2 Autoregressive Moving-Average Models

A basic idea in time series analysis is to construct more complex processes from simple ones. In the previous chapter we showed how the averaging of a white noise process leads to a process with first order autocorrelation. In this chapter we generalize this idea and consider processes which are solutions of linear stochastic difference equations. These so-called *ARMA processes* constitute the most widely used class of models for stationary processes.

Definition 2.1 (ARMA Models)**.** *A stochastic process* $\{X_t\}$ *with* $t \in \mathbb{Z}$ *is called an autoregressive moving-average process (ARMA process) of order* (p, q)*, denoted by ARMA*(p, q) *process, if the process is stationary and satisfies a linear stochastic difference equation of the form*

$$X_t - \phi_1 X_{t-1} - \ldots - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \ldots + \theta_q Z_{t-q} \tag{2.1}$$

with $Z_t \sim \mathrm{WN}(0, \sigma^2)$ *and* $\phi_p \theta_q \neq 0$*.* $\{X_t\}$ *is called an ARMA*(p, q) *process with mean* μ *if* $\{X_t - \mu\}$ *is an ARMA*(p, q) *process.*

The importance of ARMA processes is due to the fact that every stationary process can be approximated arbitrarily well by an ARMA process. In particular, it can be shown that for any given autocovariance function γ with the property $\lim_{h\to\infty} \gamma(h) = 0$ and any positive integer k there exists an autoregressive moving-average process (ARMA process) $\{X_t\}$ such that $\gamma_X(h) = \gamma(h)$, $h = 0, 1, \ldots, k$.

For an ARMA process with mean μ one often adds a constant c to the right hand side of the difference equation:

$$X_t - \phi_1 X_{t-1} - \ldots - \phi_p X_{t-p} = c + Z_t + \theta_1 Z_{t-1} + \ldots + \theta_q Z_{t-q}.$$

The mean of X_t is then: $\mu = \frac{c}{1-\phi_1-\ldots-\phi_p}$. The mean is therefore only well-defined if $\phi_1 + \ldots + \phi_p \neq 1$. The case $\phi_1 + \ldots + \phi_p = 1$ can, however, be excluded because there exists no stationary solution in this case (see Remark 2.2) and thus no ARMA process.

K. Neusser, *Time Series Econometrics*, Springer Texts in Business and Economics,
DOI 10.1007/978-3-319-32862-1_2

2.1 The Lag Operator

In times series analysis it is customary to rewrite the above difference equation more compactly in terms of the *lag operator* L. This is, however, not only a compact notation, but will open the way to analyze the inner structure of ARMA processes. The lag or back-shift operator L moves the time index one period back:

$$\mathrm{L}\{X_t\} = \{X_{t-1}\}.$$

For ease of notation we write: $\mathrm{L}X_t = X_{t-1}$. The lag operator is a linear operator with the following calculation rules:

(i) L applied to the process $\{X_t = c\}$ where c is an arbitrary constant gives:

$$\mathrm{L}c = c.$$

(ii) Applying L n times:

$$\underbrace{\mathrm{L} \ldots \mathrm{L}}_{n \text{ times}} X_t = \mathrm{L}^n X_t = X_{t-n}.$$

(iii) The inverse of the lag operator is the lead or forward operator. This operator shifts the time index one period into the future.[1] We can write L^{-1}:

$$\mathrm{L}^{-1} X_t = X_{t+1}.$$

(iv) For any integers m and n we have:

$$\mathrm{L}^m \mathrm{L}^n X_t = \mathrm{L}^{m+n} X_t = X_{t-m-n}.$$

(v) As $\mathrm{L}^{-1}\mathrm{L}X_t = X_t$ we have that

$$\mathrm{L}^0 = \mathbf{1}.$$

(vi) For any real numbers a and b, any integers m and n, and arbitrary stochastic processes $\{X_t\}$ and $\{Y_t\}$ we have:

$$(a\mathrm{L}^m + b\mathrm{L}^n)(X_t + Y_t) = aX_{t-m} + bX_{t-n} + aY_{t-m} + bY_{t-n}.$$

In this way it is possible to define *lag polynomials*: $A(\mathrm{L}) = a_0 + a_1\mathrm{L} + a_2\mathrm{L}^2 + \ldots + a_p\mathrm{L}^p$ where $a_0, a_1, \ldots, a_p$ are any real numbers. For these polynomials the usual

[1]One technical advantage of using the double-infinite index set $\mathbb{Z}$ is that the lag operators form a group.

calculation rules apply. Let, for example, $A(\mathrm{L}) = 1 - 0.5\mathrm{L}$ and $B(\mathrm{L}) = 1 + 4\mathrm{L}^2$ then $C(\mathrm{L}) = A(\mathrm{L})B(\mathrm{L}) = 1 - 0.5\mathrm{L} + 4\mathrm{L}^2 - 2\mathrm{L}^3$.

Applied to the stochastic difference equation, we define the autoregressive and the moving-average polynomial as follows:

$$\Phi(\mathrm{L}) = 1 - \phi_1\mathrm{L} - \ldots - \phi_p\mathrm{L}^p,$$
$$\Theta(\mathrm{L}) = 1 + \theta_1\mathrm{L} + \ldots + \theta_q\mathrm{L}^q.$$

The stochastic difference equation defining the ARMA process can then be written compactly as

$$\Phi(\mathrm{L})X_t = \Theta(\mathrm{L})Z_t.$$

Thus, the use of lag polynomials provides a compact notation for ARMA processes. Moreover and most importantly, $\Phi(z)$ and $\Theta(z)$, viewed as polynomials of the complex number z, also reveal much of their inherent structural properties as will become clear in Sect. 2.3.

2.2 Some Important Special Cases

Before we deal with the general theory of ARMA processes, we will analyze some important special cases first:

$q = 0$: autoregressive process of order p, AR(p) process
$p = 0$: moving-average process of order q, MA(q) process

2.2.1 The Moving-Average Process of Order q (MA(q) Process)

The MA(q) process is defined by the following stochastic difference equation:

$$X_t = \Theta(\mathrm{L})Z_t = \theta_0 Z_t + \theta_1 Z_{t-1} + \ldots + \theta_q Z_{t-q} \qquad \text{with} \quad \theta_0 = 1 \text{ and } \theta_q \neq 0$$

and $Z_t \sim \mathrm{WN}(0, \sigma^2)$. Obviously,

$$\mathbb{E}X_t = \mathbb{E}Z_t + \theta_1\mathbb{E}Z_{t-1} + \ldots + \theta_q\mathbb{E}Z_{t-q} = 0,$$

because $Z_t \sim \mathrm{WN}(0, \sigma^2)$. As can be easily verified using the properties of $\{Z_t\}$, the autocovariance function of the MA(q) processes are:

$$\begin{aligned}\gamma_X(h) &= \mathrm{cov}(X_{t+h}, X_t) = \mathbb{E}(X_{t+h}X_t)\\ &= \mathbb{E}(Z_{t+h} + \theta_1 Z_{t+h-1} + \ldots + \theta_q Z_{t+h-q})(Z_t + \theta_1 Z_{t-1} + \ldots + \theta_q Z_{t-q})\end{aligned}$$

