

Estimation non biaisée d'une diffusion Stochastique méthode parametrix

Jean Zagdoun et Steve Desmeuzes

February 23, 2018

Ce document constitue le rapport écrit du projet du cours de méthodes de Monte Carlo du master M2M0. Ce rapport a été fait grâce aux notes de cours, et aux articles [1], [2] et [3]

1 Introduction et Hypothèses

Ce rapport a pour but de présenter une méthode de simulation d'une diffusion stochastique par une approche non biaisée. Ce la se traduit de la façon suivante $\mathbb{E}[F(X_T)] = \mathbb{E}[h(\bar{X}_T)]$ X_T étant notre diffusion à estimer, \bar{X}_T est le schéma l'approchant puis $F : \mathbb{R}^d \rightarrow \mathbb{R}^d$ une fonction $C_c^\infty(\mathbb{R}^d)$.

Il s'agit d'une approche non biaisée mais pas forcément exacte dans le sens où $h(\bar{X}_T)$ n'a aucune raison de suivre la même loi que $F(X_T)$. On rappelle que la diffusion s'écrit :

$$X_T = x_0 + \int_0^T b(X_s)ds + \int_0^T \sigma(X_s)dW_s$$

avec $x_0 \in \mathbb{R}^d$, $b \in C_b^1(\mathbb{R}^d)$ et σ définie tel que $a = \sigma\sigma^T \in C_b^2(\mathbb{R}^d)$ soit uniformément elliptique, pour finir W_s représente un mouvement brownien d-dimensionnel.

La plupart des schémas vus en cours sont assez proches des méthodes classiques d'approximation d'EDP, leur principal défaut est l'erreur faible ou le biais. En effet, en décomposant l'erreur faite avec notre approximation numérique :

$$\mathbb{E}[F(X_T)] - \frac{1}{M} \sum_{i=1}^M F(\bar{X}_T^{(k)}) = \mathbb{E}[F(X_T)] - \mathbb{E}[F(\bar{X}_T)] + \mathbb{E}[F(\bar{X}_T)] - \frac{1}{M} \sum_{i=1}^M F(\bar{X}_T^{(k)})$$

On se retrouve donc avec deux erreurs. La première étant l'erreur faible et la seconde est l'erreur statistique. Cette seconde erreur est simplement due au caractère asymptotique de la loi des grands nombres.

Sans rentrer dans les détails de cet aspect, bien souvent l'erreur faible est un $o(N^{-1})$. Dans le cadre du cours, nous avons vu d'autres méthodes permettant de réduire au maximum l'erreur faible en utilisant l'extrapolation Richardson Romberg ou Multi-level Monte Carlo par exemple.

Dans ce rapport, nous allons présenter la méthode parametrix qui a été introduite par Bally & Kohatsu-Higa en 2015.

Enfin, pour des raisons de concision et de légèreté dans la lecture certaines démonstrations seront admises.

2 L'algorithme : Partie théorique

L'idée de cette méthode part de la relation suivante :

$$\partial_s \mathbb{E}[F(X_s^x)] = \mathcal{L} \mathbb{E}[F(X_s^x)]$$

Où \mathcal{L} est le générateur associé à la diffusion $(X_t)_{t \geq 0}$.

Considérons le processus parametrix $(\bar{X}_{t \geq 0}^x)$ pouvant être vu comme le processus d'approximation par le schéma d'Euler à un pas de temps : $\bar{X}_t^x = x + b(x)t + \sigma(x)W_t$ auquel on associe son générateur $\bar{\mathcal{L}}$, l'égalité ci-dessus reste alors valide. L'intuition, que l'on a, est que le processus originel ne devrait pas être trop "loin" de son schéma, pour un t suffisamment petit. Par conséquent, il est donc possible d'écrire $\mathbb{E}[F(X_t^x)] \approx \mathbb{E}[F(\bar{X}_t^x)]$.

