Méthodes de Monte Carlo pour les EDP non linéaires

Nizar TOUZI

Ecole Polytechnique Paris

8 avril 2009

Outline

Loi des Grands Nombres

$X_1, \ldots, X_n, \ldots \sim$ iid $\mu(dx)$
f une fonction $\mu-$intégrable. Alors

$$\frac{1}{n}\sum_{i=1}^{n} f(X_i) \quad \longrightarrow \quad \mathbb{E}[f(X)] = \int f(x)\mu(dx) \quad \mu - \text{p.s.}$$

• En statistique : les X_i sont des observations, et on cherche à estimer le paramètre inconnu $\mathbb{E}[f(X)]$
• Approximation numérique : on simule les X_i (ou une approximation)

Méthode de Monte Carlo : simulation de $(X_n)_{n\geq 1}$, puis approximation d'une quantité d'intérêt à partir de l'échantillon simulé

Théorème de limite centrale

On note $\varepsilon_n(f) := \frac{1}{n}\sum_{i=1}^{n} f(X_i) - \mathbb{E}[f(X)]$

Si de plus f^2 est $\mu-$intégrable, alors

$$\sqrt{n}\varepsilon_n(f) \longrightarrow \mathcal{N}\left(0, \sigma^2(f)\right) \quad \text{en loi, où} \quad \sigma^2(f) = \mathbb{V}\mathrm{ar}[\mathrm{f(X)}]$$

En particulier :

$$\mathbb{P}\left\{\frac{\sqrt{n}\varepsilon_n(f)}{\sigma(f)} \in [-u_\alpha, u_\alpha]\right\} \quad \longrightarrow \quad 1-\alpha$$

u_α étant le quantile $\alpha/2$ de la $\mathcal{N}(0,1)$

$\Longrightarrow$ Intervalle de confiance

$\Longrightarrow$ Précision de la méthode de MC caractérisée par $\sigma(f)$

Exemples

- $U_i, i \geq 1 \sim$ iid $\mathcal{U}([0,1])$, $X_i := \mathbb{1}_{\{f(U_i) \geq \lambda\}}$

$$\hat{p}_n := \frac{1}{n}\sum_{i=1}^{n} X_i \quad \longrightarrow \quad p := \mathbb{P}\{f(U_i) \geq \lambda\} \quad \text{p.s.}$$
$$\sqrt{n}(\hat{p}_n - p) \quad \longrightarrow \quad \mathcal{N}(0, p(1-p)) \quad \text{en loi}$$

Si on veut $\frac{\sqrt{p(1-p)}}{p\sqrt{n}} \leq 10^{-2}$, il faut choisir $n \geq 10^4 \frac{1-p}{p}$...

- $\theta = \mathbb{E}\left[e^{\beta N}\right]$ où $N \sim \mathcal{N}(0,1)$. Alors

$$\theta = e^{\frac{\beta^2}{2}} \quad \text{et} \quad \sigma^2 := \mathbb{V}\mathrm{ar}\left[\mathrm{e}^{\beta \mathrm{N}}\right] = \mathrm{e}^{2\beta^2} - \mathrm{e}^{\beta^2}$$

Si on veut $\frac{\sigma}{\theta\sqrt{n}} \leq 10^{-2}$, il faut choisir $n \geq 10^4\left(e^{\beta^2} - 1\right)$...

Pour $\beta = 5$... $n \geq 7 \times 10^{12}$! !

Loi du logarithme itéré

$X_1, \ldots, X_n, \ldots \sim$ iid $\mu(dx)$, f une fonction $\mu-$intégrable
On rappelle

$$\varepsilon_n(f) = \frac{1}{n}\sum_{i=1}^{n} f(X_i) - \mathbb{E}[f(X)] \quad \text{et} \quad \sigma^2(f) = \mathbb{V}\mathrm{ar}[\mathrm{f}(\mathrm{X})]$$

$$\limsup_{n\to\infty} \sqrt{\frac{n}{2\ln\ln n}}\,\frac{\varepsilon_n(f)}{\sigma(f)} \quad = \quad 1$$

L'erreur est en moyenne de l'ordre de $\sqrt{n}$ (TCL), et est au pire de l'ordre de $\sqrt{\frac{2\ln\ln n}{n}}$

Lien avec l'approximation numérique de l'intégrale

On veut calculer $\int_{[0,1]^d} f(x)dx_1 \ldots dx_d$

En dimension 1 (pour simplifier), on approxime par des formules du type :

$$\sum_{i=1}^{n} \pi_i f(x_i), \quad \pi_i \geq 0, \; \sum_i \pi_i = 1 \quad \text{et } x_i \in [0,1]$$

$\longrightarrow x_i = \frac{1}{n}$ et $\pi_0 = \pi_n = \frac{1}{2n}$, $\pi_i = \frac{1}{n}$ pour $1 \leq i \leq n-1 \Longrightarrow$ convergenc en $\frac{1}{n}$ et même mieux si f est régulière...

$\longrightarrow \pi_i = \frac{1}{n}$ pour tout $i = 0, \ldots, n$ et x_i tirés au hasard dans $\mathcal{U}([0,1]) \Longrightarrow$ convergenc en $\frac{1}{\sqrt{n}}$, indépendemment de la dimension et de la régularité de f

Réduction de variance

Idée générale : trouver une autre représentation

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

telle que $\mathrm{\mathbb{V}ar}[Y] \leq \mathrm{\mathbb{V}ar}[X]$...

