Probability and Stochastic Processes: Chapter VI: Stationary Stochastic Process

Zhenyu Liao

School of Electronic Information and Communications, HUST

December 13, 2022

Outline

Outline

1. Discrete-time random process and linear systems

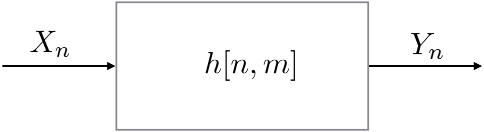
2. Continuous-time random processes and linear systems

Discrete-time linear systems

- Consider the following discrete-time linear (possibly time-varying) system with impulse response $h[n,m]$, and the input is the random process X_n.
- The output Y_n of such system is given by

$$Y_n = \sum_m h[n,m] X_{n-m}$$

and we would like to check whether the sum in (defining) Y_n exists?



Existence and WSS of the output

Theorem

Let $\{X_n : n \in \mathbb{Z}\}$ be a discrete-time random process with auto-correlation function $r_{xx}[n,m] = \mathbb{E}[X_n X_m^*]$. Let $h[n,m]$ denote the the impulse response of a discrete-time linear system with input-output relation given by $y[n] = \sum_m h[n,m] x[n-m], n \in \mathbb{Z}$. Then,

- *the output*

$$Y_n = \sum_m h[n,m] X_{n-m},$$

(for any fixed $n \in \mathbb{Z}$) exists in the mean-square (m.s.) sense if

$$\sum_{m \in \mathbb{Z}} |h[n,m]| \sqrt{r_{xx}[n-m, n-m]} < \infty.$$

- *if X_n is Wide-Sense Stationary (WSS) and the system is linear time-invariant (LTI) and BIBO-stable, then Y_n always exists in the m.s. sense and it is also WSS.*

Proof: For any given $n \in \mathbb{Z}$, define the random process $\{Z_{n,m} = h[n,m] X_{n-m} : m \in \mathbb{Z}\}$, so that $Y_n = \sum_m Z_{n,m}$. Then, the auto-correlation function of $Z_{n,m}$ (with respect to m for fixed n) is

$$r_{zz}[\ell, k] = \mathbb{E}\left[h[n,\ell] X_{n-\ell} h^*[n,k] X_{n-k}^*\right] = h[n,\ell] h^*[n,k] r_{xx}[n-\ell, n-k]$$

Proof: continuation

It is known that for a discrete-time process $\{X_n : n \in \mathbb{Z}\}$ with auto-correlation function $r_{xx}[n,m]$, then the sum $Y_n = \sum_{i=0}^{\infty} X_i$ converges in a m.s. sense

1. *if and only if* $\lim_{n,m\to\infty} \sum_{i=n+1}^{m} \sum_{j=n+1}^{m} r_{xx}[i,j] = 0$ (**Cauchy convergence**);
2. *if and only if* $\lim_{n,m\to\infty} \sum_{i=1}^{n} \sum_{j=1}^{m} r_{xx}[i,j] = r < \infty$ (**Loeve's criterion**);
3. *if* $\sum_{n=1}^{\infty} \sqrt{r_{xx}[n,n]} < \infty$ (**Cauchy–Schwartz**).

So, by Item 3, we have the convergence $Y_n = \sum_m Z_{n,m}$ if

$$\sum_{m\in\mathbb{Z}} \sqrt{r_{zz}[m,m]} = \sum_{m\in\mathbb{Z}} |h[n,m]| \sqrt{r_{xx}[n-m,n-m]} < \infty.$$

This proves the first statement.

Then, by definition of WSS (for X_n), we have $r_{xx}[n,m] = r_{xx}[n-m]$; and by definition of LTI system, we have $h[n,m] = h[m]$ (**independent** of n), so that

$$Y_n = \sum_m h[n,m] X_{n-m} = \sum_m h[m] X_{n-m} \qquad \text{discrete-time convolution}$$

Furthermore, by definition of BIBO-stable, we have $h[m]$ absolutely summable and

$$\sum_{m\in\mathbb{Z}} \sqrt{r_{zz}[m,m]} = \sum_{m\in\mathbb{Z}} |h[n,m]| \sqrt{r_{xx}[n-m,n-m]} = \sqrt{r_{xx}[0]} \sum_m |h[m]| < \infty.$$

And it remains to check Y_n is WSS to conclude the proof of the second statement.

