Rudiments de mathématiques en

Signaux Aléatoires

Julian Tugaut

FISE1

Version du 20 février 2023

Table des matières

1 Introduction

L'idée de ce fascicule est de fournir une synthèse des rudiments mathématiques du cours "Signaux Aléatoires" à partir du cours fourni par Monsieur Alata.

Ainsi, celui-ci ne saurait se substituer au cours, aux CMs ou au polycopié. Il s'agit avant tout d'une ressource supplémentaire. De même, les aspects physiques ne sont pas l'objet de ce pdf. Il convient donc de se reporter au polycopié ou de demander au responsable de cours quant au point de vue pratique.

Il est important de comprendre que des choix ont été faits dans ce fascicule. Ainsi, certains éléments importants du cours ont été passés sous silence. Nous pensons toutefois avoir mis la lumière sur les éléments les plus pertinents (et simples) du point de vue mathématique.

Le cours de "Traitement des Signaux Déterministes", le cours de "Bases Indispensables des Mathématiques" (trigonométrie, nombres complexes, intégration et algèbre linéaire ainsi que séries numériques) ainsi que le cours de "Signaux et Systèmes Discrets" font partie des pré-requis indispensables à ce cours.

Ici, le nombre complexe de module 1 et d'argument $\frac{\pi}{2}$ est noté j. Et, le conjugué du complexe z se note z^*.

2 Des variables aléatoires aux signaux aléatoires

Dans cette section, on commence doucement avec quelques rappels avant d'introduire les personnages principaux du cours.

2.1 Définitions

Définition 2.1. *Une variable aléatoire réelle est une application de Ω dans $\mathbb{R}$. Une variable aléatoire complexe est une application de Ω dans $\mathbb{C}$.*

Remarque 2.2. *Bien qu'il n'y ait fondamentalement aucune différence entre le cas réel et le cas complexe, certains calculs en sont modifiés à la marge et il est essentiel d'être précautionneux à ce sujet.*

Définition 2.3. *Un signal aléatoire est une collection de variables aléatoires indexée par un ensemble d'indices $t \in \mathbb{R}$ ou $t \in \mathbb{Z}$ voire par $t \in \mathbb{R}^2$.*

Remarque 2.4. *Un signal aléatoire indexé par $t \in \mathbb{R}$ portera le nom de signal aléatoire continu tandis qu'un signal aléatoire indexé par $t \in \mathbb{Z}$ ou $t \in \mathbb{N}$ sera dit discret. Il est important de comprendre qu'un signal aléatoire discret n'est pas nécessairement une collection de variables aléatoires discrètes. Inversement, un signal aléatoire continu peut contenir des variables aléatoires discrètes.*

Pour rappel, une variable aléatoire X est discrète si $X(\Omega)$ est fini ou dénombrable. Sinon, on dit qu'elle est continue.

Ici, on parle de collection de variables aléatoires et les termes "discret" ou "continu" portent sur le mot "signal".

Notons enfin que si l'ensemble sur lequel est indexé notre signal est $\mathbb{R}^2$ (ou même $\mathbb{Z}^2$), on parlera plutôt de "champ aléatoire".

2.2 Loi d'une variable aléatoire

La loi d'une variable aléatoire caractérise la variabilité de celle-ci. La définition précise est dans le cours de "Probabilités et Statistiques". On ne la redonne pas ici vu qu'on ne l'utilisera pas en tant que telle ; la plupart du temps du moins.

Plutôt que de considérer la loi de la variable aléatoire (si on ne dispose que d'une seule variable), on s'intéresse à ses deux premiers moments. Définissons ceux-ci.

Définition 2.5. *L'espérance d'une variable aléatoire X se note $\mathbb{E}[X]$.*

Si X est discrète, alors $\mathbb{E}[X] = \sum_{k\in\mathcal{I}} x_k \mathbb{P}(X = x_k)$ où $X(\Omega) = \{x_k \ : \ k \in \mathcal{I}\}$; où $\mathcal{I}$ est fini ou dénombrable. Pour que X admette une espérance, il faut et il suffit que la quantité $\sum_{k\in\mathcal{I}} |x_k|\mathbb{P}(X = x_k)$ soit finie.

