6. Seasonal ARIMA processes

Outline:

- Introduction
- The concept and types of seasonality
- The ARIMA seasonal model
- Simple autocorrelation function
- Partial autocorrelation function
- Generalizations

Recommended readings:

▷ Chapter 7 of D. Peña (2008).

▷ Chapter 6 of P.J. Brockwell and R.A. Davis (1996).

Introduction

▷ In this section we continue the study of non-stationary processes, analyzing a type of lack of stationarity in the mean that is frequently found in practice: seasonal behavior.

▷ Seasonality makes it so that the mean of the observations is not constant, but instead evolves according to a cyclical pattern:

- For example, in a series of monthly temperatures in Europe the mean temperature is not constant, since it varies by month, but for the same month in different years we can expect a constant average value.

▷ The most typical case is that we can incorporate seasonality into the ARIMA model multiplicatively, so that we obtain a multiplicative seasonal ARIMA model.

The concept and types of seasonality

▷ We say that a series which has no trend is seasonal when its expected value is not constant, but varies in a cyclical pattern. Specifically if

$$E(z_t) = E(z_{t+s})$$

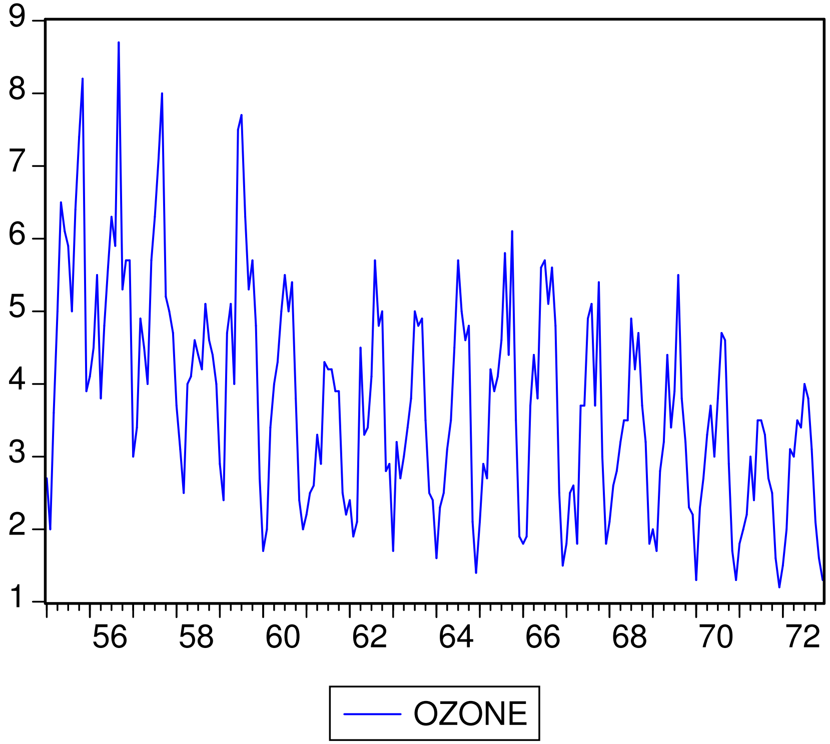
we say that the series has **seasonality** of period s.

- For example, a monthly series with no trend has seasonality if the expected values in different months of the year are different, but the expected value for the same month in different years is the same.

▷ In series with trend the seasonality is an additional cause of non-stationarity. The seasonal pattern is superimposed on the global trend, producing cyclical behavior that is repeated in the different years of the sample.

The concept and types of seasonality - Example

Example 58. *The figure shows the series of ozone contents in the atmosphere. It exhibit a very strong seasonal pattern, with peaks and valleys in the same months of the year. Moreover, the seasonal pattern is superimposed on the decreasing global trend.*



The concept and types of seasonality

▷ The **seasonal period**, s, defines the number of observations that make up a seasonal cycle. For example, $s = 12$ for monthly series, $s = 4$ for quarterly series, etc.

▷ We assume that the value of s is fixed in the series.

▷ However, this may not be precisely true: for example, if we have daily data and the seasonal period is the length of the month, s will be approximately 30, but it will vary from month to month.

▷ There may be more than one type of seasonality. For example, with daily data we can have weekly seasonality, with $s = 7$, monthly, with $s = 30$ and yearly, with $s = 365$.

▷ Initially we will assume that there is only one type of seasonality and at the end of this section we will see how to extend the methods presented here to various seasonal periods.

The concept and types of seasonality

▷ The simplest model for seasonality is when it is modelled as a constant effect that is added to the values of the series. For example, let us assume a series that, except for its seasonal effect, is stationary. We can write the series as a sum of a seasonal component $S_t^{(s)}$, and a stationary process, n_t, in such a way that the model for the series is:

$$z_t = S_t^{(s)} + n_t. \tag{117}$$

▷ This series is not stationary, since if we take expectations

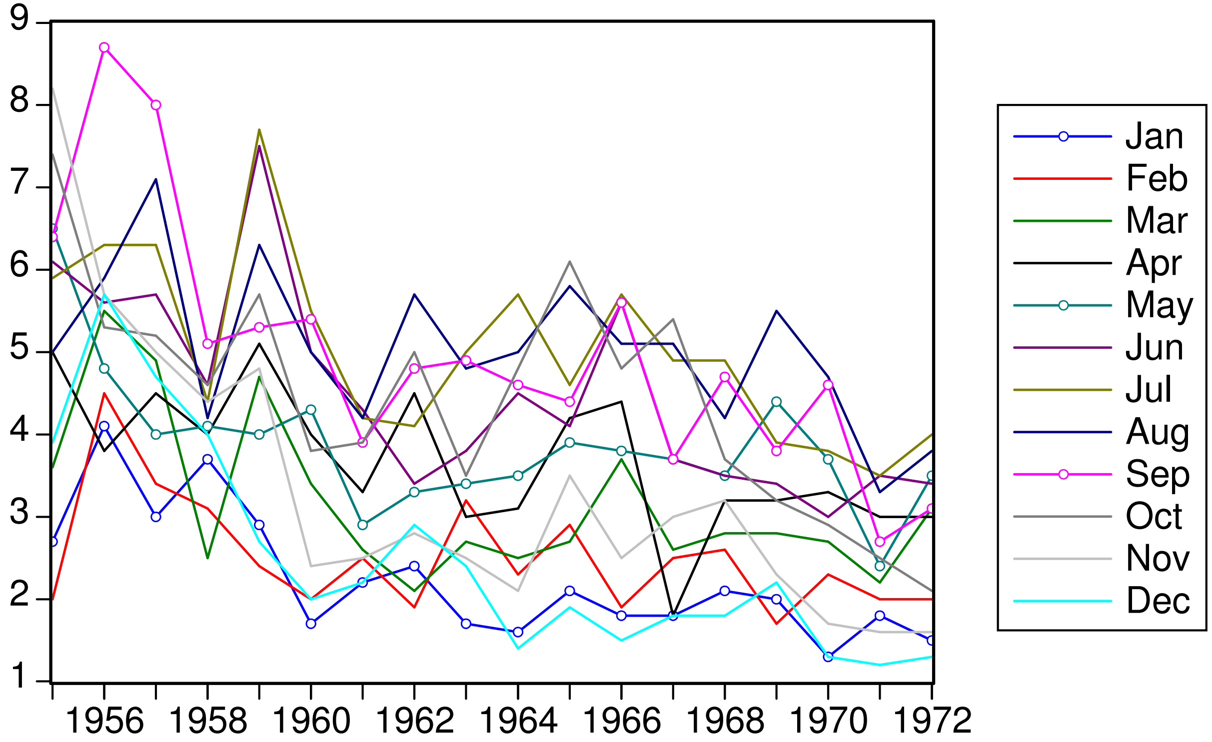
$$E(z_t) = E(S_t^{(s)}) + \mu,$$

where μ is the mean of the process n_t.

