8

Gaussian Signals, Correlation Matrices, and Sample Path Properties

In general, determination of the shape of the sample paths of a random signal $X(t)$ requires knowledge of n-point probabilities

$$\mathbf{P}(a_1 < X(t_1) < b_1, \ldots, a_n < X(t_n) < b_n)$$

for an arbitrary n, and arbitrary windows $a_1 < b_1, \ldots, a_n < b_n$. But usually this information cannot be recovered if the only signal characteristic known is the autocorrelation function. The latter depends on the two-point distributions but does not uniquely determine them. However, in the case of Gaussian signals, the autocorrelations determine not only the two-point probability distributions but also all the n-point probability distributions, so that complete information is available within the second-order theory. For example, this means that you only have to estimate means and covariances to make the model. Also, in the Gaussian universe, weak stationarity implies strict stationarity as defined in Chapter 4. For the sake of simplicity all signals in this chapter are assumed to be real-valued. The chapter ends with a more subtle analysis of sample path properties of stationary signals such as continuity and differentiability; in the Gaussian case the information is particularly complete.

Of course, faced with real-world data the proposition that they are distributed according to a Gaussian distribution must be tested rigorously. Many such tests have been developed by the statisticians.[33] In other cases, one can make an argument in favor of such a hypothesis based on the central limit theorem (3.5.5)–(3.5.6).

---

[33] See, e.g., M. Denker and W. A. Woyczyński's book mentioned in previous chapters.

8.1 Linear transformations of random vectors

In Chapter 3, we have calculated probability distributions of transformed random quantities. Repeating that procedure in the case of a linear transformation of the 1D random quantity X given by the formula

$$Y = aX, \quad a > 0, \tag{8.1.1}$$

we can obtain the cumulative distribution function (c.d.f.) $F_Y(y)$ of the random quantity Y in terms of the c.d.f. $F_X(x)$ of the random quantity X as follows:

$$F_Y(y) = P(Y \le y) = P(aX \le y) = P\left(X \le \frac{y}{a}\right) = F_X\left(\frac{y}{a}\right). \tag{8.1.2}$$

To obtain an analogous formula for the probability density functions (p.d.f.s), it suffices to differentiate both sides of (8.1.2) to see that

$$f_Y(y) = \frac{d}{dy}F_Y(y) = \frac{1}{a}f_X\left(\frac{u}{a}\right). \tag{8.1.3}$$

Example 8.1.1. Consider a standard Gaussian random quantity $X \sim N(0,1)$ with the p.d.f.

$$f_X(x) = \frac{1}{\sqrt{2\pi}}\exp\left(\frac{-x^2}{2}\right). \tag{8.1.4}$$

Then the random quantity $Y = aX$, $a > 0$, has the p.d.f.

$$f_Y(y) = \frac{1}{\sqrt{2\pi}a}\exp\left(-\frac{x^2}{2a^2}\right). \tag{8.1.5}$$

Obviously, the expectation $\mathbf{E}Y = \mathbf{E}(aX) = a\mathbf{E}X = 0$ and the variance of Y is

$$\sigma_Y^2 = \mathbf{E}(aX)^2 = a^2\mathbf{E}X^2 = a^2. \tag{8.1.6}$$

If we conduct the same argument for $a < 0$, the p.d.f. of $Y = aX$ will be

$$f_Y(y) = \frac{1}{\sqrt{2\pi}(-a)}\exp\left(-\frac{x^2}{2a^2}\right). \tag{8.1.7}$$

Thus formulas (8.1.6) and (8.1.7) can be unified in a single statement: If $X \sim N(0,1)$, then for any $a \neq 0$, random quantity $Y = aX$ has p.d.f.

$$f_Y(y) = \frac{1}{\sqrt{2\pi}|a|}\exp\left(-\frac{x^2}{2a^2}\right). \tag{8.1.8}$$

Using the above elementary reasoning as a model we will now derive the formula for a d-dimensional p.d.f.

$$f_{\vec{Y}}(\vec{y}) = f_{\vec{Y}}(y_1, \dots, y_d)$$

of a random vector

$$\vec{Y} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_d \end{pmatrix}$$

obtained by a nondegenerate (invertible) linear transformation

$$\vec{Y} = \mathbf{A}\vec{X} \tag{8.1.9}$$

consisting of multiplication of the random vector

$$\vec{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_d \end{pmatrix}$$

with a known p.d.f.

$$f_{\vec{X}}(\vec{x}) = f_{\vec{X}}(x_1, \dots, x_d)$$

by a fixed nondegenerate nonrandom matrix

$$\mathbf{A} = \begin{pmatrix} a_{11}, \cdots, a_{1d} \\ \cdots \\ a_{d1}, \cdots, a_{dd} \end{pmatrix}.$$

In other words, we assume that $\det(A) \neq 0$, or, equivalently, that the rows of the matrix A form a linearly independent system of vectors.

