Representation and statistical analyse of signals

Henrique Koji Miyamoto

1 Introduction and reminders on random variables

Definition 1.1. *Let (Ω, τ, P) be a probability space, where Ω is the sample space, τ is a σ-algebra and P is a probability measure, and the set of instants T ($\mathbb{R}$ or $\mathbb{Z}$). A (real) stochastic process is the application*

$$\begin{aligned} X:\ & T \times \Omega \to \mathbb{R} \\ & (t, \omega) \mapsto X(t, \omega). \end{aligned}$$

- For $\omega = \omega_0$ fixed, $X(t, \omega_0)$ is an ordinary function (trajectory or sample) and, for $t = t_0$ fixed, it is a random variable.
- If $T = \mathbb{R}$, the signal is time continuous and we denote the application $X(t, \omega)$. If $T = \mathbb{Z}$, we have a discrete time signal and we denote $X[k, \omega]$.

Definition 1.2. *A σ-algebra on a set Ω is a collection τ of subsets of Ω such that*

1. *It includes the empty set ($\emptyset \in \tau$);*
2. *It is closed under complement ($A \in \tau \Rightarrow \overline{A} \in \tau$);*
3. *It is closed under countable union ($A_n \in \tau \Rightarrow \cup_n A_n \in \tau$).*

The pair (Ω, τ) is said to be a measurable space or a Borel space.

Definition 1.3. *A probability measure P is an application $P : \tau \to \mathbb{R}$ which respects the Kolmogorov axioms:*

1. $0 \leq P(A) \leq 1$
2. $P(\Omega) = 1$
3. $A_i \cap A_j \Rightarrow P(\cup_i A_i) = \sum_i P(A_i)$

Definition 1.4. *Let (Ω, τ, P) be a probability space and $(\mathbb{R}, \mathfrak{B})$ a measurable space, with $\mathfrak{B}$ the Borel σ-algebra of $\mathbb{R}$[1]. A real-valued random variable is a measurable function $X : \Omega \to \mathbb{R}$, which means that $X^{-1}(B) \in \tau\ \forall B \in \mathfrak{B}$ (it relates events with number values). We write the probability of the event x be in the interval B as $Pr\{x \in B\} = P_x(B) = P(X^{-1}(B))$.*

Definition 1.5. *The cumulative distribution function (CDF) of a real-valued random variable X is the function*

$$F_X(x) = Pr\{X(\omega) \leq x\} = P_x(]-\infty, x])) = P(X^{-1}(]-\infty, x])).$$

The probability density function (PDF) $p_X(x)$ is

$$p_X(x) = F'_X(x) \Leftrightarrow F_X(x) = \int_{-\infty}^{x} p_X(\xi) d\xi.$$

The expected value of a random variable X whose CDF admits a PDF $p_X(x)$ is

$$\mathbb{E}[X] = \int_{\mathbb{R}} x p_X(x) dx.$$

- Expected value of a function: $\mathbb{E}[f(X(\omega))] = \int f(x) p_X(x) dx = \int f(x) dF_X(x) = \int f(x) dP_X(i)$

Definition 1.6. *The characteristic function of a scalar random variable X is*[2]

$$\phi_X(u) = \mathbb{E}[e^{juX}] = \int e^{juX} p_X(x) dx$$

and the second characteristic function is

$$\psi_X(u) = \ln(\phi_X(u)).$$

[1] The Borel σ-algebra on $\mathbb{R}$ is the smallest σ-algebra containing all open sets on $\mathbb{R}$.
[2] It is the Fourier transform with sign reversal in the complex exponential.

