
PROCESSUS GAUSSIENS

Master IMA 2ème année

Jean-Christophe BRETON

Université de LA ROCHELLE

Septembre–Décembre 2006

Table des matières

| | | |
|----------|--|-----------|
| 1 | Variable et vecteur Gaussiens | 1 |
| 1.1 | Variabes aléatoires | 1 |
| 1.1.1 | Définition | 1 |
| 1.1.2 | Espérance et variance probabilistes | 2 |
| 1.2 | Convergences probabilistes | 4 |
| 1.3 | Vecteurs aléatoires | 5 |
| 1.3.1 | Définition | 5 |
| 1.3.2 | Indépendance de variables aléatoires | 6 |
| 1.3.3 | Covariance et indépendance | 8 |
| 1.4 | Variabes gaussiennes | 8 |
| 1.4.1 | Vecteur gaussien | 9 |
| 1.5 | Théorèmes limites | 14 |
| 1.5.1 | Loi des grands nombres (LGN) | 14 |
| 1.5.2 | Théorème central limite (TCL) | 14 |
| 2 | Processus gaussien | 17 |
| 2.1 | Processus stochastique | 17 |
| 2.1.1 | Définition | 17 |
| 2.1.2 | Régularité des trajectoires | 18 |
| 2.1.3 | Exemples de propriétés en loi des processus. | 21 |
| 2.2 | Processus Gaussien | 21 |
| 2.3 | Exemples de processus gaussiens | 23 |
| 2.3.1 | Mouvement brownien | 23 |
| 2.3.2 | Pont Brownien | 24 |
| 2.3.3 | Processus d'Ornstein-Uhlenbeck | 25 |
| 2.3.4 | Brownien géométrique | 25 |
| 2.3.5 | Bruit blanc gaussien | 25 |
| 2.3.6 | Mouvement brownien fractionnaire | 26 |
| 3 | Mouvement brownien | 27 |
| 3.1 | Historique | 27 |
| 3.2 | Définition | 28 |
| 3.3 | Propriétés du MB | 29 |

| | | |
|----------|--|-----------|
| 3.3.1 | Propriétés en loi | 29 |
| 3.3.2 | Propriétés des trajectoires du MB | 29 |
| 3.4 | Équation de la chaleur | 33 |
| 3.4.1 | Origine physique | 33 |
| 3.4.2 | Origine mathématique | 34 |
| 3.5 | Versions multidimensionnelles du MB | 38 |
| 3.5.1 | Mouvement brownien multidimensionnel | 38 |
| 3.5.2 | Champ de Wiener-Chensov | 38 |
| 3.5.3 | Fonction brownienne de Lévy | 39 |
| 4 | Intégration stochastique | 41 |
| 4.1 | Variation quadratique | 41 |
| 4.2 | Formule d'Itô | 43 |
| 4.3 | Et l'intégrale d'Itô? | 46 |
| 4.4 | Équations différentielles stochastiques (EDS) | 46 |
| 4.4.1 | Équations différentielles et EDS | 46 |
| 4.4.2 | Exemples d'EDS | 47 |
| 5 | Mouvement brownien fractionnaire | 49 |
| 5.1 | Généralités sur l'autosimilarité | 49 |
| 5.2 | Définition du mouvement brownien fractionnaire | 51 |
| 5.3 | Régularité des trajectoires | 53 |
| 5.3.1 | Hölder | 53 |
| 5.3.2 | Exposant de Hölder | 55 |
| 5.4 | Représentations intégrales du mBf | 56 |
| 5.4.1 | Représentation en moyenne mobile | 56 |
| 5.4.2 | Représentation harmonisable | 57 |
| 5.5 | Généralisations multidimensionnelles du mBf | 60 |
| 5.5.1 | Drap brownien fractionnaire | 60 |
| 5.5.2 | Fonction brownienne fractionnaire de Lévy | 60 |
| 5.6 | Mouvement brownien multifractionnaire (mBm) | 61 |
| 5.6.1 | Définition et propriétés du mBm | 61 |
| 5.6.2 | Régularité des trajectoires | 62 |
| 5.6.3 | Simulation, estimation du mBm | 63 |
| 5.7 | Modélisation d'un signal | 64 |

Introduction

Ces notes de cours s'adressent à des étudiants Math-Info de Master 2. Elles sont largement inspirées de plusieurs sources.

Le chapitre 1 contient les rappels essentiels de variables et vecteurs aléatoires et est issu des notes de cours de L3 [JCB].

Le chapitre 2 est inspiré de la présentation de Youri Davydov pour son cours de DEA [Dav] à l'Université Lille 1 et de [Lif].

Le chapitre 3 est inspiré de [EGK] et de [Lif].

Le chapitre 4 sur l'intégration stochastique est un mélange de [EGK] et de [Dav].

Le chapitre 5 est inspiré de [ST] pour ce qui concerne le mBf mais aussi de [A-LV, LV, P-LV] pour les généralisations. Enfin la modélisation pour l'image provient essentiellement de [BCM, PPLV].

Chapitre 1

Variable et vecteur Gaussiens

Définition, propriétés, covariance, fonction caractéristique, TCL et généralisations.

1.1 Variables aléatoires

1.1.1 Définition

Définition 1.1.1 *Un espace de probabilité est un espace mesurable (Ω, \mathcal{F}) muni d'une mesure de probabilité \mathbb{P} , c'est à dire une mesure de masse totale $1 : \mathbb{P}(\Omega) = 1$.*

Les ensembles mesurables $A \in \mathcal{F}$ sont appelés les évènements (ou observables).

Définition 1.1.2 (variable aléatoire)

- *On appelle variable aléatoire (v.a.) toute application mesurable X d'un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ dans \mathbb{R} muni de la tribu borélienne $\mathcal{B}(\mathbb{R})$.*
- *On appelle loi de X la mesure de probabilité \mathbb{P}_X définie sur \mathbb{R} par*

$$\mathbb{P}_X(A) = \mathbb{P}(X \in A) = \mathbb{P}(\omega \in \Omega \mid X(\omega) \in A), \quad A \in \mathcal{B}(\mathbb{R}).$$

- *La v.a. X est discrète si elle est à valeurs dans un ensemble au plus dénombrable (en bijection avec une partie de \mathbb{N}) : $X(\Omega)$ est fini ou dénombrable (on peut compter ses éléments).*
- *La v.a. X est à densité, de densité f , si*

$$\mathbb{P}(X \in A) = \int_A f d\lambda, \quad \mathbb{P}(X \in [a, b]) = \int_a^b f(x) dx.$$

Définition 1.1.3 *La fonction caractéristique d'une v.a. X est la fonction de \mathbb{R} dans \mathbb{C}*

$$\phi_X(t) = \mathbb{E}[e^{itX}] = \int_{\Omega} e^{itX} d\mathbb{P} = \int_{\Omega} e^{itx} d\mathbb{P}_X(x). \quad (1.1)$$

Il s'agit de la transformée de Fourier de la loi \mathbb{P}_X de X . Cette fonction caractérise la loi de X .

Exemples. Une v.a. X suit la loi normale standard $\mathcal{N}(0, 1)$ si elle admet pour densité

$$t \mapsto \frac{1}{\sqrt{2\pi}} e^{-t^2/2}.$$

Elle a pour fonction caractéristique $\phi_X(t) = e^{-t^2/2}$. De façon générale, une v.a. X suit la loi normale $\mathcal{N}(m, \sigma^2)$ si elle admet pour densité

$$t \mapsto \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(t-m)^2}{2\sigma^2}\right).$$

Elle a pour fonction caractéristique $\phi_X(t) = \exp(imt - \sigma^2 t^2/2)$.

1.1.2 Espérance et variance probabilistes

Définition 1.1.4

Moment : Une v.a. $X : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow \mathbb{R}$ a un moment d'ordre $p \geq 1$ ssi

$$\mathbb{E}[|X|^p] = \int_{\Omega} |X|^p d\mathbb{P} < +\infty.$$

Espérance : Si X a un moment d'ordre 1, l'espérance d'une v.a. X est donnée par

$$\mathbb{E}[X] = \int_{\Omega} X(\omega) d\mathbb{P}(\omega) = \int X d\mathbb{P}.$$

Variance : Si X a un moment d'ordre 2, la variance de X est donnée par

$$\text{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2]. \quad (1.2)$$

Écart-type : L'écart-type d'une v.a. X qui admet une variance est donnée par $\sigma_X = \sqrt{\text{Var}(X)}$.

On note aussi $L^p(\Omega, \mathcal{F}, \mathbb{P}) = \{X : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow \mathbb{R} \mid \mathbb{E}[|X|^p] < +\infty\}$, c'est un espace vectoriel normé avec pour norme

$$\|X\|_p = (\mathbb{E}[|X|^p])^{1/p}.$$

Remarque 1.1.1 L'espérance d'une v.a. donne la valeur moyenne (au sens probabiliste) de la v.a. Sa variance (ou son écart-type) mesure la dispersion des valeurs de la v.a. autour de sa moyenne. Ou encore, l'écart-type donne l'écart moyen de la v.a. par rapport à sa moyenne.

Il est équivalent de dire que la variance de X est finie et que X admet un moment d'ordre 2 fini.

Proposition 1.1.1 (Fonction caractéristique et moments) *Si une v.a. X a un moment d'ordre p , alors sa fonction caractéristique ϕ_X est dérivable p fois et on a*

$$\phi_X^{(p)}(0) = i^p \mathbb{E}[X^p].$$

Démonstration : C'est une conséquence du théorème de dérivation sous l'espérance (par convergence dominée) avec une hypothèse de domination donnée par l'existence du moment d'ordre p . \square

Proposition 1.1.2 (Inégalité de Markov) *Soit X une v.a. positive avec un moment d'ordre 1 fini. On a pour tout $t > 0$:*

$$\mathbb{P}(X \geq t) \leq \frac{\mathbb{E}[X]}{t}.$$

Exemple. Une v.a. $X \simeq \mathcal{N}(m, \sigma^2)$ a pour espérance et variance

$$\mathbb{E}[X] = m, \quad \mathbb{E}[X^2] = \sigma^2.$$

En fait, une v.a. $X \simeq \mathcal{N}(m, \sigma^2)$ peut alors se définir comme une translatée et dilatée de $X_0 \simeq \mathcal{N}(0, 1)$ par

$$X = m + \sigma X_0.$$

Proposition 1.1.3 (Propriétés de la variance)

- $\text{Var}(X) \geq 0$.
- $\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$ (Formule de Koenig).
- $\text{Var}(aX) = a^2 \text{Var}(X)$.
- $\text{Var}(X + b) = \text{Var}(X)$ pour toute constante $b \in \mathbb{R}$.
- $\text{Var}(X) = 0$ ssi X est constante ps (et vaut alors $\mathbb{E}[X]$).

La variance est un opérateur quadratique non linéaire.

Définition 1.1.5 (Covariance) *Soient X, Y deux variables aléatoires avec des variances finies, on définit la covariance de X et de Y par*

$$\text{Cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y].$$

On a en effet

$$\begin{aligned} \text{Cov}(X, Y) &= \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] = \mathbb{E}[XY - Y\mathbb{E}[X] - X\mathbb{E}[Y] + \mathbb{E}[X]\mathbb{E}[Y]] \\ &= \mathbb{E}[XY] - \mathbb{E}[Y\mathbb{E}[X]] - \mathbb{E}[X\mathbb{E}[Y]] + \mathbb{E}[X]\mathbb{E}[Y] \\ &= \mathbb{E}[XY] - \mathbb{E}[Y]\mathbb{E}[X] - \mathbb{E}[X]\mathbb{E}[Y] + \mathbb{E}[X]\mathbb{E}[Y] \\ &= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]. \end{aligned}$$

Remarque : • $(X, Y) \mapsto \text{Cov}(X, Y)$ est une application bilinéaire symétrique.

- Si X ou Y est centrée alors $\text{Cov}(X, Y) = \mathbb{E}[XY]$.
- $\text{Cov}(X, X) = \text{Var}(X)$.

La variance Var est une forme quadratique sur $L^2(\Omega)$, d'application bilinéaire symétrique associée la covariance Cov .

Proposition 1.1.4 *Si X et Y sont deux v.a. avec des moments d'ordre 2 finis alors*

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2 \text{Cov}(X, Y) \quad (1.3)$$

$$|\text{Cov}(X, Y)| \leq \sqrt{\text{Var}(X) \text{Var}(Y)}. \quad (1.4)$$

Démonstration : Pour (1.3), il suffit de développer $\text{Var}(X + Y)$. Puis pour l'inégalité (1.4), on applique l'inégalité de Cauchy-Schwarz

$$\begin{aligned} |\text{Cov}(X, Y)| &= |\mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]| \\ &\leq \sqrt{\mathbb{E}[(X - \mathbb{E}[X])^2] \mathbb{E}[(Y - \mathbb{E}[Y])^2]} = \sqrt{\text{Var}(X) \text{Var}(Y)}. \end{aligned}$$

□

1.2 Convergences probabilistes

On rappelle brièvement les différentes convergences probabilistes.

Définition 1.2.1 *Soit $(X_n)_n$ une suite de v.a. et X une v.a. limite.*

Convergence ps. *On dit que $(X_n)_n$ converge presque sûrement vers X si la convergence est vraie avec une probabilité 1*

$$\mathbb{P}(\omega \in \Omega \mid \lim_{n \rightarrow +\infty} X_n(\omega) = X(\omega)) = 1.$$

On la note $X_n \xrightarrow{p.s.} X$.

Convergence en norme p . *On dit que $(X_n)_n$ converge en norme p vers X si $\|X_n - X\|_p \rightarrow 0$ quand $n \rightarrow +\infty$, c'est à dire*

$$\lim_{n \rightarrow +\infty} \mathbb{E}[|X_n - X|^p] = 0.$$

On la note $X_n \xrightarrow{L^p} X$.

Convergence en proba. *On dit que $(X_n)_n$ converge en probabilité vers X si*

$$\forall \varepsilon > 0, \lim_{n \rightarrow +\infty} \mathbb{P}(|X - X_n| \geq \varepsilon) = 0.$$

On la note $X_n \xrightarrow{\mathbb{P}} X$.

Convergence en loi. On dit que $(X_n)_n$ converge en loi vers X si

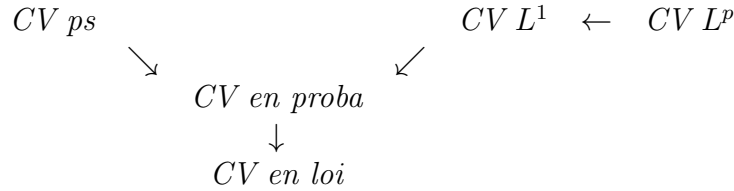
$$\mathbb{P}(X_n \in A) \rightarrow \mathbb{P}(X \in A), \quad n \rightarrow +\infty, \quad \text{pour tout } A \in \mathcal{B}(\mathbb{R})$$

On la note $X_n \Longrightarrow X$.

Théorème 1.2.1 (Paul Lévy) La convergence en loi des variables aléatoires est équivalente à la convergence simple des fonctions caractéristiques :

$$X_n \Longrightarrow X \quad \text{ssi} \quad \forall t \in \mathbb{R}, \phi_{X_n}(t) \rightarrow \phi_X(t).$$

Proposition 1.2.1 Les liens entre les différentes convergences sont donnés par le diagramme suivant :



1.3 Vecteurs aléatoires

1.3.1 Définition

Définition 1.3.1 On appelle vecteur aléatoire $X = (X_1, \dots, X_i, \dots, X_n)$ toute application de $(\Omega, \mathcal{F}, \mathbb{P})$ dans $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ mesurable. La v.a. X_i s'appelle la i ème marginale.

Définition 1.3.2 La loi du vecteur aléatoire X est la mesure image sur \mathbb{R}^n de la probabilité par X :

$$\mathbb{P}_X(A_1 \times \dots \times A_n) = \mathbb{P}(X \in A_1 \times \dots \times A_n) = \mathbb{P}(X_1 \in A_1, \dots, X_n \in A_n), A_i \in \mathcal{B}(\mathbb{R}), 1 \leq i \leq n.$$

C'est une mesure de probabilité dans $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$.

Définition 1.3.3 • Un vecteur aléatoire X est discret si l'ensemble de ses valeurs $X(\Omega)$ est discret dans \mathbb{R}^n .

- Un vecteur aléatoire X de \mathbb{R}^n est de loi à densité, de densité $f(x_1, \dots, x_n)$ si

$$d\mathbb{P}_X(x) = f(x_1, \dots, x_n) dx_1 \dots dx_n \iff \mathbb{P}_X(A) = \int_A f(x_1, \dots, x_n) dx_1 \dots dx_n, A \in \mathcal{B}(\mathbb{R}^n).$$

On vérifie sans peine que comme $\mathbb{P}_X(\mathbb{R}^n) = \mathbb{P}(X \in \mathbb{R}^n) = 1$, une densité en dimension n satisfait $f(x_1, \dots, x_n) \geq 0$ et

$$\int_{\mathbb{R}^n} f(x_1, \dots, x_n) dx_1 \dots dx_n = 1.$$

Définition 1.3.4 Soit $X = (X_1, \dots, X_n)$ un vecteur aléatoire, on définit sa fonction caractéristique comme la fonction de n variables $t = (t_1, \dots, t_n)$ de \mathbb{R}^n dans \mathbb{C} :

$$\phi_X(t) = \mathbb{E}[e^{i\langle t, X \rangle}] = \mathbb{E}[e^{i(t_1 X_1 + \dots + t_n X_n)}]$$

où $\langle t, x \rangle = t_1 x_1 + \dots + t_n x_n$ est le produit scalaire euclidien de \mathbb{R}^n .

Proposition 1.3.1 Si (X, Y) est un couple de loi $\mathbb{P}_{X,Y} = \mu$ alors les lois marginales \mathbb{P}_X et \mathbb{P}_Y de X et de Y s'obtiennent par

$$\mathbb{P}_X(A) = \mu(A \times \mathbb{R}), \quad \mathbb{P}(Y \in B) = \mu(\mathbb{R} \times B), \quad A, B \in \mathcal{B}(\mathbb{R}).$$

Démonstration : C'est évident si on remarque que $\{X \in A\} = \{(X, Y) \in A \times \mathbb{R}\}$. Idem pour la marginale Y . \square

De même si $X = (X_1, \dots, X_n)$ est de densité f , la i -ème marginale X_i est de densité

$$f_{X_i}(x_i) = \int_{\mathbb{R}^{n-1}} f(x_1, \dots, x_n) dx_1 \dots dx_{i-1} dx_{i+1} \dots dx_n.$$

1.3.2 Indépendance de variables aléatoires

Il s'agit d'une notion fondamentale en probabilité.

Définition 1.3.5 (Indépendance)

- Deux évènements $A, B \in \mathcal{F}$ d'un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ sont indépendants si

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B).$$

On note $A \perp B$.

- Deux va X et Y sont indépendantes ssi pour tout A, B , on a

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B).$$

On note encore $X \perp Y$.

En particulier, deux évènements A et B incompatibles ne peuvent pas être indépendants à moins que l'un des deux ne soit de probabilité nulle. Sinon $\mathbb{P}(A \cap B) = \mathbb{P}(\emptyset) = 0$, tandis que $\mathbb{P}(A)\mathbb{P}(B) > 0$. Il ne faut donc pas confondre les deux notions.

Proposition 1.3.2 Un vecteur aléatoire $X = (X_1, \dots, X_n)$ est à composantes indépendantes ssi sa loi \mathbb{P}_X est une loi produit (de ses lois marginales) :

$$\mathbb{P}_X = \mathbb{P}_{X_1} \otimes \dots \otimes \mathbb{P}_{X_n}. \tag{1.5}$$

Application. Si $X \perp Y$ alors (quand les espérances sont définies)

$$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$$

et plus généralement pour toutes fonctions mesurables f et g alors $\mathbb{E}[f(X)g(Y)] = \mathbb{E}[f(X)]\mathbb{E}[g(Y)]$.

En terme de densité, le critère d'indépendance s'énonce : soit (X, Y) un couple de densité $f_{(X,Y)}$ et de densité marginale $f_X(x)$ pour x , $f_Y(y)$ pour Y . Les v.a. X et Y sont indépendantes ssi

$$\forall x, y, \quad f_{(X,Y)}(x, y) = f_X(x)f_Y(y).$$

Remarque 1.3.1 Une conséquence importante : si on connaît les lois de X et de Y , des variables supposées **indépendantes**, on peut reconstruire la loi du couple (X, Y) à partir des marginales par (1.5). Ce n'est pas vrai en général quand X et Y ne sont pas indépendantes.

Théorème 1.3.1 Deux variables aléatoires X et Y sont indépendantes ssi

$$\phi_{(X,Y)}(t, s) = \phi_X(t)\phi_Y(s).$$

Démonstration : On raisonne par équivalence :

$$\begin{aligned} X \perp Y &\Leftrightarrow d\mathbb{P}_{X,Y}(dx, dy) = d\mathbb{P}_X(dx)d\mathbb{P}_Y(dy) \\ &\Leftrightarrow \int_{\mathbb{R}^2} e^{i(tx+sy)} d\mathbb{P}_{X,Y}(dx, dy) = \int_{\mathbb{R}^2} e^{i(tx+sy)} d\mathbb{P}_X(dx)d\mathbb{P}_Y(dy) \\ &\Leftrightarrow \int_{\mathbb{R}^2} e^{i(tx+sy)} d\mathbb{P}_{X,Y}(dx, dy) = \int_{\mathbb{R}} e^{itx} d\mathbb{P}_X(dx) \int_{\mathbb{R}} e^{isy} d\mathbb{P}_Y(dy) \\ &\Leftrightarrow \int_{\Omega} e^{i(tX+sY)} d\mathbb{P} = \int_{\Omega} e^{itX} d\mathbb{P} \int_{\Omega} e^{isY} d\mathbb{P} \\ &\Leftrightarrow \phi_{(X,Y)}(t, s) = \phi_X(t)\phi_Y(s) \end{aligned}$$

où on admet l'équivalence de la deuxième ligne (c'est la même que celle sur la caractérisation de la loi de X par ϕ_X) et où on a utilisé ensuite les théorèmes de Fubini, puis de transfert. \square

Proposition 1.3.3 Si deux variables aléatoires X et Y sont indépendantes alors

$$\phi_{X+Y}(t) = \phi_X(t)\phi_Y(t).$$

Plus généralement, pour n variables aléatoires X_1, \dots, X_n indépendantes on a

$$\phi_{X_1+\dots+X_n}(t) = \phi_{X_1}(t) \dots \phi_{X_n}(t).$$

Dans le cas à densité, on a le résultat suivant pour la somme de v.a. indépendantes

Proposition 1.3.4 Soit X, Y deux v.a. indépendantes et de densité f et g alors $X + Y$ est à densité, de densité

$$(f * g)(x) = \int_{-\infty}^{+\infty} f(y)g(x-y)dy.$$

1.3.3 Covariance et indépendance

Proposition 1.3.5 Soient X et Y deux vecteurs aléatoires indépendants de variances finies. Alors $\text{Cov}(X, Y) = 0$.