In terms of its coordinates the result of the linear transformation (8.1.9) can be written in the explicit form

$$\vec{Y} = \begin{pmatrix} a_{11}X_1 + a_{12}X_2 + \cdots + a_{1d}X_d \\ a_{21}X_1 + a_{22}X_2 + \cdots + a_{2d}X_d \\ \cdots \quad \cdots \quad \cdots \quad \cdots \\ a_{d1}X_1 + a_{d2}X_2 + \cdots + a_{dd}X_d \end{pmatrix}.$$

To calculate the probability distribution of $\vec{Y}$ following the 1D method, we must use the assumption that the matrix $\mathbf{A}$ is invertible, an analogue of the assumption $a \neq 0$ in the 1D case. Then, for a domain D in the d-dimensional space $\mathbf{R}^d$,

$$\mathbf{P}(\vec{Y} \in D) = \mathbf{P}(\mathbf{A}\vec{X} \in D) = \mathbf{P}(\vec{X} \in \mathbf{A}^{-1}D). \tag{8.1.10}$$

This identity can be rewritten in terms of p.d.f.s of $\vec{Y}$ and $\vec{X}$ as follows:

$$\int_D f_{\vec{Y}}(\vec{y})dy_1 \cdot \cdots \cdot dy_d = \int_{\mathbf{A}^{-1}D} f_{\vec{X}}(\vec{x})dx_1 \cdot \cdots \cdot dx_d.$$

Making a substitution $\vec{x} = \mathbf{A}^{-1}\vec{z}$ in the second integral, in view of the d-dimensional change of variables formula, we get that

$$\int_D f_{\vec{Y}}(\vec{y})dy_1 \cdot \dots \cdot dy_d = \int_D f_{\vec{X}}(\mathbf{A}^{-1}\vec{z}) \cdot |\det(\mathbf{A}^{-1})| dz_1 \cdot \dots \cdot dz_d,$$

where $\det(\mathbf{A}^{-1})$ is just the Jacobian of the substitution $\vec{x} = \mathbf{A}^{-1}\vec{z}$. Remembering that the determinant of the inverse matrix $\mathbf{A}^{-1}$ is the reciprocal of the determinant of the matrix $\mathbf{A}$, we get the identity

$$\int_D f_{\vec{Y}}(\vec{y})dy_1 \cdot \dots \cdot dy_d = \int_D \frac{f_{\vec{X}}(\mathbf{A}^{-1}\vec{z})}{|\det(\mathbf{A})|} dz_1 \cdot \dots \cdot dz_d.$$

Since this identity holds true for any domain D, the integrands on both sides must be equal, which gives the final formula for the p.d.f. of $\vec{Y}$:

$$f_{\vec{Y}}(\vec{y}) = \frac{f_{\vec{X}}(\mathbf{A}^{-1}\vec{y})}{|\det(\mathbf{A})|} \quad \text{if } \det(\mathbf{A}) \neq 0. \tag{8.1.11}$$

The 1D formula (8.1.3) is, obviously, the special case of the above general result.

8.2 Gaussian random vectors

As in the one-dimensional case, all nondegenerate zero-mean d-dimensional Gaussian random vectors can be obtained as nondegenerate linear transformations of a standard dD Gaussian random vector

$$\vec{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_d \end{pmatrix}$$

in which the coordinates $X_1, \dots, X_d$, are independent $N(0,1)$ random quantities. Because of their independence, the d-dimensional p.d.f. of $\vec{X}$ is the product of 1D $N(0,1)$ p.d.f.s and is thus of the product form

$$\begin{aligned} f_{\vec{X}}(\vec{x}) &= \frac{e^{-x_1^2/2}}{\sqrt{2\pi}} \cdot \dots \cdot \frac{e^{-x_d^2/2}}{\sqrt{2\pi}} \\ &= \frac{1}{(2\pi)^{d/2}} e^{-\frac{1}{2}(x_1^2+\dots+x_d^2)} = \frac{1}{(2\pi)^{d/2}} e^{-\frac{1}{2}\vec{x}^T\vec{x}}, \end{aligned} \tag{8.2.1}$$

where the superscript T denotes the transpose of a matrix. Indeed,

$$\vec{x}^T\vec{x} = (x_1, \dots, x_d) \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_d \end{pmatrix} = x_1^2 + \dots + x_d^2.$$

It is the latter form in (8.2.1) that will be useful now in applying formula (8.1.11). Indeed, substituting the last expression for $f_{\vec{X}}(\vec{x})$ in (8.2.1) into (8.1.11), one immediately gets, remembering that $(\mathbf{MN})^T = \mathbf{N}^T\mathbf{M}^T$, $(\mathbf{MN})^{-1} = \mathbf{N}^{-1}\mathbf{M}^{-1}$, and $(\mathbf{M}^T)^{-1} = (\mathbf{M}^{-1})^T$,

$$\begin{aligned} f_{\vec{Y}}(\vec{y}) &= \frac{1}{(2\pi)^{d/2}|\det(A)|} e^{-\frac{1}{2}(\mathbf{A}^{-1}\vec{y})^T\cdot(\mathbf{A}^{-1}\vec{y})} \\ &= \frac{1}{(2\pi)^{d/2}|\det(A)|} e^{-\frac{1}{2}\vec{y}^T(\mathbf{A}\mathbf{A}^T)^{-1}\vec{y}}. \end{aligned} \tag{8.2.2}$$

Thus formula (8.2.2) gives the general form of the d-dimensional zero-mean Gaussian p.d.f., and just as we identified the parameter a^2 in the 1D case (8.1.5)–(8.1.6) as the variance of the random quantity Y, we can identify entries of the matrix

$$\mathbf{\Gamma} = \mathbf{A}\mathbf{A}^T \tag{8.2.3}$$

appearing in the exponent in (8.2.2) as statistically significant parameters of the random vector $\vec{Y}$.