Theorem 1.1. *Let $\mathfrak{B}$ be the Borel σ-algebra of $\mathbb{R}$ and ϕ a positive bounded measure on $(\mathbb{R}, \mathfrak{B})$. Then, there exists a unique positive integrable function defined in $g \in L^1$, up to a set of measure zero, and a unique singular measure ϕ_s on $(\mathbb{R}, \mathfrak{B})$ such that*

$$\phi(B) = \int_B g(x)dx + \phi_s(B), \ \forall B \in \mathfrak{B}.$$

- A measure ϕ_s is said to be singular if $\exists S \in \mathfrak{B}$ with $\mu(S) = 0$ and $\forall B \in \mathfrak{B}, \phi_s(B) = \phi_s(B \cap S)$.
- It is often possible to write

$$\phi(B) = \int \underbrace{g(x) \sum_k \mu_k \delta(x - s_k)}_{\text{density of the measure } \phi} dx.$$

- The probability density function is a particular case of such measure:

$$P(B) = \int \underbrace{g(x) \sum_k \mu_k \delta(x - s_k)}_{\text{PDF}} dx.$$

Definition 1.7. *A temporal distribution (temporal law) of a random variable X is*

$$F_X(x_1, ..., x_n, ..., t_1, ..., t_n, ...) = P(X(t_1) \leq x_1, ..., X(t_n) \leq x_n, ...)$$

Particular cases:

- An n-order distribution (n-order law) is $F_X(x_1, ..., x_n, ..., t_1, ..., t_n) = P(X(t_1) \leq x_1, ..., X(t_n) \leq x_n)$.
- A first-order distribution is $F_X(x, t) = P(X(t_1) \leq x_1) = F_{X(t,\omega)}(x_1)$.

Proposition 1.1. *Properties of temporal distributions:*

i. Symmetry: $F_X(x_1, ..., x_n, t_1, ..., t_n) = F_X(x_n, ..., x_1, t_n, ..., t_1)$.

ii. Consistency: $F_X(x_1, ..., x_n, ..., t_1, ..., t_n, ...) = F_X(x_1, ..., x_n, x_{n+1}, ..., x_m, t_1, ..., t_n, t_{n+1}, ..., t_m)$.

Definition 1.8. *Definitions of equivalence*[3]*:*

i. Two signals are said to be wide sense equivalent when they have the same temporal distribution.

ii. Two random signals $S_1(t, \omega)$ and $S_2(t, \omega)$ are said to be strictly equivalent if $P(S_1(t, \omega) = S_2(t, \omega)) = 1, \forall t$.

iii. If these two signals are such that $P(S_1(t, \omega) = S_2(t, \omega) \forall t \in T) = 1$, they are indistinguishable.

For example, consider the following signals:

$$S_1 \equiv 0, \quad S_2 = \begin{cases} 1 \text{ in a random value in } [0, 1] \\ 0 \text{ everywhere else.} \end{cases}, \quad S_3 = \begin{cases} 1 \text{ everywhere if a random value in } [0, 1] \text{ is } 0.5 \\ 0 \text{ everywhere, otherwise.} \end{cases}$$

S_1 and S_2 are strictly equivalent, but not indistinguishable. On the other hand, S_1 and S_3 are indistinguishable.

Point processes

Definition 1.9. *A point process is a continuous-time distribution of points on T. A counting process $N(t, \omega)$ can be used to count point process as it avoids to be equivalent to a zero process.*

Definition 1.10. *A Poisson process of intensity λ is a point process which may be defined in several ways:*

1. *It is a process that follows*
   (a) *The number of points in non-overlapping intervals are independent: $N(t_k, \omega) - N(t_{k-1}, \omega), ..., N(t_1, \omega) - N(t_0, \omega)$ are independent $\forall t_0 < ... < t_k$;*
   (b) *The probability of having exactly one event in a "small" interval is proportional to the length of the interval: $P(N(t + h) - N(t) = 1) = \lambda h + o(h)$;*
   (c) *The probability of having more than one event in a "small" interval is negligible: $P(N(t+h) - N(t) > 1) = o(h)$.*
2. *It is a process that follows*

[3] Presented in a particularity growing order.

