelation structure, the AR processes allow many coefficients different from zero, but with a fixed decay pattern, whereas the MA permit a few coefficients different from zero with arbitrary values.

▷ The **ARMA processes** try to combine these properties and allow us to represent in a *reduced* form (using few parameters) those processes whose first  $q$  coefficients can be any, whereas the following ones decay according to simple rules.

# The ARMA(1,1) process

- ▶ The simplest process, the **ARMA(1,1)** is written as:

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + a_t - \theta_1 a_{t-1},$$

or, using operator notations:

$$(1 - \phi_1 B) \tilde{z}_t = (1 - \theta_1 B) a_t, \quad (84)$$

where  $|\phi_1| < 1$  for the process to be **stationary**, and  $|\theta_1| < 1$  for it to be **invertible**.

- ▶ Moreover, we assume that  $\phi_1 \neq \theta_1$ . If both parameters were identical, multiplying both parts by the operator  $(1 - \phi_1 B)^{-1}$ , we would have  $\tilde{z}_t = a_t$ , and the process would be white noise.

- ▶ In the formulation of the ARMA models we always assume that there are no common roots in the AR and MA operators.

# The ARMA(1,1) process

## The autocorrelation function

- ▶ To obtain the autocorrelation function of an ARMA(1,1), multiplying (84) by  $\tilde{z}_{t-k}$  and taking expectations, results in:

$$\gamma_k = \phi_1 \gamma_{k-1} + E(a_t \tilde{z}_{t-k}) - \theta_1 E(a_{t-1} \tilde{z}_{t-k}). \quad (85)$$

- ▶ For  $k > 1$ , the noise  $a_t$  is uncorrelated with the series history. As a result:

$$\gamma_k = \phi_1 \gamma_{k-1}, \quad k > 1. \quad (86)$$

- ▶ For  $k = 0$ ,  $E[a_t \tilde{z}_t] = \sigma^2$  and

$$E[a_{t-1} \tilde{z}_t] = E[a_{t-1} (\phi_1 \tilde{z}_{t-1} + a_t - \theta_1 a_{t-1})] = \sigma^2(\phi_1 - \theta_1)$$

replacing these results in (85), for  $k = 0$

$$\gamma_0 = \phi \gamma_1 + \sigma^2 - \theta_1 \sigma^2 (\phi_1 - \theta_1). \quad (87)$$

# The ARMA(1,1) process

## The autocorrelation function

▷ Taking  $k = 1$  in (85), results in  $E[a_t \tilde{z}_{t-1}] = 0$ ,  $E[a_{t-1} \tilde{z}_{t-1}] = \sigma^2$  and:

$$\gamma_1 = \phi_1 \gamma_0 - \theta_1 \sigma^2, \quad (88)$$

solving for (87) and (88) we obtain:

$$\gamma_0 = \sigma^2 \frac{1 - 2\phi_1\theta_1 + \theta_1^2}{1 - \phi_1^2}$$

▷ To compute the first autocorrelation coefficient, we divide (88) by the above expression:

$$\rho_1 = \frac{(\phi_1 - \theta_1)(1 - \phi_1\theta_1)}{1 - 2\phi_1\theta_1 + \theta_1^2} \quad (89)$$

▷ Observe that if  $\phi_1 = \theta_1$ , this autocorrelation is zero because, as we indicated earlier, then the operators  $(1 - \phi_1 B)$  and  $(1 - \theta_1 B)$  are cancelled out and it will result in a white noise process.



# The ARMA(1,1) process

## The autocorrelation function

- ▶ In the typical case where both coefficients are positive and  $\phi_1 > \theta_1$  it is easy to prove that the correlation increases with  $(\phi_1 - \theta_1)$ .
- ▶ The rest of the autocorrelation coefficients are obtained dividing (86) by  $\gamma_0$ , which results in:

$$\rho_k = \phi_1 \rho_{k-1} \quad k > 1 \quad (90)$$

which indicates that from the first coefficient on, the *ACF* of an ARMA(1,1) decays exponentially, determined by parameter  $\phi_1$  of the AR part.