De plus, on introduit la notation du semi-groupe: $P_t f(x) = \mathbb{E}[F(X_t^x)]$

$$P_t F(x) = \bar{P}_t F(x) + \int_0^t \partial_s (\bar{P}_{t-s} P_s F(x)) ds \quad (1)$$

$$= \bar{P}_t F(x) + \int_0^t (\bar{P}_{t-s} (\mathcal{L} - \bar{\mathcal{L}}) P_s F(x)) ds \quad (2)$$

$$= \bar{P}_t F(x) + \int_0^t S_{t-s} P_s F(x) ds \quad (3)$$

Où

$$\left\{ \begin{array}{l} S_{t-s} = \int_{\mathbb{R}^d} F(y) \theta_{t-s}(x, y) g(a(x)(t-s), y - b(x)(t-s) - x) dy \\ \theta_t(x, y) = \frac{1}{2} \sum_{i,j} \mathcal{K}_t^{ij}(x, y) - \sum_{i=1}^d \rho_t^i(x, y) \\ \mathcal{K}_t^{ij}(x, y) = \partial_{i,j}^2 a_{i,j}(y) + \partial_j a_{i,j}(y) H^i(a(x)t, y - b(x)t - x) + \partial_i a_{i,j}(y) H^j(a(x)t, y - b(x)t - x) \\ + (a_{i,j}(y) - a_{i,j}(x)) H^i j(a(x)t, y - b(x)t - x) \\ \rho_t^i(x, y) = \partial_i b^i(y) + (\partial_i b^i(y) - \partial_i b^i(x)) H^i(a(x)t, y - b(x)t - x) \\ H^i(a, x) = -(a^{-1}x) \\ H^{ij}(a, x) = (a^{-1}x)^i (a^{-1}x)^j - (a^{-1})^{ij} \end{array} \right.$$

La fonction g présente dans la première équation correspond juste à la densité d'une gaussienne multivariée provenant du schéma d'Euler.

En admettant que l'intégrale dans (3) existe. On peut, par un simple calcul itératif sur $P_s F(x)$, écrire $P_t F(x)$ de la façon suivante :

$$P_t F(x) = \sum_{n \geq 0} \int_0^t \dots \int_0^{s_{n-1}} S_{t-s_1} \dots S_{s_{n+1}-s_n} \bar{P}_{t-s_n} F(x) ds_1 \dots ds_n \quad (4)$$

$$= \sum_{n \geq 0} \int_{\Delta_n^*(t)} \left(\prod_{j=0}^{n-1} (S_{s_{j+1}-s_j}) \right) \bar{P}_{t-s_n} F(x) ds_1 \dots ds_n \quad (5)$$

$$= \sum_{n \geq 0} \int_{\Delta_n^*(t)} \mathbb{E} \left[F(\bar{X}_t^x) \left(\prod_{i=0}^{n-1} \theta_{s_{i+1}-s_i}(\bar{X}_{s_i}, \bar{X}_{s_{i+1}}) \right) \right] ds_1 \dots ds_n \quad (6)$$

Le passage de la ligne (5) à la ligne (6) se fait en admettant ce résultat :

$$\left(\prod_{i=0}^{n-1} (S_{s_{i+1}-s_i}) \right) \bar{P}_{t-s_n} F(x) = \mathbb{E} \left[F(\bar{X}_t^x) \left(\prod_{i=0}^{n-1} \theta_{s_{i+1}-s_i}(\bar{X}_{s_i}, \bar{X}_{s_{i+1}}) \right) \right]$$

Où $\Delta_n^*(t) = \{(s_1 \dots s_n) \in [0, t]^n : 0 < s_1 < \dots < s_n\}$ représente un simplexe. En introduisant un processus de Poisson $N = N(t)$ de paramètre $\lambda > 0$ ainsi que ses temps de saut : $\tau_1 < \dots < \tau_N < \tau_{N+1} = t$. Conditionnellement à $\{N = n\}$, on sait que $(\tau_1 < \dots < \tau_N)$ est distribué de manière uniforme ie $\mathbb{P}(N = n, \tau_1 = s_1, \dots, \tau_n = s_n) = \lambda^n e^{-\lambda t} \mathbb{1}_{\{(s_1, \dots, s_n) \in \Delta_n^*\}}$. Maintenant, on peut donc davantage simplifier (6) :