To see what they are, let us first calculate the entries γ_{ij}, $i, j = 1, 2, \ldots, d$, of matrix $\mathbf{\Gamma}$:

$$\gamma_{ij} = a_{i1}a_{j1} + a_{i2}a_{j2} + \cdots + a_{id}a_{jd}. \tag{8.2.4}$$

On the other hand, correlations (really, covariances, since we are working with zero-mean vectors) of different components of random vector $\vec{Y}$ are

$$\begin{aligned} \mathbf{E}(Y_iY_j) &= \mathbf{E}((a_{i1}X_1 + \cdots + a_{id}X_d)\cdot(a_{j1}X_1 + \cdots + a_{jd}X_d)) \\ &= a_{i1}a_{j1} + a_{i2}a_{j2} + \cdots + a_{id}a_{jd} \end{aligned} \tag{8.2.5}$$

because $\mathbf{E}X_iX_j = 1$ if $i = j$ and $= 0$ if $i \neq j$.

Therefore, it turns out that

$$\mathbf{\Gamma} = (\gamma_{ij}) = (\mathbf{E}Y_iY_j), \tag{8.2.6}$$

and matrix $\Gamma = (\gamma_{ij})$ is simply the correlation matrix of the general zero-mean Gaussian random vector $\vec{Y}$. Thus, since

$$\det(\mathbf{\Gamma}) = \det(\mathbf{A}\mathbf{A}^T) = \det(\mathbf{A})\cdot\det(\mathbf{A}^T) = (\det(\mathbf{A}))^2,$$

we finally get that the p.d.f. of $\vec{Y}$ can be written in the form

---

$$f_{\vec{Y}}(\vec{y}) = \frac{1}{(2\pi)^{d/2}|\det(\mathbf{\Gamma})|^{1/2}} e^{-\frac{1}{2}\vec{y}^T\mathbf{\Gamma}^{-1}\vec{y}}, \tag{8.2.7}$$

---

where $\mathbf{\Gamma}$ is the correlation matrix of $\vec{Y}$ satisfying the nondegeneracy condition $\det(\mathbf{\Gamma}) \neq 0$.

Remark 8.2.1 (*Gaussian random vectors with nonzero mean*). Of course, to get the p.d.f. of a general Gaussian random vector with nonzero expectation

$$\mathbf{E}\vec{Y} = \vec{\mu} = (\mu_1, \dots, \mu_d)^T,$$

it suffices to shift the p.d.f. (8.2.7) by $\vec{\mu}$ to obtain that

$$f_{\vec{Y}}(\vec{y}) = \frac{1}{(2\pi)^{d/2}|\det(\mathbf{\Sigma})|^{1/2}} e^{-\frac{1}{2}(\vec{y}-\vec{\mu})^T \mathbf{\Sigma}^{-1}(\vec{y}-\vec{\mu})}, \tag{8.2.8}$$

where

$$\mathbf{\Sigma} = (\sigma_{ij}) = (\mathbf{E}(Y_i - \mu_i)(Y_j - \mu_j)) \tag{8.2.9}$$

is the *covariance matrix* of $\vec{Y}$. A Gaussian random vector with joint p.d.f. given by formulas (8.2.8)–(8.2.9) is often called a normal $N(\vec{\mu}, \Sigma)$ random vector.

Example 8.2.1 (2D zero-mean Gaussian random vectors). Let us carry out the above calculation explicitly in the special case of dimension $d = 2$. Then the correlation matrix is

$$\mathbf{\Gamma} = \begin{pmatrix} \mathbf{E}Y_1Y_1 & \mathbf{E}Y_1Y_2 \\ \mathbf{E}Y_2Y_1 & \mathbf{E}Y_2Y_2 \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{pmatrix},$$

where the variances of the coordinate vectors are

$$\sigma_1^2 = \mathbf{E}Y_1^2, \qquad \sigma_2^2 = \mathbf{E}Y_2^2,$$

and the correlation coefficient of the two components is

$$\rho = \frac{\mathbf{E}Y_1Y_2}{\sigma_1\sigma_2}.$$

The determinant of the correlation matrix is

$$\det(\mathbf{\Gamma}) = \sigma_1^2\sigma_2^2(1-\rho^2),$$

and its inverse is

$$\mathbf{\Gamma}^{-1} = \frac{1}{\sigma_1^2\sigma_2^2(1-\rho^2)} \begin{pmatrix} \sigma_2^2 & -\sigma_1\sigma_2\rho \\ -\sigma_1\sigma_2\rho & \sigma_1^2 \end{pmatrix}.$$

Hence the p.d.f. of a general zero-mean Gaussian random vector is of the form

$$f_{\mathbf{Y}}(y_1, y_2) = \frac{1}{(2\pi)^{2/2}\sigma_1\sigma_2\sqrt{1-\rho^2}} \times \exp\left[-\frac{1}{2}(y_1, y_2) \frac{\begin{pmatrix} \sigma_2^2 & -\sigma_1\sigma_2\rho \\ -\sigma_1\sigma_2\rho & \sigma_1^2 \end{pmatrix}}{\sigma_1^2\sigma_2^2(1-\rho^2)} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right],$$

which, after performing prescribed matrix algebra, leads to the final expression

$$f_{\vec{\mathbf{Y}}}(y_1, y_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \times \exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{y_1^2}{\sigma_1^2} - 2\rho\frac{y_1 y_2}{\sigma_1\sigma_2} + \frac{y_2^2}{\sigma_2^2}\right)\right]. \quad (8.2.10)$$

