Chapter 2

Time series analysis and prediction

This chapter introduces the main concepts and tools for the analysis and prediction of time series, modeled as stationary stochastic processes.

2.1 Stochastic processes: definitions and properties

Definition 2.1. A *stochastic process* (hereafter, often abbreviated as s.p.) is a sequence of random variables $\boldsymbol{x}(t)$, with $t \in \mathcal{T}$, where $\mathcal{T}$ represents the time domain. If $\mathcal{T}$ is a countable set $\{t_1, t_2, \ldots, t_k, \ldots\}$, we refer to $\boldsymbol{x}(t)$ as a *discrete-time* stochastic process. Otherwise, if $\mathcal{T} = \mathbb{R}_+$, $\boldsymbol{x}(t)$ is a *continuous-time* stochastic process. In the following, we will mainly consider discrete-time stochastic processes.

By recalling the definition of random variable, a stochastic process is a function $\boldsymbol{x}(\omega, t)$ which associates to each pair $(\omega, t) \in \Omega \times \mathcal{T}$ a real number $x(t)$ [1]:

$$\boldsymbol{x} : \Omega \times \mathcal{T} \to \mathbb{R},$$

where $\mathcal{T}$ is the time domain and Ω is the space of events.

When the time instant $\bar{t} \in \mathcal{T}$ is fixed, $\boldsymbol{x}(\bar{t})$ boils down to a random variable. On the other hand, if the event ω is fixed (for instance, by carrying out a single *experiment* on the considered phenomenon), $\boldsymbol{x}(t)$ is just a deterministic function of t, known as a *realization* of the stochastic process, see Figure 2.1.

A stochastic process is completely characterized if the probability

$$\mathrm{P}\left(\boldsymbol{x}(t_1) \leq x_1, \boldsymbol{x}(t_2) \leq x_2, \ldots, \boldsymbol{x}(t_k) \leq x_k\right)$$

is known $\forall t_1, \ldots, t_k$, $\forall x_1, \ldots, x_k$ and $\forall k \in \mathbb{N}$. Similarly to what has been done for random variables, it is possible to define the cumulative distribution function

[1] For ease of notation, hereafter the dependence of $\boldsymbol{x}$ on the event ω will be omitted.

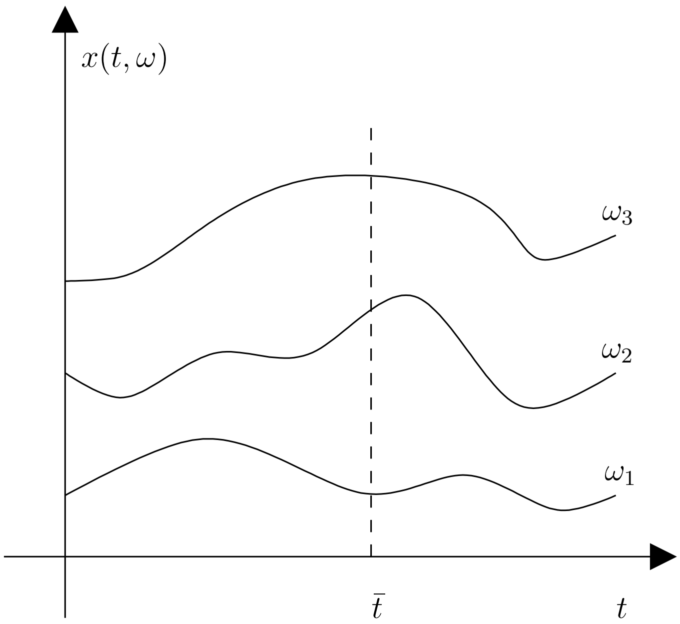


Figure 2.1: Different realizations of a stochastic process $\boldsymbol{x}(t)$, corresponding to different events ω_i.

$F_{\boldsymbol{x}}(\cdot) : \mathbb{R}^k \times \mathbb{Z}^k \to [0,1]$ and the probability density function $f_{\boldsymbol{x}}(\cdot) : \mathbb{R}^k \times \mathbb{Z}^k \to \mathbb{R}_+$ of a s.p. $\boldsymbol{x}(t)$, for $k = 1, 2, \ldots$ (also referred to as *k-th order statistics* of the s.p.):

$$F_{\boldsymbol{x}}(x_1, \ldots, x_k; t_1, \ldots, t_k) = \mathrm{P}\left(\boldsymbol{x}(t_1) \leq x_1, \boldsymbol{x}(t_2) \leq x_2, \ldots, \boldsymbol{x}(t_k) \leq x_k\right),$$
$$f_{\boldsymbol{x}}(x_1, \ldots, x_k; t_1, \ldots, t_k) = \frac{\partial^k}{\partial x_1, \ldots, x_k} F_{\boldsymbol{x}}(x_1, \ldots, x_k; t_1, \ldots, t_k).$$

In practice, such functions are rarely known for all $k \in \mathbb{N}$. Most often, only the first and second order statistics are considered. Let

$$\begin{cases} F_{\boldsymbol{x}}(x;t) = \mathrm{P}\left(\boldsymbol{x}(t) \leq x\right), \\ f_{\boldsymbol{x}}(x;t) = \dfrac{\partial}{\partial x} F_{\boldsymbol{x}}(x;t), \end{cases}$$

and

$$\begin{cases} F_{\boldsymbol{x}}(x_1, x_2; t_1, t_2) = \mathrm{P}\left(\boldsymbol{x}(t_1) \leq x_1, \boldsymbol{x}(t_2) \leq x_2\right), \\ f_{\boldsymbol{x}}(x_1, x_2; t_1, t_2) = \dfrac{\partial^2}{\partial x_1 \partial x_2} F_{\boldsymbol{x}}(x_1, x_2; t_1, t_2), \end{cases}$$

the 1st and 2nd order statistics of the s.p. $\boldsymbol{x}(t)$. Once t_1 and t_2 have been fixed, $\boldsymbol{x}(t_1)$ and $\boldsymbol{x}(t_2)$ are random variables and one can recover the first order statistics in the usual way, for example:

$$f_{\boldsymbol{x}}(x_1; t_1) = \int_{-\infty}^{+\infty} f_{\boldsymbol{x}}(x_1, x_2; t_1, t_2)\, \mathrm{d}x_2.$$

The mean and covariance function of a stochastic process are defined next.

Definition 2.2. The *mean* (or *expected value*) $m_{\boldsymbol{x}}(t)$ of the s.p. $\boldsymbol{x}(t)$ is defined as

$$m_{\boldsymbol{x}}(t) \triangleq \mathbf{E}\left[\boldsymbol{x}(t)\right] \triangleq \int_{-\infty}^{+\infty} x f_{\boldsymbol{x}}(x;t)\, \mathrm{d}x.$$

Definition 2.3. The *covariance function* (or *autocovariance*) $R_{\boldsymbol{x}}(t,s)$ of a s.p. $\boldsymbol{x}(t)$ is defined as

$$\begin{aligned} R_{\boldsymbol{x}}(t,s) &\triangleq \mathbf{E}\left[(\boldsymbol{x}(t)-m_{\boldsymbol{x}}(t))(\boldsymbol{x}(s)-m_{\boldsymbol{x}}(s))^T\right] \\ &\triangleq \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x_1-m_{\boldsymbol{x}}(t))(x_2-m_{\boldsymbol{x}}(s))^T f_{\boldsymbol{x}}\left(x_1,x_2;t,s\right)\mathrm{d}x_1\mathrm{d}x_2. \end{aligned}$$

In Definition 2.3, the notation $(\cdot)^T$ is used in the case that $\boldsymbol{x}(t)$ is a vector of stochastic processes. If $\boldsymbol{x}(t) \in \mathbb{R}^n$, then its mean is an n-dimensional vector itself, while the covariance function $R_{\boldsymbol{x}}(t,s)$ is a $n \times n$ matrix.

Definition 2.4. The *cross-covariance function* $R_{\boldsymbol{x},\boldsymbol{y}}(t,s)$ of two stochastic processes $\boldsymbol{x}(t)$, $\boldsymbol{y}(t)$ is defined as

$$\begin{aligned} R_{\boldsymbol{x},\boldsymbol{y}}(t,s) &\triangleq \mathbf{E}\left[(\boldsymbol{x}(t)-m_{\boldsymbol{x}}(t))(\boldsymbol{y}(s)-m_{\boldsymbol{y}}(s))^T\right] \\ &\triangleq \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x-m_{\boldsymbol{x}}(t))(y-m_{\boldsymbol{y}}(s))^T f_{\boldsymbol{x},\boldsymbol{y}}\left(x,y;t,s\right)\mathrm{d}x\,\mathrm{d}y. \end{aligned}$$

It is worth stressing that the mean and covariance function of a s.p. are, in general, functions of one (mean) or two (covariance) time indexes. For example, the mean $m_{\boldsymbol{x}}(t)$ of a s.p. $\boldsymbol{x}(t)$ may take on different values at different times t.

An important class of stochastic process is the one in which the statistics do not change in presence of a time shift.

Definition 2.5. A stochastic process $\boldsymbol{x}(t)$ is *strongly stationary* if all the statistics of $\boldsymbol{x}(t)$ and $\boldsymbol{x}(t+\delta)$ are identical for every δ, i.e., if

$$\mathrm{P}\left(\boldsymbol{x}(t_1+\delta) \le x_1, \ldots, \boldsymbol{x}(t_k+\delta) \le x_k\right)$$

does not depend on δ, $\forall t_1, \ldots, t_k$, $\forall x_1, \ldots, x_k$ and $\forall k$.

In other words, the cdf and pdf of a strongly stationary s.p. is invariant with respect to shifts in the time domain. Strong stationarity is often difficult to verify, as it involves all the statistics. Weaker notions are the one regarding only a subset of the statistics, like the following one.

Definition 2.6. A stochastic process $x(t)$ is *weakly stationary* if

$$\begin{aligned} m_{\boldsymbol{x}}(t) &= m_{\boldsymbol{x}}(t+\delta) \\ R_{\boldsymbol{x}}(t,s) &= R_{\boldsymbol{x}}(t+\delta, s+\delta) \end{aligned}$$

for all $\delta \in \mathbb{Z}$.

According to the above definition, the mean and the covariance function are invariant with respect to time shifts. Since Definition 2.6 holds for all δ, one has that $\boldsymbol{x}(t)$ is weakly stationary if and only if

$$\begin{aligned} m_{\boldsymbol{x}}(t) &= m_{\boldsymbol{x}} \\ R_{\boldsymbol{x}}(t,s) &= R_{\boldsymbol{x}}(t-s) \end{aligned}$$

In other words, a s.p. is weakly stationary if its mean is constant and its covariance function depends only on the difference of the two time instants t and s. For this reason, the covariance function of a stationary s.p. is often expressed as a function of the *lag* $\tau = t - s$, as

$$R_{\boldsymbol{x}}(\tau) = \mathbf{E}\left[(\boldsymbol{x}(t+\tau) - m_{\boldsymbol{x}})(\boldsymbol{x}(t) - m_{\boldsymbol{x}})^T\right]. \tag{2.1}$$

Clearly, strong stationarity implies weak stationarity, while the vice versa is not true. In these notes, we will always consider weak stationarity: a process $\boldsymbol{x}(t)$ will be simply called stationary, when it is weakly stationary.

Definition 2.7. Two stochastic processes $\boldsymbol{x}(t)$ and $\boldsymbol{y}(t)$ are *jointly weakly stationary* if they are both weakly stationary and their cross-covariance function is invariant with respect to time shifts, i.e.

$$R_{\boldsymbol{xy}}(t,s) = R_{\boldsymbol{xy}}(t+\delta, s+\delta), \quad \forall \delta.$$

Similarly to what has been done for the covariance function of a stationary s.p., we express the cross-covariance of two jointly stationary processes as a function of the lag $\tau = t - s$, i.e.

$$R_{\boldsymbol{xy}}(\tau) = \mathbf{E}\left[(\boldsymbol{x}(t+\tau) - m_{\boldsymbol{x}})(\boldsymbol{y}(t) - m_{\boldsymbol{y}})^T\right].$$

Let us stress that, when we are dealing with vectors of stochastic processes, the mean is a vector and the covariance and cross-covariance functions are matrices. For instance, if

$$\boldsymbol{x}(t) = \begin{bmatrix} \boldsymbol{x}_1(t) \\ \vdots \\ \boldsymbol{x}_n(t) \end{bmatrix} \quad , \quad \boldsymbol{y}(t) = \begin{bmatrix} \boldsymbol{y}_1(t) \\ \vdots \\ \boldsymbol{y}_p(t) \end{bmatrix}$$

one has

$$\mathbf{E}[\boldsymbol{x}(t)] = m_{\boldsymbol{x}}(t) = \begin{bmatrix} m_{\boldsymbol{x}_1}(t) \\ \vdots \\ m_{\boldsymbol{x}_n}(t) \end{bmatrix} \in \mathbb{R}^n$$

and

$$\begin{aligned} R_{\boldsymbol{x}}(t,s) &= \mathbf{E}\left[(\boldsymbol{x}(t) - m_{\boldsymbol{x}}(t))(\boldsymbol{x}(s) - m_{\boldsymbol{x}}(s))^T\right] \\ &= \begin{bmatrix} R_{\boldsymbol{x}_1}(t,s) & R_{\boldsymbol{x}_1\boldsymbol{x}_2}(t,s) & \dots & R_{\boldsymbol{x}_1\boldsymbol{x}_n}(t,s) \\ R_{\boldsymbol{x}_2\boldsymbol{x}_1}(t,s) & R_{\boldsymbol{x}_2}(t,s) & \dots & R_{\boldsymbol{x}_2\boldsymbol{x}_n}(t,s) \\ \vdots & & & \\ R_{\boldsymbol{x}_n\boldsymbol{x}_1}(t,s) & \dots & \dots & R_{\boldsymbol{x}_n}(t,s) \end{bmatrix} \in \mathbb{R}^{n \times n} \end{aligned}$$

In a similar way, one can compute $R_{\boldsymbol{xy}}(t,s) : \mathbb{Z} \times \mathbb{Z} \to \mathbb{R}^{n \times p}$.

Theorem 2.1. *Let $\boldsymbol{x}(t)$ be a stationary s.p.; then, for its covariance function $R_{\boldsymbol{x}}(\tau)$ the following properties hold:*

1. $R_{\boldsymbol{x}}(\tau) = R_{\boldsymbol{x}}^T(-\tau)$.

2. $R_{\boldsymbol{x}}(0)$ *has positive elements on the diagonal.*

3. *The matrix*

$$P(m) = \begin{bmatrix} R_{\boldsymbol{x}}(0) & R_{\boldsymbol{x}}(1) & \dots & R_{\boldsymbol{x}}(m-1) \\ R_{\boldsymbol{x}}(-1) & R_{\boldsymbol{x}}(0) & \dots & R_{\boldsymbol{x}}(m-2) \\ \vdots & \vdots & \ddots & \vdots \\ R_{\boldsymbol{x}}(1-m) & R_{\boldsymbol{x}}(2-m) & \dots & R_{\boldsymbol{x}}(0) \end{bmatrix}$$

is symmetric and positive semidefinite, $P(m) \geq 0$, $\forall m \in \mathbb{N}_+$.

4. *If $\boldsymbol{y}(t)$ is a s.p. jointly stationary with $\boldsymbol{x}(t)$; the cross-covariance function $R_{\boldsymbol{xy}}(\tau)$ satisfies*

$$R_{\boldsymbol{xy}}(\tau) = R_{\boldsymbol{yx}}^T(-\tau).$$

Proof

Without loss of generality, let $m_{\boldsymbol{x}} = 0$.

1. From (2.1) one has

$$\begin{aligned} R_{\boldsymbol{x}}(\tau) &= \mathbf{E}\left[\boldsymbol{x}(t+\tau)\boldsymbol{x}(t)^T\right] \overset{[1]}{=} \mathbf{E}\left[\boldsymbol{x}(s)\boldsymbol{x}(s-\tau)^T\right] \\ &\overset{[2]}{=} \mathbf{E}\left[\left(\boldsymbol{x}(s-\tau)\boldsymbol{x}(s)^T\right)^T\right] = R_{\boldsymbol{x}}^T(-\tau), \end{aligned}$$

in which [1] is obtained by setting $s = t + \tau$ and [2] from the property $ab^T = \left(ba^T\right)^T$, with a, b real vectors.

2. Let $\boldsymbol{x}(t) = [\boldsymbol{x}_1(t), \dots, \boldsymbol{x}_n(t)]^T \in \mathbb{R}^n$. Then, the elements on the diagonal of $R_{\boldsymbol{x}}(0)$ are the variances $R_{\boldsymbol{x}_i}(0) = \mathbf{E}\left[\boldsymbol{x}_i(t)^2\right] = \sigma_{\boldsymbol{x}_i}^2 > 0$.

3. First, observe that one can write

$$P(m) = \mathbf{E}\left[\begin{pmatrix} \boldsymbol{x}(t-1) \\ \vdots \\ \boldsymbol{x}(t-m) \end{pmatrix} \left(\boldsymbol{x}(t-1)^T, \dots, \boldsymbol{x}(t-m)^T\right)\right].$$

Let $v = (v_1^T, \ldots, v_m^T)^T$, with $v_i \in \mathbb{R}^n$. Then,

$$\begin{aligned}
v^T P(m) v &= (v_1^T, \ldots, v_m^T)\mathbf{E}\left[\begin{pmatrix} \boldsymbol{x}(t-1) \\ \vdots \\ \boldsymbol{x}(t-m) \end{pmatrix} \left(\boldsymbol{x}(t-1)^T, \ldots, \boldsymbol{x}(t-m)^T\right)\right] \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix} \\
&= \mathbf{E}\left[\left(\sum_{k=1}^{m} v_k^T \boldsymbol{x}(t-k)\right)\left(\sum_{k=1}^{m} \boldsymbol{x}(t-k)^T v_k\right)\right] \\
&= \mathbf{E}\left[\left(\sum_{k=1}^{m} v_k^T \boldsymbol{x}(t-k)\right)^2\right] \geq 0
\end{aligned}$$

Since the previous expression hold for every v, it follows that $P(m) \geq 0$.