Démonstration : Par indépendance, on a $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$. D'où

$$\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = 0.$$

□

La réciproque est fautive : si X et Y sont de covariance nulle alors ils ne sont pas nécessairement indépendants. Cependant dans le cas de variables X, Y gaussiennes, on verra que la réciproque est vraie.

Pour la somme d'une variance, on déduit de la Prop. 1.3.5 et de (1.3) :

Corollaire 1.3.1 Si X et Y sont deux v.a. indépendantes avec des moments d'ordre deux alors

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y).$$

1.4 Variables gaussiennes

Définition 1.4.1 Une v.a. X suit la loi normale standard $\mathcal{N}(0, 1)$ si elle admet pour densité

$$t \mapsto \frac{1}{\sqrt{2\pi}} e^{-t^2/2}.$$

De façon générale, une v.a. X suit la loi normale $\mathcal{N}(m, \sigma^2)$ si elle admet pour densité

$$t \mapsto \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(t-m)^2}{2\sigma^2}\right).$$

Si $\sigma^2 = 0$, la loi est dégénérée, la v.a. X est constante égale à m . Sa loi est un dirac en m : $\mathbb{P}_X = \delta_m$.

Exercice. Normalisation de la loi normale $\int_{-\infty}^{+\infty} e^{-x^2/2} dx = \sqrt{2\pi}$.

Notons $I = \int_{-\infty}^{+\infty} e^{-x^2/2} dx$ et montrons que $I^2 = 2\pi$. On a

$$\begin{aligned} I^2 &= \int_{-\infty}^{+\infty} e^{-x^2/2} dx \times \int_{-\infty}^{+\infty} e^{-y^2/2} dy \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-x^2/2} e^{-y^2/2} dx dy = \int \int_{\mathbb{R} \times \mathbb{R}} e^{-(x^2+y^2)/2} dx dy \\ &= \int_0^{2\pi} \int_0^{+\infty} e^{-r^2/2} r dr d\theta \end{aligned}$$

$$= \int_0^{2\pi} d\theta \int_0^{+\infty} r e^{-r^2/2} dr = 2\pi \left[-e^{-r^2/2} \right]_0^{+\infty} = 2\pi$$

où on a utilisé le théorème de Fubini à la 2ème ligne puis on a fait un changement de variables en polaires à la 3ème ligne.

Proposition 1.4.1 Une v.a. X de loi $\mathcal{N}(m, \sigma^2)$ a pour

- Espérance : $\mathbb{E}[X] = m$
- Variance : $\text{Var}(X) = \sigma^2$
- Fonction caractéristique : $\phi_X(t) = \exp(imt - \sigma^2 t^2/2)$.

Proposition 1.4.2 Une v.a. $X \simeq \mathcal{N}(m, \sigma^2)$ peut se voir comme la translatée et la dilatée de $X_0 \simeq \mathcal{N}(0, 1)$ par

$$X = m + \sigma X_0.$$

Autrement dit si $X \simeq \mathcal{N}(m, \sigma^2)$, on définit la variable centrée réduite $\tilde{X} = \frac{X-m}{\sigma}$. Elle suit la loi $\mathcal{N}(0, 1)$.

Proposition 1.4.3 Soient $X_1 \simeq \mathcal{N}(m_1, \sigma_1^2)$ et $X_2 \simeq \mathcal{N}(m_2, \sigma_2^2)$ indépendantes. Alors $X_1 + X_2 \simeq \mathcal{N}(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$.

Démonstration : C'est une application, dans le cas gaussien, de la Prop. 1.3.4 (avec un peu de calcul). \square

1.4.1 Vecteur gaussien

Définition 1.4.2 Un vecteur aléatoire $X = (X_1, \dots, X_n)$ est gaussien ssi toutes les combinaisons linéaires de ses coordonnées $\langle a, X \rangle = a_1 X_1 + \dots + a_n X_n$ suivent une loi gaussienne dans \mathbb{R} (pour tout $a = (a_1, \dots, a_n) \in \mathbb{R}^n$).

Définition 1.4.3 La matrice de covariance d'un vecteur aléatoire $X = (X_1, \dots, X_n)$ est la matrice carrée symétrique, positive

$$K = (\text{Cov}(X_i, X_j))_{1 \leq i, j \leq n}.$$

L'espérance de $X = (X_1, \dots, X_n)$ est le vecteur des espérances de ses marginales

$$\mathbb{E}[X] = (\mathbb{E}[X_1], \dots, \mathbb{E}[X_n]).$$

Si $\mathbb{E}[X] = 0$, le vecteur X est dit centré.

Fonction caractéristique gaussienne en dimension n

Si $X = (X_1, \dots, X_n)$ est un vecteur gaussien alors $\langle a, X \rangle = \sum_{i=1}^n a_i X_i$ suit une loi normale de paramètres

$$\begin{aligned}\mathbb{E}[\langle a, X \rangle] &= \mathbb{E}[a_1 X_1 + \dots + a_n X_n] = a_1 \mathbb{E}[X_1] + \dots + a_n \mathbb{E}[X_n] = \langle a, \mathbb{E}[X] \rangle \\ \text{Var}(\langle a, X \rangle) &= \text{Var}(a_1 X_1 + \dots + a_n X_n) = \sum_{i,j=1}^n a_i a_j \text{Cov}(X_i, X_j) = a^t \text{Cov}(X) a.\end{aligned}$$

La v.a. $\langle a, X \rangle$ suit donc la loi $\mathcal{N}(\langle a, \mathbb{E}[X] \rangle, a^t \text{Cov}(X) a)$, sa fonction caractéristique est donnée par

$$\phi_{\langle a, X \rangle}(x) = \exp\left(ix \langle a, \mathbb{E}[X] \rangle - \frac{1}{2}(a^t \text{Cov}(X) a) x^2\right).$$

D'après les définitions des fonctions caractéristiques d'une v.a. et d'un vecteur aléatoire

$$\phi_X(x) = \mathbb{E}[e^{i\langle x, X \rangle}] = \phi_{\langle x, X \rangle}(1).$$

On en déduit :

Proposition 1.4.4 *La fonction caractéristique d'un vecteur gaussien $X = (X_1, \dots, X_n)$ est donnée par*

$$\begin{aligned}\phi_X(x) &= \exp\left(i\langle x, \mathbb{E}[X] \rangle - \frac{1}{2}(x^t \text{Cov}(X) x)\right) \\ &= \exp\left(i\langle x, \mathbb{E}[X] \rangle - \frac{1}{2}\langle x, \text{Cov}(X) x \rangle\right).\end{aligned}\tag{1.6}$$

Remarque :

- La loi d'un vecteur gaussien est connue dès qu'on a le vecteur moyenne $\mathbb{E}[X]$ et la matrice de covariance $\text{Cov}(X)$.
- On parle du vecteur gaussien standard en dimension n lorsque $\mathbb{E}[X] = 0$ et $\text{Cov}(X) = I_n$. Sa fonction caractéristique est alors

$$\phi_X(x) = \exp(-\langle x, x \rangle / 2) = \exp(-\|x\|^2 / 2).$$

- Pour un vecteur gaussien centré, $\mathbb{E}[X] = 0$ et on montre que

$$\langle x, \text{Cov}(X) x \rangle = \mathbb{E}[\langle x, X \rangle^2],$$

la fonction caractéristique devient donc

$$\phi_X(x) = \exp\left(-\frac{1}{2}\langle x, \text{Cov}(X) x \rangle\right) = \exp\left(-\frac{1}{2}\mathbb{E}[\langle x, X \rangle^2]\right).$$

– En prenant $x = (x_1, 0, \dots, 0)$, on a

$$\phi_X(x) = \phi_{X_1}(x_1) = \exp(i\mathbb{E}[X_1]x_1 - \text{Var}(X_1)x_1^2/2).$$

On retrouve que $X_1 \simeq \mathcal{N}(\mathbb{E}[X_1], \text{Var}(X_1))$. Plus généralement, pour tout $1 \leq i \leq n$, on a $X_i \simeq \mathcal{N}(\mathbb{E}[X_i], \text{Var}(X_i))$.

Comme pour les v.a. gaussiennes, on peut se ramener à un vecteur gaussien standard en centrant et en réduisant un vecteur gaussien quelconque non dégénéré. On a en effet :

Proposition 1.4.5 *Soit $X \simeq \mathcal{N}(m, K)$ un vecteur gaussien non dégénéré avec $m \in \mathbb{R}^n$ et K sa matrice de covariance. Alors*

$$\sqrt{K}^{-1}(X - m) \simeq \mathcal{N}(0, I_n). \quad (1.7)$$

Démonstration : Comme le vecteur X est non dégénéré, sa matrice de covariance K est définie (c'est à dire inversible). Il existe donc une matrice $A = \sqrt{K}$ inversible telle que $K = AA^t$. Il est donc légitime d'utiliser \sqrt{K}^{-1} dans (1.7).

On montre maintenant que $\tilde{X} = \sqrt{K}^{-1}(X - m)$ est gaussien, standard :

$$\begin{aligned} \phi_{\tilde{X}}(x) &= \mathbb{E}[\exp(i\langle x, \tilde{X} \rangle)] \\ &= \mathbb{E}[\exp(i\langle x, \sqrt{K}^{-1}(X - m) \rangle)] \\ &= \mathbb{E}[\exp(i\langle (\sqrt{K}^{-1})^t x, X - m \rangle)] \\ &= \mathbb{E}[\exp(i\langle (\sqrt{K}^{-1})^t x, X \rangle)] \times e^{-i\langle (\sqrt{K}^{-1})^t x, m \rangle} \\ &= \phi_X((\sqrt{K}^{-1})^t x) \times e^{-i\langle (\sqrt{K}^{-1})^t x, m \rangle} \\ &= \exp\left(i\langle m, (\sqrt{K}^{-1})^t x \rangle - \frac{1}{2}\langle (\sqrt{K}^{-1})^t x, K(\sqrt{K}^{-1})^t x \rangle\right) \times e^{-i\langle (\sqrt{K}^{-1})^t x, m \rangle} \\ &= \exp\left(-\frac{1}{2}\langle (\sqrt{K}^{-1})^t x, K(\sqrt{K}^{-1})^t x \rangle\right) \\ &= \exp\left(-\frac{1}{2}\langle x, \sqrt{K}^{-1}K(\sqrt{K}^{-1})^t x \rangle\right) \\ &= \exp\left(-\frac{1}{2}\langle x, x \rangle\right) = \exp(-\|x\|^2/2). \end{aligned}$$

On a donc bien $\tilde{X} \simeq \mathcal{N}(0, I_n)$. □

Remarque : Une v.a. $X \simeq \mathcal{N}(m, K)$ avec K inversible peut se voir comme une translatée et dilatée du vecteur gaussien standard $X_0 \simeq \mathcal{N}(0, I_n)$:

$$X \simeq \sqrt{K}X_0 + m.$$

Indépendance de variables gaussiennes

Proposition 1.4.6 *Soient (X, Y) un couple gaussien. Alors X et Y sont indépendantes ssi $\text{Cov}(X, Y) = 0$.*

Démonstration : Le sens direct est vrai quelque soit la loi de X et de Y . Pour la réciproque, on sait que X et Y sont indépendantes ssi $\phi_{(X,Y)}(t_1, t_2) = \phi_X(t_1)\phi_Y(t_2)$. Supposons le couple centré pour simplifier. Le vecteur gaussien (X, Y) a une matrice de covariance diagonale :

$$\begin{bmatrix} \sigma_X^2 & 0 \\ 0 & \sigma_Y^2 \end{bmatrix}$$

car $\text{Cov}(X, Y) = \text{Cov}(Y, X) = 0$. On déduit de (1.6) que

$$\phi_{(X,Y)}(t_1, t_2) = \exp\left(-\frac{1}{2}(t_1^2\sigma_X^2 + t_2^2\sigma_Y^2)\right) = \exp\left(-\frac{1}{2}t_1^2\sigma_X^2\right) \times \exp\left(-\frac{1}{2}t_2^2\sigma_Y^2\right) = \phi_X(t_1)\phi_Y(t_2),$$

ce qui justifie l'indépendance. □

De la même façon, pour des vecteurs :

Proposition 1.4.7 *Soit $(X_1, \dots, X_n, Y_1, \dots, Y_p)$ un vecteur gaussien de dimension $n + p$. Les deux vecteurs aléatoires $X = (X_1, \dots, X_n)$ et $Y = (Y_1, \dots, Y_p)$ sont indépendants ssi les covariances $\text{Cov}(X_i, Y_j)$, $1 \leq i \leq n$, $1 \leq j \leq p$ sont toutes nulles.*

Densité gaussienne en dimension n

Soit $X \simeq \mathcal{N}(0, I_n)$ un vecteur gaussien standard en dimension n . Comme $\text{Cov}(X) = I_n$, les marginales X_1, \dots, X_n sont toutes indépendantes. La loi du vecteur $X = (X_1, \dots, X_n)$ est donc la loi produit de ses marginales

$$\mathbb{P}_X = \mathbb{P}_{X_1} \otimes \dots \otimes \mathbb{P}_{X_n}.$$

En terme de densité, la densité de X est donnée par le produit tensoriel

$$\begin{aligned} f_X(x_1, \dots, x_n) &= f_{X_1}(x_1) \times \dots \times f_{X_n}(x_n) \\ &= \left(\frac{1}{\sqrt{2\pi}} e^{-x_1^2/2} \right) \times \dots \times \left(\frac{1}{\sqrt{2\pi}} e^{-x_n^2/2} \right) \\ &= \frac{1}{\sqrt{2\pi}^n} \exp(-(x_1^2 + \dots + x_n^2)/2). \end{aligned}$$

Proposition 1.4.8 *La densité d'un vecteur gaussien standard en dimension n est*

$$f_X(x) = \frac{1}{\sqrt{2\pi}^n} \exp(-(x_1^2 + \dots + x_n^2)/2).$$

Pour passer au cas général d'un vecteur gaussien $X \simeq \mathcal{N}(m, K)$ non dégénéré ($K = \text{Cov}(X)$ inversible), on va utiliser le vecteur réduit donné en (1.7) : on a $X \simeq \sqrt{K}X_0 + m$ avec $X_0 \simeq \mathcal{N}(0, I_n)$.

$$\begin{aligned}
\mathbb{P}(X \in A) &= \mathbb{P}(\sqrt{K}X_0 + m \in A) \\
&= \mathbb{P}(X_0 \in \sqrt{K}^{-1}(A - m)) \\
&= \int_{\sqrt{K}^{-1}(A-m)} f(x) dx \\
&= \int_A f(\sqrt{K}^{-1}(y - m)) \frac{dy}{\det \sqrt{K}} \\
&= \int_A \frac{\exp(-\|\sqrt{K}^{-1}(y - m)\|^2/2)}{((2\pi)^n \det K)^{1/2}} dy \\
&= \int_A \frac{\exp(-\langle(x - m), K^{-1}(x - m)\rangle/2)}{((2\pi)^n \det K)^{1/2}} dx
\end{aligned}$$

où on a fait à la 4ème ligne le changement de variable $y = \sqrt{K}x + m$. On a prouvé

Proposition 1.4.9 *La densité d'un vecteur gaussien $X \simeq \mathcal{N}(m, K)$ non dégénéré est*

$$f_X(x) = \frac{\exp(-\langle(x - m), K^{-1}(x - m)\rangle/2)}{((2\pi)^n \det K)^{1/2}}.$$

Variables gaussiennes et vecteurs non gaussiens

On a déjà vu que si un vecteur $\bar{X} = (X_1, \dots, X_n)$ est gaussien alors ses marginales X_i le sont aussi, de même les combinaisons linéaires de ses marginales le sont.

La réciproque est fautive : Si des variables aléatoires sont gaussiennes alors le vecteur formé par ces variables n'est pas nécessairement gaussien.

En effet, prenons X une v.a. de loi $\mathcal{N}(0, 1)$ et Y de loi donnée, pour $a > 0$ fixé, par

$$Y = \begin{cases} X & \text{si } |X| \leq a, \\ -X & \text{si } |X| > a. \end{cases}$$

Alors Y est de loi $\mathcal{N}(0, 1)$ en effet

$$\begin{aligned}
\phi_Y(t) &= \mathbb{E}[e^{itY}] = \mathbb{E}[e^{itX} \mathbf{1}_{|X| \leq a}] + \mathbb{E}[e^{-itX} \mathbf{1}_{|X| > a}] \\
&= \mathbb{E}[e^{itX} \mathbf{1}_{|X| \leq a}] + \mathbb{E}[e^{itX} \mathbf{1}_{-X > a}] = \mathbb{E}[e^{itX} \mathbf{1}_{|X| \leq a}] + \mathbb{E}[e^{itX} \mathbf{1}_{|X| > a}] \\
&= \mathbb{E}[e^{itX} (\mathbf{1}_{|X| \leq a} + \mathbf{1}_{|X| > a})] \\
&= \mathbb{E}[e^{itX}] = e^{-t^2/2}
\end{aligned}$$

car la loi de X est symétrique : $\mathcal{L}(X) = \mathcal{L}(-X)$.

Puis, la variable $X + Y$ est donnée par

$$X + Y = \begin{cases} X + X = 2X & \text{si } |X| \leq a \\ X - X = 0 & \text{si } |X| > a \end{cases}$$

$$= 2X\mathbf{1}_{|X|\leq a}.$$

La combinaison linéaire $X + Y$ ne suit donc pas une loi gaussienne. Le couple aléatoire (X, Y) n'est donc pas gaussien, sinon on aurait $X + Y$ de loi gaussienne.

De plus, cet exemple montre aussi que la dans la proposition 1.4.6, l'hypothèse (X, Y) gaussien est nécessaire et il ne suffit pas de supposer que X et Y sont des variables gaussiennes. En effet,

$$\begin{aligned} \text{Cov}(X, Y) &= \mathbb{E}[XY] = \mathbb{E}[X^2\mathbf{1}_{|X|\leq a}] - \mathbb{E}[X^2\mathbf{1}_{|X|>a}] \\ &= \mathbb{E}[X^2] - \mathbb{E}[X^2\mathbf{1}_{|X|>a}] - \mathbb{E}[X^2\mathbf{1}_{|X|>a}] \\ &= 1 - 2\mathbb{E}[X^2\mathbf{1}_{|X|>a}]. \end{aligned}$$

La fonction $u(a) = \mathbb{E}[X^2\mathbf{1}_{|X|<a}]$ tend vers 0 en $+\infty$ par convergence dominée, est continue et vaut $\mathbb{E}[X^2] = 1$ en 0. Il existe donc a tel que $u(a) = 1/2$ et $\text{Cov}(X, Y) = 0$.

Pourtant, X et Y sont non indépendantes sinon la loi du couple (X, Y) serait

$$\mathbb{P}_{(X,Y)} = \mathbb{P}_X \otimes \mathbb{P}_Y = \gamma_{0,1} \otimes \gamma_{0,1} = \gamma_{0,2},$$

qui est gaussienne, ce qui est faux. On a donc des variables gaussiennes X et Y non corrélées mais non indépendantes.

1.5 Théorèmes limites

Dans la suite i.i.d. signifiera indépendant(e)s et identiquement distribué(e)s, c'est à dire de même loi. On notera aussi souvent *va iid* pour variables aléatoires indépendantes et identiquement distribuées.

1.5.1 Loi des grands nombres (LGN)

Théorème 1.5.1 (LGN) *Soit $(X_n)_{n \geq 1}$ une suite de va iid avec un moment d'ordre un (i.e. $\mathbb{E}[|X_1|] < +\infty$, si bien que l'espérance est bien définie). Alors*

$$\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{ps} \mathbb{E}[X_1], \quad n \rightarrow +\infty.$$

La LGN montre que la moyenne arithmétique converge vers la moyenne probabiliste. C'est grâce à ce résultat qu'on peut estimer une proportion dans une population par une proportion dans un échantillon (représentatif). C'est la base de la théorie des sondages.

1.5.2 Théorème central limite (TCL)

On déduit du théorème de Paul Lévy un résultat fondamental en probabilité et en statistiques : le théorème central limite (TCL) dû à Paul Lévy.

Théorème 1.5.2 (TCL) Soit $(X_n)_n$ une suite de v.a.i.d., d'espérance m_1 et de variance finie σ^2 . Soit $S_n = X_1 + \dots + X_n$ la somme partielle. Alors quand $n \rightarrow +\infty$

$$\frac{S_n - nm_1}{\sigma\sqrt{n}} \Longrightarrow \mathcal{N}(0, 1).$$

Remarque 1.5.1 • Le TCL complète la loi des grands nombres. En effet, la LGN donne $S_n/n \rightarrow m_1$. Le TCL donne la vitesse de cette convergence.

• En plus dans le TCL, apparaît à la limite la loi $\mathcal{N}(0, 1)$ alors que les v.a. X_i sont de lois arbitraires : ce résultat justifie le rôle universel de la loi normale. Elle modélise les petites variations de n'importe quelle loi (avec un moment d'ordre 2) par rapport à sa moyenne.

Démonstration : Posons $Y_i = X_i - m_1$, si bien que les v.a. Y_i sont indépendantes de même loi avec $\mathbb{E}[Y_i] = 0$, $\text{Var}(Y_i) = \text{Var}(X_i)$. Notons $S'_n = Y_1 + \dots + Y_n$ et $Z_n = \frac{S_n - nm_1}{\sigma\sqrt{n}} = \frac{S'_n}{\sigma\sqrt{n}}$. On a

$$\begin{aligned} \phi_{Z_n}(t) &= \mathbb{E}[\exp\{it \frac{S'_n}{\sigma\sqrt{n}}\}] = \mathbb{E}[\exp\{i \frac{t}{\sigma\sqrt{n}} S'_n\}] = \phi_{S'_n}(\frac{t}{\sigma\sqrt{n}}) \\ &= \phi_{Y_1}(\frac{t}{\sigma\sqrt{n}}) \dots \phi_{Y_n}(\frac{t}{\sigma\sqrt{n}}) \\ &= \left(\phi_{Y_1}(\frac{t}{\sigma\sqrt{n}}) \right)^n \end{aligned}$$

en utilisant $\phi_{Y_1+\dots+Y_n} = \phi_{Y_1} \dots \phi_{Y_n} = \phi_{Y_1}^n$ par indépendance et identique distribution des variables Y_i .

