Lecture 12: Second-order Processes II

Lecturer: Ziqiao Wang

In this lecture, we will continue our discussion of second-order processes. We will show how a cyclostationary (and hence non-stationary) process can be transformed into a stationary one, and we will discuss orthogonal increment processes and the properties of their autocorrelation functions. Finally, we will introduce the ergodic theorem.

1 Cyclostationarity and Wide-Sense Cyclostationarity

It is possible to transform a cyclostationary process into a stationary one by introducing a random time shift:

Theorem 1.1 (Creation of a Stationary Process from a Cyclostationary Process)**.** *Let $X(t)$ be a cyclostationary process with period T_0. Let Θ be a random variable, uniformly distributed over $\left[-\frac{T_0}{2}, \frac{T_0}{2}\right]$, and independent of the process. Define*

$$\widetilde{X}_t \triangleq X_{t+\Theta}, \qquad t \in \times. \tag{1}$$

Then $\widetilde{X}(t)$ is a stationary process.

Proof. By direct computation, we have, for any $t_1, \ldots, t_n \in \mathcal{T}$ and $x_1, \ldots, x_n \in \mathbb{R}$,

$$\begin{aligned} F_{\tilde{X}_{t_1},\ldots,\tilde{X}_{t_n}}(x_1, \ldots, x_n) &= \int_{-\infty}^{\infty} \mathbb{P}\Big(\tilde{X}_{t_1} \le x_1, \ldots, \tilde{X}_{t_n} \le x_n \;\Big|\; \Theta = \theta\Big) f_T(\theta)\, d\theta \\ &= \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} F_{X_{t_1+\theta},\ldots,X_{t_n+\theta}}(x_1, \ldots, x_n)\, d\theta, \end{aligned} \tag{2}$$

since Θ is uniformly distributed on $\left[-\frac{T_0}{2}, \frac{T_0}{2}\right]$. Because $F_{X_{t_1+\theta},\ldots,X_{t_n+\theta}}(x_1, \ldots, x_n)$ is periodic in x with period T_0, the integral above averages over complete periods of this function. Thus, it's easy to see that

$$F_{\tilde{X}_{t_1},\ldots,\tilde{X}_{t_n}}(x_1, \ldots, x_n) = F_{\tilde{X}_{t_1+\tau},\ldots,\tilde{X}_{t_n+\tau}}(x_1, \ldots, x_n), \qquad \forall\, \tau \in \times, \tag{3}$$

showing that the process $\tilde{X}(t)$ is stationary. □

Remark 1.1. *The corresponding result for discrete-time cyclostationary processes follows in an analogous manner. In addition, we can likewise show that if $X(t)$ is a wide-sense cyclostationary process with period T_0, then the process*

$$\bar{X}_t \triangleq X_{t+\Theta}, \qquad t \in \mathbb{R},$$

where Θ is uniformly distributed on $[0, T_0]$ and is independent of the process, is wide-sense stationary with

$$m_{\bar{X}} = \frac{1}{T_0} \int_0^{T_0} m_X(t)\, dt, \tag{4}$$

$$R_{\bar{X}}(\tau) = \frac{1}{T_0} \int_0^{T_0} R_X(t, t+\tau)\, dt. \tag{5}$$

Example 1. Let $X_s(t)$ be the phase-shifted version of the pulse amplitude–modulated waveform $X(t)$ introduced in Example 3 in Lecture 11. Find the mean and autocorrelation function of $X_s(t)$.

Note that $X_s(t)$ has zero mean since $X(t)$ is zero-mean. The autocorrelation of $X_s(t)$ has been obtained from Example 3 in Lecture 11, we can see that for $0 < t+\tau < T$, $R_X(t+\tau, t) = 1$ and $R_X(t+\tau, t) = 0$ otherwise. Therefore,

$$\text{for } 0 < \tau < T: \qquad R_{X_s}(\tau) = \frac{1}{T}\int_0^{T-\tau} dt = \frac{T-\tau}{T},$$

$$\text{for } -T < \tau < 0: \qquad R_{X_s}(\tau) = \frac{1}{T}\int_{-\tau}^{T} dt = \frac{T+\tau}{T}.$$

Thus, $X_s(t)$ has a triangular autocorrelation function:

$$R_{X_s}(\tau) = \begin{cases} 1 - \frac{|\tau|}{T}, & |\tau| \le T, \\ 0, & |\tau| > T. \end{cases}$$

2 Orthogonal Increment Process

To further understand wide-sense stationarity and non-stationarity, we introduce a representative class of non-stationary random processes, namely the *orthogonal increment processes*, whose theoretical and practical importance is well recognized.

