

Signaux aléatoires

Caractéristiques statistiques au second ordre

Simon Leglaive

CentraleSupélec

Dans les épisodes précédents...

- De nombreux phénomènes, notamment physiques (càd issus du monde réel), nécessitent une modélisation aléatoire.
- Un signal aléatoire est une collection, potentiellement infinie, de variables aléatoires indexées par le temps. C'est une fonction de deux variables : l'aléa et le temps.
- De la même façon qu'on peut définir la distribution de probabilité d'une variable aléatoire, on peut définir la distribution d'un processus aléatoire.

Celle-ci est complètement définie par l'ensemble des lois fini-dimensionnelles à tout ordre, ce qu'on appelle la loi temporelle.

Programme

De la même manière que l'on peut définir des valeurs moyennes pour des (fonctions de) variables aléatoires, on va s'intéresser ici à certaines caractéristiques statistiques d'un signal aléatoire :

- Moyenne : moment d'ordre 1 ;
- Variance : moment d'ordre 2 centré ;
- Autocovariance : moment d'ordre 2 entre différents instants.

Nous nous arrêtons donc au second ordre.

Moyenne et variance

Moyenne d'une variable aléatoire (rappel)

Définition

- Soit une variable aléatoire (v.a.) $X \in E$. Sa moyenne est définie comme son espérance :

$$m_X = \mathbb{E}[X] = \begin{cases} \sum_{x \in E} x p(x) & \text{si } X \text{ est une v.a. discrète à valeur dans } E ; \\ \int_E x p(x) dx & \text{si } X \text{ est une v.a. continue à valeur dans } E. \end{cases}$$

avec $p(x)$ la fonction de masse / densité de probabilité de la v.a. dans le cas discret / continu.

- Soit un vecteur aléatoire $\mathbf{X} = [X_1, \dots, X_d]^\top$ à d dimensions. Son vecteur moyenne est défini par

$$\mathbf{m}_\mathbf{X} = \mathbb{E}[\mathbf{X}] = [\mathbb{E}[X_1], \dots, \mathbb{E}[X_d]]^\top.$$

- La moyenne est déterministe : nous avons sommé/intégré sur toutes les valeurs possibles.

Intuition

- Une variable aléatoire est un objet mathématique que l'on peut échantillonner pour obtenir une valeur, c'est ce qu'on appelle une réalisation ou un échantillon.
- On peut tirer un grand nombre de réalisations, et calculer la moyenne empirique. La moyenne d'une variable aléatoire est la valeur obtenue si on tire une infinité d'échantillons.
- Soit $\{x_i\}_{i=1}^N$ un ensemble de N réalisations i.i.d de la variable X . La moyenne empirique

$$\frac{1}{N} \sum_{i=1}^N x_i$$

converge (dans un certain sens) vers $\mathbb{E}[X]$ quand N tend vers l'infini.