Exemples :

- Echantillonnage préférentiel
- Variables de contrôle
- Variables antithétiques

Outline

Mouvement brownien

$W : \mathbb{R}_+ \times \Omega \longrightarrow \mathbb{R}$ est un mouvement brownien si

- $W_0 = 0$ et $W(.,\omega)$ continu p.s.
- W est à accroissements indépendants et $W_t - W_s \sim \mathcal{N}(0, t-s)$

W est à variation infinie et à variation quadratique donnée par

$$\sum_{t_i \le t} |W_{t_i} - W_{t_{i-1}}|^2 \quad \longrightarrow \quad t \text{ en proba quand } \sup_i |t_i - t_{i-1}| \to 0$$

Conséquence directe : pour une fonction f régulière :

$$f(t, W_t) = f(0, W_0) + \int_0^t \left(\frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} \right)(s, W_s) ds + \int_0^t \frac{\partial f}{\partial x}(s, W_s) dW_s$$

Equation différentielle stochastique

W mouvement brownien dans $\mathbb{R}^d$
Soient $b : \mathbb{R}_+ \times \mathbb{R}^d \longrightarrow \mathbb{R}^d$ et $\sigma \mathbb{R}_+ \times \mathbb{R}^d \longrightarrow \mathbb{S}^d$ des fonctions Lipschitz en x uniformément en t et à croissance linéaire. Alors l'équation différentielle stochastique

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t$$

admet une unique solution Markovienne telle que

$$\mathbb{E}\left[\sup_{[0,T]} |X_t|^2\right] \leq C\left(\mathbb{E}[|X_0|^2] + e^{CT}\right)$$

Formule d'Itô :

$$f(t, X_t) = f(0, X_0) + \int_0^t \left(\frac{\partial f}{\partial t} + b \cdot Df + \frac{1}{2}\mathrm{Tr}\left[\sigma\sigma^{\mathrm{T}} D^2 f\right]\right)(s, X_s)ds + \int_0^t Df(s, X_s) \cdot \sigma(s, X_s)dW_s$$

Formule de Feynman-Kac

On note

$$\mathcal{L}v(t,x) \quad := \quad b(t,x)\cdot Dv(t,x) + \frac{1}{2}\mathrm{Tr}\left[(\sigma\sigma^T)(t,x)D^2v(t,x)\right]$$

et on considère le problème de Cauchy :

$$-\frac{\partial v}{\partial t} - \mathcal{L}v + kv - f = 0$$
$$v(T,x) = g(x)$$

S'il existe une solution régulière V à croissance polynomiale, alors

$$V(t,x) \quad = \quad \mathbb{E}_{t,x}\left[\int_t^T e^{-\int_t^s k(r,X_r)dr} f(s,X_s)ds + g(X_T)e^{-\int_t^T k(r,X_r)dr}\right]$$

où, étant donné un mouvement brownien W, X est la solution de :

$$X_s \quad = \quad X_t + \int_t^s b(r,X_r)dr + \int_t^s \sigma(r,X_r)dW_r$$

Méthode de Monte Carlo pour les EDP linéaires

La représentation donnée par la formule de Feynman-Kac ouvre la porte aux méthodes numériques probabilistes... pour simplifier, on prend $f = k = 0$

• Discrétisation d'Euler de l'EDS avec $h := \frac{T-t}{n}$, $t_i := ih$, et $\Delta^h W_{t_i} := W_{t_i} - W_{t_{i-1}}$, $i = 0, \ldots, n$

$$X_0^h = X_0 \quad \text{et} \quad X_{t_i}^h = X_{t_{i-1}}^h + b(t_{i-1}, X_{t_{i-1}}^h)h + \sigma(t_{i-1}, X_{t_{i-1}}^h)\Delta^h W_{t_i}$$

On peut montrer que

$$\limsup_{h\to 0} h^{-1}\mathbb{E} \sup_{[0,T]} |X^h - X|^2 < \infty \quad \text{et} \quad \limsup_{h\to 0} h^{-2}\mathbb{E}|X_1^h - X_1|^2 < \infty$$

• On simule des copies indépendantes de la discrétisation d'Euler $X^{h,j}$, $j = 1, \ldots, N$, et on calcule

$$\frac{1}{N}\sum_{j=1}^{N} g\left(X_{t_n}^h\right) \ldots \quad \text{LGN}, \quad \text{TCL}\ldots$$

Problème de couverture en finance

• Marché financier contenant un actif sans risque $S^0 \equiv 1$, et un actif risqué de processus de prix

$$dS_t = \mathrm{diag}[S_t]\sigma_t dW_t$$

μ, σ adaptés, σ inversible $+ \ldots$

• Portefeuille Z_t^i : montant investi dans l'actif i à la date t.
Condition d'autofinancement $\Longrightarrow$ dynamique de la valeur du portfeuille :

$$dY_t = Z_t \cdot \mathrm{diag}[S_t]^{-1} dS_t$$

Problème de surcouverture de la v.a. $G \geq 0$ $\mathcal{F}_T-$mesurable :

$$V_0 := \inf\{Y_0 : \; Y_T \geq G \text{ p.s. pour un certain } Z \in \mathcal{A}\}$$

Solution : le modèle de Black-Scholes

La valeur de sur-couverture est donnée par

$$V_0 = \mathbb{E}[G]$$

De plus, il existe un portefeuille optimal Z^* tel que

$$Y_T^{V_0,Z^*} = G \quad \text{replication parfaite}$$

• Si de plus $\mu_t = \mu(t,S_t)$, $\sigma_t = \sigma(t,S_t)$, $G = g(S_T)$, le prix de Black-Scholes $v(t,S_t) = V_t$ est caracterisé par l'EDP linéaire :

$$-Lv := -\frac{\partial v}{\partial t} - rs \cdot Dv - \frac{1}{2}\mathrm{Tr}\left[\sigma\sigma^{\mathrm{T}} D^2 v\right] + rv = 0 \quad \text{and} \quad v(T,.) = g$$

Le problème d'option américaine : une première EDP non linéaire

Etant donné un processus de gain $\{G_t, t \in [0,T]\}$, on cherche

$$\bar{V}_0 \quad := \quad \inf\left\{Y_0 : \ Y_t \geq G_t, \ t \in [0,T] \text{ p.s. pour un ceratin } Z \in \mathcal{A}\right\}$$

Par un argument similaire, ce problème de sur-couverture se réduit au problème d'arrêt optimal :

$$\bar{V}_0 \quad := \quad \sup\left\{\mathbb{E}[G_\tau] : \tau \text{ temps d'arrêt} \leq T\right\}$$

• Si de plus $\mu_t = \mu(t, S_t)$, $\sigma_t = \sigma(t, S_t)$, $G_t = g(S_t)$, le prix de de l'option américaine $v(t, S_t) = V_t$ est caracterisé par l'EDP frontière libre :

$$\min\left\{-L\bar{v} \ ; \ \bar{v} - g\right\} = 0 \quad \text{and} \quad \bar{v}(T,.) = g$$