Definition 2.1 (Orthogonal Increment Process)**.** For a second-order process $X(t), t \in \mathbb{R}$, if

$$\forall\, t_1 < t_2 \le t_3 < t_4, \quad t_1, t_2, t_3, t_4 \in \mathbb{R},$$

the condition

$$\mathbb{E}\big[(X_{t_4} - X_{t_3})(X_{t_2} - X_{t_1})\big] = 0$$

holds, then the process is called an *orthogonal increment process.*

This definition indicates that increments over different time intervals are mutually orthogonal. The term "orthogonal", with its clear geometric meaning, highlights the geometric interpretation of correlation computations.

Definition 2.2 (Independent Increment Process)**.** For a stochastic process $X(t), t \in \mathbb{R}$, if

$$\forall\, t_1 < t_2 \le t_3 < t_4, \quad t_1, t_2, t_3, t_4 \in \mathbb{R},$$

the increments $X_{t_4} - X_{t_3}$ and $X_{t_2} - X_{t_1}$ are statistically independent, then the process is called an *independent increment process.*

Remark 2.1. *If $X(t)$ is an independent increment process with zero mean, then $X(t)$ is an orthogonal increment process. Indeed, for*

$$\forall\, t_1 < t_2 \le t_3 < t_4,$$

if the increments $X_{t_4} - X_{t_3}$ and $X_{t_2} - X_{t_1}$ are statistically independent, then

$$\mathbb{E}[(X_{t_4} - X_{t_3})(X_{t_2} - X_{t_1})] = \mathbb{E}[X_{t_4} - X_{t_3}]\,\mathbb{E}[X_{t_2} - X_{t_1}] = 0.$$

Another notion that appears frequently in random process theory is that of *stationary increments.*

Definition 2.3. A process $X(t)$ is said to have stationary increments if, for any shift h, the process $\{X_{t+h} - X_t\}_{t\in\mathcal{T}}$ is stationary, namely the distribution of the increment $X(t+h) - X(t)$ depends only on the difference h.

Clearly, a stationary process has stationary increments.

The autocorrelation function of a process with orthogonal increments has a unique form.

Theorem 2.1. *Let $X(t), t \in [0, \infty)$ be a stochastic process with $X_0 = 0$. Then a* necessary and sufficient condition *for $X(t)$ to be a process with orthogonal increments is that its autocorrelation function satisfies*

$$R_X(s,t) = F(\min(s,t)),$$

where $F(\cdot)$ is a non-decreasing function.

Proof. We first prove the necessity. When $s > t$, we have

$$\begin{aligned} R_X(t,s) = \mathbb{E}(X_t X_s) =& \mathbb{E}\big((X_t - X_s + X_s)X_s\big) \\ =& \mathbb{E}\big((X_t - X_s)(X_s - X_0)\big) + \mathbb{E}(|X_s|^2) \\ =& \mathbb{E}(|X_s|^2) = F(s). \end{aligned}$$

Similarly, for $t < s$,

$$R_X(t,s) = F(t).$$

We now verify that $F(\cdot)$ is non-decreasing. When $s < t$,

$$\begin{aligned} F(t) - F(s) =& F(t) - F(s) - F(s) + F(s) \\ =& \mathbb{E}(|X_t|^2) - \mathbb{E}(|X_s|^2) \\ =& \mathbb{E}(|X_t|^2) - \mathbb{E}(X_t X_s) - \mathbb{E}(X_t X_s) + \mathbb{E}(|X_s|^2) \\ =& \mathbb{E}|X_t - X_s|^2 \geq 0. \end{aligned}$$

Hence, $F(\cdot)$ is a non-decreasing function.

Next, we prove sufficiency. If the autocorrelation function of $X(t)$ satisfies

$$R_X(t,s) = F(\min(s,t)),$$

then for all $t_1 < t_2 \leq t_3 < t_4$, where $t_1, t_2, t_3, t_4 \in \mathbb{R}$,

$$\begin{aligned} \mathbb{E}\big((X_{t_4} - X_{t_3})(X_{t_2} - X_{t_1})\big) =& \mathbb{E}(X_{t_4} X_{t_2}) - \mathbb{E}(X_{t_3} X_{t_2}) - \mathbb{E}(X_{t_4} X_{t_1}) + \mathbb{E}(X_{t_3} X_{t_1}) \\ =& F(\min(t_4, t_2)) - F(\min(t_3, t_2)) - F(\min(t_4, t_1)) + F(\min(t_3, t_1)) \\ =& F(t_2) - F(t_2) - F(t_1) + F(t_1) = 0. \end{aligned}$$

Therefore, $X(t)$ is a process with orthogonal increments. □

Example 2 (Brown Motion)**.** A stochastic process $\{B(t),\ t \geq 0\}$ is called a *Brownian motion* (or Wiener process) with variance parameter $\sigma^2 > 0$ if it satisfies: (i) $B_0 = 0$ almost surely. (ii) **Orthogonal increments**. (iii) **Gaussian increments:** For all $t > s \geq 0$, $B_t - B_s \sim \mathcal{N}\big(0,\ \sigma^2(t-s)\big)$.