$$\begin{aligned} P_t F(x) &= e^{\lambda t} \sum_{n \geq 0} \int_{\Delta_n^*(t)} \lambda^n e^{-\lambda t} \mathbb{E} \left[F(\bar{X}_t^x) \left(\prod_{i=0}^{n-1} \lambda^{-1} \theta_{s_{i+1}-s_i}(\bar{X}_{s_i}, \bar{X}_{s_{i+1}}) \right) \right] ds_1 \dots ds_n \\ &= e^{\lambda t} \sum_{n \geq 0} \mathbb{E} \left[\mathbb{1}_{\{N=n\}} F(\bar{X}_t^x) \left(\prod_{i=0}^{n-1} \lambda^{-1} \theta_{\tau_{i+1}-\tau_i}(\bar{X}_{\tau_i}, \bar{X}_{\tau_{i+1}}) \right) \right] \\ &= e^{\lambda t} \mathbb{E} \left[F(\bar{X}_t^x) \left(\prod_{i=0}^{n-1} \lambda^{-1} \theta_{\tau_{i+1}-\tau_i}(\bar{X}_{\tau_i}, \bar{X}_{\tau_{i+1}}) \right) \right] \end{aligned}$$

Finalement, on peut donc identifier l'existence d'une variable aléatoire ayant la même espérance que notre processus initial.

3 L'algorithme : Partie pratique

Cet algorithme utilise un certain nombre de fonctions annexes définies comme dans le chapitre 2. Il faut, entre autre, préciser son coefficient de diffusion sa dérivée première et seconde, son coefficient de dérive avec sa dérivée première.

On doit aussi préciser la fonction $\theta_t(x, y)$ elle même dépendant de beaucoup d'autres fonctions. Les paramètres sont: T le temps final et λ le paramètre d'intensité du processus de Poisson. L'idée principale de l'algorithme est de simuler un temps aléatoire exponentiel, et si ce temps est supérieur à l'horizon temporel on renvoie juste un schéma d'Euler simple à un pas de discrétisation. Sinon, on simule une grille temporelle aléatoire d'exponentielles iid et on calcule la fonction de poids $\theta_t(x, y)$ jusqu'à ce que la somme des temps aléatoire dépasse l'horizon temporel. Il ne sera jamais atteint exactement avec probabilité 1, car se sont des lois à densité. Il faut donc effectuer la dernière approximation à part

Dans l'algorithme ci-dessous il est supposé que toutes les fonctions annexes sont définies correctement. De plus la notation U est un appel à une simulation d'une variable aléatoire $\mathcal{U}(0, 1)$ et W est aussi un appel à une simulation d'une variable aléatoire $\mathcal{N}(0, 1)$ ou a une gaussienne multivariée.

Algorithm 1 Algorithmme parametrix

```
1: procedure PARAMETRIX( $T, \lambda$ )
2:    $X_{pred} = x_0$ 
3:    $\Theta = 1$ 
4:    $t = 0$ 
5:    $\delta_\tau = -(\frac{1}{\lambda})\log(U)$ 
6:    $t = \delta_\tau$ 
7:   if  $\delta_\tau > T$  then
8:      $X_{new} = x_0 + b(x_0)T + \sigma(x_0)\sqrt{T}W$ 
9:     return  $e^{(\lambda T)}F(X_{new})$ 
10:  else
11:    while  $t < T$  do
12:       $X_{new} = X_{pred} + b(X_{pred})\delta_\tau + \sigma(X_{pred})\sqrt{\delta_\tau}W$ 
13:       $\Theta = \Theta\theta_{\delta_\tau}(X_{pred}, X_{new})\lambda^{-1}$ 
14:       $X_{pred} = X_{new}$ 
15:       $\delta_\tau = -(\frac{1}{\lambda})\log(U)$ 
16:       $t = t + \delta_\tau$ 
17:     $\tilde{\delta} = T - t + \delta_\tau$ 
18:     $X_{last} = X_{pred} + b(X_{pred})\tilde{\delta} + \sigma(X_{pred})\sqrt{\tilde{\delta}}W$ 
19:     $\Theta = \Theta\theta_{\tilde{\delta}}(X_{pred}, X_{last})\lambda^{-1}$ 
20:  return  $e^{(\lambda T)}F(X_{last})\Theta$ 
```

3.1 Test numérique

Introduisons le processus unidimensionnel suivant :

$$dX_t = s(\sin(\omega x) + 2) dW_t + \frac{-x}{x^2 + \frac{C_1}{3C_3}} (s(\sin(\omega x) + 2))^2 dt \quad \text{avec } x_0 = 1$$

De plus, en définissant la fonction $F(x) = C_3(x^3) + C_1x + C_0$ de cette façon. $F(X_t)$ est alors une vraie martingale, ce résultat est admis. Donc $\mathbb{E}[F(X_s)] = f(x_0) = 2 \forall s \geq 0$.