- ▶ The difference with an AR(1) is that the decay starts at  $\rho_1$ , not at  $\rho_0 = 1$ , and this first value of the first order autocorrelation depends on the relative difference between  $\phi_1$  and  $\theta_1$ . We observe that if  $\phi_1 \approx 1$  and  $\phi_1 - \theta_1 = \varepsilon$  is small, we can have many coefficients different from zero but they will all be small.

# The ARMA(1,1) process

## The partial autocorrelation function

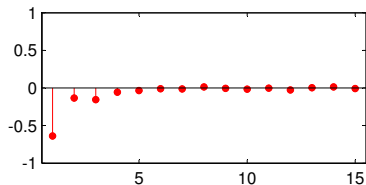
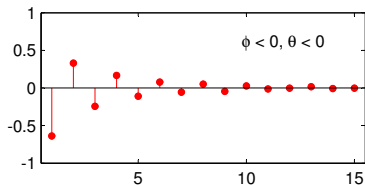
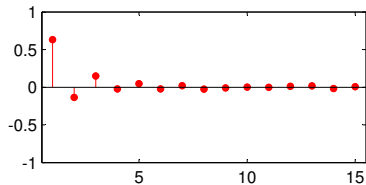
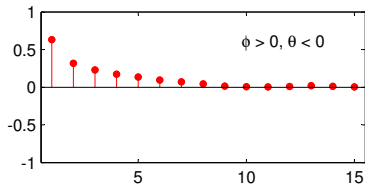
- To calculate the *PACF*, we write the ARMA(1, 1) in the AR( $\infty$ ) form:

$$(1 - \theta_1 B)^{-1} (1 - \phi_1 B) \tilde{z}_t = a_t,$$

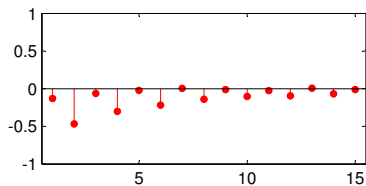
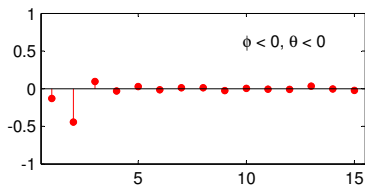
and using  $(1 - \theta_1 B)^{-1} = 1 + \theta_1 B + \theta_1^2 B^2 + \dots$ , and operating, we obtain:

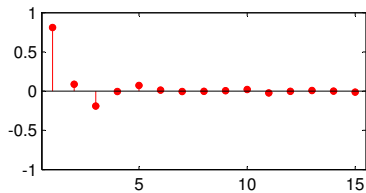
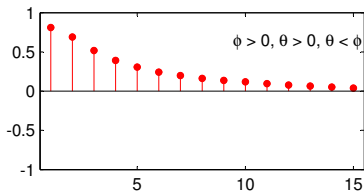
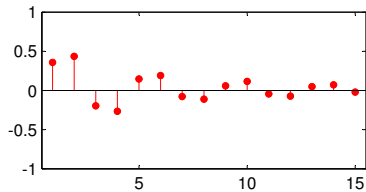
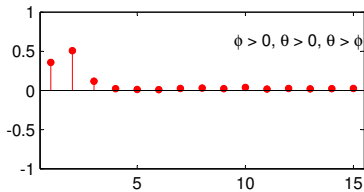
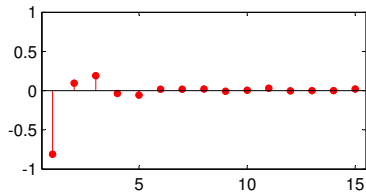
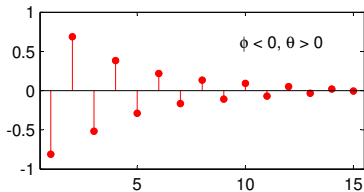
$$\tilde{z}_t = (\phi_1 - \theta_1) \tilde{z}_{t-1} + \theta_1 (\phi_1 - \theta_1) \tilde{z}_{t-2} + \theta_1^2 (\phi_1 - \theta_1) \tilde{z}_{t-3} + \dots + a_t.$$

- The direct effect of  $\tilde{z}_{t-k}$  on  $\tilde{z}_t$  decays geometrically with  $\theta_1^k$  and, therefore, the *PACF* will have a geometric decay starting from an initial value.
- In conclusion, in an ARMA(1,1) process the *ACF* and the *PACF* have a similar structure: an initial value, whose magnitude depends on  $\phi_1 - \theta_1$ , followed by a geometric decay.
- The rate of decay in the *ACF* depends on  $\phi_1$ , whereas in the *PACF* it depends on  $\theta_1$ .