Moyenne d'un signal aléatoire

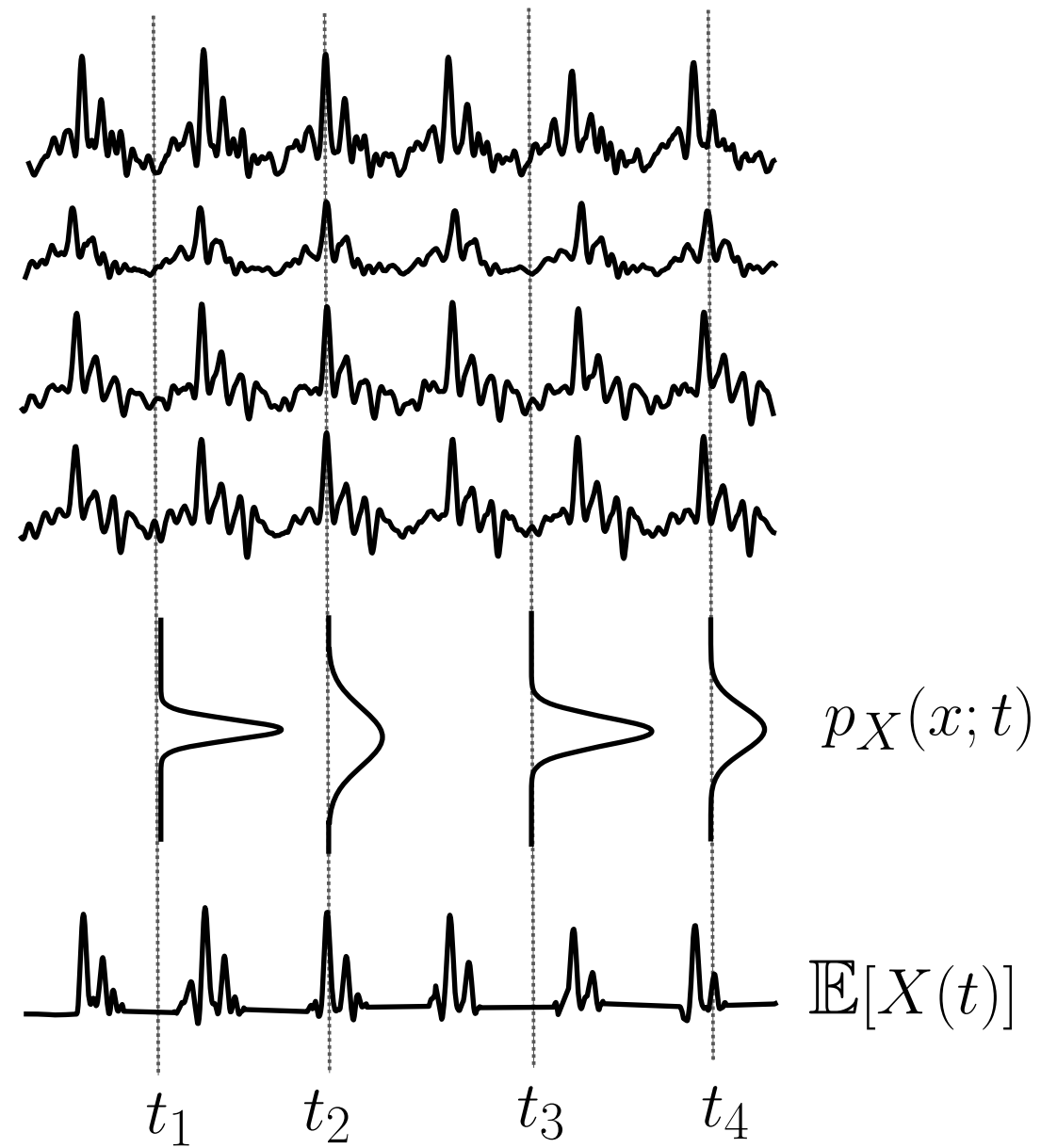
- La moyenne du processus aléatoire $\{X(t) \in \mathbb{C}\}_t$ est définie par son espérance :