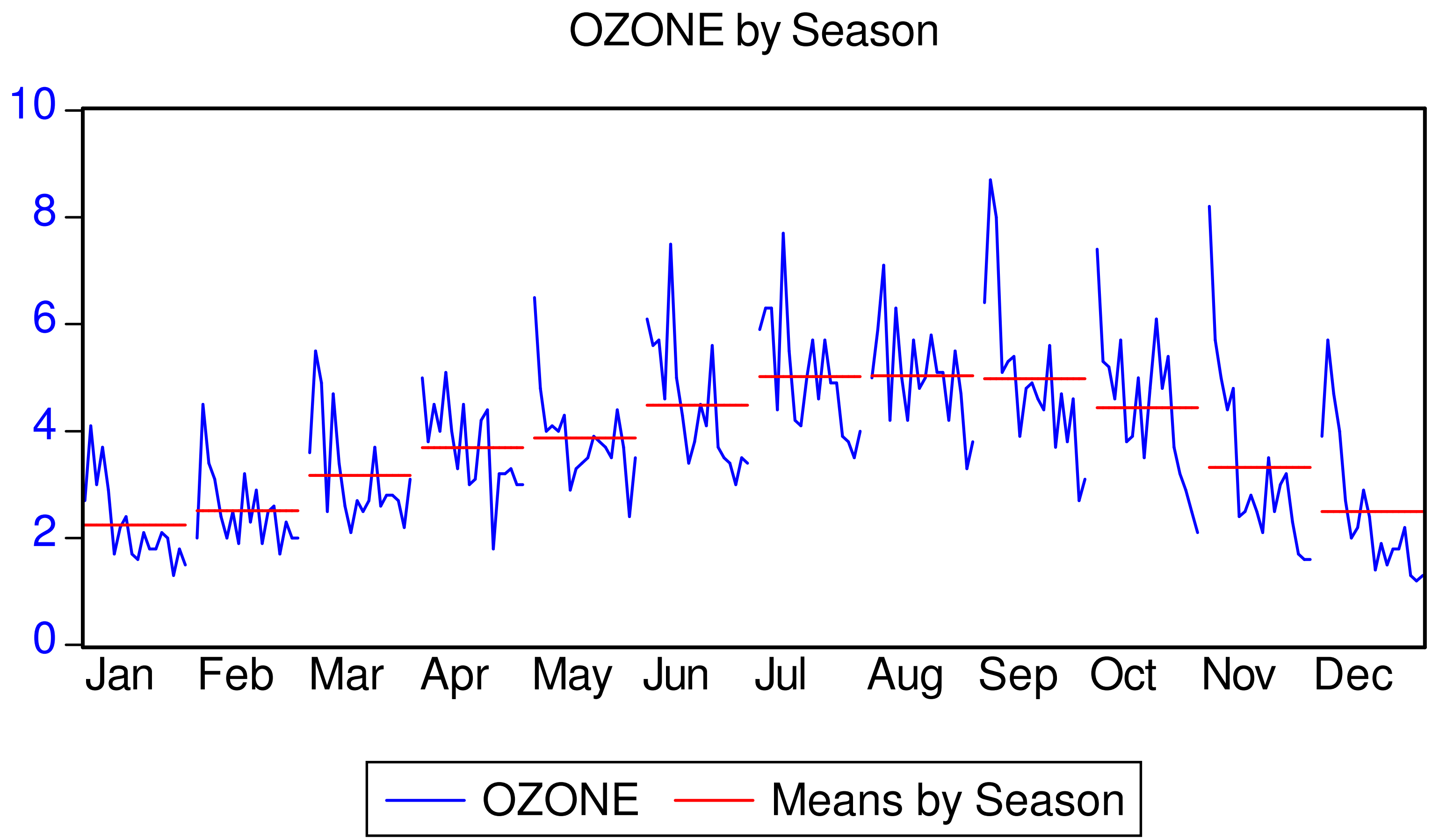
▷ Since by definition the seasonal component does not take the same value in all of the periods, the series is not stationary because it does not have a constant mean.

The concept and types of seasonality - Example

Example 59. *The figure shows the evolution of the ozone concentration by months. Each of the series represents the ozone content in that month in different years of the sample. Notice that the ozone content in January is always less than in May, and May is generally less than September.*



The concept and types of seasonality - Example



▷ Therefore, each month has a different average behavior, which is what characterizes a seasonal series.

The concept and types of seasonality

▷ We can consider different hypotheses regarding the behavior of the seasonal process $S_t^{(s)}$.

▷ The first is that seasonality is a deterministic process, that is, a constant function for the same month in different years:

$$S_t^{(s)} = S_{t+ks}^{(s)} \quad k = \pm 1, \pm 2, \ldots \tag{118}$$

▷ For example, seasonal coefficients can follow a sinusoidal function, which occurs in climate series due to the Earth's rotation. This suggests treating seasonality with sinusoidal functions but these functions are not very efficient when the seasonality follows a deterministic, but not sinusoidal, pattern.

▷ Representing this seasonality using sinusoidal functions would be very inefficient.

▷ Deterministic seasonality can always be modeled by introducing 11 dummy variables, one for each month of the year as you will see in second module.

The concept and types of seasonality

▷ Nevertheless, many time series do not have deterministic seasonality.

▷ Instead, the seasonal pattern, like other properties, also evolves over time.

▷ The second way to model seasonality is to assume that the evolution is stationary, that is, the seasonal factors are not constant, but follow a stationary process, oscillating around an average value in accordance with the representation:

$$S_t^{(s)} = \mu^{(s)} + v_t \tag{119}$$

where $\mu^{(s)}$ is a constant that depends on the month and represents the deterministic effect of the seasonality and v_t is a stationary process of zero mean that introduces variability into the seasonality of each year.

The concept and types of seasonality

▷ The third way to model seasonality is to allow it to change over time with no fixed average value. In this case, seasonality follows a non-stationary process. For example, the simplest model is to assume that it evolves according to a random walk:

$$S_t^{(s)} = S_{t-s}^{(s)} + v_t \tag{120}$$

where v_t is a stationary process of zero mean. Of course, more complicated situations are possible and the seasonality may follow any non-stationary process.

▷ We are going to prove that, in the three cases presented here, we can change a seasonal series into a stationary one by applying a **seasonal difference**. We define the seasonal difference operator of period s as:

$$\nabla_s = 1 - B^s$$

▷ Note that $\nabla_s \neq \nabla^s = (1-B)^s$).

The concept and types of seasonality

▷ If we apply this operator to a series we obtain a transformed series which is the result of replacing at each point in time, t, the value of the series with the difference between the value at time, t, and the value of the series at time $t-s$:

$$\nabla_s z_t = (1 - B^s) z_t = z_t - z_{t-s}.$$

▷ Therefore, if we apply this operator ∇_s in (117), we have:

$$\nabla_s z_t = \nabla_s S_t^{(s)} + \nabla_s n_t$$

and we are going to prove that the series $\nabla_s z_t$ is then stationary.

▷ Let us consider the three cases we have studied.

The concept and types of seasonality

Case ①: The component $S_t^{(s)}$ is deterministic.

▷ Thus, according to (118),

$$\nabla_s S_t^{(s)} = S_t^{(s)} - S_{t-s}^{(s)} = 0$$

and

$$\nabla_s z_t = \nabla_s n_t.$$

▷ By differentiating a stationary process, $\nabla_s n_t$, we always obtain another stationary process, then the difference ∇_s transforms the non-stationary series z_t given by (117) into a stationary one.

▷ We observe that in this case the model for $\nabla_s z_t$ has a moving average factor $(1 - \Theta B^s)$ with $\Theta = 1$.

The concept and types of seasonality

Case ②: The component $S_t^{(s)}$ follows a stationary process.