Méthodes numériques probabilistes pour les options américaines

Le prix d'une option américaine de payoff $\{G_t, t \geq 0\}$:

$$V_0 = \sup\left\{\mathbb{E}\left[G_\tau\right] : \tau \text{ stopping time } \leq T\right\}$$

peut être approximé par l'enveloppe de Snell :

$$V_T^n := G_T \quad \text{et} \quad V_{t_k}^n := \max\left\{G_{t_k}, \mathbb{E}\left[V_{t_{k+1}}^n | \mathcal{F}_{t_k}\right]\right\}$$

où $t_k := kT/n$, $k = 0, \ldots, n$

Dans le cas Markov, l'espérance conditionnelle $\mathbb{E}\left[\cdot | \mathcal{F}_{t_k}\right]$ est une regression, et peut être approximée une approximation implémentable $\hat{\mathbb{E}}_k^N\left[\cdot\right]$, ce qui conduit au schema implementable :

$$\hat{V}_T^{n,N} := G_T \quad \text{et} \quad \hat{V}_{t_k}^{n,N} := \max\left\{G_{t_k}, \hat{\mathbb{E}}_k^N\left[\hat{V}_{t_{k+1}}^{n,N}\right]\right\}$$

Méthodes d'estimation de regression

Méthodes classiques en statistique :

- méthode des noyaux (suggérée dans le contexte des options US par Carrière)
- Projection sur des bases de $\mathbb{L}^2(\mathbb{P})$ <Longstaff-Schwarz, Gobet-Lemor-Warin AAP05>

Méthodes numériques probabilistes

- Markov chain approximation <Kushner>
- Stochastic mesh <Broadie-Glasserman>
- quantification... <Bally-Pagès SPA03>

Integration par parties <Lions-Reigner 00, Bouchard-T. SPA04>

Outline

The Pontryagin approach for stochastic control

- On considère le problème de contrôle stochastique

$$v(t,x) \quad := \quad \sup_{\nu \in \mathcal{U}} \mathbb{E}_{t,x}\left[\int_t^T f\left(r, X_r{}^{\nu}, \nu_r\right) dr + g\left(X_T{}^{\nu}\right)\right]$$

où X^ν est un processus détat contrôlé

$$dX_t{}^{\nu} \quad = \quad b\left(r, X_t{}^{\nu}, \nu_t\right) dt + \sigma\left(r, X_t{}^{\nu}, \nu_t\right) dW_t$$

- On définit le Hamiltonien :

$$H(t,x,u,y,z) \quad := \quad f(t,x,u) + b(t,x,u) \cdot y + \mathrm{Tr}\left[\sigma(tr,x,u)^T z\right]$$

- Pour tout contrôle ν, on introduit l'EDR rétrograde :

$$dY_t \;=\; -\nabla_x H\left(t, X_t, \nu_t, Y_t, Z_t\right) dt + Z_t \cdot dW_t \quad \text{with} \quad Y_T = \nabla g(X_T)$$

The Pontryagin Maximum Principle

- Soit $\hat{\nu} \in \mathcal{U}$ et $\hat{X} := X^{\hat{\nu}}$ l'état contrôlé correspondant, i.e.

$$d\hat{X}_t = \nabla_y H\left(t, \hat{X}_t, \hat{\nu}_t, \hat{Y}_t, \hat{Z}_t\right) dt + \nabla_z H\left(t, \hat{X}_t, \hat{\nu}_t, \hat{Y}_t, \hat{Z}_t\right) \cdot dW_t$$

$\left(\hat{Y}, \hat{Z}\right)$: solution adaptée de l'EDS rétrograde (état adjoint) :

$$d\hat{Y}_t = -\nabla_x H\left(t, \hat{X}_t, \hat{\nu}_t, \hat{Y}_t, \hat{Z}_t\right) dt + \hat{Z}_t \cdot dW_t \quad \text{avec} \quad \hat{Y}_T = \nabla g(\hat{X}_T)$$

- On suppose de plus pour tout $t \in [0, T]$:

$$H\left(t, \hat{X}_t, \hat{\nu}_t, \hat{Y}_t, \hat{Z}_t\right) = \max_{u \in U} H\left(t, \hat{X}_t, u, \hat{Y}_t, \hat{Z}_t\right)$$

- **Sous des conditions de convexité, $\hat{\nu}$ est un contrôle optimal**
- Le triplet $(\hat{X}, \hat{Y}, \hat{Z})$ est solution du systèmes d'EDS Forward-Backward

Preuve du principe du maximum de Pontryagin

- $f \equiv 0$ pour simplifier, $h(t,x,y,z) := \sup_{u \in U} H(t,x,u,y,z)$.
- **Hypothèse :** g et h concave en x Alors, pour $\nu \in \mathcal{U}$:

$$
\begin{aligned}
\mathbb{E}\left[g\left(\hat{X}_T\right)\right] - \mathbb{E}\left[g\left(X_T^\nu\right)\right] &\geq \mathbb{E}\left[\left(\hat{X}_T - X_T^\nu\right)\cdot\nabla g\left(\hat{X}_T\right)\right] \\
&= \mathbb{E}\left[\left(\hat{X}_T - X_T^\nu\right)\cdot\hat{Y}_T - \left(\hat{X}_0 - X_0^\nu\right)\cdot\hat{Y}_0\right] \\
&= \mathbb{E}\left[\int_0^T d\left\{\left(\hat{X}_t - X_t^\nu\right)\cdot\hat{Y}_t\right\}\right] \\
&= \mathbb{E}\int_0^T\Big[H\left(t,\hat{X}_t,\hat{\nu}_t,\hat{Y}_t,\hat{Z}_t\right) - H\left(t,X_t^\nu,\nu_t,\hat{Y}_t,\hat{Z}_t\right) \\
&\qquad - \left(\hat{X}_t - X_t^\nu\right)\cdot\nabla_x H\left(t,\hat{X}_t,\hat{\nu}_t,\hat{Y}_t,\hat{Z}_t\right)\Big]\,dt \\
&\geq \mathbb{E}\int_0^T\Big[h\left(t,\hat{X}_t,\hat{Y}_t,\hat{Z}_t\right) - h\left(t,X_t^\nu,\hat{Y}_t,\hat{Z}_t\right) \\
&\qquad - \left(\hat{X}_t - X_t^\nu\right)\cdot\nabla_x h\left(t,\hat{X}_t,\hat{Y}_t,\hat{Z}_t\right)\Big]\,dt \\
&\geq 0
\end{aligned}
$$

