# Chapter 3

# ARIMA Models

Classical regression is often insufficient for explaining all of the interesting dynamics of a time series. For example, the ACF of the residuals of the simple linear regression fit to the price of chicken data (see Example 2.4) reveals additional structure in the data that regression did not capture. Instead, the introduction of correlation that may be generated through lagged linear relations leads to proposing the *autoregressive (AR)* and *autoregressive moving average (ARMA)* models that were presented in Whittle [209]. Adding nonstationary models to the mix leads to the *autoregressive integrated moving average (ARIMA)* model popularized in the landmark work by Box and Jenkins [30]. The *Box–Jenkins method* for identifying ARIMA models is given in this chapter along with techniques for *parameter estimation* and *forecasting* for these models. A partial theoretical justification of the use of ARMA models is discussed in Sect. B.4.

## 3.1 Autoregressive Moving Average Models

The classical regression model of Chap. 2 was developed for the static case, namely, we only allow the dependent variable to be influenced by current values of the independent variables. In the time series case, it is desirable to allow the dependent variable to be influenced by the past values of the independent variables and possibly by its own past values. If the present can be plausibly modeled in terms of only the past values of the independent inputs, we have the enticing prospect that forecasting will be possible.

### Introduction to Autoregressive Models

Autoregressive models are based on the idea that the current value of the series, $x_t$, can be explained as a function of $p$ past values, $x_{t-1}, x_{t-2}, \dots, x_{t-p}$, where $p$ determines



R.H. Shumway, D.S. Stoffer, *Time Series Analysis and Its Applications*,
Springer Texts in Statistics, DOI 10.1007/978-3-319-52452-8_3

the number of steps into the past needed to forecast the current value. As a typical case, recall Example 1.10 in which data were generated using the model

$$x_t = x_{t-1} - .90x_{t-2} + w_t,$$

where $w_t$ is white Gaussian noise with $\sigma_w^2 = 1$. We have now assumed the current value is a particular *linear* function of past values. The regularity that persists in Fig. 1.9 gives an indication that forecasting for such a model might be a distinct possibility, say, through some version such as

$$x_{n+1}^n = x_n - .90x_{n-1},$$

where the quantity on the left-hand side denotes the forecast at the next period $n + 1$ based on the observed data, $x_1, x_2, \dots, x_n$. We will make this notion more precise in our discussion of forecasting (Sect. 3.4).

The extent to which it might be possible to forecast a real data series from its own past values can be assessed by looking at the autocorrelation function and the lagged scatterplot matrices discussed in Chap. 2. For example, the lagged scatterplot matrix for the Southern Oscillation Index (SOI), shown in Fig. 2.8, gives a distinct indication that lags 1 and 2, for example, are linearly associated with the current value. The ACF shown in Fig. 1.16 shows relatively large positive values at lags 1, 2, 12, 24, and 36 and large negative values at 18, 30, and 42. We note also the possible relation between the SOI and Recruitment series indicated in the scatterplot matrix shown in Fig. 2.9. We will indicate in later sections on transfer function and vector AR modeling how to handle the dependence on values taken by other series.

The preceding discussion motivates the following definition.

**Definition 3.1** *An* **autoregressive model** *of order $p$, abbreviated* **AR**($p$), *is of the form*

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \phi_p x_{t-p} + w_t, \tag{3.1}$$

*where $x_t$ is stationary, $w_t \sim wn(0, \sigma_w^2)$, and $\phi_1, \phi_2, \dots, \phi_p$ are constants ($\phi_p \neq 0$). The mean of $x_t$ in (3.1) is zero. If the mean, $\mu$, of $x_t$ is not zero, replace $x_t$ by $x_t - \mu$ in (3.1),*

$$x_t - \mu = \phi_1(x_{t-1} - \mu) + \phi_2(x_{t-2} - \mu) + \cdots + \phi_p(x_{t-p} - \mu) + w_t,$$

*or write*

$$x_t = \alpha + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \phi_p x_{t-p} + w_t, \tag{3.2}$$

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

We note that (3.2) is similar to the regression model of Sect. 2.1, and hence the term auto (or self) regression. Some technical difficulties, however, develop from applying that model because the regressors, $x_{t-1}, \dots, x_{t-p}$, are random components, whereas $z_t$ was assumed to be fixed. A useful form follows by using the backshift operator (2.29) to write the AR($p$) model, (3.1), as

$$(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p)x_t = w_t, \tag{3.3}$$

or even more concisely as

$$\phi(B)x_t = w_t. \tag{3.4}$$

The properties of $\phi(B)$ are important in solving (3.4) for $x_t$. This leads to the following definition.