Si X est continue, salors $\mathbb{E}[X] = \int_{\mathbb{R}} x f_X(x)\mathrm{d}x$ où f_X est sa densité de probabilité. Pour que X admette une espérance, il faut et il suffit que l'intégrale $\int_{\mathbb{R}} |x| f_X(x)\mathrm{d}x$ soit finie.

On rappelle également la formule de transfert.

Théorème 2.6. *Soit ψ une fonction de $\mathbb{R}$ dans $\mathbb{R}$. Alors $\mathbb{E}[\psi(X)] = \int_{\mathbb{R}} \psi(x) f_X(x)\mathrm{d}x$ si le signal est continu et $\mathbb{E}[\psi(X)] = \sum_{k\in\mathcal{I}} \psi(x_k)\mathbb{P}(X = x_k)$ s'il est discret.*

On peut ainsi définir la notion de moment d'ordre n.

Définition 2.7. *Le moment d'ordre n d'une variable aléatoire X, s'il existe, est $\mathbb{E}[X^n]$.*

Définition 2.8. *Le moment centré d'ordre deux d'une variable aléatoire réelle est aussi appelé la variance. Il s'agit de $\mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$.*

Remarque 2.9. *Si la variable aléatoire X est complexe, sa variance est $\mathbb{E}\left[|X - \mathbb{E}[X]|^2\right] = \mathbb{E}[|X|^2] - |\mathbb{E}[X]|^2$. Il convient de garder à l'esprit que c'est la différence principale entre les variables aléatoires réelles et les variables aléatoires complexes.*

On s'intéresse également à l'entropie de Shannon. Si X est discrète, l'entropie de X est $\mathbb{E}[-\log(P_X(X))] = -\sum_{k\in\mathcal{I}} p_k \log(p_k)$ où $p_k := \mathbb{P}(X = x_k)$. Si X est continue, son entropie est $\mathbb{E}[-\log(f_X(X))] = -\int_{\mathbb{R}} f_X(x)\log(f_X(x))\mathrm{d}x$.

Bien que cette notion soit cruciale (je l'ai personnellement utilisée pour montrer la convergence en temps long de diffusions de McKean-Vlasov en paysage non-convexe), nous ne l'approfondirons pas ici.

2.3 Extension à deux variables aléatoires

Dans le paragraphe précédent, nous avons regardé comment étudier les variables aléatoires prises séparément.

Ici, nous verrons comment on étudie leurs relations entre elles.

Nous invitons les étudiants à revoir le cours sur les vecteurs aléatoires en “Probabilités et Statistiques” avant de lire cette partie.

Définition 2.10. *Soient p et q deux entiers positifs ou nuls. Alors le moment d'ordre (p, q) d'un couple de variables aléatoires (X, Y) est*

$$\mathbb{E}[X^p Y^q] \,.$$

Si (X, Y) est constitué de variables aléatoires discrètes alors

$$\mathbb{E}[X^p Y^q] = \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} x_i^p y_j^q \mathbb{P}(X = x_i, Y = y_j)$$

où $X(\Omega) = \{x_i \ : \ , i \in \mathcal{I}\}$ et $Y(\Omega) = \{y_j \ : \ , j \in \mathcal{J}\}$ avec $\mathcal{I}$ et $\mathcal{J}$ finis ou dénombrables.

Si au contraire le couple (X, Y) est à densité, alors $\mathbb{E}[X^p Y^q] = \iint_{\mathbb{R}^2} x^p y^q f_{(X,Y)}(x, y)\mathrm{d}x\mathrm{d}y$.