$$m_X(t) = \mathbb{E}[X(t)] \in \mathbb{C}.$$

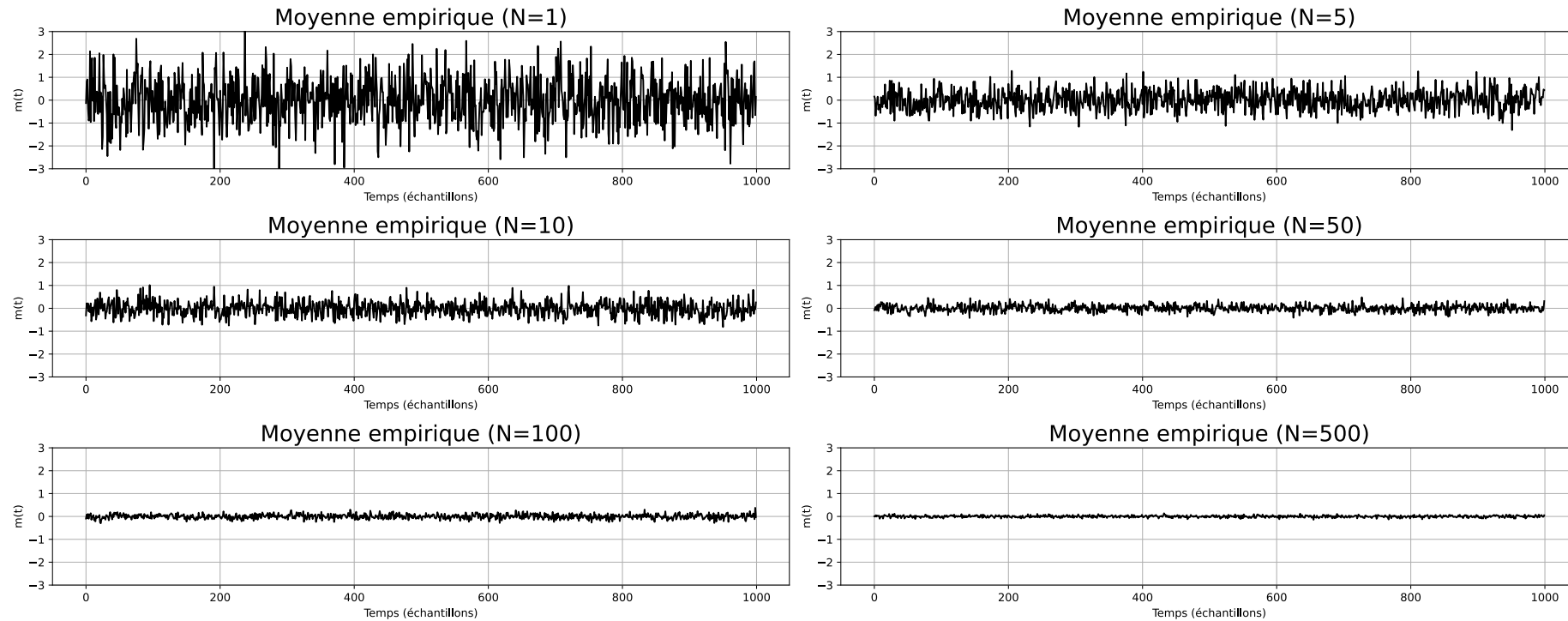
- Le vecteur moyenne du processus aléatoire $\{\mathbf{X}(t) \in \mathbb{C}^d\}_t$ est définie par

$$\mathbf{m}_X(t) = \mathbb{E}[\mathbf{X}(t)] \in \mathbb{C}^d.$$

On s'intéressera essentiellement par la suite aux signaux à valeurs continues, qu'on prendra complexes par soucis de généralité.



Moyenne empirique du bruit blanc pour différentes valeurs de N :



```
1 import numpy as np
2 import matplotlib.pyplot as plt
3
4 # Paramètres du signal
5 T = 1000 # Nombre d'échantillons temporels
6 N_values = [1, 5, 10, 50, 100, 500] # Différentes valeurs de N (nombre de réalisations)
7
8 # Création de la figure
9 plt.figure(figsize=(20, 8))
```

- Si la **moyenne est nulle** on dit que le signal / processus est **centré**.
- **Le processus et sa moyenne ont la même nature géométrique :**
 - scalaire ou vectoriel ;
 - discret, réel ou complexe.

En revanche, **la moyenne est déterministe**.

Variance d'une variable aléatoire (rappel)

La variance

$$\sigma_X^2 = \text{Var}(X) = \mathbb{E}[|X - m_X|^2] = \mathbb{E}[|X|^2] - |m_X|^2$$

d'une v.a. $X \in \mathbb{C}$ caractérise dans quelle mesure les valeurs qu'elle peut prendre s'écartent de la moyenne $m_X = \mathbb{E}[X]$.

Covariance (rappel)