Lag





# The ARMA(p,q) processes

- ▷ The **ARMA** (p, q) process is defined by:

$$(1 - \phi_1 B - \dots - \phi_p B^p) \tilde{z}_t = (1 - \theta_1 B - \dots - \theta_q B^q) a_t \quad (91)$$

or, in compact notation,

$$\phi_p(B) \tilde{z}_t = \theta_q(B) a_t.$$

- ▷ The process is **stationary** if the roots of  $\phi_p(B) = 0$  are outside the unit circle, and **invertible** if those of  $\theta_q(B) = 0$  are.
- ▷ We also assume that there are no common roots that can be cancelled between the AR and MA operators.

# The ARMA(p,q) processes

- ▶ To obtain the coefficients  $\psi_i$  of the general representation of the MA( $\infty$ ) model we write:

$$\tilde{z}_t = \phi_p(B)^{-1} \theta_q(B) a_t = \psi(B) a_t$$

and we equate the powers of  $B$  in  $\psi(B) \phi_p(B)$  to those of  $\theta_q(B)$ .

- ▶ Analogously, we can represent an ARMA( $p, q$ ) as an AR( $\infty$ ) model making:

$$\theta_q^{-1}(B) \phi_p(B) \tilde{z}_t = \pi(B) \tilde{z}_t = a_t$$

and the coefficients  $\pi_i$  will be the result of  $\phi_p(B) = \theta_q(B) \pi(B)$ .

# The ARMA(p,q) processes

## Autocorrelation function

► To calculate the autocovariances, we multiply (91) by  $\tilde{z}_{t-k}$  and take expectations,

$$\begin{aligned} & \gamma_k - \phi_1 \gamma_{k-1} - \dots - \phi_p \gamma_{k-p} = \\ & = E[a_t \tilde{z}_{t-k}] - \theta_1 E[a_{t-1} \tilde{z}_{t-k}] - \dots - \theta_q E[a_{t-q} \tilde{z}_{t-k}] \end{aligned}$$

► For  $k > q$  all the terms on the right are cancelled, and dividing by  $\gamma_0$ :

$$\rho_k - \phi_1 \rho_{k-1} - \dots - \phi_p \rho_{k-p} = 0,$$

that is:

$$\phi_p(B) \rho_k = 0 \quad k > q, \quad (92)$$

► We conclude that the autocorrelation coefficients for  $k > q$  follow a decay determined only in the autoregressive part.

# The ARMA( $p,q$ ) processes

## Autocorrelation function

- ▶ The first  $q$  coefficients depend on the MA and AR parameters and of those,  $p$  provide the initial values for the later decay (for  $k > q$ ) according to (92). Therefore, if  $p > q$  all the  $ACF$  will show a decay dictated by (92).
- ▶ To summarize, the  $ACF$ :
  - have  $q - p + 1$  initial values with a structure that depends on the AR and MA parameters;
  - they decay starting from the coefficient  $q - p$  as a mixture of exponentials and sinusoids, determined exclusively by the autoregressive part.
- ▶ It can be proved that the  $PACF$  have a similar structure.



# Summary

► The *ACF* and *PACF* of the ARMA processes are the result of superimposing their AR and MA properties:

- In the *ACF* certain initial coefficients that depend on the order of the MA part and later a decay dictated by the AR part.
- In the *PACF* initial values dependent on the AR order followed by the decay due to the MA part.
- This complex structure makes it difficult in practice to identify the order of an ARMA process.