Input-output second-order statistics

- Consider $Y_n = \sum_m h[n,m] X_{n-m}$ and assume that it exists in the m.s. sense. Then,

$$\begin{aligned}
\mu_y[n] &= \sum_m h[n,m]\mu_x[n-m] \\
r_{yy}[n,m] &= \sum_\ell \sum_k h[n,\ell] r_{xx}[n-\ell, m-k] h^*[m,k] \\
r_{yx}[n,m] &= \sum_\ell h[n,\ell] r_{xx}[n-\ell, m] \\
r_{xy}[n,m] &= \sum_k r_{xx}[n, m-k] h^*[m,k]
\end{aligned}$$

- Suppose that the transformation is LTI, BIBO stable, and in the input is WSS, then

$$Y_n = \sum_m h[m] X_{n-m}$$

- Constant mean function of the output:

$$\mu_y = \sum_m h[m]\mu_x = \mu_x \sum_m h[m]$$

- Auto-correlation function of the output

$$r_{yy}[m] = \mathbb{E}[Y_n Y_{n-m}^*] = \sum_\ell \sum_k h[\ell] r_{xx}(m-\ell+k) h^*[k] = h[m] \otimes h^*[-m] \otimes r_{xx}[m]$$

- Output-input cross-correlation function

$$r_{yx}[m] = \mathbb{E}[Y_n X_{n-m}^*] = \sum_{\ell} h[\ell] r_{xx}[m-\ell] = h[m] \otimes r_{xx}[m]$$

- Input-output cross-correlation function

$$r_{xy}[m] = \mathbb{E}[X_n Y_{n-m}^*] = \sum_{\ell} h^*[-\ell] r_{xx}[m-\ell] = h^*[-m] \otimes r_{xx}[m]$$

- Notice also:

$$r_{yy}[m] = h[m] \otimes r_{xy}[m] = h^*[-m] \otimes r_{yx}[m]$$

- Discrete-time signals ($\mathcal{I} = \mathbb{Z}$):

$$\check{x}(f) = \sum_n x_n e^{-\jmath 2\pi f n} \qquad \text{dt-Fourier transform}$$

$$x_n = \int_{-1/2}^{1/2} \check{x}(f) e^{\jmath 2\pi f n} df \qquad \text{representation}$$

- Discrete-time finite duration signals (or periodic signals) ($\mathcal{I} = \mathbb{Z}_N$):

$$\check{x}_k = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x_n e^{-\jmath \frac{2\pi}{N} nk} \qquad \text{DFT}$$

$$x_n = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \check{x}_k e^{\jmath \frac{2\pi}{N} nk} \qquad \text{representation}$$

Meaning of the equalities

- In all above relations, the $=$ sign establishes a correspondence between x and its transform $\check{x}$.
- If we restrict to $\mathcal{L}_2$ and ℓ_2 functions and sequences, the correspondence is one-to-one in the Hilbert spaces of squared summable functions and sequences with inner product

$$\langle x, y \rangle = \int x(t) y^*(t) dt, \qquad \langle x, y \rangle = \sum_n x_n y_n^*$$

- Parseval identity ensures that the Fourier transform operator maps $\mathcal{L}_2$ (resp., ℓ_2) into $\mathcal{L}_2$ (resp., ℓ_2), in particular, in the discrete-time case we have:

$$\langle x, x \rangle = \sum_n |x_n|^2 = \int |\check{x}(f)|^2 df$$

Energy spectral density

- Consider a random process X_n with auto-correlation function $r_{xx}[n,m]$ such that $\sum_{n,m} r_{xx}[n,m] < \infty$ (Notice: this process is generally NOT WSS).
- Its Fourier transform exists in the m.s. sense

$$\check{X}(f) = \sum_n X_n e^{-\jmath 2\pi f n}$$

- The auto-correlation function of $\check{X}(f)$ is defined as

$$r_{\check{x}\check{x}}(f_1, f_2) = \mathbb{E}[\check{X}(f_1)\check{X}^*(f_2)] = \sum_n \sum_m r_{xx}[n,m] e^{-\jmath 2\pi (f_1 n - f_2 m)}$$

- In electrical engineering, the second moment $\mathbb{E}[|X_n|^2] = r_{xx}[n,n]$ represents the (ensemble average) energy per sample of the process.
- Summed over all n, the total (ensemble average) energy is

$$\mathcal{E}_x = \mathbb{E}\left[\sum_n |X_n|^2\right] = \sum_n r_{xx}[n,n]$$

- By using Parseval's identity, we have

$$\mathcal{E}_x = \mathbb{E}\left[\sum_n |X_n|^2\right] = \mathbb{E}\left[\int_{-1/2}^{1/2} |\check{X}(f)|^2 df\right] = \int r_{\check{x}\check{x}}(f,f) df$$

- We define the energy spectral density (ESD) function as

$$E_x(f) = \mathbb{E}[|\check{X}(f)|^2] = r_{\check{x}\check{x}}(f,f)$$

- The ESD is non-negative real, and when integrated over $f \in [-1/2, 1/2]$ yields the process average energy.
- The quantity $E_x(f)df$ can be interpreted as the average amount of energy that the process allocates to its frequency component at frequency f.