8.3 Gaussian stationary signals

By definition, a nondegenerate zero-mean random signal $X(t)$ is Gaussian if, for any positive integer N, and any selection of sampling times $t_1 < t_2 < \cdots < t_N$, the random vector

$$\vec{X}_{(t_1,\ldots,t_N)} = \begin{pmatrix} X(t_1) \\ X(t_2) \\ \vdots \\ X(t_N) \end{pmatrix} \quad (8.3.1)$$

is a Gaussian zero-mean random vector with a nondegenerate correlation matrix. Thus, in view of results of Section 8.2, its N-dimensional joint p.d.f. $f_{(t_1,\ldots,t_N)}(x_1,\ldots,x_N)$ is given by the formula[34]

$$f_{(t_1,\ldots,t_N)}(x_1,\ldots,x_N) = \frac{1}{(2\pi)^{N/2}|\det(\mathbf{\Gamma})|^{1/2}} \cdot e^{-\frac{1}{2}\vec{x}^T\mathbf{\Gamma}^{-1}\vec{x}}, \quad \det(\mathbf{\Gamma}) \neq 0, \quad (8.3.2)$$

where $\mathbf{\Gamma}$ is the $N \times N$ correlation matrix

$$\mathbf{\Gamma} = \mathbf{\Gamma}_{(t_1,\ldots,t_N)} = (\gamma_X(t_i, t_j)) = (\mathbf{E}X(t_i)X(t_j)). \quad (8.3.3)$$

Thus, in view of (8.3.1)–(8.3.2), *the only information needed to completely determine all finite-dimensional joint probability distributions of a zero-mean Gaussian random signal $X(t)$ is the knowledge of its autocorrelation function*

$$\gamma_X(s,t) = \mathbf{E}X(t)X(s).$$

For stationary Gaussian signals the situation is simpler yet as the autocorrelation function $\gamma_X(s,t)$ is just a function of a single variable:

[34] Note that for some simple Gaussian stationary signals, like, e.g., $X(t) = X \cdot e^{jt}$, where $X \sim N(0,1)$, one can choose the t_is so that the determinant of the correlation matrix is zero; take, for example, $N = 2$ and $t_1 = \pi$, $t_2 = 2\pi$. Then the joint p.d.f. of the Gaussian random vector $(X(t_1),\ldots,X(t_N))^T$ is not of the form (8.3.2). Such signals are called degenerate.

$$\gamma_X(t,s) = \gamma_X(t-s).$$

Thus the correlation matrix $\mathbf{\Gamma}$ for a stationary random signal $X(t)$ sampled at $t_1, t_2, \dots, t_N$, is of the form

$$\mathbf{\Gamma}_{(t_1,\dots,t_N)} = \begin{pmatrix} \gamma_X(0) & \gamma_X(t_2-t_1) & \gamma_X(t_3-t_1) & \cdots & \gamma_X(t_N-t_1) \\ \cdots & \gamma_X(0) & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \gamma_X(t_N-t_1) & \gamma_X(t_N-t_2) & \cdots & \cdots & \gamma_X(0) \end{pmatrix},$$

and it is obviously invariant under translations, that is, for any t,

$$\mathbf{\Gamma}_{(t_1,\dots,t_N)} = \mathbf{\Gamma}_{(t_1+t,\dots,t_N+t)}, \tag{8.3.4}$$

which, in view of (8.3.2)–(8.3.3), implies that all finite-dimensional p.d.f.s of $X(t)$ are also invariant under translations; that is, for any positive integer N, any sampling times $t_1, \dots, t_N$, and any time shift t,

$$f_{(t_1,\dots,t_N)}(x_1,\dots,x_N) = f_{(t_1+t,\dots,t_N+t)}(x_1,\dots,x_N). \tag{8.3.5}$$

In other words,

$$\boxed{\text{a Gaussian weakly stationary signal is strictly stationary}}.$$

In the particular case when the sampling times are uniformly spaced with the intersampling time interval Δt, the correlation matrix $\mathbf{\Gamma}$ of the signal $X(t)$ sampled at times

$$t, \quad t+\Delta t, \quad t+2\Delta t, \quad \dots, \quad t+(N-1)\Delta t,$$

is

$$\mathbf{\Gamma} = \begin{pmatrix} \gamma_X(0) & \gamma_X(\Delta t) & \gamma_X(2\Delta t) & \cdots & \gamma_X((N-1)\Delta t) \\ \gamma_X(\Delta t) & \gamma_X(0) & \gamma_X(\Delta t) & \cdots & \gamma_X((N-2)\Delta t) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \gamma_X((N-1)\Delta t) & \gamma_X((N-2)\Delta t) & \cdots & \cdots & \gamma_X(0) \end{pmatrix}.$$

Example 8.3.1. Consider a Gaussian signal $X(t)$ with autocorrelation function

$$\gamma_X(t) = e^{-0.3|t|}.$$

We are interested in finding the joint p.d.f. of the signal at times $t_1 = 1$, $t_2 = 2$, and the probability that the signal has values between -0.6 and 1.4 at t_1 and between 0.7 and 2.6 at t_2.