4. See the proof of item 1. □

If $\boldsymbol{x}(t)$ is a scalar s.p., property 1 in Theorem 2.1 becomes

$$R_{\boldsymbol{x}}(\tau) = R_{\boldsymbol{x}}(-\tau), \tag{2.2}$$

i.e., the covariance function of a stationary scalar s.p. is a symmetric function with respect to τ. Moreover, consider the linear combination $\alpha_1 \boldsymbol{x}(t+\tau) + \alpha_2 \boldsymbol{x}(t)$, with $\tau \in \mathbb{N}$, and assume $m_{\boldsymbol{x}} = 0$. then,

$$\begin{aligned}
0 &\leq \mathbf{E}\left[(\alpha_1 \boldsymbol{x}(t+\tau) + \alpha_2 \boldsymbol{x}(t))^2\right] \\
&= \mathbf{E}\left[\left(\alpha_1^2 \boldsymbol{x}(t+\tau)^2 + 2\alpha_1\alpha_2 \boldsymbol{x}(t+\tau)\boldsymbol{x}(t) + \alpha_2^2 \boldsymbol{x}(t)^2\right)\right] \\
&= \alpha_1^2 \mathbf{E}\left[\boldsymbol{x}(t+\tau)^2\right] + 2\alpha_1\alpha_2 \mathbf{E}\left[\boldsymbol{x}(t+\tau)\boldsymbol{x}(t)\right] + \alpha_2^2 \mathbf{E}\left[\boldsymbol{x}(t)^2\right] \\
&= \alpha_1^2 R_{\boldsymbol{x}}(0) + 2\alpha_1\alpha_2 R_{\boldsymbol{x}}(\tau) + \alpha_2^2 R_{\boldsymbol{x}}(0) = (\alpha_1^2 + \alpha_2^2) R_{\boldsymbol{x}}(0) + 2\alpha_1\alpha_2 R_{\boldsymbol{x}}(\tau).
\end{aligned}$$

Therefore, $(\alpha_1^2 + \alpha_2^2) R_{\boldsymbol{x}}(0) + 2\alpha_1\alpha_2 R_{\boldsymbol{x}}(\tau) \geq 0$. It is straightforward to verify that the latter inequality can be rewritten as

$$\begin{bmatrix} \alpha_1 & \alpha_2 \end{bmatrix} \begin{bmatrix} R_{\boldsymbol{x}}(0) & R_{\boldsymbol{x}}(\tau) \\ R_{\boldsymbol{x}}(\tau) & R_{\boldsymbol{x}}(0) \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} \geq 0. \tag{2.3}$$

(notice also that (2.3) stems directly from the fact that $P(m) \geq 0$). Since (2.3) hold for every $\alpha_1,\ \alpha_2 \in \mathbb{R}$, the matrix $\begin{bmatrix} R_{\boldsymbol{x}}(0) & R_{\boldsymbol{x}}(\tau) \\ R_{\boldsymbol{x}}(\tau) & R_{\boldsymbol{x}}(0) \end{bmatrix}$ is positive semidefinite. Hence, its determinant is nonnegative

$$\det \begin{bmatrix} R_{\boldsymbol{x}}(0) & R_{\boldsymbol{x}}(t) \\ R_{\boldsymbol{x}}(\tau) & R_{\boldsymbol{x}}(0) \end{bmatrix} = R_{\boldsymbol{x}}(0)^2 - R_{\boldsymbol{x}}(\tau)^2 \geq 0.$$

Being $R_{\boldsymbol{x}}(0) > 0$, one gets

$$|R_{\boldsymbol{x}}(\tau)| \leq R_{\boldsymbol{x}}(0), \quad \forall \tau \in \mathbb{N}.$$

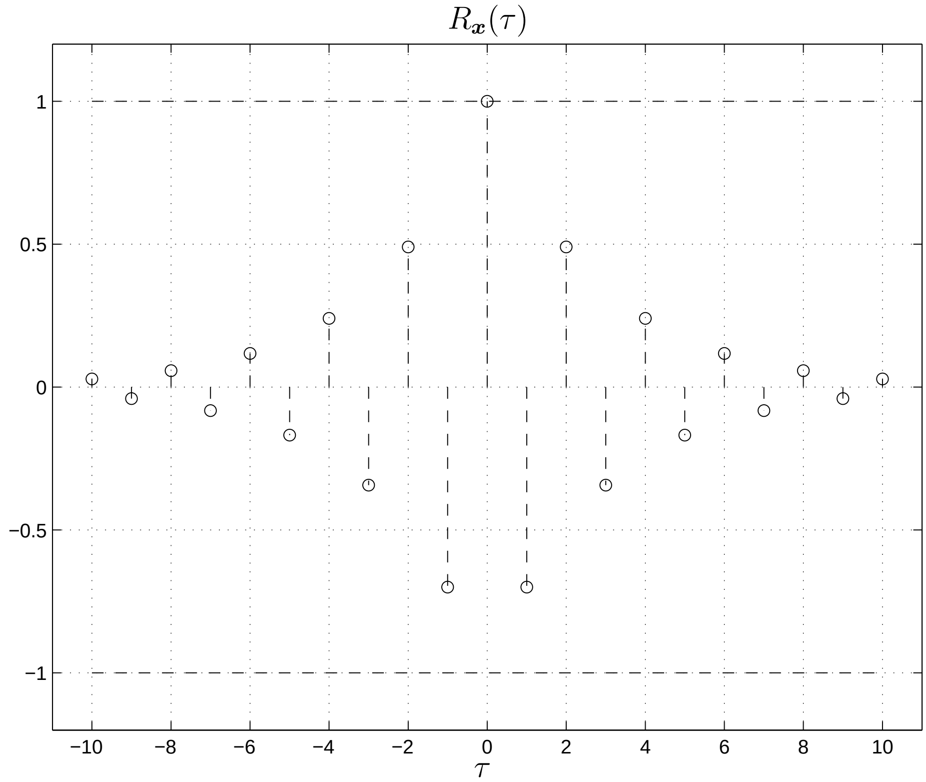


Figure 2.2: An example of covariance function of a scalar stochastic process.

Therefore, we can conclude that the covariance function of a scalar s.p. is bounded in the interval $[-R_{\boldsymbol{x}}(0),\ R_{\boldsymbol{x}}(0)]$ and takes on its maximum value for $\tau = 0$ (an example is shown in Figure 2.2).

Notice that $R_{\boldsymbol{x}}(0)$ represents the variance of the stationary s.p. $\boldsymbol{x}(t)$ (or its covariance matrix, if $\boldsymbol{x}(t)$ is a vector).

Definition 2.8. A stochastic process is *Gaussian* if its statistics

$$F_{\boldsymbol{x}}(x_1, \ldots, x_k; t_1, \ldots, t_k) \qquad f_{\boldsymbol{x}}(x_1, \ldots, x_k; t_1, \ldots, t_k)$$

are Gaussian, for all $k \in \mathbb{N}_+$. A Gaussian s.p. is completely characterized if its mean and covariance functions are known. Moreover, for Gaussian processes, weak stationarity coincides with strong stationarity.

Definition 2.9. A stochastic process is *white* (also referred to as *white process*) if it is a sequence of independent random variables. If the random variables are also identically distributed, we refer to it as an *i.i.d.* stochastic process.

2.2 Examples of stochastic processes

In this section, some examples of stochastic processes are presented.

2.2.1 Purely deterministic stochastic processes

Let us consider a stochastic process of the form

$$\boldsymbol{x}(t) = \sum_{i=1}^{m} \boldsymbol{x}_i g_i(t),$$

in which $\boldsymbol{x}_i$ are random variables and $g_i(\cdot)$ deterministic functions of time. Once the realizations of the random variables $\boldsymbol{x}_i$ have been observed, such a process is completely predictable in the future because, that is the s.p. $\boldsymbol{x}(t)$ is known $\forall t \in \mathcal{T}$.

Let us see some examples of purely deterministic stochastic processes.

- $\boldsymbol{x}(t) = \boldsymbol{x}, \quad \forall t \in \mathcal{T}$,
  in which $\boldsymbol{x}$ is a random variable with mean $m_{\boldsymbol{x}}$ and variance $\sigma^2_{\boldsymbol{x}}$. The mean of the s.p. $\boldsymbol{x}(t)$ is clearly

  $$\mathbf{E}\left[\boldsymbol{x}(t)\right] = \mathbf{E}\left[\boldsymbol{x}\right] = m_{\boldsymbol{x}},$$

  while the covariance function is given by

  $$R_{\boldsymbol{x}}\left(t,s\right) = \mathbf{E}\left[(\boldsymbol{x}(t) - m_{\boldsymbol{x}})(\boldsymbol{x}(s) - m_{\boldsymbol{x}})\right] = \mathbf{E}\left[(\boldsymbol{x} - m_{\boldsymbol{x}})^2\right] = \sigma^2_{\boldsymbol{x}}.$$

  Hence, $R_{\boldsymbol{x}}\left(t,s\right)$ does not depend on t and s, but it is constant

  $$R_{\boldsymbol{x}}(\tau) = \sigma^2_{\boldsymbol{x}}, \quad \forall \tau.$$

  Therefore, the s.p. $\boldsymbol{x}(t)$ is (weakly) stationary.

- $\boldsymbol{x}(t) = A\cos(\omega t + \boldsymbol{\varphi})$,
  with $\boldsymbol{\varphi}$ random variable uniformly distributed in the interval $[0,\ 2\pi]$. The mean of the s.p. $\boldsymbol{x}(t)$ takes on the value

  $$\mathbf{E}\left[\boldsymbol{x}(t)\right] = \frac{1}{2\pi}\int_0^{2\pi} A\cos(\omega t + \varphi)\mathrm{d}\varphi = 0,$$

  while the covariance function is

  $$R_{\boldsymbol{x}}\left(t,s\right) = \mathbf{E}\left[\boldsymbol{x}(t)\boldsymbol{x}(s)\right] = \frac{1}{2\pi}\int_0^{2\pi} A^2\cos(\omega t + \varphi)\cos(\omega s + \varphi)\mathrm{d}\varphi.$$

  Since $\cos(\alpha)\cos(\beta) = \frac{1}{2}\left[\cos(\alpha+\beta) + \cos(\alpha-\beta)\right]$, one has

  $$\begin{aligned} R_{\boldsymbol{x}}\left(t,s\right) &= \frac{A^2}{4\pi}\int_0^{2\pi} \left[\cos(\omega t + \omega s + 2\varphi) + \cos(\omega t - \omega s)\right]\mathrm{d}\varphi \\ &= \frac{A^2}{4\pi}2\pi\cos[\omega(t-s)] = \frac{A^2}{2}\cos(\omega\tau) \triangleq R_{\boldsymbol{x}}(\tau). \end{aligned}$$

  Therefore, the s.p. $\boldsymbol{x}(t)$ is (weakly) stationary.

- $\boldsymbol{x}(t) = \boldsymbol{A}\sin(\omega t)$,
  with $\boldsymbol{A} \sim f_{\boldsymbol{A}}(a)$ random variable with mean $m_{\boldsymbol{A}} = 0$ and variance $\sigma^2_{\boldsymbol{A}}$. The mean of the s.p. $\boldsymbol{x}(t)$ is

$$\begin{aligned}\mathbf{E}[\boldsymbol{x}(t)] &= \int_{-\infty}^{+\infty} a\sin(\omega t) f_{\boldsymbol{A}}(a)\,\mathrm{d}a \\ &= \sin(\omega t)\int_{-\infty}^{+\infty} a f_{\boldsymbol{A}}(a)\,\mathrm{d}a = \sin(\omega t) m_{\boldsymbol{A}} = 0,\end{aligned}$$

  while the covariance function

$$\begin{aligned}R_{\boldsymbol{x}}(t,s) = \mathbf{E}[\boldsymbol{x}(t)\boldsymbol{x}(s)] &= \int_{-\infty}^{+\infty} a^2 \sin(\omega t)\sin(\omega s) f_{\boldsymbol{A}}(a)\,\mathrm{d}a \\ &= \sin(\omega t)\sin(\omega s)\int_{-\infty}^{+\infty} a^2 f_{\boldsymbol{A}}(a)\,\mathrm{d}a = \sigma^2_{\boldsymbol{A}}\sin(\omega t)\sin(\omega s) \\ &= \frac{\sigma^2_{\boldsymbol{A}}}{2}\left[\cos(\omega(t-s)) - \cos(\omega(t+s))\right],\end{aligned}$$

  In which the last equality stems from $\sin(\alpha)\sin(\beta) = \frac{1}{2}[\cos(\alpha-\beta) - \cos(\alpha+\beta)]$. Therefore, the s.p. $\boldsymbol{x}(t)$ is not stationary.

The last example shows that a purely deterministic process may not be stationary.

2.2.2 White process

A white process $\boldsymbol{x}(t)$ is a sequence of independent random variables. Therefore, $\boldsymbol{x}(t_1)$ and $\boldsymbol{x}(t_2)$ are independent whenever $t_1 \neq t_2$. Such a process is completely unpredictable in the future.

By assuming the variables $\boldsymbol{x}(t_i)$ are *i.i.d.* with zero mean and variance $\sigma^2_{\boldsymbol{x}}$, one has

$$m_{\boldsymbol{x}}(t) = \mathbf{E}[\boldsymbol{x}(t)] = 0,$$

and

$$R_{\boldsymbol{x}}(t,s) = \begin{cases}\mathbf{E}\left[\boldsymbol{x}(t)^2\right] = \sigma^2_{\boldsymbol{x}} & \text{if } t = s \\ 0 & \text{if } t \neq s\end{cases} = \sigma^2_{\boldsymbol{x}}\delta_{t-s},$$

in which δ_τ denotes the Kronecker unit impulse

$$\delta_\tau = \begin{cases}1 & \text{if } \tau = 0 \\ 0 & \text{if } \tau \neq 0\end{cases}.$$

Hence, by setting $\tau = t - s$, one can write the covariance function as

$$R_{\boldsymbol{x}}(\tau) = \sigma^2_{\boldsymbol{x}}\delta_\tau$$

and therefore $\boldsymbol{x}(t)$ is a stationary process.

An example of white process is given by binary transmission. Let $\boldsymbol{x}(t) \in \{-1, 1\}$, with $t \in \mathbb{N}$ and $\boldsymbol{x}(t_1)$, $\boldsymbol{x}(t_2)$ independent if $t_1 \neq t_2$. By assuming that both values

$-1, 1$ can occur with the same probability, the first order statistic of $\boldsymbol{x}(t)$ is given by:

$$f_{\boldsymbol{x}}(x;t) = \frac{1}{2}\delta(x-1) + \frac{1}{2}\delta(x+1).$$

where $\delta(t)$ is the Dirac delta function. The mean of the s.p. $\boldsymbol{x}(t)$ takes on the value

$$\mathbf{E}[\boldsymbol{x}(t)] = \int_{-\infty}^{+\infty} x\left(\frac{1}{2}\delta(x-1) + \frac{1}{2}\delta(x+1)\right)\mathrm{d}x = \frac{1}{2} - \frac{1}{2} = 0,$$

in which we have exploited the property of the Dirac function

$$\int_{-\infty}^{+\infty} g(x)\delta(x-\alpha)\mathrm{d}x = g(\alpha).$$

The variance of $\boldsymbol{x}(t)$ is equal to

$$R_{\boldsymbol{x}}(t,t) = \mathbf{E}\left[\boldsymbol{x}(t)^2\right] = \int_{-\infty}^{+\infty} x^2\left(\frac{1}{2}\delta(x-1) + \frac{1}{2}\delta(x+1)\right)\mathrm{d}x = \frac{1}{2} + \frac{1}{2} = 1.$$

Being $\boldsymbol{x}(t_1)$ and $\boldsymbol{x}(t_2)$ independent whenever $t_1 \neq t_2$, they are also uncorrelated, and hence

$$R_{\boldsymbol{x}}(t_1,t_2) = \mathbf{E}[\boldsymbol{x}(t_1)\boldsymbol{x}(t_2)] = 0.$$

Summing up, the covariance function of $\boldsymbol{x}(t)$ is given by

$$R_{\boldsymbol{x}}(\tau) = \begin{cases} 1 & \text{se } \tau = 0 \\ 0 & \text{se } \tau \neq 0 \end{cases}$$

and hence the process $\boldsymbol{x}(t)$ is stationary.

White stochastic processes play a fundamental role in the following treatment. In particular, we will see that a broad class of stationary stochastic processes can be obtained by filtering a white process with a suitable linear system. Hereafter, a white process $\boldsymbol{e}(t)$ with constant mean $m_{\boldsymbol{e}}$ and variance $\sigma_{\boldsymbol{e}}^2$, will be referred to as

$$\boldsymbol{e}(t) \sim WP(m_{\boldsymbol{e}}, \sigma_{\boldsymbol{e}}^2).$$

2.2.3 Wiener process

A Wiener process $\boldsymbol{w}(t)$, also referred to as *Brownian motion* or *random walk*, has the following characteristics

- $\boldsymbol{w}(0) = 0$;
- $\mathbf{E}[\boldsymbol{w}(t)] = 0$;
- $R_{\boldsymbol{w}}(t_1,t_2) = \alpha \min\{t_1, t_2\} = \begin{cases} \alpha t_1 & \text{if } t_1 \leq t_2 \\ \alpha t_2 & \text{if } t_1 \geq t_2 \end{cases}$

Notice that the variance of $\boldsymbol{w}(t)$ is equal to $R_{\boldsymbol{w}}(t,t) = \alpha t$, which means that the "spread" of the realizations grows linearly with time, and hence the s.p. is not stationary. Let $t_1 > t_2 > t_3$; the following property holds

$$\begin{aligned} &\mathbf{E}\left[(\boldsymbol{w}(t_1) - \boldsymbol{w}(t_2))\,(\boldsymbol{w}(t_2) - \boldsymbol{w}(t_3))\right] \\ &\quad = R_{\boldsymbol{w}}(t_1,t_2) - R_{\boldsymbol{w}}(t_1,t_3) - R_{\boldsymbol{w}}(t_2,t_2) + R_{\boldsymbol{w}}(t_2,t_3) \\ &\quad = \alpha t_2 - \alpha t_3 - \alpha t_2 + \alpha t_3 = 0. \end{aligned}$$

For this reason, a Wiener process is also referred to as a process with uncorrelated increments. Such a process can be obtained by summing a white process, i.e.

$$\boldsymbol{w}(t) = \sum_{k=1}^{t} \boldsymbol{e}(k)$$

with $\boldsymbol{e}(t) \sim WP(0,1)$.

2.2.4 Exponentially correlated processes

Let us consider the stationary s.p. $\boldsymbol{x}(t)$ whose covariance function takes on the form

$$R_{\boldsymbol{x}}(\tau) = \sigma_{\boldsymbol{x}}^2 a^{|\tau|}, \quad \tau \in \mathbb{Z}, \tag{2.4}$$

with $|a| < 1$. The variance of the s.p. $\boldsymbol{x}(t)$ is $R_{\boldsymbol{x}}(0) = \sigma_{\boldsymbol{x}}^2$. Notice that:

- if $a \to 1$, $x(t)$ tends to a purely deterministic process;
- if $a \to 0$, $x(t)$ tends to a process of uncorrelated variables (which is also a white process if $\boldsymbol{x}(t)$ is Gaussian).

The covariance function depicted in Figure 2.2 corresponds to (2.4), with $\sigma_{\boldsymbol{x}}^2 = 1$ and $a = -0.7$.