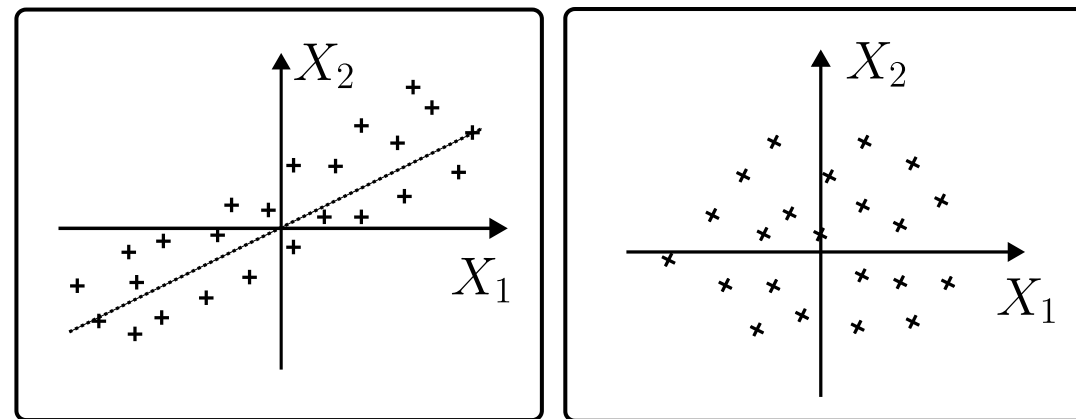
- On définit la covariance entre deux v.a. $X_1, X_2 \in \mathbb{C}$ de moyenne $m_1, m_2 \in \mathbb{C}$ par

$$\text{Cov}(X_1, X_2) = \mathbb{E}[(X_1 - m_1)(X_2 - m_2)^*].$$

- Intuitivement, la covariance caractérise comment, en espérance càd en moyenne statistique, varie une variable si on fait varier l'autre (après les avoir centré).
- Si $\text{Cov}(X_1, X_2) = 0$ les deux variables sont **décorrélées** (⚠ à ne pas confondre avec l'indépendance).

Covariance (intuition)

- Soient X_1 et X_2 deux v.a. réelles de moyenne nulle.
- On suppose qu'il existe un coefficient $a \in \mathbb{R}$ tel que $X_2 \approx aX_1$.
- On estime a au sens du minimum de l'erreur quadratique moyenne :



$$a^* = \arg \min_a \mathbb{E} \left[(X_2 - aX_1)^2 \right] = \frac{\text{Cov}(X_1, X_2)}{\text{Var}(X_1)}.$$

Intuition: On choisit l'erreur quadratique entre X_2 et son approximation aX_1 . Il faut bien comprendre que ce sont deux variables aléatoires, donc des objets mathématiques sans valeur précise (à la différence d'une réalisation). L'erreur quadratique $(X_2 - aX_1)^2$ n'a donc pas de valeur précise non plus. Nous n'avons d'autre choix que de moyenner cette quantité sur l'ensemble des réalisations possibles du couple (X_1, X_2) , c'est-à-dire prendre l'espérance.

Coefficient de corrélation (rappel)

- On définit le coefficient de corrélation entre deux v.a. $X_1, X_2 \in \mathbb{C}$ par :

$$\rho = \frac{\text{Cov}(X_1, X_2)}{\sqrt{\text{Var}(X_1)}\sqrt{\text{Var}(X_2)}}.$$

- L'avantage du coefficient de corrélation est qu'il est normalisé donc insensible à tout effet d'échelle. En appliquant l'inégalité de Schwarz pour les v.a.¹ on montre que : $-1 \leq \rho \leq 1$.
- Deux v.a. dont le coef. de corrélation est :
 - nul sont dites décorrélées;
 - positif sont dites corrélées positivement (une augmentation de l'une entraîne une augmentation de l'autre);
 - négatif sont dites corrélées négativement (une augmentation de l'une entraîne une diminution de l'autre).

1. Soient X et Y deux v.a. de carré intégrable (i.e., $\mathbb{E}[|X|^2], \mathbb{E}[|Y|^2] < +\infty$), on a $|\mathbb{E}[XY^*]| \leq \sqrt{\mathbb{E}[|X|^2]} \sqrt{\mathbb{E}[|Y|^2]}$.

Matrice de covariance d'un vecteur aléatoire (rappel)

Soit un vecteur aléatoire $\mathbf{X} \in \mathbb{C}^d$ et H l'opérateur de transposition hermitienne (transposée du conjugué).

La matrice de covariance

$$\mathbf{\Sigma} = \mathbb{E} \left[(\mathbf{X} - \mathbf{m})(\mathbf{X} - \mathbf{m})^H \right] = \begin{pmatrix} \text{Var}(X_1) & \text{Cov}(X_1, X_2) & \cdots & \text{Cov}(X_1, X_d) \\ \text{Cov}(X_2, X_1) & \ddots & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(X_d, X_1) & \cdots & \cdots & \text{Var}(X_d) \end{pmatrix}$$

caractérise comment chaque dimension de $\mathbf{X}$ covarie avec toutes les autres.