(a) The number of points in non-overlapping intervals are independent: $N(t_k,\omega) - N(t_{k-1},\omega), .., .N(t_1,\omega) - N(t_0,\omega)$ *are independent* $\forall t_0 < ... < t_k$*;*

(b) $P(N(t+T) - N(t) = k) = e^{-\lambda T}\frac{(\lambda T)^k}{k!}$

3. *The intervals* T_i *between the occurrence of two events are i.i.d. (independent and identically distributed) with probability density function*

$$p(t) = \lambda e^{-\lambda t}, \quad t \geq 0.$$

Reminder:

- Independence: $P(N_1 \cap N_2) = P(N_1)P(N_2)$
- Uncorrelation: $\mathbb{E}[N_1 \cap N_2] = \mathbb{E}[N_1]\mathbb{E}[N_2]$

2 Partial characterisation of stochastic processes and temporal properties

Definition 2.1. *Important definitions:*

	Continuous-time	Discrete-time
First-order moment	$m(t) := \mathbb{E}[X(t,\omega)]$	$m[k] := \mathbb{E}[X[k,\omega]]$
Centred signal	$X_C(t,\omega) := X(t,\omega) - m(t)$	$X_C[k,\omega] := X[k,\omega] - m[k]$
Second-order moment[4] or (auto)correlation	$\gamma(t_1,t_2) := \mathbb{E}[X(t_1,\omega)X^*(t_2,\omega)]$	$\gamma[k_1,k_2] := \mathbb{E}[X[k_1,\omega]X^*[k_2,\omega]]$
Covariance	$c(t_1,t_2) := \gamma_{X_C}(t_1,t_2) = \mathbb{E}[X_C(t_1)X_C^*(t_2)]$	$c[k_1,t_2] := \gamma_{X_c}[k_1,k_2] = \mathbb{E}[X_C[k_1]X_C^*[k_2]]$
Variance	$\sigma(t) := \sqrt{c(t,t)} = \sqrt{\mathbb{E}[\lvert X_C(t)\rvert^2]}$	$\sigma[k] := \sqrt{c[k,k]} = \sqrt{\mathbb{E}[\lvert X_C[k]\rvert^2]}$
Power	$P(t) := \mathbb{E}[\lvert X(t)\rvert^2]$	$P[k] := \mathbb{E}[\lvert X[k]\rvert^2]$

Other types of (random) power:

- Random instant power: $|X(t,\omega)|^2$
- Random temporal mean power: $\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}|X(t,\omega)|^2 dt$

Proposition 2.1. *Properties:*

i. $c(t_1,t_2) = \gamma_X(t_1,t_2) - m(t_1)m^*(t_2)$

ii. $\sigma^2(t) = \mathbb{E}[|X(t,\omega)|^2] - |m(t)|^2$

Definition 2.2. *In the vector case, we have the vector of random variables*

$$\mathbf{X}(t,\omega) = \begin{bmatrix} X_1(t,\omega) \\ \vdots \\ X_n(t,\omega) \end{bmatrix},$$

the first-order moment

$$\mathbf{m}(t) = \mathbb{E}[\mathbf{X}(t,\omega)] = \begin{bmatrix} m_1(t) \\ \vdots \\ m_n(t) \end{bmatrix},$$

the second-order (correlation) matrix

$$\Gamma(t_1,t_2) = \mathbb{E}[\mathbf{X}(t_1,\omega)\mathbf{X}^\dagger(t_2,\omega)],$$

the cross-correlation

$$\Gamma_{X_1,X_2}(t_1,t_2) = \mathbb{E}[X_1(t_1)X_2^*(t_2)]$$

and the covariance matrix

$$\mathbf{c}(t_1,t_2) = \mathbb{E}[\mathbf{X_C}(t_1,\omega)\mathbf{X_C^\dagger}(t_2,\omega)] = \Gamma(t_1,t_2) - \mathbf{m}(t_1)\mathbf{m}^\dagger(t_2).$$

[4] Generalisation: n-order moment (it is not unique as there is no rule for the application of complex conjugate).