Let us fix a time instant t and consider the random variables $\boldsymbol{x}(t)$ and $\boldsymbol{x}(t+1)$. Being $\boldsymbol{x}(t)$ stationary, the mean and variance of the two random variables are the same, i.e.

$$\mathbf{E}\left[\boldsymbol{x}(t)\right] = \mathbf{E}\left[\boldsymbol{x}(t+1)\right] = m_{\boldsymbol{x}}$$
$$\mathbf{E}\left[(\boldsymbol{x}(t) - m_{\boldsymbol{x}})^2\right] = \mathbf{E}\left[(\boldsymbol{x}(t+1) - m_{\boldsymbol{x}})^2\right] = \sigma_{\boldsymbol{x}}^2$$

Let us now compute the cross-covariance of $\boldsymbol{x}(t+1)$ and $\boldsymbol{x}(t)$:

$$\mathbf{E}\left[(\boldsymbol{x}(t+1) - m_{\boldsymbol{x}})(\boldsymbol{x}(t) - m_{\boldsymbol{x}})\right] \triangleq R_{\boldsymbol{x}}(1) = \sigma_{\boldsymbol{x}}^2 a,$$

in which the last equality stems from (2.4) with $\tau = 1$. Then, the correlation coefficient between $\boldsymbol{x}(t+1)$ and $\boldsymbol{x}(t)$ is equal to

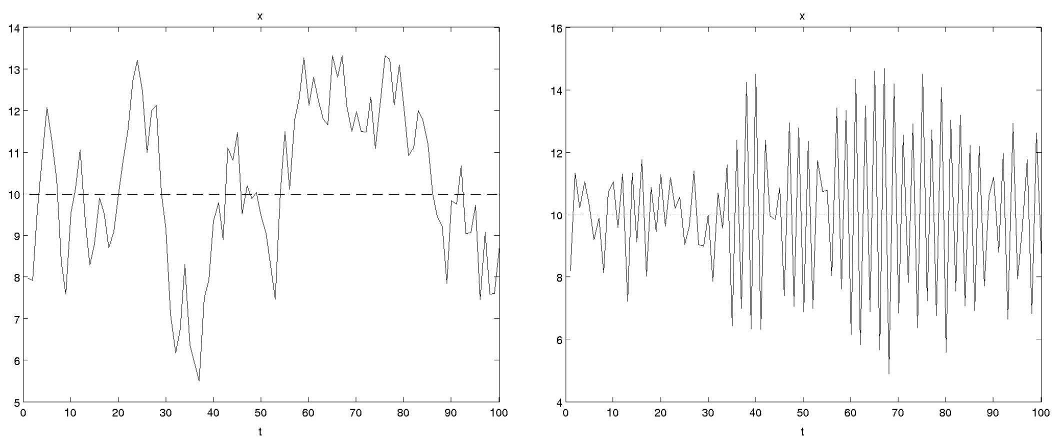
$$\rho = \frac{R_{\boldsymbol{x}}(1)}{\sigma_{\boldsymbol{x}}^2} = \frac{\sigma_{\boldsymbol{x}}^2 a}{\sigma_{\boldsymbol{x}}^2} = a.$$

Hence, the parameter a in (2.4) corresponds to the correlation coefficient between the values of the stochastic process $\boldsymbol{x}(t)$ at two consecutive time instants. From a qualitative point of view, a consequence of this property is that

- if $a > 0$, it is more likely that the sign of $\boldsymbol{x}(t) - m_{\boldsymbol{x}}$ does not change in two consecutive time instants;

- if $a < 0$, it is more likely that the sign of $\boldsymbol{x}(t) - m_{\boldsymbol{x}}$ changes in two consecutive time instants .

Assume $x(t) > m_{\boldsymbol{x}}$. As a approaches 1, the probability that $\boldsymbol{x}(t+1) > m_{\boldsymbol{x}}$ grows. Similarly, as a approaches -1, the probability that $\boldsymbol{x}(t+1) < m_{\boldsymbol{x}}$ will increase. In the special case $a = 1$, we have that $\boldsymbol{x}(t)$ and $\boldsymbol{x}(t+1)$ are fully correlated, meaning that by observing $\boldsymbol{x}(t)$ we can infer the exact value taken by $\boldsymbol{x}(t+1)$.



(a) (b)

Figure 2.3: Realizations of exponentially correlated stochastic processes with: (a) $a = 0.9$; (b) $a = -0.9$.

Figure 2.3 reports two realizations of two different exponentially correlated s.p., both with mean equal to 10, but with opposite correlation coefficients. It is apparent that when $a = -0.9$, the difference $\boldsymbol{x}(t) - m_{\boldsymbol{x}}$ changes its sign much more often than what happens when $a = 0.9$.

The class of exponentially correlated stochastic process is much broader than the example shown above. In general, all stationary processes whose covariance function decreases exponentially with the lag τ belong to this class. Moreover, such processes can be obtained *asymptotically* as outputs of linear time-invariant (LTI) dynamic systems, whose input is a white process, as it is shown in the next example.

Example 2.1. Let the s.p. $\boldsymbol{y}(t)$ be the solution of the difference equation:

$$y(t+1) = ay(t) + e(t), \quad |a| < 1,$$

where $\boldsymbol{e}(t)$ is a white process with zero mean and variance $\sigma_{\boldsymbol{e}}^2$. We can model $\boldsymbol{y}(t)$ as the output of an LTI system fed by the input $\boldsymbol{e}(t)$ (see Figure 2.4), with transfer function $G(z) = \dfrac{1}{z-a}$.

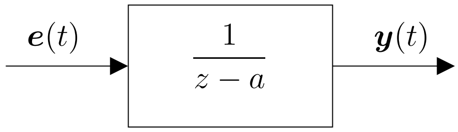


Figure 2.4. A linear stochastic system

Let us compute the mean $m_{\boldsymbol{y}}(t)$ and the covariance function $R_{\boldsymbol{y}}(t+\tau, t)$, in order to verify if the s.p. $\boldsymbol{y}(t)$ is stationary. From systems theory, the output of an LTI system is given by the convolution between the impulse response and the input. In the considered example, one has

$$\boldsymbol{y}(t) = \sum_{i=0}^{\infty} a^i \boldsymbol{e}(t-i-1).$$

Notice that, since $|a| < 1$, the series $\sum_{i=0}^{\infty} a^i$ converges. Then, by recalling that $m_{\boldsymbol{e}}=0$, one has

$$m_{\boldsymbol{y}}(t) = \mathbf{E}\left[\sum_{i=0}^{\infty} a^i \boldsymbol{e}(t-i-1)\right] = \sum_{i=0}^{\infty} a^i \mathbf{E}\left[\boldsymbol{e}(t-i-1)\right] = 0.$$

Concerning the covariance function, one has

$$R_{\boldsymbol{y}}(t+\tau, t) = \mathbf{E}\left[\left(\sum_{i=0}^{\infty} a^i \boldsymbol{e}(t+\tau-i-1)\right)\left(\sum_{j=0}^{\infty} a^j \boldsymbol{e}(t-j-1)\right)\right].$$

Once again, from the linearity of the expectation, we get

$$\begin{aligned} R_{\boldsymbol{y}}(t+\tau, t) &= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} a^i a^j \mathbf{E}\left[\boldsymbol{e}(t+\tau-i-1)\boldsymbol{e}(t-j-1)\right] \\ &= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} a^{i+j} R_{\boldsymbol{e}}(\tau-i+j), \end{aligned} \tag{2.5}$$

Being $\boldsymbol{e}(t) \sim WP(0, \sigma_{\boldsymbol{e}}^2)$, its covariance function is

$$R_{\boldsymbol{e}}(\nu) = \begin{cases} \sigma_{\boldsymbol{e}}^2 & \text{se } \nu = 0 \\ 0 & \text{se } \nu \neq 0 \end{cases}$$

Therefore, in the second sum in (2.5) the only nonnegative terms are those with indexes i, j such that $\tau - i + j = 0$, i.e., $j = i - \tau$. Moreover, since the index j can take only nonnegative values, if $\tau \geq 0$, one may have $j = i - \tau$ only if $i \geq \tau$. Hence, one has

$$
\begin{aligned}
R_{\boldsymbol{y}}(t+\tau, t) &= \begin{cases} \displaystyle\sum_{i=\tau}^{\infty} a^{2i-\tau}\sigma_{\boldsymbol{e}}^2 & \text{if } \tau \geq 0 \\ \displaystyle\sum_{i=0}^{\infty} a^{2i-\tau}\sigma_{\boldsymbol{e}}^2 & \text{if } \tau < 0 \end{cases} = \begin{cases} a^{\tau}\sigma_{\boldsymbol{e}}^2 \displaystyle\sum_{i=\tau}^{\infty} a^{2i-2\tau} & \text{if } \tau \geq 0 \\ a^{-\tau}\sigma_{\boldsymbol{e}}^2 \displaystyle\sum_{i=0}^{\infty} a^{2i} & \text{if } \tau < 0 \end{cases} \\
&\overset{[1]}{=} \begin{cases} a^{\tau}\sigma_{\boldsymbol{e}}^2 \displaystyle\sum_{k=0}^{\infty} \left(a^2\right)^k & \text{if } \tau \geq 0 \\ a^{-\tau}\sigma_{\boldsymbol{e}}^2 \displaystyle\sum_{i=0}^{\infty} \left(a^2\right)^i & \text{if } \tau < 0 \end{cases} \overset{[2]}{=} \begin{cases} \dfrac{\sigma_{\boldsymbol{e}}^2}{1-a^2} a^{\tau} & \text{if } \tau \geq 0 \\ \dfrac{\sigma_{\boldsymbol{e}}^2}{1-a^2} a^{-\tau} & \text{if } \tau < 0 \end{cases} \\
&= \frac{\sigma_{\boldsymbol{e}}^2}{1-a^2} a^{|\tau|}
\end{aligned}
$$

Equality [1] is obtained by setting $k = i - \tau$, while [2] stems from $\sum_{i=0}^{\infty} p^i = \frac{1}{1-p}$, if $|p| < 1$.

Summing up:

- $m_{\boldsymbol{y}}(t) = 0$;
- $R_{\boldsymbol{y}}(t+\tau, t) = R_{\boldsymbol{y}}(\tau) = \dfrac{\sigma_{\boldsymbol{e}}^2}{1-a^2} a^{|\tau|}$;

and therefore $\boldsymbol{y}(t)$ is a stationary process. △

Observation 2.1. It is worth stressing that in the previous analysis we have implicitly assumed that the output $y(t)$ depends on all past values of the input $e(t)$ up to time $t = -\infty$. This amounts to consider the system at steady state, neglecting the effect of the transient, or equivalently to assume that the initial condition has occurred a very long time before the current time instant t. This is admissible, due to the fact that the considered LTI system is asymptotically stable: therefore, after a sufficiently long time, the effect of the initial condition is indeed negligible. If one is interested in the analysis of the transient behavior (by setting the initial condition at a finite time, say $t = 0$), the output $y(t)$ is only *asymptotically stationary*, which means that:

- $\lim_{t \to \infty} m_{\boldsymbol{y}}(t)$ exists and is finite;
- $\lim_{t \to \infty} R_{\boldsymbol{y}}(t+\tau, t)$ exists and is finite for every τ.

2.3 Frequency analysis

In this paragraph, some useful tools for the analysis of stochastic processes in the frequency domain are introduced.

Definition 2.10. Let $\boldsymbol{x}(t)$ be a stationary stochastic process with covariance function $R_{\boldsymbol{x}}(\tau)$. The *spectrum* of $\boldsymbol{x}(t)$ is the function of the complex variable $z \in \mathbb{C}$:

$$\phi_{\boldsymbol{x}}(z) = \sum_{\tau=-\infty}^{\infty} R_{\boldsymbol{x}}(\tau) z^{-\tau}.$$

The spectrum $\phi_{\boldsymbol{x}}(z)$ is the bilateral Z-transform of the covariance function of $\boldsymbol{x}(t)$.

Definition 2.11. The *spectral density* of the stationary s.p. $\boldsymbol{x}(t)$ is the function of the real variable $\omega \in \mathbb{R}$

$$\phi_{\boldsymbol{x}}(e^{j\omega}) = \sum_{\tau=-\infty}^{\infty} R_{\boldsymbol{x}}(\tau) e^{-j\omega\tau}.$$

The spectral density is the spectrum computed on the unit circumference. Hence, the spectral density is a periodic function with period 2π. For this reason, we will analyze it in the interval $\omega \in [-\pi,\ \pi]$.

Definition 2.12. Let $\boldsymbol{x}(t)$ and $\boldsymbol{y}(t)$ two jointly stationary stochastic processes and let $R_{\boldsymbol{xy}}(\tau)$ be their cross-covariance function. Then, the *cross-spectrum* of $\boldsymbol{x}(t)$ and $\boldsymbol{y}(t)$ is the function of the complex variable $z \in \mathbb{C}$:

$$\phi_{\boldsymbol{xy}}(z) = \sum_{\tau=-\infty}^{\infty} R_{\boldsymbol{xy}}(\tau) z^{-\tau}.$$

Given the spectrum of a s.p. $\boldsymbol{x}(t)$, it is possible to recover its covariance function $R_{\boldsymbol{x}}(\tau)$. Let us compute the integral

$$\frac{1}{2\pi j} \oint \phi_{\boldsymbol{x}}(z) z^{k-1} \mathrm{d}z. \tag{2.6}$$

where the symbol $\oint$ denotes the integral of a function in the complex domain, computed along the unit circumference. Therefore, we can set $z = e^{j\omega}$, with $\omega \in [-\pi,\ \pi]$. Since $\mathrm{d}z = je^{j\omega}\mathrm{d}\omega$, (2.6) becomes

$$\begin{aligned} \frac{1}{2\pi j} \oint \phi_{\boldsymbol{x}}(z) z^{k-1} \mathrm{d}z &= \frac{1}{2\pi j} \int_{-\pi}^{\pi} \phi_{\boldsymbol{x}}(e^{j\omega}) e^{j\omega(k-1)} j e^{j\omega} \mathrm{d}\omega \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \phi_{\boldsymbol{x}}(e^{j\omega}) e^{j\omega k} \mathrm{d}\omega \end{aligned} \tag{2.7}$$

$$
\begin{aligned}
&\overset{[1]}{=} \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{\tau=-\infty}^{\infty} R_{\boldsymbol{x}}(\tau) e^{-j\omega\tau} e^{j\omega k} \mathrm{d}\omega \\
&= \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} R_{\boldsymbol{x}}(\tau) \int_{-\pi}^{\pi} e^{j\omega(k-\tau)} \mathrm{d}\omega \\
&\overset{[2]}{=} \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} R_{\boldsymbol{x}}(\tau) 2\pi \delta_{k-\tau} \\
&= R_{\boldsymbol{x}}(k).
\end{aligned}
$$

Notice that [1] stems from Definition 2.11, while [2] comes from the fact that, for $n \in \mathbb{N}$, one has

$$
\int_{-\pi}^{\pi} e^{j\omega n} \mathrm{d}\omega = \begin{cases} 2\pi & \text{se } n = 0 \\ 0 & \text{se } n \neq 0 \end{cases}.
$$

Therefore, we can conclude that from the spectrum one can compute the covariance function as

$$
R_{\boldsymbol{x}}(\tau) = \frac{1}{2\pi j} \oint \phi_{\boldsymbol{x}}(z) z^{\tau-1} \mathrm{d}z, \quad \tau \in \mathbb{Z}.
$$

Equivalently, the covariance function can be derived also from the spectral density (see the second equality in (2.7)), as

$$
R_{\boldsymbol{x}}(\tau) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \phi_{\boldsymbol{x}}(e^{j\omega}) e^{j\omega\tau} \mathrm{d}\omega, \quad \tau \in \mathbb{Z}.
$$

If $\tau = 0$, the previous relationship becomes

$$
R_{\boldsymbol{x}}(0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \phi_{\boldsymbol{x}}(e^{j\omega}) \mathrm{d}\omega. \tag{2.8}
$$

Equation (2.8) provides a motivation for the name "spectral density" for $\phi_{\boldsymbol{x}}(e^{j\omega})$. Indeed, by recalling that $R_{\boldsymbol{x}}(0)$ is the variance of $\boldsymbol{x}(t)$ (which amounts to the "power" of the signal $x(t)$), we can conclude that $\phi_{\boldsymbol{x}}(e^{j\omega})$ indicates the distribution of such power at the different frequencies ω.

Theorem 2.2. *The spectrum and spectral density enjoy the following properties:*

1. $\phi_{\boldsymbol{x}}(z) = \phi_{\boldsymbol{x}}^T(z^{-1})$;
2. $\phi_{\boldsymbol{x}}(e^{j\omega}) = \phi_{\boldsymbol{x}}^T(e^{-j\omega})$;
3. $\phi_{\boldsymbol{x}}(e^{j\omega}) \geq 0$;
4. $\phi_{\boldsymbol{xy}}(z) = \phi_{\boldsymbol{yx}}^T(z^{-1})$.

Proof

1. From the definition of spectrum, one has

$$\begin{aligned}\phi_{\boldsymbol{x}}(z) = \sum_{k=-\infty}^{\infty} R_{\boldsymbol{x}}(k) z^{-k} \overset{[1]}{=} \sum_{k=-\infty}^{\infty} R_{\boldsymbol{x}}^T(-k) z^{-k} = \\ \overset{[2]}{=} \left(\sum_{\tau=-\infty}^{\infty} R_{\boldsymbol{x}}(\tau) \left(z^{-1}\right)^{-\tau} \right)^T = \phi_{\boldsymbol{x}}^T(z^{-1}),\end{aligned}$$

   in which [1] stems form property 1 in Theorem 2.1 (see page 29), while [2] is obtained by setting $\tau = -k$.

