Chapter 7

ARIMA Models

A generalization of ARMA models which incorporates a wide class of nonstationary TS is obtained by introducing the differencing into the model. The simplest example of a nonstationary process which reduces to a stationary one after differencing is Random Walk. As we have seen in Section 4.5.2 Random Walk is a nonstationary AR(1) process with the value of the parameter ϕ equal to 1, that is the model is given by

$$X_t = X_{t-1} + Z_t, \quad \text{where } Z_t \sim WN(0, \sigma^2).$$

Its autocovariances depend on time as well as on lag. However, the first difference

$$\nabla X_t = X_t - X_{t-1}$$

is a stationary process, as it is just the White Noise Z_t. So, if we include WN in the ARMA class then ∇X_t is an ARMA(0,0) process, or in ARIMA notation it is ARIMA(0,1,0) process as it is obtained after first order differencing of X_t.

More generally, consider a TS model

$$X_t = m_t + Y_t,$$

where m_t is a polynomial of order k and Y_t is a stationary process. Then X_t is nonstationary (having a polynomial trend). However, we can detrend such a process by calculating the difference of order k. For a polynomial

$$m(t) = \beta_0 + \beta_1 t + \ldots + \beta_k t^k$$

we have, see Section 2.1.3,

$$\nabla^k X_t = k!\beta_k + \nabla^k Y_t,$$

where $\nabla^k Y_t$ is a linear combination of a stationary process, so it is stationary.

For example, let $m_t = \beta_0 + \beta_1 t$. Then differencing of order 1 will produce a stationary process

$$\nabla X_t = X_t - X_{t-1} = \beta_0 + \beta_1 t + Y_t - \beta_0 - \beta_1(t-1) - Y_{t-1} = \beta_1 + \nabla Y_t.$$

Hence we obtain a stationary process with mean $\mathrm{E}(\nabla X_t) = \beta_1$ which can be modelled as an ARMA time series.

This leads to a wider family of models which are ARMA models **after** differencing. Below is the formal definition.

Definition 7.1. *A process $\{X_t\}$ is said to follow an **Integrated ARMA model**, denoted by **ARIMA(p,d,q)**, if*

$$\nabla^d X_t = (1-B)^d X_t \tag{7.1}$$

is ARMA(p,q). We write the model as

$$\phi(B)(1-B)^d X_t = \theta(B) Z_t, \qquad Z_t \sim WN(0, \sigma^2). \tag{7.2}$$

The integration parameter d is a nonnegative integer.

Remark 7.1. When $d = 0$ we have the usual ARMA model, that is

$$ARIMA(p, 0, q) \equiv ARMA(p, q).$$

Remark 7.2. If $\mathrm{E}(\nabla^d X_t) = \mu \neq 0$, then we can write the model as

$$\phi(B)(1-B)^d X_t = \alpha + \theta(B) Z_t,$$

where $\alpha = \mu(1 - \phi_1 - \ldots - \phi_p)$.

Remark 7.3. The associated polynomial

$$\phi(z)(1-z)^d$$

has a unit root with multiplicity d, so even if all the roots of $\phi(z)$ are different than 1, the process X_t is nonstationary, however $\nabla^d X_t$ is stationary.

7.1 Building ARIMA Models

The basic steps in fitting ARIMA models to TS data are

Plotting the Data Plotting x_t versus t and inspecting the graph may reveal some unusual features, outstanding observations, may indicate if the series is stationary and/or seasonal and if the variance is stable.

Transforming the Data If it is necessary to transform the data, we may use Box-Cox power transformation

$$y_t = \begin{cases} \frac{x_t^\lambda - 1}{\lambda}, & \text{if } \lambda \neq 0, \\ \ln x_t, & \text{if } \lambda = 0. \end{cases}$$

MINITAB calculates optimum λ for a simpler version of the transformation, namely

$$y_t = \begin{cases} x_t^\lambda, & \text{if } \lambda \neq 0, \\ \ln x_t, & \text{if } \lambda = 0. \end{cases}$$

It finds lambda value which minimizes the standard deviation of a standardized transformed variable.