Définition 2.11. *La corrélation entre X et Y est le moment d'ordre $(1, 1)$ de (X, Y). En d'autres termes, c'est $\mathbb{E}[XY]$.*

Remarque 2.12. *Dans le cas de variables complexes, la corrélation est $\mathbb{E}[XY^*]$.*

Définition 2.13. *La covariance est le moment centré d'ordre $(1, 1)$, c'est-à-dire $\mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$. Et, si les variables sont complexes, c'est $\mathbb{E}[XY^*] - \mathbb{E}[X]\mathbb{E}[Y]^*$.*

2.4 Fonction d'autocorrélation

La fonction d'autocorrélation du signal aléatoire X en (s, t) donne la corrélation entre X_s et X_t. Plus formellement, on a :

Définition 2.14. *La fonction d'autocorrélation d'un signal $X := (X_t)_{t \in \mathcal{E}}$ (où $\mathcal{E} = \mathbb{R}$ ou $\mathcal{E} = \mathbb{Z}$) est la fonction de $\mathcal{E}^2$ dans $\mathbb{R}$ définie par*

$$r_{XX}(s, t) := \mathbb{E}[X_s X_t] \,.$$

Remarque 2.15. *Dans le cas complexe, on a $r_{XX}(s, t) := \mathbb{E}[X_s X_t^*]$ et donc la fonction est à valeurs complexes.*

La fonction d'autocovariance s'en déduit :

$$c_{XX}(s, t) := r_{XX}(s, t) - \mathbb{E}[X_s]\mathbb{E}[X_t] = \mathbb{E}[X_s X_t] - \mathbb{E}[X_s]\mathbb{E}[X_t] \,,$$

si X est un signal composé de variables aléatoires réelles et

$$c_{XX}(s, t) := r_{XX}(s, t) - \mathbb{E}[X_s]\mathbb{E}[X_t]^* = \mathbb{E}[X_s X_t^*] - \mathbb{E}[X_s]\mathbb{E}[X_t^*] \,,$$

si X est un signal composé de variables aléatoires complexes.

Dorénavant, on se place dans le cas complexe. En effet, le conjugué d'un réel étant lui-même, le cas réel s'ensuit du cas complexe.

On termine cette section en introduisant la fonction d'intercorrélation :

Définition 2.16. *Soient deux signaux X et Y indexés par un même ensemble $\mathcal{E}$. Alors, la fonction d'intercorrélation de X et Y est la fonction de $\mathcal{E}^2$ dans $\mathbb{C}$ définie par*

$$r_{XY}(s,t) := \mathbb{E}[X_s Y_t^*] .$$

On peut également définir l'intercovariance comme suit.

$$c_{XY}(s,t) := r_{XY}(s,t) - \mathbb{E}[X_s]\mathbb{E}[Y_t]^* = \mathbb{E}[X_s Y_t^*] - \mathbb{E}[X_s]\mathbb{E}[Y_t]^* .$$

3 Stationnarité

Dans cette section, nous allons introduire la notion de stationnarité. D'abord, on parlera des signaux avec des variables aléatoires IID (Indépendantes et Identiquement Distribuées). Puis on s'intéressera aux signaux SSS (Stationnaires au Sens Strict) et enfin aux signaux SSL (Stationnaires au Sens Large).

On note que IID implique SSS qui implique SSL. En effet, dans IID, on suppose l'indépendance des différentes variables aléatoires de la collection, ce qui n'est pas le cas pour SSS ou SSL. Il convient par ailleurs de commencer à se familiariser avec les différents acronymes.

Enfin, on terminera la section par l'introduction de la DSP (Densité Spectrale de Puissance) ; qui n'est rien d'autre que la transformée de Fourier d'une fonction qui dérive de la fonction d'autocorrélation dans le cas où le signal est SSL.

3.1 IID

Un signal $X = (X_t)_{t\in\mathcal{E}}$ est constitué de variables aléatoires IID si pour tout $s \in \mathcal{E}$ et pour tout $t \in \mathcal{E}$, la loi de X_s est identique à la loi de X_t et de plus les variables aléatoires X_s et X_t sont indépendantes si $s \neq t$.