Lien avec l'approche de Bellman

L'approche de Bellman est basée sur le Principe de la Programmation Dynamique :

$$v(t,x) \;:=\; \sup_{\nu\in\mathcal{U}} \mathbb{E}_{t,x}\left[\int_t^{t+h} f\left(r, X_r^{\nu}, \nu_r\right) dr + v\left(t+h, X^{\nu}{}_{t+h}\right)\right]$$

dont la version infinitésimale est :

$$\frac{\partial v}{\partial t} + \sup_{u\in U}\left\{ f(t,x,u) + b(t,x,u)Dv + \frac{1}{2}\mathrm{Tr}[\sigma\sigma^T(t,x,u)D^2 v]\right\} \;=\; 0$$

avec la condition finale $v(T,x) = g(x)$

Sous quelques conditions de régularité, on a :

$$\hat{Y}_t = Dv(t,\hat{X}_t) \quad \text{et} \quad \hat{Z}_t = \sigma(t,\hat{X}_t)^{\mathrm{T}} D^2 v(t,\hat{X}_t)$$

Definition (Pardoux and Peng)

Given $X_t = X_0 + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s,\ t \geq 0$, find an $\mathbb{F}^W-$**adapted** (Y, Z) satisfying :

$$
\begin{aligned}
Y_t &= g(X_T) + \int_t^T f(r, X_r, Y_r, Z_r)dr - \int_t^T Z_r \cdot dW_r \\
\text{i.e.} \qquad dY_t &= -f(t, X_t, Y_t, Z_t)dt + Z_t \cdot dW_t \quad \text{and} \quad Y_T = g(X_T)
\end{aligned}
$$

where the generator $f \ : \ [0, T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \longrightarrow \mathbb{R}$ is Lipschitz in (x, y, z) uniformly in t, and $g(X_T) \in \mathbb{L}^2(\mathbb{P})$. Then there is a unique solution satisfying

$$
\mathbb{E} \sup_{t \leq T} |Y_t|^2 + \mathbb{E} \int_0^T |Z_t|^2 dt \ < \ \infty
$$

Moreover, there exists a measurable function V :

$$
Y_t = V(t, X_t) , \qquad 0 \leq t \leq T
$$

Connection with PDE's

- By definition,

$$\begin{aligned} Y_{t+h} - Y_t &= V(t+h, X_{t+h}) - V(t, X_t) \\ &= -\int_t^{t+h} f(X_r, Y_r, Z_r)dr + \int_t^{t+h} Z_r \cdot \sigma(X_r)dW_r \end{aligned}$$

- If $V(t,x)$ is smooth, it follows from Itô's formula that :

$$\int_t^{t+h} \mathcal{L}V(r, X_r)dr + \int_t^{t+h} DV(r, X_r) \cdot \sigma(X_r)dW_r$$
$$= -\int_t^{t+h} f(X_r, Y_r, Z_r)dr + \int_t^{t+h} Z_r \cdot dW_r$$

where $\mathcal{L}V := \partial_t V + b \cdot DV + \frac{1}{2}\mathrm{Tr}[\sigma\sigma^T D^2 V]$. Then

$$-\mathcal{L}V(t,x) - f\left(t, x, V(t,x), \sigma^{\mathrm{T}} DV(t,x)\right) = 0$$

Feynman-Kac representation for semilinear PDE's

Assume that the semilinear PDE

$$-\mathcal{L}V(t,x) - f\left(t,x,V(t,x),\sigma^{\mathrm{T}}DV(t,x)\right) = 0, \qquad V(T,.) = g$$

has a classical solution with polynomial growth. Then
$v(t,x) = Y_t^{t,x}$, where (Y,Z) is the solution of the BSDE :

$$X_s^{t,x} = x + \int_t^s b(r,X_r^{t,x})dr + \int_t^s \sigma(r,X_r^{t,x})dW_r$$
$$Y_s^{t,x} = g(X_T^{t,x}) + \int_s^T f\left(r,X_r^{t,x},Y_r^{t,x},Z_r^{t,x}\right)dr - \int_s^T Z_r^{t,x}dW_r$$

• For $f \equiv 0 \Longrightarrow v(t,x) = Y_t^{t,x} = \mathbb{E}\left[g(X_T^{t,x})\right]$
• Extension to fully nonlinear PDEs : Cheridito, Soner, T., Victoir (CPAM 06)...
• This representation suggests numerical probabilistic schemes...

Discrete-time approximation of BSDEs

<Bally-Pagès SPA03, Zhang AAP04, Bouchard-T. SPA04>

Numerical solution of a semi-linear PDE by simulating the associated backward SDE by means of Monte Carlo methods

Start from Euler discretization : $Y^n_{t_n} = g\left(X^n_{t_n}\right)$ is given, and

$$Y^n_{t_{i+1}} - Y^n_{t_i} = -f\left(X^n_{t_i}, Y^n_{t_i}, Z^n_{t_i}\right)\Delta t_i + Z^n_{t_i} \cdot \sigma\left(X^n_{t_i}\right)\Delta W_{t_{i+1}}$$

Discrete-time approximation of BSDEs

<Bally-Pagès SPA03, Zhang AAP04, Bouchard-T. SPA04>

Numerical solution of a semi-linear PDE by simulating the associated backward SDE by means of Monte Carlo methods

Start from Euler discretization : $Y_{t_n}^n = g\left(X_{t_n}^n\right)$ is given, and

$$\mathbb{E}_i^n[\qquad \longrightarrow Y_{t_{i+1}}^n - Y_{t_i}^n = -f\left(X_{t_i}^n, Y_{t_i}^n, Z_{t_i}^n\right)\Delta t_i + Z_{t_i}^n \cdot \sigma\left(X_{t_i}^n\right)\Delta W_{t_{i+1}}$$

$\Longrightarrow$ Discrete-time approximation : $Y_{t_n}^n = g\left(X_{t_n}^n\right)$ and

$$Y_{t_i}^n = \mathbb{E}_i^n\left[Y_{t_{i+1}}^n\right] + f\left(X_{t_i}^n, Y_{t_i}^n, Z_{t_i}^n\right)\Delta t_i \qquad , \; 0 \le i \le n-1 \, ,$$