For Brownian motion with variance parameter σ^2, the autocovariance function is

$$R_B(t,s) = \mathbb{E}\Big(B^2_{\min(t,s)}\Big) = \mathrm{Var}(B_{\min(t,s)}) = \sigma^2 \min(t,s).$$

Let $\{B(t),\ t \geq 0\}$ be a Brownian motion with variance parameter σ^2. Define the *white noise process* formally as the generalized derivative

$$Y(t) = \frac{d}{dt} B(t).$$

Although $B(t)$ is almost surely nowhere differentiable, the process $Y(t)$ is well-defined in the sense of generalized stochastic processes. Its autocorrelation function can be computed formally as follows.

We compute

$$R_Y(t,s) = \mathbb{E}[Y_t Y_s] = \mathbb{E}\left[\frac{d}{dt}B_t \frac{d}{ds}B_s\right].$$

Using the fact that differentiation under expectation is valid for generalized processes,

$$R_Y(t,s) = \frac{\partial^2}{\partial t\,\partial s}\,\mathbb{E}[B_t B_s] = \frac{\partial^2}{\partial t\,\partial s}\,R_B(t,s) = \frac{\partial^2}{\partial t\,\partial s}\big(\sigma^2 \min(t,s)\big).$$

We then use the fact that $\min(t,s) = \frac{1}{2}\big(t+s-|t-s|\big)$, hence,

$$\begin{aligned}
R_Y(t,s) =& \frac{\partial^2}{\partial t\,\partial s}\left[\frac{\sigma^2}{2}\big(t+s-|t-s|\big)\right]\\
=& -\frac{\sigma^2}{2}\,\frac{\partial^2}{\partial t\,\partial s}\,|t-s|\\
=& -\frac{\sigma^2}{2}\,\frac{\partial}{\partial s}\,\operatorname{Sgn}(t-s)\\
=& -\frac{\sigma^2}{2}\,\frac{\partial}{\partial s}\,\big(U(t-s)-U(s-t)\big)\\
=& \sigma^2\,\delta(t-s).
\end{aligned}$$

That is,

$$\boxed{R_Y(t,s) = \sigma^2 \delta(t-s).}$$

Thus the derivative of Brownian motion is a **Gaussian white noise process** with intensity σ^2.

Notice that we **turn an orthogonal increment process to a W.S.S. process**.

3 Ergodic Theorem

In many situations, to estimate statistical quantities of a random process $X(t,\omega)$, we repeat the random experiment that generates the process a large number of times and take the arithmetic average of the quantities of interest. For example, to estimate the mean $m_X(t)$ of a random process $X(t,\omega)$, we repeat the experiment and compute the empirical average:

$$\hat{m}_X(t) = \frac{1}{N}\sum_{i=1}^{N} X(t,\omega_i),$$

where N is the number of repetitions of the experiment, and $X(t,\omega_i)$ denotes the realization observed in the i-th repetition.

In some situations, we are interested in estimating the mean or the autocorrelation function from the *time average* of a single realization. That is,

$$\langle X(t)\rangle_T = \frac{1}{2T}\int_{-T}^{T} X(t,\omega)\,dt.$$

An *ergodic theorem* provides conditions under which a time average converges to the ensemble average as the observation interval becomes large.

The strong law of large numbers is one of the most important ergodic theorems. It states that if X_n is an i.i.d. discrete-time random process with finite mean $\mathbb{E}[X_n] = m$, then the time average of the samples converges to the ensemble average with probability one:

$$P\left[\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} X_i = m\right] = 1.$$

This result allows us to estimate m by taking the time average of a single realization of the process. We are interested in obtaining results of this type for a larger class of random processes, that is, for non-i.i.d. discrete-time random processes, and for continuous-time random processes.

Example 3. Let $X(t) = A$ for all t, where A is a zero-mean, unit-variance random variable. Find the limiting value of the time average.

The mean of the process is

$$m_X(t) = \mathbb{E}\,[X(t)] = \mathbb{E}\,[A] = 0.$$

However,

$$\langle X(t)\rangle_T = \frac{1}{2T}\int_{-T}^{T} A\,dt = A.$$

Thus, the time-average mean does *not* always converge to $m_X(t) = 0$.

Remark 3.1. *Note that this process is stationary. Therefore, this example shows that* **stationary processes need not be ergodic**.

Consider the estimate $\langle X(t)\rangle_T$ for $\mathbb{E}\,[X(t)] = m_X(t)$. The estimate is independent of t, so obviously it only makes sense to consider processes for which $m_X(t) = m$, a constant. We now develop an ergodic theorem for the time average of WSS process.