3.2 SSS

Définition 3.1. *Un signal aléatoire X est SSS si les caractéristiques probabilistes ne changent pas pour peu que l'on translate le temps. En d'autres termes, pour tout $n \in \mathbb{N}$, pour tous $t_1, \cdots, t_n \in \mathcal{E}^n$, la loi de $(X_{t_1}, \cdots, X_{t_n})$ est la même que celle de $(X_{t_1-t_0}, \cdots, X_{t_n-t_0})$.*

En particulier, pour tous $s,t \in \mathcal{E}$, la loi du couple (X_s, X_t) ne dépend que de $t - s =: \tau$. Par exemple, si $0 \in \mathcal{E}$, alors $\mathcal{L}\left((X_s,X_t)\right) = \mathcal{L}\left((X_0,X_\tau)\right)$. Ainsi, l'espérance de X_t est égale à l'espérance de X_0 et il en est de même pour la variance. De plus, $r_{XX}(t_1,t_2) = r_{XX}(t_1 - t_0, t_2 - t_0)$.

3.3 SSL

La stationnarité au sens stricte est parfois trop restrictive donc on regarde aussi la stationnarité au sens large.

Définition 3.2. *Un signal aléatoire X est SSL s'il est stationnaire pour les moments jusqu'au second ordre. Trois hypothèses doivent être satisfaites :*
— *L'espérance de X_t ne dépend pas de t.*
— *L'espérance de $|X_t|^2$ est finie et ne dépend pas de t.*
— *La fonction d'autocorrélation r_{XX} est telle qu'il existe une fonction r^s_{XX} de $\mathcal{E}$ dans $\mathbb{C}$ satisfaisant $r_{XX}(t_1, t_2) = r^s_{XX}(\tau)$ où $\tau := t_2 - t_1$.*

Dans ce cas, on a également $c_{XX}(t_1, t_2) = r_{XX}(t_1, t_2) - \mathbb{E}[X_{t_1}]\mathbb{E}[X_{t_2}]^* = r^s_{XX}(\tau) - |\mathbb{E}[X_0]|^2$.

En posant $c^s_{XX}(\tau) := r^s_{XX}(\tau) - |\mathbb{E}[X_0]|^2$, il s'ensuit

$$c_{XX}(t_1, t_2) = c^s_{XX}(\tau) \,.$$

3.4 DSP

Avant d'attaquer ce paragraphe, il est conseillé de revoir le cours de "Signaux et Systèmes Discrets".

Définition 3.3. *Soit X un signal SSL. Sa densité spectrale de puissance est la transformée de Fourier de la fonction r^s_{XX}. On la note $\widehat{R^s_{XX}}$.*

L'intérêt de la DSP réside principalement dans le filtrage des signaux aléatoires ; que nous verrons subséquemment.

4 Ergodicité

Comprendre la notion d'ergodicité dans sa généralité recquiert de maîtriser celle de tribu et celle d'espérance conditionnelle. Néanmoins, ici, nous présentons un cas particulier d'ergodicité ou plus exactement, nous nous focalisons surtout sur le cas où la tribu sous-jacente est triviale (ne contient que des évènements de probabilités 1 ou 0).

Pour commencer, on introduit la notion de statistiques temporelles.

Définition 4.1. *Soit un signal aléatoire X. Pour tout $m \in \mathbb{N}^*$, la statistique (ou moment) temporelle d'ordre m est :*

$$\overline{x} := \lim_{N \to +\infty} \frac{1}{2N} \sum_{n=-N}^{n=+N} x_n^m \,,$$

si X est un signal discret et

$$\overline{x} := \lim_{T \to +\infty} \frac{1}{2T} \int_{t=-T}^{t=+T} x_t^m \mathrm{d}t \,,$$

si X est est un signal continu.

En opposition à la notion de statistique temporelle, on a le moment d'ordre m à t fixé de X_t ; à savoir $\mathbb{E}[X_t^m]$.

On peut aussi définir une version "temporelle" de la fonction d'autocorrélation :

$$\overline{xx_{(t_1,t_2)}} := \lim_{N\to+\infty} \frac{1}{2N} \sum_{n=-N}^{n=+N} x_{t_1+n} x^*_{t_2+n}\,,$$

si le signal est discret et

$$\overline{xx_{(t_1,t_2)}} := \lim_{T\to+\infty} \frac{1}{2T} \int_{t=-T}^{t=+T} x_{t_1+t} x^*_{t_2+t} \mathrm{d}t\,,$$

s'il est continu.