Discrete-time approximation of BSDEs

<Bally-Pagès SPA03, Zhang AAP04, Bouchard-T. SPA04>

Numerical solution of a semi-linear PDE by simulating the associated backward SDE by means of Monte Carlo methods
Start from Euler discretization : $Y^n_{t_n} = g\left(X^n_{t_n}\right)$ is given, and

$$\mathbb{E}^n_i[\Delta W_{t_{i+1}} \quad \rightarrow \quad Y^n_{t_{i+1}} - Y^n_{t_i} \;=\; -f\left(X^n_{t_i}, Y^n_{t_i}, Z^n_{t_i}\right)\Delta t_i + Z^n_{t_i}\cdot\sigma\left(X^n_{t_i}\right)\Delta W_{t_{i+1}}$$

$\Longrightarrow$ Discrete-time approximation : $Y^n_{t_n} \;=\; g\left(X^n_{t_n}\right)$ and

$$\begin{aligned} Y^n_{t_i} &= \mathbb{E}^n_i\left[Y^n_{t_{i+1}}\right] + f\left(X^n_{t_i}, Y^n_{t_i}, Z^n_{t_i}\right)\Delta t_i \qquad , \; 0 \le i \le n-1 \\ Z^n_{t_i} &= \mathbb{E}^n_i\left[Y^n_{t_{i+1}} \frac{\Delta W_{t_{i+1}}}{\Delta t_i}\right] \end{aligned}$$

Discrete-time approximation of BSDEs

<Bally-Pagès SPA03, Zhang AAP04, Bouchard-T. SPA04>

Numerical solution of a semi-linear PDE by simulating the associated backward SDE by means of Monte Carlo methods

Start from Euler discretization : $Y^n_{t_n} = g\left(X^n_{t_n}\right)$ is given, and

$$\mathbb{E}^n_i[\Delta W_{t_{i+1}} \quad \rightarrow \quad Y^n_{t_{i+1}} - Y^n_{t_i} \;=\; -f\left(X^n_{t_i}, Y^n_{t_i}, Z^n_{t_i}\right)\Delta t_i + Z^n_{t_i}\cdot\sigma\left(X^n_{t_i}\right)\Delta W_{t_{i+1}}$$

$\Longrightarrow$ Discrete-time approximation : $Y^n_{t_n} \;=\; g\left(X^n_{t_n}\right)$ and

$$\begin{aligned} Y^n_{t_i} &= \mathbb{E}^n_i\left[Y^n_{t_{i+1}}\right] + f\left(X^n_{t_i}, Y^n_{t_i}, Z^n_{t_i}\right)\Delta t_i \qquad , \; 0 \le i \le n-1 \\ Z^n_{t_i} &= \mathbb{E}^n_i\left[Y^n_{t_{i+1}} \frac{\Delta W_{t_{i+1}}}{\Delta t_i}\right] \end{aligned}$$

$\Longrightarrow$ Similar to numerical computation of **American options**

Discrete-time approximation of BSDEs, continued

$$0 = t_0 < t_1 < \ldots < t_n = T, \; t_i := i\frac{T}{n}$$

Theorem (Zhang 04, Bouchard-T. 04) Assume f and g are Lipschitz. Then :

$$\limsup_{n\to\infty} n^{1/2} \left\{ \sup_{0\le t\le 1} \|Y_t^n - Y_t\|_{\mathbb{L}^2} + \|Z^n - Z\|_{\mathbb{H}^2} \right\} < \infty$$

Theorem (Gobet-Labart 06) Under additional conditions :

$$\limsup_{n\to\infty} n\|Y_0^n - Y_0\|_{\mathbb{L}^2} \quad < \quad \infty$$

Weak error...

Simulation of Backward SDE's

1. Simulate trajectories of the forward process X (well understood)

2. Backward scheme :

$$\left\| \begin{array}{rcl} \hat{Y}^n_{t_n} & = & g\left(X^n_{t_n}\right) \\ \hat{Y}^n_{t_{i-1}} & = & \widehat{\mathbb{E}}^n_{t_{i-1}}\left[\hat{Y}^n_{t_i}\right] + f\left(X^n_{t_{i-1}}, \hat{Y}^n_{t_{i-1}}, \hat{Z}^n_{t_{i-1}}\right)\Delta t_i\ , \quad 1 \le i \le n\ , \\ \hat{Z}^n_{t_{i-1}} & = & \widehat{\mathbb{E}}^n_{t_{i-1}}\left[\hat{Y}^n_{t_i}\dfrac{\Delta W_{t_i}}{\Delta t_i}\right] \end{array} \right.$$

Simulation of BSDEs : bound on the rate of convergence

Error estimate for the Malliavin-based scheme

Theorem For $p > 1$:

$$\limsup_{n\to\infty} \max_{0\le i\le n} n^{-1-d/(4p)} N^{1/2p} \left\| \hat{Y}^n_{t_i} - Y^n_{t_i} \right\|_{\mathbb{L}^p} < \infty$$

For the time step $\frac{1}{n}$, and limit case $p = 1$:

$$\text{rate of convergence of } \frac{1}{\sqrt{n}}$$
$$\text{if and only if}$$
$$n^{-1-\frac{d}{4}} N^{1/2} = n^{1/2}, \quad \text{i.e. } N = n^{3+\frac{d}{2}}$$

Recent developments in Crisan, Manolarakis and T.

Outline

Introducing the scheme without reference to BSDEs

• Isolate a diffusion part in the equation :

$$0 = -v_t(t,x) - \frac{1}{2}1\Delta v(t,x) - f\left(.,v,Dv,D^2v\right)(t,x)$$

• The Monte Carlo component   Let $X_s = x + 1W_{s-t+h}$, $s \geq t-h$, evaluate at (s, X_s), and take expectations :

$$\begin{aligned} 0 &= \mathbb{E}\left[\int_{t-h}^{t} -(v_t + \frac{1}{2}\Delta v)(s,X_s)ds - \int_{t-h}^{t} f\left(.,v,Dv,D^2v\right)(s,X_s)ds\right] \\ &= v(t-h,x) - \mathbb{E}\left[v(t,X_t) + \int_{t-h}^{t} f\left(.,v,Dv,D^2v\right)(s,X_s)ds\right] \end{aligned}$$

• The Finite Differences component   Natural approximation

$$v^h(t-h,x) = \mathbb{E}\left[v^h(t,X_t)\right] + h\,f\left(t,x,\left(\mathbb{E}[D^i v^h(t,X_t)]\right)_{0\leq i\leq 2}\right)$$

for $v^h(t-h,x)$, need $v^h(t,.)$, $Dv^h(t,.)$, $D^2v^h(t,.)$! !