	<b>ACF</b>	<b>PACF</b>
AR(p)	Many non-null coefficients	first p non-null, the rest 0
MA(q)	first q non-null, the rest 0	Many non-null coefficients
ARMA(p,q)	Many non-null coefficients	Many non-null coefficients

# ARMA processes and the sum of stationary processes

▶ One reason that explains why the ARMA processes are frequently found in practice is that summing AR processes results in an ARMA process.

▶ To illustrate this idea, we take the simplest case where we add white noise to an AR(1) process. Let

$$z_t = y_t + v_t \quad (93)$$

where  $y_t = \phi y_{t-1} + a_t$  follows an AR(1) process of zero mean and  $v_t$  is white noise independent of  $a_t$ , and thus of  $y_t$ .

▶ Process  $z_t$  can be interpreted as the result of observing an AR(1) process with a certain measurement error. The variance of this addition process is:

$$\gamma_z(0) = E(z_t^2) = E[(y_t^2 + v_t^2 + 2y_tv_t)] = \gamma_y(0) + \sigma_v^2, \quad (94)$$

since, as the summands are independent, the variance is the sum of the variance of the components.

# ARMA processes and the sum of stationary processes

► To calculate the autocovariance we take into account that the autocovariances of process  $y_t$  verify  $\gamma_y(k) = \phi^k \gamma_y(0)$  and those of process  $v_t$  are null. Thus,  $k \geq 1$ ,

$$\gamma_z(k) = E(z_t z_{t-k}) = E[(y_t + v_t)(y_{t-k} + v_{t-k})] = \gamma_y(k) = \phi^k \gamma_y(0),$$

since, due to the independence of the components,  $E[y_t v_{t-k}] = 0$  for any  $k$  and since  $v_t$  is white noise  $E[v_t v_{t-k}] = 0$ . Specifically, replacing the variance  $\gamma_y(0)$  with its expression (94) for  $k = 1$ , we obtain:

$$\gamma_z(1) = \phi \gamma_z(0) - \phi \sigma_v^2, \quad (95)$$

whereas for  $k \geq 2$

$$\gamma_z(k) = \phi \gamma_z(k-1). \quad (96)$$

► If we compare equation (95) with (88), and equation (96) with (86) we conclude that process  $z_t$  follows an ARMA(1,1) model with an AR parameter equal to  $\phi$ . Parameter  $\theta$  and the variance of the innovations of the ARMA(1,1) depend on the relationship between the variances of the summands.

# ARMA processes and the sum of stationary processes

▷ Indeed, letting  $\lambda = \sigma_v^2 / \gamma_y(0)$  denote the quotient of variances between the two summands, according to equation (95) the first autocorrelation is:

$$\rho_z(1) = \phi - \phi \frac{\lambda}{1 + \lambda}$$

whereas by (96) the remainders verify, for  $k \geq 2$ ,

$$\rho_z(k) = \phi \rho_z(k-1).$$

▷ If  $\lambda$  is very small, which implies that the variance of the additional noise or measurement error is small, the process will be very close to an AR(1), and parameter  $\theta$  will be very small.

▷ If  $\lambda$  is not very small, we have the ARMA(1,1) and the value of  $\theta$  depends on  $\lambda$  and on  $\phi$ .

▷ If  $\lambda \rightarrow \infty$ , such that the white noise is dominant, the parameter  $\theta$  will be equal to the value of  $\phi$  and we have a white noise process.

# ARMA processes and the sum of stationary processes

- ▶ The above results can be generalized for any  $AR(p)$  process. It can be proved that:

$$AR(p) + AR(0) = ARMA(p, p),$$

and also that:

$$AR(p) + AR(q) = ARMA(p + q, \max(p, q))$$

- ▶ For example, if we add two independent  $AR(1)$  processes we obtain a new process,  $ARMA(2,1)$ .
- ▶ The sum of MA processes is simple: by adding independent MA processes we obtain new MA processes.

# ARMA processes and the sum of stationary processes

- Let us assume that

$$z_t = x_t + y_t$$

where the two processes  $x_t$ ,  $y_t$  have zero mean and follow independent MA(1) processes with covariances  $\gamma_x(k)$ ,  $\gamma_y(k)$ , that are zero for  $k > 1$ .

