Mean Ergodicity

X_t is WSS and $\mu_X = \mathbb{E}[X_t]$. X_t is called mean ergodic in an appropriate sense if

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t X_t dt = \mu_X \text{ in an appropriate sense}$$

or

$$\lim_{K\to\infty} \frac{1}{K}\sum_{k=0}^{K} X_k dt = \mu_X \text{ in an appropriate sense}$$

Example

$\{X_k\}$ i.i.d. with $\mathbb{E}[X_k] = \mu$ and $Var(X_k) < C$ for some $C > 0$. By SLLN X is mean ergodic in the a.s. sense.

Example

X_1 is uniform over $[0, 1]$. $X_k = X_1$ for any $k > 1$. Therefore, we have

$$\mathbb{E}[X_k] = \mathbb{E}[X_1] = \frac{1}{2} \quad \text{and} \quad R_X(t, t+s) = \mathbb{E}[X_t, X_{t+s}] = \mathbb{E}[X_1^2].$$

Therefore, X_t is WSS. On the other hand,

$$\frac{1}{N} \sum_{k=1}^{N} X_k = X_1 \neq \frac{1}{2} \text{ in any sense.}$$

Therefore X_t is not mean ergodic.

Sufficient conditions for mean ergodicity in the m.s. sense

X_t is WSS and X_t is mean ergodic in the m.s. sense if one of the following conditions holds

1. $$\int_0^\infty |C_X(\tau)| d\tau < \infty.$$

2. $$\int_0^\infty |R_X(\tau)| d\tau < \infty.$$

3. $$\lim_{\tau \to \infty} R_X(\tau) = 0$$

4. $$\lim_{\tau \to \infty} C_X(\tau) = 0$$

Proof

To prove mean ergodicity in the m.s. sense, we need

$$\lim_{T\to\infty} \mathbb{E}\left[\left(\frac{1}{T}\int_0^T X_t dt - \mu_X\right)^2\right] = 0.$$

$$\mathbb{E}\left[\left(\frac{1}{T}\int_0^T X_t dt - \mu_X\right)^2\right] = \mathbb{E}\left[\frac{1}{T^2}\left(\int_0^T (X_t - \mu_X) dt\right)^2\right]$$

$$= \frac{1}{T^2}\int_0^T \int_0^T C_X(t,s) dt ds \quad (*)$$

$$= \frac{1}{T^2}\int_0^T \int_0^T C_X(t-s) dt ds \quad (\text{WSS})$$

*: Proposition 7.15 in “Random Processes for Engineers” by Bruce Hajek.

$$
\begin{aligned}
&\mathbb{E}[(\frac{1}{T}\int_0^T X_t dt - \mu_X)^2] \\
=&\frac{1}{T^2}\int_{s=0}^{T}\int_{\tau=-s}^{T-s} C_X(\tau)d\tau ds \quad (\tau = t - s) \\
=&\frac{1}{T^2}\int_0^T\int_{s=0}^{T-\tau} C_X(\tau)dsd\tau + \frac{1}{T^2}\int_{\tau=-T}^{0}\int_{s=-\tau}^{T} C_X(\tau)dsd\tau \\
=&\frac{2}{T^2}\int_0^T (T-\tau)C_X(\tau)d\tau
\end{aligned}
$$

$\lim_{T\to\infty} \frac{2}{T^2}\int_0^T (T-\tau)C_X(\tau)d\tau = 0$ implies mean erdocicity in the m.s. sense.

Condition (1):

$$\left|\frac{2}{T^2}\int_0^T (T-\tau)C_X(\tau)d\tau\right| \leq \left|\frac{2}{T}\int_0^T C_X(\tau)d\tau\right| \leq \frac{2}{T}\int_0^T |C_X(\tau)|d\tau,$$

which converges to zero if

$$\int_0^\infty |C_X(\tau)|d\tau < \infty.$$

Condition (4): For any $\epsilon > 0$, there exists T_ϵ such that $|C_X(\tau)| \leq \epsilon$ for $\tau \geq T_\epsilon$.

$$\begin{aligned}
&\lim_{T\to\infty} \frac{1}{T}\int_0^T \left(1-\frac{\tau}{T}\right) C_X(\tau)d\tau \\
\leq &\lim_{T\to\infty} \frac{1}{T}\int_0^{T_\epsilon} \left(1-\frac{\tau}{T}\right) C_X(\tau)d\tau + \frac{1}{T}\int_{T_\epsilon}^T \left(1-\frac{\tau}{T}\right) \epsilon d\tau \\
\leq &\epsilon.
\end{aligned}$$

Condition (2) and (3) imply $\mu_X = 0$.