Intuition From Greeks Calculation

- Using the approximation $f'(x) \sim_{h=0} \mathbb{E}[f'(x+W_h)]$:

$$\begin{aligned} f'(x) &\sim \int f'(x+y)\frac{e^{-y^2/(2h)}}{\sqrt{2\pi}}dy \\ &= \int f(x+y)\frac{y}{h}\frac{e^{-y^2/(2h)}}{\sqrt{2\pi}}dy \\ &= \mathbb{E}\left[f(x+W_h)\frac{W_h}{h}\right] \end{aligned}$$

- Similarly, by an additional integration by parts :

$$\begin{aligned} f''(x) &= \int f(x+y)\frac{y^2-h}{h^2}\frac{e^{-y^2/(2h)}}{\sqrt{2\pi}}dy \\ &= \mathbb{E}\left[f(x+W_h)\left(\frac{W_h^2-h}{h^2}\right)\right] \end{aligned}$$

A probabilistic numerical scheme for fully nonlinear PDEs

We have then reduced the scheme to :

$$v^h(t-h,x) = \mathbb{E}\left[v^h(t,X_t)\right]$$
$$+\, hf\left(t,x,\mathbb{E}\Big[v^h(t,X_t)\Big],\mathbb{E}\Big[v^h(t,X_t)\frac{W_h}{h}\Big],\mathbb{E}\Big[v^h(t,X_t)\frac{W_hW_h^{\mathrm{T}}-hI_d}{h^2}\Big]\right)$$

Or, in BSDE notations $Y^n_{t_n} = g\left(X^n_{t_n}\right)$, and for $1 \le i \le n$:

$$\begin{aligned}
Y^n_{t_{i-1}} &= \mathbb{E}^n_{i-1}\left[Y^n_{t_i}\right] + \Delta t_i\, f\left(t_i, X^n_{t_{i-1}}, Y^n_{t_{i-1}}, Z^n_{t_{i-1}}, \Gamma^n_{t_{i-1}}\right)\\
Z^n_{t_{i-1}} &= \mathbb{E}^n_{i-1}\left[Y^n_{t_i}\frac{\Delta W_{t_i}}{\Delta t_i}\right]\\
\Gamma^n_{t_{i-1}} &= \mathbb{E}^n_{i-1}\left[Y^n_{t_i}\frac{\Delta W_{t_i}\Delta W_{t_i}^{\mathrm{T}}-\Delta t_i I_d}{|\Delta t_i|^2}\right]
\end{aligned}$$

Connection with Finite Differences : $X_h := x + W_h$

- Consider the binomial approximation of the Brownian motion

$$``W_h \sim \sqrt{h}\left(\frac{1}{2}\delta_{\{1\}} + \frac{1}{2}\delta_{\{-1\}}\right)" \text{ Then :}$$

$$\mathbb{E}\left[\psi'(X_h)\right] = \mathbb{E}\left[\psi(X_h)\frac{W_h}{h}\right] \sim \frac{\psi(x+\sqrt{h}) - \psi(x-\sqrt{h})}{2\sqrt{h}}$$

- With the trinomial approximation of the Brownian motion

$$``W_h \sim \sqrt{3h}\left(\frac{1}{6}\delta_{\{1\}} + \frac{2}{3}\delta_{\{0\}} + \frac{1}{6}\delta_{\{-1\}}\right)" \text{ Then :}$$

$$\mathbb{E}\left[\psi''(X_h)\right] = \mathbb{E}\left[\psi(X_h)\frac{W_h^2 - h}{h^2}\right] \sim \frac{\psi(x+\sqrt{3h}) - 2\psi(x) + \psi(x-\sqrt{3h})}{3h}$$

Convergence of the Discretization

<Fahim, T. and Warin 09>

Theorem (i) *Suppose that f is Lipschitz uniformly in x and $\varepsilon I \leq \nabla_\gamma f \leq \sigma\sigma^{\mathrm{T}}$. Then*

$$Y_0^n(t,x) \longrightarrow v(t,x) \qquad \textit{uniformly on compacts}$$

where v is the unique viscosity solution of the nonlinear PDE.
(ii) *If f is either convex in (y,z,γ), i.e. HJB operator,*

$$-Ch^{1/10} \leq v - v^h \leq Ch^{1/4}$$

- Proof by means of the viscosity theory for monotonic schemes
- In contrast with FD (Bonnans and Zidani), the present scheme is automatically monotic...

An Implementable MC-FD Scheme

In order to define an implementable scheme, we consider an approximation $\hat{\mathbb{E}}^N$ of $\mathbb{E}$, and :

$$
\begin{aligned}
\hat{Y}^n_{t_n} &= g\left(X^n_{t_n}\right) \text{ and for } 1 \le i \le n : \\
\hat{Y}^n_{t_{i-1}} &= \hat{\mathbb{E}}^N_{i-1}\left[\hat{Y}^n_{t_i}\right] + \Delta t_i \, f\left(t_i, X^n_{t_{i-1}}, \hat{Y}^n_{t_{i-1}}, \hat{Z}^n_{t_{i-1}}, \hat{\Gamma}^n_{t_{i-1}}\right) \\
\hat{Z}^n_{t_{i-1}} &= \hat{\mathbb{E}}^N_{i-1}\left[\hat{Y}^n_{t_i} \frac{\Delta W_{t_i}}{\Delta t_i}\right] \\
\hat{\Gamma}^n_{t_{i-1}} &= \hat{\mathbb{E}}^N_{i-1}\left[\hat{Y}^n_{t_i} \frac{|\Delta W_{t_i}|^2 - \Delta t_i}{|\Delta t_i|^2}\right]
\end{aligned}
$$