Comme Y_1 a un moment d'ordre 2, ϕ_{Y_1} est dérivable 2 fois avec $\phi_{Y_1}(0) = 1$, $\phi'_{Y_1}(0) = i\mathbb{E}[Y_1] = 0$ et $\phi''_{Y_1}(0) = i^2\mathbb{E}[Y_1^2] = -\sigma^2$. La formule de Taylor à l'ordre 2 en 0 donne alors

$$\begin{aligned} \phi_{Y_1}(x) &= \phi_{Y_1}(0) + x\phi'_{Y_1}(0) + \frac{x^2}{2}\phi''_{Y_1}(0) + x^2\epsilon(x) \\ &= 1 - \frac{\sigma^2 x^2}{2} + x^2\epsilon(x) \end{aligned}$$

où la fonction ϵ vérifie $\lim_{x \rightarrow 0} \epsilon(x) = 0$. On a donc

$$\begin{aligned} \phi_{Z_n}(t) &= \left(\phi_{Y_1}(\frac{t}{\sigma\sqrt{n}}) \right)^n = \left(1 - \frac{\sigma^2 t^2}{2\sigma^2 \sqrt{n}^2} + \frac{t}{\sigma^2 \sqrt{n}^2} \epsilon(t/(\sigma\sqrt{n})) \right)^n = \left(1 - \frac{t^2}{2n} + \frac{1}{n} \epsilon(1/\sqrt{n}) \right)^n \\ &= \exp\left(n \ln\left(1 - \frac{t^2}{2n} + \frac{1}{n} \epsilon(1/\sqrt{n})\right)\right) = \exp\left(n\left(-\frac{t^2}{2n} + \frac{1}{n} \epsilon(1/\sqrt{n})\right)\right) \\ &= \exp\left(-\frac{t^2}{2} + \epsilon(1/\sqrt{n})\right). \end{aligned}$$

D'où pour chaque $t \in \mathbb{R}$,

$$\lim_{n \rightarrow +\infty} \phi_{Z_n}(t) = \exp\left\{-\frac{t^2}{2}\right\} = \phi_{\mathcal{N}(0,1)}(t).$$

Le théorème de Paul Lévy affirme alors que Z_n converge en loi vers $\mathcal{N}(0, 1)$, c'est la conclusion du TCL. \square

Remarque 1.5.2 En général, lorsque n est grand, on approxime la loi d'une somme de v.a.i.d de $L^2(\Omega)$ par une loi normale grâce au TCL de la façon suivante. Soit $S_n = X_1 + \dots + X_n$ la somme de v.a.i.d X_i avec $\sigma_X^2 < +\infty$, on a d'après le TCL

$$\frac{X_1 + \dots + X_n - n\mathbb{E}[X_1]}{\sigma\sqrt{n}} \Longrightarrow \mathcal{N}(0, 1).$$

Quand n est grand on approxime alors la loi de $\frac{X_1 + \dots + X_n - n\mathbb{E}[X_1]}{\sigma\sqrt{n}}$ par celle de $Y = \mathcal{N}(0, 1)$. Si bien que la loi de la somme $S_n = X_1 + \dots + X_n$ est approximée par celle de

$$n\mathbb{E}[X_1] + \sigma\sqrt{n}Y \simeq \mathcal{N}(n\mathbb{E}[X_1], \sigma^2 n).$$

Règle statistique : La somme S_n d'une suite de v.a.i.i.d. L^2 de moyenne m_1 et de variance σ^2 s'approxime par

$$S_n \approx \mathcal{N}(nm_1, \sigma^2 n).$$

Application. Comme une v.a. X_n de loi binomiale $\mathcal{B}(n, p)$ peut se voir comme la somme de n v.a. ϵ_i indépendantes de loi de Bernoulli $b(p)$, $X_n = \epsilon_1 + \dots + \epsilon_n$, la remarque précédente montre qu'on peut approcher la loi $\mathcal{B}(n, p)$ par la loi normale $\mathcal{N}(np, np(1-p))$.

Bien sûr, il existe des versions vectorielles de la LGN et du TCL.

Chapitre 2

Processus gaussien

2.1 Processus stochastique

2.1.1 Définition

Définition 2.1.1 *Un processus stochastique $X = (X_t)_{t \in T}$ est une famille de variables aléatoires X_t indexée par un ensemble T .*

En général $T = \mathbb{R}$ ou \mathbb{R}_+ et on considère que le processus est indexé par le temps t .

Si T est un ensemble fini, le processus est un vecteur aléatoire. Si $T = \mathbb{N}$ alors le processus est une suite de variables aléatoires. Plus généralement quand $T \subset \mathbb{Z}$, le processus est dit discret.

Pour $T \subset \mathbb{R}^d$, on parle de champ aléatoire (drap quand $d = 2$).

Un processus dépend de deux paramètres : $X_t(\omega)$ dépend de t (en général le temps) et de l'aléatoire $\omega \in \Omega$.

Pour $t \in T$ fixé, $\omega \in \Omega \mapsto X_t(\omega)$ est une variable aléatoire sur l'espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$.

Pour $\omega \in \Omega$ fixé, $t \in T \mapsto X_t(\omega)$ est une fonction à valeurs réelles, appelée trajectoire du processus. C'est un enjeu que de savoir si un processus admet des trajectoires mesurables, continues, dérivables ou encore plus régulières.

Dans la suite, sauf mention contraire, on prendra $T = \mathbb{R}$ ou \mathbb{R}_+ .

Définition 2.1.2 *Étant donné un processus stochastique $(X_t)_{t \in T}$, les lois fini dimensionnelles de X sont les lois de tous les vecteurs $(X_{t_1}, \dots, X_{t_n})$ pour $t_1, \dots, t_n \in T$ et $n \in \mathbb{N}$.*

L'ensemble des lois fini dimensionnelles caractérise la loi \mathbb{P}_X du processus X . Dans la suite quand nous écrirons $X \stackrel{\mathcal{L}}{=} Y$ égalité en loi de deux processus, nous signifierons l'égalité de toutes les loi fini dimensionnelles de X et de Y .

$$(X_{t_1}, \dots, X_{t_n}) \stackrel{\mathcal{L}}{=} (Y_{t_1}, \dots, Y_{t_n})$$

pour tout t_1, \dots, t_n et $n \in \mathbb{N}$.

Il y a plusieurs façons pour des processus stochastiques X et Y d'être égaux :

Définition 2.1.3 – Deux processus X et Y sont dits équivalents s'ils ont même loi (égalité de toutes les lois fini dimensionnelles). On écrira $X \stackrel{\mathcal{L}}{=} Y$.

- On dira que Y est une version du processus X si pour tout $t \in T$, $\mathbb{P}(X_t = Y_t)$. On parle encore d'équivalence au sens fort.
- Deux processus X et Y sont dit indistinguables si $\mathbb{P}(X_t = Y_t, \forall t \in T) = 1$.

Il est facile de voir que pour deux processus stochastiques X et Y :

Proposition 2.1.1 *indistinguishable \Rightarrow équivalence forte \Rightarrow équivalence.*

L'équivalence forte définit une relation d'équivalence pour les processus stochastiques : deux processus fortement équivalent sont équivalents pour cette relation.

2.1.2 Régularité des trajectoires

Souvent lorsqu'on considère un processus stochastique X , on en cherche une version Y dont les trajectoires ont de bonnes propriétés de régularité. Ce n'est pas toujours possible comme le montre l'exemple ci-dessous :

Exemple. Soit l'espace de probabilité $([0, 1], \mathcal{B}([0, 1]), \lambda)$ et $T = [0, 1]$. Considérons D la diagonale de $[0, 1] \times [0, 1]$ et définissons

$$X(t, \omega) = 0 \quad \forall (t, \omega), \quad Y(t, \omega) = \mathbf{1}_D(t, \omega).$$

Pour t fixé, on a $X(t, \omega) = 0$ et $Y(t, \omega) = 0$ pour $\omega \neq t$, 1 pour $\omega = t$. On a donc $X(t, \omega) = Y(t, \omega)$ pour tout $\omega \neq t$, c'est à dire ps. On a donc X, Y fortement équivalents (versions du même processus). Pourtant, les trajectoires de X sont continues tandis que celles de Y ne le sont pas.

Théorème 2.1.1 (Kolmogorov) *Soit $(X_t)_{t \in T}$ un processus tel qu'il existe $a, b, c > 0$ vérifiant pour tout s, t :*

$$\mathbb{E}[|X_t - X_s|^a] \leq c|t - s|^{1+b}. \quad (2.1)$$

Alors il existe une version continue \tilde{X} de X .

En fait, les trajectoires de \tilde{X} sont mêmes γ -höldériennes pour tout $\gamma < b/a$.

Remarque :

- La condition du théorème porte sur les lois de dimension 2 :

$$\mathbb{E}[|X_t - X_s|^a] = \int_{\mathbb{R}^2} |x - y|^a d\mathbb{P}_{(X,Y)}(dx, dy),$$

ce qui en pratique n'est pas trop difficile à calculer.

- A priori, dans le théorème, a et b sont non liés. En réalité, on peut toujours prendre $a \geq 1 + b$. En effet, si $a < 1 + b$, alors (2.1) se réécrit

$$E \left[\left| \frac{X_t - X_s}{t - s} \right|^a \right] \leq c |t - s|^{1+b-a}$$

avec $1 + b - a > 0$. En faisant $s \rightarrow t$, la dérivée dans le sens L^a de $(X_t)_t$ est nulle et $(X_t)_t$ est donc constant. Ce n'est donc pas très intéressant d'utiliser le théorème dans un tel cas. Puisque le processus initial est en fait constant, il est évident qu'il est aussi continue.

- La condition $b > 0$ est capitale : Pour $b = 0$, on a un contre exemple avec le processus de Poisson : Soit $X_t = \Pi_t - t$ où $(\Pi_t)_t$ est un processus de Poisson $\Pi_t \simeq \mathcal{P}(t)$, $\mathbb{E}[\Pi_t] = t$ et $\mathbb{E}[(\Pi_t - t)] = t = \text{Var}(\Pi_t)$. On a

$$\mathbb{E}[|X_t - X_s|^2] = \text{Var}(\Pi_t - \Pi_s) = \text{Var}(\Pi_{t-s}) = t - s.$$

On a donc (2.1) avec $a = 2$, $b = 0$ et $c = 1$. Or les trajectoires du processus de Poisson sont en escalier avec des sauts.

Démonstration : L'idée est de définir une suite de processus polygonaux par morceaux qui converge uniformément vers \tilde{X} , processus de trajectoire continue.

On considère la partition de $[0, 1]$ donnée par $t_{n,k} = k/2^n$, $0 \leq k \leq 2^n$. On définit $X_n(t)$ sur $[k/2^n, (k+1)/2^n]$ par interpolation linéaire entre $X(k/2^n)$ et $X((k+1)/2^n)$.

[Dessin des trajectoires typiques.]

On compare facilement $X_n(t)$ et $X_{n-1}(t)$: clairement, sur $[k/2^{n-1}, (k+1)/2^{n-1}]$, le sup de $X_n - X_{n-1}$ a lieu en $\frac{2k+1}{2^n}$. D'où

$$\begin{aligned} \|X_n - X_{n-1}\|_\infty &= \max_{1 \leq k \leq 2^{n-1}} \sup_{t \in [k/2^{n-1}, (k+1)/2^{n-1}]} |X_n(t) - X_{n-1}(t)| \\ &= \max_{1 \leq k \leq 2^{n-1}} \left| X_n\left(\frac{2k+1}{2^n}\right) - X_{n-1}\left(\frac{2k+1}{2^n}\right) \right|. \end{aligned}$$

Or

$$X_n\left(\frac{2k+1}{2^n}\right) = X\left(\frac{2k+1}{2^n}\right) \quad \text{et} \quad X_{n-1}\left(\frac{2k+1}{2^n}\right) = \frac{1}{2} \left(X\left(\frac{k}{2^{n-1}}\right) + X\left(\frac{k+1}{2^{n-1}}\right) \right).$$

D'où

$$\left| X_n\left(\frac{2k+1}{2^n}\right) - X_{n-1}\left(\frac{2k+1}{2^n}\right) \right| \leq \max \left(X\left(\frac{2k+1}{2^n}\right) - X\left(\frac{2k}{2^n}\right), X\left(\frac{2k+1}{2^n}\right) - X\left(\frac{2k+2}{2^n}\right) \right).$$

Il suit

$$\|X_n - X_{n-1}\|_\infty \leq \max_{0 \leq j \leq 2^n - 1} |X\left(\frac{j}{2^n}\right) - X\left(\frac{j+1}{2^n}\right)|.$$

On cherche à estimer $\mathbb{P}(\|X_n - X_{n-1}\|_\infty > \delta_n)$ pour une suite $\delta_n \rightarrow 0$ à choisir ultérieurement : Donc

$$\begin{aligned} \mathbb{P}(\|X_n - X_{n-1}\|_\infty > \delta_n) &\leq \mathbb{P}(\max_{0 \leq j \leq 2^n - 1} |X(\frac{j}{2^n}) - X(\frac{j+1}{2^n})| > \delta_n) \\ &\leq \sum_{j=0}^{2^n - 1} \mathbb{P}(|X(\frac{j}{2^n}) - X(\frac{j+1}{2^n})| > \delta_n) \\ &\leq \sum_{j=0}^{2^n - 1} \frac{1}{\delta_n^a} \mathbb{E}[|X(\frac{j}{2^n}) - X(\frac{j+1}{2^n})|^a] \end{aligned}$$

en utilisant l'inégalité de Markov. On a alors d'après (2.1) :

$$\mathbb{P}(\|X_n - X_{n-1}\|_\infty > \delta_n) \leq \sum_{j=0}^{2^n - 1} \frac{c}{\delta_n^a} (1/2^n)^{1+b} \leq \frac{c}{\delta_n^a} 2^{-nb}.$$

On choisit $\delta_n = 2^{-n\gamma}$ avec $0 < \gamma < b/a$. On a bien $\delta_n \rightarrow 0$ puis

$$\mathbb{P}(\|X_n - X_{n-1}\|_\infty > \delta_n) \leq c 2^{-n(b-\gamma a)}.$$

Cela assure que $\sum_n \mathbb{P}(\|X_n - X_{n-1}\|_\infty > \delta_n) < +\infty$. Le lemme de Borel-Cantelli s'applique donc et donne

$$\mathbb{P}(\limsup_{n \rightarrow +\infty} \{\|X_n - X_{n-1}\|_\infty > \delta_n\}) = 0,$$

c'est à dire ps, il existe n_0 tel que pour tout $n \geq n_0$,

$$\|X_n - X_{n-1}\|_\infty \leq \delta_n.$$

Comme $\sum_n \delta_n < +\infty$, on a aussi ps $\sum_n \|X_n - X_{n-1}\|_\infty < +\infty$.

Cela assure que la suite de fonctions $\{X_n\}_n$ est ps une suite de Cauchy dans l'espace (complet) des fonctions continues $(C^0(\mathbb{R}), \|\cdot\|_\infty)$. Elle converge donc uniformément vers une limite \tilde{X} continue (car dans $(C^0(\mathbb{R}), \|\cdot\|_\infty)$).

Soit $t = j/2^m$. Par définition, on a $X_n(t) = X(t)$ pour $n \geq m$. Et donc pour tous les dyadiques t , on a ps $\tilde{X}(t) = X(t)$.

Comme une réunion de négligeables est encore négligeable, on a ps, pour tout dyadique $\tilde{X}(t) = X(t)$.

Soit maintenant s non dyadique. Par densité des dyadiques, il existe une suite de dyadiques $(t_n)_n$ qui converge vers s . Comme \tilde{X} est continu, on a $\tilde{X}(t_n) \rightarrow \tilde{X}(s)$.

Mais l'hypothèse (2.1) implique la continuité en probabilité de X . En effet, par l'inégalité de Markov

$$\mathbb{P}(|X_{t_n} - X_s| \geq \varepsilon) \leq \frac{\mathbb{E}[|X_{t_n} - X_s|^a]}{\varepsilon^a} \leq \frac{c(t_n - s)^{1+b}}{\varepsilon^a}.$$

D'où $X(t_n) \xrightarrow{\mathbb{P}} X(s)$. Il y a donc aussi convergence presque sûre d'une sous suite $X(t_{n'}) \rightarrow X(s)$. Mais comme par ailleurs, on a vu $\tilde{X}(t_{n'}) \rightarrow \tilde{X}(s)$, on a ps $X(s) = \tilde{X}(s)$. \square

2.1.3 Exemples de propriétés en loi des processus.

Il existe de nombreuses classes de processus particuliers : les processus de Markov (et notamment les chaînes de Markov quand T est discret), les martingales, les processus gaussiens, les processus de Poisson, les processus stables ou encore les processus de Lévy.

Exemples de propriétés des processus. Les lois fini dimensionnelles des processus vérifient parfois des propriétés qui peuvent être utiles pour modéliser des phénomènes réels.

Définition 2.1.4 *Un processus X est dit à accroissements indépendants si pour tout $0 < t_1 < t_2 < \dots < t_n$, les v.a. $X_{t_1}, X_{t_2} - X_{t_1}, \dots, X_{t_n} - X_{t_{n-1}}$ sont indépendantes.*

Un processus est dit stationnaire si pour tout h , $(X_{t+h})_t \stackrel{\mathcal{L}}{=} (X_t)_t$ ne dépend pas de $h > 0$, c'est à dire pour tout h et tout t_1, \dots, t_n , on a $(X_{t_1+h}, \dots, X_{t_n+h}) \stackrel{\mathcal{L}}{=} (X_{t_1}, \dots, X_{t_n})$.

Un processus est dit à accroissements stationnaires si la loi des accroissements $X_{t+h} - X_t$ ne dépend pas de $h > 0$.

Exemple. $T = \mathbb{N}$. Soit $(X_i)_i$ une suite de v.a.i. On considère $S_n = \sum_{i=0}^n X_i$ le processus de sommes partielles. On parle de marche aléatoire. Alors $(S_n)_{n \in \mathbb{N}}$ est un processus à accroissements indépendants. Si en plus les v.a. X_i sont de même loi (des v.a.i.d), le processus est à accroissements indépendants et stationnaires.

2.2 Processus Gaussien

Définition 2.2.1 *Un processus est dit gaussien si toutes ses lois fini dimensionnelles $\mathcal{L}(X_{t_1}, \dots, X_{t_n})$ sont gaussiennes ($\forall n \in \mathbb{N}, \forall t_1, \dots, t_n \in T$).*

Autrement dit $X = (X_t)_t$ est gaussien si toute combinaison linéaire $a_1 X_{t_1} + \dots + a_n X_{t_n}$ suit une loi gaussienne (pour tout $n \in \mathbb{N}, t_1, \dots, t_n \in T$ et $a_1, \dots, a_n \in \mathbb{R}$).

Il est connu que la loi d'un vecteur gaussien $(X_{t_1}, \dots, X_{t_n})$ est connue (via sa fonction caractéristique) par le vecteur moyenne $(\mathbb{E}[X_{t_1}], \dots, \mathbb{E}[X_{t_n}])$ et la matrice de covariance $(\text{Cov}(X_{t_i}, X_{t_j})_{1 \leq i, j \leq n})$.

On comprend dès lors que toute la loi d'un processus gaussien est connue dès qu'on se donne la fonction moyenne $a(t) = \mathbb{E}[X_t]$ et l'opérateur de covariance $K(s, t) = \text{Cov}(X_s, X_t)$. En effet, la loi fini dimensionnelle de $(X_{t_1}, \dots, X_{t_n})$ est alors la loi normale de dimension n $\mathcal{N}(a_n, K_n)$ avec $a_n = (a(t_1), \dots, a(t_n))$ et $K_n = (K(t_i, t_j))_{1 \leq i, j \leq n}$. Les fonctions a et K définissent donc toutes les lois fini dimensionnelles de X et donc aussi sa loi.

Remarques :

Toutes les marginales d'un processus gaussien sont bien sûr gaussiennes.

Toute combinaison linéaire de marginales d'un processus gaussien est encore gaussienne.

Des bonnes conditions pour avoir une version assez régulière d'un processus gaussien sont données dans le résultat suivant dû au théorème 2.1.1.

Théorème 2.2.1 (Régularité) *Soit X un processus gaussien centré ($\mathbb{E}[X_t] = 0$), de fonction de covariance $r(s, t)$. On suppose qu'il existe $\alpha > 0$ tel que pour tout s, t :*

$$r(t, t) + r(s, s) - 2r(s, t) \leq c|t - s|^\alpha.$$

Alors il existe une version continue \tilde{X} de X . De plus, pour tout $\gamma < \alpha/2$, les trajectoires de \tilde{X} sont ps höldériennes de coefficient γ .

Démonstration : On a

$$\mathbb{E}[|X_t - X_s|^2] = \mathbb{E}[X_t^2] + \mathbb{E}[X_s^2] - 2\mathbb{E}[X_t X_s] \leq c|t - s|^\alpha.$$

On ne peut pas appliquer directement le théorème 2.1.1 car $\alpha > 1$ n'est pas garanti. On s'intéresse plutôt à $\mathbb{E}[|X_t - X_s|^{2m}]$.

Lemme 2.2.1 *Soit $X \simeq \mathcal{N}(0, 1)$. On a $\mathbb{E}[X^{2m}] = \frac{(2m)!}{2^m m!} \text{Var}(X)^m$.*

Démonstration : Des ipp. □

On déduit du lemme que

$$\mathbb{E}[|X_t - X_s|^{2m}] \leq c \frac{(2m)!}{2^m m!} |t - s|^{m\alpha}.$$

On choisit m tel que $m\alpha > 1$. D'après le théorème 2.1.1 de Kolmogorov avec

$$b = m\alpha - 1, \quad a = 2m, \quad \frac{b}{a} = \frac{m\alpha - 1}{2m} \rightarrow \frac{\alpha}{2}, \quad m \rightarrow +\infty$$

on a une version γ -höldérienne de X pour tout $\gamma < \alpha/2$. □

Soit $T = \mathbb{R}$. Dans le cas gaussien, on caractérise facilement la stationnarité d'un processus.

Proposition 2.2.1 *Un processus gaussien X est stationnaire ssi $\mathbb{E}[X_t]$ est constante et $K(s, t) = K(s - t)$ (on parle de stationnarité faible).*

Démonstration : Il est clair que ces conditions sont nécessaires, que le processus soit gaussien ou pas (exo).

Elles sont suffisantes seulement dans le cas gaussien. En effet, dans ce cas, la loi est caractérisé par $t \mapsto \mathbb{E}[X_t]$ et par $K(s, t)$. Il est facile alors de voir dans ce cas qu'une translation

dans les indices ne modifie pas la loi. \square

Dans le cas stationnaire (par le théorème de Bochner), la fonction de covariance $K(t)$ s'écrit

$$K(t) = \int_{-\infty}^{+\infty} e^{itu} \nu(du)$$

où ν est une mesure (dite mesure spectrale du processus) vérifiant $\nu(\mathbb{R}) < +\infty$. Cette mesure spectrale possède beaucoup d'informations décrivant le processus. Par exemple, on a $K(0) = \text{Var}(X_t) = \nu(\mathbb{R})$.

Exemple : cf ci-dessous le processus d'Ornstein-Uhlenbeck.

2.3 Exemples de processus gaussiens

2.3.1 Mouvement brownien

Soit $T = \mathbb{R}_+$, le mouvement brownien (standard) $(W_t)_{t \geq 0}$ est le processus gaussien défini par $\mathbb{E}[W_t] = 0$ et $K(s, t) = \min(s, t)$. On l'appelle aussi processus de Wiener.