▷ Then, taking a seasonal difference in (119), since $\nabla_s S_t^{(s)} = \nabla_s v_t$ is stationary, we again have a stationary process, with:

$$\nabla_s z_t = \nabla_s v_t + \nabla_s n_t$$

▷ Using the rules for summing two MA(1) processes, we conclude that the sum process, $\nabla_s z_t$, is an MA(1) in operator B^s with coefficient $\Theta = 1$.

The concept and types of seasonality

Case ③: The component $S_t^{(s)}$ follows a non-stationary process.

▷ If we assume the simplest case of model (120), we have:

$$\nabla_s z_t = v_t + \nabla_s n_t$$

and, again, $\nabla_s z_t$ is a stationary process sum of two terms: the first is white noise and the second an MA(1) of type $(1 - \Theta B^{12})$ with $\Theta = 1$.

▷ Since the autocorrelation of the sum process is a weighted mean of the autocorrelations of the summands, the sum will yield an MA(1), $(1 - \Theta B^{12})$, with an invertible moving average coefficient $\Theta < 1$.

▷ In conclusion, in fairly general conditions for the seasonal structure, and for processes with deterministic as well as stochastic seasonality, the operator ∇_s transforms a seasonal process into a stationary one.

The ARIMA seasonal model

▷ We have seen that we can convert non-stationary series into stationary ones by taking regular differences, that is, the difference from one period with respect to the next.

▷ We also saw that we can eliminate seasonality by means of seasonal differences.

▷ Combining both results we conclude that, in general, we can convert a non-stationary series with seasonality into a stationary one by using the transformation:

$$w_t = \nabla_s^D \nabla^d z_t,$$

where D is the number of seasonal differences (if there is seasonality we almost always have $D = 1$, if there is no seasonality $D = 0$) and d is the number of regular differences $(d \leq 3)$.

The ARIMA seasonal model

▷ When seasonal dependence exists we can generalize the ARMA model for stationary series incorporating both the regular dependence, which is that associated with the measurement intervals of the series, as well as the seasonal dependence, which is that associated with observations separated by s periods.

▷ We will discuss next how to model these two types of dependence:

- The first solution is to incorporate seasonal dependence into the regular, adding B^s terms to the AR or MA operators in the B operator, in order to represent the dependence between observations separated by s periods.

- The inconvenience of this formulation is that it would lead to very large polynomials in the AR and MA part:

  ◇ For example, with monthly data, $s = 12$, if a month is related to the same month in three previous years we need an AR or MA of order 36 to represent this dependence.

The ARIMA seasonal model

▷ A simpler approach, and one which works well in practice, is to model the regular and seasonal dependence separately, and then construct the model incorporating both multiplicatively. Thus a **multiplicative seasonal ARIMA model** is obtained which has the form:

$$\Phi_P\left(B^s\right)\phi_p\left(B\right)\nabla_s^D\nabla^d z_t = \theta_q\left(B\right)\Theta_Q\left(B^s\right)a_t \qquad (121)$$

where

- $\Phi_P\left(B^s\right) = \left(1 - \Phi_1 B^s - \ldots - \Phi_P B^{sP}\right)$ is the seasonal AR operator of order P;
- $\phi_p = \left(1 - \phi_1 B - \ldots - \phi_p B^P\right)$ is the regular AR operator of order p;
- $\nabla_s^D = (1 - B^s)^D$ represents the seasonal differences and $\nabla^d = (1 - B)^d$ the regular differences;
- $\Theta_Q\left(B^s\right) = \left(1 - \Theta_1 B^s - \ldots - \Theta_Q B^{sQ}\right)$ is the seasonal moving average operator of order Q;
- $\theta_q\left(B\right) = (1 - \theta_1 B - \ldots - \theta_q B^q)$ is the regular moving average operator of order q;
- a_t is a white noise process.

The ARIMA seasonal model

▷ This class of models, introduced by Box and Jenkins (1976), offers a good representation of many seasonal series that we find in practice and we will write model (121) in simplified form as the ARIMA model $(P, D, Q)_s \times (p, d, q)$.

▷ To justify (121) let us consider a seasonal series z_t with period s. We denote these yearly series by $y_\tau^{(j)}$, where $j = 1, ..., 12$ indicates the month that defines the series and the time index of the series, τ, is the year, which varies between $\tau = 1, ..., h$.

▷ Indeed, these series always refer to the same month, j, and relate the value of this month in one year to the month in previous years. The way to obtain these annual series starting from the original monthly series, z_t, is by means of:

$$y_\tau^{(j)} = z_{j+12(\tau-1)} \qquad (\tau = 1, ..., h) \tag{122}$$

The ARIMA seasonal model

▷ Now let us assume that each of these 12 yearly series that follow an ARIMA model:

$$\Phi^{(j)}(B)\,(1-B)^D\, y_\tau^{(j)} = c_j + \Theta^{(j)}(B)u_\tau^{(j)}; \qquad \tau = 1, ..., h \tag{123}$$

where

$$\Phi^{(j)}(B) = \left(1 - ... - \Phi_P^{(j)} B^P\right) \text{ and } \Theta^{(j)}(B) = \left(1 - ... - \Theta_Q^{(j)} B^Q\right)$$

have the same common model, i.e., they do not depend on j.

▷ Then, this model must necessarily be non-stationary with $D = 1$.

▷ Indeed, if $D = 0$ the series of each month will oscillate around a value $c_j \neq 0$ and for that model to be common $c_j = c$.

▷ However, in a stationary model the constant is proportional to the mean of the process and as the means of the months are different by hypothesis the constants c_j must be different. Therefore, we cannot have $D = 0$ and the same model.

The ARIMA seasonal model

▷ Nevertheless, if the common model has $D = 1$, by taking this difference in each series the means of the months disappear, and the differences between the same month in one year and the previous year follow a stationary process of zero mean.

▷ The common model of the yearly series can be written in terms of the original data as:

$$\left(1 - \Phi_1 B^{12} - ... - \Phi_P B^{12P}\right)\left(1 - B^{12}\right) z_t =$$
$$= \left(1 - \Theta_1 B^{12} - ... - \Theta_Q B^{12Q}\right) \alpha_t \qquad (124)$$

where $t = 1, ..., T$ and now the ARIMA model is formulated in B^{12} since we are relating months from different years.

The ARIMA seasonal model

▷ As the model is the same for all the months we can apply this model to the original series z_t and obtain the series of residuals α_t. This residuals series will not generally be a white noise process since we have not taken into account the dependence between one data point and those immediately prior.

▷ Assuming that α_t follows a regular ARIMA process:

$$\phi_p(B)\nabla^d \alpha_t = \theta_q(B) a_t, \tag{125}$$

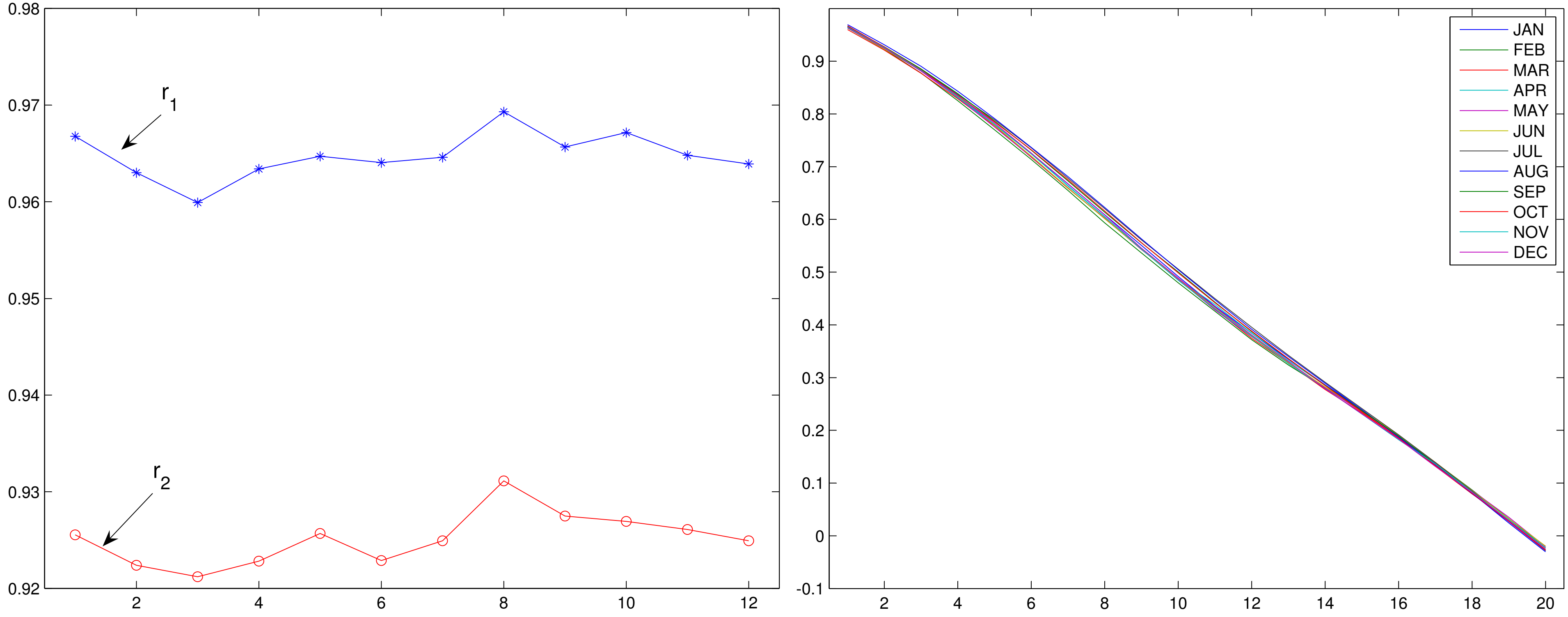
substituting this regular model (125) in the seasonal model (124) we obtain the complete model for the observed process, which is the one given in (121).

▷ To summarize, the multiplicative seasonal ARIMA model is based on the central hypothesis that *the relationship of seasonal dependence* (124) *is the same for all periods.*

▷ Experience indicates that this situation, although frequent, is not always true so it is advisable, whenever sufficient data are available, to test it by constructing the models (124) and checking to see if all of them are equal.

The ARIMA seasonal model - Examples

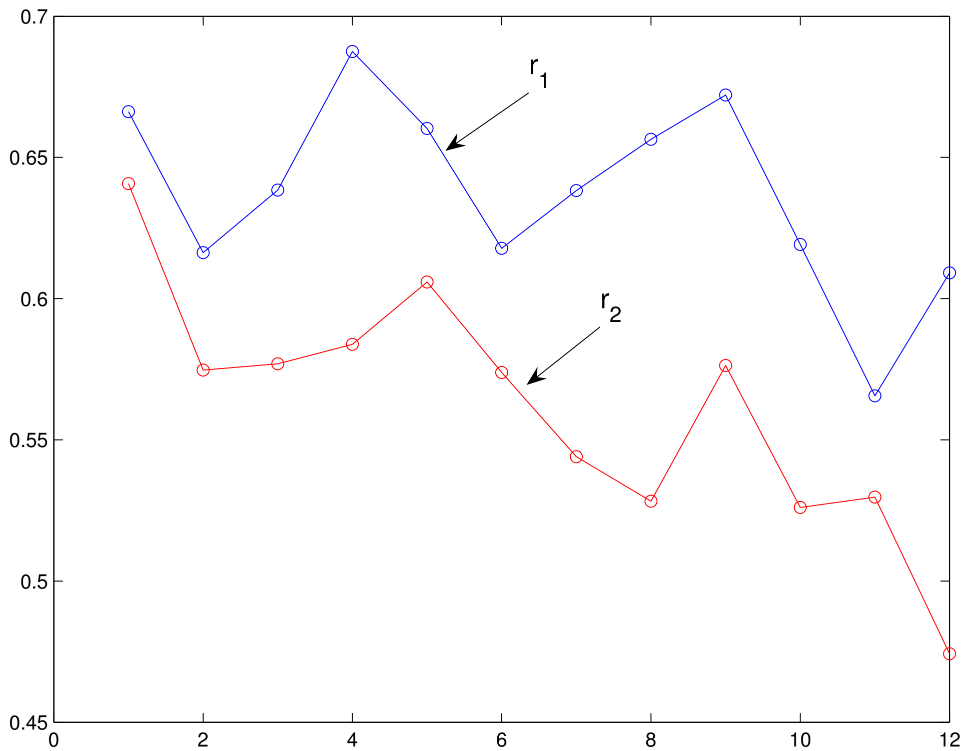
Example 60. *The figures give the first autocorrelation coefficients calculated using the 12 yearly series of each month of the year for gasoline consumption in Spain.*

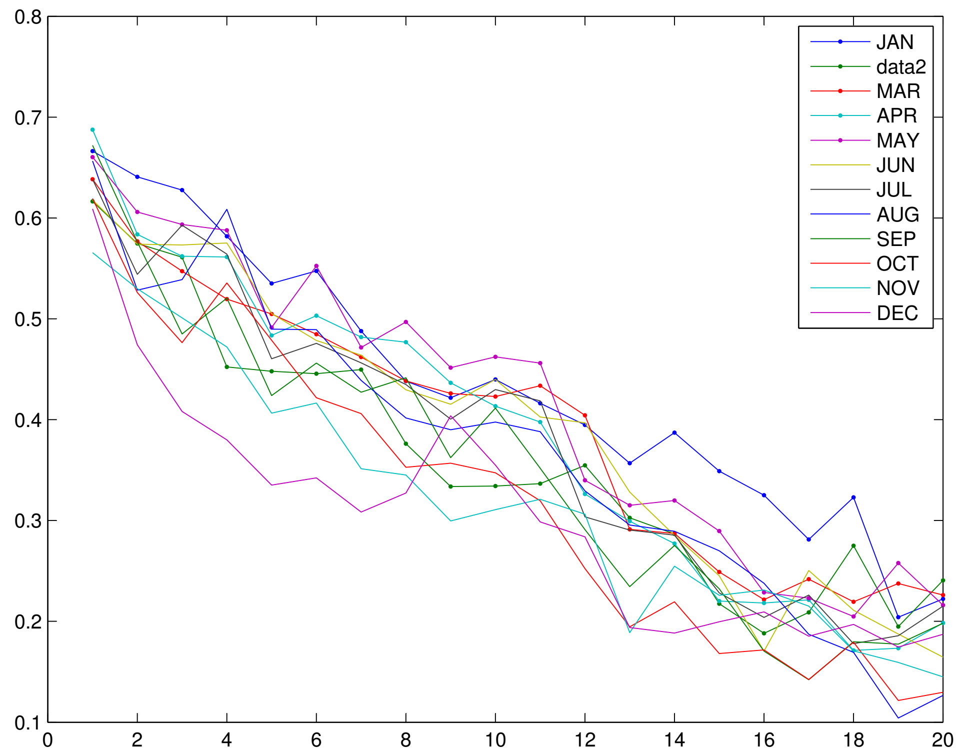


▷ We can conclude that the autocorrelations are similar.

The ARIMA seasonal model - Examples

Example 61. *The figures give the first autocorrelation coefficients calculated using the 12 yearly series of each month of the year for World temperature.*





▷ The approximated standard deviation from estimating each coefficient is $1\sqrt{T}$ and in this case $1/\sqrt{122} = .09$, thus we can conclude that the autocorrelations are different.