2. From the previous property, by setting $z = e^{j\omega}$.

3. We prove this for a scalar process $\boldsymbol{x}(t)$, whose covariance function is exponentially bounded, i.e. there exist real constants $C > 0$ and $\alpha \in (0,\ 1)$ such that $|r_{\boldsymbol{x}}(\tau)| \leq C\alpha^{|\tau|}$. From property 3 in Theorem 2.1 (see page 29), the matrix

$$P(N) = \begin{bmatrix} r_{\boldsymbol{x}}(0) & r_{\boldsymbol{x}}(1) & \dots & r_{\boldsymbol{x}}(N-1) \\ r_{\boldsymbol{x}}(-1) & r_{\boldsymbol{x}}(0) & \dots & r_{\boldsymbol{x}}(N-2) \\ \vdots & \vdots & \ddots & \vdots \\ r_{\boldsymbol{x}}(-N+1) & r_{\boldsymbol{x}}(-N+2) & \dots & r_{\boldsymbol{x}}(0) \end{bmatrix}$$

   is positive semidefinite. This means that for every complex vector $v \in \mathbb{C}^N$, one has $v^* P(N) v \geq 0$, in which v^* denotes the conjugate transpose of v, $v^* = [\bar{v}_1, \bar{v}_2, \dots, \bar{v}_N]$. Therefore, by choosing $v = [1, z^{-1}, \dots, z^{-N+1}]^T$, with $z = e^{j\omega}$, one gets

$$\begin{aligned} 0 &\leq \frac{1}{N} \left(1, \overline{z^{-1}}, \dots, \overline{z^{-N+1}}\right) P(N) \begin{pmatrix} 1 \\ z^{-1} \\ \vdots \\ z^{-N+1} \end{pmatrix} \\ &\overset{[1]}{=} \frac{1}{N} \left(1, z, \dots, z^{N-1}\right) \begin{pmatrix} r_{\boldsymbol{x}}(0) + r_{\boldsymbol{x}}(1) z^{-1} + \dots + r_{\boldsymbol{x}}(N-1) z^{-N+1} \\ r_{\boldsymbol{x}}(-1) + r_{\boldsymbol{x}}(0) z^{-1} + \dots + r_{\boldsymbol{x}}(N-2) z^{-N+1} \\ \vdots \\ r_{\boldsymbol{x}}(-N+1) + r_{\boldsymbol{x}}(-N+2) z^{-1} + \dots + r_{\boldsymbol{x}}(0) z^{-N+1} \end{pmatrix} \\ &= \frac{1}{N} \left[N r_{\boldsymbol{x}}(0) + \sum_{k=1}^{N} (N-k)(r_{\boldsymbol{x}}(k) z^{-k} + r_{\boldsymbol{x}}(-k) z^k) \right] \\ &= \frac{1}{N} \left[N r_{\boldsymbol{x}}(0) + \sum_{k=1}^{N} N(r_{\boldsymbol{x}}(k) z^{-k} + r_{\boldsymbol{x}}(-k) z^k) - \sum_{k=1}^{N} k(r_{\boldsymbol{x}}(k) z^{-k} + r_{\boldsymbol{x}}(-k) z^k) \right] \\ &\overset{[2]}{=} \sum_{k=-N}^{N} r_{\boldsymbol{x}}(k) z^{-k} - \frac{1}{N} \sum_{k=-N}^{N} |k|\, r_{\boldsymbol{x}}(k) z^{-k}, \end{aligned} \tag{2.9}$$

in which [1] comes from the property of complex numbers $\overline{e^{-j\omega k}} = e^{j\omega k}$, while [2] stems from the symmetry of the covariance function, see equation (2.2). Let us consider the modulus of the last sum in (2.9). From the triangle inequality

$$\left| \frac{1}{N} \sum_{k=-N}^{N} |k| \, r(k) z^{-k} \right| \leq \frac{1}{N} \sum_{k=-N}^{N} \left| |k| \, r(k) z^{-k} \right| = \frac{1}{N} \sum_{k=-N}^{N} |k| \, |r(k)| \left| z^{-k} \right|$$

By setting $z = e^{j\omega}$, and since $|e^{j\omega}| = 1$, we get

$$\begin{aligned} \left| \frac{1}{N} \sum_{k=-N}^{N} |k| \, r(k) e^{-j\omega k} \right| &\leq \frac{1}{N} \sum_{k=-N}^{N} |k| \, |r(k)| \\ &\leq \frac{1}{N} \sum_{k=-N}^{N} |k| \, C\alpha^{|k|} = \frac{2C}{N} \sum_{k=1}^{N} k\alpha^k \end{aligned} \tag{2.10}$$

where we have exploited the assumption $|r_{\boldsymbol{x}}(\tau)| \leq C\alpha^{|\tau|}$. We show that the quantity on the right in (2.10) tends to zero when N goes to infinity. To do this, observe that

$$\begin{aligned} (1-\alpha)^2 \sum_{k=1}^{N} k\alpha^k &= \sum_{k=1}^{N} k\alpha^k - 2\alpha \sum_{k=1}^{N} k\alpha^k + \alpha^2 \sum_{k=1}^{N} k\alpha^k \\ &= \alpha + 2\alpha^2 - 2\alpha^2 + \sum_{k=3}^{N} [k - 2(k-1) + k - 2]\alpha^k \\ &\qquad - 2N\alpha^{N+1} + (N-1)\alpha^{N+1} + N\alpha^{N+2} \\ &= \alpha - (N+1)\alpha^{N+1} + N\alpha^{N+2}. \end{aligned}$$

and dividing both members by $(1-\alpha)^2$, we get

$$\sum_{k=1}^{N} k\alpha^k = \frac{\alpha - (N+1)\alpha^{N+1} + N\alpha^{N+2}}{(1-\alpha)^2}.$$

Recalling that, by assumption, $0 < \alpha < 1$, as N goes to infinity one gets

$$\sum_{k=1}^{N} k\alpha^k \to \frac{\alpha}{(1-\alpha)^2}$$

and hence

$$\frac{2C}{N} \sum_{k=1}^{N} k\alpha^k \to 0.$$

Therefore, also the term on the left in (2.10) tends to zero. By evaluating (2.9) for $z = e^{j\omega}$ and letting N tend to infinity, we can conclude that

$$\phi_{\boldsymbol{x}}(e^{j\omega}) = \sum_{k=-\infty}^{\infty} r_{\boldsymbol{x}}(k) \, e^{-j\omega k} \geq 0.$$

4. From the definition of cross-spectrum, one has

$$\phi_{\boldsymbol{xy}}(z) = \sum_{k=-\infty}^{\infty} R_{\boldsymbol{xy}}(k)\, z^{-k} \overset{[1]}{=} \sum_{k=-\infty}^{\infty} R^T_{\boldsymbol{yx}}(-k) z^{-k} =$$
$$\overset{[2]}{=} \left(\sum_{\tau=-\infty}^{\infty} R_{\boldsymbol{yx}}(\tau) \left(z^{-1}\right)^{-\tau} \right)^T = \phi^T_{\boldsymbol{yx}}(z^{-1}),$$

in which [1] stems form property 4 in Theorem 2.1 (see page 29), while [2] is obtained by setting $\tau = -k$. □

If $\boldsymbol{x}(t)$ is a scalar stochastic process, its spectrum satisfies

$$\phi_{\boldsymbol{x}}(z) = \phi_{\boldsymbol{x}}(z^{-1}),$$

while its spectral density is an even nonnegative function of ω

$$\phi_{\boldsymbol{x}}(e^{j\omega}) = \phi_{\boldsymbol{x}}(e^{-j\omega}),$$
$$\phi_{\boldsymbol{x}}(e^{j\omega}) \geq 0, \quad \forall \omega.$$

By recalling that $\phi_{\boldsymbol{x}}(e^{j\omega})$ is also periodic of period 2π, we can analyze it only in the interval $\omega \in [0,\ \pi]$.

Example 2.2. Consider a white process $\boldsymbol{x}(t)$, with covariance function

$$R_{\boldsymbol{x}}(\tau) = \begin{cases} \sigma^2_{\boldsymbol{x}} & \text{se } \tau = 0 \\ 0 & \text{se } \tau \neq 0 \end{cases}$$

From the definitions of spectrum and spectral density, we get

$$\phi_{\boldsymbol{x}}(z) = \sigma^2_{\boldsymbol{x}}, \qquad \forall z \in \mathbb{C},$$

and

$$\phi_{\boldsymbol{x}}(e^{j\omega}) = \sigma^2_{\boldsymbol{x}}, \qquad \forall \omega \in [-\pi, \pi].$$

Hence, the spectral density of a white process is constant. This means that the "power" of the signal is equally distributed at all frequencies. This motivates the attribute *white*, which recalls the property of the white light, containing all the frequencies of the visible spectrum. △

Example 2.3. Consider the exponentially correlated stochastic process $\boldsymbol{x}(t)$, whose covariance function is

$$R_{\boldsymbol{x}}(\tau) = \sigma^2_{\boldsymbol{x}} a^{|\tau|}, \quad \tau \in \mathbb{Z}, \tag{2.11}$$

with $|a| < 1$. Let us compute the spectrum and spectral density. From the definition,

$$\begin{aligned}
\phi_{\boldsymbol{x}}(z) &= \sum_{\tau=-\infty}^{\infty} R_{\boldsymbol{x}}(\tau) z^{-\tau} = \sum_{\tau=-\infty}^{\infty} \sigma_{\boldsymbol{x}}^2 a^{|\tau|} z^{-\tau} \\
&= \sigma_{\boldsymbol{x}}^2 \left(\sum_{\tau=-\infty}^{-1} a^{-\tau} z^{-\tau} + \sum_{\tau=0}^{\infty} a^{\tau} z^{-\tau} \right) \\
&= \sigma_{\boldsymbol{x}}^2 \left(\sum_{k=1}^{\infty} a^k z^k + \sum_{\tau=0}^{\infty} a^{\tau} z^{-\tau} \right) \\
&= \sigma_{\boldsymbol{x}}^2 \left(\sum_{k=0}^{\infty} (az)^k + \sum_{\tau=0}^{\infty} \left(az^{-1}\right)^{\tau} - 1 \right).
\end{aligned}$$

Since the series $\sum_{k=0}^{\infty} p^k$ converges to $\frac{1}{1-p}$, if $|p| < 1$, in order for both series to converge we need

$$\begin{aligned}
|az| &< 1, \\
\left|az^{-1}\right| &< 1.
\end{aligned}$$

The previous inequalities define the region of the complex plane in which the spectrum is defined

$$|a| < |z| < \frac{1}{|a|}.$$

Figure 2.5 shows such a region.

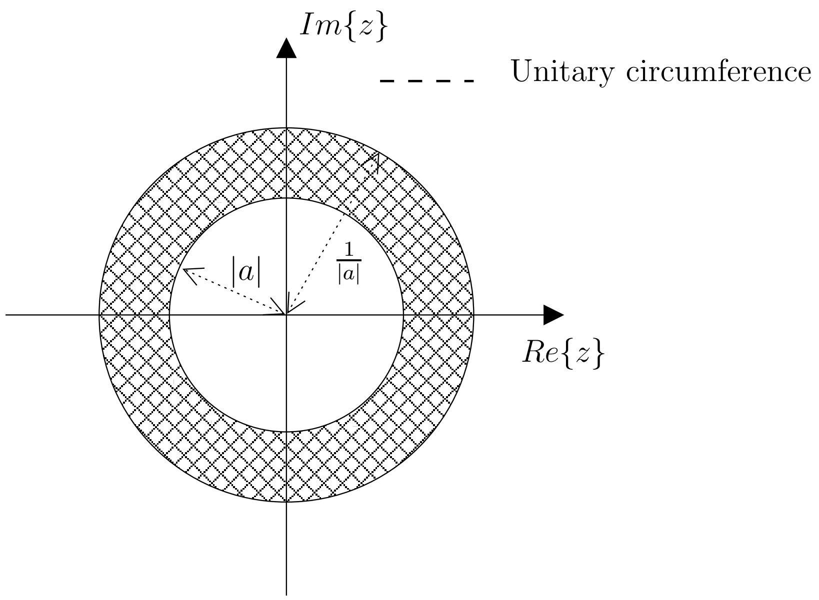


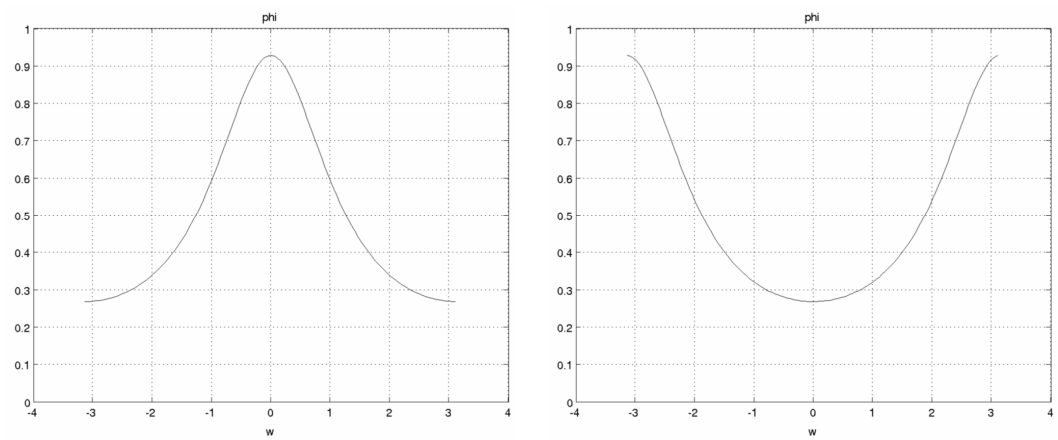
Figure 2.5: Convergence region of the spectrum $\phi_{\boldsymbol{x}}(z)$.

In its convergence region, the spectrum takes on the value

$$\begin{aligned}\phi_{\boldsymbol{x}}(z) &= \sigma_{\boldsymbol{x}}^2\left(\frac{1}{1-az}+\frac{1}{1-az^{-1}}-1\right)\\ &= \frac{\sigma_{\boldsymbol{x}}^2(1-az^{-1}+1-az-1-a^2+az+az^{-1})}{(1-az^{-1})\,(1-az)}\\ &= \frac{\sigma_{\boldsymbol{x}}^2(1-a^2)}{(1-az^{-1})\,(1-az)}.\end{aligned}$$

Since the convergence region always contains the unit circumference, it is always possible to define the spectral density

$$\begin{aligned}\phi_{\boldsymbol{x}}(e^{j\omega}) &= \frac{\sigma_{\boldsymbol{x}}^2(1-a^2)}{(1-ae^{-j\omega})\,(1-ae^{j\omega})} = \frac{\sigma_{\boldsymbol{x}}^2(1-a^2)}{1+a^2-a(e^{j\omega}+e^{-j\omega})}\\ &= \frac{\sigma_{\boldsymbol{x}}^2(1-a^2)}{1+a^2-2a\cos\omega}\ .\end{aligned} \tag{2.12}$$



(a) (b)

Figure 2.6: Spectral density of an exponentially correlated process for (a) $a = 0.4$; (b) $a = -0.4$.

Figure 2.6 shows the plots of the spectral density (2.12) for $a = 0.4$ and $a = -0.4$. Observe that when the correlation coefficient is positive, the signal power is concentrated at low frequencies, while the opposite occurs for negative values of a. In fact, for an exponentially correlated process, one has

$$R_{\boldsymbol{x}}(1) = \mathbf{E}\left[\boldsymbol{x}(t+1)\boldsymbol{x}(t)\right] = \sigma_{\boldsymbol{x}}^2\, a.$$

This means that for positive values of a, the stochastic process will present on average less oscillations, with respect to the case of negative a (see also Section 2.2.4), since we expect that $x(t+1)$ and $x(t)$ have often the same sign. △

2.4 Linear stochastic systems

In this section, we aim to analyze what happens when a stochastic process is used as the input of a dynamic system. In particular, we will concentrate on the input-output representation of *asymptotically stable* LTI systems, fed by stationary stochastic processes. We will refer to this case as a *linear stochastic system.*

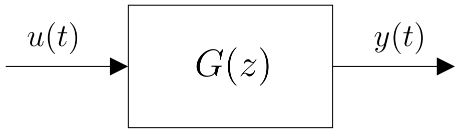


Figure 2.7: A linear stochastic system with input $u(t)$, output $y(t)$ and transfer function matrix $G(z)$.

With reference to Figure 2.7, let

$$G(z) = \sum_{k=0}^{\infty} g_k z^{-k}$$

be the transfer function matrix of an LTI system.[2] Recall that $G(z)$ is the Z-transform of the impulse response $\{g_k\}_{k=0}^{\infty}$ of the system. Therefore, the output $y(t)$ can be written as

$$y(t) = \sum_{k=0}^{\infty} g_k u(t-k), \tag{2.13}$$

where $u(t)$ is the input signal.

If the input $\boldsymbol{u}(t)$ is a stochastic process, then also the output $\boldsymbol{y}(t)$ will be a stochastic process. Equation (2.13) relates the specific realization of the output to the corresponding realization of the input. We aim at finding relationships between the statistics of the stochastic processes $\boldsymbol{y}(t)$ and $\boldsymbol{u}(t)$. The following result provides an answer for the mean, covariance function and spectrum of the two processes.

Theorem 2.3. *Let $G(z)$ be the transfer function matrix of an asymptotically stable LTI system. Let the input $\boldsymbol{u}(t)$ be a stationary stochastic process with mean $m_{\boldsymbol{u}}$ and spectrum $\phi_{\boldsymbol{u}}(z)$. Then, the output $\boldsymbol{y}(t)$ is an asymptotically stationary stochastic process, whose mean and spectrum are given by*

1. $m_{\boldsymbol{y}} = G(1)\, m_{\boldsymbol{u}}$;

2. $\phi_{\boldsymbol{y}}(z) = G(z)\phi_{\boldsymbol{u}}(z)G^T(z^{-1})$.

Moreover, $\boldsymbol{y}(t)$ and $\boldsymbol{u}(t)$ are jointly stationary with cross-covariance and cross-spectrum satisfying the following relationships

[2]In general, if $u \in \mathbb{R}^m$ and $y \in \mathbb{R}^p$, then $g_k \in \mathbb{R}^{p\times m}$. For single-input single-output systems, $m = p = 1$, one has $g_k \in \mathbb{R}$ and $G(z)$ is called *transfer function* of the system.

3. $R_{\boldsymbol{yu}}(\tau)=\sum_{k=0}^{\infty} g_k R_{\boldsymbol{u}}(\tau-k)\,,\qquad R_{\boldsymbol{y}}(\tau)=\sum_{k=0}^{\infty} R_{\boldsymbol{yu}}(\tau+k)\,g_k^T$*;*

4. $\phi_{\boldsymbol{yu}}(z)=G(z)\phi_{\boldsymbol{u}}(z),\qquad \phi_{\boldsymbol{y}}(z)=\phi_{\boldsymbol{yu}}(z)G^T(z^{-1})$.

Proof

1. By using (2.13), we get

$$m_{\boldsymbol{y}}=\mathbf{E}\left[\boldsymbol{y}(t)\right]=\sum_{k=0}^{\infty} g_k\mathbf{E}\left[\boldsymbol{u}(t-k)\right]\overset{[1]}{=}\left(\sum_{k=0}^{\infty} g_k\right)m_{\boldsymbol{u}}\overset{[2]}{=}G(1)\,m_{\boldsymbol{u}}$$

where [1] comes from the stationarity of $\boldsymbol{u}(t)$, while [2] follows from the definition of transfer function (2.4). Notice that the infinite sum of g_k converges to the finite value $G(1)$ thanks to the asymptotic stability of the system.

2. For ease of exposition we assume, without loss of generality, that $m_{\boldsymbol{u}}=m_{\boldsymbol{y}}=0$ (otherwise, what follows still holds by replacing $\boldsymbol{y}(t)$ and $\boldsymbol{u}(t)$ with $\boldsymbol{y}(t)-m_{\boldsymbol{y}}$ and $\boldsymbol{u}(t)-m_{\boldsymbol{u}}$, respectively). First, let us verify that the s.p. $\boldsymbol{y}(t)$ is asymptotically stationary. In item 1, we established that its mean is constant. Let us compute its covariance function by using (2.13):

$$\begin{aligned}R_{\boldsymbol{y}}(t+\tau,t)&=\mathbf{E}\left[\left(\sum_{k=0}^{\infty} g_k\boldsymbol{u}(t+\tau-k)\right)\left(\sum_{l=0}^{\infty} g_l\boldsymbol{u}(t-l)\right)^T\right]\\&=\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} g_kR_{\boldsymbol{u}}(\tau-k+l)\,g_l^T.\end{aligned}$$

Hence, $R_{\boldsymbol{y}}(t+\tau,t)=R_{\boldsymbol{y}}(\tau)$, and therefore $\boldsymbol{y}(t)$ is asymptotically stationary, in the sense of Observation 2.1. From the expression of $R_{\boldsymbol{y}}(\tau)$, one can derive the spectrum of $y(t)$

$$\phi_{\boldsymbol{y}}(z)=\sum_{\tau=-\infty}^{\infty} R_{\boldsymbol{y}}(\tau)z^{-\tau}=\sum_{\tau=-\infty}^{\infty}\left(\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} g_kR_{\boldsymbol{u}}(\tau-k+l)\,g_l^T\right)z^{-\tau}.$$

Observing that we can rewrite $z^{-\tau}=z^{-(\tau-k+l)}z^{-k+l}$, the previous equation becomes

$$\begin{aligned}\phi_{\boldsymbol{y}}(z)&=\sum_{\tau=-\infty}^{\infty}\left(\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} g_kR_{\boldsymbol{u}}(\tau-k+l)\,g_l^T\right)z^{-(\tau-k+l)}z^{-k+l}\\&=[\text{by setting } m=\tau-k+l]\\&=\sum_{m=-\infty}^{\infty}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} g_kR_{\boldsymbol{u}}(m)\,g_l^Tz^{-m}z^{-k+l}\\&=\sum_{k=0}^{\infty} g_kz^{-k}\sum_{m=-\infty}^{\infty} R_{\boldsymbol{u}}(m)\,z^{-m}\sum_{l=0}^{\infty} g_l^T(z^{-1})^{-l}\\&=G(z)\phi_{\boldsymbol{u}}(z)G^T(z^{-1}),\end{aligned}$$

in which the last equality stems from (2.4) and Definition 2.10.