Propriétés immédiates

- 1) $W_0 = 0$ car la loi de W_0 est $\mathcal{N}(0, 0) = \delta_0$, la loi dégénérée en 0.
- 2) $(W_t)_t$ est un processus à accroissements indépendants. En effet soit $0 \leq t_1 < t_2 < t_3 < t_4$, on a

$$\begin{aligned} \text{Cov}(W_{t_2} - W_{t_1}, W_{t_4} - W_{t_3}) &= \mathbb{E}[(W_{t_2} - W_{t_1})(W_{t_4} - W_{t_3})] \\ &= \mathbb{E}[W_{t_2}W_{t_4}] - \mathbb{E}[W_{t_2}W_{t_3}] - \mathbb{E}[W_{t_1}W_{t_4}] + \mathbb{E}[W_{t_1}W_{t_3}] \\ &= t_2 - t_2 - t_1 + t_1 = 0. \end{aligned}$$

Les variables $W_{t_2} - W_{t_1}$ et $W_{t_4} - W_{t_3}$ sont donc non corrélées. Comme elles sont gaussiennes, elles sont indépendantes. On justifie de même l'indépendance de n accroissements.

- 3) $W_t \simeq \mathcal{N}(0, t)$ car $\mathbb{E}[W_t] = 0$ et $\text{Var}(W_t) = K(t, t) = t$.
- 4) Si $s \leq t$, on a $W_t - W_s \simeq W_{t-s}$. En effet

$$\begin{aligned} \text{Var}(W_t - W_s) &= \text{Cov}(W_t - W_s, W_t - W_s) \\ &= \text{Cov}(W_t, W_t) - 2\text{Cov}(W_t, W_s) + \text{Cov}(W_s, W_s) \\ &= t - 2s + s = t - s. \end{aligned}$$

Donc $W_t - W_s \simeq \mathcal{N}(0, t - s) \simeq W_{t-s}$.

- 5) Autosimilarité : $\tilde{W}_t = \frac{1}{\sqrt{c}} W_{ct}$ est encore un mouvement brownien (standard).

- 6) Comportement analogue en 0 et en $+\infty$: $\tilde{W}_t = tW_{1/t}$ est encore un mouvement brownien standard.

Le processus W_t a donc le même comportement en 0 et en $+\infty$

7) Localisation : pour tout t_0 , $\tilde{W}_t = W_{t+t_0} - W_{t_0}$ est encore un mouvement brownien standard.

Le processus W_t a donc le même comportement en 0 et en t_0 .

8) $(W_t)_t$ a des trajectoires ps holdériennes d'ordre $\gamma < 1/2$ mais non ps dérivables.

En effet, $\mathbb{E}[|W_t - W_s|^2] = |t - s|$, donc la continuité höldérienne suit du théorème 2.2.1. On admet (pour l'instant ?) la non-dérivabilité des trajectoires.

[Graphe typique des trajectoires.]

2.3.2 Pont Brownien

Soit $T = [0, 1]$, le pont brownien $(W_t^\circ)_{t \in [0,1]}$ est le processus gaussien centré défini par la fonction de covariance $K(s, t) = \min(s, t) - st$.

[Graphe typique des trajectoires.]

Proposition 2.3.1 *On peut définir directement un pont brownien W° à partir d'un mouvement brownien W par*

$$W_t^\circ = W_t - tW_1.$$

Démonstration : En effet, d'abord $(W_t - tW_1)_{t \in [0,1]}$ est gaussien, centré puis pour $s, t \in [0, 1]$, on a

$$\begin{aligned} & \text{Cov}(W_t - tW_1, W_s - tW_1) \\ &= \text{Cov}(W_t, W_s) - t \text{Cov}(W_1, W_s) - t \text{Cov}(W_s, W_1) + ts \text{Cov}(W_1, W_1) \\ &= \min(t, s) - ts - st + ts \\ &= \min(t, s) - ts \end{aligned}$$

Comme le processus $(W_t - tW_1)_{t \in [0,1]}$ est gaussien, centré avec la bonne covariance, il s'agit d'un pont brownien. \square

Réciproquement, on peut construire le mouvement brownien W sur $T = [0, 1]$ à partir du pont brownien W° et d'une loi normale $X \simeq \mathcal{N}(0, 1)$ indépendante de W° par

$$W_t = W_t^\circ + tX.$$

Exercice. Vérifier par un calcul de covariance qu'on définit ainsi un mouvement Brownien.

Propriétés.

1. $\tilde{W}_t^\circ = W_{1-t}^\circ$ est encore un pont brownien. Le pont brownien est donc symétrique en 0 et en 1 par retournement du temps.
2. $(W_t^\circ)_t$ a des trajectoires ps holdériennes d'ordre $\gamma < 1/2$ mais non ps dérivables. L'argument est le même que pour le mouvement brownien avec $\mathbb{E}[(W_t^\circ)^2] = t - t^2$.
3. W° est un MB W conditionné à valoir 0 à la date $t = 1$.

2.3.3 Processus d'Ornstein-Uhlenbeck

Soit $T = \mathbb{R}$, le processus d'O.U. est le processus gaussien centré défini par

$$U_t = e^{-t/2}W(e^t)$$

où W est un MB. On montre facilement que $U_t \simeq \mathcal{N}(0, 1)$ car $\text{Var}(U_t) = 1$, ce processus est donc stationnaire.

Sa fonction de covariance est donnée par

$$K(s, t) = e^{-|t-s|/2}$$

Elle ne dépend que de la différence $(t - s)$, il s'agit bien d'un processus stationnaire de fonction de covariance plus simplement donnée par $K(t) = e^{-|t|/2}$ (exo). Elle est donnée sous forme intégrale avec la mesure spectrale $\nu(du) = \frac{1}{\pi} \frac{du}{1+u^2}$.

2.3.4 Brownien géométrique

Ce n'est pas un processus gaussien mais l'exponentiel d'un processus gaussien. Il s'agit de

$$S_t = x \exp(\mu t + \sigma W_t - \sigma^2 t/2).$$

Un tel processus modélise le cours d'un actif S_t soumis à un taux d'intérêt μ et à une volatilité σ et qui vaut x au temps 0.

On le trouve en supposant comme Samuelson que les rendements entre deux périodes sont mesurés par les logarithmes des cours S_t .

On suppose de plus que les rendements entre 0 et t suivent un mouvement brownien de tendance (drift) $\mu - \sigma^2/2$ et de coefficient de diffusion (volatilité) σ . Cela se traduit par les propriétés suivantes sur les prix $(S_t)_t$:

- $S_0 = x$.
- Les rendements $\log S_t - \log S_s$ suivent une loi gaussienne de moyenne $(\mu - \sigma^2/2)(t - s)$ et de variance $\sigma^2(t - s)$.
- Pour tout $t_0 = 0 \leq t_1 \leq \dots \leq t_n$, les accroissements relatifs $S_{t_{i+1}}/S_{t_i}$, $0 \leq i \leq n - 1$, sont indépendants.

On en déduit qu'il existe un MB $(W_t)_{t \geq 0}$ tel que en t , $S_t = x \exp(\mu t + \sigma W_t - \sigma^2 t/2)$.

2.3.5 Bruit blanc gaussien

Soit (\mathcal{A}, μ) un espace mesuré et $\mathcal{U} = \{A \in \mathcal{A}, \mu(A) < +\infty\}$.

Le bruit blanc est un processus gaussien $(X_A)_{A \in \mathcal{A}}$ indexé par l'ensemble des mesurables \mathcal{A} défini par $\mathbb{E}[X_A] = 0$ et $\text{Cov}(X_A, X_B) = \mu(A \cap B)$.

Il faut appréhender le bruit blanc comme une mesure aléatoire $A \mapsto X_A(\omega)$. Elle est aléatoire car X_A dépend de ω . On connaît quand même la loi de $X(A) \simeq \mathcal{N}(0, \mu(A))$.

Attention cependant, un bruit blanc n'est pas une vraie mesure car $A \mapsto X_A$ n'est pas σ -additif.

2.3.6 Mouvement brownien fractionnaire

Soit $T = \mathbb{R}_+$. Le mouvement brownien fractionnaire (mBf) $(B^H(t))_{t \geq 0}$ est le processus gaussien centré défini par la fonction de covariance

$$K(s, t) = \frac{1}{2}(|s|^{2H} + |t|^{2H} - |s - t|^{2H}).$$

Propriétés immédiates.

1. Pour $H = 1/2$, le mBf devient le MB standard.
2. On a $\mathbb{E}[|B^H(t) - B^H(s)|^2] = |t - s|^{2H}$.
3. Autosimilarité : $B^H(t) = c^{-H/2}B^H(ct)$ est encore un mBf d'indice H .
4. Les accroissements du mBf ne sont indépendants que lorsque $H = 1/2$ (c'est à dire dans le cas du MB).

Chapitre 3

Mouvement brownien

Dans cette section, on développe la présentation du mouvement brownien (MB).

3.1 Historique

Le MB est associé à l'analyse de mouvements qui évoluent au cours du temps de manière si désordonnée qu'il semble difficile de prévoir leur évolution, même dans un intervalle de temps très court. Il joue un rôle central dans la théorie des processus stochastiques, notamment parce que dans de nombreux problèmes théoriques ou appliqués, le mouvement brownien ou les diffusions (qui lui sont associées) fournissent des modèles simples sur lesquels de nombreux calculs peuvent être faits.

Un botaniste anglais, Robert Brown, décrit en 1827 le mouvement de fines particules organiques en suspension dans un gaz ou un fluide. Au 19^{ème} siècle, après lui, plusieurs physiciens reconnaissent que ce mouvement est très irrégulier et ne semble pas admettre de tangente. On ne pourrait donc pas parler de sa vitesse, ni a fortiori lui appliquer les lois de la mécanique !

En 1900, Louis Bachelier introduit le mouvement Brownien pour modéliser la dynamique des prix des actions à la bourse, mais sa démarche est ensuite oubliée jusque vers les années 1960.

En 1905, Albert Einstein construit un modèle probabiliste pour décrire le mouvement d'une particule qui diffuse : il trouve que la loi de la position à l'instant t de la particule, sachant que l'état initial est x admet une densité qui vérifie l'équation de la chaleur et de ce fait est gaussienne. Sa théorie sera rapidement confortée par des mesures expérimentales de constantes de diffusions satisfaisantes.

La même année qu'Einstein, le physicien polonais Smoluchowski décrit le mouvement brownien comme limite de promenades aléatoires.

En 1923, l'américain Norbert Wiener construit rigoureusement la « fonction aléatoire » du mouvement brownien. Il établit en particulier que les trajectoires sont continues.

Le français Paul Lévy découvre ensuite avec d'autres mathématiciens de nombreuses propriétés du mouvement brownien et introduit une première forme des équations diffé-

rentielles stochastiques, dont l'étude sera ensuite systématisée par le japonais K. Itô. Ses travaux sont rassemblés dans un traité paru en 1948 et devenu très célèbre

3.2 Définition

Le caractère très erratique des trajectoires qui caractérise le mouvement brownien est en général associé à l'observation que le phénomène, bien que très désordonné, présente une certaine homogénéité dans le temps, au sens où la date d'origine des observations n'a pas d'importance. Ces propriétés sont reprises dans la définition qui suit.

Définition 3.2.1 *Un mouvement brownien (standard) réel est un processus gaussien centré $(W_t)_{t \in \mathbb{R}^+}$ à trajectoires continues de fonction de covariance*

$$K(s, t) = \min(s, t).$$

Remarque 3.2.1 *On a déjà vu les propriétés immédiates suivantes*

- $W_0 = 0$.
- $W_t \simeq \mathcal{N}(0, t)$.
- Tout accroissement $W_t - W_s$ (où $0 \leq s < t$) suit une loi gaussienne centrée $\mathcal{N}(0, t-s)$.
- Pour tout $0 \leq t_1 < \dots < t_n$, les accroissements $(W_{t_{i-1}} - W_{t_i}, 0 \leq i \leq n-1)$ sont indépendants.

[Graphe des trajectoires typiques du MB]

La probabilité que W_t appartienne à un petit intervalle $[x, x + dx]$ est donc donnée par la densité gaussienne

$$\mathbb{P}(W_t \in [x, x + dx]) = \frac{1}{\sqrt{2\pi t}} \exp(-x^2/2t) dx.$$

En particulier, la v.a. W_t qui est une v.a. gaussienne de variance t est comprise entre les nombres $f_1(t) = 2\sqrt{t}$ et $f_2(t) = -2\sqrt{t}$ avec une probabilité 95%.

On peut montrer que cette propriété est vraie pour toute la trajectoire brownienne qui est donc comprise entre les deux courbes de f_1 et de f_2 avec une proba comparable (c'est vrai globalement et pas seulement « t par t »).

En général, les phénomènes observés ne sont pas aussi normalisés. On appelle encore mouvement brownien issu de x et de dérive (ou drift) b et de coefficient de diffusion σ , le processus $X_t = x + \sigma W_t + \mu t$.

Proposition 3.2.1 *Le mouvement brownien X est encore un processus à accroissements indépendants stationnaires et gaussiens. Il est non centré et tel que $X_0 = x$. De plus $X_t \simeq \mathcal{N}(x + \mu t, \sigma^2 t)$.*

Sauf mention contraire, quand on parlera du mouvement brownien, il s'agira du mouvement brownien standard W .

3.3 Propriétés du MB

3.3.1 Propriétés en loi

Systématiquement, pour vérifier qu'on a un MB, il s'agit de vérifier qu'on a un processus gaussien, centré, à trajectoires continues et avec la bonne fonction de covariance. En général, il est facile (et omis) de constater que le processus est gaussien, centré et à trajectoires continues.

Symétrie. Si W est un MB alors $-W$ en est un aussi.

Propriété d'échelle (autosimilarité) Si W est un MB alors pour tout $c > 0$, $W_t^c = \frac{1}{\sqrt{c}}W_{ct}$ est encore un mouvement brownien (standard).

En effet dans cas,

$$\text{Cov}(W_t^c, W_s^c) = \frac{1}{\sqrt{c}^2} \text{Cov}(W_{ct}, W_{cs}) = \frac{1}{c} \min(ct, cs) = \min(t, s).$$

Cette propriété est particulièrement importante, elle montre que ε fois W_t se comporte comme un mouvement brownien lu en $\varepsilon^2 t$.

Comportement analogue en 0 et en $+\infty$. $\tilde{W}_t = tW_{1/t}$ est encore un mouvement brownien standard.

En effet dans cas,

$$\text{Cov}(\tilde{W}_t, \tilde{W}_s) = ts \text{Cov}(W_{1/t}, W_{1/s}) = ts \min(1/t, 1/s) = ts / \max(t, s) = \min(t, s).$$

Le processus W_t a donc le même comportement en 0 et en $+\infty$

Invariance par translation. Le MB translaté de $h > 0$ $\bar{W}_t^h = W_{t+h} - W_h$ est encore un mouvement brownien standard, indépendant du MB arrêté en h $(W_t)_{0 \leq t \leq h}$.

En effet pour $s \leq t$:

$$\begin{aligned} \text{Cov}(\bar{W}_t^h, \bar{W}_s^h) &= \text{Cov}(W_{t+h} - W_h, W_{s+h} - W_h) \\ &= K(t+h, s+h) - K(t+h, h) - K(h, s+h) + K(h) \\ &= s+h-h-h+h = s = \min(s, t). \end{aligned}$$

Le processus W_t a donc le même comportement en 0 et en h .

Retournement du temps. Le processus retourné à l'instant T , $\hat{W}_t^T = W_T - W_{T-t}$ est encore un MB sur $[0, T]$.

3.3.2 Propriétés des trajectoires du MB

Il est naturel d'étudier les propriétés des trajectoires browniennes en tant que fonction du temps. Le premier point est évidemment le comportement à l'infini.

Proposition 3.3.1 *Le MB oscille entre $+\infty$ et $-\infty$ lorsque $t \rightarrow +\infty$, c'est à dire ps*

$$\limsup_{t \rightarrow +\infty} W_t = +\infty, \quad \liminf_{t \rightarrow +\infty} W_t = -\infty.$$

De plus, la vitesse de convergence est moins grande que celle de $t : W_t/t \rightarrow 0$, quand $t \rightarrow +\infty$.

Démonstration : On a vu que $W'(t) = tW(1/t)$ est encore un MB. On se ramène donc à étudier en 0, $W_{1/t}$. Comme W_t prend des valeurs des deux signes au voisinage de 0, tout suit. La vitesse est admise. \square

Proposition 3.3.2 *Presque sûrement, les propriétés suivantes sont vérifiées :*

- $\sup_{t \in [0,1]} W_t > 0$.
- $\inf_{t \in [0,1]} W_t < 0$.
- *il existe $t \in]0, 1[$ tel que $W_t = 0$.*

Démonstration : Les deux premiers points sont équivalents par symétrie $\mathcal{L}(W) = \mathcal{L}(-W)$. S'ils sont faux, ils le sont donc simultanément. Mais alors on aurait $W_t \equiv 0$, ce qui est absurde. Ils sont donc vrais. Par continuité, le troisième est une conséquence du théorème des valeurs intermédiaires et des deux premiers points. \square

Proposition 3.3.3 *Pour tout $\varepsilon > 0$, W_t a un zéro ps sur $]0, \varepsilon[$.*

Démonstration : Le processus $\tilde{W}_t = \sqrt{\varepsilon}^{-1}W_{\varepsilon t}$, $t \in [0, 1]$ est un MB pour lequel il existe $t_0 \in]0, 1[$ tel que $\tilde{W}_{t_1} = 0$, c'est à dire $W(\varepsilon t_0) = 0$. Le point $t_1 = \varepsilon t_0 \in]0, \varepsilon[$ répond à la question. \square

On a une conséquence immédiate de ce résultat :

Proposition 3.3.4 *Presque sûrement, il existe $t_n \rightarrow 0$ tels que $W_{t_n} = 0$.*

Par translation, en tout point point, il existe une suite de zéro du MB qui converge vers ce point.

Le processus W est par définition continue. En fait grâce au théorème de régularité de Kolmogorov pour les gaussiens, on a même plus

Proposition 3.3.5 *Le MB $(W_t)_t$ a des trajectoires ps holdériennes d'ordre $\gamma < 1/2$.*

Démonstration : En effet, on a

$$r(t, t) + r(s, s) - 2r(s, t) = t + s - 2\min(t, s) = |t - s|.$$

donc le théorème 2.2.1 s'applique avec $\alpha = c = 1$. Il donne la continuité höldérienne pour tout $\gamma < \alpha/2 = 1/2$. \square

Proposition 3.3.6 *Les trajectoires du MB ne sont pas dérivables.*

Démonstration : Par la propriété de translation, il suffit de montrer la non dérivabilité en 0, c'est à dire montrer que

$$\lim_{t \rightarrow 0} \frac{W_t - W_0}{t - 0} = \lim_{t \rightarrow 0} \frac{W_t}{t}$$

n'existe pas. Or par retournement du temps $\frac{W_t}{t} = \tilde{W}_{1/t}$ où \tilde{W} est encore un MB. Mais d'après la proposition précédente,

$$\limsup_{s \rightarrow +\infty} \tilde{W}_s = +\infty, \quad \liminf_{s \rightarrow +\infty} \tilde{W}_s = -\infty$$

avec $s = 1/t$, ce qui montre que la limite cherchée n'existe pas. \square

Il est possible d'avoir une description un peu plus précise du comportement du mouvement brownien au voisinage de l'infini (ou par retournement au voisinage de l'origine) grâce à la loi du log itéré.

Théorème 3.3.1 (Loi du log itéré) *Soit W un MB standard. Alors on a*

$$\limsup_{t \rightarrow +\infty} \frac{W_t}{h(t)} = 1 \text{ ps}, \quad \liminf_{t \rightarrow +\infty} \frac{W_t}{h(t)} = 1 \text{ ps}$$

où $h(t) = \sqrt{2t \log \log(1/t)}$.

Démonstration : Amis. Faire un dessin.

Variation quadratique

Définition 3.3.1 *Soit $f : [0, 1] \rightarrow \mathbb{R}$. On définit la α -variation de f par*

$$Var(f, \alpha) = \sup_{t_k} \sum_k |f(t_{k+1}) - f(t_k)|^\alpha$$

où $\{t_k\}$ est une partition de $[0, 1]$ et le sup est pris sur l'ensemble de ces partitions.

Pour $\alpha = 1$, on parle de la variation. On dit que f est à variations bornées si $Var(f, 1) < +\infty$.

Pour $\alpha = 2$, on parle de la variation quadratique.

Exo. Montrer que la variation quadratique d'une fonction dérivable tend vers 0 (appliquer le théorème des accroissements finis).

Proposition 3.3.7 *Soit $v_n = \sum_{k=0}^{n-1} |W(\frac{k+1}{n}) - W(\frac{k}{n})|^2$. Alors $v_n \rightarrow 1$ ps.*

Heuristiquement : la variation quadratique du MB tend vers 1.

Démonstration : On a $v_n = \sum_{k=0}^{n-1} \Delta_k(W)^2$ avec $\Delta_k(W) = W(\frac{k+1}{n}) - W(\frac{k}{n})$.

On a $v_n - 1 = \sum_{k=0}^{n-1} (\Delta_k(W)^2 - \frac{1}{n}) = \sum_{k=0}^{n-1} \alpha_{n,k}$ avec $\alpha_{n,k} = \Delta_k(W)^2 - \frac{1}{n}$.

Notons $\zeta = W_1^2 - 1$ et remarquons

$$\Delta_0(W)^2 - \frac{1}{n} = W^2(\frac{1}{n}) - \frac{1}{n} \stackrel{\mathcal{L}}{=} (\frac{1}{\sqrt{n}}W_1)^2 - \frac{1}{n} = \frac{1}{n}\zeta.$$

On a

- Pour $k = 0, \dots, n$, les $\alpha_{n,k}$ sont iid (indépendance et stationarité des accroissements de W)
- $\mathbb{E}[\alpha_{n,k}] = \mathbb{E}[\Delta_0(W)^2] - \frac{1}{n} = \mathbb{E}[W(\frac{1}{n})^2] - \frac{1}{n} = 0$.
- $\mathbb{E}[\alpha_{n,k}^2] = \mathbb{E}[\alpha_{n,0}^2] = \mathbb{E}[\zeta^2/n^2] = \frac{1}{n^2}\mathbb{E}[\zeta^2]$,
- $\mathbb{E}[\alpha_{n,k}^4] = \mathbb{E}[\alpha_{n,0}^4] = \mathbb{E}[\zeta^4/n^4] = \frac{1}{n^4}\mathbb{E}[\zeta^4]$

On utilise maintenant le résultat suivant :

Lemme 3.3.1

$$\mathbb{E}[(\sum_{k=0}^{n-1} \alpha_{n,k})^4] \leq Cn^2\mathbb{E}[\alpha_{n,0}^4].$$

On a alors par l'inégalité de Markov

$$\mathbb{P}(|v_n - 1| \geq \delta_n) \leq \frac{\mathbb{E}[(\sum_{k=0}^{n-1} \alpha_{n,k})^4]}{\delta_n^4} \leq \frac{Cn^2\mathbb{E}[\zeta^4]}{n^4\delta_n^4} = \frac{c'}{n^2\delta_n^4}.$$

Avec le choix $\delta = 1/\ln n$, on a $\sum_n \mathbb{P}(|v_n - 1| \geq \delta_n) < +\infty$.