L'idée de l'ergodicité est que, couplée avec la stationnarité, elle permet d'estimer les statistiques spatiales à partir des statistiques temporelles. En d'autres termes, plutôt que d'avoir à intégrer sur l'espace Ω sous-jacent ce qui est à la fois coûteux et parfois impossible, il suffit de regarder une trajectoire aléatoire pour reconstituer les statistiques spatiales du signal aléatoire.

Définir rigoureusement l'ergodicité n'est pas simple. On va donc plutôt définir l'hypothèse d'ergodicité forte, qui est plus simple à appréhender.

Définition 4.2. *Un signal aléatoire est dit fortement ergodique à l'ordre $m \in \mathbb{N}^*$ si pour tout $k \in [\![1; m]\!]$, la variable aléatoire $\overline{\mu_k}$ est une constante. De plus, la convergence de $\frac{1}{2N}\sum_{n=-N}^{n=+N} X_n^k$ (si le signal est discret) ou de $\frac{1}{2T}\int_{t=-T}^{t=+T} X_t^k \mathrm{d}t$ (si le signal est continu) vers $\overline{\mu_k}$ a lieu presque sûrement.*

Remarque 4.3. *On dit d'une suite de variables aléatoires $(Y_n)_n$ qu'elle converge presque sûrement vers une constante a si et seulement s'il existe un évènement $\Lambda \subset \Omega$ de probabilité 1 tel que pour tout $\omega \in \Lambda$, $Y_n(\omega)$ converge vers a.*

Il est essentiel de comprendre qu'un signal ergodique n'est pas forcément fortement ergodique. De même, l'ergodicité n'est pas la stationnarité et la stationnarité n'est pas l'ergodicité.

L'ergodicité signifie en effet qu'on peut appliquer une loi des grands nombres tandis que la stationnarité est plutôt liée à la loi du processus que l'on considère.

Dans la suite de cette section, on considère un signal SSL et fortement ergodique jusqu'à l'ordre 2. Alors, si l'on note $\mu_X = \mathbb{E}[X_t]$ pour tout t :

$$\mu_X = \overline{x}\,.$$

De même, on a

$$r^s_{XX}(\tau) = \overline{xx_{(t_0,t_0+\tau)}}\,,$$

pour tous $t_0, \tau \in \mathcal{E}$. Il convient de noter qu'on récupère la variance en considérant la quantité $r^s_{XX}(0) - \mu_x^2$.

5 Bruit Blanc

Un Bruit Blanc (BB) est un signal X SSL dont le premier moment vaut 0 et tel que

$$r^s_{XX}(\tau) = 0 \quad \text{si } \tau \neq 0\,.$$

On fera un abus de notation en considérant $r^s_{XX}(\tau) = \sigma_X^2 \delta(\tau)$ où δ est la distribution de Dirac. Ceci serait faux dans le cas analogique mais de manière générale, on ne travaille qu'en numérique si bien que confondre Kronecker et Dirac ne porte plus à confusion.

Il convient de noter qu'un BB n'est pas forcément un signal IID tandis qu'un signal IID et centré est toujours un BB. En effet, les variables sont décorrélées les unes des autres mais ne sont pas nécessairement indépendantes.

6 Filtrage de SA

Dans cette section, on met des signaux aléatoires en entrée de systèmes LIT (Linéaires et Invariants dans le Temps), voir le cours de "Signaux et Systèmes Discrets". La fonction de transfert sera notée $\widehat{H}$ et la réponse impulsionnelle est h.