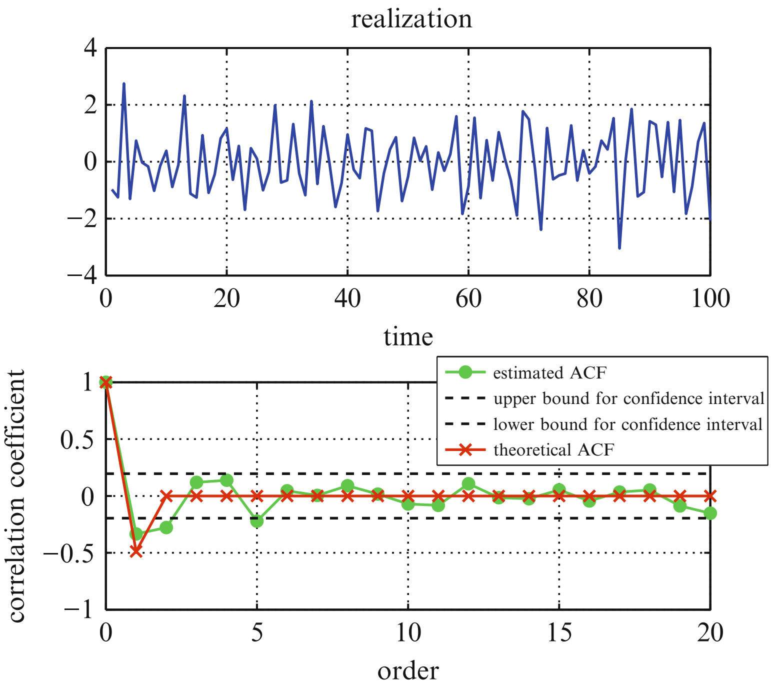

Fig. 2.1 Realization and estimated ACF of a MA(1) process: $X_t = Z_t - 0.8Z_{t-1}$ with $Z_t \sim$ IID N(0, 1)

$$= \begin{cases} \sigma^2 \sum_{i=0}^{q-|h|} \theta_i \theta_{i+|h|}, & |h| \le q; \\ 0, & |h| > q. \end{cases}$$

This implies the following autocorrelation function:

$$\rho_X(h) = \text{corr}(X_{t+h}, X_t) = \begin{cases} \frac{1}{\sum_{i=0}^{q} \theta_i^2} \sum_{i=0}^{q-|h|} \theta_i \theta_{i+|h|}, & |h| \le q; \\ 0, & |h| > q. \end{cases}$$

Every MA(q) process is therefore stationary irrespective of its parameters $\theta_0, \theta_1, \ldots, \theta_q$. Because the correlation between X_t and X_s is equal to zero if the two time points t and s are more than q periods apart, such processes are sometimes called processes with *short memory* or processes with *short range dependence*.

Figure 2.1 displays an MA(1) process and its autocorrelation function.

2.2.2 The First Order Autoregressive Process (AR(1) Process)

The AR(p) process requires a more thorough analysis as will already become clear from the AR(1) process. This process is defined by the following stochastic difference equation:

$$X_t = \phi X_{t-1} + Z_t, \qquad Z_t \sim \text{WN}(0, \sigma^2) \text{ and } \phi \neq 0. \tag{2.2}$$

The above stochastic difference equation has in general several solutions. Given a sequence $\{Z_t\}$ and an arbitrary distribution for X_0, it determines all random variables X_t, $t \in \mathbb{Z} \setminus \{0\}$, by applying the above recursion. The solutions are, however, not necessarily stationary. But, according to the Definition 2.1, only stationary processes qualify for ARMA processes. As we will demonstrate, depending on the value of ϕ, there may exist no or just one solution.

Consider first the case of $|\phi| < 1$. Inserting into the difference equation several times leads to:

$$\begin{aligned} X_t &= \phi X_{t-1} + Z_t = \phi^2 X_{t-2} + \phi Z_{t-1} + Z_t \\ &= \quad \ldots \\ &= Z_t + \phi Z_{t-1} + \phi^2 Z_{t-2} + \ldots + \phi^k Z_{t-k} + \phi^{k+1} X_{t-k-1}. \end{aligned}$$

If $\{X_t\}$ is a stationary solution, $\mathbb{V}X_{t-k-1}$ remains constant independently of k. Thus

$$\mathbb{V}\left(X_t - \sum_{j=0}^{k} \phi^j Z_{t-j}\right) = \phi^{2k+2}\mathbb{V}X_{t-k-1} \to 0 \text{ for } k \to \infty.$$

This shows that $\sum_{j=0}^{k} \phi^j Z_{t-j}$ converges in the mean square sense, and thus also in probability, to X_t for $k \to \infty$ (see Theorem C.8 in Appendix C). This suggests to take

$$X_t = Z_t + \phi Z_{t-1} + \phi^2 Z_{t-2} + \ldots = \sum_{j=0}^{\infty} \phi^j Z_{t-j} \tag{2.3}$$

as the solution to the stochastic difference equation. As $\sum_{j=0}^{\infty} |\phi^j| = \frac{1}{1-\phi} < \infty$ this solution is well-defined according to Theorem 6.4 and has the following properties:

$$\mathbb{E}X_t = \sum_{j=0}^{\infty} \phi^j \mathbb{E}Z_{t-j} = 0,$$

$$\gamma_X(h) = \operatorname{cov}(X_{t+h}, X_t) = \lim_{k\to\infty} \mathbb{E}\left(\sum_{j=0}^{k} \phi^j Z_{t+h-j}\right)\left(\sum_{j=0}^{k} \phi^j Z_{t-j}\right)$$

$$= \sigma^2 \phi^{|h|} \sum_{j=0}^{\infty} \phi^{2j} = \frac{\phi^{|h|}}{1-\phi^2}\,\sigma^2, \qquad h \in \mathbb{Z},$$

$$\rho_X(h) = \phi^{|h|}.$$

Thus the solution $X_t = \sum_{j=0}^{\infty} \phi^j Z_{t-j}$ is stationary and fulfills the difference equation as can be easily verified. It is also the only stationary solution which is compatible with the difference equation. Assume that there is second solution $\{\tilde{X}_t\}$ with these properties. Inserting into the difference equation yields again

$$\mathbb{V}\left(\tilde{X}_t - \sum_{j=0}^{k} \phi^j Z_{t-j}\right) = \phi^{2k+2} \mathbb{V} \tilde{X}_{t-k-1}.$$

This variance converges to zero for k going to infinity because $|\phi| < 1$ and because $\{\tilde{X}_t\}$ is stationary. The two processes $\{\tilde{X}_t\}$ and $\{X_t\}$ with $X_t = \sum_{j=0}^{\infty} \phi^j Z_{t-j}$ are therefore identical in the mean square sense and thus with probability one.