The first step is then to find the correlation matrix

$$\mathbf{\Gamma}_{(1,2)} = \begin{pmatrix} \gamma_X(0) & \gamma_X(1) \\ \gamma_X(1) & \gamma_X(0) \end{pmatrix} = \begin{pmatrix} e^0 & e^{-0.3} \\ e^{-0.3} & e^0 \end{pmatrix} = \begin{pmatrix} 1 & 0.74 \\ 0.74 & 1 \end{pmatrix}.$$

The correlation coefficient of $X(1)$ and $X(2)$ is then

$$\rho = \frac{\gamma_X(1)}{\gamma_X(0)} = 0.74$$

and, in view of Example 8.2.1 (see (8.2.10)), the joint p.d.f. of $X(1)$ and $X(2)$ is of the form

$$\begin{aligned} f_{(1,2)}(x_1, x_2) &= \frac{1}{2\pi\sqrt{1-0.74^2}} \\ &\quad \cdot \exp\left[\frac{-1}{2(1-0.74^2)}(x_1^2 - 2\cdot 0.74 x_1 x_2 + x_2^2)\right] \\ &= 0.24 \cdot \exp[-1.11(x_1^2 - 1.48x_1x_2 + x_2^2)]. \end{aligned}$$

Finally, the desired probability is

$$\begin{aligned} &\mathbf{P}(-0.6 \le X(1) \le 1.4 \text{ and } 0.7 \le X(2) \le 2.6) \\ &\quad = \int_{-0.6}^{1.4}\int_{0.7}^{2.6} 0.24 \cdot e^{-1.11(x_1^2-1.48x_1x_2+x_2^2)}\, dx_1 dx_2 = 0.17, \end{aligned}$$

where the last integral has been evaluated numerically in *Mathematica* with a two-digit precision.

8.4 Sample path properties of general and Gaussian stationary signals

Mean-square continuity and differentiability. It is clear that the local properties of the autocorrelation function $\gamma_X(\tau)$ of a stationary signal $X(t)$ affect properties of the sample paths of the signal itself in the mean-square sense, that is in terms of the behavior of the expectation of the square of signal's increments, i.e., the variances of the increments.[35] Indeed, with no distributional assumptions on $X(t)$, we have

$$\sigma^2(\tau) = \mathbf{E}(X(t+\tau) - X(t))^2 = 2(\gamma_X(0) - \gamma_X(\tau));$$

the variance of the increment is independent of t. Hence we have the following result:

A stationary signal $X(t)$ is continuous in the mean-square sense, that is, for any $t > 0$,

$$\lim_{\tau\to 0} \mathbf{E}(X(t+\tau) - X(t))^2 = 0,$$

if and only if the autocorrelation function $\gamma_X(\tau)$ is continuous at $\tau = 0$,

[35] Recall that the sequence (X_n) of random quantities is said to converge to X, in the mean-square, if $\mathbf{E}|X_n - X|^2 \to 0$, as $n \to \infty$.

that is,

$$\lim_{\tau\to 0} \gamma_X(\tau) = \gamma_X(0).$$

In particular, signals with autocorrelation functions $\gamma_X(\tau) = e^{|\tau|}$ or $\gamma_X(\tau) = \frac{1}{1+\tau^2}$ are mean-square continuous.

A similar, mean-square analysis of the limit at $\tau = 0$ of the differential ratio,

$$\mathbf{E}\left(\frac{X(t+\tau)-X(t)}{\tau}\right)^2 = 2\frac{\gamma_X(0)-\gamma_X(\tau)}{\tau^2},$$

shows that a stationary signal with autocorrelation function $\gamma_X(\tau) = e^{|\tau|}$ cannot be possibly mean-square differentiable because in this case

$$\lim_{\tau\to 0}\frac{\gamma_X(0)-\gamma_X(\tau)}{\tau^2} = \lim_{\tau\to 0}\frac{1-e^{-|\tau|}}{\tau^2} = \infty,$$

whereas the differentiability cannot be excluded for the signal with autocorrelation $\gamma_X(\tau) = \frac{1}{1+\tau^2}$ because in this case

$$\lim_{\tau\to 0}\frac{\gamma_X(0)-\gamma_X(\tau)}{\tau^2} = \lim_{\tau\to 0}\frac{1-\frac{1}{1+\tau^2}}{\tau^2} = 1.$$

Of course, the above brief discussion just verifies the boundedness of the variance of the signal's differential ratio as $\tau \to 0$, not whether the latter has a limit. So, let us take a closer look at the issue of the mean-square differentiability of a stationary signal, that is the existence of the random quantity $X'(t)$, for a fixed t. First, observe that this existence is equivalent to the statement that[36]

$$\lim_{\tau_1\to 0}\lim_{\tau_2\to 0}\mathbf{E}\left(\frac{X(t+\tau_1)-X(t)}{\tau_1}-\frac{X(t+\tau_2)-X(t)}{\tau_2}\right)^2 = 0$$