Le lemme de Borel-Cantelli s'applique et donne : ps, pour n assez grand on a $|v_n - 1| \leq \delta_n \rightarrow 0$. D'où ps $v_n \rightarrow 1$. □

Lemme 3.3.2 (Borel-Cantelli) Si $(A_n)_n$ est une suite d'évènements telle que $\sum_n \mathbb{P}(A_n) < +\infty$ alors

$$\mathbb{P}(\limsup_n A_n) = 0$$

c'est à dire presque sûrement, seul un nombre fini d'évènements A_n est réalisé.

Preuve du lemme. Soit $(\alpha_i)_i$ iid tels que $\mathbb{E}[\alpha_i] = 0$ et $\mathbb{E}[\alpha_i^4] < +\infty$. On montre que

$$\mathbb{E}[(\sum_{i=1}^n \alpha_i)^4] \leq Cn^2\mathbb{E}[\alpha_1^4].$$

Pour cela, on a

$$(\sum_{i=1}^n \alpha_i)^4 = \sum_i \alpha_i^4 + C_4^2 \sum_{i \neq j} \alpha_i^2 \alpha_j^2 + C_4^1 \sum_{i \neq j} \alpha_i^3 \alpha_j + C' \sum_{i \neq j \neq k} \alpha_i^2 \alpha_j \alpha_k + C''' \sum_{i \neq j \neq k \neq l} \alpha_i \alpha_j \alpha_k \alpha_l.$$

Par indépendance et centrage des v.a., il vient

$$\begin{aligned}\mathbb{E}\left[\left(\sum_{i=1}^n \alpha_i\right)^4\right] &= \sum_i \mathbb{E}[\alpha_i^4] + C_4^2 \sum_{i \neq j} \mathbb{E}[\alpha_i^2 \alpha_j^2] \\ &= n\mathbb{E}[\alpha_1^4] + 6n(n-1)\mathbb{E}[\alpha_1^2]^2.\end{aligned}$$

Or d'après l'inégalité de Cauchy Schwarz, on a $(\mathbb{E}[\alpha_i^2])^2 \leq \mathbb{E}[\alpha_i^4]$, ce qui conclut le lemme. \square

Proposition 3.3.8 *Presque sûrement, les trajectoires du MB sont à trajectoires à variations non bornées.*

Ce résultat justifie, que si les trajectoires browniennes sont continues, elles oscillent quand même beaucoup... tellement que les trajectoires ne sont pas dérivables.

Démonstration : En effet, il suffit de justifier la remarque suivante : si f est continue et

$$\sum_{k=0}^n \left|f\left(\frac{k+1}{n}\right) - f\left(\frac{k}{n}\right)\right|^2 \rightarrow a \in]0, +\infty[$$

alors $Var(f, 1) = +\infty$.

Pour cela, supposons que $Var(f, 1) < +\infty$ et notons

$$w_f(u) = \sup_{|x-y|<u} |f(x) - f(y)|$$

le module de continuité de f . Alors

$$\sum_{k=0}^n \left|f\left(\frac{k+1}{n}\right) - f\left(\frac{k}{n}\right)\right|^2 \leq w_f\left(\frac{1}{n}\right) \sum_{k=0}^n \left|f\left(\frac{k+1}{n}\right) - f\left(\frac{k}{n}\right)\right| \leq w_f\left(\frac{1}{n}\right) Var(f, 1) \rightarrow 0$$

quand $n \rightarrow +\infty$ car par continuité de f : $w_f(1/n) \rightarrow 0$.

Il y a donc une absurdité à supposer que $Var(f, 1) < +\infty$. \square

3.4 Équation de la chaleur

3.4.1 Origine physique

L'équation de la chaleur est l'EDP qui décrit la propagation de la chaleur en donnant la température $T(t, x)$ dans un milieu en fonction du temps t et du lieu x .

Le flux d'énergie thermique qui traverse une surface unitaire par unité de temps est donné par la loi de Fourier :

$$J = -k \frac{\partial T}{\partial x}$$

où k est le coefficient de conductivité thermique du milieu (en $mkgs^{-3}K^{-1}$). Cette loi stipule qu'une différence de température engendre un flux d'énergie dans la direction des températures décroissantes.

Si on calcule le gain d'énergie par unité de temps d'un volume d'épaisseur dx et de section S , on remarque que cette puissance est égale à la différence entre le flux entrant et le flux sortant :

$$P = J(x)S - J(x + dx)S \quad \text{c'est à dire} \quad P = -\frac{\partial J}{\partial x}Sdx.$$

Cette puissance assimilée à cet élément de volume Sdx est supposée élever sa température T . On pose une relation linéaire simple

$$P = cm\frac{\partial T}{\partial t}$$

où c est la chaleur spécifique de la matière considérée. Comme $m = \rho Sdx$ (où ρ est la masse volumique du milieu), on a alors

$$\rho c\frac{\partial T}{\partial t} = -\frac{\partial J}{\partial x}$$

qui est l'équation de continuité d'un flux thermique d'énergie. En utilisant la loi de Fourier, on obtient l'équation de la chaleur

$$\frac{\partial T}{\partial t} = \frac{k}{\rho c} \frac{\partial^2 T}{\partial x^2}.$$

Cette EDP se généralise facilement en dimension supérieure.

3.4.2 Origine mathématique

Les premières propriétés du mouvement brownien mises en évidence par Bachelier et Einstein concernent le lien entre la distribution du mouvement brownien issu de x et l'équation de la chaleur. On note $p_t(x, \cdot)$ la densité de la v.a. qui modélise le phénomène $x + W_t$ à la date t qui part de x à la date 0. Par stationarité du phénomène, $p_t(x, y)$ ne dépend de x, y que par $y - x$. Puis, ces auteurs déduisent de la propriété d'accroissements indépendants que la densité $p_t(x, \cdot)$ de la loi de $x + W_t$ (qui n'est pas supposée gaussienne a priori) vérifie l'équation de convolution

$$\int_{\mathbb{R}} p_t(x, y)p_h(y, z)dy = p_{t+h}(x, z) \quad (3.1)$$

En effet, on écrit $x + W_{t+h} = x + W_t + W_{t+h} - W_t$ et on note que

- $W_t = W_t - W_0$ est indépendant de $W_{t+h} - W_t$,
- si $x + W_t = y$, la loi de $x + W_t + W_{t+h} - W_t$ est la même que celle de $y + \tilde{W}_h$ de densité $p_h(y, \cdot)$,

- on récupère la densité du tout, en intégrant par rapport à la loi de $x + W_t$ de densité $p_t(s, \cdot)$.

Bachelier conclut en montrant que l'équation (3.1) est vérifiée par les fonctions de la forme $A_t e^{-\pi A_t^2 x^2}$ pour lesquelles A_t^2 est proportionnelle au temps. Il n'envisage pas a priori d'autre type de fonctions. Ce résultat montre la grande généralité des situations qui peuvent être modélisées par un MB.

Revenons à la densité gaussienne

$$g(t, x, y) := g(t, y - x) = \frac{1}{\sqrt{2\pi t}} \exp(-(y - x)^2/(2t))$$

qui traduit que $x + W_{t+h}$ est la somme des v.a. gaussiennes indépendantes $x + W_t$ et $W_{t+h} - W_h$.

Un calcul direct montre que le noyau gaussien est solution de l'équation de la chaleur, c'est à dire de l'EDP

$$\begin{cases} g'_t(t, x, y) &= \frac{1}{2} g''_{yy}(t, x, y) \\ g'_t(t, x, y) &= \frac{1}{2} g''_{xx}(t, x, y). \end{cases}$$

La densité gaussienne standard satisfait donc l'équation de la chaleur par rapport aux variables x et y . Cette propriété est étendue à une vaste classe de fonctions construites à partir du mouvement brownien.

Théorème 3.4.1 a) *Considérons la fonction*

$$u(t, x, f) = \mathbb{E}[f(x + W_t)] = \int_{\mathbb{R}} g(t, x, y) f(y) dy$$

où f est une fonction borélienne bornée. La fonction u est C^∞ en espace et en temps pour $t > 0$ et vérifie l'équation de la chaleur

$$u'_t(t, x, f) = \frac{1}{2} u''_{xx}(t, x, f), \quad u(0, x) = f(x). \quad (3.2)$$

b) *Lorsque le point de départ du mouvement X_0 est tiré aléatoirement suivant une loi de densité $\pi(x)$, indépendante du MB, la densité de la loi de $X_0 + W_t$ est égale à $q(t, y) = \int_{\mathbb{R}} g(t, y - x) \pi(x) dx$ et vérifie l'équation de la chaleur*

$$q'_t(t, y) = \frac{1}{2} q''_{yy}(t, y), \quad q(0, y) = \pi(y).$$

Démonstration : a) La fonction $u(t, x, f) = \mathbb{E}[f(x + W_t)] = \int_{\mathbb{R}} g(t, x, y) f(y) dy$ est très régulière pour $t > 0$, car la densité gaussienne (le noyau de la chaleur) est C^∞ , à dérivées bornées pour $t > a$. Il vient par dérivation sous le signe intégral que

$$u'_t(t, x, f) = \int_{\mathbb{R}} g'_t(t, x, y) f(y) dy, \quad u''_{xx}(t, x) = \int_{\mathbb{R}} g''_{xx}(t, x, y) f(y) dy.$$

On utilise maintenant que $g(t, x, y)$ satisfait l'équation de la chaleur en x et on en déduit (3.2)

b) Supposons que la condition initiale soit tirée au hasard de manière indépendante du MB et donc de W_t . La loi de $X_0 + W_t$ admet une densité qui est la convolée de $\pi(x)$ et de $g(t, x)$. C'est le résultat du théorème. \square

La formule précédente peut être étendue sous certaines conditions à d'autres fonctions que les fonctions bornées, par exemple pour les fonctions $f(x) = e^{\lambda x}$. La fonction $u(t, x, e^{\lambda \cdot})$ est la transformée de Laplace de $x + W_t$. Des calculs gaussiens (classiques) montrent que

$$u(t, x, e^{\lambda \cdot}) = \mathbb{E}[e^{\lambda(x+W_t)}] = e^{\lambda x + \frac{1}{2}\lambda^2 t}.$$

Lorsque la fonction considérée est régulière, une autre formulation peut être donnée à cette relation qui jouera un rôle important dans la suite :

Proposition 3.4.1 *Si f est une fonction C_b^1 en temps et C_b^2 en espace, c'est à dire à dérivées bornées en temps et en espace, on a*

$$u'_t(t, x, f) = u(t, x, f'_t + \frac{1}{2}f''_{xx})$$

soit sous une forme plus probabiliste

$$\mathbb{E}[f(t, x + W_t)] = f(0, x) + \int_0^t \mathbb{E}[\frac{1}{2}f''_{xx}(s, x + W_s) + f'_t(s, x + W_t)]ds. \quad (3.3)$$

Démonstration : On représente la fonction $u(t, x, f)$ de la façon suivante

$$u(t, x, f) = \mathbb{E}[f(t, x + W_t)] = \int_{\mathbb{R}} g(t, y)f(t, x + y)dy = \int_{\mathbb{R}} g(t, x, z)f(t, z)dz$$

où $g(t, y)$ est la densité de $W_t \simeq \mathcal{N}(0, t)$ et $g(t, x, z) = g(t, z - x)$ est la densité de $x + W_t \simeq \mathcal{N}(x, t)$.

Par convergence dominée, on dérive sous le signe intégral, pour une fonction f deux fois dérivable, à dérivées bornées.

$$\begin{aligned} u''_{xx}(t, x, f) &= \int_{\mathbb{R}} g(t, y)f''_{xx}(t, x + y)dy = u(t, x, f''_{xx}) \\ &= \int_{\mathbb{R}} g''_{xx}(t, x, z)f(t, z)dz \quad \text{avec deux ipp.} \\ u'_t(t, x, f) &= \int_{\mathbb{R}} g(t, x, z)f'_t(t, z)dz + \int_{\mathbb{R}} g'_t(t, x, z)f(t, z)dz \\ &= u(t, x, f'_t) + \frac{1}{2} \int_{\mathbb{R}} g''_{xx}(t, x, z)f(t, z)dz = u(t, x, f'_t) + \frac{1}{2}u(t, x, f''_{xx}). \end{aligned}$$

Dans ce calcul on a utilisé que g satisfait l'équation de la chaleur. Pour avoir l'équation intégrale (3.3), il suffit d'intégrer par rapport à t et d'expliciter les fonctions $u(t, x, f'_t)$ et

$u(t, x, f''_{xx})$ comme des espérances. \square

Le MB décentré $X_t^x = x + bt + \sigma W_t$ joue un rôle important dans les applications. Les équations aux dérivées partielles (EDP) précédentes s'étendent sans difficulté à partir de l'EDP satisfaite par la densité de X_t^x .

$$g_{b,\sigma^2}(t, x, y) = \frac{1}{\sqrt{2\pi\sigma^2 t}} \exp\left(-\frac{(y-x-bt)^2}{2\sigma^2 t}\right) = g(\sigma^2 t, x+bt, y) = g(\sigma^2 t, x, y-bt).$$

Nous introduisons le générateur associé à ce processus, c'est à dire l'opérateur du 2d ordre défini par

$$L_{b,\sigma^2}\phi(x) = \frac{1}{2}\sigma^2\phi''_{xx}(x) + b\phi'_x(x).$$

Puisque $g(t, x, y)$ satisfait l'équation de la chaleur, la fonction $x \mapsto g_{b,\sigma^2}(t, x, y)$ vérifie

$$\partial_t g_{b,\sigma^2}(t, x, y) = \frac{1}{2}\sigma^2 g''_{xx}(\sigma^2 t, x+bt, y) + b g'_x(\sigma^2 t, x+bt, y) = L_{b,\sigma^2} g_{b,\sigma^2}(t, x, y).$$

Proposition 3.4.2 *Les fonctions $u(t, x, f) = \mathbb{E}[f(x+bt+\sigma W_t)] = \int_{\mathbb{R}} g_{b,\sigma^2}(t, x, y) f(y) dy$ satisfont l'EDP*

$$\begin{cases} u'_t(t, x, f) &= L_{b,\sigma^2} u(t, x, f) = \frac{1}{2}\sigma^2 u''_{xx}(t, x, f) + b u'_x(t, x, f) \\ u(0, x, f) &= f(x). \end{cases} \quad (3.4)$$

De plus, si f est une fonction de classe C_b^1 en temps sur $\mathbb{R}^+ \setminus \{0\}$ et C_b^2 en espace, alors

$$\mathbb{E}[f(t, X_t^x)] = f(0, x) + \int_0^t \mathbb{E}[L_{b,\sigma^2} f(s, X_s^x) + f'_t(s, X_s^x)] ds. \quad (3.5)$$

Démonstration : Pour (3.4), il suffit d'intégrer par rapport à $f(y)dy$ l'EDP satisfaite par la densité $g_{b,\sigma^2}(t, x, y)$ considérée comme fonction de x . Puis, pour (3.5), il suffit de faire des intégrations par parties. \square

Remarque. La représentation de la solution de l'équation de la chaleur comme $\mathbb{E}[f(x+W_t)]$ montre que la trajectoire brownienne joue pour cette équation le même rôle que les caractéristiques, solutions d'équations différentielles du premier ordre, pour la résolution des EDP du premier ordre. L'équation (3.4) montre que ce résultat peut être étendu aux EDP elliptiques à coefficients constants à condition de se référer à un MB décentré.

Nous montrerons dans la suite de ce cours que si les coefficients de l'EDP ne sont plus constants, le résultat reste valable à condition de se référer à un processus stochastique, solution d'une EDS.

Mais l'étape la plus importante sera de passer de ce calcul, vrai en moyenne, à un calcul trajectorien. Cette étape a été amorcée par Paul Lévy dans les années 30 et complétée par K. Itô dans les années 50. Plus précisément, Itô interprète la quantité

$$T_t(f) = f(t, x + W_t) - f(0, x) - \int_0^t \frac{1}{2} f''_{xx}(s, x + W_s) + f'(s, x + W_s) ds$$

qui mesure la différence trajectorielle entre les deux termes de l'équation (3.3) sans espérance E comme une intégrale stochastique. Ce faisant, il introduit un calcul différentiel stochastique, le calcul d'Itô, vrai sur les trajectoires et non plus seulement en moyenne. Nous le développerons dans les chapitres suivants.

3.5 Versions multidimensionnelles du MB

Il y a deux types de généralisation de $(B_t)_{t \in \mathbb{R}^+}$ en dimension d : celle pour laquelle le processus est à valeurs dans \mathbb{R}^d (cf. le mouvement brownien multidimensionnel) et celles pour lesquelles le processus est indexé par $T = \mathbb{R}^d$ mais reste à valeurs réelles (cf. champ de Wiener-Chensov et fonction brownienne de Lévy).

3.5.1 Mouvement brownien multidimensionnel

Il s'agit du processus à valeurs dans \mathbb{R}^d $B_t = (B_t^1, \dots, B_t^d)$ où $(B_t^j)_t$ sont d mouvements browniens réels indépendants. Il suffit de considérer ce champ composante par composante et d'appliquer les résultats classiques du cadre unidimensionnel. Cette généralisation n'a donc pas beaucoup d'intérêt.

3.5.2 Champ de Wiener-Chensov

Soit $T = \mathbb{R}_+^d = \{t = (t_1, \dots, t_d), t_j \geq 0, 1 \leq j \leq d\}$ et le champ de Wiener-Chensov $\{B(t), t \in \mathbb{R}_+^d\}$ est le processus gaussien (en dimension d) défini par

$$\mathbb{E}[B(t)] = 0, \quad K(s, t) = \prod_{j=1}^d (s_j \wedge t_j).$$

Propriétés

- Pour $d = 1$, on a un mouvement brownien classique.
- Le champ est autosimilaire : $\tilde{B}(t) = (1/c^{d/2})B(ct)$ est encore un champ de Wiener-Chensov.
- Le champ a des accroissements stationnaires.
- Le champ a des accroissements indépendants (quand on les définit bien).

Quand $d = 2$, soit $\Delta = \Delta B = B(s_1, s_2) + B(t_1, t_2) - B(s_1, t_2) - B(t_1, s_2)$ associé à un rectangle. Si on a un système de rectangles Δ^i disjoints alors les variables aléatoires associées Δ^i sont indépendantes.

De même en dimension $d \geq 2$ en associant à des pavés des sommes d'accroissements.

- Inconvénient : la définition du champ dépend du choix du système de coordonnées (par exemple, il est nul sur les axes). En particulier, il n'est pas invariant par translation.
- Inconvénient : la définition ne peut pas être généralisée à la dimension infinie ($d = +\infty$).

3.5.3 Fonction brownienne de Lévy

Soit $T = \mathbb{R}^d$. La fonction brownienne de Lévy $\{B(t), t \in \mathbb{R}^d\}$ est le processus gaussien défini par

$$B(0) = 0, \quad \mathbb{E}[B(t)] = 0, \quad \mathbb{E}[(B(s) - B(t))^2] = \|s - t\|.$$

On a donc $\mathbb{E}[(B(t))^2] = \|t\|$ et $\mathbb{E}[B(s)B(t)] = (1/2)(\|s\| + \|t\| - \|t - s\|)$

Propriétés

- Pour $d = 1$, on retrouve le mouvement brownien classique.
- Autosimilarité : $\tilde{B}(t) = (1/c^{1/2})B(ct)$ est encore une fonction brownienne de Lévy.
- $\tilde{B}(t) = B(Ut)$ où $U : \mathbb{R}^d \rightarrow \mathbb{R}^d$ est une matrice unitaire est encore une fonction brownienne de Lévy.
- Inconvénient : par contre, il n'y a pas de généralisation de la propriété d'indépendance des accroissements.

Chapitre 4

Intégration stochastique

Dans ce chapitre, on définit l'intégrale stochastique par rapport à un MB et on introduit au calcul stochastique d'Itô. Il s'agit d'un calcul différentiel trajectorien (c'est à dire sur les trajectoires des processus) et pas en loi.

4.1 Variation quadratique

Si nous considérons une fonction $x : t \in [0, T] \mapsto x_t$ dérivable (à variations finies suffit) et f une fonction de classe C^1 , par la règle de dérivation composée, on a

$$f(x_t) = f(x_0) + \int_0^t f'(x_u)x'_u du = f(x_0) + \int_0^t f'(x_u)dx_u.$$

Lorsque la même chose est tentée avec un mouvement brownien W à la place de la fonction x , il apparait un terme avec un dérivée seconde. Cela est lié à la non dérivabilité du MB et à sa variation quadratique non nulle.

Proposition 4.1.1 *Soit $T > 0$ et $\{0 = t_0 < t_1 < \dots < t_n = T\}$ une partition de $[0, T]$. Par définition, la variation quadratique d'un mouvement brownien W associée à cette partition est la v.a.*

$$V(t_0, t_1, \dots, t_n) = \sum_{j=1}^n (W_{t_j} - W_{t_{j-1}})^2.$$

$V(t_0, t_1, \dots, t_n)$ converge dans L^2 vers T lorsque le pas de la subdivision $\delta := \max(t_j - t_{j-1})$ tend vers 0. De plus la convergence est ps lorsque la subdivision est uniforme.

Remarque. Pour une fonction x de classe C^1 , la variation quadratique tend vers 0, en effet

$$V(t_0, t_1, \dots, t_n) = \sum_{j=1}^n (x_{t_j} - x_{t_{j-1}})^2 = \sum_{j=1}^n (x'_{t_j^*} (t_j - t_{j-1}))^2$$

$$\leq \delta \|x'\|_\infty \sum_{j=1}^n |t_j - t_{j-1}| \leq \delta \|x'\|_\infty T$$

où $t_i^* \in]t_i, t_{i+1}[$ est donné par le théorème des accroissements finis appliqué à la fonction dérivable x .

Démonstration : Notons d'abord que le carré d'une v.a. gaussienne X centré de variance σ^2 est une v.a. d'espérance σ^2 et de variance $\text{Var}(X^2) = 3\sigma^4 - \sigma^4 = 2\sigma^4$ (par ipp).

1) On a pour la convergence dans L^2 :

$$\begin{aligned} \mathbb{E}[(V(t_0, t_1, \dots, t_n) - T)^2] &= E \left[\left(\sum_{j=1}^n (W_{t_j} - W_{t_{j-1}})^2 - (t_j - t_{j-1}) \right)^2 \right] \\ &= \sum_{j=1}^n \text{Var}((W_{t_j} - W_{t_{j-1}})^2) \end{aligned}$$

car les v.a. $(W_{t_j} - W_{t_{j-1}})^2 - (t_j - t_{j-1})$ sont centrées et indépendantes. On a donc

$$\mathbb{E}[(V(t_0, t_1, \dots, t_n) - T)^2] = 2 \sum_{j=1}^n (t_j - t_{j-1})^2 \leq 2T\delta \rightarrow 0$$

quand le pas $\delta \rightarrow 0$.