Let $X(t)$ be a WSS process. The expected value of $\langle X(t)\rangle_T$ is

$$\mathbb{E}\,[\langle X(t)\rangle_T] = \mathbb{E}\left[\frac{1}{2T}\int_{-T}^{T} X(t)\,dt\right] = \frac{1}{2T}\int_{-T}^{T}\mathbb{E}\,[X(t)]\,dt = m.$$

This equation states that $\langle X(t)\rangle_T$ is an unbiased estimator for m.

Consider the variance of $\langle X(t)\rangle_T$:

$$\begin{aligned}
\mathrm{Var}[\langle X(t)\rangle_T] =& \mathbb{E}\left[(\langle X(t)\rangle_T - m)^2\right] \\
=& \mathbb{E}\left[\left\{\frac{1}{2T}\int_{-T}^{T}(X(t)-m)\,dt\right\}\left\{\frac{1}{2T}\int_{-T}^{T}(X(t')-m)\,dt'\right\}\right] \\
=& \frac{1}{4T^2}\int_{-T}^{T}\int_{-T}^{T} E[(X(t)-m)(X(t')-m)]\,dt\,dt' \\
=& \frac{1}{4T^2}\int_{-T}^{T}\int_{-T}^{T} C_X(t,t')\,dt\,dt'.
\end{aligned}$$

Since the process $X(t)$ is WSS, the equation becomes

$$\mathrm{Var}[\langle X(t)\rangle_T] = \frac{1}{4T^2}\int_{-T}^{T}\int_{-T}^{T} C_X(t-t')\,dt\,dt'.$$

Using basic calculus (e.g., let $\tau = t - t'$), we have

$$\begin{aligned}
\mathrm{Var}[\langle X(t)\rangle_T] =& \frac{1}{4T^2}\int_{-2T}^{2T}(2T - |\tau|)C_X(\tau)\,d\tau \\
=& \frac{1}{2T}\int_{-2T}^{2T}\Big(1 - \frac{|\tau|}{2T}\Big)C_X(\tau)\,d\tau.
\end{aligned}$$

Therefore,

$$\langle X(t)\rangle_T \longrightarrow m \quad \text{in the mean–square sense, that is,} \quad \mathbb{E}\left[(\langle X(t)\rangle_T - m)^2\right] \to 0,$$

provided that the expression in the equation above approaches zero as T increases. We have just proved the following ergodic theorem.

Theorem 3.1. *Let $X(t)$ be a WSS process with $m_X(t) = m$. Then*

$$\lim_{T\to\infty} \langle X(t)\rangle_T = m$$

in the mean square sense, if and only if

$$\lim_{T\to\infty} \frac{1}{2T}\int_{-2T}^{2T} \left(1 - \frac{|\tau|}{2T}\right) C_X(\tau)\, d\tau = 0.$$

We say that a WSS process is **mean ergodic** if it satisfies the conditions of the above theorem.

The above theorem can be used to obtain ergodic theorems for the time average of other quantities. For example, if we replace $X(t)$ with $Y(t+\tau)Y(t)$, we obtain a time-average estimate for the autocorrelation function of the process $Y(t)$:

$$\langle Y(t+\tau)Y(t)\rangle_T = \frac{1}{2T}\int_{-T}^{T} Y(t+\tau)Y(t)\, dt. \tag{9.105}$$

It is easily shown that

$$\mathbb{E}\left[\, Y(t+\tau)Y(t)\,\right] = R_Y(\tau) \quad \text{if } Y(t) \text{ is WSS.}$$

The above ergodic theorem then implies that the time-average autocorrelation converges to $R_Y(\tau)$ in the mean square sense if the term with $X(t)$ replaced by $Y(t)Y(t+\tau)$ converges to zero.

If the random process under consideration is discrete-time, then the time-average estimate for the mean and the autocorrelation functions of X_n are given by

$$\langle X_n\rangle_T = \frac{1}{2T+1}\sum_{n=-T}^{T} X_n,$$

$$\langle X_{n+k}X_n\rangle_T = \frac{1}{2T+1}\sum_{n=-T}^{T} X_{n+k}X_n.$$

If X_n is a WSS random process, then $\mathbb{E}\left[\langle X_n\rangle_T\right] = m$, and so $\langle X_n\rangle_T$ is an unbiased estimate for m. It is also easy to show that the variance of $\langle X_n\rangle_T$ is

$$\mathrm{Var}\left[\langle X_n\rangle_T\right] = \frac{1}{2T+1}\sum_{k=-2T}^{2T} \left(1 - \frac{|k|}{2T+1}\right) C_X(k).$$

Therefore, $\langle X_n\rangle_T$ approaches m in the mean square sense and is mean ergodic if the expression in the equation above approaches zero with increasing T.