ESD and LTI systems

- Consider $Y_n = \sum_m h[m] X_{n-m}$ where $h[m]$ is the impulse response of a BIBO-stable LTI system. Then

$$\boxed{\check{Y}(f) = \check{h}(f)\check{X}(f)}$$

- Direct calculation shows immediately that

$$r_{\check{y}\check{y}}(f_1, f_2) = \check{h}(f_1)\check{h}^*(f_2) r_{\check{x}\check{x}}(f_1, f_2)$$

- It follows that

$$\boxed{E_y(f) = r_{\check{y}\check{y}}(f, f) = \left|\check{h}(f)\right|^2 E_x(f)}$$

- In plain words, the LTI system acts on the energy density of the process by re-weighting the frequency components by the squared magnitude of the system transfer function $|\check{h}(f)|^2$.

Power spectral density

- Consider a random process X_n such that $\sum_{n,m} r_{xx}[n,m]$ may not converge, but such that its truncation to the finite support $[-N,N]$ is finite for any finite N.
- Define

$$\check{X}_N(f) = \sum_{n=-N}^{N} X_n e^{-\jmath 2\pi f n}$$

and the corresponding ESD $E_x^{(N)}(f) = \mathbb{E}[|\check{X}_N(f)|^2]$, with $\int E_x^{(N)}(f) df = \mathcal{E}_x^{(N)}$

- The power of X_n is defined as the average energy per index/symbol/time

$$\mathcal{P}_x = \lim_{N\to\infty} \frac{1}{2N+1} \mathcal{E}_x^{(N)} = \lim_{N\to\infty} \frac{1}{2N+1} \int_{-1/2}^{1/2} E_x^{(N)}(f) df$$

when this limit exists.

- Assuming that we can exchange limit and integration (e.g., with Lebesgue's dominated convergence theorem), we define the power spectral density (PSD) of X_n as

$$P_x(f) = \lim_{N\to\infty} \frac{E_x^{(N)}(f)}{2N+1} = \lim_{N\to\infty} \frac{1}{2N+1} \mathbb{E}\left[\left| \check{X}_N(f) \right|^2 \right]$$

- It follows that, if the PSD exists (exchanging limit with integration is valid), then the signal power is given by

$$\mathcal{P}_x = \int_{-1/2}^{1/2} P_x(f) df$$

Wiener-Khintchine Theorem

In plain words, auto-correlation function of a wide-sense-stationary (WSS) random process has a spectral decomposition given by the power spectrum of that process.

Theorem

If X_n is WSS with absolutely summable auto-correlation function $r_{xx}[m]$, then

$$P_x(f) = \sum_m r_{xx}[m] e^{-\jmath 2\pi f m}$$

i.e., the PSD is the Fourier transform of the auto-correlation function.

- $P_x(f)$ is real non-negative valued since $r_{xx}[m]$ is Hermitian symmetric and positive semi-definite.
- If X_n is real-valued, then $P_x(f)$ is an even function (i.e., $P_x(f) = P_x(-f)$).
- By the inverse Fourier transform, we have

$$r_{xx}[m] = \int_{-1/2}^{1/2} P_x(f) e^{\jmath 2\pi f m} df$$

Proof of Wiener-Khintchine Theorem

- We write

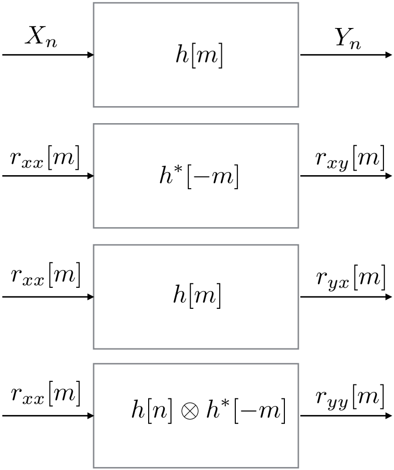
$$\begin{aligned}\frac{1}{2N+1}\mathbb{E}\left[\left|\check{X}_N(f)\right|^2\right] &= \frac{1}{2N+1}\mathbb{E}\left[\left|\sum_{n=-N}^{N} X_n e^{-\jmath 2\pi f n}\right|^2\right] \\ &= \frac{1}{2N+1}\sum_{n=-N}^{N}\sum_{m=-N}^{N} r_{xx}[n-m]e^{-\jmath 2\pi f(n-m)} \\ &= \sum_{\ell=-2N}^{2N} r_{xx}[\ell]\left(1-\frac{|\ell|}{2N+1}\right)e^{-\jmath 2\pi f \ell}\end{aligned}$$

  where the last equality follows from the fact that $r_{xx}[n-m]$ is constant for each diagonal summation path $n-m=\ell$ in the rectangle $[-N,N]\times[-N,N]$.
- If $r_{xx}[m]$ is absolutely summable, we can let $N\to\infty$ and have the result.