3. Let us verify that $\boldsymbol{y}(t)$ and $\boldsymbol{u}(t)$ are jointly stationary:

$$\begin{aligned} R_{\boldsymbol{yu}}(t+\tau, t) &= \mathbf{E}\left[\left(\sum_{k=0}^{\infty} g_k \boldsymbol{u}(t+\tau-k)\right)\boldsymbol{u}^T(t)\right] \\ &= \sum_{k=0}^{\infty} g_k \mathbf{E}\left[\boldsymbol{u}(t+\tau-k)\boldsymbol{u}^T(t)\right] \\ &= \sum_{k=0}^{\infty} g_k R_{\boldsymbol{u}}(\tau-k) = R_{\boldsymbol{yu}}(\tau). \end{aligned}$$

4. From the previous expression of $R_{\boldsymbol{yu}}(\tau)$, one gets immediately

$$\begin{aligned} \phi_{\boldsymbol{yu}}(z) &= \sum_{\tau=-\infty}^{\infty}\left(\sum_{k=0}^{\infty} g_k R_{\boldsymbol{u}}(\tau-k)\right) z^{-\tau} \\ &= \sum_{k=0}^{\infty} g_k z^{-k}\left(\sum_{\tau=-\infty}^{\infty} R_{\boldsymbol{u}}(\tau-k)\, z^{-(\tau-k)}\right) \\ &= \sum_{k=0}^{\infty} g_k z^{-k}\phi_{\boldsymbol{u}}(z) = G(z)\phi_{\boldsymbol{u}}(z). \end{aligned}$$

The other relationships in items 3 and 4 can be derived in the same way. □

Some noteworthy consequences of Theorem 2.3 are:

$$\begin{aligned} \phi_{\boldsymbol{y}}(e^{j\omega}) &= G(e^{j\omega})\phi_{\boldsymbol{u}}(e^{j\omega})G^T(e^{-j\omega}) \\ \phi_{\boldsymbol{uy}}(z) &= \phi_{\boldsymbol{u}}^T(z^{-1})G^T(z^{-1}) \end{aligned} \tag{2.14}$$

and in the *scalar case*:

$$\begin{aligned} \phi_{\boldsymbol{y}}(z) &= G(z)G(z^{-1})\phi_{\boldsymbol{u}}(z) \\ \phi_{\boldsymbol{y}}(e^{j\omega}) &= \left|G(e^{j\omega})\right|^2 \phi_{\boldsymbol{u}}(e^{j\omega}) \\ \phi_{\boldsymbol{uy}}(z) &= G(z^{-1})\phi_{\boldsymbol{u}}(z) \end{aligned}$$

Let us now consider the special case in which the input $\boldsymbol{u}(t)$ is a white process $\boldsymbol{e}(t)$, with zero mean and variance $R_{\boldsymbol{e}}$. Then,

$$\begin{aligned} m_{\boldsymbol{e}} &= 0 \\ R_{\boldsymbol{e}}(\tau) &= \begin{cases} R_{\boldsymbol{e}} & \text{if } \tau = 0 \\ 0 & \text{if } \tau \neq 0 \end{cases} \\ \phi_{\boldsymbol{e}}(e^{j\omega}) &= R_{\boldsymbol{e}} \end{aligned}$$

From equation (2.14) one gets

$$\phi_{\boldsymbol{y}}(e^{j\omega}) = G(e^{j\omega})R_{\boldsymbol{e}}G^T(e^{-j\omega}).$$

By applying (2.8), we obtain the following version of Parseval's theorem

$$R_{\boldsymbol{y}}(0) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \phi_{\boldsymbol{y}}(e^{j\omega})\mathrm{d}\omega = \frac{1}{2\pi}\int_{-\pi}^{\pi} G(e^{j\omega})R_{\boldsymbol{e}}G^T(e^{-j\omega})\mathrm{d}\omega,$$

that in the scalar case becomes

$$R_{\boldsymbol{y}}(0) = \sigma_{\boldsymbol{e}}^2 \frac{1}{2\pi}\int_{-\pi}^{\pi} \left|G(e^{j\omega})\right|^2 \mathrm{d}\omega. \tag{2.15}$$

Equation (2.15) shows that the variance of the stochastic process $\boldsymbol{y}(t)$ can be obtained by integrating the squared modulus of the transfer function along the unit circumference.

2.5 Linear models of stochastic processes

In this paragraph, we will analyze scalar stochastic processes obtained by filtering a white process through an asymptotically stable system. To this aim, let $\boldsymbol{e}(t) \sim WP(0, \sigma_{\boldsymbol{e}}^2)$ be the input of an LTI system with output $\boldsymbol{y}(t)$, see Figure 2.8. We define different classes of linear stochastic processes, depending on the structure of the transfer function $G(z)$.

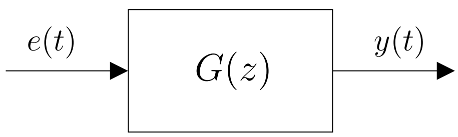


Figure 2.8: A s.p. $\boldsymbol{y}(t)$ obtained by filtering a white process $\boldsymbol{e}(t)$.

2.5.1 MA process

Consider the scalar s.p. $\boldsymbol{y}(t)$, defined by the equation

$$y(t) = c_0 e(t) + c_1 e(t-1) + \cdots + c_m e(t-m). \tag{2.16}$$

where $\boldsymbol{e}(t) \sim WP(0, \sigma_{\boldsymbol{e}}^2)$ and $c_i \in \mathbb{R}$, $i = 0, 1, \ldots, m$. Such a process is said *Moving Average* of order m, (briefly denoted as MA(m)), because $y(t)$ is obtained as the weighted average of the last $m+1$ samples of the white process $e(t)$.

Since $\boldsymbol{e}(t)$ has zero mean, so it will be for $\boldsymbol{y}(t)$:

$$m_{\boldsymbol{y}} = \mathbf{E}\left[\boldsymbol{y}(t)\right] = c_0\mathbf{E}\left[\boldsymbol{e}(t)\right] + c_1\mathbf{E}\left[\boldsymbol{e}(t-1)\right] + \cdots + c_m\mathbf{E}\left[\boldsymbol{e}(t-m)\right] = 0.$$

Let us compute the variance of $\boldsymbol{y}(t)$:

$$\begin{aligned}\sigma_{\boldsymbol{y}}^2 &= \mathbf{E}\left[\boldsymbol{y}(t)^2\right] = \mathbf{E}\left[(c_0\boldsymbol{e}(t) + c_1\boldsymbol{e}(t-1) + \cdots + c_m\boldsymbol{e}(t-m))^2\right] \\ &= (c_0^2 + c_1^2 + \cdots + c_m^2)\sigma_{\boldsymbol{e}}^2\end{aligned}$$

in which the last equality comes from

$$\mathbf{E}\left[\boldsymbol{e}(t_1)\boldsymbol{e}(t_2)\right] = \begin{cases} \sigma_{\boldsymbol{e}}^2 & \text{if } t_1 = t_2 \\ 0 & \text{if } t_1 \neq t_2 \end{cases}. \tag{2.17}$$

Then, let us compute the covariance function for a generic lag $\tau > 0$:

$$\begin{aligned} R_{\boldsymbol{y}}(\tau) &= \mathbf{E}\left[\boldsymbol{y}(t+\tau)\boldsymbol{y}(t)\right] \\ &= \mathbf{E}[(c_0\boldsymbol{e}(t+\tau) + c_1\boldsymbol{e}(t+\tau-1) + \cdots + c_m\boldsymbol{e}(t+\tau-m)) \times \\ &\qquad \times (c_0\boldsymbol{e}(t) + c_1\boldsymbol{e}(t-1) + \cdots + c_m\boldsymbol{e}(t-m))]. \end{aligned}$$

Being $\boldsymbol{e}(t)$ white, the only products returning a non zero contribution are those involving the signal e at the same time instant: $e(\bar{t})e(\bar{t})$. Hence, with reference to Figure 2.9, one has

τ	$R_{\boldsymbol{y}}(\tau)$
1	$(c_1c_0 + c_2c_1 + \cdots + c_mc_{m-1})\sigma_{\boldsymbol{e}}^2$
2	$(c_2c_0 + c_3c_1 + \cdots + c_mc_{m-2})\sigma_{\boldsymbol{e}}^2$
$\vdots$	$\vdots$
$m-1$	$(c_{m-1}c_0 + c_mc_1)\sigma_{\boldsymbol{e}}^2$
m	$c_mc_0\sigma_{\boldsymbol{e}}^2$

Finally, if $\tau > m$, the covariance function is identically zero.

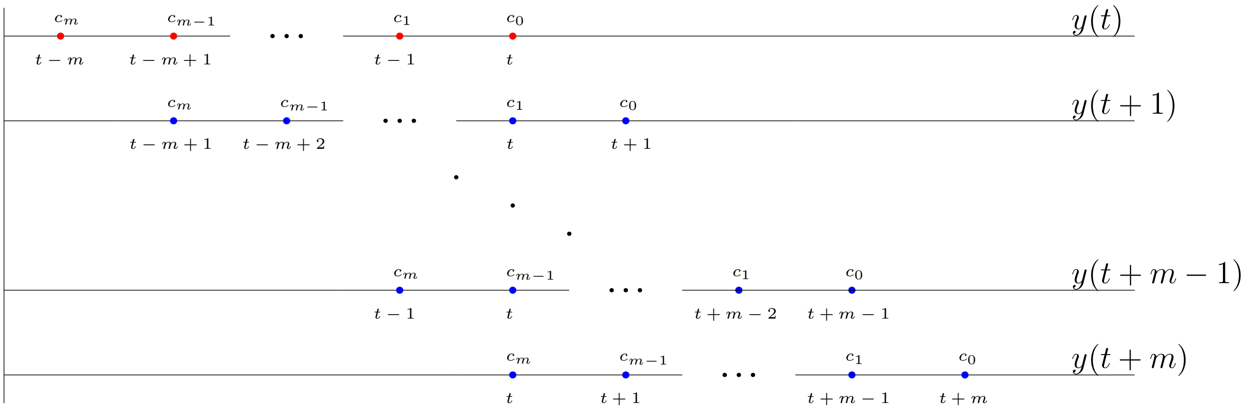


Figure 2.9: Computation of the covariance function of an MA(m) process.

By interpreting z^{-1} as the backward shift operator[3], such that

$$z^{-1}x(t) = x(t-1),$$

it is straightforward to derive the transfer function of the system generating $\boldsymbol{y}(t)$ from the input $\boldsymbol{e}(t)$. Indeed, (2.16) can be rewritten as

$$\boldsymbol{y}(t) = C(z)\boldsymbol{e}(t), \qquad \text{with } C(z) = c_0 + c_1z^{-1} + \cdots + c_mz^{-m}.$$

[3]With an abuse of notation, we use z to denote both the forward shift operator, $z\,x(t) = x(t+1)$, and the complex variable in the domain of the Z-transform, like, e.g., in the Definitions 2.10 and 2.12. The meaning however, should be clear from the context.

Then, by multiplying and dividing $C(z)$ by z^m, we can write the transfer function from $\boldsymbol{e}(t)$ to $\boldsymbol{y}(t)$ in the usual rational form

$$G(z) = C(z) = \frac{c_0 z^m + c_1 z^{m-1} + \cdots + c_m}{z^m}.$$

Notice that all the poles of $G(z)$ are in $z = 0$. Therefore, the system is asymptotically stable (it is also finite time stable, in the sense that its impulse response $\{c_0, c_1, \ldots, c_m, 0, 0, \ldots\}$ goes to zero in $m+1$ time instants). Such kind of systems is called *Finite Impulse Response (FIR)* systems. Therefore, thanks to Theorem 2.3, $\boldsymbol{y}(t)$ is asymptotically stationary.

Summing up, an MA(m) s.p. $\boldsymbol{y}(t)$ has the following properties:

- it is asymptotically stationary;
- it has zero mean, $m_{\boldsymbol{y}} = 0$ (if $\boldsymbol{e}(t)$ has zero mean);
- at steady state, its covariance function is equal to

$$R_{\boldsymbol{y}}(\tau) = \begin{cases} \left(\displaystyle\sum_{k=|\tau|}^{m} c_k c_{k-|\tau|} \right) \sigma_{\boldsymbol{e}}^2 & \text{se } |\tau| \leq m \\ 0 & \text{se } |\tau| > m \end{cases}$$

i.e., $y(t_1)$ and $y(t_2)$ are uncorrelated, if $|t_1 - t_2| > m$.

Example 2.4. Consider the MA(1) process

$$\boldsymbol{y}(t) = \boldsymbol{e}(t) + \frac{1}{2}\boldsymbol{e}(t-1)$$

with $\boldsymbol{e}(t) \sim WP(0, 1)$. Clearly, one has $m_{\boldsymbol{y}} = 0$. Let us compute the variance of $\boldsymbol{y}(t)$

$$\begin{aligned} R_{\boldsymbol{y}}(0) &= \mathbf{E}\left[\boldsymbol{y}(t)^2\right] = \mathbf{E}\left[\left(\boldsymbol{e}(t) + \frac{1}{2}\boldsymbol{e}(t-1)\right)^2\right] \\ &\overset{[1]}{=} \mathbf{E}\left[\boldsymbol{e}(t)^2\right] + \frac{1}{4}\mathbf{E}\left[\boldsymbol{e}(t-1)^2\right] = \left(1 + \frac{1}{4}\right)\sigma_{\boldsymbol{e}}^2 \\ &= \frac{5}{4}, \end{aligned}$$

where [1] stems from the property (2.17) of a white process. The covariance function of $\boldsymbol{y}(t)$ is equal to

$$\begin{aligned} R_{\boldsymbol{y}}(1) &= \mathbf{E}\left[\boldsymbol{y}(t+1)\boldsymbol{y}(t)\right] \\ &= \mathbf{E}\left[\left(\boldsymbol{e}(t+1) + \frac{1}{2}\boldsymbol{e}(t)\right)\left(\boldsymbol{e}(t) + \frac{1}{2}\boldsymbol{e}(t-1)\right)\right] \\ &= \frac{1}{2}\sigma_{\boldsymbol{e}}^2 = \frac{1}{2}, \end{aligned}$$

$$\begin{aligned} R_{\boldsymbol{y}}(2) &= \mathbf{E}\left[\boldsymbol{y}(t+2)\boldsymbol{y}(t)\right] \\ &= \mathbf{E}\left[\left(\boldsymbol{e}(t+2) + \frac{1}{2}\boldsymbol{e}(t+1)\right)\left(\boldsymbol{e}(t) + \frac{1}{2}\boldsymbol{e}(t-1)\right)\right] \\ &= 0 \end{aligned}$$

and the same occurs for $R_{\boldsymbol{y}}(\tau)$, $\tau > 2$. Summing up:

$$R_{\boldsymbol{y}}(\tau) = \begin{cases} \frac{5}{4} & \text{se } |\tau| = 0 \\ \frac{1}{2} & \text{se } |\tau| = 1 \\ 0 & \text{se } |\tau| > 1 \end{cases}.$$

For a MA process, it is straightforward to compute the spectrum and spectral density, from Definitions 2.10-2.11:

$$\phi_{\boldsymbol{y}}(z) = \sum_{\tau=-\infty}^{\infty} R_{\boldsymbol{y}}(\tau)z^{-\tau} = \frac{5}{4} + \frac{1}{2}z^{-1} + \frac{1}{2}z,$$

$$\phi_{\boldsymbol{y}}(e^{j\omega}) = \frac{5}{4} + \frac{1}{2}e^{-j\omega} + \frac{1}{2}e^{j\omega} = \frac{5}{4} + \frac{e^{j\omega} + e^{-j\omega}}{2} = \frac{5}{4} + \cos\omega.$$

Notice that we can arrive at the same results by applying Theorem 2.3. Being the transfer function from $\boldsymbol{e}(t)$ to $\boldsymbol{y}(t)$

$$G(z) = \frac{z + \frac{1}{2}}{z}$$

and the spectrum and spectral density of $\boldsymbol{e}(t)$

$$\phi_{\boldsymbol{e}}(z) = \phi_{\boldsymbol{e}}(e^{j\omega}) = 1,$$

the spectrum and spectral density of $\boldsymbol{y}(t)$ can be derived as

$$\phi_{\boldsymbol{y}}(z) = G(z)\phi_{\boldsymbol{e}}(z)G(z^{-1}) = 1 + \frac{1}{4} + \frac{1}{2}z^{-1} + \frac{1}{2}z = \frac{5}{4} + \frac{1}{2}z^{-1} + \frac{1}{2}z,$$

$$\begin{aligned} \phi_{\boldsymbol{y}}(e^{j\omega}) &= \left|G(e^{j\omega})\right|^2 \phi_{\boldsymbol{e}}(e^{j\omega}) = \left|G(e^{j\omega})\right|^2 = \left|\frac{e^{j\omega} + \frac{1}{2}}{e^{j\omega}}\right|^2 \\ &= \frac{\left|e^{j\omega} + \frac{1}{2}\right|^2}{\left|e^{j\omega}\right|^2} \overset{[1]}{=} \left|\frac{1}{2} + \cos\omega + j\sin\omega\right|^2 \\ &= \left(\frac{1}{2} + \cos\omega\right)^2 + \sin^2\omega \\ &= \frac{1}{4} + \cos^2\omega + \cos\omega + \sin^2\omega \\ &= \frac{5}{4} + \cos\omega, \end{aligned}$$

in which in [1] the relationships $|e^{j\omega}| = 1$ and $e^{j\omega} = \cos\omega + j\sin\omega$ have been exploited. △

2.5.2 AR process

Consider the s.p. $\boldsymbol{y}(t)$, defined by the equation

$$y(t) + a_1 y(t-1) + \cdots + a_n y(t-n) = e(t) \tag{2.18}$$

where $\boldsymbol{e}(t) \sim WP(0, \sigma_{\boldsymbol{e}}^2)$ and $a_i \in \mathbb{R}$, $i = 0, 1, \ldots, n$. Such a s.p. is said *Auto-Regressive* of order n (or briefly, AR(n)).