2) On considère maintenant la partition uniforme de $[0, T]$ avec les points $t_k = \frac{k}{n}$. Notons $U_j = \sqrt{\frac{n}{T}}(W(\frac{jT}{n}) - W(\frac{(j-1)T}{n}))$, ce sont des v.a. gaussiennes centrées réduites et indépendantes. D'après la LGN, on a :

$$Z_n = \sum_{j=1}^n \left(W\left(\frac{jT}{n}\right) - W\left(\frac{(j-1)T}{n}\right) \right)^2 = \frac{T}{n} \sum_{j=1}^n U_j^2$$

converge ps vers T . □

Comme conséquence de cette propriété, nous avons maintenant la formule suivante :

Proposition 4.1.2 *Soit W un MB standard. On a ps, pour tout t :*

$$W_t^2 = 2 \int_0^t W_s dW_s + t$$

où l'intégrale stochastique $\int_0^t W_s dW_s$ est à comprendre comme la limite ps de $\sum_{j=1}^n W_{t_{j-1}^n} (W_{t_j^n} - W_{t_{j-1}^n})$. La convergence a lieu aussi dans L^2 .

Ce résultat est à comparer avec l'analogie pour une fonction x de classe C^1 et nulle en 0 :

$$x(t)^2 = 2 \int_0^t x(s)x'(s)ds = 2 \int_0^t x(s)dx(s).$$

il n'y a pas de terme d'ordre 2 car dans ce cas la variation quadratique de $x \in C^1$ est nulle.

Démonstration : On considère la partition $\{t_i^n; i = 0, \dots, n\}$ de $[0, t]$. On a par une transformation d'Abel

$$\begin{aligned} W_t^2 &= \sum_{j=1}^n W_{t_j^n}^2 - W_{t_{j-1}^n}^2 \\ &= 2 \sum_{j=1}^n W_{t_{j-1}^n} (W_{t_j^n} - W_{t_{j-1}^n}) + \sum_{j=1}^n (W_{t_j^n} - W_{t_{j-1}^n})^2. \end{aligned}$$

Le deuxième terme converge (ps et dans L^2) vers la variation quadratique de W sur $[0, t]$:

Par suite, le premier terme converge ps et dans L^2 . Sa limite est par définition l'intégrale stochastique $\int_0^t W_s dW_s$. \square

4.2 Formule d'Itô

Le calcul différentiel s'étend à d'autres fonctions que la fonction $f(x) = x^2$ pour laquelle la Prop. 4.1.2 donne une formule d'Itô (noter bien qu'avec $f(x) = x^2$, $f(W_t) = W_t^2$). Ce calcul est nécessaire pour représenter trajectoriellement les variations infinitésimales de $f(x + W_t)$. Mais à la formule classique du cas déterministe, il faut ajouter un terme supplémentaire, dû au fait que la variation quadratique limite n'est pas nulle.

Théorème 4.2.1 Soit $f \in C_b^2(\mathbb{R})$ une fonction bornée à dérivées premières et secondes continues et bornées. Alors ps pour tout $t \in \mathbb{R}^+$, on a

$$f(x + W_t) = f(x) + \int_0^t f'_x(x + W_s) dW_s + \frac{1}{2} \int_0^t f''_{xx}(x + W_s) ds.$$

Le terme $\int_0^t f'_x(x + W_s) dW_s$ s'appelle l'intégrale stochastique de $f'(x + W_s)$ par rapport au MB W . C'est un processus stochastique à trajectoires continues. Il est défini comme la limite dans L^2 des sommes de type sommes de Riemann

$$S^\tau(f, W) = \sum_{i=1}^n f'_x(x + W_{t_{i-1}^n}) (W_{t_i^n} - W_{t_{i-1}^n})$$

lorsque le pas de la subdivision $\tau_n = \{0 = t_0 < \dots < t_n = t\}$ tend vers 0.

Démonstration :

a) On considère un temps $t > 0$ fixé et une partition $\tau_n = \{t_i^n : 0 \leq i \leq n\}$ de $[0, t]$. D'après la formule de Taylor-Lagrange à l'ordre 2, on peut écrire

$$f(x + W_t) = f(x) + \sum_{i=1}^n f'_x(x + W_{t_{i-1}^n}) (W_{t_i^n} - W_{t_{i-1}^n}) + \frac{1}{2} \sum_{i=1}^n f''_{xx}(\zeta_i) (W_{t_i^n} - W_{t_{i-1}^n})^2 \quad (4.1)$$

où $\zeta_i \in (x + W_{t_i^n}, x + W_{t_{i-1}^n})$ est un point aléatoire de l'intervalle aléatoire $(x + W_{t_i^n}, x + W_{t_{i-1}^n})$.

On étudie d'abord la limite du 3ème terme

$$\begin{aligned} & \left| \sum_{i=1}^n f''_{xx}(\zeta_i)(W_{t_i^n} - W_{t_{i-1}^n})^2 - \sum_{i=1}^n f''_{xx}(x + W_{t_{i-1}^n})(W_{t_i^n} - W_{t_{i-1}^n})^2 \right| \\ &= \left| \sum_{i=1}^n (f''_{xx}(\zeta_i) - f''_{xx}(x + W_{t_{i-1}^n}))(W_{t_i^n} - W_{t_{i-1}^n})^2 \right| \\ &\leq \sup_{1 \leq i \leq n} |f''_{xx}(\zeta_i) - f''_{xx}(x + W_{t_{i-1}^n})| \sum_{i=1}^n (W_{t_i^n} - W_{t_{i-1}^n})^2 \end{aligned}$$

Mais par uniforme continuité de f''_{xx}

$$\lim_{n \rightarrow +\infty} \sup_{1 \leq i \leq n} |f''_{xx}(\zeta_i) - f''_{xx}(x + W_{t_{i-1}^n})| = 0$$

et comme la variation quadratique de W (converge vers t), cette majoration tend vers 0 ps.

La convergence a lieu aussi dans L^2 car la variation quadratique converge aussi dans L^2 .

b) Finalement, il reste à étudier

$$A_n = \sum_{i=1}^n f''_{xx}(x + W_{t_{i-1}^n})(W_{t_i^n} - W_{t_{i-1}^n})^2$$

en comparant cette expression à

$$B_n = \sum_{i=1}^n f''_{xx}(x + W_{t_{i-1}^n})(t_i^n - t_{i-1}^n).$$

Pour cela, notons $M = \sup_{s \in [0, t]} |f''_{xx}(s)|$ la borne de la dérivée de f . Noter que

$$\begin{aligned} A_n &= A_{n-1} + f''_{xx}(x + W_{t_{n-1}^n})(W_{t_n^n} - W_{t_{n-1}^n})^2 \\ B_n &= B_{n-1} + f''_{xx}(x + W_{t_{n-1}^n})(t_n^n - t_{n-1}^n) \\ A_n - B_n &= A_{n-1} - B_{n-1} + f''_{xx}(x + W_{t_{n-1}^n})((W_{t_n^n} - W_{t_{n-1}^n})^2 - (t_n^n - t_{n-1}^n)). \end{aligned}$$

Noter que $(W_{t_n^n} - W_{t_{n-1}^n})^2 - (t_n^n - t_{n-1}^n)$ est une v.a. centrée, de variance $2(t_n^n - t_{n-1}^n)^2$. De plus, par indépendance des accroissements de W , elle est indépendante de $f''_{xx}(x + W_{t_{n-1}^n})$ et de $(A_{n-1} - B_{n-1})$.

On a alors

$$\begin{aligned} & \mathbb{E}[(A_n - B_n)^2] \\ &= \mathbb{E}[(A_{n-1} - B_{n-1})^2] + E \left[f''_{xx}(x + W_{t_{n-1}^n})^2 ((W_{t_n^n} - W_{t_{n-1}^n})^2 - (t_n^n - t_{n-1}^n))^2 \right] \end{aligned}$$

$$\begin{aligned}
& + 2E \left[f''_{xx}(x + W_{t_{n-1}^n}) [(W_{t_n^n} - W_{t_{n-1}^n})^2 - (t_n^n - t_{n-1}^n)] (A_{n-1} - B_{n-1}) \right] \\
& \leq \mathbb{E}[(A_{n-1} - B_{n-1})^2] + 2M^2(t_n^n - t_{n-1}^n)^2
\end{aligned}$$

où l'espérance du double produit est nulle par indépendance des v.a. et centrage de $(W_{t_n^n} - W_{t_{n-1}^n})^2 - (t_n^n - t_{n-1}^n)$.

Par récurrence sur n , il vient

$$\mathbb{E}[(A_n - B_n)^2] \leq 2M^2 \sum_{i=1}^n (t_i^n - t_{i-1}^n)^2 \leq 2M^2 \delta_n \sum_{i=1}^n (t_i^n - t_{i-1}^n) = 2M^2 T \delta_n \rightarrow 0$$

qui converge vers 0 puisque c'est la variation quadratique de la fonction $x(t) = t$ qui est à variation finie.

c) La convergence de B_n vers $\int_0^t f''_{xx}(x + W_s) ds$ est une simple conséquence de la convergence des sommes de Darboux vers l'intégrale de Riemann classique. Cette convergence est ps d'après les propriétés de l'intégrale de Riemann.

Finalement, le 3ème terme de (4.1) converge vers $\int_0^t f''_{xx}(W_s) ds$.

d) Pour montrer que la convergence a lieu aussi dans L^2 , il suffit de noter que les sommes partielles B_n et $\int_0^t f''_{xx}(x + W_s) ds$ sont majorées en module par MT pour appliquer le théorème de convergence dominée et conclure à la convergence dans L^2 .

e) Comme le 3ème terme de (4.1) converge ps et dans L^2 , le 2ème terme converge aussi. On note $\int_0^t f'_x(x + W_s) dW_s$ ce terme. Il s'agit d'une intégrale stochastique. \square

Remarque. Cette formule s'étend sans problème (autre que technique) à des fonctions $f(t, x)$ dépendant du temps de façon régulière. La formule d'Itô devient

$$f(t, x + W_t) = f(0, x) + \int_0^t f'_x(s, x + W_s) dW_s + \int_0^t f'_t(s, x + W_s) ds + \frac{1}{2} \int_0^t f''_{xx}(s, x + W_s) ds. \quad (4.2)$$

Les fonctions $u(t, x, f) = \mathbb{E}[f(x + W_t)]$ introduites dans le chapitre 3 en relation avec l'équation de la chaleur jouent un rôle très important.

Corollaire 4.2.1 Soit $f \in C_b^2$ une fonction bornée, deux fois dérivables à dérivées bornées et $u(t, x, f) = \mathbb{E}[f(x + W_t)]$ la fonction de classe C_b^2 en espace et C_b^1 en temps sur tout compact de $\mathbb{R}^+ \setminus \{0\}$ étudiée au chapitre 3.

La v.a. $f(x + W_T)$ admet une représentation comme intégrale stochastique

$$f(x + W_T) = \mathbb{E}[f(x + W_T)] + \int_0^T u'_x(T - s, x + W_s, f) dW_s$$

Il s'agit de ce qu'on appelle un théorème de représentation d'une famille de v.a. comme intégrale stochastique. Il convient de noter que dans cette représentation la fonction intégrée

dans l'intégrale stochastique n'est pas la dérivée de la fonction f mais celle de la solution de l'EDP, $u(t, x, f)$.

Démonstration : Nous appliquons la formule d'Itô à la fonction régulière $v(t, x, f) = u(T - t, x, f)$ où $u(t, x, f) = \mathbb{E}[f(x + W_t)]$ est solution de l'EDP $u'_t(t, x, f) = \frac{1}{2}u''_{xx}(t, x, f)$. Ceci nous donne que pour tout $t \leq T$

$$\begin{aligned} v(t, x + W_t, f) &= v(0, x, f) + \int_0^t v'_t(s, x + W_s, f)ds + \int_0^t v'_x(s, x + W_s, f)dW_s \\ &\quad + \frac{1}{2} \int_0^t v''_{xx}(s, x + W_s, f)ds \end{aligned}$$

c'est à dire

$$\begin{aligned} u(T - t, x + W_t, f) &= u(T, x, f) - \int_0^t u'_t(T - s, x + W_s, f)ds + \int_0^t u'_x(T - s, x + W_s, f)dW_s \\ &\quad + \frac{1}{2} \int_0^t u''_{xx}(T - s, x + W_s, f)ds. \end{aligned}$$

Soit compte tenu de l'EDP $u'_t(t, x, f) = \frac{1}{2}u''_{xx}(t, x, f)$:

$$u(T - t, x + W_t, f) = u(T, x, f) + \int_0^t u'_x(T - s, x + W_s, f)dW_s.$$

Il suffit maintenant de faire $t = T$ et de noter que $u(0, x, f) = f(x)$ pour conclure. \square

Ce théorème est particulièrement important, en finance notamment, car il permet de construire des stratégies de couverture pour des actifs financiers.

4.3 Et l'intégrale d'Itô ?

La notion d'intégrale stochastique $\int_0^t f(W_s)dW_s$ par rapport à un MB peut s'étendre à une classe plus générale que les fonctions déterministes f de carré intégrable ou de la forme $f'_x(x + W_s)$. De façon similaire, la formule d'Itô se généralise à ces situations.

Il faut introduire les notions de martingale, de processus d'Itô, de processus adapté.

Nous ne le ferons pas dans ce cours.

4.4 Équations différentielles stochastiques (EDS)

4.4.1 Équations différentielles et EDS

Les équations différentielles sont des équations d'évolution du type

$$\dot{x}(t) = b(t, x(t)) \tag{4.3}$$

où l'inconnue est une fonction $x(t)$ qui doit vérifier une équation impliquant sa dérivée \dot{x} et elle-même. Les cas les plus simples sont les équations différentielles d'ordre 1 comme en (4.3) (seule la dérivée 1ère est impliquée) avec $b(t, x) = a + bx$ indépendant de t et affine par rapport à x . Symboliquement, l'équation (4.3) se réécrit

$$dx(t) = b(t, x(t))dt. \quad (4.4)$$

Par exemple, l'équation $dx(t) = \mu(t)x(t)dt$ modélise le cours d'un actif financier $x(t)$ soumis au taux d'intérêt variable $\mu(t)$. Il est bien connu que la solution est

$$x(t) = x_0 \exp\left(\int_0^t \mu(s)ds\right).$$

Les EDS sont des généralisations des équations (4.4) où la dynamique déterministe d'évolution b est perturbée par un terme aléatoire (stochastique). On parle alors d'équation différentielle stochastique. En général la perturbation aléatoire est considéré comme un bruit. Par un argument du type TCL, il est légitime de considérer que ce bruit est un processus gaussien et en général il est modélisé par un mouvement brownien B :

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dB_t \quad (4.5)$$

où σ est une fonction du temps t et de l'inconnue au temps t X_t mais pourrait juste dépendre du temps (σ_t) ou de X ($\sigma(X_t)$) ou encore être constante σ .

La solution d'une EDS est aléatoire. Il s'agit donc d'un processus qu'on note X_t . En fait l'écriture (4.5) est symbolique car dB_t n'a pas vraiment de sens (le mouvement brownien n'est pas dérivable). Il faudrait écrire (4.5) sous la forme

$$X_t = X_0 + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dB_s.$$

En général (dans les cas simples), pour trouver la solution d'une EDS, on intuite la forme de la solution et on vérifie que l'EDS de départ est bien satisfaite en appliquant la formule d'Itô.

Il existe des résultats d'existence et d'unicité des EDS (du genre théorème de Cauchy-Lipschitz) mais nous n'en parlerons pas ici.

4.4.2 Exemples d'EDS

Black et Scholes

Considérons l'EDS

$$dX_t = \mu X_t dt + \sigma X_t dB_t \quad (4.6)$$

c'est à dire avec $b(t, x) = \mu x$ et $\sigma(t, x) = \sigma x$. Cette EDS modélise l'évolution d'un cours X soumis à un taux d'intérêt déterministe μ et à une perturbation stochastique $\sigma X_t dB_t$.

Le coefficient σ est appelé diffusion ou volatilité. Noter que la partie déterministe de l'accroissement de X_t (μX_t) et sa partie aléatoire (σX_t) sont toutes les deux proportionnelles à la valeur de l'actif X_t en t .

La solution de (4.6) est

$$X_t = X_0 \exp\left(\mu t - \frac{\sigma^2}{2}t + \sigma B_t\right).$$

En effet supposons $X_0 = 1$ et appliquons la formule d'Itô à $g(t, B_t)$ où $g(t, x) = \exp(\mu t - \frac{\sigma^2}{2}t + \sigma x)$. On a $g'_t(t, x) = (\mu - \frac{\sigma^2}{2})g(t, x)$ et $g'_x(t, x) = \sigma g(t, x)$ et $g''_{xx}(t, x) = \sigma^2 g(t, x)$. On a alors

$$\begin{aligned} X_t &= g(t, B_t) \\ &= g(0, 0) + \int_0^t g'_t(s, B_s)ds + \int_0^t g'_x(s, B_s)dB_s + \frac{1}{2} \int_0^t g''_{xx}(s, B_s)ds \\ &= 1 + \int_0^t (\mu - \frac{\sigma^2}{2})g(s, B_s)ds + \int_0^t \sigma g(s, B_s)dB_s + \frac{1}{2} \int_0^t \sigma^2 g(s, B_s)ds \\ &= 1 + \int_0^t (\mu X_s ds + \sigma X_s dB_s) \end{aligned}$$

c'est à dire sous forme symbolique $dX_t = \mu X_t dt + \sigma X_t dB_t$.

Comme dans le cas déterministe, les équations affines à coefficients éventuellement aléatoires mais bornés admettent des solutions explicites obtenues par la méthode de variation des constantes. Leur importance vient de ce qu'elles peuvent être considérées comme les linéarisées d'équations plus complexes.

Ornstein-Uhlenbeck

Il s'agit de l'équation

$$dX_t = -aX_t dt + \sigma dB_t \quad (4.7)$$

c'est à dire avec $b(t, x) = -ax$, et $\sigma(x) = \sigma$. La solution est donnée par

$$X_t = X_0 e^{-at} + \sigma \int_0^t e^{-a(t-s)} dB_s.$$

C'est immédiat en dérivant $X_t = X_0 e^{-at} + \sigma e^{-at} \int_0^t e^{as} dB_s$:

$$\begin{aligned} dX_t &= X_0(-ae^{-at})dt + \sigma(-ae^{-at})dt \int_0^t e^{as} dB_s + \sigma e^{-at} e^{at} dB_t \\ &= -a(X_0 e^{-at} + \sigma e^{-at} \int_0^t e^{as} dB_s)dt + \sigma dB_t \\ &= -aX_t dt + \sigma dB_t. \end{aligned}$$

Chapitre 5

Mouvement brownien fractionnaire

Les processus autosimilaires sont invariants en loi sous une certaine échelle d'espace et de temps. Ils sont particulièrement importants en modélisation. Parmi eux, nous présenterons en particulier le mouvement brownien fractionnaire (mBf).

En terme de modèle, dans la nature, l'autosimilarité est liée à un phénomène d'invariance et de persistance (appelé *effet Joseph* par Mandelbrot). Il traduit un phénomène cyclique et de persistance dans un état.

Historiquement, ce phénomène a été observé la première fois pour les crues du Nil dans les années 522-1284. Ces observations ont été formalisées par Tousson en 1925 et montrent que les cumuls du niveau du Nil recentré témoignent d'une certaine persistance.

Ces observations ont été poursuivies par un hydrologue Hurst qui a mis en évidence des règles statistiques en hydrologie, fondée sur l'autosimilarité des processus sous-jacents.

5.1 Généralités sur l'autosimilarité

On trouvera plus de détail sur cette notion dans [ST] dont la section est inspirée. Le coefficient d'autosimilarité ou indice d'autosimilarité est un nombre $0 \leq H \leq 1$.

Définition 5.1.1 *Un processus X est autosimilaire d'indice H si pour tout $a > 0$:*

$$\{X(at), t \in \mathbb{R}\} \stackrel{\mathcal{L}}{=} \{a^H X(t), t \in \mathbb{R}\}$$

au sens de l'égalité des lois fini-dimensionnelles.

Cette propriété montre qu'un changement d'échelle dans le temps est équivalent (en loi) à un changement d'échelle en espace. Attention, cependant au fait qu'il s'agit d'une égalité en loi et pas en trajectoire.

Exemple. Un mouvement brownien B est autosimilaire d'indice $1/2$. En effet, il suffit de comparer les fonctions de covariance des deux processus $(B(at))_t$ et $(a^{1/2}B(at))_t$ pour le constater, car ils ont alors même loi (centrés et même covariance).

En plus, le mouvement brownien a des accroissements indépendants. On verra des exemples

où les accroissements ne sont pas indépendants (mais restent stationnaires) mettant en évidence le phénomène de longue mémoire du processus.

Dans la suite, on prendra $T = \mathbb{R}$ ou \mathbb{R}^+ .

– On a en particulier

$$X(t) \stackrel{\mathcal{L}}{=} |t|^H X(\text{sign } t) = \begin{cases} t^H X_1 & \text{si } t > 0 \\ |t|^H X_{-1} & \text{si } t < 0. \end{cases}$$

– $X(0) = 0$ ps car $X(0) = X(a0) \stackrel{\mathcal{L}}{=} a^H X(0)$.

Dans la suite, on s'intéresse à des processus autosimilaires d'indice $H > 0$.

Remarque : Un processus autosimilaire ne peut pas être en plus stationnaire car on aurait

$$X(t) \stackrel{\mathcal{L}}{=} X(at) \stackrel{\mathcal{L}}{=} a^H X(t).$$

On a en particulier $\mathbb{E}[X(t)] = a^H \mathbb{E}[X(t)]$, ce qui donne une contradiction quand on fait tendre $a^H \rightarrow +\infty$ ($H > 0$).

Cependant, il existe un lien entre les processus autosimilaires et les processus stationnaires

Proposition 5.1.1 *Soit $(X_t)_{t \geq 0}$ H -autosimilaire, alors*

$$Y_t = e^{-tH} X(e^t), \quad t \in \mathbb{R}$$

est stationnaire. Réciproquement, si $(Y_t)_{t \in \mathbb{R}}$ est stationnaire alors $X(t) = t^H Y(\ln t)$ est H autosimilaire.

Démonstration : Exercice facile □

Par exemple le processus stationnaire ainsi associé au MB (1/2 autosimilaire) est le processus d'O.U.

Le résultat suivant établit le lien entre l'intégrabilité d'un processus autosimilaire et son indice d'autosimilarité.

Proposition 5.1.2 *Soit $(X_t)_t$ un processus autosimilaire et à accroissements stationnaires tel que $\mathbb{P}(X(1) \neq 0) > 0$. On suppose que $\mathbb{E}[|X(1)|^\gamma] < +\infty$ alors*

$$\begin{aligned} 0 < H < 1/\gamma & \quad \text{si} \quad 0 < \gamma < 1, \\ 0 < H \leq 1 & \quad \text{si} \quad \gamma \geq 1. \end{aligned}$$

Démonstration : Cf. [ST] □

5.2 Définition du mouvement brownien fractionnaire

Le mouvement brownien fractionnaire a été introduit par Kolmogorov en 1940. Son étude a été reprise et approfondie par Mandelbrot et van Ness en 1968.