$$m(t_1,...,t_n) = \mathbb{E}[X(t_1),...,X(t_n)], \quad n \geq 3.$$

A fact about real random variables The moments are coefficients of a series expansion of the characteristic function ϕ_X. Let $m_n = \mathbb{E}[X^n]$:

$$\phi_X(u) = \mathbb{E}[e^{juX}] = \mathbb{E}\left[\sum_k j^k u^k \frac{X^k}{k!}\right] = \sum_k j^k \frac{u^k}{k!}\mathbb{E}\left[X^k\right] = \sum_k j^k \frac{m_n}{k!} u^k$$

Proposition 2.2. *Properties:*

i. The autocorrelation $\gamma(t_1, t_2)$ exists $\forall (t_1, t_2) \in T^2$ if and only if $\gamma(t, t) = \mathbb{E}[|X(t)|^2]$ exists ($< \infty$) $\forall t \in T$.

ii. The autocorrelation $\gamma(t_1, t_2) = \mathbb{E}[X(t_1)X^*(t_2)]$ defines a pseudo inner product.

iii. The Cauchy-Schwarz inequality holds: $|\gamma(t_1, t_2)|^2 \leq \gamma(t_1, t_1)\gamma(t_2, t_2)$.

iv. Existency of $\gamma(t_1, t_2)$ implies existency of $m(t) = \mathbb{E}[X(t)]$.

v. $\gamma(t_1, t_2)$ is a non-negative definite function (NND), i.e., $\sum_i \sum_j \lambda_i \lambda_j^* \gamma(t_i, t_j) \geq 0,\ \forall \lambda_i, t_i \in \mathbb{C}^n \times T^n$.

Proof.

ii. The expectation defines a pseudo inner product: $\langle X, Y\rangle := \mathbb{E}[XY^*]$, i.e., it is symmetric, bilinear and almost surely positive-definite: $\langle X, X\rangle = 0 \Leftrightarrow P(X = 0) = 1$. We can then define the autocorrelation inner product by setting $X = X(t_1)$ and $Y = X(t_2)$.

iii. Apply the Cauchy-Schwarz inequality to the inner product defined above.

$$\mathbb{E}[XY^*] \leq \mathbb{E}[|X|^2]\mathbb{E}[|Y|^2]$$

iv. Applpy the expectation Cauchy-Schwarz inequality to $X = X(t)$ and $Y = 1$:

$$|\mathbb{E}[X(t) \cdot 1]|^2 \leq \mathbb{E}[|X(t)^2|]\mathbb{E}[1] = \gamma(t, t).$$

v. Take $Z(t_i) = \sum_i \lambda_i X(t_i)$ and calculate

$$\mathbb{E}[|Z|^2] = \mathbb{E}\left[\left(\sum \lambda_i X(t_i)\right)\left(\sum \lambda_j X(t_j)\right)^*\right] = \sum\sum \lambda_i \lambda_j^* \mathbb{E}\left[X(t_i)X^*(t_j)\right] \geq 0.$$

□

Stationarity

Definition 2.3. *A random process $X(t)$ is said to be strict sense stationary (SSS) if its temporal law is invariant by time shift, i.e.*

$$F_X(x_1, ..., x_n, t_1, ..., t_n) = F_X(x_1, ..., x_n, t_1 + h, ..., t_n + h) \quad \forall n \forall x_i \forall t_i \forall h.$$

Definition 2.4. *A random process is said to be stationary of order n if its moments up to order n are stationary.*

- *In particular, a stationary process of order 1 is such that them mean $m(t)$ is constant, i.e.,*

$$m(t) = m(t + h)\ \forall t \forall h.$$

- *A stationary process of order 2 or wide sense stationary (WSS) has $m(t)$ constant and also*

$$\gamma(t_1, t_2) = \gamma(t_1 + h, t_2 + h)\ \forall t \forall h.$$

This means that $\gamma(t_1, t_2) = \gamma(t, t + \tau)$ depends only on $\tau = t_2 - t_1$. We can write

$$\gamma(\tau) = \mathbb{E}[X(t + \tau), X^*(t)] = \mathbb{E}[X(t), X^*(t - \tau)].$$