Finally, note that the recursion (2.2) will only generate a stationary process if it is initialized with X_0 having the stationary distribution, i.e. if $\mathbb{E}X_0 = 0$ and $\mathbb{V}X_0 = \sigma^2/(1-\phi^2)$. If the recursion is initiated with an arbitrary variance of X_0, $0 < \sigma_0^2 < \infty$, Eq. (2.2) implies the following difference equation for the variance of X_t, σ_t^2:

$$\sigma_t = \phi^2 \sigma_{t-1}^2 + \sigma^2.$$

The solution of this difference equation is

$$\sigma_t^2 - \sigma_*^2 = (\sigma_0^2 - \sigma_*^2)(\phi^2)^t$$

where $\sigma_*^2 = \sigma^2/(1-\phi^2)$ denotes the variance of the stationary distribution. If $\sigma_0^2 \neq \sigma_*^2$, σ_t^2 is not constant implying that the process $\{X_t\}$ is not stationary. However, as $|\phi| < 1$, the variance of X_t, σ_t^2, will converge to the variance of the stationary distribution.[2]

Figure 2.2 shows a realization of such a process and its estimated autocorrelation function.

In the case $|\phi| > 1$ the solution (2.3) does not converge. It is, however, possible to iterate the difference equation forward in time to obtain:

[2]Phillips and Sul (2007) provide an application and an in depth discussion of the hypothesis of economic growth convergence.

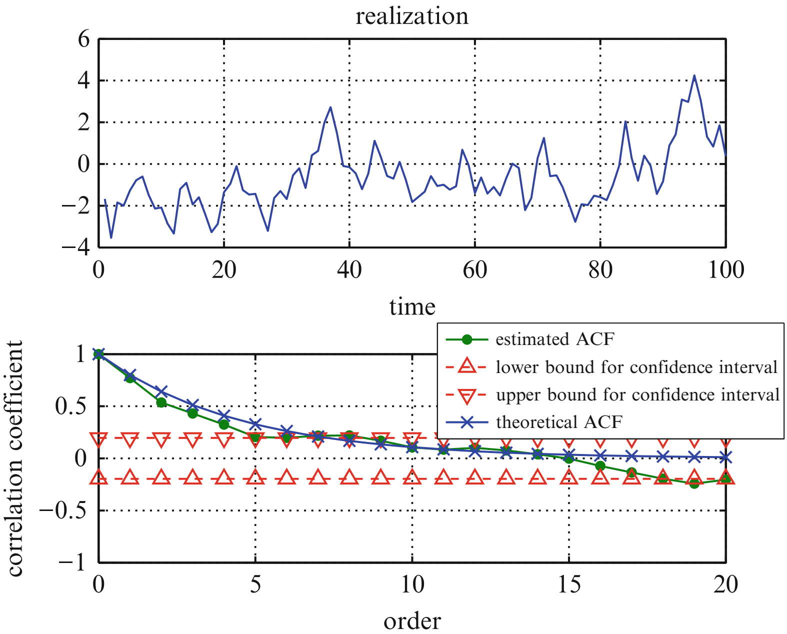

Fig. 2.2 Realization and estimated ACF of an AR(1) process: $X_t = 0.8X_{t-1} + Z_t$ with $Z_t \sim \text{IIN}(0, 1)$

$$X_t = \phi^{-1}X_{t+1} - \phi^{-1}Z_{t+1}$$
$$= \phi^{-k-1}X_{t+k+1} - \phi^{-1}Z_{t+1} - \phi^{-2}Z_{t+2} - \ldots - \phi^{-k-1}Z_{t+k+1}.$$

This suggests to take

$$X_t = -\sum_{j=1}^{\infty} \phi^{-j}Z_{t+j}$$

as the solution. Going through similar arguments as before it is possible to show that this is indeed the only stationary solution. This solution is, however, viewed to be inadequate because X_t depends on future shocks $Z_{t+j}, j = 1, 2, \ldots$ Note, however, that there exists an AR(1) process with $|\phi| < 1$ which is observationally equivalent, in the sense that it generates the same autocorrelation function, but with a new shock or forcing variable $\{\widetilde{Z}_t\}$ (see next section).

In the case $|\phi| = 1$ there exists no stationary solution (see Sect. 1.4.4) and therefore, according to our definition, no ARMA process. Processes with this property are called random walks, unit root processes or integrated processes. They play an important role in economics and are treated separately in Chap. 7.

2.3 Causality and Invertibility

If we interpret $\{X_t\}$ as the state variable and $\{Z_t\}$ as an impulse or shock, we can ask whether it is possible to represent today's state X_t as the outcome of current and past shocks $Z_t, Z_{t-1}, Z_{t-2}, \ldots$ In this case we can view X_t as being *caused* by past shocks and call this a *causal representation*. Thus, shocks to current Z_t will not only influence current X_t, but will propagate to affect also future X_t's. This notion of causality rest on the assumption that the past can cause the future but that the future cannot cause the past. See Sect. 15.1 for an elaboration of the concept of causality and its generalization to the multivariate context.

In the case that $\{X_t\}$ is a moving-average process of order q, X_t is given as a weighted sum of current and past shocks $Z_t, Z_{t-1}, \ldots, Z_{t-q}$. Thus, the moving-average representation is already the causal representation. In the case of an AR(1) process, we have seen that this is not always feasible. For $|\phi| < 1$, the solution (2.3) represents X_t as a weighted sum of current and past shocks and is thus the corresponding causal representation. For $|\phi| > 1$, no such representation is possible. The following Definition 2.2 makes the notion of a causal representation precise and Theorem 2.1 gives a general condition for its existence.

Definition 2.2 (Causality). *An ARMA(p, q) process $\{X_t\}$ with $\Phi(\mathrm{L})X_t = \Theta(\mathrm{L})Z_t$ is called* causal *with respect to $\{Z_t\}$ if there exists a sequence $\{\psi_j\}$ with the property $\sum_{j=0}^{\infty} |\psi_j| < \infty$ such that*

$$X_t = Z_t + \psi_1 Z_{t-1} + \psi_2 Z_{t-2} + \ldots = \sum_{j=0}^{\infty} \psi_j Z_{t-j} = \Psi(\mathrm{L})Z_t \quad \text{with } \psi_0 = 1.$$

where $\Psi(\mathrm{L}) = 1 + \psi_1 \mathrm{L} + \psi_2 \mathrm{L}^2 + \ldots = \sum_{j=0}^{\infty} \psi_j \mathrm{L}^j$. The above equation is referred to as the causal representation *of $\{X_t\}$ with respect to $\{Z_t\}$.*

The coefficients $\{\psi_j\}$ are of great importance because they determine how an impulse or a shock in period t propagates to affect current and future X_{t+j}, $j = 0, 1, 2 \ldots$ In particular, consider an impulse e_{t_0} at time t_0, i.e. a time series which is equal to zero except for the time t_0 where it takes on the values e_{t_0}. Then, $\{\psi_{t-t_0} e_{t_0}\}$ traces out the time history of this impulse. For this reason, the coefficients ψ_j with $j = t - t_0$, $t = t_0, t_0 + 1, t_0 + 2, \ldots$, are called the *impulse response function*. If $e_{t_0} = 1$, it is called a unit impulse. Alternatively, e_{t_0} is sometimes taken to be equal to σ, the standard deviation of Z_t. It is customary to plot ψ_j as a function of j, $j = 0, 1, 2 \ldots$