Identifying the Orders (p,d,q) of the Model Inspection of the TS plot may help to identify the differencing order d, while inspection of the ACF and PACF of the differenced data $\nabla^d x_t$ may help to identify the AR order p and MA order q. Differencing has to be done with care as it may unnecessarily introduce correlation; for example if Z_t is uncorrelated, $\nabla Z_t = Z_t - Z_{t-1}$ is correlated.

Estimation of the Model Parameters $\boldsymbol{\phi}$ and $\boldsymbol{\theta}$ This can be performed using Yule-Walker equations, the Maximum Likelihood method or the Least Squares method. The estimates are usually calculated numerically. Here we could also calculate confidence intervals for the parameters.

Residuals Diagnostics If the diagnostics, such as graphs of the residuals' ACF, PACF, Q-Q plot, histogram do not indicate a Gaussian White Noise, we should repeat the model estimation for another set of orders (p, d, q). In addition to the visual inspection of the graphs MINITAB offers the Ljung-Box-Pierce statistic

$$Q = n(n+2)\sum_{\tau=1}^{k} \frac{\widehat{\rho}_e^2(\tau)}{n-\tau} \sim \chi^2_{k-p-q}$$

for testing the hypothesis that the group of autocorrelations $\rho(1)$ to $\rho(k)$ is nonsignificant.

Having the model well fitting to the TS data we may perform prediction of future values.

7.2 Seasonal ARMA

Let us assume that there is seasonality in the data, but no trend. Then we could model the data as

$$X_t = s_t + Y_t, \tag{7.3}$$

where Y_t is a stationary process. The seasonality component is such that

$$s_t = s_{t-h},$$

where h denotes the length of the period and

$$\sum_{k=1}^{h} s_k = 0.$$

In Section 2.2.3 we have discussed removing the seasonal effect from the data by differencing at lag h. We have introduced the lag-h operator

$$\nabla_h X_t = X_t - X_{t-h} = X_t - B^h X_t = (1 - B^h) X_t,$$

which, for (7.3), gives

$$\nabla_h X_t = s_t + Y_t - s_{t-h} - Y_{t-h} = \nabla_h Y_t. \tag{7.4}$$

Hence, this operation removes the seasonality effect. This fact leads to introducing the **seasonal ARMA model**, denoted by $ARMA(P, Q)_h$, which is of the form

$$\Phi(B^h) X_t = \Theta(B^h) Z_t, \tag{7.5}$$

where

$$\Phi(B^h) = 1 - \Phi_1 B^h - \Phi_2 B^{2h} - \ldots - \Phi_P B^{Ph},$$

and

$$\Theta(B^h) = 1 + \Theta_1 B^h + \Theta_2 B^{2h} + \ldots + \Theta_Q B^{Qh}$$

are, respectively, the seasonal AR operator and the seasonal MA operator, with seasonal period of length h.

Remark 7.4. Analogously to $ARMA(p, q)$, the $ARMA(P, Q)_h$ model is causal only when the roots of $\Phi(z^h)$ lie outside the unit circle, and it is invertible only when the roots of $\Theta(z^h)$ lie outside the unit circle.

Example 7.1. Seasonal $ARMA(1,1)_{12}$.
Such a model can be written as

$$(1-\Phi B^{12})X_t = (1+\Theta B^{12})Z_t,$$

or

$$X_t - \Phi X_{t-12} = Z_t + \Theta Z_{t-12},$$

which is a generalization of (7.4).

When written as

$$X_t = \Phi X_{t-12} + Z_t + \Theta Z_{t-12},$$

and compared to $ARMA(1,1)$

$$X_t = \phi X_{t-1} + Z_t + \theta Z_{t-1}$$

we see that the seasonal ARMA presents the series in terms of its past values at lag equal to the length of the period (here h=12), while the non-seasonal ARMA does it in terms of its past values at lag 1. Seasonal ARMA incorporates the seasonality into the model.

Similarly as for the non-seasonal ARMA, here too, we require $|\Phi| < 1$ for the causality and $|\Theta| < 1$ for invertibility of the model.

Remark 7.5. Note that seasonal $ARMA(0,Q)_h$ is a seasonal $MA(Q)_h$, and seasonal $ARMA(P,0)_h$ is a seasonal $AR(P)_h$.