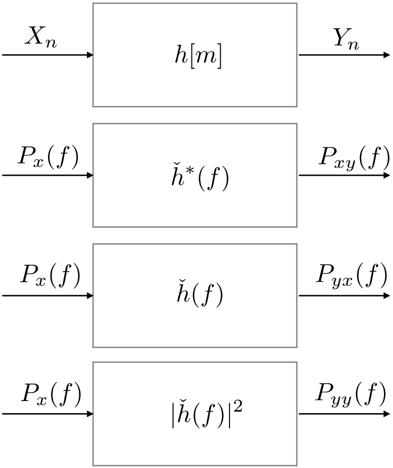
- Let X_n, Y_n be the WSS input and output of a LTI BIBO-stable system with impulse response $h[m]$.
- Recalling the convolution relation between $r_{xx}[m], r_{yx}[m]$ and $r_{yy}[m]$, we have

$$
\begin{aligned}
r_{yy}[m] = h[m] \otimes h^*[-m] \otimes r_{xx}[m] &\quad \Leftrightarrow \quad P_y(f) = |\check{h}(f)|^2 P_x(f) \\
r_{yx}[m] = h[m] \otimes r_{xx}[m] &\quad \Leftrightarrow \quad P_{yx}(f) = \check{h}(f) P_x(f) \\
r_{xy}[m] = h^*[-m] \otimes r_{xx}[m] &\quad \Leftrightarrow \quad P_{xy}(f) = \check{h}^*(f) P_x(f) \\
r_{yy}[m] = h[m] \otimes r_{xy}[m] &\quad \Leftrightarrow \quad P_y(f) = \check{h}(f) P_{xy}(f)
\end{aligned}
$$

Input-output relations for correlation



Input-output relations for PSD



PSD of more general processes

- For a wide class of processes, the ensemble-averaged power

$$\mathcal{P}_x = \lim_{N\to\infty} \frac{1}{2N+1} \mathbb{E}\left[\sum_{n=-N}^{N} |X_n|^2 \right] = \lim_{N\to\infty} \frac{1}{2N+1} \mathbb{E}\left[\int_{-1/2}^{1/2} |\check{X}_N(f)|^2 df \right]$$

exists.
- In addition, it happens that $\lim_{N\to\infty} \frac{1}{2N+1} \mathbb{E}[|\check{X}_N(f)|^2] = P_x(f)$ exists for all $f \in \mathbb{R}$ and that $\mathcal{P}_x = \int_{-1/2}^{1/2} P_x(f) df$.
- In this case, we wish to calculate the PSD $P_x(f)$ even though X_n is not WSS.

Theorem

For sufficiently well-behaved $r_{xx}[n,m]$*,*

$$P_x(f) = \sum_m \bar{r}_{xx}[m] e^{-\jmath 2\pi f m}$$

where

$$\bar{r}_{xx}[m] = \lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} r_{xx}[n, n-m]$$

is the time-averaged auto-correlation function.

Corollary

The family of WSC processes has mean function $\mu_x[n]$ *periodic with period* T *and auto-correlation function* $r_{xx}[n, n-m]$ *periodic with respect to* n *with period* T *for all* m*. In this case, the time-averaged auto-correlation function is easily obtained by averaging over one period:*

$$\bar{r}_{xx}[m] = \frac{1}{T} \sum_{n=0}^{T-1} r_{xx}[n, n-m]$$

Proof: as in the proof of the Wiener-Khintchine Theorem.

A reminder on continuous-time random processes

- Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and an interval $\mathcal{T} \subset \mathbb{R}$, a continuous-time random process is the collection of random variables $\{X(\omega, t) : t \in \mathcal{T}\}$ where for any n and n-tuple of indices $t_1, \ldots, t_n \in \mathcal{T}$ we have that $(X(\cdot, t_1), \cdots, X(\cdot, t_n)) : \Omega \to \mathbb{R}^n$ is a random vector with respect to the given probability space.
- As usual, we generally neglect the explicit dependence on ω, and write $\{X(t) : t \in \mathcal{T}\}$, or even just $X(t)$ when $\mathcal{T}$ is clear from the context.
- We shall indicate by $x(\omega, t)$ a sample path of the process $X(t)$, that is, a particular realization, or trajectory of the process in correspondence of the abstract random experiment outcome ω.

Outline

Extension of definitions (from dt to ct)

- We have defined and discussed stationarity, cyclostationarity, ergodicity, etc., for discrete-time random processes
- All what said for discrete-time random processes holds almost verbatim for continuous-time random processes, with the following replacements:

$$\begin{aligned} \sum_n \quad &\to \quad \int \; dt \\ \text{dt-Fourier Transform} \quad &\to \quad \text{ct-Fourier Transform} \\ \text{dt-convolution} \quad &\to \quad \text{ct-convolution} \\ \text{dt-frequency domain } [-1/2, 1/2] \quad &\to \quad \text{ct-frequency domain } \mathbb{R} \end{aligned}$$

- Transform

$$\check{x}(f) = \int_{\mathcal{T}} x(t) e^{-\jmath 2\pi f t} dt$$

- Inverse transform

$$x(t) = \int_{-\infty}^{+\infty} \check{x}(f) e^{\jmath 2\pi f t} df$$

- Convolution

$$y(t) = \int h(\tau) x(t-\tau) d\tau \leftrightarrow \check{y}(f) = \check{h}(f)\check{x}(f)$$