Note that the notion of causality is not an attribute of $\{X_t\}$, but is defined relative to another process $\{Z_t\}$. It is therefore possible that a stationary process is causal with respect to one process, but not with respect to another process. In order to make this point more concrete, consider again the AR(1) process defined by the equation $X_t = \phi X_{t-1} + Z_t$ with $|\phi| > 1$. As we have seen, the only stationary solution is given by $X_t = -\sum_{j=1}^{\infty} \phi^{-j} Z_{t+j}$ which is clearly not causal with respect $\{Z_t\}$. Consider as an alternative the process

$$\widetilde{Z}_t = X_t - \frac{1}{\phi}X_{t-1} = \phi^{-2}Z_t + (\phi^{-2} - 1)\sum_{j=1}^{\infty}\phi^{-j}Z_{t+j}. \qquad (2.4)$$

This new process is white noise with variance $\tilde{\sigma}^2 = \phi^{-2}\sigma^2$.[3] Because $\{X_t\}$ fulfills the difference equation

$$X_t = \frac{1}{\phi}X_{t-1} + \tilde{Z}_t,$$

$\{X_t\}$ is causal with respect to $\{\tilde{Z}_t\}$. This remark shows that there is no loss of generality involved if we concentrate on causal ARMA processes.

Theorem 2.1. *Let $\{X_t\}$ be an ARMA(p,q) process with $\Phi(\mathrm{L})X_t = \Theta(\mathrm{L})Z_t$ and assume that the polynomials $\Phi(z)$ and $\Theta(z)$ have no common root. $\{X_t\}$ is causal with respect to $\{Z_t\}$ if and only if $\Phi(z) \neq 0$ for $|z| \leq 1$, i.e. all roots of the equation $\Phi(z) = 0$ are outside the unit circle. The coefficients $\{\psi_j\}$ are then uniquely defined by identity :*

$$\Psi(z) = \sum_{j=0}^{\infty}\psi_j z^j = \frac{\Theta(z)}{\Phi(z)}.$$

Proof. Given that $\Phi(z)$ is a finite order polynomial with $\Phi(z) \neq 0$ for $|z| \leq 1$, there exits $\epsilon > 0$ such that $\Phi(z) \neq 0$ for $|z| \leq 1 + \epsilon$. This implies that $1/\Phi(z)$ is an analytic function on the circle with radius $1 + \epsilon$ and therefore possesses a power series expansion:

$$\frac{1}{\Phi(z)} = \sum_{j=0}^{\infty}\xi_j z^j = \Xi(z), \qquad \text{for } |z| < 1 + \epsilon.$$

This implies that $\xi_j(1+\epsilon/2)^j$ goes to zero for j to infinity. Thus there exists a positive and finite constant C such that

$$|\xi_j| < C(1 + \epsilon/2)^{-j}, \qquad \text{for all } j = 0, 1, 2, \ldots$$

This in turn implies that $\sum_{j=0}^{\infty}|\xi_j| < \infty$ and that $\Xi(z)\Phi(z) = 1$ for $|z| \leq 1$. Applying $\Xi(\mathrm{L})$ on both sides of $\Phi(\mathrm{L})X_t = \Theta(\mathrm{L})Z_t$, gives:

$$X_t = \Xi(\mathrm{L})\Phi(\mathrm{L})X_t = \Xi(\mathrm{L})\Theta(\mathrm{L})Z_t.$$

[3] The reader is invited to verify this.

Theorem 6.4 implies that the right hand side is well-defined. Thus $\Psi(L) = \Xi(L)\Theta(L)$ is the sought polynomial. Its coefficients are determined by the relation $\Psi(z) = \Theta(z)/\Phi(z)$.

Assume now that there exists a causal representation $X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}$ with $\sum_{j=0}^{\infty} |\psi_j| < \infty$. Therefore

$$\Theta(L)Z_t = \Phi(L)X_t = \Phi(L)\Psi(L)Z_t.$$

Take $\eta(z) = \Phi(z)\Psi(z) = \sum_{j=0}^{\infty} \eta_j z^j$, $|z| \leq 1$. Multiplying the above equation by Z_{t-k} and taking expectations shows that $\eta_k = \theta_k$, $k = 0, 1, 2, \ldots, q$, and that $\eta_k = 0$ for $k > q$. Thus we get $\Theta(z) = \eta(z) = \Phi(z)\Psi(z)$ for $|z| \leq 1$. As $\Theta(z)$ and $\Phi(z)$ have no common roots and because $|\Psi(z)| < \infty$ for $|z| \leq 1$, $\Phi(z)$ cannot be equal to zero for $|z| \leq 1$. □

Remark 2.1. If the AR and the MA polynomial have common roots, there are two possibilities:

- No common roots lies on the unit circle. In this situation there exists a unique stationary solution which can be obtained by canceling the common factors of the polynomials.
- If at least one common root lies on the unit circle then more than one stationary solution may exist (see the last example below).

Some Examples

We concretize the above Theorem and Remark by investigating some examples starting from the ARMA model $\Phi(L)X_t = \Theta(L)Z_t$ with $Z_t \sim \text{WN}(0, \sigma^2)$.

$\Phi(L) = 1 - 0.05L - 0.6L^2$ and $\Theta(L) = 1$: The roots of the polynomial $\Phi(z)$ are $z_1 = -4/3$ and $z_2 = 5/4$. Because both roots are absolutely greater than one, there exists a causal representation with respect to $\{Z_t\}$.

$\Phi(L) = 1 + 2L + 5/4L^2$ and $\Theta(L) = 1$: In this case the roots are conjugate complex and equal to $z_1 = -4/5 + 2/5\imath$ and $z_2 = -4/5 - 2/5\imath$. The modulus or absolute value of z_1 and z_2 equals $|z_1| = |z_2| = \sqrt{20/25}$. This number is smaller than one. Therefore there exists a stationary solution, but this solution is not causal with respect to $\{Z_t\}$.

$\Phi(L) = 1 - 0.05L - 0.6L^2$ and $\Theta(L) = 1 + 0.75L$: $\Phi(z)$ and $\Theta(z)$ have the common root $z = -4/3 \neq 1$. Thus one can cancel both $\Phi(L)$ and $\Theta(L)$ by $1 + \frac{3}{4}L$ to obtain the polynomials $\tilde{\Phi}(L) = 1 - 0.8L$ and $\tilde{\Theta}(L) = 1$. Because the root of $\tilde{\Phi}(z)$ equals 5/4 which is greater than one, there exists a unique stationary and causal representation with respect to $\{Z_t\}$.

$\Phi(L) = 1 + 1.2L - 1.6L^2$ and $\Theta(L) = 1 + 2L$: The roots of $\Phi(z)$ are $z_1 = 5/4$ and $z_2 = -0.5$. Thus one root is outside the unit circle whereas one is inside. This would suggest that there is no causal solution. However, the root $-0.5 \neq 1$ is shared by $\Phi(z)$ and $\Theta(z)$ and can therefore be canceled to obtain $\tilde{\Phi}(L) = 1-0.8L$

and $\tilde{\Theta}(\mathrm{L}) = 1$. Because the root of $\tilde{\Phi}(z)$ equals $5/4 > 1$, there exists a unique stationary and causal solution with respect to $\{Z_t\}$.