Simple autocorrelation function

▷ Let $w_t = \nabla^d \nabla_s^D z_t$ denote the stationary process obtained by differentiating the series regularly d times and D times seasonally. Thus, ω_t follows a multiplicative seasonal ARMA process:

$$\Phi_P(B^s)\,\phi_p(B)\,\omega_t = \theta_q(B)\,\Theta_Q(B^s)\,a_t \tag{126}$$

▷ The autocorrelation function of this process is a mixture of the AC functions corresponding to the regular and seasonal parts. It can be proved that if we denote by r_i the AC coefficients of the "regular" ARMA(p,q) process:

$$\phi_p(B)\,x_t = \theta_q(B)\,u_t, \tag{127}$$

R_{si} the AC coefficients in the lags s, $2s$, $3s$, ... of the seasonal ARMA (P,Q) process:

$$\Phi_P(B^s)\,y_t = \Theta_Q(B^s)\,v_t, \tag{128}$$

and ρ_j the AC coefficients of the complete process (126), it is found that:

$$\rho_j = \frac{r_j + \sum_{i=1}^{\infty} R_{si}\left(r_{si+j} + r_{si-j}\right)}{1 + 2\sum_{i=1}^{\infty} r_{si} R_{si}}. \tag{129}$$

Simple autocorrelation function

▷ If we assume $s = 12$ and allow $r_j \simeq 0$ for large lags (for example, for $j \geq 8$), the denominator of (129) is the unit and the autocorrelation function is:

1. In small lags ($j = 1, ..., 6$) only the regular part is observed, that is:

$$\rho_j \simeq r_j \quad j = 1, ..., 6.$$

2. In seasonal lags basically the seasonal part is observed, that is:

$$\rho_{12i} \simeq R_{12i}\left(r_{24i} + r_0\right) + R_{24i}\left(r_{36i} + r_{12i}\right)$$

and assuming that $r_{12i} \simeq 0$ for $i \geq 1$, with $r_0 = 1$ this expression is reduced to:

$$\rho_{12i} \simeq R_{12i} \quad (i = 1, 2, ...)\,.$$

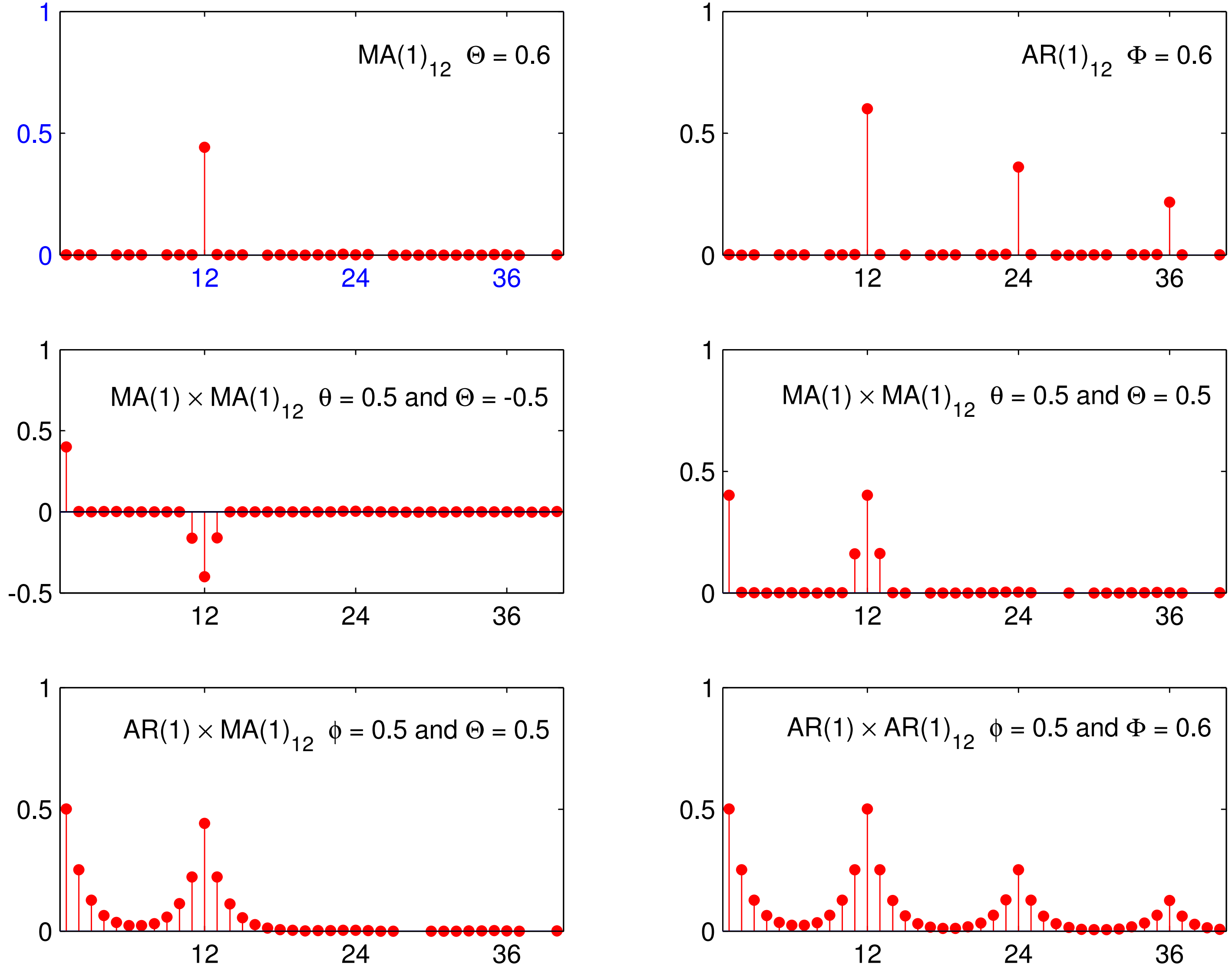
Simple autocorrelation function

3. Around the seasonal lags we observe the *interaction* between the regular and seasonal part, which shows up in the repetition of the regular part of the autocorrelation function on both sides of each seasonal lag.

   - Specifically, if the regular part is a moving average of order q, on both sides of each non-null seasonal lag there will be q coefficients different from zero.
   - If the regular part is autoregressive, we will observe the decrease imposed by the AR structure on both sides of the seasonal lags.

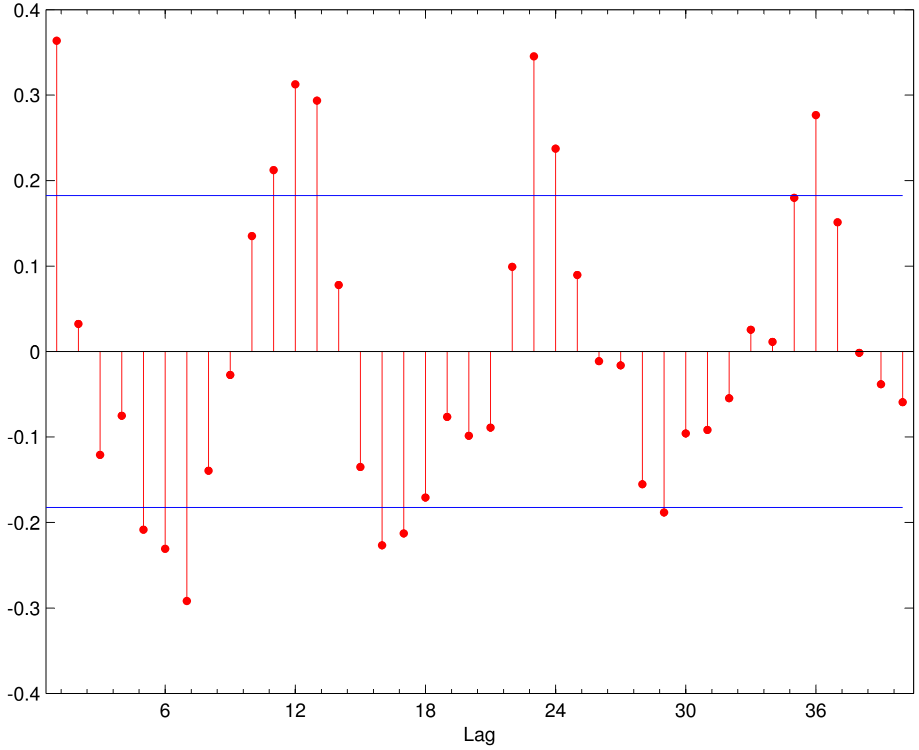
▷ The figure gives some examples of seasonal AR(1) and MA(1) models with different regular structures. If we combine these structures with a regular MA(1), we observe interaction only in the lags adjoining seasonal structures, whereas if we combine them with a regular AR, we have a long interaction structure where the regular structure is repeated on both sides of the seasonal lags.