Mean and autocorrelation functions

- Consider the complex proper random process $X(t)$ defined over $\mathcal{T} = \mathbb{R}$.
- The mean function $\mu : \mathcal{T} \to \mathbb{R}$ is defined as

$$\mu(t) = \mathbb{E}[X(t)]$$

- The (auto)covariance function $c_{xx} : \mathcal{T} \times \mathcal{T} \to \mathbb{R}$ is defined as

$$c_{xx}(t_1, t_2) = \mathrm{Cov}(X(t_1), X(t_2)) = \mathbb{E}[X(t_1)X^*(t_2)] - \mu(t_1)\mu^*(t_2)$$

- The (auto)correlation function $r_{xx} : \mathcal{T} \times \mathcal{T} \to \mathbb{R}$ is defined as

$$r_{xx}(t_1, t_2) = \mathbb{E}[X(t_1)X^*(t_2)] = c_{xx}(t_1, t_2) + \mu(t_1)\mu^*(t_2)$$

- Usually these are referred to as “covariance” and “correlation” functions...

Definition

A random process $X(t)$ is called Wide-Sense Stationary (WSS) if its second-order statistics are invariant with respect to time-shifts τ for any $\tau \in \mathcal{T}$.

- The mean function of a WSS process $X(t)$ satisfies, for all τ:

$$\mathbb{E}[X(t-\tau)] = \mathbb{E}[X(t)] \quad \Rightarrow \quad \mu(t) = \mu \quad \text{(constant function)}$$

- The autocorrelation function (and covariance function) of a WSS process $X(t)$ satisfies, for all τ:

$$\mathbb{E}[X(t_1-\tau)X^*(t_2-\tau)] = \mathbb{E}[X(t_1)X^*(t_2)] \quad \Rightarrow \quad r_{xx}(t_1-\tau, t_2-\tau) = r_{xx}(t_1, t_2)$$

- Letting $\tau = t_2$ we have that $r_{xx}(t_1, t_2) = r_{xx}(t_1 - t_2, 0)$, that is the autocorrelation function depends only on the time difference.
- For WSS processes, with some abuse of notation, we define $r_{xx}(\tau) = r_{xx}(t, t-\tau)$ and $c_{xx}(\tau) = c_{xx}(t, t-\tau)$.

Wide-Sense Cyclostationarity

Definition

A random process $X(t)$ is called Wide-Sense Cyclostationary (WSC) of period T if its second-order statistics are periodic functions of period T.

- The mean function of a WSC process $X(t)$ satisfies:

$$\mathbb{E}[X(t-T)] = \mathbb{E}[X(t)] \quad \Rightarrow \quad \mu(t-T) = \mu(t) \quad \text{(periodic)}$$

- The auto-correlation function (and covariance function) of a WSC process $X(t)$ satisfies:

$$\mathbb{E}[X(t_1-T)X^*(t_2-T)] = \mathbb{E}[X(t_1)X^*(t_2)] \quad \Rightarrow \quad r_{xx}(t_1-T, t_2-T) = r_{xx}(t_1, t_2)$$

- Letting $t_1 = t$ and $t_2 = t - \tau$ we have that $r_{xx}(t, t-\tau) = r_{xx}(t-T, t-T-\tau)$ is a periodic function in the variable t, for any time difference τ.

- In order to study second-order processes in continuous time we need to develop a theory for continuity, differentiability, and integrability.
- This allows us to study the effect of continuous-time random processes as input and output of linear systems (i.e., convoluted with a system impulse response), or as input/output of system of differential equations.
- We develop tools for the existence of limits in the m.s. sense very similar to what already done for dt-processes.

And many other things holds as in the discrete case

- Loeve's criterion on the existence of the limit of $X(t)$ for $t \to \infty$ (e.g., in the mean-square sense) "controlled" by the (existence of the) limit of auto-correlation function;
- mean-square continuity and uniform mean-square continuity
- mean-square differentiability
- etc.

Loeve's criterion

Theorem

Let $\{X(t) : t \in \mathcal{T}\}$ be a random process with autocorrelation function $r(t_1, t_2)$. The limit of $X(t)$ for $t \to \infty$ exists in the mean-square sense if and only if $\lim_{t_1,t_2\to\infty} r(t_1, t_2) = r \in \mathbb{R}_+$ (a constant, independent of how t_1, t_2 go to infinity). □

Proof: it is analogous to what already done for the discrete-time case.