$\Phi(\mathrm{L}) = 1 + \mathrm{L}$ and $\Theta(\mathrm{L}) = 1 + \mathrm{L}$: $\Phi(z)$ and $\Theta(z)$ have the common root -1 which lies on the unit circle. As before one might cancel both polynomials by $1 + \mathrm{L}$ to obtain the trivial stationary and causal solution $\{X_t\} = \{Z_t\}$. This is, however, not the only solution. Additional solutions are given by $\{Y_t\} = \{Z_t + A(-1)^t\}$ where A is an arbitrary random variable with mean zero and finite variance σ_A^2 which is independent from both $\{X_t\}$ and $\{Z_t\}$. The process $\{Y_t\}$ has a mean of zero and an autocovariance function $\gamma_Y(h)$ which is equal to

$$\gamma_Y(h) = \begin{cases} \sigma^2 + \sigma_A^2, & h = 0; \\ (-1)^h \sigma_A^2, & h = \pm 1, \pm 2, \ldots \end{cases}$$

Thus this new process is therefore stationary and fulfills the difference equation.

Remark 2.2. If the AR and the MA polynomial in the stochastic difference equation $\Phi(\mathrm{L})X_t = \Theta(\mathrm{L})Z_t$ have no common root, but $\Phi(z) = 0$ for some z on the unit circle, there exists no stationary solution. In this sense the stochastic difference equation does no longer define an ARMA model. Models with this property are said to have a unit root and are treated in Chap. 7. If $\Phi(z)$ has no root on the unit circle, there exists a unique stationary solution.

As explained in the previous Theorem, the coefficients $\{\psi_j\}$ of the causal representation are uniquely determined by the relation $\Psi(z)\Phi(z) = \Theta(z)$. If $\{X_t\}$ is a MA process, $\Phi(z) = 1$ and the coefficients $\{\psi_j\}$ just correspond to the coefficients of the MA polynomial, i.e. $\psi_j = \theta_j$ for $0 \leq j \leq q$ and $\psi_j = 0$ for $j > q$. Thus in this case no additional computations are necessary. In general this is not the case. In principle there are two ways to find the coefficients $\{\psi_j\}$. The first one uses polynomial division or partial fractions, the second one uses the method of undetermined coefficients. This book relies on the second method because it is more intuitive and presents some additional insides. For this purpose let us write out the defining relation $\Psi(z)\Phi(z) = \Theta(z)$:

$$\left(\psi_0 + \psi_1 z + \psi_2 z^2 + \ldots\right)\left(1 - \phi_1 z - \phi_2 z^2 - \ldots \phi_p z^p\right) \\ = 1 + \theta_1 z + \theta_2 z^2 + \ldots + \theta_q z^q$$

Multiplying out the left hand side one gets:

$$\begin{aligned} &\psi_0 - \psi_0\phi_1 z - \psi_0\phi_2 z^2 - \psi_0\phi_3 z^3 - \cdots - \psi_0\phi_p z^p \\ &\qquad \psi_1 z - \psi_1\phi_1 z^2 - \psi_1\phi_2 z^3 - \cdots - \psi_1\phi_p z^{p+1} \\ &\qquad\quad + \psi_2 \quad z^2 - \psi_2\phi_1 z^3 - \cdots - \psi_2\phi_p z^{p+2} \end{aligned}$$

$$\cdots$$
$$= 1 + \theta_1 z + \quad \theta_2 z^2 + \quad \theta_3 z^3 + \cdots + \quad \theta_q z^q$$

Equating the coefficients of the powers of z, z^j, $j = 0, 1, 2, \ldots$, one obtains the following equations:

$$\begin{aligned}
z^0 : &\qquad \psi_0 = 1, \\
z^1 : &\qquad \psi_1 = \theta_1 + \phi_1\psi_0 = \theta_1 + \phi_1, \\
z^2 : &\qquad \psi_2 = \theta_2 + \phi_2\psi_0 + \phi_1\psi_1 = \theta_2 + \phi_2 + \phi_1\theta_1 + \phi_1^2, \\
&\qquad \ldots
\end{aligned}$$

As can be seen, it is possible to solve recursively for the unknown coefficients $\{\psi_j\}$. This is convenient when it comes to numerical computations, but in some cases one wants an analytical solution. Such a solution can be obtained by observing that, for $j \geq \max\{p, q+1\}$, the recursion leads to the following difference equation of order p:

$$\psi_j = \sum_{k=1}^{p} \phi_k \psi_{j-k} = \phi_1\psi_{j-1} + \phi_2\psi_{j-2} + \ldots + \phi_p\psi_{j-p}, \quad j \geq \max\{p, q+1\}.$$

This is a linear homogeneous difference equation with constant coefficients. The solution of such an equation is of the form (see Eq. (B.1) in Appendix B):

$$\psi_j = c_1 z_1^{-j} + \ldots + c_p z_p^{-j}, \qquad j \geq \max\{p, q+1\} - p, \tag{2.5}$$

where $z_1, \ldots, z_p$ denote the roots of $\Phi(z) = 1 - \phi_1 z - \ldots - \phi_p z^p = 0$.[4] Note that the roots are exactly those which have been computed to assess the existence of a causal representation. The coefficients $c_1, \ldots, c_p$ can be obtained using the p boundary conditions obtained from $\psi_j - \sum_{0<k\leq j} \phi_k \psi_{j-k} = \theta_j$, $\max\{p, q+1\} - p \leq j < \max\{p, q+1\}$. Finally, the values for ψ_j, $0 \leq j < \max\{p, q+1\} - p$, must be computed from the first $\max\{p, q+1\} - p$ iterations (see the example in Sect. 2.4).

As mentioned previously, the coefficients $\{\psi_j\}$ are of great importance as they quantify the effect of a shock to Z_{t-j} on X_t, respectively of Z_t on X_{t+j}. In macroeconomics they are sometimes called *dynamic multipliers* of a transitory or temporary shock. Because the underlying ARMA process is stationary and causal, the infinite sum $\sum_{j=0}^{\infty} |\psi_j|$ converges. This implies that the effect ψ_j converges to

[4] In the case of multiple roots one has to modify the formula according to Eq. (B.2).

zero as $j \to \infty$. Thus the effect of a shock dies out eventually[5]:

$$\frac{\partial X_{t+j}}{\partial Z_t} = \psi_j \to 0 \text{ for } j \to \infty.$$

As can be seen from Eq. (2.5), the coefficients $\{\psi_j\}$ even converge to zero exponentially fast to zero because each term $c_i z_i^{-j}$, $i = 1, \ldots, p$, goes to zero exponentially fast as the roots z_i are greater than one in absolute value. Viewing $\{\psi_j\}$ as a function of j one gets the so-called *impulse response function* which is usually displayed graphically.