But the expression under the limit signs is equal to

$$\mathbf{E}\left(\frac{X(t+\tau_1)-X(t)}{\tau_1}\right)^2 + \mathbf{E}\left(\frac{X(t+\tau_2)-X(t)}{\tau_2}\right)^2 - 2\mathbf{E}\left(\frac{X(t+\tau_1)-X(t)}{\tau_1}\cdot\frac{X(t+\tau_2)-X(t)}{\tau_2}\right).$$

So, the existence of the derivative $X'(t)$ in the mean-square is equivalent to the fact that the first two terms converge to $\gamma_{X'}(0)$ and the third to

[36] This argument relies on the so-called Cauchy criterion of convergence for random quantities with finite variance: *A sequence X_n converges in the mean-square as $n \to \infty$, that is, there exists a random quantity X such that* $\lim_{n\to\infty}\mathbf{E}(X_n - X)^2 = 0$, *if and only if* $\lim_{n\to\infty}\lim_{m\to\infty}\mathbf{E}(X_n - X_m)^2 = 0$. This criterion permits the verification of the convergence without knowing what the limit is; see, e.g., Theorem 11.4.2 in W. Rudin, *Principles of Mathematical Analysis*, McGraw–Hill, New York, 1976.

$-2\gamma_{X'}(0)$. But the convergence of the last term means the existence of the limit

$$\begin{aligned}
&\lim_{\tau_1\to 0}\lim_{\tau_2\to 0}\frac{1}{\tau_1\tau_2}\mathbf{E}((X(t+\tau_1)-X(t))\cdot(X(t+\tau_2)-X(t)))\\
&\quad=\lim_{\tau_1\to 0}\lim_{\tau_2\to 0}\frac{1}{\tau_1\tau_2}(\gamma_X(\tau_2-\tau_1)-\gamma_X(\tau_1)-\gamma_X(\tau_2)+\gamma_X(0))\\
&\quad=\lim_{\tau_1\to 0}\lim_{\tau_2\to 0}\frac{1}{\tau_1\tau_2}\Delta_{-\tau_1}\Delta_{\tau_2}\gamma_X(0),
\end{aligned}$$

where $\Delta_\tau f(t) := f(t+\tau)-f(t)$ is the usual difference operator. Indeed,

$$\begin{aligned}
\Delta_{-\tau_1}\Delta_{\tau_2}\gamma_X(0) &= \Delta_{\tau_1}(\gamma_X(\tau_2)-\gamma_X(0))\\
&= (\gamma_X(\tau_2-\tau_1)-\gamma_X(-\tau_1))-(\gamma_X(\tau_2)-\gamma_X(0)).
\end{aligned}$$

Since the existence of the last limit appearing above means twice differentiability of the autocorrelation function of X at $\tau = 0$ we arrive at the following criterion:

A stationary signal $X(t)$ is mean-square differentiable if and only if its autocorrelation function $\gamma_X(\tau)$ is twice differentiable at $\tau = 0$. Moreover, in this case, the cross-correlation of the signal $X(t)$ and its derivative $X'(t)$ is

$$\mathbf{E}X(t)X'(s) = \lim_{\tau\to 0}\frac{\gamma_X(t+\tau-s)-\gamma_X(t-s)}{\tau} = \frac{\partial}{\partial t}\gamma_X(t-s), \tag{8.4.1}$$

and the autocorrelation of the derivative signal is

$$\mathbf{E}X'(t)X'(s) = \lim_{\tau\to 0}\left(\frac{\partial}{\partial t}\gamma_X(t+\tau-s)-\frac{\partial}{\partial t}\gamma_X(t-s)\right) = \frac{\partial^2}{\partial t\partial s}\gamma_X(t-s). \tag{8.4.2}$$

In a similar fashion one can calculate the cross-correlation of higher derivatives of the signal $X(t)$ to obtain that[37]

$$\mathbf{E}X^{(n)}(t)X^{(m)}(s) = \frac{\partial^{n+m}}{\partial t^n\partial s^m}\gamma_X(t-s). \tag{8.4.3}$$

Sample path continuity. A study of properties of the individual sample paths (trajectories, realizations) of stationary random signals is a more delicate matter, with the most precise results obtainable only in the case of Gaussian signals. Indeed, we have observed in the previous sections that, for a Gaussian signal, the autocorrelation function determines all the finite-dimensional probability distributions of the signal, meaning

[37] For details, see M. Loeve, *Probability Theory*, Van Nostrand, Princeton, NJ, 1963, Section 34.3.

that for any finite sequence of windows, $[a_1, b_1], [a_2, b_2], \ldots, [a_N, b_N]$, and any collection of time instants $t_1, t_2, \ldots, t_N$, we can find the probability that the signal fits into those windows at prescribed times, that is,

$$\mathbf{P}(a_1 < t_1 < b_1,\ a_2 < t_2 < b_2, \ldots,\ a_N < t_N < b_N).$$

So it seems that by taking N to ∞, and making the time instants closer to each other and the windows narrower, one could find the probability that the signal's sample path has any specific shape or property. This idea is, roughly speaking, correct, but only in a subtle sense that will be explained below.

The discussion of the sample path properties of stationary signals will be based here on the following theorem of the theory of general random signals (stochastic processes) due to N. N. Kolmogorov.

