

On Brownian Motion, Ito Calculus and the Feynman-Kac Formula

27 janvier 2016

Dans cette partie nous décrirons une méthode probabiliste pour évaluer le noyau k_β que nous avons décrit dans la section précédente. Avant de commencer le calcul de k_β , nous allons introduire quelques notions utiles.

1 Definitions

Definition 1.1. Processus de Markov

La séquence $(X_n)_{n \in \mathbb{N}}$ est un processus de Markov défini sur \mathbb{R}^d par rapport à la mesure de probabilité φ si pour toute fonction f de classe \mathcal{C}^0 tendant vers 0 à l'infini, la limite suivante existe (presque sûrement) :

$$\frac{1}{N} \sum_{n=1}^N f(X_n) \xrightarrow[N \rightarrow \infty]{p.s.} \mathbb{E}_\varphi(f) \quad (1)$$

où l'on a noté $\mathbb{E}_\varphi(f)$ l'espérance de f par rapport à la mesure φ .

Definition 1.2. Ergodicité

Un processus est ergodique s'il satisfait les conditions suivantes :

- φ est une mesure de probabilité invariante par le processus de Markov.
- **Condition d'accessibilité** :

$$\forall B \in \mathcal{B}(\mathbb{R}^d) \quad \text{t.q.} \quad \varphi(B) > 0, \quad \forall x \in \mathbb{R}^d, \quad \exists n \in \mathbb{N} \quad \text{t.q.} \quad \mathbb{P}(X_n \in B | X_0 = x) > 0 \quad (2)$$

Remark. 1. Nous avons défini la notion d'ergodicité sur \mathbb{R}^d , mais les mêmes définitions restent valables pour $\mathbb{L}\mathbb{T}^d$.

2. Au moment de simuler avec l'algorithme de Metropolis, il suffira de vérifier que la probabilité de transfert pour cet algorithme satisfait la condition d'accessibilité dans la définition d'ergodicité.

Definition 1.3. Chaînes de Markov

Une chaîne de Markov (discrète) est un processus de Markov défini par :

$$\begin{cases} X_{n+1}^{\Delta t} &= X_n^{\Delta t} + b(X_n^{\Delta t}) + \sqrt{\Delta t} \xi_n \\ X_0^{\Delta t} &= f(x) \end{cases} \quad (3)$$

avec $x \in \mathbb{R}^d$ et $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$ une fonction régulière. Ici $\xi_j \sim \mathcal{N}(0, 1)_{\mathbb{R}^d}$, i.e. ξ_j est un vecteur gaussien de dimension d .

À partir d'une chaîne de Markov discrète nous pouvons construire une chaîne de Markov en temps continu en reliant les instants de temps t_n et t_{n+1} de manière affine. Dans ce cas, nous avons avec les mêmes conditions que ci-dessus :

$$\forall t \in [t_n, t_{n+1}], \quad \tilde{X}_t^{\Delta t} = X_n^{\Delta t} + (t - t_n) X_{n+1}^{\Delta t} \quad (4)$$

Nous avons de même

$$\tilde{X}_t^{\Delta t} \xrightarrow{\Delta t \rightarrow 0} X_t \quad (5)$$

où X_t suit l'équation différentielle stochastique donnée par :

$$dX_t = b(X_t)dt + dW_t \quad (6)$$

COMMENT : ou dW représente quoi ??? Un bruit blanc ? Il faut l'expliciter !

2 Formule d'Itô

À la fin de la section précédente, nous avons abouti à une équation différentielle stochastique qui détermine la chaîne de Markov en temps continu. Pour faciliter la résolution de cette équation, nous mettrons en place des formules de calcul différentiel stochastique qui nous permettront d'aboutir à la formule de Feynman-Kac.

Considérons d'abord l'équation 6 sans le terme stochastique dW_t . Dans ce cas, l'équation devient déterministe et sa solution Y_t suit l'équation :

$$\frac{dY_t}{dt} = b(Y_t) \quad (7)$$

Pour rester dans un cadre simple, nous travaillerons dans \mathbb{R} , mais les résultats sont bien évidemment généralisables sur \mathbb{R}^d . Soit h une fonction mesurable allant de Ω un ouvert de \mathbb{R} dans \mathbb{R} . Alors nous avons :

$$\begin{aligned} d(h(Y_t)) &= \nabla h(Y_t) dY_t \\ &= \nabla h(Y_t) b(Y_t) dt \end{aligned}$$

Dans le cas de X_t , solution de l'équation différentielle stochastique, nous devons développer la dernière équation jusqu'à deuxième ordre afin de faire apparaître le terme d'Itô.

$$d(h(X_t)) = \nabla_x h(X_t) dX_t + \frac{1}{2} \Delta_x h(X_t) dt + \frac{\partial h(t, X_t)}{\partial t} dt \quad (8)$$

$$h(t + \Delta t, X(t + \Delta t)) = h(t + \Delta t, X_t + b(X_t)\Delta t + \sqrt{\Delta t}\xi_n) \quad (9)$$

$$= h(t, X_t) + \frac{dh}{dt}(t, X_t)\Delta t + \nabla_x h(t, X_t)b(X_t)\Delta t \quad (10)$$

$$+ \nabla_x(t, \Delta t)\xi_n\sqrt{\Delta t} + \frac{\Delta t}{2} \sum_{i,j} \xi_i \xi_j \frac{\partial^2 h}{\partial x_i \partial x_j} + o(\Delta t) \quad (11)$$