Example 7.2. ACF of $MA(1)_{12}$
A seasonal MA model with the period length $h=12$ can be written as

$$X_t = Z_t + \Theta Z_{t-12}.$$

It is a zero mean stationary model and it is easy to calculate its autocovariance, namely

$$\begin{aligned}
\gamma(\tau) &= \operatorname{cov}[Z_t + \Theta Z_{t-12}, Z_{t+\tau} + \Theta Z_{t+\tau-12}] \\
&= \operatorname{E}[(Z_t + \Theta Z_{t-12})(Z_{t+\tau} + \Theta Z_{t+\tau-12})] \\
&= \operatorname{E}(Z_t Z_{t+\tau}) + \Theta \operatorname{E}(Z_t Z_{t+\tau-12}) + \Theta \operatorname{E}(Z_{t-12} Z_{t+\tau}) + \Theta^2 \operatorname{E}(Z_{t-12} Z_{t+\tau-12}) \\
&= \begin{cases} (1+\Theta^2)\sigma^2 & \text{for } \tau = 0, \\ \Theta\sigma^2 & \text{for } \tau = \pm 12, \\ 0 & \text{otherwise} \end{cases}
\end{aligned}$$

Thus, the only non-zero correlations are $\rho(0) = 1$ and

$$\rho(\pm 12) = \frac{\Theta}{1+\Theta^2},$$

which is of the same form as $\rho(\pm 1)$ for a non-seasonal MA(1).

Example 7.3. ACF of $AR(1)_h$
Using the techniques for calculating ACVF and ACF of the non-seasonal AR(1) we obtain

$$\gamma(\tau) = \begin{cases} \frac{\sigma^2}{1-\Phi^2} & \text{for } \tau = 0, \\ \frac{\sigma^2\Phi^k}{1-\Phi^2} & \text{for } \tau = \pm hk,\ k = 1, 2, \dots, \\ 0 & \text{otherwise.} \end{cases}$$

This give the ACF similar to the ACF of a non-seasonal AR(1), namely

$$\rho(\tau) = \begin{cases} 1 & \text{for } \tau = 0, \\ \Phi^k & \text{for } \tau = \pm hk,\ k = 1, 2, \dots, \\ 0 & \text{otherwise.} \end{cases}$$

The following table summarizes the behaviour of the ACF and PACF of the causal and invertible seasonal ARMA models (see R.H.Shumway and Stoffer (2000)).

	$AR(P)_h$	$MA(Q)_h$	$ARMA(P,Q)_h$
ACF	Tails off at lags kh,	Cuts off after lag Qh	Tails off at lags kh
PACF	Cuts off after lag Ph	Tails off at lags kh	Tails off at lags kh

where h is the length of the seasonal period, $k = 1, 2, \dots$ and the values of ACF and PACF are zero at non-seasonal lags $\tau \neq kh$.

7.2.1 Mixed Seasonal ARMA

When we combine seasonal and non-seasonal operators we obtain a model

$$\Phi(B^h)\phi(B)X_t = \Theta(B^h)\theta(B)Z_t,$$

which is called **mixed seasonal ARMA** and it is denoted by

$$ARMA(p,q) \times (P,Q)_h.$$

The behavior of the ACF and PACF for such models is a combination of behavior of the seasonal and nonseasonal parts of the model.

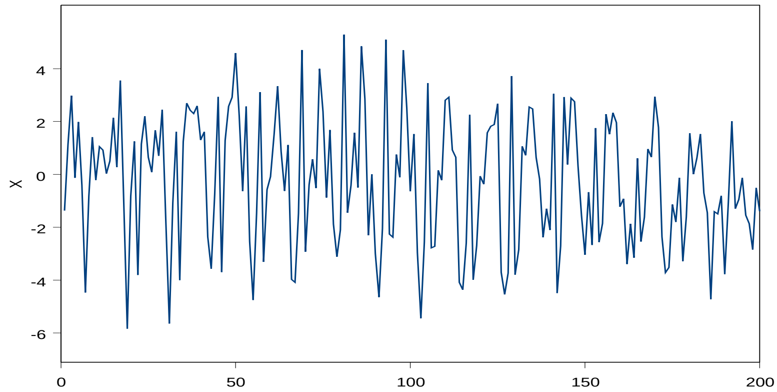

Figure 7.1: Simulated $ARMA(0,1)(1,0)_{12}$ process.