Importance of Ergodicity

Consider a discrete-time random process $(X_n : n \in \mathbb{Z})$, and a function h which is bounded and Borel measurable on $\mathbb{R}^k$.

Question: When

$$\frac{1}{n}\sum_{j=1}^{n} h(X_j, X_{j+1}, \ldots X_{j+k-1}) \approx \mathbb{E}[h(X_1, \ldots, X_k)]?$$

In other words, can we learn the finite-dimensional statistics from data samples?

Definition: Ergodic

A stationary random process $\{X_n : n \in \mathbb{Z}\}$ is defined to be ergodic if

$$\lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} h(X_j, \ldots, X_{j+k-1}) = \mathbb{E}[h(X_1, \ldots, X_k)], \quad \forall k \text{ and } \forall h$$

where limit is taken in any of the three senses (a.s., m.s. or p.).

Importance of Ergodicity

Consider a discrete-time random process $(X_n : n \in \mathbb{Z})$, and a function h which is bounded and Borel measurable on $\mathbb{R}^k$.

Question: When

$$\frac{1}{n}\sum_{j=1}^{n} h(X_j, X_{j+1}, \ldots X_{j+k-1}) \approx \mathbb{E}[h(X_1, \ldots, X_k)]?$$

In other words, can we learn the finite-dimensional statistics from data samples?

Answer

If X_n is ergodic, then all of its finite dimensional distributions are determined as time averages.

Example

Consider an ergodic process $\{X_k\}$ and function

$$h(X_{k-1}, X_k) = \begin{cases} 1 \text{ if } X_{k-1} > 0 \geq X_k \\ 0 \text{ otherwise} \end{cases}$$

$h(\cdot) = 1$ if X_k makes a "down crossing" of level 0.
Since X_k is ergodic, we have

$$\lim_{n\to\infty} \frac{\text{number of downcrossings between 1 and n+1}}{n}$$
$$= \mathbb{E}[h(X_1, X_2)] = \mathbb{P}(X_1 > 0 \geq X_2)$$

Therefore, we can compute $\mathbb{P}(X_1 > 0 \geq X_2)$ by computing the long time-average down-crossing rate.

Two ergodic random processes

- $\{X_k\}$ i.i.d.
- $\{X_t\}$: Stationary Gaussian random process with $\lim_{\tau \to \infty} C_X(\tau) = 0$.

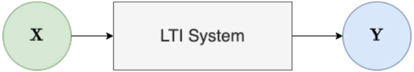
Linear time invariant (LTI) systems

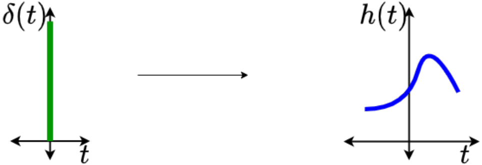
System process input signals $x(t)$ to produce output signals $y(t)$. The output at time t given by $y(t)$ can be dependent on past/future values of input signal.

A system is linear time invariant if it satisfies the following properties:

1. $x_i(t) \to y_i(t) \implies \sum_i c_i x_i(t) \to \sum_i c_i y_i(t)$
2. $x(t) \to y(t) \implies x(t-T) \to y(t-T)$

If a system is LTI then the output can be written as a convolution of the input signal and the impulse response $h(t)$, which is the output of the system to a $\delta(t)$ input.





$$x(t) \longrightarrow \begin{aligned} y(t) &= (x * h)(t) \\ &= \int_{-\infty}^{\infty} h(t-\tau)x(\tau)d\tau \end{aligned}$$

WSS process through LTI system

Definition: Joint Wide Sense Stationary (J-WSS)

Two process $\{X_t\}$ and $\{Y_t\}$ are said to be Joint-WSS if both the following conditions hold:

1. $\{X_t\}$ and $\{Y_t\}$ are both WSS.
2. Their cross correlation function $R_{XY}(t_1, t_2) := \mathbb{E}[X(t_1)Y(t_2)]$ depends on t_1 and t_2 only via their difference.