Similarly to what has been done for the MA process, we can derive the transfer function of the system generating $y(t)$ from the input $e(t)$. Equation (2.18) can be rewritten as

$$A(z)y(t) = e(t), \qquad \text{with } A(z) = 1 + a_1 z^{-1} + \cdots + a_n z^{-n}.$$

By multiplying and dividing $A(z)$ by z^n, the transfer function from $e(t)$ to $y(t)$ is

$$G(z) = \frac{1}{A(z)} = \frac{z^n}{z^n + a_1 z^{n-1} + \cdots + a_n}.$$

If the roots of the polynomial $A(z)$ have modulus strictly smaller than 1, the system is asymptotically stable. Therefore, thanks to Theorem 2.3, $\boldsymbol{y}(t)$ is an asymptotically stationary stochastic process.[4]

Let now $\{g_k\}_{k=0}^{\infty}$ be the impulse response of the LTI system with transfer function $G(z)$. From (2.13) we know that

$$y(t) = \sum_{k=0}^{\infty} g_k e(t-k).$$

This shows that an AR process can be seen as a MA process of infinite order. It follows that

$$m_{\boldsymbol{y}} = \mathbf{E}\left[\boldsymbol{y}(t)\right] = \left(\sum_{k=0}^{\infty} g_k\right) m_{\boldsymbol{e}} = G(1)\, m_{\boldsymbol{e}} = 0 \tag{2.19}$$

$$R_{\boldsymbol{y}}(0) = \mathbf{E}\left[\boldsymbol{y}(t)^2\right] = \left(\sum_{k=0}^{\infty} g_k^2\right) \sigma_{\boldsymbol{e}}^2, \tag{2.20}$$

$$R_{\boldsymbol{y}}(\tau) = \mathbf{E}\left[\boldsymbol{y}(t+\tau)\boldsymbol{y}(t)\right] = \left(\sum_{k=0}^{\infty} g_{k+\tau} g_k\right) \sigma_{\boldsymbol{e}}^2, \tag{2.21}$$

in which we exploited the properties of the white process $\boldsymbol{e}(t)$, and in particular (2.17). It should be remarked that all the infinite sums in (2.19)-(2.21) converge if and only if $G(z)$ is an asymptotically stable transfer function, i.e., the polynomial $A(z)$ has all the roots strictly inside the unit circle.

[4] In the following discussion, we will consider the process $\boldsymbol{y}(t)$ at steady state, i.e., after the transient has died out, see Observation 2.1.

Observation 2.2. Equations (2.20)-(2.21) provide a practical way to compute the covariance function of an AR process. In fact, by picking a sufficiently large integer M such that $g_k \simeq 0$, $\forall k > M$ (notice that such an M exists due to the asymptotic stability of the system), one can approximate the covariance function by computing

$$R_{\boldsymbol{y}}(\tau) \simeq \left(\sum_{k=0}^{M} g_{k+\tau} g_k \right) \sigma_{\boldsymbol{e}}^2.$$

The larger is M, the more precise will be the obtained approximate value of the covariance function.

Example 2.5. Consider the AR(1) s.p.

$$y(t) - \frac{1}{2} y(t-1) = e(t)$$

with $\boldsymbol{e}(t) \sim WP(0,1)$. We already know from Example 2.1 that the covariance function of a generic AR(1) process

$$y(t) = ay(t-1) + e(t)$$

is equal to

$$R_{\boldsymbol{y}}(\tau) = \frac{\sigma_{\boldsymbol{e}}^2}{1-a^2} a^{|\tau|}.$$

Since in this example $a = 0.5$, one has

$$R_{\boldsymbol{y}}(\tau) = \frac{4}{3} 0.5^{|\tau|}.$$

Moreover, in Example 2.3 we computed the spectrum and spectral density of a generic AR(1) process. By setting $a = 0.5$ and $\sigma_{\boldsymbol{y}}^2 = \frac{4}{3}$ in (2.11), one gets

$$\phi_{\boldsymbol{y}}(z) = \frac{1}{\left(1 - \frac{1}{2} z^{-1}\right)\left(1 - \frac{1}{2} z\right)}$$

and

$$\phi_{\boldsymbol{y}}(e^{j\omega}) = \frac{1}{\frac{5}{4} - \cos\omega}.$$

Clearly, the same expressions can be obtained by direct application of Theorem 2.3, with $G(z) = \dfrac{z}{z - \frac{1}{2}}$. △

2.5.3 ARMA process

Consider the scalar s.p. $\boldsymbol{y}(t)$, defined by the equation

$$y(t) + a_1 y(t-1) + \cdots + a_n y(t-n) = c_0 e(t) + c_1 e(t-1) + \cdots + c_m e(t-m) \quad (2.22)$$

where $\boldsymbol{e}(t) \sim WP(0, \sigma_{\boldsymbol{e}}^2)$; $a_i \in \mathbb{R}$, $i = 0, 1, \ldots, n$, and $c_j \in \mathbb{R}$, $j = 0, 1, \ldots, m$. Such a s.p. is said *Auto-Regressive Moving Average* of order (n, m) (or briefly, ARMA(n, m)). Clearly, the MA and AR processes examined before are special cases of this general class.

The transfer function of the system can be derived by rewriting equation (2.22) as

$$A(z)y(t) = C(z)e(t),$$

with

$$A(z) = 1 + a_1 z^{-1} + \cdots + a_n z^{-n} \quad \text{e} \quad C(z) = c_0 + c_1 z^{-1} + \cdots + c_m z^{-m}.$$

Multiplying both $A(z)$ and $C(z)$ by z^n, we obtain the transfer function from $e(t)$ to $y(t)$

$$G(z) = \frac{C(z)}{A(z)} = \frac{c_0 z^n + c_1 z^{n-1} + \cdots + c_m z^{n-m}}{z^n + a_1 z^{n-1} + \cdots + a_n}.$$

If all the roots of $A(z)$ have modulus strictly smaller than 1, the system is asymptotically stable. Hence, thanks to Theorem 2.3, $\boldsymbol{y}(t)$ is asymptotically stationary.

The mean and covariance function of $\boldsymbol{y}(t)$ can be obtained by employing the same procedure proposed for the AR process. By computing the inverse Z-transform of $G(z)$ one gets the impulse response g_k, satisfying

$$y(t) = \sum_{k=0}^{\infty} g_k e(t-k).$$

Then, $m_{\boldsymbol{y}}$ and $R_{\boldsymbol{y}}(\tau)$ can be derived as in (2.19)-(2.21) (and practically computed according to Observation 2.2).

Example 2.6. Consider the ARMA(1,1) s.p.

$$y(t) - 0.9y(t-1) = e(t) + 0.5e(t-1),$$

with $\boldsymbol{e}(t) \sim WP(0,1)$. The transfer function from $\boldsymbol{e}(t)$ to $\boldsymbol{y}(t)$ is equal to

$$G(z) = \frac{C(z)}{A(z)} = \frac{z + 0.5}{z - 0.9}.$$

Being $p = 0.9$ the unique root of $A(z)$, the s.p. $\boldsymbol{y}(t)$ is asymptotically stationary. From Theorem 2.3:

$$m_{\boldsymbol{y}} = G(1)\, m_{\boldsymbol{e}} = 15 \cdot 0 = 0.$$

In order to derive the covariance function $R_{\boldsymbol{y}}(\tau)$, let us first find the impulse response of the system. Being

$$G(z) = \frac{z + 0.5}{z - 0.9} = 1 + \frac{1.4}{z - 0.9}$$

one has

$$g_k = \begin{cases} 1 & \text{if } k = 0 \\ 1.4 \cdot (0.9)^{k-1} & \text{if } k \geq 1 \end{cases}.$$

Then, by applying (2.20) one has

$$R_{\boldsymbol{y}}(0) = \left(\sum_{k=0}^{\infty} g_k^2\right)\sigma_{\boldsymbol{e}}^2 = 1 + (1.4)^2 \sum_{k=1}^{\infty} (0.9)^{2(k-1)} = 1 + \frac{(1.4)^2}{0.19} \simeq 11.32.$$

Finally, from (2.21) with $\tau > 0$ we get

$$\begin{aligned} R_{\boldsymbol{y}}(\tau) &= \left(\sum_{k=0}^{\infty} g_{k+\tau} g_k\right)\sigma_{\boldsymbol{e}}^2 = 1.4 \cdot (0.9)^{\tau-1} + (1.4)^2 \sum_{k=1}^{\infty} (0.9)^{k+\tau-1}(0.9)^{k-1} \\ &= \frac{1.4}{0.9}(0.9)^{\tau} + (1.4)^2 (0.9)^{\tau} \sum_{k=1}^{\infty} (0.9)^{2(k-1)} = \left\{\frac{1.4}{0.9} + \frac{(1.4)^2}{0.19}\right\}(0.9)^{\tau}. \end{aligned}$$

In order to compute the spectrum and the spectral density, one can exploit Theorem 2.3 to get

$$\phi_{\boldsymbol{y}}(z) = G(z)G(z^{-1})\phi_{\boldsymbol{e}}(z) = \frac{(1+0.5z^{-1})(1+0.5z)}{(1-0.9z^{-1})(1-0.9z)}$$

and

$$\phi_{\boldsymbol{y}}(e^{j\omega}) = \left|G(e^{j\omega})\right|^2 \phi_{\boldsymbol{e}}(e^{j\omega}) = \left|\frac{(1+0.5e^{j\omega})}{(1-0.9e^{j\omega})}\right|^2 = \frac{1.25+\cos\omega}{1.81-1.8\cos\omega}.$$

△

2.6 Ergodic processes

Let us assume that we observe a single realization $\bar{x}(t)$, for $t = 0, 1, \ldots, N$, of a stochastic process $\boldsymbol{x}(t)$. Is it possible to estimate the statistics of $\boldsymbol{x}(t)$ (e.g., the mean and covariance function) from the observed realization?

We say that a s.p. is *ergodic* if its statistics can be recovered from a single realization of the process itself, or equivalently if the "time averages" (computed on a single realization) are equal to the "ensemble averages" (the corresponding expected values computed on all possible realizations). By limiting our attention to the first and second order statistics, let us define the *sample mean* of $\bar{x}(t)$ as

$$\bar{m}_{\boldsymbol{x}}^N = \frac{1}{N}\sum_{t=1}^{N} \bar{x}(t)$$

and the *sample covariance* as

$$\bar{R}_x^N(\tau) = \frac{1}{N-\tau}\sum_{t=1}^{N-\tau} (\bar{x}(t+\tau) - \bar{m}_{\boldsymbol{x}})(\bar{x}(t) - \bar{m}_{\boldsymbol{x}})^T$$

Now we are ready to state the following simplified notion of ergodicity.

Definition 2.13. A stationary s.p. $x(t)$ is *ergodic* if

$$\lim_{N\to\infty} \bar{m}_{\boldsymbol{x}}^N = \mathbf{E}\left[\boldsymbol{x}(t)\right] = m_{\boldsymbol{x}} \tag{2.23}$$

$$\lim_{N\to\infty} \bar{R}_x^N(\tau) = \mathbf{E}\left[(\boldsymbol{x}(t+\tau) - m_{\boldsymbol{x}})(\boldsymbol{x}(t) - m_{\boldsymbol{x}})^T\right] = R_{\boldsymbol{x}}(\tau) \tag{2.24}$$

The convergence of the limits (2.23)-(2.24) must be intended *in probability*. In other words, $\boldsymbol{x}(t)$ is ergodic if $\bar{m}_{\boldsymbol{x}}^N$ and $\bar{R}_x^N(\tau)$ are consistent estimates of $m_{\boldsymbol{x}}$ and $R_{\boldsymbol{x}}(\tau)$, respectively.

An example of non-ergodic process is the Wiener process introduced in Section 2.2.3. In fact, its expected value is always zero, while its sample mean is a random variable whose variance goes to infinity with the number of samples N.

In the following, we will implicitly assume that we deal with ergodic processes and we will use sample estimates to approximate the statistics of a stationary stochastic process.

2.7 Spectral factorization

In Section 2.4, we have seen how the spectrum of a stochastic process is modified when it is filtered by an asymptotically stable linear system. Now we address the inverse problem: given a scalar stationary s.p. $\boldsymbol{y}(t)$, with a rational spectrum $\phi_{\boldsymbol{y}}(z)$, find a transfer function $H(z)$ and a variance $\sigma_{\boldsymbol{e}}^2$, such that

$$\phi_{\boldsymbol{y}}(z) = H(z)H(z^{-1})\sigma_{\boldsymbol{e}}^2.$$

This problem is known as *spectral factorization* and, as we will see, it is instrumental to derive the optimal prediction of a stochastic process.

The spectral factorization problem in general has not a unique solution, as it is shown by the following examples.

Example 2.7. Let $H_1(z)$ be a solution of the spectral factorization problem, with variance $\sigma_{\mathbf{1}}^2$, i.e.

$$\phi_{\boldsymbol{y}}(z) = H_1(z)H_1(z^{-1})\sigma_{\mathbf{1}}^2.$$

Then, also

$$H_2(z) = \beta H_1(z), \qquad \beta \neq 0$$

is a solution, with variance $\sigma_{\mathbf{2}}^2 = \dfrac{\sigma_{\mathbf{1}}^2}{\beta^2}$. In fact,

$$H_2(z)H_2(z^{-1})\sigma_{\mathbf{2}}^2 = \beta H_1(z)\beta H_1(z^{-1})\frac{\sigma_{\mathbf{1}}^2}{\beta^2} = H_1(z)H_1(z^{-1})\sigma_{\mathbf{1}}^2 = \phi_{\boldsymbol{y}}(z).$$

△

Example 2.8. Let $H_1(z)$ be a solution of the spectral factorization problem, with variance $\sigma_{\mathbf{1}}^2$ and assume that $H_1(z)$ has a zero in $z = \alpha$, i.e.

$$\phi_{\boldsymbol{y}}(z) = H_1(z)H_1(z^{-1})\sigma_{\mathbf{1}}^2 = \tilde{H}(z)(z-\alpha)\tilde{H}(z^{-1})(z^{-1}-\alpha)\sigma_{\mathbf{1}}^2.$$

Then, also

$$H_2(z) = \tilde{H}(z)\left(z - \frac{1}{\alpha}\right)$$

is a solution, with variance $\sigma_{\mathbf{2}}^2 = \alpha^2\sigma_{\mathbf{1}}^2$. In fact,

$$\begin{aligned} H_2(z)H_2(z^{-1})\sigma_{\mathbf{2}}^2 &= \tilde{H}(z)\left(z - \frac{1}{\alpha}\right)\tilde{H}(z^{-1})\left(z^{-1} - \frac{1}{\alpha}\right)\alpha^2\sigma_{\mathbf{1}}^2 \\ &= \tilde{H}(z)\frac{z}{\alpha}(\alpha - z^{-1})\tilde{H}(z^{-1})\frac{z^{-1}}{\alpha}(\alpha - z)\alpha^2\sigma_{\mathbf{1}}^2 \\ &= \tilde{H}(z)(z-\alpha)\tilde{H}(z^{-1})(z^{-1}-\alpha)\sigma_{\mathbf{1}}^2 = \phi_{\boldsymbol{y}}(z) \end{aligned}$$

Therefore, if a transfer function with a zero in α is a solution of the spectral factorization problem, with variance σ^2, then also the transfer function that we obtain by replacing the zero in α with its reciprocal $\frac{1}{\alpha}$ is a solution with variance $\alpha^2\sigma^2$. Clearly, the same reasoning holds for the poles of the transfer function. △

In order to guarantee uniqueness of the solution of the spectral factorization problem, it is necessary to enforce some constraints on the transfer function $H(z)$.

Theorem 2.4. *Given a rational spectrum*

$$\phi_{\boldsymbol{y}}(z) = \frac{\displaystyle\sum_{k=-m}^{m} \eta_k z^{-k}}{\displaystyle\sum_{h=-n}^{n} \gamma_h z^{-h}}$$

with $\eta_k = \eta_{-k}$ *and* $\gamma_h = \gamma_{-h}$ *(i.e.,* $\phi_{\boldsymbol{y}}(z) = \phi_{\boldsymbol{y}}(z^{-1})$*), that does not have zeros and poles with modulus exactly equal to 1, the spectral factorization*

$$\phi_{\boldsymbol{y}}(z) = H(z)H(z^{-1})\sigma_{\boldsymbol{e}}^2,$$

with

$$H(z) = \frac{C(z)}{A(z)}$$

exists and is unique if the following conditions hold:

1. *the polynomials* $A(z)$ *and* $C(z)$ *have the structure*

$$A(z) = 1 + a_1 z^{-1} + \cdots + a_n z^{-n}$$

$$C(z) = 1 + c_1 z^{-1} + \cdots + c_m z^{-m}$$

   and they are coprime, i.e., they do not have common roots;

2. *the roots of* $A(z)$ *and* $C(z)$ *are strictly inside the unit circle.*

The transfer function $H(z)$ satisfying the conditions of Theorem 2.4 is called *canonical spectral factor* of $\phi_{\boldsymbol{y}}(z)$. Observe that the second condition guarantees that $H(z)$ is asymptotically stable. Therefore, the s.p. $\boldsymbol{y}(t)$, with spectrum $\phi_{\boldsymbol{y}}(z)$, can be seen as the output of a linear system with transfer function $H(z)$, with all poles and zeros inside the unit circle, fed by a white process $\boldsymbol{e}(t)$ with variance $\sigma_{\boldsymbol{e}}^2$.

Example 2.9. Let

$$\phi_{\boldsymbol{y}}(z) = \frac{1}{1.25 - 0.5z - 0.5z^{-1}}$$

be the spectrum of the s.p. $\boldsymbol{y}(t)$. We aim at computing the canonical spectral factor

$$H(z) = \frac{C(z)}{A(z)}$$

and its variance $\sigma_{\boldsymbol{e}}^2$. Clearly, $C(z) = 1$. In order to compute $A(z)$, observe that the denominator of $\phi_{\boldsymbol{y}}(z)$ can be rewritten as

$$1.25 - 0.5z - 0.5z^{-1} = (1 - 0.5z^{-1})(1 - 0.5z),$$

and hence

$$A(z) = 1 - 0.5z^{-1}.$$

Therefore, the canonical spectral factor is given by

$$H(z) = \frac{1}{1 - 0.5z^{-1}},$$

with variance $\sigma_{\boldsymbol{e}}^2 = 1$. △

Example 2.10. Consider the s.p. $\boldsymbol{y}(t)$ generated as in Figure 2.10, where the input $\boldsymbol{u}(t)$ is a s.p. such that $u(t) = w(t) + 2w(t-1)$ and $G(z) = \frac{1}{1+0.5z^{-1}}$. Assuming that $\boldsymbol{w}(t) \sim WP(0,1)$, $\boldsymbol{v}(t) \sim WP(0,2)$, with $\boldsymbol{w}(t)$ and $\boldsymbol{v}(t)$ independent, let us compute the canonical spectral factor of $\phi_{\boldsymbol{y}}(z)$.

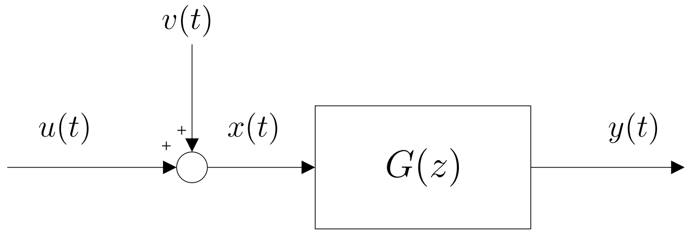


Figure 2.10: System of Example 2.10.