Theorem 8.4.1. *Let $g(h)$ be an even function, nondecreasing for $h > 0$, and such that $g(h) \to 0$ as $h \to 0$. Furthermore, suppose that $X(t)$ is a random signal such that*

$$\mathbf{P}(|X(t+h) - X(t)| > g(h)) \le q(h), \tag{8.4.4}$$

for a function $q(h)$ satisfying the following three conditions:

$$q(h) \to 0 \quad \text{as } h \to 0; \tag{8.4.5}$$

$$\sum_{n=1}^{\infty} 2^n q(2^{-n}) < \infty; \tag{8.4.6}$$

$$\sum_{n=1}^{\infty} g(2^{-n}) < \infty; \tag{8.4.7}$$

Then, with probability 1, the sample paths of the signal $X(t)$ are continuous.

Although the proof of the above theorem is beyond the scope of this book,[38] the intuitive meaning of the assumptions (8.4.4)–(8.4.7) is clear: for the signal to have continuous sample paths, the increments of the signal over small time intervals can be permitted to be large only with a very small probability.

Applied to the second-order (not necessarily stationary) signals, Theorem 8.4.1 immediately gives the following.

[38] For a more complete discussion of this theorem and its consequences for sample path continuity and differentiability of random signals, see, for example, M. Loève, *Probability Theory*, Van Nostrand, Princeton, NJ, 1963, Section 35.3.

Corollary 8.4.1. *If there exists a* τ_0 *such that, for all* τ, $0 \le \tau < \tau_0$, *and all* t *in a finite time interval,*

$$\mathbf{E}(X(t+\tau) - X(t))^2 \le C|\tau|^{1+\epsilon} \tag{8.4.8}$$

for some constants C, $\epsilon > 0$, *then the sample paths of the signal* $X(t)$ *are continuous with probability* 1.

To see how Corollary 8.4.1 follows from Theorem 8.4.1,[39] observe first that for any random quantity Z and any constant $a > 0$.

$$\mathbf{P}(Z > a) \le \int_a^\infty f_Z(z)dz \le \int_a^\infty \frac{z^2}{a^2} f_Z(z)dz \le \frac{\mathbf{E}Z^2}{a^2}.$$

Condition (8.4.8) then implies that

$$\mathbf{P}(X(t+\tau) - X(t)| > g(\tau)) \le \frac{C|\tau|^{1+\epsilon}}{g^2(\tau)},$$

so that by selecting $g(\tau) = |\tau|^{\epsilon/4}$, and

$$q(\tau) = \frac{C|\tau|^{1+\epsilon}}{g^2(\tau)} = C|\tau|^{1+\epsilon/2},$$

we easily see that $g(\tau)$ and $q(\tau)$ are continuous functions vanishing at $\tau = 0$, and that the conditions (8.4.6)–(8.4.7) of the theorem are also satisfied. Indeed,

$$\sum_{n=1}^{\infty} 2^n q(2^{-n}) = C \sum_{n=1}^{\infty} 2^n (2^{-n})^{1+\epsilon/2} = C \sum_{n=1}^{\infty} 2^{-n\epsilon/2} < \infty,$$

and

$$\sum_{n=1}^{\infty} g(2^{-n}) = \sum_{n=1}^{\infty} 2^{-n\epsilon/4} < \infty.$$

In the special case of a stationary signal we have $\mathbf{E}(X(t+\tau) - X(t))^2 = 2(\gamma_X(0) - \gamma_X(\tau))$, so the sample path continuity is guaranteed by the following condition on the autocorrelation function:

$$|\gamma_X(0) - \gamma_X(\tau)| \le C|\tau|^{1+\epsilon}, \tag{8.4.9}$$

[39] This inequality is known as the Chebyshev inequality, and its proof here has been carried out only in the case of absolutely continuous probability distributions. The proof in the discrete case is left to the reader as an exercise; see Section 8.5.

for some constant $\epsilon > O$, and small enough τ.

In particular, for the autocorrelation function $\gamma_X(\tau) = \frac{1}{1+\tau^2}$,

$$|\gamma_X(0) - \gamma_X(\tau)| = 1 - \frac{1}{1+\tau^2} = \frac{\tau^2}{1+\tau^2} \le \tau^2,$$

and the condition (8.4.8) is satisfied, thus giving the sample path continuity.

However, for a signal with autocorrelation function $\gamma_X(\tau) = e^{-|\tau|}$, the difference $\gamma_X(0) - \gamma_X(\tau)$ behaves asymptotically like τ for $\tau \to 0$. Therefore, there is no positive ϵ for which condition (8.4.9) is satisfied and we cannot claim the continuity of the sample path in this case—not a surprising result if one remembers that the exponential autocorrelation was first encountered in the context of the obviously sample path discontinuous switching signal. Nevertheless, as we observed at the beginning of this section, a signal with an exponential autocorrelation is mean-square continuous.

For a Gaussian stationary signal $X(t)$, Theorem 8.4.1 can be applied in a more precise fashion since the probabilities $\mathbf{P}(X(t+\tau) - X(t) > a)$ are known exactly. Indeed, since for any positive z,

$$\int_z^\infty e^{-x^2/2} dx \le \int_z^\infty \frac{x}{z} e^{-x^2/2} dx = \frac{1}{z} e^{-z^2/2},$$

because $\frac{x}{z} \ge 1$ in the interval of integration, we have, for any nonnegative function $g(\tau)$ and positive constant C,

$$\mathbf{P}(|X(t+\tau) - X(t)| > Cg(\tau)) \le \sqrt{\frac{2}{\pi}} \frac{\sigma(\tau)}{Cg(\tau)} \exp\left(-\frac{1}{2}\frac{C^2 g^2(\tau)}{\sigma^2(\tau)}\right), \tag{8.4.10}$$

where $\sigma^2(\tau) = \mathbf{E}(X(t+\tau) - X(t))^2 = 2(\gamma_X(0) - \gamma_X(\tau))$. This estimate yields the following result.