Théorème 6.1. *Supposons que le signal d'entrée, X, soit SSL. On note $\mu_X := \mathbb{E}[X_t]$, $\sigma_X^2 := \mathrm{Var}[X_t]$ et $r^s_{XX}(\tau) := \mathbb{E}[X_t X^*_{t+\tau}]$. Alors le signal de sortie, Y, est aussi SSL. De plus :*

- $\mu_Y := \mathbb{E}[Y_t] = \mu_X \sum_{m\in\mathbb{Z}} h[m]$ *si le système est discret et* $\mathbb{E}[Y_t] = \mu_X \int_{\mathbb{R}} h(t)\mathrm{d}t$.
- $r^s_{YY}(\tau) = \left[\left(h * \overline{h}\right) * r^s_{XX}\right](\tau)$ *où* $\overline{h}(t) := h(-t)$.

En particulier, pour $\tau = 0$, $\mathrm{Var}[Y_t] = \left[\left(h * \overline{h}\right) * r^s_{XX}\right](0) - |\mu_Y|^2$.

Corollaire 6.2. *En mettant en entrée un Bruit Blanc B, on récupère un signal SSL en sortie dont la moyenne est nulle et dont la fonction d'autocorrélation est*

$$r^s_{YY}(\tau) = \sigma_B^2 \left(h * \overline{h}\right)(\tau)\,.$$

De manière plus générale, on dispose de la formule des interférences.

Théorème 6.3. *Soient deux signaux SSL, $X^{(1)}$ et $X^{(2)}$ mis en entrée de deux systèmes de réponses impulsionnelles respectives $h^{(1)}$ et $h^{(2)}$. Alors, l'intercorrélation de $Y^{(1)}$ et $Y^{(2)}$ est*

$$r^s_{Y^{(1)}Y^{(2)}}(\tau) = \left[\left(h^{(1)} * \overline{h^{(2)}}\right) * r^s_{X^{(1)}X^{(2)}}\right](\tau)\,,$$

où $\overline{h^{(2)}}(t) := h^{(2)}(t)$.

On en déduit le résultat suivant portant sur la relation entre la DSP d'entrée et celle de sortie.

Pour rappel, la TF transforme le produit de convolution en produit simple. Il vient :

$$\widehat{R^s_{YY}}(f) = \left|\hat{H}(f)\right|^2 \widehat{R^s_{XX}}(f)\,.$$

Il est ainsi aisé de comprendre pourquoi le module de la fonction de transfert est si intéressant et utilisé en traitement du signal.

7 Filtrage adapté

Dans la plupart des cas d'intérêt pratique, les signaux que l'on considère sont bruités.

On cherche ainsi à déterminer un filtre dont la RI permette de détecter un signal donné, de longueur K, noyé dans un bruit additif. Le RSB est alors utilisé en tant que ratio ratio signal utile sur bruit. Pour ce faire, on procède à ce que l'on appelle du filtrage adapté.

En entrée, on a $x + B$ où B est un BB dit additif et x est un signal déterministe et numérique. La réponse impulsionnelle est h. En sortie, on récupère ainsi :

$$h * x + h * B\,.$$

Le ratio que l'on cherche à maximiser est le suivant :

$$\mathrm{RSB}[k] := \frac{\left|(h * x)[k]\right|^2}{\mathbb{E}\left[\left|(h * B)[k]\right|^2\right]}\,.$$

Or, $\mathbb{E}\left[\left|(h * B)[k]\right|^2\right] = \sigma_B^2 \sum_{k\in\mathbb{Z}} |h[k]|^2$ d'après la formule des interférences. Puis, on a également :

$$\mathbb{E}\left[\left|(h * B)[k]\right|^2\right] = \sigma_B^2 \int_{-\frac{1}{2}}^{\frac{1}{2}} \left|\hat{H}(\nu)\right|^2 \mathrm{d}\nu\,.$$

Enfin, l'inégalité de Cauchy-Schwarz implique $\left|(h * x)[k]\right|^2 \leq \left(\sum_{n\in\mathbb{Z}} |h[n]|^2\right)\left(\sum_{n\in\mathbb{Z}} |x[n]|^2\right)$. De plus, le cas d'égalité est atteint à condition que l'on ait $h[n] = x[K - n]^*$ où K est le nombre d'échantillons pris.

Ainsi, pour maximiser, il faut se ramener à $h[n] := x[K - n]^*$.