Denoting by $\phi_{\boldsymbol{x}}(z)$ the spectrum of the s.p.

$$\boldsymbol{x}(t) = \boldsymbol{u}(t) + \boldsymbol{v}(t),$$

the spectrum of $\boldsymbol{y}(t)$ can be obtained as

$$\phi_{\boldsymbol{y}}(z) = G(z)G(z^{-1})\phi_{\boldsymbol{x}}(z).$$

Since $\boldsymbol{w}(t)$ and $\boldsymbol{v}(t)$ are independent, we have

$$\phi_{\boldsymbol{x}}(z) = \phi_{\boldsymbol{u}}(z) + \phi_{\boldsymbol{v}}(z).$$

Clearly, being $\boldsymbol{v}(t)$ a white process, its spectrum is equal to

$$\phi_{\boldsymbol{v}}(z) = \sigma_{\boldsymbol{v}}^2 = 2,$$

while the spectrum of $\boldsymbol{u}(t)$ (which is a MA(1) process) is given by

$$\phi_{\boldsymbol{u}}(z) = 2z + 5 + 2z^{-1}.$$

Then,

$$\begin{aligned}\phi_{\boldsymbol{y}}(z) &= G(z)G(z^{-1})(\phi_{\boldsymbol{u}}(z) + \phi_{\boldsymbol{v}}(z)) \\ &= \frac{2z + 7 + 2z^{-1}}{(1 + 0.5z^{-1})(1 + 0.5z)}.\end{aligned}$$

We aim at finding the canonical spectral factor

$$H(z) = \frac{C(z)}{A(z)}$$

of $\phi_{\boldsymbol{y}}(z)$ and its related variance $\sigma_{\boldsymbol{e}}^2$. The denominator is already factorized, hence

$$A(z) = 1 + 0.5z^{-1}.$$

In order to obtain the numerator and the variance $\sigma_{\boldsymbol{e}}^2$, assume that $C(z)$ takes on the form

$$C(z) = 1 + cz^{-1},$$

where c is a coefficient to be determined. By equaling the polynomials

$$C(z)C(z^{-1})\sigma_{\boldsymbol{e}}^2 = (1 + cz^{-1})(1 + cz)\sigma_{\boldsymbol{e}}^2 = 2z + 7 + 2z^{-1}$$

we get the equations

$$\begin{cases} \sigma_{\boldsymbol{e}}^2 c = 2 \\ \sigma_{\boldsymbol{e}}^2(1 + c^2) = 7 \end{cases}$$

from which

$$\begin{cases} \sigma_{\boldsymbol{e}}^2 = \dfrac{2}{c} \\ 2c^2 - 7c + 2 = 0 \end{cases}$$

The second equation admits two solutions

$$c_1 = 0.3139, \qquad c_2 = 3.1861.$$

Notice that one is the reciprocal of the other (as it has to be, being $-c$ a zero of the spectrum). To comply with the conditions of Theorem 2.4, we select the solution corresponding to the zero of $C(z)$ inside the unit circle, i.e.

$$c = 0.3139.$$

Therefore,

$$\sigma_{\boldsymbol{e}}^2 = \frac{2}{c} = 6.3715$$

and

$$H(z) = \frac{1 + 0.3139z^{-1}}{1 + 0.5z^{-1}}.$$

△

2.8 Time series prediction

In this section, we address the following fundamental problem. Let $\boldsymbol{y}(t)$ be a stochastic process and assume we observe its values up to a certain time t. The objective is to estimate the value of $\boldsymbol{y}(t+l)$, with $l \in \mathbb{N}_+$. This is the *l-step ahead prediction problem for the time series $\boldsymbol{y}(t)$*. Since we want to estimate the value of the random variable (or vector) $\boldsymbol{y}(t+l)$, on the basis of the observations of the random variables $Y_{-\infty}^t = \{y(t), y(t-1), y(t-2), \dots\}$, this is clearly a Bayesian estimation problem. We will denote the estimate as $\hat{y}(t+l|Y_{-\infty}^t)$, or briefly $\hat{y}(t+l|t)$ (meaning the estimate of $\boldsymbol{y}(t+l)$ based on the data up to time t).

As a criterion to derive an optimal estimator, we choose the Mean Square Error (MSE). In order to minimize the MSE, one has to compute the conditional expected value of $\boldsymbol{y}(t+l)$, given the measurements $Y_{-\infty}^t$, i.e., $E[\boldsymbol{y}(t+l)|Y_{-\infty}^t]$.

Hereafter, we will assume that $\boldsymbol{y}(t)$ is a scalar stationary stochastic process, with spectrum $\phi_{\boldsymbol{y}}(z)$. Let

$$H(z) = \frac{C(z)}{A(z)}$$

be the canonical spectral factor of $\phi_{\boldsymbol{y}}(z)$, with variance $\sigma_{\boldsymbol{e}}^2$. This means that, if we consider a white process $\boldsymbol{e}(t)$ with variance $\sigma_{\boldsymbol{e}}^2$, we can write

$$\boldsymbol{y}(t) = H(z)\boldsymbol{e}(t). \tag{2.25}$$

It is always possible to rewrite $H(z)$ as

$$\frac{C(z)}{A(z)} = Q(z) + \frac{R(z)}{A(z)}z^{-l} \tag{2.26}$$

in which $Q(z)$ is the quotient and $R(z)z^{-l}$ is the reminder of the polynomial division $\frac{C(z)}{A(z)}$. To derive the polynomials $Q(z)$ and $R(z)$, one has to proceed with the polynomial division (in the variable z^{-1}) until in the remainder it is possible to collect the common factor z^{-l} in all the terms of the polynomial (this occurs after *at most* l steps of the division). Then, the polynomial $Q(z)$ will take on the form

$$Q(z) = q_0 + q_1 z^{-1} + \dots q_{l-1} z^{-(l-1)}.$$

Moreover, it can be verified that $R(z) = r_0 + r_1 z^{-1} + \cdots + r_h z^{-h}$, with h smaller than the degree of the polynomial $A(z)$. Now, from (2.25)-(2.26) one gets

$$\begin{aligned} y(t+l) &= H(z)e(t+l) = \frac{C(z)}{A(z)}e(t+l) \\ &= \left(Q(z) + \frac{R(z)}{A(z)}z^{-l}\right)e(t+l) \\ &= Q(z)e(t+l) + \frac{R(z)}{A(z)}e(t). \end{aligned} \tag{2.27}$$

Moreover, due to the properties of the canonical spectral factor, $H(z)$ is asymptotically stable and also its inverse $H^{-1}(z)$ is asymptotically stable (recall that also the zeros of the canonical spectral factor are inside the unit circle). Therefore, it is possible to invert the relationship (2.25), thus obtaining

$$e(t) = H^{-1}(z)y(t) = \frac{A(z)}{C(z)}y(t).$$

By substituting the previous equation in (2.27), one gets

$$y(t+l) = Q(z)e(t+l) + \frac{R(z)}{C(z)}y(t). \tag{2.28}$$

Notice that the first term on the right corresponds to

$$Q(z)e(t+l) = q_0 e(t+l) + q_1 e(t+l-1) + \cdots + q_{l-1}e(t+1)$$

and it depends on values taken by the white process $\boldsymbol{e}(\cdot)$ at future time instants, with respect to the current time t. Being the process white, its values $e(t+i)$, $i = 1, \ldots, l$ are independent from all past observations $Y_{-\infty}^t = \{y(t), y(t-1), y(t-2), \ldots\}$. Therefore, their conditional expectation is equal to the mean value, i.e.

$$\hat{e}(t+i|t) = \mathbf{E}\left[\boldsymbol{e}(t+i)|Y_{-\infty}^t\right] = \mathbf{E}\left[\boldsymbol{e}(t+i)\right] = 0.$$

Conversely, the second term on the right in (2.28) is known, being a linear combination of the available data $Y_{-\infty}^t$. Hence, we can conclude that the minimum MSE estimator of $\boldsymbol{y}(t+l)$, based on the measurements available up to time t, is given by

$$\hat{y}(t+l|t) = \frac{R(z)}{C(z)}y(t). \tag{2.29}$$

The transfer function $P(z) = \dfrac{R(z)}{C(z)}$ is known as the *l-step ahead Wiener predictor*[5]. Notice that it is a linear function of the data.

[5]Norbert Wiener solved the problem of estimating a continuous-time signal $s(t)$ from observations of another signal $y(t)$ in the 1940s. In the same years, Andrey Kolmogorov solved the problem independently in the discrete-time setting.

In order to compute the prediction MSE, by subtracting (2.29) from (2.28) we get

$$\begin{aligned}\tilde{y}(t+l) &= y(t+l) - \hat{y}(t+l|t) = Q(z)e(t+l)\\ &= q_0 e(t+l) + q_1 e(t+l-1) + \cdots + q_{l-1} e(t+1)\\ &= \sum_{i=0}^{l-1} q_i e(t+l-i).\end{aligned}$$

Therefore,

$$MSE^* = \mathbf{E}\left[\tilde{y}(t+l)^2\right] = \mathbf{E}\left[\left(\sum_{i=0}^{l-1} q_i \boldsymbol{e}(t+l-i)\right)^2\right] = \sigma_{\boldsymbol{e}}^2 \sum_{i=0}^{l-1} q_i^2, \tag{2.30}$$

being the cross products all equal to zero, due to whiteness of $\boldsymbol{e}(t)$. Notice that, the MSE in (2.30) grows with l, i.e., with the prediction step.

For the special case $l = 1$ (1-step ahead predictor), (2.26) boils down to

$$\frac{C(z)}{A(z)} = 1 + \frac{C(z) - A(z)}{A(z)} = 1 + \frac{R(z)z^{-1}}{A(z)}$$

where

$$\begin{aligned}C(z) - A(z) &= (1 + c_1 z^{-1} + \cdots + c_m z^{-m}) - (1 + a_1 z^{-1} + \cdots + a_n z^{-n})\\ &= (c_1 - a_1)z^{-1} + \cdots + (c_m - a_m)z^{-m} - a_{m+1}z^{-(m+1)} - \ldots - a_n z^{-n}\end{aligned}$$

and hence

$$R(z) = (C(z) - A(z))z.$$

So, the 1-step ahead Wiener predictor becomes

$$P(z) = \frac{R(z)}{C(z)} = \frac{(C(z) - A(z))z}{C(z)} = z\left(1 - \frac{A(z)}{C(z)}\right) = z\left(1 - \frac{1}{H(z)}\right)$$

where $H(z)$ is the canonical spectral factor of $\boldsymbol{y}(t)$. Then,

$$\hat{y}(t+1|t) = z\left(1 - \frac{1}{H(z)}\right) y(t) = \left(1 - \frac{1}{H(z)}\right) y(t+1)$$

Summing up, the Wiener predictor of a time series $\boldsymbol{y}(t)$ can be computed according to the following procedure:

1. Compute the spectrum $\phi_{\boldsymbol{y}}(z)$ of $\boldsymbol{y}(t)$.

2. Compute the canonical spectral factor $H(z)$, with variance $\sigma_{\boldsymbol{e}}^2$.

3. Compute $Q(z)$ and $R(z)$ according to (2.26).

4. The minimum MSE predictor is given by (2.29), with prediction MSE equal to (2.30).

Example 2.11. Consider the ARMA(1,1) process, generated by the equation

$$y(t) - \frac{1}{2}y(t-1) = w(t) - 4w(t-1),$$

where $\boldsymbol{w}(t) \sim WP(0,2)$. We want to compute the 2-step ahead Wiener predictor $\hat{y}(t+2|t)$. The transfer function $W(z)$ of the system with input $\boldsymbol{w}(t)$ and output $\boldsymbol{y}(t)$ is

$$W(z) = \frac{1-4z^{-1}}{1-\frac{1}{2}z^{-1}},$$

and hence the spectrum of $\boldsymbol{y}(t)$ is obtained as

$$\phi_{\boldsymbol{y}}(z) = W(z)W(z^{-1})\sigma_{\boldsymbol{w}}^2 = \frac{(1-4z^{-1})(1-4z)}{(1-\frac{1}{2}z^{-1})(1-\frac{1}{2}z)} \cdot 2$$

In order to compute the canonical spectral factor

$$H(z) = \frac{C(z)}{A(z)}$$

we observe that from the factorization of the denominator, one gets

$$A(z) = 1 - \frac{1}{2}z^{-1}.$$

Conversely, the polynomial in z^{-1} in the numerator of $\phi_{\boldsymbol{y}}(z)$ has an unstable zero at $z = 4$. In order to compute $C(z)$ we exploit the observation in Example 2.8, so that we can exchange the unstable zero with its reciprocal, provided that the variance of the input signal is set to $\sigma_{\boldsymbol{e}}^2 = 16 \cdot \sigma_{\boldsymbol{w}}^2$. This means that the spectrum can be rewritten as

$$\phi_{\boldsymbol{y}}(z) = \frac{(1-\frac{1}{4}z^{-1})(1-\frac{1}{4}z)}{(1-0.5z^{-1})(1-0.5z)} \cdot 32$$

from which

$$C(z) = 1 - \frac{1}{4}z^{-1}$$

and $\sigma_{\boldsymbol{e}}^2 = 32$. Then, the 2-step ahead predictor is given by

$$\hat{y}(t+2|t) = \frac{R(z)}{C(z)}y(t),$$

where

$$\frac{C(z)}{A(z)} = Q(z) + \frac{R(z)}{A(z)}z^{-2}.$$

Let us perform the polynomial division $\frac{C(z)}{A(z)}$ until there is a common factor z^{-2} in the remainder:

$$\begin{array}{ll|l}
1-\frac{1}{4}z^{-1} & & 1-\frac{1}{2}z^{1} \\ \cline{3-3}
1-\frac{1}{2}z^{-1} & & 1+\frac{1}{4}z^{-1} \\ \cline{1-1}
\frac{1}{4}z^{-1} & & \\
\frac{1}{4}z^{-1} & -\frac{1}{8}z^{-2} & \\ \cline{1-2}
 & \frac{1}{8}z^{-2} &
\end{array}$$

Hence,

$$R(z) = \frac{1}{8} \qquad Q(z) = 1 + \frac{1}{4}z^{-1}$$

and the 2-step ahead predictor takes on the form

$$\hat{y}(t+2|t) = \frac{\frac{1}{8}}{1 - \frac{1}{4}z^{-1}} y(t).$$

In the time domain, this corresponds to the difference equation

$$\hat{y}(t+2|t) = \frac{1}{4}\hat{y}(t+1|t-1) + \frac{1}{8}y(t).$$

The prediction MSE is equal to

$$MSE^* = (q_0^2 + q_1^2)\sigma_{\boldsymbol{e}}^2 = \left(1 + \frac{1}{16}\right) 32 = 34.$$

Example 2.12. Consider the system depicted in Figure 2.11, where $\boldsymbol{u}(t)$ is a s.p. satisfying the equation

$$u(t) = w(t) + \frac{1}{5}w(t-1),$$

with $\boldsymbol{w}(t) \sim WP(0, 4)$, and

$$W_1(z) = \frac{1}{1 + \frac{1}{4}z^{-1}}, \qquad W_2(z) = \frac{1 + \frac{3}{10}z^{-1}}{1 + \frac{1}{4}z^{-1}}.$$

We want to compute the 2-step ahead Wiener predictor of $\boldsymbol{y}_2(t)$, $\hat{y}_2(t+2|t)$.

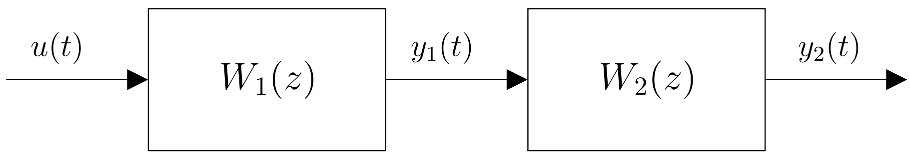


Figure 2.11.

Let us first compute the spectrum of $\boldsymbol{y}_2$. Let

$$W(z) = W_1(z)W_2(z) = \frac{(1 + \frac{3}{10}z^{-1})}{(1 + \frac{1}{4}z^{-1})^2}$$

be the transfer function from $u(t)$ to $y_2(t)$. Then,

$$\phi_{\boldsymbol{y_2}}(z) = W(z)W(z^{-1})\phi_{\boldsymbol{u}}(z),$$

where

$$\phi_{\boldsymbol{u}}(z) = \left(1 + \frac{1}{5}z^{-1}\right)\left(1 + \frac{1}{5}z\right) 4.$$

Hence,

$$\begin{aligned}\phi_{\boldsymbol{y_2}}(z) &= \frac{(1 + \frac{1}{5}z^{-1})(1 + \frac{3}{10}z^{-1})}{(1 + \frac{1}{4}z^{-1})^2} \frac{(1 + \frac{1}{5}z)(1 + \frac{3}{10}z)}{(1 + \frac{1}{4}z)^2} \cdot 4 \\ &\triangleq \frac{C(z)}{A(z)} \frac{C(z^{-1})}{A(z^{-1})}\sigma_{\boldsymbol{e}}^2.\end{aligned}$$

Observing the spectrum, it can be noticed that the polynomials in z^{-1} already satisfy the conditions of Theorem 2.4; therefore, the canonical spectral factor turns out to be

$$H(z) = \frac{C(z)}{A(z)} = \frac{(1+\frac{1}{5}z^{-1})(1+\frac{3}{10}z^{-1})}{(1+\frac{1}{4}z^{-1})^2} = \frac{1+\frac{1}{2}z^{-1}+\frac{3}{50}z^{-2}}{1+\frac{1}{2}z^{-1}+\frac{1}{16}z^{-2}}$$

with variance $\sigma_e^2 = 4$. In order to compute the predictor, let us perform the division $\dfrac{C(z)}{A(z)}$:

$$\begin{array}{r|l} 1+\frac{1}{2}z^{-1}+\frac{3}{50}z^{-2} & 1+\frac{1}{2}z^{-1}+\frac{1}{16}z^{-2} \\ \cline{2-2} 1+\frac{1}{2}z^{-1}+\frac{1}{16}z^{-2} & 1 \\ \hline -\frac{1}{400}z^{-2} & \end{array}$$

from which

$$R(z) = -\frac{1}{400} \qquad Q(z) = 1.$$

Notice that, despite we are looking for the 2-step ahead predictor, one division step has been sufficient in this case to isolate the common factor z^{-2} in the remainder. In general, with a prediction horizon of length l, one has to do at most l division steps. Hence, the predictor transfer function is given by

$$P(z) = \frac{R(z)}{C(z)} = \frac{-\frac{1}{400}}{1+\frac{1}{2}z^{-1}+\frac{3}{50}z^{-2}},$$

which corresponds to the time-domain equation

$$\hat{y}_2(t+2|t) = -\frac{1}{2}\hat{y}_2(t+1|t-1) - \frac{3}{50}\hat{y}_2(t|t-2) - \frac{1}{400}y(t).$$

Finally, the prediction error is equal to

$$MSE^* = q_0^2\sigma_e^2 = 4.$$

2.9 Exercises

2.1. Consider the s.p. $\boldsymbol{z}(t)$ obtained as the sum of two s.p. $\boldsymbol{x}(t),\ \boldsymbol{y}(t)$:

$$\boldsymbol{z}(t) = \boldsymbol{x}(t) + \boldsymbol{y}(t),$$

with

$$\begin{aligned}
\mathbf{E}\left[\boldsymbol{x}(t)\right] &= m_{\boldsymbol{x}};\\
\mathbf{E}\left[\boldsymbol{y}(t)\right] &= m_{\boldsymbol{y}};\\
\mathbf{E}\left[(\boldsymbol{x}(t+\tau) - m_{\boldsymbol{x}})(\boldsymbol{x}(t) - m_{\boldsymbol{x}})^T\right] &= R_{\boldsymbol{x}}(\tau);\\
\mathbf{E}\left[(\boldsymbol{y}(t+\tau) - m_{\boldsymbol{y}})(\boldsymbol{y}(t) - m_{\boldsymbol{y}})^T\right] &= R_{\boldsymbol{y}}(\tau);\\
\mathbf{E}\left[(\boldsymbol{x}(t+\tau) - m_{\boldsymbol{x}})(\boldsymbol{y}(t) - m_{\boldsymbol{y}})^T\right] &= R_{\boldsymbol{xy}}(\tau).
\end{aligned}$$

Verify that:

- $R_{\boldsymbol{z}}(\tau) = R_{\boldsymbol{x}}(\tau) + R_{\boldsymbol{y}}(\tau) + R_{\boldsymbol{xy}}(\tau) + R_{\boldsymbol{yx}}(\tau)$;
- $\phi_{\boldsymbol{z}}(z) = \phi_{\boldsymbol{x}}(z) + \phi_{\boldsymbol{y}}(z) + \phi_{\boldsymbol{xy}}(z) + \phi_{\boldsymbol{yx}}(z)$.