En choisissant la grille de paramètre suivante :

$$s = 0.1 \quad C_0 = 0 \quad C_1 = 1 \quad C_3 = 1 \quad \omega = 0.1$$

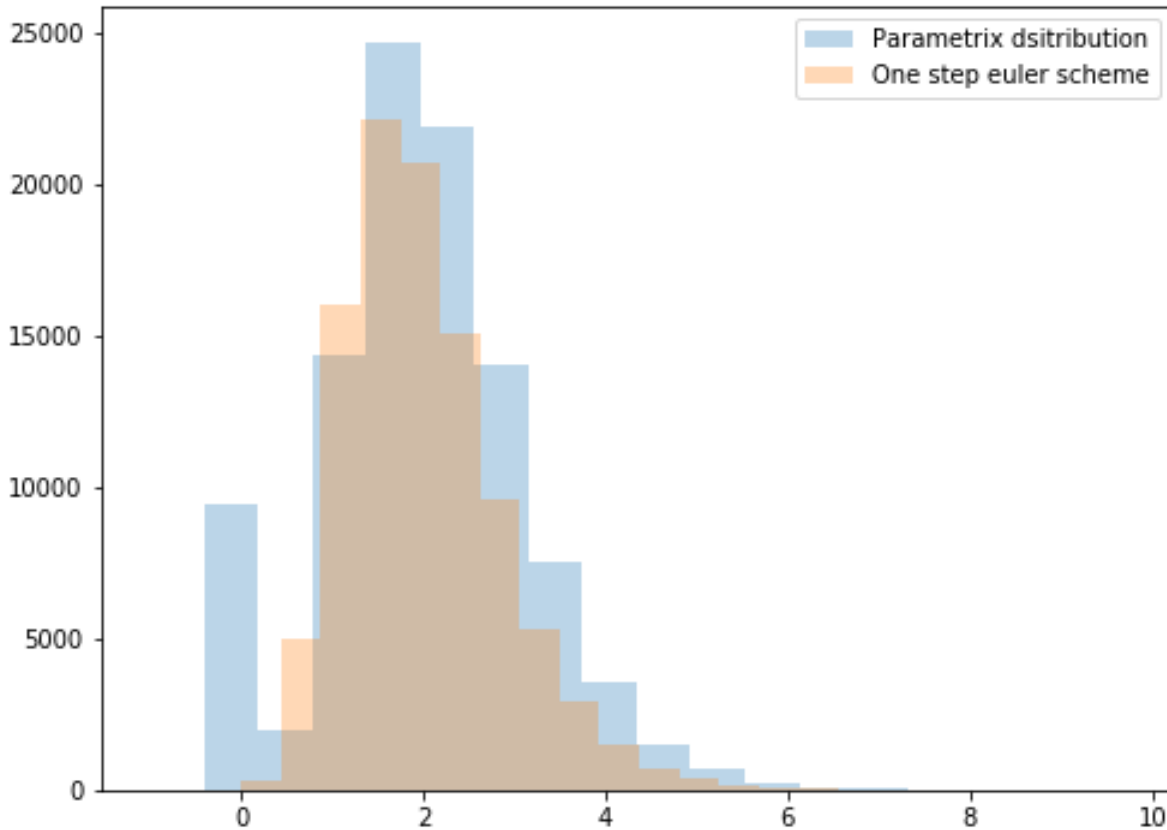
La méthode parametrix a un paramètre λ qui reste à chercher. Une idée, est de faire une recherche quadrillé puis de prendre le λ qui minimise la variance :

λ	0.1	0.2	0.3	0.6	0.7
valeurs numérique	2.0037	1.9963	2.0029	2.0171	2.0052
variance estimée	1.2164	1.7727	2.3739	4.6078	5.5095
Intervalle de confiance	[1.9969,2.0106]	[1.9881,2.0046]	[1.9934,2.0125]	[2.0038,2.0304]	[1.9907,2.0198]
taille IC	0.01370	0.01649	0.01910	0.02660	0.02910

Ainsi un choix cohérent, pour cette diffusion, est de prendre $\lambda^* = 0.1$.