Definition 2.5. *Two random processes $X(t, \omega)$ and $Y(t, \omega)$ are said to be jointly stationary (of order 2) if $\begin{bmatrix} X(t, \omega) \\ Y(t, \omega) \end{bmatrix}$ is stationary (of order 2). Then, $\gamma_X = \gamma_X(\tau)$, $\gamma_Y = \gamma_Y(\tau)$, $\gamma_X Y = \gamma_X Y(\tau)$, $m_Y(t) = m_Y$ and $m_Y(t) = m_Y$.*

Definition 2.6. *A random process $X(t, \omega)$ is said to be cyclostationary if there exists T such that*

$$\mathbb{E}[X(t + kT) = \mathbb{E}[X(t)]] \text{ and } \gamma(t_1 + kT, t_2 + kT) = \gamma(t_1, t_2) \quad \forall k.$$

- If $X(t)$ is cyclostationary, then $Y(t) = X(t + \mathcal{O}(\omega))$, with $\mathcal{O}$ uniformly distributed over $[0, T]$ is stationary.

Properties of WSS signals[5]

Proposition 2.3. *Basic properties. Let $X(t,\omega)$ be a random process. Then:*

i. $\gamma(0) \geq |\gamma(\tau)|\ \forall\tau$

ii. $\gamma(0) = P(t)$

iii. $\gamma(\tau) = \gamma^*(-\tau)$

Proof.

i. Use the Cauchy-Schwarz inequality for expectation inner product:
$$|\mathbb{E}[X(t+\tau)X^*(t)]|^2 \leq \mathbb{E}[|X(t+\tau)|^2]\mathbb{E}[|X(t)|^2] \Rightarrow |\gamma(\tau)|^2 \leq \gamma(0)\gamma(0) \Rightarrow |\gamma(\tau)| \leq \gamma(0).$$

ii. By definition, $\gamma(0) = \mathbb{E}[X(t+0)X^*(t)] = \mathbb{E}[|X(t)|^2] = P(t) = P$.

iii. $\gamma(\tau) = \mathbb{E}[X(t+\tau)X^*(t)] = \mathbb{E}[X(t)X^*(t-\tau)] = \mathbb{E}[X(t-\tau)X^*(t)]^* = \gamma^*(-\tau)$.

□

Proposition 2.4. *Basic properties for vector case:*

i. $\Gamma(\tau) = \mathbb{E}[\mathbf{X}(t+\tau)\mathbf{X}^*(t)] = \Gamma^\dagger(-\tau)$

ii. If $\Gamma(0)$ is hermitian, then it is orthogonally diagonalisable and has real eigenvalues.

Proposition 2.5. *Periodicity:*

i. If $\gamma(0) = \gamma(\tau_1)$ with $\tau_1 \neq 0$, then γ is τ_1-periodic.

ii. $\gamma(\tau)$ is τ_1-periodic $\Leftrightarrow \gamma(0) = \gamma(\tau_1) \Leftrightarrow P(X(t) = X(t+\tau_1)) = 1$.

Proposition 2.6. *If $\gamma(0) = \gamma(\tau_1) = \gamma(\tau_2)$ with $\frac{\tau_1}{\tau_2} \notin \mathbb{Q}$ and $\tau_1, \tau_2 \neq 0$, then $\gamma(\tau) = \gamma$ is constant (as long as γ is continuous).*

Proposition 2.7. *The autocorrelation $\gamma(\tau)$ is uniformely continuous if and only if $\gamma(\tau)$ is continuous at $\tau = 0$.*

Proposition 2.8. *For non-WSS case, $\gamma(t_1, t_2) \in C^0$ on the diagonal of T (i.e., $(t,t) \in T^2$) if and only if γ is continuous on every $(t_i, t_j) \in T^2$.*

Markov process

Definition 2.7. *A Markov process is a stochastic process whose future probabilities are determined by its most recent values, i.e.,*
$$P(X(t_n) \in B_n | X(t_{n-1}) \in B_{n-1}, ..., X(t_1) \in B_1) = P(X(t_n) \in B_n | X(t_{n-1}) \in B_{n-1}), \quad t_n > t_{n-1} > \cdots > t_1.$$

In a Markov process, the n-order distribution $P(x_1, ..., x_n, t_1, ..., t_n)$ depends only on the second order distribution $P(x_1, x_2, t_1, t_2)$.