What happens if $\boldsymbol{x}(t)$ e $\boldsymbol{y}(t)$ are uncorrelated?

2.2. Establish which of the following functions $R_i(\tau),\ i = 1, \ldots, 3$ can be the covariance function of a stationary stochastic process, motivating the answer. Say which type of process it is and compute its variance.

$$\begin{array}{lll}
R_1(0) = -1 & R_2(0) = 1 & R_3(0) = 1\\
R_1(\pm 1) = 0.5 & R_2(\pm 1) = 0.5 & R_3(1) = 0.5\\
R_1(\tau) = 0\ |\tau| > 1 & R_2(\tau) = 0,\ |\tau| > 1 & R_3(-1) = 0.7\\
 & & R_3(\tau) = 0,\ |\tau| > 1
\end{array}$$

2.3. Let $\boldsymbol{e}(t) \sim WP(0,1)$. Show that the s.p.

$$\boldsymbol{w}(t) = \sum_{k=1}^{t} \boldsymbol{e}(k)$$

is a Wiener process.

2.4. Consider the MA(2) process

$$y(t) = e(t) + 2e(t-1) + e(t-2), \qquad \boldsymbol{e} \sim WP(0,1).$$

Compute the covariance function $R_y(t_1, t_2)$ and show that it depends only on the lag $\tau = t_1 - t_2$. Derive the transfer function from $\boldsymbol{e}(t)$ to $\boldsymbol{y}(t)$, the spectral density $\phi_{\boldsymbol{y}}(e^{j\omega})$ and plot it for $\omega \in [-\pi, \pi]$.

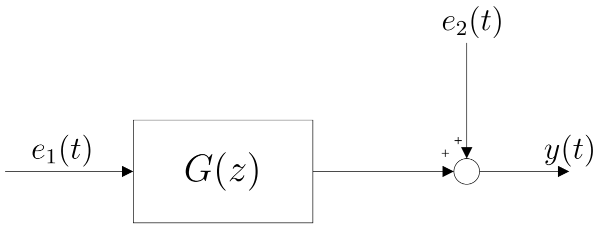


Figure 2.12.

2.5. Consider the system in Figure 2.12, where $\boldsymbol{e}_1 \sim WP(0,1)$ and

$$G(z) = \frac{z}{z - \frac{1}{4}}.$$

Compute the mean and spectral density of $\boldsymbol{y}(t)$ in the following cases:

a) $\boldsymbol{e}_2(t) = 0, \ \forall t$;

b) $\boldsymbol{e}_2(t) \sim WP(0,1)$ independent from $\boldsymbol{e}_1(t)$;

c) $\boldsymbol{e}_2(t) = \boldsymbol{e}_1(t), \ \forall t$.

2.6. Consider the stochastic process

$$y(t) = 0.3y(t-1) + e(t) + 0.6e(t-1)$$

where $\boldsymbol{e}(t) \sim WP(0,2)$.

a) Compute the transfer function $G(z)$ from $\boldsymbol{e}(t)$ to $\boldsymbol{y}(t)$.

b) Compute $E[\boldsymbol{y}(t)]$ and the covariance function $R_{\boldsymbol{y}}(\tau)$ of the process $\boldsymbol{y}(t)$ at steady state.

c) Compute the spectrum $\phi_{\boldsymbol{y}}(z)$ and the spectral density $\phi_{\boldsymbol{y}}(e^{j\omega})$ of the process $\boldsymbol{y}(t)$.

2.7. Consider a stationary s.p. $\boldsymbol{y}(t)$ with covariance function $R_{\boldsymbol{y}}(\tau)$. The normalized covariance function $r_{\boldsymbol{y}}(\tau) = R_{\boldsymbol{y}}(\tau)/R_{\boldsymbol{y}}(0)$ is depicted in Fig. 2.13.

1. Establish which one, among the following systems, might have generated the process $\boldsymbol{y}(t)$:

   a) $y(t) = e(t) + 0.5e(t-1) + 1.8e(t-2)$

   b) $y(t) = 0.7y(t-1) + e(t) + e(t-1)$

   c) $y(t) = -0.7y(t-1) + e(t) + e(t-1)$

   in which $\boldsymbol{e} \sim WP(0,1)$.

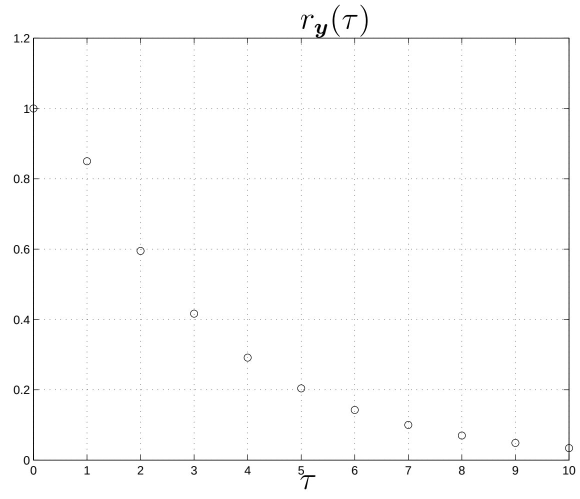


Figure 2.13.

2. Plot the spectral density $\phi_{\boldsymbol{y}}(e^{j\omega})$.

2.8. Consider the stochastic process $\boldsymbol{y}(t)$ generated as in Figure 2.14, where the input $\boldsymbol{u}(t)$ is a stochastic process such that $u(t) = e(t) + 2e(t-1)$ and

$$G(z) = \frac{1}{1 + 0.5z^{-1}}.$$

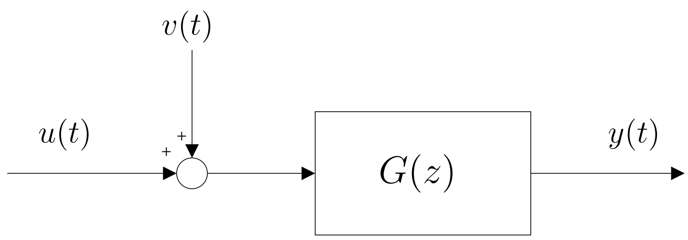


Figure 2.14.

Assume that $\boldsymbol{e}(t) \sim WP(0,1)$, $\boldsymbol{v}(t) \sim WP(0,1)$, and they are independent. Find the spectral density $\phi_{\boldsymbol{y}}(e^{j\omega})$ of $\boldsymbol{y}(t)$ and plot it.

2.9. Consider the interconnection in Figure 2.15, in which $\boldsymbol{e}(t) \sim WP(0,16)$ and K is a real constant.

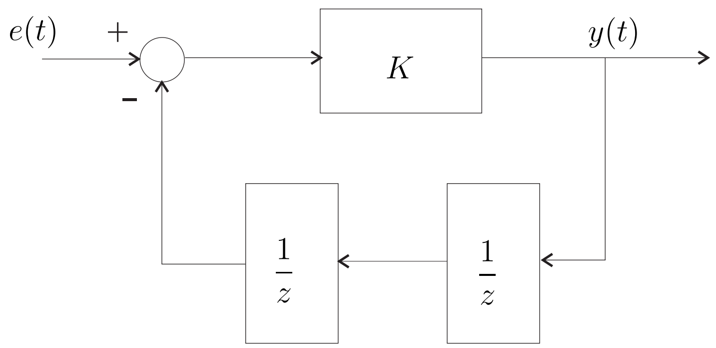


Figure 2.15.

a) Find the values of $K \in \mathbb{R}$ for which $\boldsymbol{y}(t)$ is asymptotically stationary.

b) Assuming $K = -\frac{1}{3}$, compute the covariance function $R_{\boldsymbol{y}}(\tau)$ of the process $\boldsymbol{y}(t)$.

c) Assuming $K = \frac{1}{5}$, find the equation of the 2-step ahead Wiener predictor $\hat{y}(t+2|t)$ for the process $\boldsymbol{y}(t)$ and the corresponding prediction $MSE = \mathbf{E}\left[\{\boldsymbol{y}(t+2) - \hat{y}(t+2|t)\}^2\right]$.

d) Find the values of K, for which the MSE of the 4-step ahead Wiener predictor, $\mathbf{E}\left[\{\boldsymbol{y}(t+4) - \hat{y}(t+4|t)\}^2\right]$, is smaller than 60.

2.10. Consider the s.p. described by the equation

$$y(t) = e(t) - \frac{3}{2}e(t-2) - e(t-4)$$

in which $\boldsymbol{e}(t) \sim WP(0,1)$. Find the l-step ahead Wiener predictor, for the values $l = 1, 2, 3, 4$, and compute the corresponding prediction MSE.

2.11. Consider the system shown in Figure 2.16, in which $\boldsymbol{u}(t)$ is a s.p. satisfying the equation

$$u(t) = 0.6u(t-1) + e(t),$$

with $\boldsymbol{e}(t) \sim WP(0,1)$, and

$$G(z) = 1 + 4z^{-1}.$$

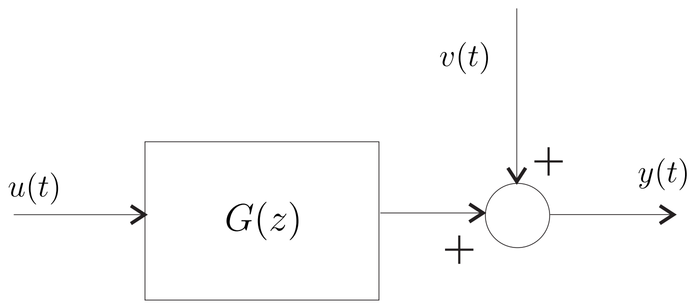


Figure 2.16.

a) Assuming $v(t) = 0$, $\forall t$, find the 2-step ahead Wiener predictor $\hat{y}(t+2|t)$ for the s.p. $\boldsymbol{y}(t)$, and the corresponding prediction error, $MSE = \mathbf{E}\left[\{\boldsymbol{y}(t+2) - \hat{y}(t+2|t)\}^2\right]$.

b) Repeat item a), assuming $\boldsymbol{v}(t) \sim WP(0,10)$, independent from $\boldsymbol{e}(t)$.

c) Now, assume that $\boldsymbol{y}(t)$ is an AR(1) process, whose equation is

$$y(t) = 0.6y(t-1) + w(t),$$

with $\boldsymbol{w}(t) \sim WP(0,19)$. Assuming that $\boldsymbol{v}(t)$ and $\boldsymbol{e}(t)$ are independent, compute the spectrum $\phi_{\boldsymbol{v}}(z)$ of $\boldsymbol{v}(t)$ and write a model of the s.p. $\boldsymbol{v}(t)$.

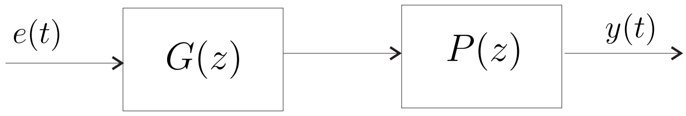


Figure 2.17.

2.12. Consider the system shown in Figure 2.17, in which $\boldsymbol{e}(t) \sim WP(0,3)$ and $P(z) = 1 - 2z^{-1}$.

a) Assuming $G(z) = 1$, find the equation of the 1-step ahead Wiener predictor $\hat{y}(t+1|t)$ of $\boldsymbol{y}(t)$, and the corresponding prediction error, $MSE = \mathbf{E}\left[\{\boldsymbol{y}(t+1) - \hat{y}(t+1|t)\}^2\right]$.

b) Assuming $G(z) = K$, with $K > 0$, find for which values of K the MSE of the 1-step ahead Wiener predictor is smaller than 10.

c) Assuming $G(z) = \frac{1}{5 - z^{-1}}$, repeat the exercise in item a).

2.13. Consider the system depicted in Figure 2.18, in which $\boldsymbol{e}_1(t) \sim WP(0,2)$, $\boldsymbol{e}_2(t) \sim WP(0,1)$, and $\boldsymbol{e}_1(t)$, $\boldsymbol{e}_2(t)$ are independent. Moreover,

$$G_1(z) = 1 - \frac{1}{3}z^{-1} \quad , \quad G_2(z) = 1 + 2z^{-1} \ .$$

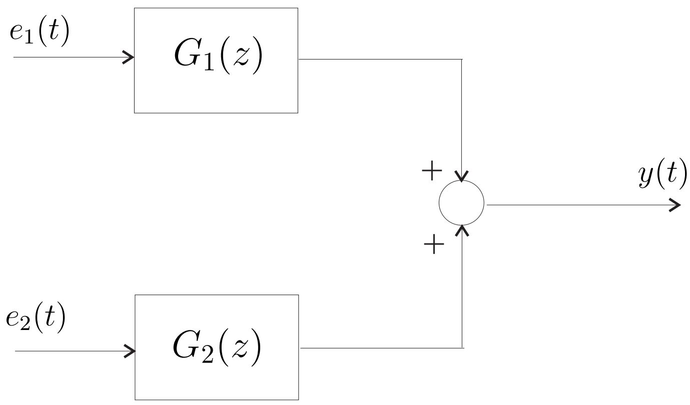


Figure 2.18.

a) Find the spectrum $\phi_{\boldsymbol{y}}(z)$ of the process $\boldsymbol{y}(t)$.

b) Derive the equation of the 1-step ahead Wiener predictor $\hat{y}(t+1|t)$ for $\boldsymbol{y}(t)$ and the corresponding prediction error, $MSE = \mathbf{E}\left[\{\boldsymbol{y}(t+1) - \hat{y}(t+1|t)\}^2\right]$.

c) Consider the "trivial" predictor $\tilde{y}(t+1|t) = \boldsymbol{y}(t)$. Compute the mean square error $MSE_{trivial} = \mathbf{E}\left[\{\boldsymbol{y}(t+1) - \tilde{y}(t+1|t)\}^2\right]$ and compare it with the MSE computed in item b).

d) Let $\boldsymbol{r}(t)$ be the stochastic process generated as the output of a linear system with transfer function $W(z)$, whose input is the process $\boldsymbol{y}(t)$. Determine $W(z)$ in such a way that $\boldsymbol{r}(t)$ is a white process, with variance equal to $\frac{1}{4}$.

2.14. Consider again the system in Figure 2.18. Now assume $\boldsymbol{e}_1(t) \sim WP(0,6)$, $\boldsymbol{e}_2(t) \sim WP(0,2)$, and $\boldsymbol{e}_1(t)$, $\boldsymbol{e}_2(t)$ are independent. Moreover,

$$G_1(z) = \frac{1}{1-\frac{1}{2}z^{-1}} \quad , \quad G_2(z) = \frac{1}{1-p\,z^{-1}} \ .$$

a) Find the values of $p \in \mathbb{R}$ for which the process $\boldsymbol{y}(t)$ is asymptotically stationary.

b) Assuming $p = \frac{1}{3}$, find the equation of the 2-step ahead Wiener predictor $\hat{y}(t+2|t)$ of $\boldsymbol{y}(t)$, and the corresponding prediction error $MSE = \mathbf{E}\left[\{\boldsymbol{y}(t+2) - \hat{y}(t+2|t)\}^2\right]$.

c) Assuming $p = \frac{1}{2}$, find the maximum value of the variance $\sigma^2_{\mathbf{2}}$ of $\boldsymbol{e}_2(t)$, such that the MSE of the 2-step ahead Wiener predictor is not larger than 20.

2.15. Consider the system depicted in Figure 2.19, in which $\boldsymbol{u}(t)$ and $\boldsymbol{v}(t)$ are independent white processes, with zero mean and variance equal to 1, and

$$G(z) = \frac{1+\frac{1}{2}z^{-1}}{1-\frac{1}{4}z^{-1}} \quad , \quad H(z) = \frac{1+2z^{-1}}{1-\frac{1}{3}z^{-1}}.$$

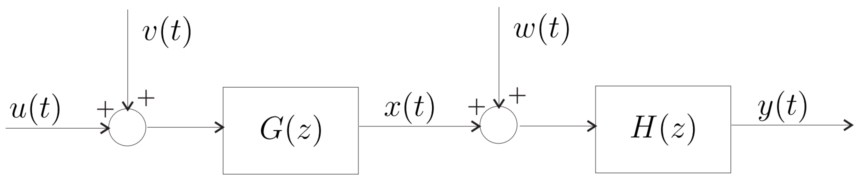


Figure 2.19.

a) Find the 1-step ahead Wiener predictor $\hat{x}(t+1|t)$ of the process $\boldsymbol{x}(t)$, and the prediction error $MSE = \mathbf{E}\left[\{\boldsymbol{x}(t+1) - \hat{x}(t+1|t)\}^2\right]$.

b) Assuming $\boldsymbol{w}(t) \equiv 0$, find the 2-step ahead Wiener predictor $\hat{y}(t+2|t)$ for the process $\boldsymbol{y}(t)$, and the corresponding prediction error $MSE = \mathbf{E}\left[\{\boldsymbol{y}(t+2) - \hat{y}(t+2|t)\}^2\right]$.

c) Find the spectrum $\phi_{\boldsymbol{y}}(z)$ of the process $\boldsymbol{y}(t)$ in the following cases:

1) $\boldsymbol{w}(t)$ is a white process, independent from $\boldsymbol{u}(t)$ and $\boldsymbol{v}(t)$, with variance $\sigma^2_{\boldsymbol{w}} = 9$;

2) $\boldsymbol{w}(t) \equiv \boldsymbol{v}(t)$, $\forall t$.