3.2 Visualiser le schéma

Pour bien nous rendre compte de l'inexactitude de cette méthode, nous avons simulé $h(\bar{X}_T)$ et nous avons simulé un schéma d'euler à un pas \tilde{X}_T 100000 fois puis nous avons comparé les résultats (même si le schéma d'Euler n'est pas non plus exact à un temps t fixé il l'est asymptotiquement en temps et en espace grâce à une propriété, vue en cours, sur l'erreur forte):



On remarque alors, que les distributions finales ne coïncident pas. Il faut noter que cette approche n'est pas une approche rigoureuse de comparaison.

3.3 Limitations de cette méthode

Cet algorithme semble très bien marcher pour notre diffusion mais le choix des paramètres y est pour beaucoup. En effet, comme vu en cours ou dans [1], cette méthode est connue pour avoir une variance infinie. Pour s'en rendre compte, il suffit de prendre $s = 0.3$ et l'on trouve une valeur approchée de 2.0257 avec une variance de 148.912 et donc un intervalle de confiance relativement large [1.9501, 2.1013] et ce malgré le choix de λ^* .

Si l'on veut obtenir un intervalle de confiance de meilleure qualité et donc améliorer le modèle, il faut appliquer une technique de réduction de la variance comme par exemple importance sampling.

De plus, un autre point faible de cette méthode provient du fait que la fonction F doit

être $\mathcal{C}_c^\infty(\mathbb{R}^d)$, notamment pour les applications en mathématiques financières. Néanmoins ces hypothèses peuvent être affaiblies grâce à la remarque 5.2 de [1].

Les conditions de régularité des coefficients de la diffusion sont aussi très contraignantes. Cependant une amélioration existe pour relaxer les contraintes sur les coefficients à juste de la continuité hölderienne.

4 Une amélioration du modèle: réduction de variance

Une amélioration du modèle présenté par [1], aussi mentionné dans [2], consiste à remplacer les temps exponentiels de la grille d'approximation du schéma par des temps aléatoires qui suivent une loi Bêta amplifiée. Ensuite, il faut utiliser une méthode d'importance sampling venant de l'égalité suivante :

$$\mathbb{E}[F(X_T)] = \mathbb{E} \left[\frac{F(\bar{X}_T^x)}{p_N(\tau_1, \dots, \tau_N)} \left(\prod_{j=0}^{n-1} \theta_{\tau_{i+1}-\tau_i}(\bar{X}_{\tau_i}, \bar{X}_{\tau_{i+1}}) \right) \right]$$

Où $N = \inf\{n : \tau_n < T < \tau_{n+1}\}$, $\tau_n = \sum_{i=1}^n \xi_i$ où $(\xi_i)_i$ iid de loi Bêta

$\bar{\tau}\mathcal{B}(1-\gamma, 1)$, où $\bar{\tau} > T$ sera un paramètre à optimiser et enfin :

$$p_n(s_1, \dots, s_n) = \left(1 - \left(\frac{T - s_n}{\bar{\tau}} \right)^{1-\gamma} \right) \prod_{i=1}^{n-1} \frac{1-\gamma}{\bar{\tau}^{1-\gamma}(s_{i+1} - s_i)^\gamma}$$

Pour que la variance soit finie il faut $-1 < \gamma < 1$, (en tout cas dans le cadre qui nous intéresse) et si $\gamma = \frac{1}{2}$ alors tous les moments sont finis (cf proposition 7.3 de [1])

Tous ses résultats découlent des propositions et théorème de la section 7.2 de [1] et du corollaire 4.2 de [1] aucuns ne seront démontré ici par faute de place.