Définition 5.2.1 (mBf) Soit $0 < H \leq 1$. Le mouvement brownien fractionnaire (mBf) d'indice H $(B^H(t))_t$ est le processus gaussien centré de fonction de covariance

$$K_H(t_1, t_2) = \frac{1}{2}(|t_1|^{2H} + |t_2|^{2H} - |t_1 - t_2|^{2H}) \text{Var}(X_1).$$

Le coefficient d'autosimilarité H est aussi appelé **paramètre de Hurst**. On verra qu'il est lié à la régularité des trajectoires.

Noter que la variance de $B^H(t)$ est $t^{2H} \text{Var}(X_1)$ si bien que

$$B^H(t) \sim \mathcal{N}(0, t^{2H} \text{Var}(X_1)).$$

Proposition 5.2.1 Le mBf d'indice H est H -autosimilaire et à accroissements stationnaires. Réciproquement, si $(X_t)_{t \in \mathbb{R}}$ est un processus gaussien H autosimilaire et à accroissements stationnaires. Alors

$$0 < H \leq 1, \quad X(0) = 0 \quad ps$$

et

$$\text{Cov}(X(t_1), X(t_2)) = \frac{1}{2}(|t_1|^{2H} + |t_2|^{2H} - |t_1 - t_2|^{2H}) \text{Var}(X_1).$$

De plus pour tout $t \in \mathbb{R}$, on a

- i) Si $0 < H < 1$: $\mathbb{E}[X(t)] = 0$,
- ii) Si $H = 1$: $X(t) = tX(1)$ ps.

Ce résultat montre que le mBf est le seul processus gaussien H -autosimilaire à accroissements stationnaires.

Démonstration : Il est facile de vérifier que le mBf est H -autosimilaire et à accroissements stationnaires : il suffit de faire les calculs avec la fonction de covariance.

$$\begin{aligned} \text{Cov}(X(at_1), X(at_2)) &= \frac{1}{2}(|at_1|^{2H} + |at_2|^{2H} - |at_1 - at_2|^{2H}) \text{Var}(X_1) \\ &= \frac{a^{2H}}{2}(|t_1|^{2H} + |t_2|^{2H} - |t_1 - t_2|^{2H}) \text{Var}(X_1) \\ &= a^{2H} \text{Cov}(X(t_1), X(t_2)) \\ &= \text{Cov}(a^H X(t_1), a^H X(t_2)). \end{aligned}$$

Les processus $(X(at))_t$ et $(a^H X(t))_t$ ont même covariance donc même loi. On vérifie aussi la stationarité des accroissements par la fonction de covariance :

$$\begin{aligned} \text{Var}(X(t_2) - X(t_1)) &= \text{Cov}(X(t_2) - X(t_1), X(t_2) - X(t_1)) \\ &= \text{Cov}(X(t_2), X(t_2)) - 2\text{Cov}(X(t_2), X(t_1)) + \text{Cov}(X(t_1), X(t_1)) \\ &= t_2^{2H} - (t_1^{2H} + t_2^{2H} - |t_1 - t_2|^{2H}) + t_1^{2H} \\ &= |t_1 - t_2|^{2H}, \end{aligned}$$

ce qui établit le sens direct. Pour la réciproque, notons d'abord que, par autosimilarité, on a $X(0) = 0$ ps : en effet comme $X(0) \stackrel{\mathcal{L}}{=} a^H X(0)$, on a $\mathbb{E}[X(0)^2] = a^{2H} \mathbb{E}[X(0)^2]$ et donc $\mathbb{E}[X(0)^2] = 0$, c'est à dire $X(0) = 0$ ps. Puis encore avec l'autosimilarité, on a :

$$\begin{aligned} \mathbb{E}[X(t_1)X(t_2)] &= \frac{1}{2} (\mathbb{E}[X(t_1)^2] + \mathbb{E}[X(t_2)^2] - \mathbb{E}[(X(t_1) - X(t_2))^2]) \\ &= \frac{1}{2} (\mathbb{E}[X(t_1)^2] + \mathbb{E}[X(t_2)^2] - \mathbb{E}[(X(t_1 - t_2) - X(0))^2]) \\ &= \frac{1}{2} (|t_1|^{2H} + |t_2|^{2H} - |t_1 - t_2|^{2H}) \mathbb{E}[X(1)^2]. \end{aligned} \quad (5.1)$$

On admet que $0 < H \leq 1$. Supposons d'abord $0 < H < 1$. Comme

$$\mathbb{E}[X(1)] = \mathbb{E}[X(2) - X(1)] = 2^H \mathbb{E}[X(1)] - \mathbb{E}[X(1)],$$

on en déduit $\mathbb{E}[X(1)] = 0$ puis par autosimilarité $\mathbb{E}[X(t)] = 0$ pour $t \geq 0$. Mais comme par stationarité

$$X(1) = X(1) - X(0) \stackrel{\mathcal{L}}{=} X(0) - X(-1) = X(-1),$$

les variables aléatoires $X(1)$ et $X(-1)$ ont en particulier même espérance $\mathbb{E}[X(-1)] = -\mathbb{E}[X(1)]$ si bien que $\mathbb{E}[X(t)] = |t|^H \mathbb{E}[X(\text{sign } t)]$. On en déduit que X est un processus centré avec la bonne fonction de covariance.

Supposons maintenant $H = 1$. D'après (5.1), on a $\mathbb{E}[X(t_1)X(t_2)] = t_1 t_2 \mathbb{E}[X^2(1)]$ et donc

$$\begin{aligned} \mathbb{E}[(X(t) - X(1))^2] &= \mathbb{E}[X(t)^2] - 2t\mathbb{E}[X(t)X(1)] + t^2\mathbb{E}[X(1)^2] \\ &= (t^2 - 2t \times t + t^2)\mathbb{E}[X(1)^2] \\ &= 0. \end{aligned}$$

On a donc $X(t) = tX(1)$ ps pour tout t . La fonction d'autocovariance suit facilement. \square

Remarque :

– Pour $H = 1/2$, le mBf est le mouvement brownien. En effet dans ce cas

$$\mathbb{E}[B_{1/2}(t_1)B_{1/2}(t_2)] = \begin{cases} \sigma_0^2(t_1 \wedge t_2) & \text{si } t_1, t_2 \text{ sont de mêmes signes} \\ 0 & \text{si } t_1, t_2 \text{ sont de signes opposés.} \end{cases}$$

où σ_0^2 est la variance du mouvement brownien.

- Pour $H = 1$, le mBf est une droite aléatoire :

$$B_1(t) = t(X + \mu)$$

où $Z \simeq \mathcal{N}(0, \sigma_0^2)$ et $\mu \in \mathbb{R}$.

- Le seul mBf avec des accroissements indépendants est le mouvement brownien (pour $H = 1/2$). En effet pour $t_1 < t_2 < t_3 < t_4$, on a

$$\begin{aligned} & \text{Cov}(X(t_2) - X(t_1), X(t_4) - X(t_3)) \\ &= \text{Cov}(X(t_2), X(t_4)) - \text{Cov}(X(t_2), X(t_3)) - \text{Cov}(X(t_1), X(t_4)) + \text{Cov}(X(t_1), X(t_3)) \\ &= \frac{1}{2}(|t_2|^{2H} + |t_4|^{2H} - (t_4 - t_2)^{2H} - |t_2|^{2H} - |t_3|^{2H} + (t_3 - t_2)^{2H} \\ & \quad - |t_1|^{2H} - |t_4|^{2H} + (t_4 - t_1)^{2H} + |t_1|^{2H} + |t_3|^{2H} - (t_3 - t_1)^{2H}) \\ &= \frac{1}{2}((t_3 - t_2)^{2H} + (t_4 - t_1)^{2H} - (t_3 - t_1)^{2H} - (t_4 - t_2)^{2H}). \end{aligned}$$

Cette covariances ne s'annule que pour $2H = 1$, ie. dans le cas du mouvement brownien.

Le cas $H = 1$ est un cas dégénéré. Dans la suite, on supposera donc $0 < H < 1$.

On a défini le mBf par ses lois fini dimensionnelles. Toutefois, cela ne donne pas une idée très précise du processus. On en donnera dans les sections suivantes une représentation sous forme d'intégrale stochastique.

5.3 Régularité des trajectoires

5.3.1 Hölder

Définition 5.3.1 (Hölder) *On rappelle qu'une fonction $f : \mathbb{R}^p \rightarrow \mathbb{R}^d$ est dite α -höldérienne s'il existe $C < +\infty$ tel que*

$$\|f(x) - f(y)\| \leq C\|x - y\|^\alpha$$

où $\|\cdot\|$ désigne la norme ambiantes de \mathbb{R}^p ou de \mathbb{R}^d .

Remarque :

- Pour $\alpha = 1$, on retrouve les fonctions lipschitziennes.
- Si $\alpha > 1$, il est facile de voir que f est constante (dérivable de dérivée nulle). Le caractère α -höldérien n'a donc d'intérêt que pour $\alpha \leq 1$.

Proposition 5.3.1 *Le mBf B^H d'indice H est α -höldérien pour tout $\alpha < H$.*

Démonstration : Par stationnarité, on a

$$\text{Var}(B_t^H - B_s^H) = \mathbb{E}[(B_t^H - B_s^H)^2] = \mathbb{E}[(B_{t-s}^H)^2] = (t - s)^{2H}.$$

Le théorème de régularité pour les processus gaussiens (th. 2.2.1) est satisfait avec $c = 1$ et $a = 2H$. Il assure que les trajectoires sont ps α -Hölderienne pour tout $\alpha < a/2 = H$. \square

Proposition 5.3.2 (Variation quadratique du mBf) *Soit B^H un mBf d'indice H . Alors on a ps*

$$\lim_{n \rightarrow +\infty} n^{2H-1} \sum_{k=0}^{n-1} (B_{\frac{k+1}{n}}^H - B_{\frac{k}{n}}^H)^2 = 1.$$

Démonstration : On se contente d'une justification heuristique : comme B^H est α -höldérien pour tout $\alpha < H$, on fait comme s'il était H -höldérien. On a alors $B_{\frac{k+1}{n}}^H - B_{\frac{k}{n}}^H \asymp (\frac{k+1}{n} - \frac{k}{n})^H$ et on comprend donc que

$$\sum_{k=0}^{n-1} (B_{\frac{k+1}{n}}^H - B_{\frac{k}{n}}^H)^2 \asymp n(1/n)^{2\alpha} = n^{1-2H}.$$

Comme s'est vrai pour tout $\alpha < H$, on a $n^{2H-1} \sum_{k=0}^{n-1} (B_{\frac{k+1}{n}}^H - B_{\frac{k}{n}}^H)^2 \asymp 1$.

En fait, une justification rigoureuse de ce résultat repose sur une LGN pour des variables aléatoires non indépendantes mais dont on contrôle la corrélation en $+\infty$ (c'est le cas pour le mBf). \square

Finalement, pour la variation quadratique, on peut annoncer que quand $n \rightarrow +\infty$:

$$v_n := \sum_{k=0}^{n-1} (B_{\frac{k+1}{n}}^H - B_{\frac{k}{n}}^H)^2 \simeq n^{1-2H}.$$

On retrouve que la variation quadratique

- est 0 si $H > 1/2$,
- est 1 si $H = 1/2$ (dans ce cas, il s'agit du mouvement brownien),
- est $+\infty$ si $H < 1/2$.

On constate donc que la régularité du mBf d'indice H augmente avec H .

On déduit de la Prop. 5.3.2 un estimateur du paramètre de Hurst H à partir de la variation quadratique (qui est facilement calculable à partir d'une trajectoire) :

Proposition 5.3.3 $\hat{H}_n = \frac{1}{2} \left(1 - \frac{\ln(v_{2n}/v_n)}{\ln 2} \right)$ est un estimateur de H .

Démonstration : En effet, on a

$$\frac{v_{2n}}{v_n} \asymp \frac{(2n)^{1-2H}}{n^{1-2H}} = 2^{1-2H} = e^{(1-2H)\ln 2},$$

ce qui permet de conclure facilement. \square

À l'aide du TCL, on peut donner des vitesses de convergence de l'estimateur \hat{H}_n vers H .

5.3.2 Exposant de Hölder

Le caractère α -höldérien des trajectoires du mBf est une propriété globale qu'on souhaite localiser pour étudier la régularité localement. Pour cela, on mesure la régularité locale d'un processus X par :

Définition 5.3.2 (Régularité locale) *On peut mesurer la régularité d'un processus $(X(t))_t$ en un point t par*

– *l'exposant de Hölder ponctuel :*

$$\alpha_X(t, \omega) = \sup \left(\alpha : \lim_{h \rightarrow 0} \frac{X(t+h, \omega) - X(t, \omega)}{|h|^\alpha} = 0 \right),$$

– *l'exposant de Hölder local :*

$$\tilde{\alpha}_X(t, \omega) = \sup \left(\tilde{\alpha} : \exists \rho > 0, \sup_{s, s' \in]t-\rho, t+\rho[} \frac{|X(s, \omega) - X(s', \omega)|}{|s - s'|^{\tilde{\alpha}}} < +\infty \right).$$

Ce coefficient prend en compte le comportement de X au voisinage de t et pas seulement en t .

Noter qu'on a toujours $\tilde{\alpha}_X(t, \omega) \leq \alpha_X(t, \omega)$. Et pour le mBf, on comprend qu'on a, par la stationnarité de ses accroissements :

Proposition 5.3.4 *Pour le mBf B^H avec un paramètre de Hurst H constant, on a ps*

$$\tilde{\alpha}_{B^H}(t, \omega) = \alpha_{B^H}(t, \omega) = H.$$

Plus H est grand, plus le mBf est régulier (comparer déjà les cas $H = 1/2$: brownien et $H = 1$: droite aléatoire). Dans le cas extrême ($H = 1$), on a vu que le mBf est une droite aléatoire. La trajectoire (une droite) est alors parfaitement régulière.

Cependant, pour un mBf donné, son indice H est fixé et la régularité des trajectoires du mBf est ps constante. Cela restreint l'intérêt du mBf dans les modélisations et justifie l'introduction d'une généralisation du mBf : le mouvement brownien multifractionnaire dont l'indice H varie avec t .

Une motivation historique (due à Voss R. F. (1985)) est la modélisation des irrégularités de l'allure d'une montagne par un mBf B_t^H où t représente ici le lieu. Pour H proche de 1, la trajectoire (donc l'allure de la montagne) a un aspect assez régulier et on modélise donc un massif montagneux assez vieux (et érodé). Par contre pour H petit, la trajectoire (l'allure de la montagne) a un aspect plus abrupt et on modélise ainsi un massif montagneux assez jeune (peu érodé).

Mais en utilisant le mBf, on a un modèle qui présente une irrégularité uniforme tout le long de la montagne. Cela ne tient pas compte de la nature géologique qui sur une même montagne peut varier et entraîner des allures différentes selon le lieu. Cela motive l'introduction du mouvement brownien multifractionnaire (mBm) pour lequel le paramètre de H devient une fonction $H(t)$. Dans le cas de la modélisation de l'allure d'une montagne, la fonction H prend en compte les spécificités géologiques du terrain au point t .

Avec le mBm, on pourra modéliser des trajectoires dont la régularité varie au cours du temps (ou de l'espace selon l'interprétation du paramètre t) : allure d'une montagne soumise à l'érosion, image radiologique non uniforme, analyse financière (cf. [LV]). On renvoie à la section 5.6 pour une introduction au mBm.

5.4 Représentations intégrales du mBf

5.4.1 Représentation en moyenne mobile

On considère M un bruit blanc gaussien sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda/2)$. Nous avons vu au chapitre ?? la construction des intégrales stochastiques pour $f \in L^2(\mathbb{R}, \lambda)$

$$I(f) = \int_{-\infty}^{+\infty} f dM.$$

Notons $u_+ = \max(u, 0)$ et $u_- = \max(-u, 0)$ si bien que $u = u_+ - u_-$ avec $u_+, u_- \geq 0$.

Proposition 5.4.1 *Soit $0 < H < 1$. Le mBf $(B^H(t))_t$ a la représentation intégrale*

$$\frac{1}{C_1(H)} \int_{-\infty}^{+\infty} (((t-s)_+)^{H-1/2} - ((-s)_+)^{H-1/2}) M(ds) \quad (5.2)$$

$$\text{où } C_1(H) = \left(\int_0^{+\infty} ((1+s)^{H-1/2} - s^{H-1/2})^2 ds + \frac{1}{2H} \right)^{1/2}.$$

Remarque : Quand $H = 1/2$, $C_1(1/2) = 1$ et la représentation (5.2) est à interpréter comme $\int_0^t M(ds)$ si $t \geq 0$ et $-\int_t^0 M(ds)$ si $t < 0$, on obtient alors la représentation du mouvement brownien comme primitive de bruit blanc gaussien.

Les propriétés trajectoires du mBf sont plus facilement étudiables avec ce type de représentation intégrale.

Esquisse de preuve. Il s'agit d'abord de vérifier que les intégrands proposés sont bien dans $L^2(\mathbb{R})$.

Puis on calcule la fonction de covariance de (5.2).

On conclut par unicité en appliquant la Prop. 5.2.1.

La constante $C_1(H)$ est là pour assurer que la variance est bien réduite à 1. \square

La représentation (5.2) n'est pas unique. Par exemple, avec $u_- = \max(-u, 0)$, on montre que pour a et b quelconques,

$$\int_{-\infty}^{+\infty} \left\{ a \left[(t-s)_+^{H-1/2} - (-s)_+^{H-1/2} \right] + b \left[(t-s)_-^{H-1/2} - (-s)_-^{H-1/2} \right] \right\} M(ds)$$

est aussi une représentation du mBf (à une constante multiplicative près). Celle vue en (5.2) correspond à $a = 1/C_1(H)$ et $b = 0$. Celle en (5.2) est non-anticipative car l'intégrale se calcule pour les $s \leq t$.

Pour $a = b = 1$, on a une représentation dite « équilibrée » du mBf

$$\int_{-\infty}^{+\infty} (|t-s|^{H-1/2} - |s|^{H-1/2}) M(ds).$$

5.4.2 Représentation harmonisable

On propose une représentation harmonisable du mBf du type

$$I(\tilde{f}_t) = \int_{-\infty}^{+\infty} \tilde{f}_t(x) \tilde{M}(dx)$$

où \tilde{f}_t et \tilde{M} sont maintenant à valeurs complexes. En général, $I(\tilde{f}_t)$ est aussi à valeurs complexes. Mais sous de bonnes conditions sur la mesure aléatoire complexe \tilde{M} , l'intégrale est bien à valeurs réelles.

De façon à avoir un résultat réel, on va considérer une mesure aléatoire complexe $\tilde{M} = M_1 + iM_2$ qui n'est pas orthogonale : même si $A \cap B = \emptyset$, $M_i(A)$ et $M_i(B)$ ne sont pas nécessairement indépendantes $1 \leq i \leq 2$.

Heuristiquement, il faut voir \tilde{f}_t et \tilde{M}_t comme des transformées de Fourier resp. de f_t et de M_t , à valeurs réelles. Par une relation de type Parseval sur les covariances, on aura

$$\int_{-\infty}^{+\infty} f_t(x) M(dx) \stackrel{\mathcal{L}}{=} \int_{-\infty}^{+\infty} \tilde{f}_t(x) \tilde{M}(dx).$$

Pour cela, on doit avoir

$$\tilde{f}_t(x) = \overline{\tilde{f}_t(-x)}, \quad \tilde{M}(A) = \overline{\tilde{M}(-A)}.$$

Concrètement, on prend

- $\widetilde{M} = M_1 + iM_2$ où M_1 et M_2 sont des mesures aléatoires gaussiennes indépendantes contrôlées par $\lambda/4$. Ainsi $\mathbb{E}[|\widetilde{M}(A)|^2] = \lambda(A)$ tandis que $\mathbb{E}[\widetilde{M}(A)^2] = 0$.
- $\tilde{f} = f_1 + if_2$ avec $f_1(x) = f_1(-x)$, $f_2(x) = -f_2(-x)$ et $\int_{-\infty}^{+\infty} f_i(x)^2 dx < +\infty$, c'est à dire \tilde{f} est dans

$$\mathcal{F} = \{\tilde{f} \mid \tilde{f}(-x) = \overline{\tilde{f}(x)} \text{ et } \int_{-\infty}^{+\infty} |\tilde{f}(x)|^2 dx < +\infty\}.$$

On montre alors que (cf. [ST, p. 326])

$$I(\tilde{f}) = \int_{-\infty}^{+\infty} \tilde{f}(x) \widetilde{M}(dx) = \int_{-\infty}^{+\infty} f_1(x) M_1(dx) - \int_{-\infty}^{+\infty} f_2(x) M_2(dx).$$

Puis on montre que $I(\tilde{f})$ est gaussien réel centré et de variance

$$\mathbb{E}[I(\tilde{f})^2] = \int_{-\infty}^{+\infty} |\tilde{f}(x)|^2 dx < +\infty$$

et même de covariance

$$\mathbb{E}[I(\tilde{f}_1)I(\tilde{f}_2)] = \int_{-\infty}^{+\infty} \tilde{f}_1(x) \overline{\tilde{f}_2(x)} dx. \quad (5.3)$$

Dans la suite, on prend des fonctions

$$f_t \in \mathcal{F}_0 = \left\{ f_t \mid \int_{-\infty}^{+\infty} |f_t(x)|^2 dx < +\infty, f_t \in \mathbb{R} \right\}.$$

On considère la transformée de Fourier $\tilde{f}_t(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{ixu} f_t(u) du$ d'une telle fonction $f_t \in \mathcal{F}_0$. On a bien $\tilde{f} \in \mathcal{F}$. Puis

Proposition 5.4.2

$$\left\{ I(f_t) = \int_{-\infty}^{+\infty} f_t(x) M(dx), t \in T \right\} \stackrel{\mathcal{L}}{=} \left\{ I(\tilde{f}_t) = \int_{-\infty}^{+\infty} \tilde{f}_t(x) \widetilde{M}(dx), t \in T \right\}.$$

Démonstration : Comme $\{I(f_t), t \in T\}$ et $\{I(\tilde{f}_t), t \in T\}$ sont gaussiens et centrés, il suffit de montrer l'égalité des fonctions de covariance, c'est à dire pour $f_{t_1}, f_{t_2} \in \mathcal{F}_0$: $\mathbb{E}[I(f_{t_1})I(f_{t_2})] = \mathbb{E}[I(\tilde{f}_{t_1})I(\tilde{f}_{t_2})]$.