- The variance of the summed process is:

$$\gamma_z(0) = \gamma_x(0) + \gamma_y(0), \quad (97)$$

and the autocovariance of order  $k$

$$E(z_t z_{t-k}) = \gamma_z(k) = E[(x_t + y_t)(x_{t-k} + y_{t-k})] = \gamma_x(k) + \gamma_y(k).$$

- Therefore, all the covariances  $\gamma_z(k)$  of order higher than one will be zero because  $\gamma_x(k)$  and  $\gamma_y(k)$  are zero.

# ARMA processes and the sum of stationary processes

▷ Dividing the equation (44) by  $\gamma_z(0)$  and using (97), shows that the autocorrelations verify:

$$\rho_z(k) = \rho_x(k)\lambda + \rho_y(k)(1 - \lambda)$$

where:

$$\lambda = \frac{\gamma_x(0)}{\gamma_x(0) + \gamma_y(0)}$$

is the relative variance of the first summand.

▷ In the particular case in which one of the processes is white noise we obtain an MA(1) model whose autocorrelation is smaller than that of the original process. In the same way it is easy to show that:

$$MA(q_1) + MA(q_2) = MA(\max(q_1, q_2)).$$

# ARMA processes and the sum of stationary processes

- ▶ For ARMA processes it is also proved that:

$$ARMA(p_1, q_1) + ARMA(p_2, q_2) = ARMA(a, b)$$

where

$$a \leq p_1 + p_2, \quad b \leq \max(p_1 + q_1, p_2 + q_2)$$

- ▶ These results suggest that whenever we observe processes that are the sum of others, and some of them have an AR structure, we expect to observe ARMA processes.

- ▶ This result may seem surprising at first because the majority of real series can be considered to be the sum of certain components, which would mean that all real processes should be ARMA.

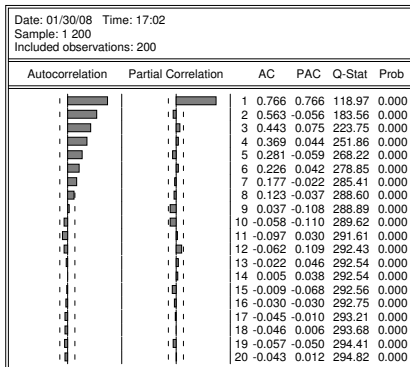
- ▶ Nevertheless, in practice many real series are approximated well by means of AR or MA series. The explanation for this paradox is that an  $ARMA(q + h, q)$  process with  $q$  similar roots in the AR and MA parts can be well approximated by an  $AR(h)$ , due to the near cancellation of similar roots in both members.



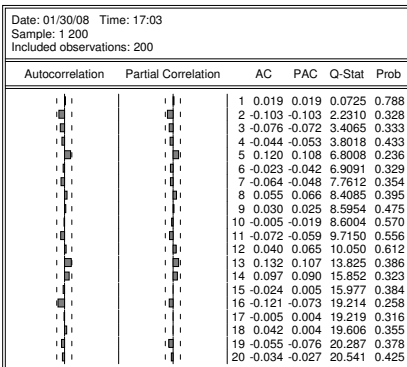
## Example 43

The figures show the autocorrelation functions of an  $AR(1)$  and an  $AR(0)$ .