De plus,

$$\mathbf{\Sigma} = \mathbb{E} [\mathbf{X}\mathbf{X}^H] - \mathbf{m}\mathbf{m}^H.$$

Une matrice de covariance est symétrique (hermitienne) et (semi-définie) positive.

Variance d'un signal aléatoire

La variance du processus aléatoire $\{X(t) \in \mathbb{C}\}_t$ est définie par

$$\sigma_X^2(t) = \mathbb{E}[|X(t) - m_X(t)|^2] = \mathbb{E}[|X(t)|^2] - |m_X(t)|^2$$

Puissance instantanée d'un signal aléatoire

La puissance instantanée du processus aléatoire $\{X(t) \in \mathbb{C}\}_t$ est définie par

$$P_X(t) = \mathbb{E}[|X(t)|^2]$$

Matrice de covariance d'un signal aléatoire multidimensionnel

La matrice de covariance du processus aléatoire multidimensionnel $\{\mathbf{X}(t) \in \mathbb{C}^d\}_t$ est définie par

$$\mathbf{\Sigma}(t) = \mathbb{E} \left[(\mathbf{X}(t) - \mathbf{m}(t))(\mathbf{X}(t) - \mathbf{m}(t))^H \right]$$

Dans la suite du cours on s'intéressera essentiellement aux processus aléatoires scalaires car

1. L'objectif n'est pas de vous assommer avec des définitions ;
2. Les concepts dans le cas multidimensionnel se déduisent assez naturellement de ceux définis dans le cas scalaire.

Fonctions de covariance

- La moyenne et la variance d'un processus aléatoire ont toutes deux été définies à un instant particulier t :

$$m_X(t) = \mathbb{E}[X(t)], \quad \sigma_X^2(t) = \mathbb{E}[|X(t) - m_X(t)|^2].$$

- Considérons maintenant deux instants t_1 et t_2 . Il est assez naturel de se demander comment le processus à l'instant t_1 **covarie avec lui même** à l'instant t_2 .

Cette question nous amène à définir la **fonction d'autocovariance** du processus, une quantité très importante pour décrire statistiquement un signal aléatoire.

Fonction d'autocovariance : définition

- On appelle fonction d'autocovariance du processus $\{X(t)\}_t$ la fonction définie par :

$$R_{XX}(t_1, t_2) = \mathbb{E}[X_c(t_1)X_c(t_2)^*],$$

où $X_c = X(t) - \mathbb{E}[X(t)]$ désigne le processus centré.

- On choisira parfois une autre paramétrisation, avec $t_1 = t + \tau$ et $t_2 = t$:

$$R_{XX}(t + \tau, t) = \mathbb{E}[X_c(t + \tau)X_c(t)^*].$$

- La variance du processus se déduit naturellement de la fonction d'autocovariance :

$$\sigma_X^2(t) = R_{XX}(t, t).$$

- Pour simplifier les notations et quand cela ne portera pas à confusions, on notera la fonction d'autocovariance simplement $R(t_1, t_2)$, sans indiquer XX en indice.

Fonction d'autocorrélation : définition

- On appelle fonction d'autocorrélation du processus $\{X(t)\}_t$ la fonction définie par :

$$C_{XX}(t_1, t_2) = \mathbb{E}[X(t_1)X(t_2)^*].$$

On remarque qu'il s'agit de la même définition que la fonction d'autocovariance, à une différence importante près : **le processus n'est pas centré**.