Il suffit donc de changer, seulement, quelques lignes de l'algorithme précédent:

Algorithm 2 Algorithme parametrix 2.0 : importance sampling

```
1: Les parties du code qui diffèrent de la procédure parametrix du chapitre 3 sont indiquées
   en rouge. Cet algorithme dépend de  $T$  l'horizon temporel,  $\gamma$  et  $\bar{\tau}$  les paramètres de la loi
   bêta
2: procédure BÊTA PARAMETRIX( $T, \bar{\tau}, \gamma$ )
3:    $X_{pred} = x_0$ 
4:    $\Theta = 1$ 
5:    $t = 0$ 
6:    $E = -\log(U)$ 
7:    $\delta_\tau = \bar{\tau}e^{-(1-\gamma)E}$ 
8:    $t = \delta_\tau$ 
9:    $S = [t]$ 
10:  if  $\delta_\tau > T$  then
11:     $X_{new} = x_0 + b(x_0)T + \sigma(x_0)\sqrt{T}W$ 
12:    return  $\frac{F(X_{new})}{p_0(0)}$ 
13:  else
14:    while  $t < T$  do
15:       $X_{new} = X_{pred} + b(X_{pred})\delta_\tau + \sigma(X_{pred})\sqrt{\delta_\tau}W$ 
16:       $\Theta = \Theta\theta_{\delta_\tau}(X_{pred}, X_{new})$ 
17:       $X_{pred} = X_{new}$ 
18:       $E = -\log(U)$ 
19:       $\delta_\tau = \bar{\tau}e^{-(1-\gamma)E}$ 
20:       $t = t + \delta_\tau$ 
21:      if  $t < T$  then
22:         $S = [S, t]$ 
23:     $\tilde{\delta} = T - t + \delta_\tau$ 
24:     $X_{last} = X_{pred} + b(X_{pred})\tilde{\delta} + \sigma(X_{pred})\sqrt{\tilde{\delta}}W$ 
25:     $\Theta = \Theta\theta_{\tilde{\delta}}(X_{pred}, X_{last})$ 
26:  return  $\frac{F(X_{last})\Theta}{p_N(\tau_1, \dots, \tau_N)}$ 
```

Ainsi avec les mêmes paramètres que précédemment et en fixant $\bar{\tau} = 5$.

On trouve les résultats suivants :

γ	0.6	0.7	0.9	0.999
valeur numérique	1.9880	1.9618	1.9984	2.0004
variance estimée :	0.7247	0.7642	0.7197	0.7205
Intervalle de confiance :	[1.9827,1.9932]	[1.9563,1.9672]	[1.9931,2.0037]	[1.9951,2.0056]
taille IC :	0.0105	0.0109	0.0105	0.0105

Notons que cette méthode dépend de deux paramètres : $\bar{\tau}$ et γ . Pour effectuer une recherche opérationnelle cohérente il faut faire une recherche quadrillée sur les deux paramètres. Pour des raisons de lisibilité, nous avons choisi de fixer $\bar{\tau}$.

De manière générale, on remarque que cette méthode donne des intervalles de confiances de meilleure qualité que ceux précédemment obtenus.

5 Conclusion

Nous avons donc vu dans ce rapport une méthode d'estimation d'une diffusion stochastique non-biaisée et une amélioration de cette méthode pour avoir une variance finie. Il faut remarquer que beaucoup d'autres améliorations existent:

- Importance sampling avec des temps aléatoires qui suivent une loi gamma
- Utiliser un système d'interaction de particules

Des implémentations numériques pour comparer ces deux méthodes peuvent être trouvées dans [2]. De plus nous n'avons pas étudié la complexité de l'algorithme améliorée avec des temps aléatoire bêta en fonction de ces paramètres. Toujours dans [2], il est question de cette problématique et ils définissent un score pour étudier l'utilité de l'algorithme en fonction de la variance et de la complexité.

References

- [1] Patrik ANDERSSON, Arturo KOHATSU-HIGA. *Unbiased simulation of stochastic differential equations using parametrix expansions*. Submitted to Bernoulli.
- [2] Mahamadou DOUMBIA, Nadia OUDJANE, Xavier WARIN. *Computing expectations for general SDE with pure Monte Carlo method*. 2016 .
- [3] Vlad BALLY, Arturo KOHATSU-HIGA. *A probabilistic interpretation of the parametrix method* 2015