Mean-square continuity

Definition

Let $X(t)$ denote a real or complex-valued random process defined over $\mathcal{T} \subseteq \mathbb{R}$. We say that $X(t)$ is continuous in the m.s. sense at t_0 if

$$\lim_{t \to t_0} \mathbb{E}[|X(t) - X(t_0)|^2] = 0$$

or, more explicitly, for all $\epsilon > 0$ there exist $\delta(t_0, \epsilon)$ such that

$$\mathbb{E}[|X(t) - X(t_0)|^2] < \epsilon, \;\; \forall \;\; |t - t_0| < \delta(t_0, \epsilon)$$

Furthermore, if for all $t_0 \in \mathcal{T}$ there exists some $\delta(t_0, \epsilon) = \delta(\epsilon)$ satisfying the above condition that does not depend on t_0, we say that $X(t)$ is uniformly continuous in the m.s. sense.

- Stated in terms of the correlation function $r_{xx}(t_1, t_2)$, the condition for m.s. continuity yields

$$r_{xx}(t,t) - r_{xx}(t,t_0) - r_{xx}(t_0,t) + r_{xx}(t_0,t_0) < \epsilon, \quad \forall \; |t - t_0| < \delta(t_0, \epsilon)$$

Lemma

$X(t)$ is m.s. continuous at t_0 if and only if its auto-correlation function $r_{xx}(t_1, t_2)$ is continuous at the point (t_0, t_0). □

Corollary

If $X(t)$ is WSS, then it is m.s. continuous if and only if $r_{xx}(\tau)$ is continuous at $\tau = 0$. Then, a WSS process $X(t)$ is either uniformly m.s. continuous at all t, or discontinuous at all t.

Mean-square differentiability

Definition

Let $X(t)$ denote a real or complex-valued random process defined over $\mathcal{T} \subseteq \mathbb{R}$. We say that $X(t)$ is differentiable at $t_0 \in \mathcal{T}$ in the mean-square sense if the sequence of random variables

$$Y_h = \frac{X(t_0 + h) - X(t_0)}{h}$$

converges in mean-square to some limit Y, as $h \to 0$. In this case, the limit $\dot{X}(t_0)$ is the m.s. derivative of $X(t)$ at t_0.

Lemma

$X(t)$ is m.s. differentiable at t_0 if and only if $\frac{\partial^2 r_{xx}(t_1,t_2)}{\partial t_1 \partial t_2}$ exists and it is finite at the point (t_0, t_0).

Correlation of a process and its derivative

- Suppose that $X(t)$ is differentiable in the m.s. sense and let $\dot{X}(t) = \frac{d}{dt}X(t)$ denote the derivative process.
- Mean-square differentiability ensures that we can exchange expectation with the differentiation operation.

$$\mu_{\dot{x}}(t) = \frac{d}{dt}\mu_x(t)$$

$$r_{\dot{x}\dot{x}}(t_1, t_2) = \frac{\partial^2 r_{xx}(t_1, t_2)}{\partial t_1 \partial t_2}$$

$$r_{\dot{x}x}(t_1, t_2) = \frac{\partial r_{xx}(t_1, t_2)}{\partial t_1}$$

$$r_{x\dot{x}}(t_1, t_2) = \frac{\partial r_{xx}(t_1, t_2)}{\partial t_2}$$

- When $X(t)$ is WSS, we define the auto-correlation function $r_{xx}(t, t-\tau) = r_{xx}(\tau)$ and its derivatives $\dot{r}_{xx}(\tau) = \frac{d}{d\tau} r_{xx}(\tau)$ and $\ddot{r}_{xx}(\tau) = \frac{d^2}{d\tau^2} r_{xx}(\tau)$, and obtain

$$\mu_{\dot{x}}(t) = \frac{d}{dt}\mu_x = 0$$

$$r_{\dot{x}\dot{x}}(t_1, t_2) = -\ddot{r}_{xx}(t_1 - t_2)$$

$$r_{\dot{x}x}(t_1, t_2) = \dot{r}_{xx}(t_1 - t_2)$$

$$r_{x\dot{x}}(t_1, t_2) = -\dot{r}_{xx}(t_1 - t_2)$$

- We conclude that $X(t)$ and $\dot{X}(t)$ are jointly WSS, and the derivative process has mean zero and auto-covariance function $-\ddot{r}_{xx}(\tau)$.

Riemann integration of random processes

- The Riemann integral of $X(t)$ over $[a,b]$ is defined as the limit:

$$S[a,b] = \int_a^b X(t)dt = \lim_{m\to\infty} \sum_{i=1}^{m} X(t_i')(t_{i+1} - t_i)$$

  where $\mathcal{T}_m = \{t_0, t_1, \ldots, t_m\}$ is a grid of non-decreasing indices such that $t_0 = a$ and $t_m = b$, $\mathcal{T}_m' = \{t_1', \ldots, t_m'\}$ is a sequence of indices such that $t_{i-1} \leq t_i' < t_i$ for all i, and for all sufficiently large m, $\mathcal{T}_m$ satisfies

$$\max_{1\leq i\leq m} |t_i - t_{i-1}| \leq \delta_m \qquad (*)$$

  where $\delta_m \to 0$ as $m \to \infty$.
- The limit with respect to m is a short-hand notation to indicate the limit along any sequence of sets $\mathcal{T}_m$, $\mathcal{T}_m'$, satisfying the above conditions, and the limit must exist and be the same irrespectively of the sequence of sets, as long as condition $(*)$ is satisfied.