Theorem

Let $\{X_t\}$ be a WSS process which is passed a LTI system with impulse response h. The output process $\{Y_t\}$ and $\{X_t\}$ are J-WSS.

WSS process through LTI system

Theorem

Let $\{X_t\}$ be a WSS process which is passed a LTI system with impulse response h. The output process $\{Y_t\}$ and $\{X_t\}$ are J-WSS.

Proof: a) Mean function is constant

$$
\begin{aligned}
m_Y(t) = \mathbb{E}[Y_t] &= \mathbb{E}\left[\int_{-\infty}^{\infty} h(t-\tau)X(\tau)d\tau\right] \\
&= \int_{-\infty}^{\infty} h(t-\tau)\mathbb{E}[X(\tau)]d\tau = \int_{-\infty}^{\infty} h(t-\tau)m_X(\tau)d\tau \\
&= \int_{-\infty}^{\infty} h(t-\tau)\cdot c\, d\tau = c\int_{-\infty}^{\infty} h(\tau)d\tau
\end{aligned}
$$

Thus, $m_Y(t)$ is independent of t.

WSS process through LTI system

Theorem

Let $\{X_t\}$ be a WSS process which is passed a LTI system with impulse response h. The output process $\{Y_t\}$ and $\{X_t\}$ are J-WSS.

Proof: b) Cross correlation function $R_{XY}(t_1, t_2)$ depends only on $t_1 - t_2$

Before looking at the auto-correlation function R_Y, let's calculate the cross-correlation function $R_{XY}(t_1, t_2) := \mathbb{E}[X_{t_1} Y_{t_2}]$

$$R_{XY}(t_1, t_2) = \mathbb{E}\left[X(t_1)\int_{-\infty}^{\infty} h(\tau)X(t_2-\tau)d\tau\right] = \int_{-\infty}^{\infty} h(\tau)\mathbb{E}[X(t_1)X(t_2-\tau)]d\tau$$

$$= \int_{-\infty}^{\infty} h(\tau)R_X(t_1 - t_2 + \tau)d\tau = (\bar{h} * R_X)(t_1 - t_2) =: R_{XY}(t_1 - t_2)$$

where $\bar{h}(x) = h(-x)$.

The cross correlation function depends on t_1 and t_2 only via their difference $t_1 - t_2$.

WSS process through LTI system

Theorem

Let $\{X_t\}$ be a WSS process which is passed a LTI system with impulse response h. The output process $\{Y_t\}$ and $\{X_t\}$ are J-WSS.

Proof: c) Correlation function $R_Y(t_1, t_2)$ depends only on $t_1 - t_2$

$$
\begin{aligned}
R_Y(t_1, t_2) = \mathbb{E}[Y_{t_1} Y_{t_2}] &= \mathbb{E}\left[\left(\int_{-\infty}^{\infty} h(\tau) X(t_1 - \tau) d\tau\right) Y(t_2)\right] \\
&= \int_{-\infty}^{\infty} h(\tau)\mathbb{E}[X(t_1 - \tau) Y(t_2)] d\tau = \int_{-\infty}^{\infty} h(\tau) R_{XY}(t_1 - \tau, t_2) d\tau \\
&= \int_{-\infty}^{\infty} h(\tau) R_{XY}(t_1 - t_2 - \tau) d\tau = (h * R_{XY})(t_1 - t_2) =: R_Y(t_1 - t_2)
\end{aligned}
$$

Like the cross correlation function, the auto-correlation function too depends on t_1 and t_2 only via their difference $t_1 - t_2$.
Thus, $\{Y_t\}$ and $\{X_t\}$ are J-WSS.