The effect of a *permanent shock* in period t on X_{t+j} is defined as the cumulative effect of a transitory shock. Thus, the effect of a permanent shock to X_{t+j} is given by $\sum_{i=0}^{j} \psi_i$. Because $\sum_{i=0}^{j} \psi_i \le \sum_{i=0}^{j} |\psi_i| \le \sum_{i=0}^{\infty} |\psi_i| < \infty$, the cumulative effect remains finite.

In time series analysis we view the observations as realizations of $\{X_t\}$ and treat the realizations of $\{Z_t\}$ as unobserved. It is therefore of interest to know whether it is possible to recover the unobserved shocks from the observations on $\{X_t\}$. This idea leads to the concept of invertibility.

Definition 2.3 (Invertibility). *An ARMA(p,q) process for $\{X_t\}$ satisfying $\Phi(\mathrm{L})X_t = \Theta(\mathrm{L})Z_t$ is called* invertible *with respect to $\{Z_t\}$ if and only if there exists a sequence $\{\pi_j\}$ with the property $\sum_{j=0}^{\infty} |\pi_j| < \infty$ such that*

$$Z_t = \sum_{j=0}^{\infty} \pi_j X_{t-j}.$$

Note that like causality, invertibility is not an attribute of $\{X_t\}$, but is defined only relative to another process $\{Z_t\}$. In the literature, one often refers to invertibility as the strict miniphase property.[6]

Theorem 2.2. *Let $\{X_t\}$ be an ARMA(p,q) process with $\Phi(\mathrm{L})X_t = \Theta(\mathrm{L})Z_t$ such that polynomials $\Phi(z)$ and $\Theta(z)$ have no common roots. Then $\{X_t\}$ is invertible with respect to $\{Z_t\}$ if and only if $\Theta(z) \neq 0$ for $|z| \le 1$. The coefficients $\{\pi_j\}$ are then uniquely determined through the relation:*

[5]The use of the partial derivative sign actually represents an abuse of notation. It is inspired by an alternative definition of the impulse responses: $\psi_j = \frac{\partial \widetilde{\mathbb{P}}_t X_{t+j}}{\partial x_t}$ where $\widetilde{\mathbb{P}}_t$ denotes the optimal (in the mean squared error sense) linear predictor of X_{t+j} given a realization back to infinite remote past $\{x_t, x_{t-1}, x_{t-2}, \ldots\}$ (see Sect. 3.1.3). Thus, ψ_j represents the sensitivity of the forecast of X_{t+j} with respect to the observation x_t. The equivalence of alternative definitions in the linear and especially nonlinear context is discussed in Potter (2000).

[6]Without the qualification strict, the miniphase property allows for roots of $\Theta(z)$ on the unit circle. The terminology is, however, not uniform in the literature.

$$\Pi(z) = \sum_{j=0}^{\infty} \pi_j z^j = \frac{\Phi(z)}{\Theta(z)}.$$

Proof. The proof follows from Theorem 2.1 with X_t and Z_t interchanged. □

The discussion in Sect. 1.3 showed that there are in general two MA(1) processes compatible with the same autocorrelation function $\rho(h)$ given by $\rho(0) = 1$, $\rho(1) = \rho$ with $|\rho| \leq \frac{1}{2}$, and $\rho(h) = 0$ for $h \geq 2$. However, only one of these solutions is invertible because the two solutions for θ are inverses of each other. As it is important to be able to recover Z_t from current and past X_t, one prefers the invertible solution. Section 3.2 further elucidates this issue.

Remark 2.3. If $\{X_t\}$ is a stationary solution to the stochastic difference equation $\Phi(\mathrm{L})X_t = \Theta(\mathrm{L})Z_t$ with $Z_t \sim \mathrm{WN}(0, \sigma^2)$ and if $\Phi(z)\Theta(z) \neq 0$ for $|z| \leq 1$ then

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j},$$

$$Z_t = \sum_{j=0}^{\infty} \pi_j X_{t-j},$$

where the coefficients $\{\psi_j\}$ and $\{\pi_j\}$ are determined for $|z| \leq 1$ by $\Psi(z) = \frac{\Theta(z)}{\Phi(z)}$ and $\Pi(z) = \frac{\Phi(z)}{\Theta(z)}$, respectively. In this case $\{X_t\}$ is causal and invertible with respect to $\{Z_t\}$.

Remark 2.4. If $\{X_t\}$ is an ARMA process with $\Phi(\mathrm{L})X_t = \Theta(\mathrm{L})Z_t$ such that $\Phi(z) \neq 0$ for $|z| = 1$ then there exists polynomials $\tilde{\Phi}(z)$ and $\tilde{\Theta}(z)$ and a white noise process $\{\tilde{Z}_t\}$ such that $\{X_t\}$ fulfills the stochastic difference equation $\tilde{\Phi}(\mathrm{L})X_t = \tilde{\Theta}(\mathrm{L})\tilde{Z}_t$ and is causal with respect to $\{\tilde{Z}_t\}$. If in addition $\Theta(z) \neq 0$ for $|z| = 1$ then $\tilde{\Theta}(\mathrm{L})$ can be chosen such that $\{X_t\}$ is also invertible with respect to $\{\tilde{Z}_t\}$ (see the discussion of the AR(1) process after the definition of causality and Brockwell and Davis (1991, p. 88)). Thus, without loss of generality, we can restrict the analysis to causal and invertible ARMA processes.

2.4 Computation of the Autocovariance Function of an ARMA Process

Whereas the autocovariance function summarizes the external and directly observable properties of a time series, the coefficients of the ARMA process give information of its internal structure. Although there exists for each ARMA model a corresponding autocovariance function, the converse is not true as we have seen in Sect. 1.3 where we showed that two MA(1) processes are compatible with the same autocovariance function. This brings up a fundamental identification problem. In order to shed some light on the relation between autocovariance function and ARMA models it is necessary to be able to compute the autocovariance function for a given ARMA model. In the following, we will discuss three such procedures. Each procedure relies on the assumption that the ARMA process $\Phi(\mathrm{L})X_t = \Theta(\mathrm{L})Z_t$ with $Z_t \sim \mathrm{WN}(0, \sigma^2)$ is causal with respect to $\{Z_t\}$. Thus there exists a representation of X_t as a weighted sum of current and past Z_t's: $X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}$ with $\sum_{j=0}^{\infty} |\psi_j| < \infty$.

2.4.1 First Procedure

Starting from the causal representation of $\{X_t\}$, it is easy to calculate its autocovariance function given that $\{Z_t\}$ is white noise. The exact formula is proved in Theorem (6.4).

$$\gamma(h) = \sigma^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+|h|},$$

where

$$\Psi(z) = \sum_{j=0}^{\infty} \psi_j z^j = \frac{\Theta(z)}{\Phi(z)} \qquad \text{for } |z| \leq 1.$$

The first step consists in determining the coefficients ψ_j by the method of undetermined coefficients. This leads to the following system of equations:

$$\psi_j - \sum_{0<k\leq j} \phi_k \psi_{j-k} = \theta_j, \qquad 0 \leq j < \max\{p, q+1\},$$

$$\psi_j - \sum_{0<k\leq p} \phi_k \psi_{j-k} = 0, \qquad j \geq \max\{p, q+1\}.$$

This equation system can be solved recursively (see Sect. 2.3):

$$\psi_0 = \theta_0 = 1,$$

$$\psi_1 = \theta_1 + \psi_0\phi_1 = \theta_1 + \phi_1,$$

$$\psi_2 = \theta_2 + \psi_0\phi_2 + \psi_1\phi_1 = \theta_2 + \phi_2 + \phi_1\theta_1 + \phi_1^2,$$

$$\ldots$$

Alternatively one may view the second part of the equation system as a linear homogeneous difference equation with constant coefficients (see Sect. 2.3). Its solution is given by Eq. (2.5). The first part of the equation system delivers the necessary initial conditions to determine the coefficients $c_1, c_2, \ldots, c_p$. Finally one can insert the ψ's in the above formula for the autocovariance function.