- La puissance instantanée du processus se déduit naturellement de la fonction d'autocorrélation :

$$P(t) = C_{XX}(t, t).$$

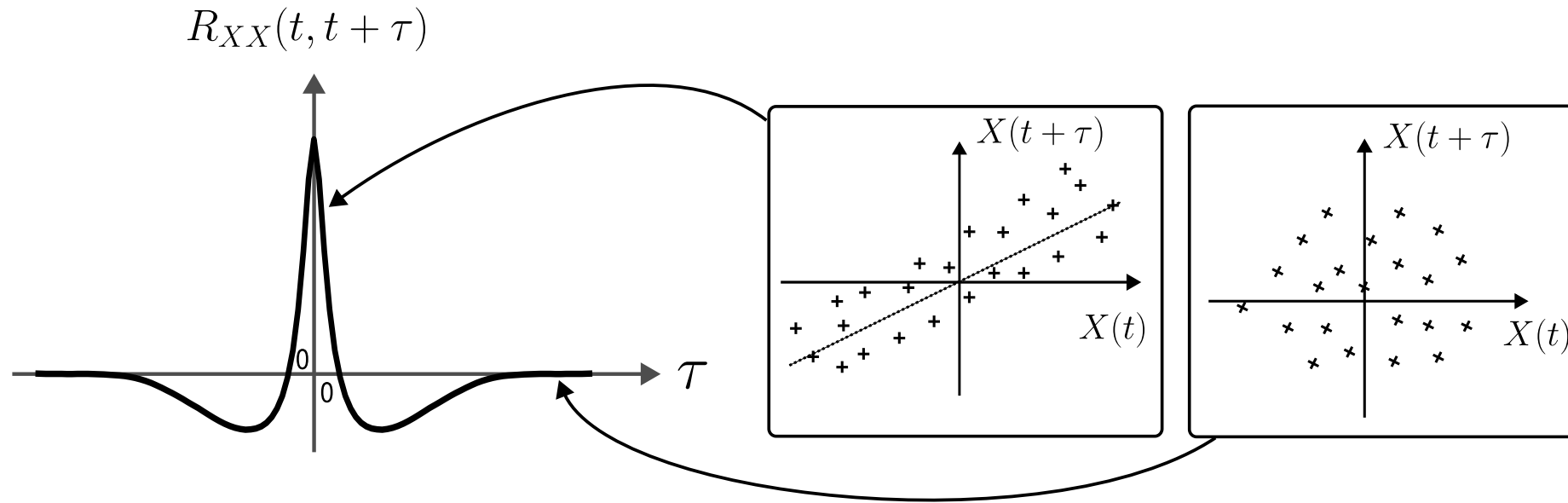
Interprétation en matière de prédictibilité ou d'aléa

- De nombreux problèmes en traitement du signal consistent à prédire la valeur future d'un signal (restauration de signal dégradé, suivi de trajectoire d'un objet, spéculation boursière, etc.)
- Considérons un signal aléatoire réel de moyenne nulle et cherchons à prédire $X(t + \tau)$ en fonction de $X(t)$. On suppose qu'il existe un coefficient $a(t, \tau) \in \mathbb{R}$ tel que $X(t + \tau) \approx a(t, \tau)X(t)$.
- Comme précédemment, on estime $a(t, \tau)$ au sens du minimum de l'erreur quadratique moyenne :

$$a(t, \tau)^* = \arg \min_{a(t, \tau)} \mathbb{E}[(X(t + \tau) - a(t, \tau)X(t))^2] = \frac{R_{XX}(t, t + \tau)}{\sigma_X^2(t)}.$$

- Le prédicteur optimal (au sens du min. de l'erreur quadratique moyenne) est donc donné par :

$$\hat{X}(t + \tau) = \frac{R_{XX}(t, t + \tau)}{\sigma_X^2(t)} X(t).$$



- La fonction d'autocovariance est donc directement liée au caractère prédictible (ou non-prédictible, donc aléatoire) d'un processus aléatoire, sous un modèle linéaire.
- Les valeurs non nulles de la fonction d'autocovariance $R_{XX}(t, t + \tau)$ montrent jusqu'à quel horizon τ cette prédiction est possible, et avec quelle précision.
- Plus $R_{XX}(t, t + \tau)$ est proche de 0, moins la prédiction est fiable.
- Dans le cas limite $R_{XX}(t, t + \tau) = 0$ la prédiction est impossible (avec le modèle linéaire).