- When the integrand function is a random process, we have to specify in which sense the sequence of RVs $S_m = \sum_{i=1}^{m} X(t_i')(t_{i+1} - t_i)$ converges.
- We use Loeve's criterion: $\mathbb{E}[S_m S_n^*]$ must converge to some real limit for $m, n \to \infty$ irrespectively of the path.
- We have

$$\mathbb{E}\left[S_m S_n^*\right] = \sum_{i=1}^{m} \sum_{j=1}^{n} r_{xx}(t_i', s_j')(t_{i+1} - t_i)(s_{j+1} - s_j)$$

- Taking the limit for $m, n \to \infty$ we arrive at the necessary and sufficient condition: the Riemann integral $\int_a^b X(t)dt$ exists in the m.s. sense if and only if

$$\int_a^b \int_a^b r_{xx}(t, s) dt ds < \infty$$

- The output of an LTI system with impulse response $h(t)$ is written as

$$y(t) = \int h(\tau)x(t-\tau)d\tau$$

- When the input is a random process $X(t)$, then the output exists in a mean-square sense if the convolution integral

$$Y(t) = \int h(\tau)X(t-\tau)d\tau$$

  exists in the m.s. sense, that is, the process $Z_t(\tau) = h(\tau)X(t-\tau)$ must be integrable in the m.s. sense for all t.
- Using the necessary and sufficient condition seen before (and restricting to well-behaved processes and systems for which Riemann integration applies), we have the necessary and sufficient condition

$$\int\int h(\tau)h(\tau')r_{xx}(t-\tau, t-\tau')d\tau d\tau' < \infty$$

  for every $t \in \mathcal{T}$.
- If $X(t)$ is WSS we have $r_{xx}(t-\tau, t-\tau') = r_{xx}(\tau'-\tau)$, therefore the condition becomes

$$\int\int h(\tau)h(\tau')r_{xx}(\tau'-\tau)d\tau d\tau' < \infty$$

Lemma

The output $Y(t)$ to the LTI system with impulse response $h(\tau)$ and WSS input $X(t)$ exists in the m.s. sense if the system is BIBO-stable.

Proof: We can write

$$\begin{aligned} \int\int h(\tau)h(\tau')r_{xx}(\tau'-\tau)d\tau d\tau' &\leq \int\int |h(\tau)h(\tau')||r_{xx}(\tau'-\tau)|d\tau d\tau' \\ &\leq r_{xx}(0)\left(\int |h(\tau)|d\tau\right)^2 \end{aligned}$$

where $|r_{xx}(\tau'-\tau)| \leq r_{xx}(0)$ follows from the Cauchy-Schwartz inequality. If the system is BIBO-stable, then its impulse response is absolutely integrable.

Input-output second-order statistics

- Consider a WSS input X to an LTI BIBO-stable system with impulse response $h(\tau)$, and let $Y(t) = \int h(\tau)X(t-\tau)d\tau$ denote the output.
- Mean function of the output:

$$\mu_y = \int h(\tau)\mu_x d\tau = \mu_x \int h(\tau)d\tau$$

- Auto-correlation function of the output

$$r_{yy}(t_1 - t_2) = \int \int h(t_1 - t')r_{xx}(t' - t'')h^*(t_2 - t'')dt'dt'' =$$

$$= \int \int h(\tau')r_{xx}(t_1 - t_2 - \tau' + \tau'')h^*(\tau'')d\tau'd\tau''$$

- Output-input cross-correlation function

$$r_{yx}(t_1 - t_2) = \int h(t_1 - t')r_{xx}(t' - t_2)dt' = \int h(\tau)r_{xx}(t_1 - t_2 - \tau)d\tau$$

- Input-output cross-correlation function (using the fact that $r_{xy}(t_1, t_2) = r_{yx}(t_2, t_1)$)

$$r_{xy}(t_1 - t_2) = \int h^*(t_2 - t')r_{xx}(t' - t_1)dt' = \int h^*(-\tau)r_{xx}(t_1 - t_2 - \tau)d\tau$$

- Written in a more compact way, we have

$$\begin{aligned} r_{yy}(\tau) &= h(\tau) \otimes h^*(-\tau) \otimes r_{xx}(\tau) \\ r_{yx}(\tau) &= h(\tau) \otimes r_{xx}(\tau) \\ r_{xy}(\tau) &= h^*(-\tau) \otimes r_{xx}(\tau) \end{aligned}$$

- Notice: these are completely analogous to the discrete-time case.