Simple autocorrelation function - Examples



Simple autocorrelation function - Examples

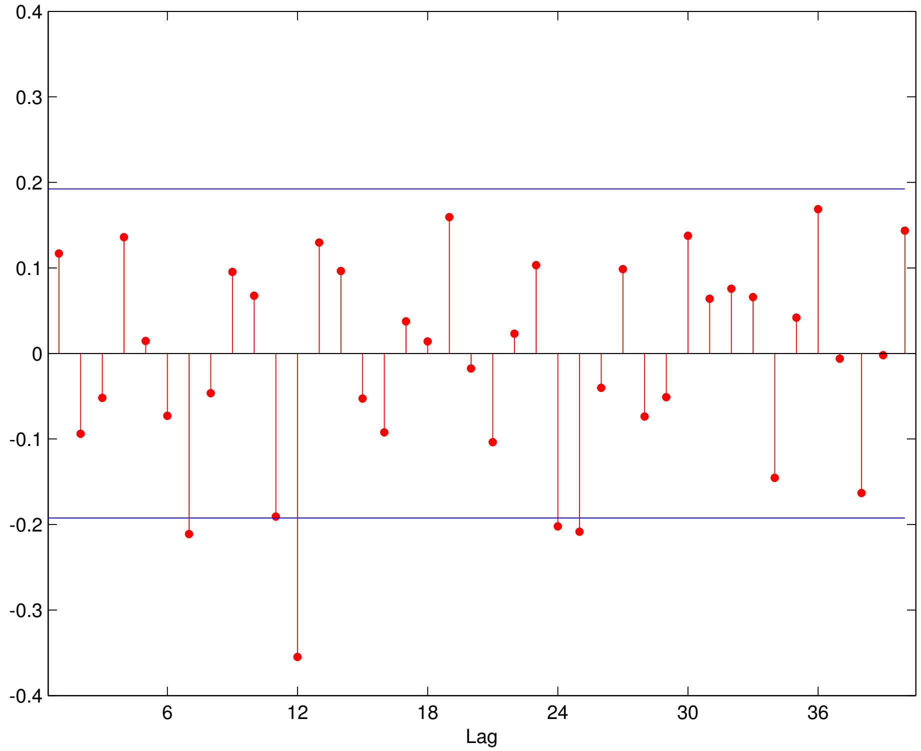
Example 62. *The figure gives the* ACF *for the Santiago rainfall series. Notice that in the regular part there is decay in the AR structure, whereas in lags* $12, 24, 36$*, a slow decay is observed in the coefficients, indicating the presence of a 12 period seasonal component.*



▷ The figure gives the *ACF* of the series with one seasonal difference.

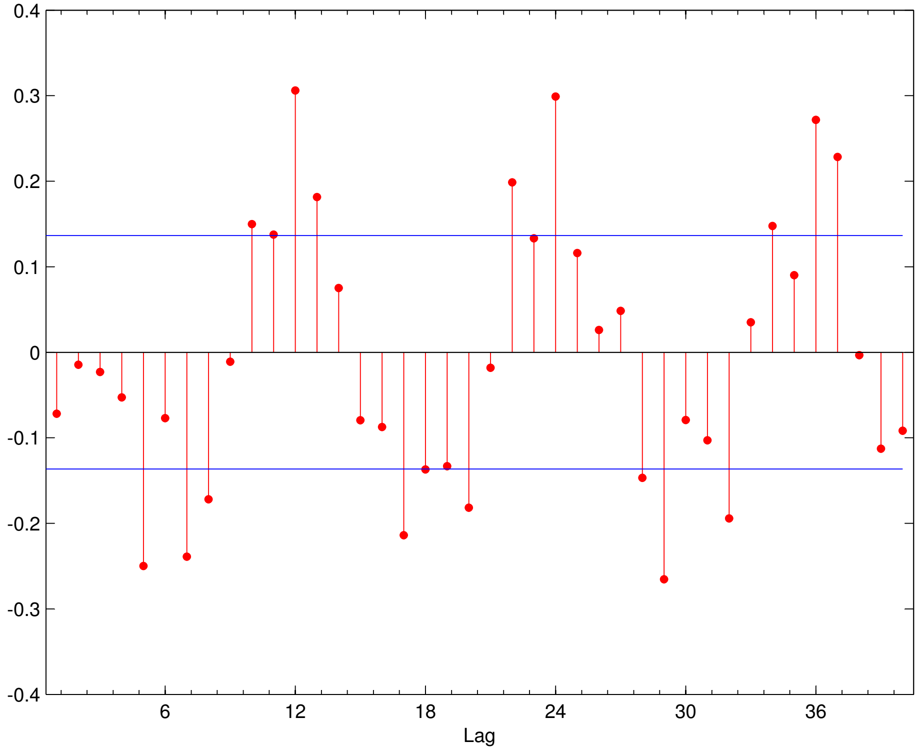
▷ In lags $1, 2, 3$, a weak structure is now seen that may be an AR(1) or MA(1) with a small coefficient and, in the seasonal part there are significant coefficients in $12, 24$ and 36, indicating an AR structure that must be at least of order 2, due to the fact that coefficients 12 and 24 are negative and 36 is positive.

▷ The high coefficients that appear around the seasonal lags confirm the existence of a regular structure, particularly the high coefficients in lags 11 and 23 are a clear sign of interaction between the regular and seasonal part.



Simple autocorrelation function - Examples

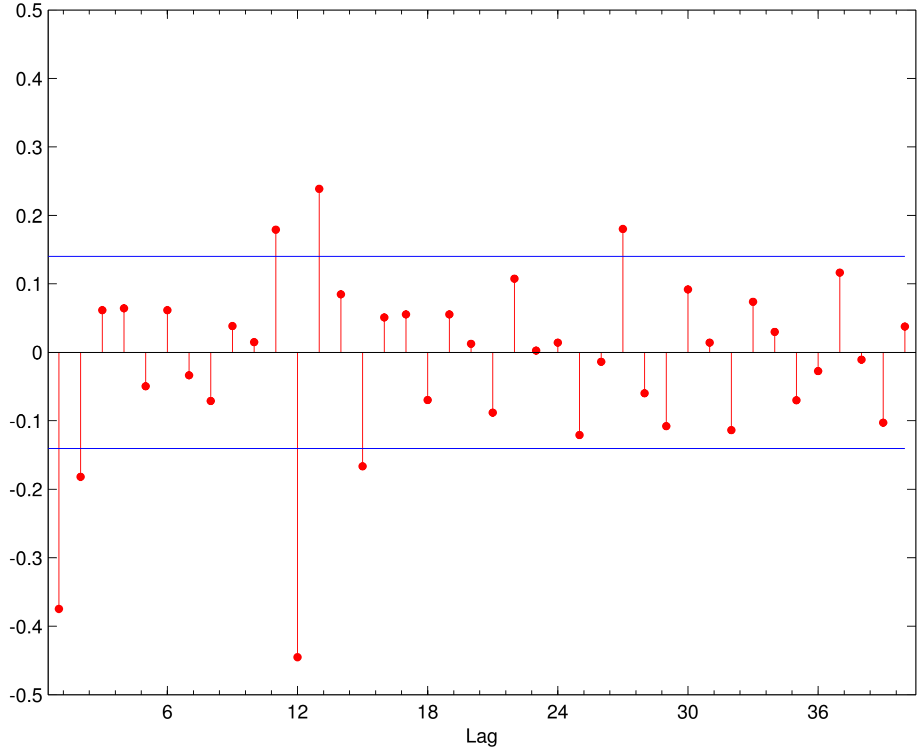
Example 63. *The figure gives the* ACF *of the first regular difference in the ozone series. This graph clearly shows the existence of strong seasonal dependency, with high coefficients in 12, 24, 36,. . . which fade slowly with the lag, suggesting the need for a seasonal difference.*



▷ The figure gives the *ACF* of the series with one regular difference and another seasonal.