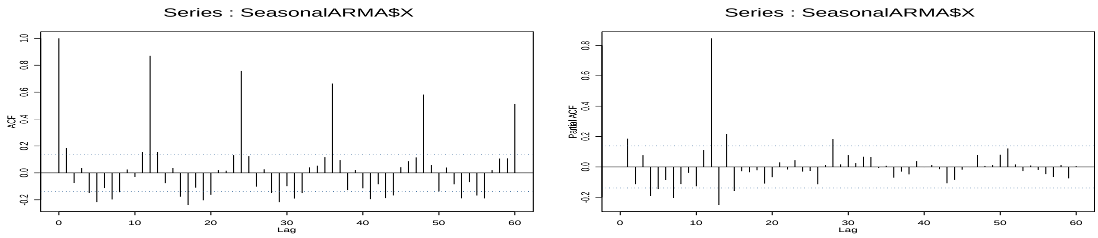

Figure 7.2: ACF and PACF of the above $ARMA(0,1)(1,0)_{12}$ process.

Example 7.4. $ARMA(0,1) \times (1,0)_{12}$

Such a model has the following form

$$X_t - \Phi X_{t-12} = Z_t + \theta Z_{t-1},$$

where $|\Phi| < 1$ and $|\theta| < 1$. Here we obtain

$$\rho(\tau) = \begin{cases} \Phi^k & \text{for } \tau = 12k, \; k = 1, 2, \ldots, \\ \frac{\theta}{1+\theta^2}\Phi^k & \text{for } \tau = 12k \pm 1, \; k = 1, 2, \ldots, \\ 0 & \text{otherwise.} \end{cases}$$

7.2.2 Seasonal ARIMA

Mixed seasonal ARMA is a stationary process. In practice however we often have nonstationary processes. Seasonal nonstationarity can occur when the process is nearly periodic in the season and the seasonal component varies slowly from period to period (say from year to year) according to a random walk, that is

$$s_t = s_{t-h} + V_t,$$

where V_t is a white noise. We can subtract the effect of the season (say month) using the backshift operator B^h to obtain seasonal stationarity

$$X_t - X_{t-h} = (1 - B^h)X_t.$$

This is a seasonal difference of order 1. In general we define a seasonal difference of order D as

$$\nabla_h^D X_t = (1 - B^h)^D X_t,$$

where $D = 1, 2, \ldots$. Usually $D = 1$ is sufficient to obtain seasonal stationarity. This leads to a very general **seasonal autoregressive integrated moving average (SARIMA)** model written as follows

$$\Phi(B^h)\phi(B)\nabla_H^D \nabla^d X_t = \alpha + \Theta(B^h)\theta(B)Z_t, \tag{7.6}$$

and denoted by $ARIMA(p, d, q) \times (P, D, Q)_h$.

Example 7.5. The model $ARIMA(0, 1, 1) \times (0, 1, 1)_{12}$ with $\alpha = 0$ is often applied for various economic data. Using formula (7.6) we obtain

$$(1 - B^{12})(1 - B)X_t = (1 + \Theta B^{12})(1 + \theta B)Z_t,$$

or, when expanded, we get the following form

$$(1 - B - B^{12} + B^{13})X_t = (1 + \theta B + \Theta B^{12} + \Theta\theta B^{13})Z_t,$$

or

$$X_t = X_{t-1} + X_{t-12} - X_{t-13} + Z_t + \theta Z_{t-1} + \Theta Z_{t-12} + \Theta\theta Z_{t-13}.$$

Bibliography

Box, G. and G. Jenkins (1976). *Time Series Analysis: Forecasting and Control. Revised Edition*. Holden Day.

Brockwell, P. and R. Davis (2002). *An Introduction to Time Series and Forecasting. Second Edition*. Springer-Verlag.

Castella, G. and R. Berger (1990). *Statistical Inference*. Duxbury Press.

Chatfield, C. (2004). *The Analysis of Time Series: An Introduction. Sixth Edition*. Chapman and Hall.

J.Hansen and S.Lebedeff (1987). Global trends of measured surface air temperature. *Journal of Geophysical Research 92*, 13.345–13.372.

R.H.Shumway and D. Stoffer (2000). *Time Series Analysis and Its Applications*. New York: Springer-Verlag.