Convergence of the MC-FD Scheme

Assumption There exist constants $C, \lambda, \nu > 0$ such that for any bounded function ψ :

$$\left\| (\hat{\mathbb{E}}^N - \mathbb{E})\psi(W_h)H_i(W_h) \right\|_p \quad \leq \quad C\, h^{-\lambda} N^{-\nu}$$

where $H_0(w) = 1,\ H_1(w) = w,\ H_2(w) = w^{\mathrm{T}} w - I$

Theorem (i) *Suppose that f is Lipschitz uniformly in x and $\varepsilon I \leq \nabla_\gamma f \leq \sigma\sigma^{\mathrm{T}}$, and $\lim_{h\to 0} h^{\lambda+2} N_h^\nu = \infty$. Then*

$$\hat{v}^h(t,x) \longrightarrow v(t,x) \qquad \text{uniformly on compacts}$$

where v is the unique viscosity solution of the nonlinear PDE.
(ii) If f is either convex or concave in (y, z, γ), i.e. HJB operator, and $\lim_{h\to 0} h^{\lambda+\frac{21}{10}} N_h^\nu > 0$

$$-Ch^{1/10} \leq \|v - \hat{v}^h\|_p \leq Ch^{1/4}$$

Comments on the MC-FD scheme

• in BSDEs the drift coefficient μ of the forward SDE can be changed arbitrarily by Girsanov theorem : importance sampling...

• in the present setiing (fully nonlinear, 2BSDEs) both μ and σ can be changed : previous theorem and numerical results however recommend prudence...

Portfolio optimization

$\{S_t, t \in [0, T]\}$ price process of n financial securities
$\{\theta_t, t \in [0, T]\}$ with values in $\mathbb{R}^n$: portfolio
$\{r_t, t \in [0, T]\}$: interest rates process

Wealth process $X_t^\theta = X_0 + \int_0^t \theta_u \cdot \mathrm{diag}[S_u]^{-1} dS_u + (X_u^\theta - \theta_u \cdot 1) r_u du$

Portfolio optimization problem :

$$v_0 \quad := \quad \sup_{\theta \in \mathcal{A}} \mathbb{E}\left[- \exp\left(-\eta X_T^\theta \right) \right]$$

One dimensional implementation

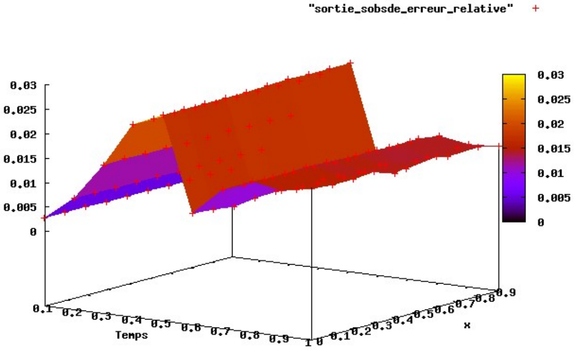
$r_t = 0$ and $dS_t = S_t\sigma(dW_t + \lambda dt)$

An explicit solution is available...

- V is the characterized by the fully nonlinear PDE

$$-V_t + \frac{1}{2}\lambda^2 \frac{(V_y)^2}{V_{yy}} = 0 \quad \text{and} \quad V(T,.) = U$$

Numerical results, one-dimensional market



Varying the drift of the FSDE

Drift FSDE	Relative error (Regression)
-1	0,0648429
-0,8	0,0676044
-0,6	0,0346846
-0,4	0,0243774
-0,2	0,0172359
0	0,0124126
0,2	0,00880041
0,4	0,00656142
0,6	0,00568952
0,8	0,00637239

Varying the volatility of the FSDE

Diffusion FSDE	Relative error (Regression)	Relative error (Quantization)
0,2	0,581541	0,526552
0,4	0,42106	0,134675
0,6	0,0165435	0,0258884
0,8	0,0170161	0,00637319
1 0,	0124126	0,0109905
1,2	0,0211604	0,0209174
1,4	0,0360543	0,0362259
1,6	0,0656076	0,0624566

A two-dimensional example

$r_t = 0$, and security price process defined by the Heston model :

$$
\begin{aligned}
dS_t &= S_t\left(\mu dt + \sqrt{Y_t}\, dW_t^{(1)}\right) \\
dY_t &= k(m - Y_t)dt + c\sqrt{Y_t}\left(\rho dW_t^{(1)} + \sqrt{1-\rho^2}\, dW_t^{(2)}\right)
\end{aligned}
$$

Then $V(t,x,s,y) = V(t,x,y)$ is characterized by :

$$
\begin{aligned}
-v_t - k(m-y)v_y - \frac{1}{2}c^2 y v_{yy} + \frac{(\mu v_x + \rho c y v_{xy})^2}{2 y v_{xx}}. &= 0 \\
V(T,x,y) = -e^{-\eta x}
\end{aligned}
$$

A quasi explicit solution by T. Zariphopoulou :

$$
\begin{aligned}
v(t,x,y) &= -e^{-\eta x}\left\| \exp\left(-\frac{1}{2}\int_t^T \frac{\mu^2}{\tilde{Y}_s}ds\right)\right\|_{\mathbb{L}^{1-\rho^2}} \\
d\tilde{Y}_t &= \left(k(m-\tilde{Y}_t) - \mu c\rho\right)dt + c\sqrt{\tilde{Y}_t}\, dW_t
\end{aligned}
$$

Numerical results : two-dimensional portfolio optimization

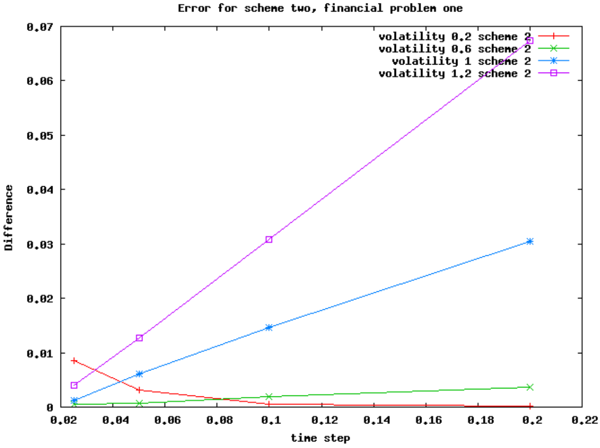


Fig.: Difference between calculation and reference for the problem of portfolio optimization under stochastic volatility