A Numerical Example

Consider the ARMA(2,1) process with $\Phi(\mathrm{L}) = 1 - 1.3\mathrm{L} + 0.4\mathrm{L}^2$ and $\Theta(\mathrm{L}) = 1 + 0.4\mathrm{L}$. Writing out the defining equation for $\Psi(z)$, $\Psi(z)\Phi(z) = \Theta(z)$, gives:

$$
\begin{aligned}
1 + \psi_1 z + \psi_2 z^2 + \psi_3 z^3 + \ldots \\
- 1.3z - 1.3\psi_1 z^2 - 1.3\psi_2 z^3 - \ldots \\
+ 0.4z^2 + 0.4\psi_1 z^3 + \ldots \\
\ldots \qquad = 1 + 0.4z.
\end{aligned}
$$

Equating the coefficients of the powers of z leads to the following equation system:

$$
\begin{aligned}
z^0 &: \quad \psi_0 = 1, \\
z &: \quad \psi_1 - 1.3 = 0.4, \\
z^2 &: \quad \psi_2 - 1.3\psi_1 + 0.4 = 0, \\
z^3 &: \quad \psi_3 - 1.3\psi_2 + 0.4\psi_1 = 0, \\
&\qquad \ldots \\
&\quad \psi_j - 1.3\psi_{j-1} + 0.4\psi_{j-2} = 0, \quad \text{for } j \geq 2.
\end{aligned}
$$

The last equation represents a linear difference equation of order two. Its solution is given by

$$
\psi_j = c_1 z_1^{-j} + c_2 z_2^{-j}, \qquad j \geq \max\{p, q+1\} - p = 0,
$$

whereby z_1 and z_2 are the two distinct roots of the characteristic polynomial $\Phi(z) = 1 - 1.3z + 0.4z^2 = 0$ (see Eq. (2.5)) and where the coefficients c_1 and c_2 are determined from the initial conditions. The two roots are $\frac{1.3 \pm \sqrt{1.69 - 4 \times 0.4}}{2 \times 0.4} = 5/4 = 1.25$ and 2. The general solution to the homogeneous equation therefore is $\psi_j = c_1 0.8^j + c_2 0.5^j$. The constants c_1 and c_2 are determined by the equations:

$$
\begin{aligned}
j = 0 : \quad & \psi_0 = 1 \quad = c_1 0.8^0 + c_2 0.5^0 = c_1 + c_2 \\
j = 1 : \quad & \psi_1 = 1.7 = c_1 0.8^1 + c_2 0.5^1 = 0.8c_1 + 0.5c_2.
\end{aligned}
$$

Solving this equation system in the two unknowns c_1 and c_2 gives: $c_1 = 4$ and $c_2 = -3$. Thus the solution to the difference equation is given by:

$$\psi_j = 4(0.8)^j - 3(0.5)^j.$$

Inserting this solution for ψ_j into the above formula for $\gamma(h)$ one obtains after using the formula for the geometric sum:

$$\begin{aligned}
\gamma(h) &= \sigma^2 \sum_{j=0}^{\infty} \left(4 \times 0.8^j - 3 \times 0.5^j\right) \left(4 \times 0.8^{j+h} - 3 \times 0.5^{j+h}\right) \\
&= \sigma^2 \sum_{j=0}^{\infty} (16 \times 0.8^{2j+h} - 12 \times 0.5^j \times 0.8^{j+h} \\
&\qquad\qquad - 12 \times 0.8^j \times 0.5^{j+h} + 9 \times 0.5^{2j+h}) \\
&= 16\sigma^2 \frac{0.8^h}{1-0.64} - 12\sigma^2 \frac{0.8^h}{1-0.4} - 12\sigma^2 \frac{0.5^h}{1-0.4} + 9\sigma^2 \frac{0.5^h}{1-0.25} \\
&= \frac{220}{9}\sigma^2 (0.8)^h - 8\sigma^2 (0.5)^h.
\end{aligned}$$

Dividing $\gamma(h)$ by $\gamma(0)$, one gets the autocorrelation function:

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \frac{55}{37} \times 0.8^j - \frac{18}{37} \times 0.5^j$$

which is represented in Fig. 2.3.

2.4.2 Second Procedure

Instead of determining the ψ_j coefficients first, it is possible to compute the autocovariance function directly from the ARMA model. To see this multiply the ARMA equation successively by $X_{t-h}, h = 0, 1, \ldots$ and apply the expectations operator:

$$\begin{aligned}
&\mathbb{E}X_t X_{t-h} - \phi_1 \mathbb{E}X_{t-1} X_{t-h} - \cdots - \phi_p \mathbb{E}X_{t-p} X_{t-h} \\
&\qquad = \mathbb{E}Z_t X_{t-h} + \theta_1 \mathbb{E}Z_{t-1} X_{t-h} + \cdots + \theta_q \mathbb{E}Z_{t-q} X_{t-h}.
\end{aligned}$$

This leads to an equation system for the autocovariances $\gamma(h), h = 0, 1, 2, \ldots$:

$$\begin{aligned}
&\gamma(h) - \phi_1 \gamma(h-1) - \ldots - \phi_p \gamma(h-p) = \sigma^2 \sum_{h \le j \le q} \theta_j \psi_{j-h}, && h < \max\{p, q+1\} \\
&\gamma(h) - \phi_1 \gamma(h-1) - \ldots - \phi_p \gamma(h-p) = 0, && h \ge \max\{p, q+1\}.
\end{aligned}$$

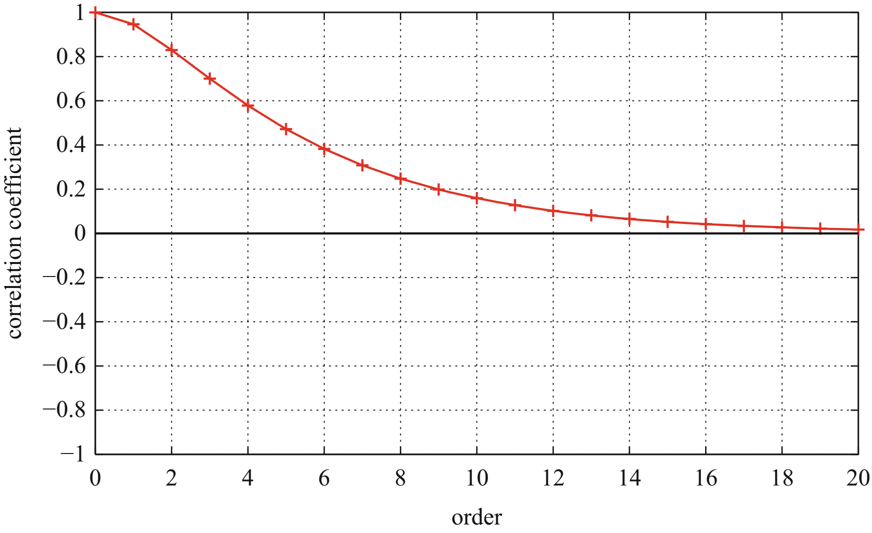

Fig. 2.3 Autocorrelation function of the ARMA(2,1) process: $(1 - 1.3\mathrm{L} + 0.4\mathrm{L}^2)X_t = (1 + 0.4\mathrm{L})Z_t$