Corollary 8.4.2. *If there exists τ_0 such that, for all τ, $0 \le \tau \le \tau_0$, the autocorrelation function $\gamma_X(\tau)$ of a stationary Gaussian signal $X(t)$ satisfies the condition*

$$\gamma_X(0) - \gamma_X(\tau) \le \frac{K}{|\ln|\tau||^\delta}, \tag{8.4.11}$$

for some constants $K > 0$ and $\delta > 3$, then the signal $X(t)$ has continuous sample paths with probability 1.

The proof of the corollary is completed by selecting

$$g(\tau) = |\ln|\tau||^{-\nu},$$

with any number ν satisfying condition $1 < \nu < \frac{\delta-1}{2}$, choosing

$$q(C,\tau) = \frac{K'}{C|\ln|\tau||^{\delta/2-\nu}} \exp\left(-\frac{C^2}{2K}|\ln|\tau||^{\delta-2\nu}\right)$$

and verifying the convergence of the two series in conditions (8.4.6)–(8.4.7); see the exercise in Section 8.5.

Returning to the case of a stationary random signal with an exponential autocorrelation function, we see that if the signal is Gaussian, then Corollary 8.4.2 guarantees the continuity of its sample paths with probability 1. Indeed, condition (8.4.11) is obviously satisfied since (e.g., picking $\delta = 4$) we have

$$\lim_{\tau\to 0}(\gamma_X(0) - \gamma_X(\tau)) \cdot |\ln|\tau||^4 = \lim_{\tau\to 0}(1 - e^{-|\tau|}) \cdot |\ln|\tau||^4 = 0$$

in view of l'Hospital's rule.

8.5 Problems and exercises

8.5.1. A zero-mean Gaussian random signal has the autocorrelation function of the form

$$\gamma_X(\tau) = 10e^{-0.1|\tau|}\cos 2\pi\tau.$$

Write the covariance matrix for the signal sampled at four time instants separated by 0.5 seconds.

8.5.2. Find the joint p.d.f. of the signal from Exercise 8.5.1 at $t_1 = 1$ and $t_2 = 2.5$. Write the integral formula for

$$P(0 \le X(1) \le 1, 0 \le X(2.5) \le 2).$$

Evaluate the above probability numerically.

8.5.3. Find the joint p.d.f. of the signal from Exercise 8.5.1 at $t_1 = 1$, $t_2 = 1.5$, and $t_3 = 2.5$. Write the integral formula for

$$P(0 \le X(1) \le 1, -1 \le X(1.5) \le 3, 0 \le X(2.5) \le 2).$$

Evaluate the above probability numerically.

8.5.4. Show that if a 2D Gaussian random vector $\vec{Y} = (Y_1, Y_2)$ has uncorrelated components Y_1, Y_2, then those components are statistically independent random quantities.

8.5.5. Produce 2D surface plots for p.d.f.s of three Gaussian random vectors: $(X(1.0), X(1.1))^T$, $(X(1.0), X(2.0))^T$, $(X(1.0), X(5.0))^T$, where $X(t)$ is the stationary signal described in Example 8.3.1. Comment on similarities and differences in the three plots.

8.5.6. Prove that if there exists a τ_0 such that, for all $\tau < \tau_0$ and all t in a finite time interval,

$$\mathbf{E}(X(t+\tau) - X(t))^2 \leq C\frac{|\tau|}{|\ln|\tau||^{1+\delta}},$$

for some $C > 0$ and $\delta > 2$, then the sample paths of the signal $X(t)$ are continuous with probability 1. *Hint*: This result is a little more delicate than Corollary 8.4.1, but the idea of the proof is similar: take $g(\tau) = |\ln|\tau||^{-\beta}$, for a β between 1 and $\frac{\delta}{2}$, from which we have

$$q(\tau) = \frac{|\tau|}{|\ln|\tau||^{1+\delta-2\beta}},$$

and check conditions (8.4.4)–(8.4.7) in Theorem 8.4.1.

8.5.7. Verify the Chebyshev inequality, $\mathbf{P}(|Z| > a) \leq \frac{\mathbf{E}Z^2}{a^2}$, $a > 0$, for a discrete random quantity Z.

8.5.8. Produce plots of several 2D Gaussian densities with selected means and covariance matrices.

8.5.9. Random signal $X(t)$ has an autocorrelation function of the form $\gamma_X(\tau) = \exp(-|\tau|^\alpha)$ with $0 < \alpha \leq 2$. For which values of parameter α can you claim the continuity of sample paths of $X(t)$ with probability 1?

8.5.10. Verify formula (8.4.3) for the cross-correlation of higher derivatives of a stationary signal.

8.5.11. Verify the convergence of the series (8.4.6)–(8.4.7) in the proof of Corollary 8.4.2.