A five-dimensional example

Stochastic interest rate process

$$dr_t = \kappa(b - r_t)dt + \zeta dW_t^{(0)}$$

Two tradable assets defined by :

$$\begin{cases} dS_t^{(1)} = \mu_1 S_t^{(1)} dt + \sigma_1 \sqrt{Y_t^{(1)}} \sqrt{S_t^{(1)}} dW_t^{(1,1)} \\ dY_t^{(1)} = k_1 \left(m_1 - Y_t^{(1)} \right) dt + c_1 \sqrt{Y_t^{(1)}} dW_t^{(1,2)} \end{cases}$$

$$\begin{cases} dS_t^{(2)} = S_t^{(2)} \left(\mu_2 dt + \sigma_2 \sqrt{Y_t^{(2)}} dW_t^{(2,1)} \right) \\ dY_t^{(2)} = k_2 \left(m_2 - Y_t^{(2)} \right) dt + c_2 \sqrt{Y_t^{(2)}} dW_t^{(2,2)} \end{cases}$$

$\left(W^{(0)}, W^{(1,1)}, W^{(1,2)}, W^{(2,1)}, W^{(2,2)}\right)$ is a Brownian motion in $\mathbb{R}^5$

Numerical results : five-dimensional portfolio optimization

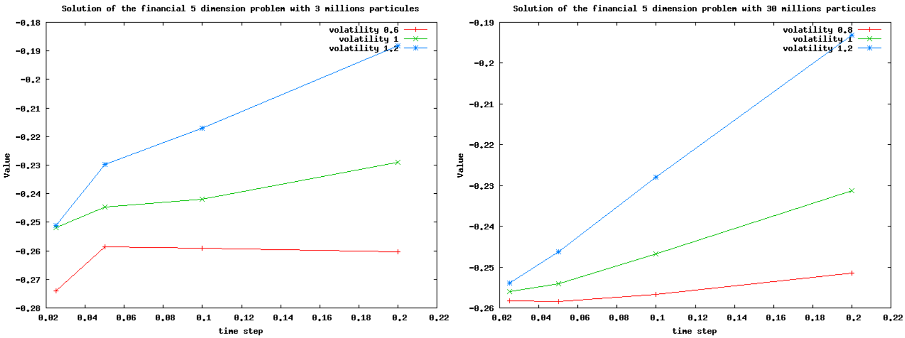


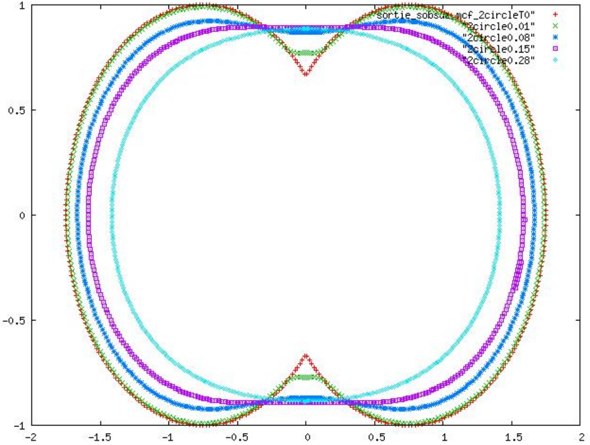
Fig.: Five dimensional portfolio optimization problem for different volatilities with 3 millions and 30 millions particles

Mean Curvature Flow : two connected circles

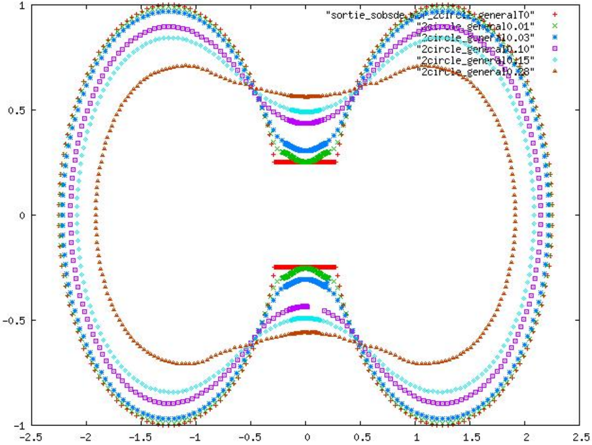
Mouvement of a codimension 1 submanifold along the inward normal with speed proportional to the mean curvature (codimension 1) :

$$u_t - \Delta u - \frac{D^2 u \; Du \cdot Du}{|Du|^2} = 0 \quad \text{and} \quad u(0,.) = g$$

Mean Curvature Flow : two intersecting circles



Mean Curvature Flow : two connected circles I



Mean Curvature Flow : two connected circles II

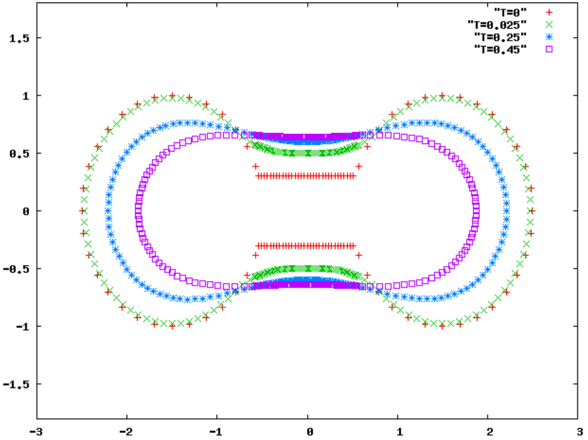


Fig.: Mean curvature flow problem in 2D

Mean Curvature Flow : a sphere in dimension 3

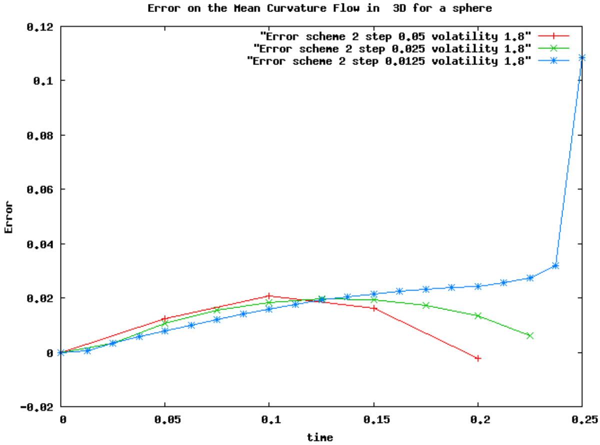


Fig.: Mean curvature flow problem for different time step and diffusions