Propriétés des fonctions d'autocovariance (voir exercice de TD)

La fonction d'autocovariance $R(t_1, t_2)$ d'un processus aléatoire $\{X(t)\}_t$ vérifie les propriétés suivantes :

1. Symétrie hermitienne :

$$R(t_1, t_2) = R(t_2, t_1)^*.$$

2. Inégalité de Schwarz :

$$|R(t_1, t_2)|^2 \leq R(t_1, t_1)R(t_2, t_2).$$

3. Non-négativité : Pour tout entier $k > 0$, pour toutes suites arbitraires d'instants $\{t_1, \dots, t_k\} \in \mathbb{Z}^k$ et de valeurs complexes $\{\lambda_1, \dots, \lambda_k\} \in \mathbb{C}^k$ on a :

$$\sum_{i,j=1}^k \lambda_i \lambda_j^* R(t_i, t_j) \geq 0.$$

Fonction d'intercovariance / intercorrélation : définition

- Pour l'étude de deux processus $\{X(t)\}_t$ et $\{Y(t)\}_t$, on définit également la fonction d'intercovariance par :

$$R_{XY}(t_1, t_2) = \mathbb{E}[X_c(t_1)Y_c(t_2)^*],$$

où $X_c = X(t) - \mathbb{E}[X(t)]$ et $Y_c = Y(t) - \mathbb{E}[Y(t)]$ désignent les processus centrés.

- La fonction d'intercorrélation est définie par :

$$C_{XY}(t_1, t_2) = \mathbb{E}[X(t_1)Y(t_2)^*].$$

Propriétés des fonctions d'intercovariance

La fonction d'intercovariance $R_{XY}(t_1, t_2)$ des processus aléatoires $\{X(t)\}_t$ et $\{Y(t)\}_t$ vérifie les propriétés suivantes :

1. Symétrie :

$$R_{XY}(t_1, t_2) = R_{YX}^*(t_2, t_1).$$

2. Inégalité de Schwarz :

$$|R_{XY}(t_1, t_2)|^2 \leq R_{XX}(t_1, t_1)R_{YY}(t_2, t_2).$$

Elle ne vérifie pas de propriété de positivité.

Espaces de signaux aléatoires

- Nous avons défini des quantités (moyenne, variance, fonctions de covariance) sans nous soucier des conditions de leur existence. Nous avons implicitement supposé que toutes ces quantités existent.
- En toute rigueur, ces quantités sont uniquement définies pour certaines classes de processus aléatoires.
- Par exemple :
 - La moyenne $\mathbb{E}[X(t)]$ existe si la variable $X(t)$ est de **module intégrable** pour tout t , c'est à dire $\mathbb{E}[|X(t)|] < +\infty$.
 - La variance et l'autocovariance existent si la variable $X(t)$ est de **carré intégrable** pour tout t , c'est à dire $\mathbb{E}[|X(t)|^2] < +\infty$.
- On pourrait donc définir la classe des processus aléatoires tels que telle ou telle condition est vérifiée.
- En pratique on supposera presque toujours des **processus du second ordre**.

Espace L^2

- On appelle **espace des variables aléatoires de carré intégrable** l'ensemble des variables admettant un moment (brut, non centré) d'ordre deux, càd tel que $\mathbb{E}[|X|^2] < \infty$.
- On note cet espace $L^2(\Omega, \mathcal{F}, \mathbb{P})$ pour faire apparaître l'espace probabilisé sur lequel la variable aléatoire est définie.
- L'espace $L^2(\Omega, \mathcal{F}, \mathbb{P})$ muni du produit scalaire

$$\langle X, Y \rangle = \mathbb{E}[XY^*]$$

est un espace de Hilbert.

- $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ implique $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$.

Si X est de carré intégrable alors elle est de module intégrable et donc son espérance $E[X]$ est bien définie (voir exercice de TD).

On en déduit que sa variance $\text{Var}(X) = \mathbb{E}[|X|^2] - |E[X]|^2$ est également bien définie.

Processus du second ordre

- On dit que le processus $\{X(t)\}_t$ défini sur $(\Omega, \mathcal{F}, \mathbb{P})$ et à valeur dans $\mathbb{C}$ est un **processus du second ordre** si $X(t)$ est de carré intégrable pour tout t .
- C'est un processus pour lequel tout ce qu'on vient de voir est bien défini 🙌.