3 Power Spectral Density (PSD)

We shall consider WSS signals here.

Harmonic signals (...)

Expansion of Kawhuen-Loeve (...)

Continuous case

Theorem 3.1. *(Bochner) A function $\gamma : \mathbb{R} \to \mathbb{C}$ is continuous and non-negative definite if, and only if, there exists a positive bounded measure φ on $(\mathbb{R}, \mathfrak{B})$ such that*
$$\gamma(\tau) = \int_{-\infty}^{+\infty} e^{j2\pi f\tau} d\varphi(f) \quad \forall\tau \in \mathbb{R}.$$

Applying the Theorem 1.1 to autocorrelation, we have
$$\gamma(\tau) = \int_{-\infty}^{+\infty} e^{j2\pi f\tau}\Gamma(f) df$$
and $P = \gamma(0) = \int_{-\infty}^{+\infty} d\phi(f) = \int_{-\infty}^{+\infty} \Gamma(f) df$.

[5]All signals considered in this section are WSS, unless explicitly indicated.

Discrete case

Theorem 3.2. *A function $\gamma : \mathbb{Z} \to \mathbb{C}$ is non-negative definite if and only if there exists a positive bounded measure ϕ on $(\mathbb{R}, \mathfrak{B})$ such that*

$$\gamma[m] = \int_{-1/2}^{1/2} e^{j2\pi\nu m} d\phi(\nu) \quad \forall m \in \mathbb{Z}.$$

Applying an analogous theorem, we have

$$\gamma[m] = \int_{-1/2}^{+1/2} e^{j2\pi\nu m} \Gamma(\nu) d\nu$$

and $P = \gamma[0] = \int_{-\infty}^{+\infty} d\phi(\nu) = \int_{-\infty}^{+\infty} \Gamma(\nu) d\nu$.

Definition 3.1. *(Short definition)*
Continuous time:

$$\Gamma(f) = \mathcal{F}\{\gamma(\tau)\} = \int_{-\infty}^{+\infty} \gamma(\tau) e^{j2\pi f \tau} d\tau.$$

Discrete time:

$$\Gamma[\nu] = \mathcal{F}\{\gamma[k]\} = \sum_{k=-\infty}^{+\infty} \gamma[k] e^{j2\pi k \nu}.$$

Proposition 3.1. *Properties of the PSD:*

i. $\Gamma(f)$ *is real and positive*[6].

ii. *If the signal $X(t, \omega)$ is real, then its PSD $\Gamma(f)$ is even.*

iii. *Some PSD examples:*

$\gamma(\tau)$	$\Gamma(f)$
$\frac{N_0}{2}\gamma(\tau)$	$\frac{N_0}{2}$
$N_0 B \operatorname{sinc}(2\pi B \tau)$	$\frac{N_0}{2} \operatorname{rect}_{[-B,B]}(f)$
$N_0 B \operatorname{sinc}(2\pi B \tau) \cos(2\pi f_0 \tau)$	$\frac{N_0}{4} \left[\operatorname{rect}_{[-B,B](f-f_0)+\operatorname{rect}_{[-B,B]}}(f+f_0)\right]$

Sampling a signal

4 White noise

Definition 4.1. *A (wide sense) white signal is a signal for which the PSD is constant (WSS signal).*

Definition 4.2. *A (strict sense) white signal is a signal $X(t, \omega)$ such that*

i. $X(t, \omega)$ *is centred and WSS*

ii. $X(t_1, \omega)$ *and* $X(t_2, \omega)$ *are independent* $\forall t_1 \neq t_2$

Definition 4.3. *A band-limited white signal is a signal for which the support of the PSD is limited in frequency.*

[6] The Fourier transform of non-negative definite function in non-negative. In addition, if it had a negative part, we could filter it and get a negative signal, but it would have negative power, which is impossible!