En faisant un développement de Taylor au premier ordre de $h(t + \Delta t, X(t + \Delta t))$ nous trouvons :

$$\begin{aligned} h(t + \Delta t, X(t + \Delta t)) &= h(t, X_t) + \frac{dh}{dt}(t, X_t)\Delta t + \nabla_x h(t, X_t)b(X_t)\Delta t \\ &\quad + \nabla_x(t, \Delta t)\xi_n\sqrt{\Delta t} + \frac{\Delta t}{2} \sum_{i,j} \xi_i \xi_j \frac{\partial^2 h}{\partial x_i \partial x_j} + o(\Delta t) \end{aligned} \quad (12)$$

Nous remarquons que le terme stochastique qui sera dominant est en $\sqrt{\Delta t}$ et le terme dominant déterministe est de l'ordre Δt . Nous ne prenons donc pas en compte les dérivées d'ordre 2 et aboutissons à l'équation suivante :

$$\begin{aligned} h(t + \Delta t, X(t + \Delta t)) &= h(t, X_t) + \frac{dh}{dt}(t, X_t)\Delta t \\ &\quad + \nabla_x h(t, X_t)b(X_t)\Delta t + \nabla_x(t, \Delta t)\xi_n\sqrt{\Delta t} \end{aligned} \quad (13)$$

3 Le Formule de Feynman-Kac

Munis des règles du calcul différentiel, nous sommes maintenant en mesure de dériver la formule de Feynman-Kac.

Pour ce faire, nous introduisons deux équations : les équations de Kolmogorov forward et backward

Definition 3.1. L'équation de Kolmogorov forward Elle est définie par :

$$\begin{aligned} \frac{\partial u}{\partial t} &= \frac{1}{2} \Delta u - \text{div}(bu) - Vu \\ u(0) &= f(x) \end{aligned} \quad (14)$$

Soit alors X_t un processus stochastique défini par

$$\begin{aligned} dX_t &= b(X_t) + dW_t \\ X_t(0) &= f(x)dx \end{aligned} \tag{15}$$

Donc, nous introduisons l'équation adjointe de Kolmogorov forward :

Definition 3.2. L'Equation de Kolmogorov Backward

Soit g une fonction allant de Ω ouvert de \mathbb{R} dans \mathbb{R} régulière bornée. Soit β qui noté comme un instant de temps finale.

$$\begin{aligned} \frac{\partial v}{\partial t} + \frac{1}{2}\Delta v + b\nabla v - Vv &= 0 \\ v(\beta, x) &= g(x) \end{aligned} \tag{16}$$

Nous considérons la dérivée totale $d(v(t, X_t)e^{-\int_0^t V(X_s)ds})$ où X_t est défini en (15). En utilisant l'équation (8), nous trouvons que :

$$\begin{aligned} d(v(t, X_t)e^{-\int_0^t V(X_s)ds}) &= \left(\frac{\partial v}{\partial t} + b\nabla v + \frac{1}{2}v - Vv \right) e^{-\int_0^t V(X_s)ds} \\ &\quad + \nabla v(t, X_t)dW_t e^{-\int_0^t V(X_s)ds} \end{aligned} \tag{17}$$

le premier terme est nul d'après (16). Donc :

$$d(v(t, X_t)e^{-\int_0^t V(X_s)ds}) = \nabla v(t, X_t)dW_t e^{-\int_0^t V(X_s)ds} \tag{18}$$

Enfin nous remarquons que l'espérance de $\nabla v(t, X_t)dW_t e^{-\int_0^t V(X_s)ds}$ est nulle, et en utilisant la commutation des différentes opérations avec l'espérance, nous avons ainsi que pour tout $\beta > 0$:

$$\mathbb{E}(v(\beta, X_\beta)e^{-\int_0^\beta V(X_s)ds}) = \text{cte} \tag{19}$$

Et enfin :

$$\mathbb{E}(v(\beta, X_\beta)e^{-\int_0^\beta V(X_s)ds}) = \mathbb{E}(v(0, X_0)) \tag{20}$$

$$\mathbb{E}(g(x_\beta))e^{-\int_0^\beta V(X_s)ds} = \int_{\Omega} v(0, x)f(x)dx \tag{21}$$

où nous avons utilisé la condition initiale de (16) et où l'intégration porte sur l'ouvert Ω considéré.