Power spectral density

- The Fourier transform of the truncated process $\{X(t) : t \in [-T/2, T/2]\}$ is

$$\check{X}_T(f) = \int X_T(f) e^{-\jmath 2\pi f t} dt = \int_{-T/2}^{T/2} X(t) e^{-\jmath 2\pi f t} dt$$

- The power of $X(t)$ is defined as

$$\mathcal{P}_x = \lim_{T\to\infty} \frac{1}{T} \mathbb{E}\left[\int_{-T/2}^{T/2} |X(t)|^2 dt \right] = \lim_{T\to\infty} \frac{1}{T} \mathbb{E}\left[\int \left|\check{X}_T(f)\right|^2 df \right]$$

  when this limit exists.
- Assuming that we can exchange limit and integration, we define the power spectral density (PSD) of $X(t)$ as

$$P_x(f) = \lim_{T\to\infty} \frac{1}{T} \mathbb{E}\left[\left|\check{X}_T(f)\right|^2 \right]$$

- It follows that, if the PSD exists (exchanging limit with integration is valid), then the signal power is given by

$$\mathcal{P}_x = \int P_x(f) df$$

Wiener-Khintchine Theorem: continuous time

Theorem

If $X(t)$ is WSS with absolutely integrable auto-correlation function $r_{xx}(\tau)$, then

$$P_x(f) = \int r_{xx}(\tau) e^{-\jmath 2\pi f \tau} d\tau$$

□

Proof: as in the discrete setting with a change of variable, skipped here.

Cross-Spectrum and output PSD

- A simple corollary of the Wiener-Khintchine theorem concerns the case of jointly WSS processes $X(t)$ and $Y(t)$ with absolutely integrable cross-correlation function $r_{xy}(\tau) = \mathbb{E}[X(t+\tau)Y^*(t)]$.
- In this case, we can define the cross-spectrum as the Fourier transform

$$P_{xy}(f) = \int r_{xy}(\tau) e^{-\jmath 2\pi f \tau} d\tau$$

- When X and Y are the input and output of a stable LTI system with transfer function $\check{h}(f)$, we have

$$P_{xy}(f) = \check{h}^*(f) P_x(f), \quad P_{yx}(f) = \check{h}(f) P_x(f)$$

and

$$P_y(f) = |\check{h}(f)|^2 P_x(f)$$

PSD of more general processes

- For a wide class of processes, the time-average power

$$\mathcal{P}_x = \lim_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\int_{-T/2}^{T/2} |X(t)|^2 dt\right] = \lim_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\int_{-\infty}^{\infty} |\check{X}_T(f)|^2 df\right]$$

  exists.
- In addition, it happens that $\lim_{T\to\infty} \frac{1}{T}\mathbb{E}[|\check{X}_T(f)|^2] = P_x(f)$ exists for all $f \in \mathbb{R}$ and that $\mathcal{P}_x = \int P_x(f)df$.
- In this case, we wish to calculate the PSD $P_x(f)$ even though $X(t)$ is not WSS.

Theorem

For sufficiently well-behaved $r_{xx}(t_1, t_2)$*,*

$$P_x(f) = \int \bar{r}_{xx}(\tau) e^{-\jmath 2\pi f \tau} d\tau$$

where

$$\bar{r}_{xx}(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} r_{xx}(\tau + \theta, \theta) d\theta$$

is the time-averaged auto-correlation function.

Corollary

The family of WSC processes has mean function $m_x(t)$ *periodic with period* T *and auto-correlation function* $r_{xx}(t + \tau, t)$ *periodic with respect to* t *with period* T *for all* τ*. In this case, the time-averaged auto-correlation function is easily obtained by averaging over one period:*

$$\bar{r}_{xx}(\tau) = \frac{1}{T} \int_0^T r_{xx}(\tau + \theta, \theta) d\theta$$

From ct to dt: Shannon sampling theorem

Theorem

Let $x(t)$ be a function with Fourier transform $\check{x}(f)$ with support strictly inside the interval $[-B/2, B/2]$. Then, the following equality holds pointwise

$$x(t) = \sum_{n=-\infty}^{\infty} x(n/B)\,\mathrm{sinc}(B(t - n/B))$$

Notice: The set of functions $\psi_n(t) = \sqrt{B}\,\mathrm{sinc}(B(t - n/B))$ for $n \in \mathbb{Z}$ forms an orthonormal basis. This is a complete basis for the set of functions with bandwidth strictly limited in $[-B/2, B/2]$.

Sampling of band-limited processes

Theorem

Let $X(t)$ be a ct WSS process with PSD $P_x(f)$, with support strictly inside the interval $[-B/2, B/2]$ (strictly band-limited WSS process). Then, the following equality holds in the m.s. sense

$$X(t) = \sum_{n=-\infty}^{\infty} X(n/B)\,\mathrm{sinc}(B(t - n/B))$$

Notice: This theorem says that band-limited continuous-time processes can be essentially identified with discrete-time processes, obtained by sampling at an appropriate rate B samples per unit time.
It follows that almost all processes relevant in system-theory problems with finite bandwidth can be safely studied by looking at their discrete-time equivalent.

Thank you! Q & A?