Mais cela vient de la Prop. ??, de (5.3) et de l'identité de Parseval pour la transformée de Fourier. \square

On en déduit une représentation harmonisable du mBf :

Proposition 5.4.3 (Représentation harmonisable) Soit $0 < H < 1$. Alors le mBf standard $(B^H(t))_{t \in \mathbb{R}}$ a la représentation intégrale

$$\left\{ \frac{1}{C_2(H)} \int_{-\infty}^{+\infty} \frac{e^{ixt} - 1}{ix} |x|^{-(H-1/2)} \widetilde{M}(dx), \quad t \in \mathbb{R} \right\} \quad (5.4)$$

$$\text{où } C_2(H) = \left(\frac{\pi}{H\Gamma(2H) \sin(H\pi)} \right)^{1/2}.$$

Démonstration : Soit $X(t)$ l'intégrale donnée en (5.4) et $\tilde{f}_t(x)$ son intégrand. Comme $|\frac{e^{ixt}-1}{ix}|$ est bornée en 0 et se comporte comme $1/x$ en $\pm\infty$, on a $\int_{-\infty}^{+\infty} |\tilde{f}_t(x)|^2 dx < +\infty$. $X(t)$ est donc bien définie. Il est facile de voir qu'il s'agit d'un processus gaussien H -autosimilaire et à accroissements stationnaires (faire les calculs avec la fonction de covariance). C'est donc un mBf d'après la Proposition 5.1.1. Il reste à vérifier que $\mathbb{E}[X(1)^2] = 1$, c'est à dire

$$\int_{-\infty}^{+\infty} |\tilde{f}_1(x)|^2 dx = C_2^2(H).$$

C'est un calcul technique que l'on admettra. □

Remarque :

- Avec $H = 1/2$, $C_2(1/2) = \sqrt{2\pi}$, on en déduit une représentation harmonisable du mouvement brownien :

$$\left\{ B_t = \sqrt{2\pi} \int_{-\infty}^{+\infty} \frac{e^{itx} - 1}{ix} \widetilde{M}(dx) : t \geq 0 \right\}.$$

- Pour tout réel a, b , le processus

$$\left\{ \int_{-\infty}^{+\infty} \frac{e^{itx} - 1}{ix} (a(x_+)^{-(H-1/2)} + b(x_-)^{-(H-1/2)}) \widetilde{M}(dx) : t \in \mathbb{R} \right\}$$

est aussi un mBf (à un facteur multiplicatif près) car il est gaussien H -autosimilaire et à accroissements stationnaires.

la représentation (5.4) correspond au choix $a = b = 1/C_2(H)$.

- La représentation (5.4) (ou la précédente) a une interprétation physique. Le facteur $(e^{ixt} - 1)/(ix)$ est la transformée de Fourier de la fonction indicatrice de $[0, t]$ et assure la stationnarité des accroissements.

Le terme $|x|^{-(H-1/2)}$ contribue au terme $\| |x|^{-(H-1/2)} \|^2 = x^{-(2H-1)}$ dans la densité spectrale, qui

- diverge en basse fréquence ($x \rightarrow 0$) quand $1/2 < H < 1$,
- est sans influence quand $H = 1/2$,
- tend vers 0 en basse fréquence quand $0 < H < 1/2$.

5.5 Généralisations multidimensionnelles du mBf

D'abord on considère les versions multidimensionnelles du mBf, c'est à dire indexées par des paramètres multi-indices. Comme pour le mouvement brownien en section 3.5, il y a deux types de généralisation possible : une du type Wiener-Chensov, l'autre de type Lévy.

5.5.1 Drap brownien fractionnaire

C'est l'extension de type Wiener-Chensov. Soit $T = \mathbb{R}_+^d = \{t = (t_1, \dots, t_d) : t_j \geq 0, 1 \leq j \leq d\}$ et $H = (H_1, \dots, H_d) \in]0, 1[^d$. Le drap brownien fractionnaire $\{B^H(t), t \in \mathbb{R}_+^d\}$ est le processus gaussien (en dimension d) défini par

$$\mathbb{E}[B^H(t)] = 0, \quad \text{Cov}(B^H(s), B^H(t)) = \frac{1}{2^d} \prod_{j=1}^d (|t_j|^{2H_j} + |s_j|^{2H_j} - |t_j - s_j|^{2H_j}).$$

Propriétés

- Pour $d = 1$, on a un mBf classique.
- Pour $H = 1/2$, on a un champ de Wiener-Chensov.
- Il s'agit d'une extension anisotrope du mBf (le paramètre de Hurst H_i n'est pas le même dans chaque direction).
- Le champ est autosimilaire dans chaque direction axiale. Il ne l'est globalement que si tous les H_i sont égaux.
- Le champ a des accroissements stationnaires.
- Inconvénient : la définition du champ dépend du choix du système de coordonnées (par exemple, il est nul sur les axes). En particulier, il n'est pas invariant par translation ou rotation.
- Inconvénient : la définition ne peut pas être généralisée à la dimension infinie ($d = +\infty$).

5.5.2 Fonction brownienne fractionnaire de Lévy

Soit $T = \mathbb{R}^d$ et $H \in]0, 1[$. Le champ brownien fractionnaire (de Lévy) $\{B^H(t), t \in \mathbb{R}^d\}$ est le processus gaussien défini dans \mathbb{R}^d par

$$B^H(0) = 0, \quad \mathbb{E}[B^H(t)] = 0, \quad \mathbb{E}[B^H(s)B^H(t)] = (1/2)(\|s\|^{2H} + \|t\|^{2H} - \|t - s\|^{2H}).$$

Propriétés

- Pour $d = 1$, on retrouve le mouvement brownien classique.
- Pour $H = 1/2$, on retrouve la fonction brownienne fractionnaire de Lévy.
- Il s'agit d'une extension isotrope : dans toutes les directions le paramètre de Hurst est H .
- Autosimilarité : $\tilde{B}^H(t) = (1/c^H)B^H(ct)$ est encore un champ brownien de Lévy.
- $\hat{B}^H(t) = B^H(Ut)$ où $U : \mathbb{R}^d \rightarrow \mathbb{R}^d$ est une matrice unitaire est encore une fonction fractionnaire de Lévy.

5.6 Mouvement brownien multifractionnaire (mBm)

Il s'agit d'un mBf avec un paramètre de Hurst H qui est une fonction du temps (ou de l'espace) assez régulière : $H : \mathbb{R}_+ \rightarrow]0, 1[$ β -höldérienne.

5.6.1 Définition et propriétés du mBm

Définition 5.6.1 *Le mouvement brownien multifractionnaire noté mBm est le processus gaussien $(B^H(t))_{t \in \mathbb{R}_+}$ centré défini par*

$$\mathbb{E}[B^H(s)B^H(t)] = (\sigma^2/2)(|s|^{2H_t+2H_s} + |t|^{2H_t+2H_s} - |t-s|^{2H_t+2H_s}).$$

On a à nouveau des représentations de ce processus sous forme d'intégrales stochastiques. En particulier, on a une représentation moyenne mobile :

$$B^{H_t}(t) = \frac{1}{\Gamma(H_t + 1/2)} \int_{-\infty}^0 [(t-s)^{H_t-1/2} - (-s)^{H_t-1/2}] dM(s) + \int_0^t (t-s)^{H_t-1/2} dM(s) \quad (5.5)$$

où M est un bruit blanc gaussien standard (contrôlé par la mesure de Lebesgue λ)

Remarque : Le mBm est une extension naturelle du mBf mais perd certaines de ses propriétés :

- le mBm n'est plus autosimilaire ;
- les accroissements du mBm ne sont plus stationnaires.

Cependant, on a :

Proposition 5.6.1 *Presque sûrement, le mBm $B^{H_t}(t)$ est une fonction continue de t .*

Démonstration : Cf. [P-LV]. □

Pour le mBf, on a même mieux : le mBf $(H, t) \rightarrow B^H(t)$ est uniformément continue par rapport à H :

Proposition 5.6.2 *Soit $B^H(t)_{t \geq 0}$ un mBf d'indice $H \in]0, 1[$. Alors pour tout intervalle $[a, b] \subset]0, 1[$ et $K > 0$, on a ps*

$$\lim_{h \rightarrow 0} \sup_{\substack{a \leq H, H' \leq b \\ |H' - H| < h}} \sup_{t \in [0, K]} |B^H(t) - B^{H'}(t)| = 0.$$

Remarque :

- Comme H varie localement, on peut étudier des propriétés de régularité locale (cf. Section 5.3).
- Le mBm est utilisé en modélisation du trafic sur internet, pour décrire les fonds marins.
- On a les mêmes généralisations en dimensions supérieures que pour le mBf : le drap multifractionnaire et le champ multifractionnaire.

- Il existe encore des généralisations du mBm (avec des fonctions de Hurst H de plus en plus générale (non continues)).

5.6.2 Régularité des trajectoires

Comme le mBm n'est pas stationnaire, on ne s'intéresse pas aux propriétés globales des trajectoires mais aux propriétés locales.

Proposition 5.6.3 *Presque sûrement, l'exposant de Hölder en t_0 du mBm B^H est H_{t_0} .*

Démonstration : Cf. [P-LV]. □

La dimension de Hausdorff (1918) est un nombre réel positif (éventuellement $+\infty$) associé à tout espace métrique. Cette notion étend celle de dimension d'un espace vectoriel réel.

Définition 5.6.2 (Dimension de Hausdorff) *On définit pour un ensemble E et $s > 0$, $\delta > 0$:*

- $H_\delta^s(E) = \inf \{ \sum_{i=1}^{+\infty} \text{diam}(A_i)^s \text{ pour les partitions } (A_i)_{i \geq 1} \text{ de } E \text{ avec } \text{diam}(A_i) \leq \delta \}$.
- $H_s(E) = \lim_{\delta \rightarrow 0} H_\delta^s(E)$.

La dimension de Hausdorff de E est alors

$$\dim_H(E) = \inf(s : H^s(E) = 0) = \sup(s | H^s(E) = +\infty).$$

Exemples :

- Un sous espace vectoriel de dimension n a pour dimension de Hausdorff n . Idem pour un sous espace affine.
- Le flocon de Koch a pour dimension de Hausdorff $\ln 4 / \ln 2 \simeq 1,26$.
- Le triangle de Sierpinski a pour dimension de Hausdorff $\ln 3 / \ln 2 \simeq 1,585$.
- La zone délimitée par un MB en dimension 2 a pour dimension de Hausdorff $4/3 \simeq 1,33$ (Werner et al.)
- La courbe de Peano a pour dimension de Hausdorff 2.
- Les côtes britanniques ont pour dimension de Hausdorff 1,24.
- Les côtes norvégiennes ont pour dimension de Hausdorff 1,52.
- La dimension de Hausdorff de la surface d'un chou-fleur est $\simeq 2,33$.
- La dimension de Hausdorff d'une boule de papier froisé est $\simeq 2,4 \pm 0,2$.
- La dimension de Hausdorff de la surface pulmonaire est $\simeq 2,97$.

Proposition 5.6.4 *Presque sûrement, pour tout intervalle $[a, b] \subset \mathbb{R}_+$, le graphe du mBm $(Y_t)_{t \in [a, b]}$ vérifie*

$$\dim_H\{Y_t, t \in [a, b]\} = 2 - \min \{H_t, t \in [a, b]\}$$

où \dim_H est la dimension de Hausdorff. Pour un H -mBf, on a donc $\dim_H\{B_t^H, t \in [a, b]\} = 2 - H$.

5.6.3 Simulation, estimation du mBm

Estimation de la fonction de Hurst H

Comme pour le mBf, on peut estimer le paramètre de Hurst H . Il s'agit pour le mBm d'une fonction. On définit une variation quadratique généralisée au voisinage du point $0 \leq x_0 \leq 1$. On commence par définir

$$\mathcal{V}_{\varepsilon,n}(x_0) = \left\{ p \in \mathbb{Z}, \left| \frac{p}{n} - x_0 \right| \leq \varepsilon \right\}.$$

On pose alors comme variation quadratique généralisée

$$V_{\varepsilon,n}(x_0) = \sum_{p \in \mathcal{V}_{\varepsilon,n}(x_0)} \left(B_h\left(\frac{p+1}{n}\right) - 2B_h\left(\frac{p}{n}\right) + B_h\left(\frac{p-1}{n}\right) \right)^2.$$

On obtient l'estimateur suivant de $H(x_0)$ en x_0 :

$$H_{\varepsilon,n}(x_0) = \frac{1}{2} \left(1 + \frac{\ln(V_{\varepsilon,n/2}(x_0)/V_{\varepsilon,n}(x_0))}{2} \right).$$

Simulation. La méthode de simulation est fondée sur la Prop. 5.6.2 et est tirée de [P-LV]. On suppose qu'est donnée une fonction H et la taille N désirée de l'échantillon. Pour chaque valeur de $H(i/N)$ ($1 \leq i \leq N$), on génère un mBf $B^{H(i/N)}$ d'exposant $H(i/N)$ par n'importe quelle méthode classique. On simule alors le mBm par $B_H(i/N) = B^{H(i/N)}(i/N)$ pour $1 \leq i \leq N$.

Estimation. Soit $(X_H(t))_t$ un mBf standard et $\{X_{i,n} = X_H(i/N), 0 \leq i \leq N\}$ un échantillon de ce processus. Soit

$$S_n = \frac{1}{n-1} \sum_{i=1}^{n-1} |X_{i+1,n} - X_{i,n}|,$$

et

$$H_n = -\frac{\log(\sqrt{\pi/2}S_n)}{\log(n-1)}.$$

Alors ps

$$\lim_{n \rightarrow +\infty} H_n = H.$$

Explication heuristique. D'après la Prop. 5.6.2, et la continuité de H , pour chaque $t_0 \in [0, 1]$, il existe un voisinage $V(t_0)$ sur lequel on peut considérer que la fonction H est (presque) constante égale à H_{t_0} .

Soit n la taille de l'échantillon du mBm et $k \in [1, n]$ la taille du voisinage utilisé pour estimer le paramètre fonctionnel. On estimera alors $H(t)$ seulement pour $t \in [k/n, 1 - k/n]$. Sans perte de généralité, on supposera $m = n/k$ entier. Alors l'estimateur de $H(i/(n-1))$ est

$$\tilde{H}_{i/(n-1)} = -\frac{\log(\sqrt{\pi/2}S_{k,n}(i))}{\log(n-1)}$$

où

$$S_{k,n}(i) = \frac{m}{n-1} \sum_{j \in [i-k/2, i+k/2]} |X_{j+1,n} - X_{j,n}|.$$

5.7 Modélisation d'un signal

On peut modéliser un signal par un processus aléatoire $(X_t)_t$, en général un mouvement brownien, un mBf ou un mBm en fonction du type de signal. Le signal à la date t vaut donc X_t .

Régularité d'un signal

On mesure la régularité du signal en t par l'exposant de Hölder $\alpha_X(t)$ du signal en t . Ainsi, si $\alpha_X(t)$ est petit en t alors le signal est très irrégulier. Dans la modélisation d'une image, l'étude de la régularité permet d'étudier les variations de niveau de gris (donc de contraste) et d'analyser l'image en détectant de cette façon les contours.

Image numérique

Une image numérique se présente comme un rectangle divisé en petits carrés, appelés pixels. Chaque pixel est le résultat d'une mesure, généralement faite par une matrice de capteurs appelée CCD (pour Charge Couple Device). Un pixel correspond à un petit carré de la matrice CDD dans lequel le nombre de photons arrivant est compté. Chaque pixel contient trois nombres allant d'habitude de 0 à 255 et indiquant la quantité de rouge, de vert et de bleu contenues dans le pixel. Une combinaison adéquate de ces trois nombres permet de reproduire n'importe quelle couleur sur un ordinateur.

Dans toute image numérique, les trois valeurs de couleur observées présentent une incertitude due au bruit. Cette incertitude est due aux aléas de comptage des photons arrivant dans chaque capteur. Les valeurs mesurées sont perturbées car les capteurs reçoivent des photons parasites et subissent des fluctuations électrostatiques lors de leurs charges et décharges. Quand un capteur reçoit beaucoup de photons venant d'une scène bien éclairée, les parasites sont négligeables par rapport au flux de vrais photons. Mais même dans une photo d'exposition suffisante, les pixels sombres reçoivent très peu de photons et sont donc sensible au bruit.

Si le temps d'exposition d'une photo est trop court, c'est toute l'image qui est bruitée. En effet, une photo papier sous-exposée serait simplement sombre. Mais une caméra digitale peut la corriger en augmentant le contraste numérique. L'ennui, c'est que cette opération augmente dans la même proportion le bruit. L'utilisateur voit alors une image claire mais bruitée (donc floue). En fait d'autres opérations de restauration des images digitales amplifient le bruit et demandent donc à être couplées avec un débruitage. C'est typiquement le cas pour les opérations enlevant le flou dans les images.

La réduction du bruit dans les images est liée à plusieurs enjeux technologiques. En effet, la taille des capteurs CDD est, elle-même, dictée en partie par l'exigence de réduction de bruit. Le nombre de photons qu'un capteur reçoit dans une prise de vue est proportionnel à sa surface. Un capteur trop petit recevrait donc peu de photons venant de la scène photographiée et serait bruité. Si on disposait d'un procédé de débruitage plus efficace que les actuels, on pourrait construire des caméras encore plus petites, en proportion de la diminution de taille du CDD, ou des caméras avec encore plus de pixels. Mais peut-être l'amélioration la plus importante serait-elle de pouvoir prendre des photos de bonnes qualités avec un éclairage médiocre et sans flash. Ces enjeux technologiques sont présents dans pratiquement tout dispositif créant des images : en microscopie, en imagerie médicale ou satellitaire et bien sûr pour les caméras grand public.

Une image en noir en blanc est modélisée par des pixels qui ont un certain niveau de gris. On peut supposer pour simplifier que l'image est une fonction $X : [0, 1] \times [0, 1] \rightarrow \mathbb{R}$ qui indique le niveau de gris en chaque point $(s, t) \in [0, 1] \times [0, 1]$ par $X(s, t)$.

Comme l'image digitale est (par nature) bruitée, elle est modélisée par un champ aléatoire $(X_{s,t})_{s,t \in [0,1] \times [0,1]}$ indexé par $[0, 1] \times [0, 1]$ (ou par une partie de \mathbb{R}^2).

Bien sûr si l'image représente une scène bien précise, le champ qui la représente n'est pas du tout (ou pas complètement) aléatoire. On utilise la modélisation par un champ aléatoire

- lorsque l'image représente une scène a priori peu ordonnée mais dans laquelle on recherche un semblant de structure,
- lorsqu'on modélise le bruit aléatoire (et inconnu) B qui s'ajoute à la fonction déterministe (inconnue aussi) f représentant l'image. Ce qu'on observe est alors $X = f + B$.

La modélisation la plus courante est alors par des champs gaussiens isotropes (si aucune direction n'est privilégiée) ou anisotropes (si une direction semble privilégiée dans l'image) : le drap brownien ou le drap brownien multifractionnaire.

Détection de contours dans une image

Il n'est pas pertinent d'étudier directement la valeur des niveaux de gris $X(t)$ car les contours sont invariants par changement de contraste ou de luminosité.

Toutefois, il y a une correspondance entre les contours et une variation brusque des niveaux de gris. C'est donc en étudiant les variations brusques du signal $X(t)$ qu'on détecte les contours de l'image. Ces variations sont détectables en observant l'exposant de Hölder $\alpha_X(t)$ du signal : quand $\alpha_X(t)$ est petit, le signal est très irrégulier et il y a sûrement un contour.

Pour un mBf, $\alpha_{B^H}(t) = H$ ps pour tout t . L'exposant est constant et ne permet pas de détecter quoi que ce soit. L'image modélisée par un mBf est uniforme.

Noter qu'en dimension $d \geq 2$, pour un drap fractionnaire, la régularité est H_i dans la direction i alors que pour un champ, elle est H dans toutes les directions.

Pour un mBm, $\alpha_{B^H}(t) = H_t$ ps pour tout t . On cherche les t correspondants aux petites valeurs de t .

Dans le cas d'une image, le signal est un champ et on considère donc des mBf ou mBm en dimension 2 (drap ou champ).

Pour $\alpha_X \simeq 0$, le signal est très irrégulier : l'image est bruitée (poluée par du bruit dû à la transmission)

Pour $\alpha_X \simeq 1$, le signal est très régulier : la texture de l'image est uniforme.

Décomposition multifractale

Il s'agit d'étudier les lignes de niveaux $E_\alpha = \{t : \alpha_X(t) = \alpha\}$ de l'exposant de Hölder. Décomposer ainsi le support de X revient à regrouper les points t qui ont la même régularité (l'ensemble E_α).

En mesurant les ensembles E_α par $f(\alpha)$ (en général par la dimension de Hausdorff $\dim_H(E_\alpha)$), on se donne une décomposition multifractale de l'image.

Noter qu'il n'est pas nécessaire de supposer que l'image vérifie des propriétés fractales pour définir sa décomposition multifractale.

Application à la détection de contours

Cette décomposition permet aussi de faire de la détection de contour : il faut pour cela détecter les lignes de niveaux E_α qui sont de dimension de Hausdorff 1 (car un contour est une courbe, donc un espace affine de dimension 1). Dans ce cas, on mesure les ensembles E_α par leur dimension de Hausdorff.

Application au débruitage des images

Observons d'abord que pour une scène très bruitée, la plupart des points ont une faible régularité : $f(\alpha)$ sera grand pour α petit ; tandis que pour une image lisse, la plupart des points ont une grande régularité : $f(\alpha)$ est grand pour α grand.

Principe du débruitage

On modifie l'image bruitée de telle sorte que son spectre multifractal soit translaté vers les grandes valeurs de α . Ce faisant, on augmente la régularité de chaque point tout en gardant la forme du spectre.

Cela permet de rendre l'image plus lisible tout en conservant la force respective de chaque singularité (un point sur un contour restera après cette opération plus irrégulier qu'un point d'une zone uniforme, car on a tout translaté).

Cette méthode est utilisée par exemple pour restaurer des images SAR (Radar à Ouverture Synthétique)

Bibliographie

- [A-LV] A. Ayache, J. Lévy Vehel. *The generalized multifractional Brownian motion*. Statistical Inference for Stochastic processes, vol. 3, p. 7-18, 2000.
- [JCB] J-C Breton. *Cours de Probabilité*. Université de La Rochelle, L3.
<http://perso.univ-lr.fr/jcbreton/Fichiers/proba-L3.pdf>.
- [Bil] P. Billingsley. *Convergence of probability measures*. Wiley series in Probability, 1968.
- [BCM] A. Baudes, B. Coll, J-M. Morel. *La chasse au bruit dans les images*, Gazette de la SMF, juin 2006.
- [Dav] Y. Davydov. *Cours de DEA "Processus Stochastiques"*. Université Lille 1, DEA, 1999.
- [EGK] N. El Karoui, E. Gobet, E. Pardoux. *Introduction au calcul stochastique*. École Polytechnique.
- [LV] J. Lévy Vehel. *Introduction to multifractal analysis of images* in Y. Fisher (ed.) *Fractal image encoding and analysis*. Springer Verlag, 1996.
- [Lif] M. Lifshits. *Gaussian random functions*, Kluwer. 1995.
- [P-LV] R. Peltier, J. Lévy Vehel. *Multifractional Brownian motion : definition and preliminary results*. Rapport de recherche de l'INRIA no. 2645 1995.
- [PPLV] B. Pesquet-Popescu, J. Lévy-Vehel. *Stochastic fractal models for image processing*.
- [ST] G. Samorodnitsky, M. Taqqu. *Stable Non gaussian random processes*. Chapman & Hall.