Correlogram of AR1



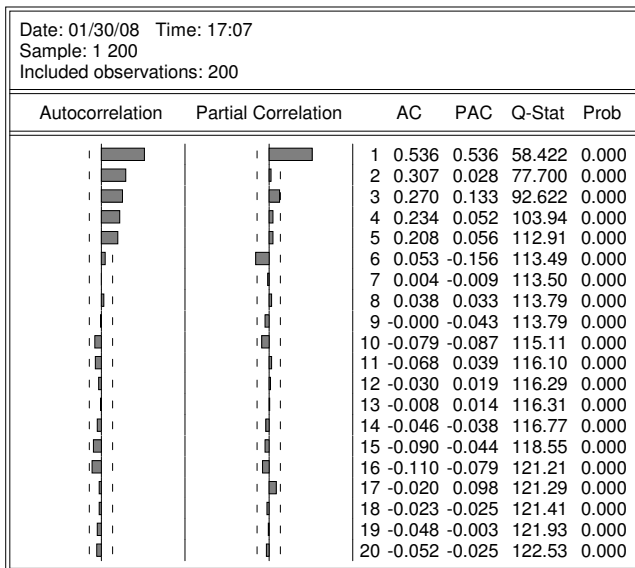
Correlogram of E2



Datafile sumofst.xls

► The figure shows the autocorrelation functions of the sum of  $AR(1)+AR(0)$ .

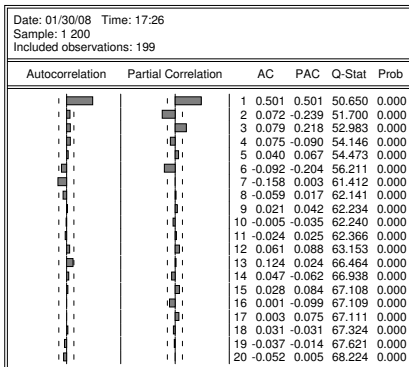
Correlogram of SUM1



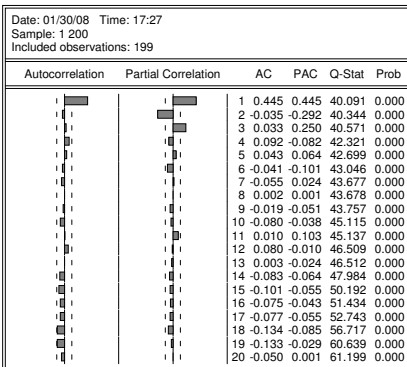
## Example 44

*The figures show the autocorrelation functions of two MA(1).*

Correlogram of MA1A



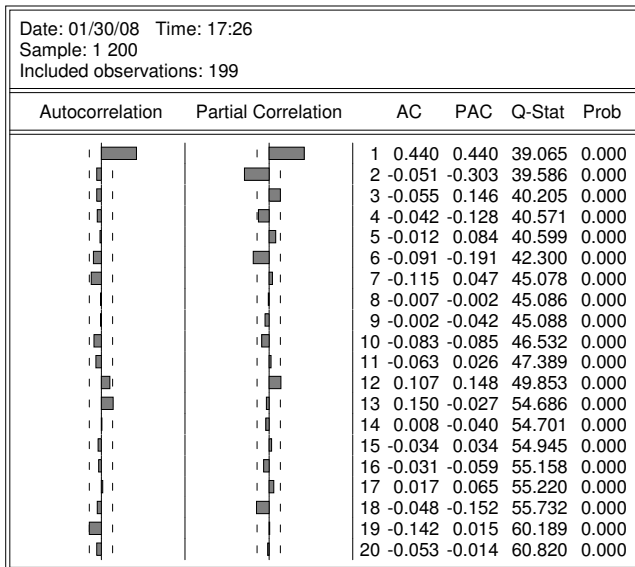
Correlogram of MA1B



Datafile sumofst.xls

► The figure shows the autocorrelation functions of the sum of MA(1)+MA(1).

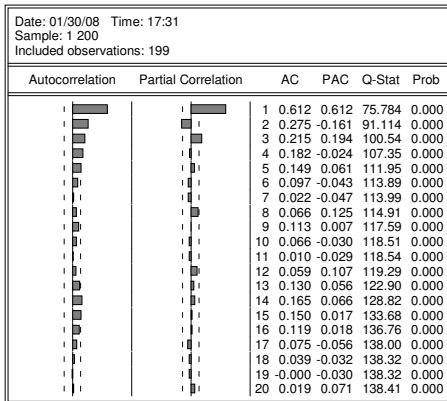
Correlogram of SUM2



## Example 45

*The figures show the autocorrelation functions of the two sum of  $AR(1)+MA(1)$ .*

Correlogram of SUM3



Correlogram of SUM4