The second part of the equation system consists again of a linear homogeneous difference equation in $\gamma(h)$ whereas the first part can be used to determine the initial conditions. Note that the initial conditions depend $\psi_1, \ldots, \psi_q$ which have to be determined before hand. The general solution of the difference equation is:

$$\gamma(h) = c_1 z_1^{-h} + \ldots + c_p z_p^{-h} \tag{2.6}$$

where $z_1, \ldots, z_p$ are the distinct roots of the polynomial $\Phi(z) = 1 - \phi_1 z - \ldots - \phi_p z^p = 0$.[7] The constants $c_1, \ldots, c_p$ can be computed from the first p initial conditions after the $\psi_1, \ldots \psi_q$ have been calculated like in the first procedure. The form of the solution shows that the autocovariance and hence the autocorrelation function converges to zero exponentially fast.

A Numerical Example

We consider the same example as before. The second part of the above equation system delivers a difference equation for $\gamma(h)$: $\gamma(h) = \phi_1 \gamma(h-1) + \phi_2 \gamma(h-2) = 1.3\gamma(h-1) - 0.4\gamma(h-2)$, $h \geq 2$. The general solution of this difference equation is (see Appendix B):

$$\gamma(h) = c_1 (0.8)^h + c_2 (0.5)^h, \qquad h \geq 2$$

[7] In case of multiple roots the formula has to be adapted accordingly. See Eq. (B.2) in the Appendix.

where 0.8 and 0.5 are the inverses of the roots computed from the same polynomial $\Phi(z) = 1 - 1.3z - 0.4z^2 = 0$.

The first part of the system delivers the initial conditions which determine the constants c_1 and c_2:

$$\gamma(0) - 1.3\gamma(-1) + 0.4\gamma(-2) = \sigma^2(1 + 0.4 \times 1.7)$$
$$\gamma(1) - 1.3\gamma(0) + 0.4\gamma(-1) = \sigma^2 0.4$$

where the numbers on the right hand side are taken from the first procedure. Inserting the general solution in this equation system and bearing in mind that $\gamma(h) = \gamma(-h)$ leads to:

$$0.216c_1 + 0.450c_2 = 1.68\sigma^2$$
$$-0.180c_1 - 0.600c_2 = 0.40\sigma^2$$

Solving this equation system in the unknowns c_1 and c_2 one gets finally gets: $c_1 = (220/9)\sigma^2$ and $c_2 = -8\sigma^2$.

2.4.3 Third Procedure

Whereas the first two procedures produce an analytical solution which relies on the solution of a linear difference equation, the third procedure is more suited for numerical computation using a computer. It rests on the same equation system as in the second procedure. The first step determines the values $\gamma(0), \gamma(1), \ldots, \gamma(p)$ from the first part of the equation system. The following $\gamma(h),\ h > p$ are then computed recursively using the second part of the equation system.

A Numerical Example

Using again the same example as before, the first of the equation delivers $\gamma(2), \gamma(1)$ and $\gamma(0)$ from the equation system:

$$\gamma(0) - 1.3\gamma(-1) + 0.4\gamma(-2) = \sigma^2(1 + 0.4 \times 1.7)$$
$$\gamma(1) - 1.3\gamma(0) + 0.4\gamma(-1) = \sigma^2 0.4$$
$$\gamma(2) - 1.3\gamma(1) + 0.4\gamma(0) = 0$$

Bearing in mind that $\gamma(h) = \gamma(-h)$, this system has three equations in three unknowns $\gamma(0), \gamma(1)$ and $\gamma(2)$. The solution is: $\gamma(0) = (148/9)\sigma^2$, $\gamma(1) = (140/9)\sigma^2$, $\gamma(2) = (614/45)\sigma^2$. This corresponds, of course, to the same numerical values as before. The subsequent values for $\gamma(h),\ h > 2$ are then determined recursively from the difference equation$\gamma(h) = 1.3\gamma(h-1) - 0.4\gamma(h-2)$.

2.5 Exercises

Exercise 2.5.1. *Consider the AR(1) process* $X_t = 0.8X_{t-1} + Z_t$ *with* $Z_t \sim \mathrm{WN}(0, \sigma^2)$. *Compute the variance of* $(X_1 + X_2 + X_3 + X_4)/4$.

Exercise 2.5.2. *Check whether the following stochastic difference equations possess a stationary solution. If yes, is the solution causal and/or invertible with respect to* $Z_t \sim \mathrm{WN}(0, \sigma^2)$*?*

(i) $X_t = Z_t + 2Z_{t-1}$
(ii) $X_t = 1.3X_{t-1} + Z_t$
(iii) $X_t = 1.3X_{t-1} - 0.4X_{t-2} + Z_t$
(iv) $X_t = 1.3X_{t-1} - 0.4X_{t-2} + Z_t - 0.3Z_{t-1}$
(v) $X_t = 0.2X_{t-1} + 0.8X_{t-2} + Z_t$
(vi) $X_t = 0.2X_{t-1} + 0.8X_{t-2} + Z_t - 1.5Z_{t-1} + 0.5Z_{t-2}$

Exercise 2.5.3. *Compute the causal representation with respect to* $Z_t \sim \mathrm{WN}(0, \sigma^2)$ *for the following ARMA processes:*

(i) $X_t = 1.3X_{t-1} - 0.4X_{t-2} + Z_t$
(ii) $X_t = 1.3X_{t-1} - 0.4X_{t-2} + Z_t - 0.2Z_{t-1}$
(iii) $X_t = \phi X_{t-1} + Z_t + \theta Z_{t-1}$ *with* $|\phi| < 1$

Exercise 2.5.4. *Compute the autocovariance function of the ARMA processes:*

(i) $X_t = 0.5X_{t-1} + 0.36X_{t-2} + Z_t$
(ii) $X_t = 0.5X_{t-1} + 0.36X_{t-2} + Z_t + 0.5Z_{t-1}$

Thereby $Z_t \sim \mathrm{WN}(0, \sigma^2)$.

Exercise 2.5.5. *Verify that the process* $\{\widetilde{Z}_t\}$ *defined in Eq.* (2.4) *is white noise with* $\widetilde{Z}_t \sim \mathrm{WN}(0, \theta^{-2}\sigma^2)$.