▷ In the initial lags corresponding to the regular part we can observe either an MA(2) or an AR of at least order 2.

▷ In the seasonal lags the coefficient in 12 is significant and that of 24 is at the limit, suggesting a seasonal structure similar to a regular one.



Simple autocorrelation function - Examples

Example 64. *Write the simple theoretical autocorrelation function of the series ω_t obtained by applying a regular difference and a weekly difference (of period seven) to the daily series, z_t, with weekly seasonality that follows the model:*

$$\omega_t = \nabla\nabla_7 z_t = (1 - 0.5B)\left(1 - 0.8B^7\right) a_t.$$

▷ *The regular component is MA(1) and produces a single non-null autocorrelation coefficient. The coefficients generated by the regular part are:*

$$r_1 = \frac{-0.5}{1 + 0.5^2} = -0.4; \quad r_j = 0 \quad j \geq 2.$$

▷ *The seasonal component is also MA(1)$_7$ with generated coefficients:*

$$R_7 = \frac{-0.8}{1.64} = -0.49; \quad R_j = 0 \quad j \neq 7$$

▷ To obtain the coefficients of the process, which are the superimposition of both parts, using the equation (129) for $j = 1$:

$$\rho_1 = \frac{r_1 - R_7\left(r_8 + r_6\right) + R_{14}\left(r_{15} + r_{13}\right) + \ldots}{1 + 2\left(r_7 R_7 + r_{14} R_{14} + \ldots\right)}$$

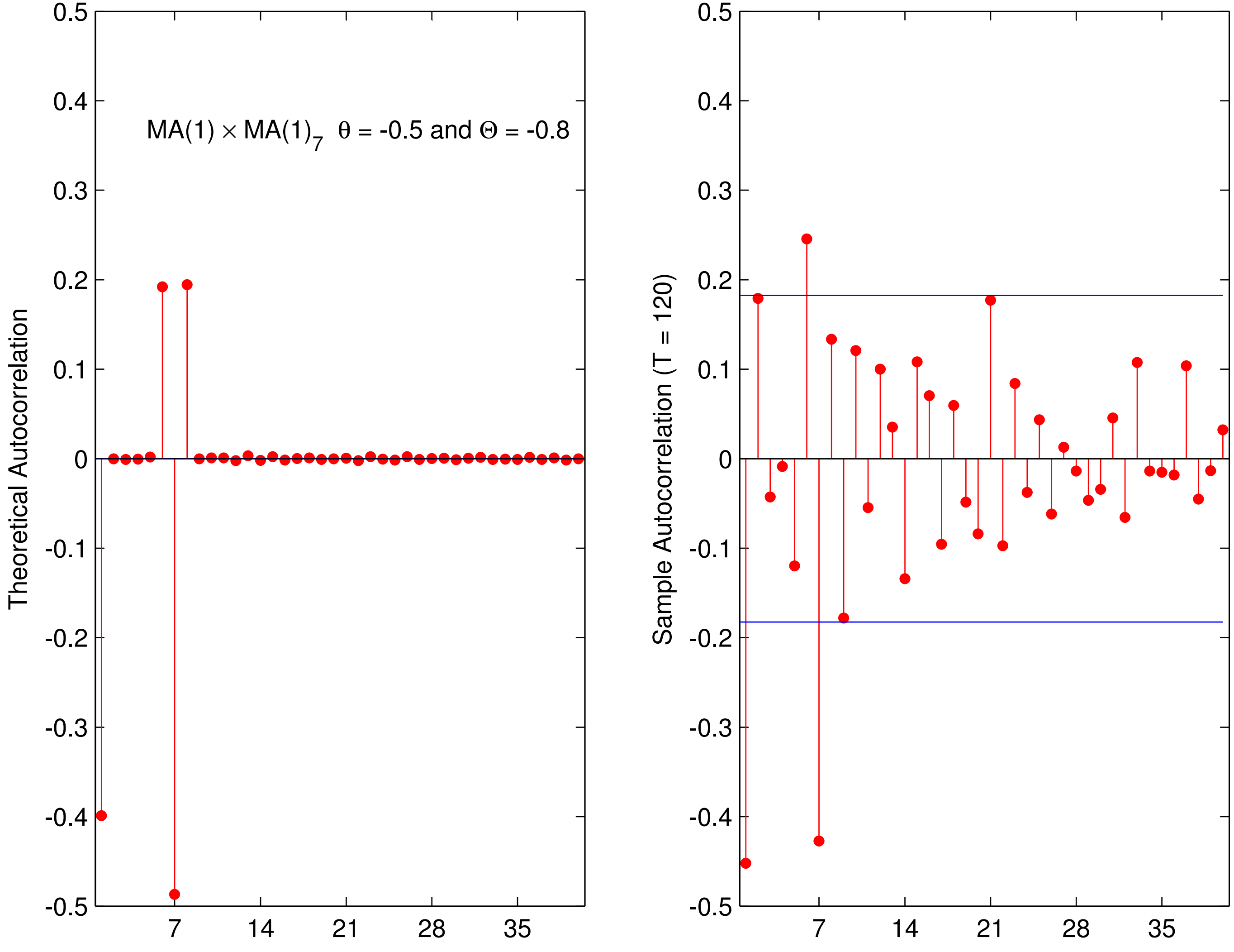
and since $r_j = 0$ for $j > 1$, we obtain:

$$\rho_1 = r_1 = -0.4.$$

▷ Applying the general formula it is analogously obtained that:

$$\begin{array}{l}
\rho_j = 0; \qquad j = 2, \ldots, 5 \\
\rho_6 = r_6 + R_7\left(r_{13} + r_1\right) + R_{14}(r_{20} + r_8) + \ldots = R_7 r_1 = 0.195 \\
\rho_7 = r_7 + R_7\left(r_{14} + r_0\right) + \ldots = R_7 = -0.49 \\
\rho_8 = r_8 + R_7\left(r_{15} + r_1\right) + \ldots = R_7 r_1 = 0.195 \\
\rho_j = 0; \qquad j > 8
\end{array} .$$

▷ The figure gives the *ACF* of the series ω_t:



Partial autocorrelation function

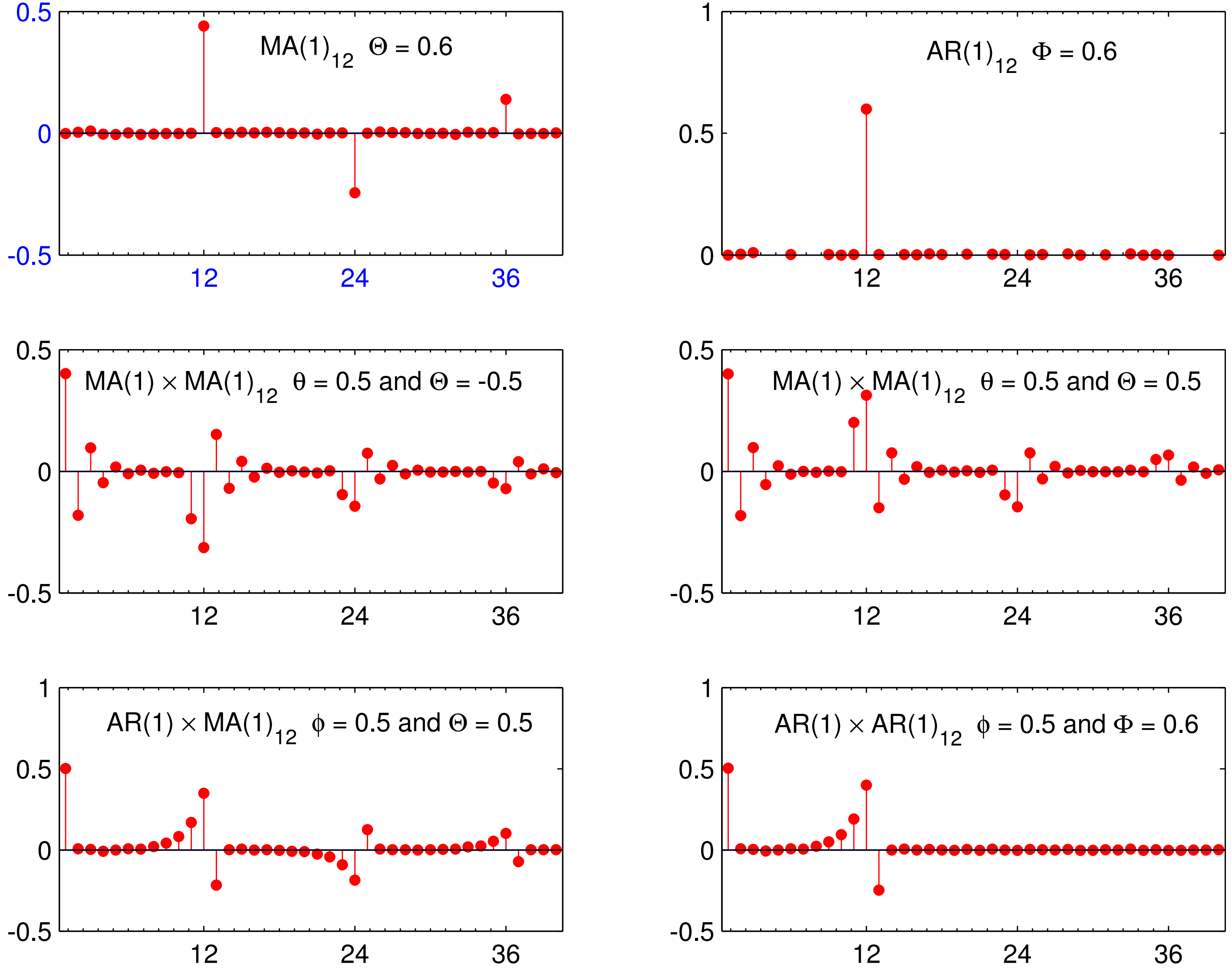
▷ The partial autocorrelation function of a multiplicative seasonal process is complex because it depends on the partial autocorrelation functions of the regular and seasonal parts (127) and (128) as well as on the simple autocorrelation of the regular part:

1. In the first lags the *PACF* of the regular structure appears and in the seasonal lags the *PACF* of the seasonal structure appears.

2. To the right of each seasonal coefficient (lags $js+1$, $js+2...$) the *PACF* of the regular part will appear. If the seasonal coefficient is positive the regular *PACF* appears inverted, whereas if it is negative the *PACF* appears with its sign.

3. To the left of the seasonal coefficients (lags $js-1$, $js-2$), we observe the autocorrelation function of the regular part.

Partial autocorrelation function - Examples



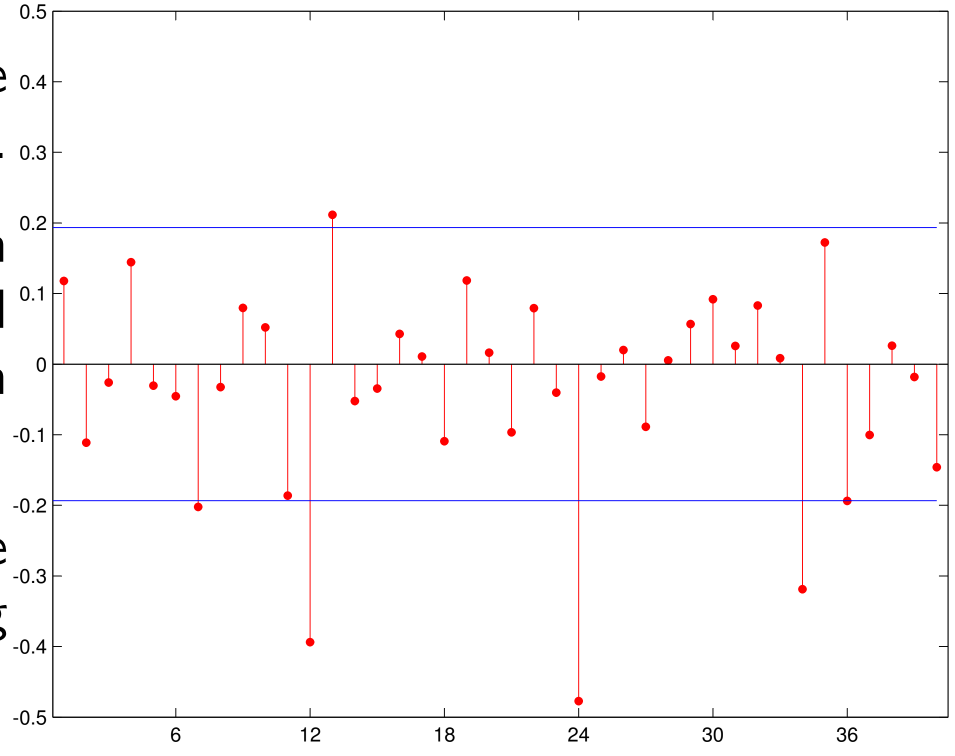
Partial autocorrelation function - Examples

Example 65. *The figure gives the PACF of the Santiago rainfall data.*

▷ In the regular part that the structure is AR or MA of a low order.

▷ In the seasonal lags two high and significant coefficients in 12 and 24 are observed which suggests an AR(2) for the seasonal part.

▷ The interactions explain the relatively high coefficients around lag 12.



Generalizations

▷ The multiplicative model can be easily generalized for more than one type of seasonality, but the estimation of these models requires special programs since the commercially available programs usually only assume one type of seasonality.

▷ The multiplicative models, while useful in practice are not suitable for all series and there are situations where the autocorrelation is different in the different points of time that make up a seasonal period.

- For example, with monthly data we may find that there is very little correlation between some months whereas there may be very high correlation with others.
- In these cases we have to model each month independently and later add the dependency structure month by month. These are called periodic models.

▷ One field not covered here in this book but which is important in economic applications is the seasonal adjustment of series.

Additional references

Periodic models:

- *Periodic Time Series Model*, by P.H. Franses and R. Paap, Oxford University Press, 2004.

- *Periodicity and Stochastic Trends in Economic Time Series* by P.H. Franses, Oxford University Press, 1996.

Seasonal adjustment:

- Chapter 8 in the book by Peña, Tiao and Tsay (2001).

- Chapter 2 in the book by Franses (1996).

- *Seasonal Adjustment with the X-11 Method*, by D. Ladiray and B. Quenneville, Lecture Notes in Statistics, Springer, 2001.