Considérons maintenant l'intégrale $\int_{\Omega} v(0, x) f(x) dx$. Par la condition initiale de (14), nous avons par conséquent :

$$\int_{\Omega} v(0, x) f(x) dx = \int_{\Omega} v(0, x) u(0, x) dx \quad (22)$$

D'où :

$$\begin{aligned} \int_{\Omega} v(0, x) u(0, x) dx &= \int_{\Omega} v(\beta, x) u(\beta, x) dx - \int_0^{\beta} \frac{d}{dt} \left(\int_{\Omega} v(t, x) u(t, x) dx \right) dt \\ &= \int_{\Omega} v(\beta, x) u(\beta, x) dx - \int_0^{\beta} \left(\int_{\Omega} \frac{\partial v}{\partial t} u dx + \frac{\partial u}{\partial v} dx \right) dt \end{aligned} \quad (23)$$

Nous remarquons alors que les problèmes (14) et (16) sont l'adjoint l'un de l'autre. En d'autres termes, si L est un opérateur différentiel et L^* son opérateur adjoint, nous avons :

$$\begin{aligned} \frac{\partial u}{\partial t} &= Lu \\ \frac{\partial v}{\partial t} &= -L^*v \end{aligned} \quad (24)$$

Les intégrales sur Ω sont ainsi les normes L^2 des opérateurs. En d'autres termes :

$$- \int_0^{\beta} \left(\int_{\Omega} \frac{\partial v}{\partial t} u dx + \frac{\partial u}{\partial v} dx \right) dt = - \int_0^{\beta} ((-L^*v, u)_{L^2} + (v, Lu)_{L^2}) dt = 0 \quad (25)$$

Ainsi :

$$\begin{aligned} \int_{\Omega} v(\beta, x) u(\beta, x) dx &= \int_{\Omega} v(0, x) f(x) dx \\ \int_{\Omega} g(x) u(\beta, x) dx &= \int_{\Omega} v(0, x) f(x) dx \end{aligned}$$

et on déduit la formule de Feynman-Kac :

$$\boxed{\int_{\Omega} g(x) u(\beta, x) dx = \mathbb{E} \left(g(x_{\beta}) e^{-\int_0^{\beta} V(x_s) ds} \right)} \quad (26)$$

4 Application à notre problème

Revenons sur notre problème de départ, qui est de trouver le noyau de l'opérateur $e^{-\beta H}$. Nous rappelons que le noyau est une fonction k_β telle que :

$$\left(e^{-\beta H} f\right)(x) = \int_{\Omega} k_\beta(x, y) f(y) dy \quad (27)$$

Nous remarquons que, pour $b = 0$, $e^{-\beta H} f$ est la solution de (14). D'où en appliquant (26) :

$$\int_{\Omega \times \Omega} k_\beta(x, y) f(y) g(x) dx dy = \mathbb{E} \left(g(x_\beta) e^{-\int_0^\beta V(x_s) ds} \right) \quad (28)$$

Cette dernière condition étant vraie pour tout f , nous prenons $f(y) = \delta(y)$, le delta de Dirac. Dans ce cas, nous avons pour le processus X_t :

$$X_\beta = y + W_\beta \quad (29)$$

D'où :

$$\begin{aligned} \int_{\Omega} k_\beta(x, y) g(x) dx &= \mathbb{E} \left(g(y + W_\beta) e^{-\int_0^\beta V(y + W_s) ds} \right) \\ &= \int_{\Omega} g(x) \mathbb{E}_{y + W_\beta = x} \left(e^{-\int_0^\beta V(y + W_s) ds} \right) dx \end{aligned} \quad (30)$$

Nous identifions alors le noyau comme étant :

$$\boxed{k_\beta(x, y) = \mathbb{E}_{y + W_\beta = x} \left(e^{-\int_0^\beta V(y + W_s) ds} \right)} \quad (31)$$

Dans notre cas, nous nous intéressons au calcul de la fonction de partition $Z = \text{Tr}(e^{-\beta H})$, qui d'après ce qui précède s'écrit :

$$\begin{aligned} Z &= \text{Tr}(e^{-\beta H}) = \int_{\Omega} k_\beta(x, x) dx \\ &= \int_{\Omega} \mathbb{E}_{W_\beta = 0} \left(e^{-\int_0^\beta V(x + W_s) ds} \right) \end{aligned} \quad (32)$$

Remark. Ici, nous devons interpréter les deux termes dans la dernière relation ci-dessus. Tout d'abord, prenons \mathbb{E}_{W_β} . Ce terme est le terme d'espérance sur les ponts browniens, c'est à dire une espérance sur tous chemins allant de point de départ x_0 à temps $t = 0$ jusqu'au un moment β où encore une fois $x_\beta = x_0$.

Le deuxième terme en exponentiel représente la création ou annihilation des particules à un point $x + W_s$ donné. Donc, nous formons une image de ce processus comme suivant :

1. La particule commence son trajet à $t=0$ à $x = x_0 = 0$. Nous allons regarder seulement les trajets qui finissent à un temps $t = \beta$ de telle sorte que $x_\beta = 0$.
2. Pendant ce trajet, la particule peut être détruite ou créée selon le poids $e^{-\int_0^\beta V(x+W_s)}$.
3. L'espérance sur tous ces trajets nous donne la fonction de partition $Z = \text{Tr}(e^{-\beta H})$.