Multilevel Monte Carlo for PDE solutions based on Feynman-Kac theorem

Mike Giles, Francisco Bernal

Mathematical Institute, University of Oxford, UK

Instituto Superior Tecnico, Portugal

ICIAM 2015

August 10, 2015

Outline

- Feynman-Kac formula
- prior work – Gobet & Menozzi
- multilevel Monte Carlo
- prior work – Higham *et al*
- new idea – approximating a conditional expectation
- outline numerical analysis

Feynman-Kac formula

Suppose that $u(x,t)$ satisfies the parabolic PDE

$$\frac{\partial u}{\partial t} + \sum_j a_j \frac{\partial u}{\partial x_j} + \tfrac{1}{2} \sum_{j,k,l} b_{jl} b_{kl} \frac{\partial^2 u}{\partial x_j \partial x_k} - V\, u + f = 0$$

in bounded domain D, subject to $u(x,t) = g(x,t)$ on the boundary ∂D.

It will be assumed that $f(x,t), V(x,t), a(x,t), b(x,t)$ are all Lipschitz continuous, and $g(x,t)$ is continuously twice-differentiable.

Feynman-Kac formula

Feynman and Kac proved that $u(x,t)$ can also be expressed as

$$u(x,t) = \mathbb{E}\left[\int_t^\tau E(t,s)\, f(X_s,s)\,\mathrm{d}s + E(t,\tau)\, g(X_\tau,\tau) \mid X_t = x \right]$$

where X_t satisfies the SDE

$$\mathrm{d}X_t = a(X_t,t)\,\mathrm{d}t + b(X_t,t)\,\mathrm{d}W_t,$$

with W_t being a Brownian motion with independent components, τ is the first time at which X_t leaves D, and

$$E(t_0,t_1) = \exp\left(- \int_{t_0}^{t_1} V(X_t,t)\,\mathrm{d}t \right).$$

Note: in the special case in which $f(x,t)\!=\!0$, $g(x,t)\!=\!t$, $V(x,t)\!=\!0$ $u_{exit}(x,t)$ is the expected exit time.

Numerical approximation

An Euler-Maruyama discretisation with uniform timestep h gives

$$\widehat{X}_{t_{n+1}} = \widehat{X}_{t_n} + a(\widehat{X}_{t_n}, t_n)\, h + b(\widehat{X}_{t_n}, t_n)\, \Delta W_n,$$

with initial data $\widehat{X}_0 = x$ at time t.

If $\widehat{X}(t)$ is the piecewise-constant interpolant, we then have

$$\widehat{u}(x,t) = \mathbb{E}\left[\int_t^{\widehat{\tau}} \widehat{E}(t,s)\, f(\widehat{X}(s),s)\, \mathrm{d}s + \widehat{E}(t,\widehat{\tau})\, g(\widehat{X}(\tau),\widehat{\tau}) \right].$$

with $\widehat{\tau}$ being the exit time, and

$$\widehat{E}(t_0,t_1) = \exp\left(- \int_{t_0}^{t_1} V(\widehat{X}_t, t)\, \mathrm{d}t \right).$$

Prior work – Gobet & Menozzi

The Euler-Maruyama method has strong accuracy

$$\left(\mathbb{E}\left[\sup_{[0,\min(\tau,\widehat{\tau})]} \|X_t - \widehat{X}(t)\|^2 \right] \right)^{1/2} = O(h^{1/2} |\log h|^{1/2}),$$

and Gobet & Menozzi (2007) proved that it has weak error

$$u(x,t) - \widehat{u}(x,t) = O(h^{1/2}).$$

For standard Monte Carlo method, ε RMS accuracy needs $O(\varepsilon^{-2})$ paths, each with $h = O(\varepsilon^2)$, so total cost is $O(\varepsilon^{-4})$

Gobet & Menozzi (2010) reduced this to $O(\varepsilon^{-3})$ by shifting the boundary by $O(h^{1/2})$ to improve the weak error to $O(h)$.

Multilevel Monte Carlo

MLMC is based on the telescoping sum

$$\mathbb{E}[\widehat{P}_L] = \mathbb{E}[\widehat{P}_0] + \sum_{\ell=1}^{L} \mathbb{E}[\widehat{P}_\ell - \widehat{P}_{\ell-1}]$$

where $\widehat{P}_\ell$ represents an approximation using timestep $h_\ell = 2^{-\ell}\, h_0$, with weak convergence

$$\mathbb{E}[\widehat{P}_\ell - P] = O(2^{-\alpha\,\ell})$$

If $\widehat{Y}_\ell$ is an unbiased estimator for $\mathbb{E}[\widehat{P}_\ell - \widehat{P}_{\ell-1}]$, based on N_ℓ samples, with variance

$$\mathbb{V}[\widehat{Y}_\ell] = O(N_\ell^{-1}\, 2^{-\beta\,\ell})$$

and expected cost

$$\mathbb{E}[C_\ell] = O(N_\ell\, 2^{\gamma\,\ell}), \quad \ldots$$

Multilevel Monte Carlo

... then the finest level L and the number of samples N_ℓ on each level can be chosen to achieve an RMS error of ε at an expected cost

$$C = \begin{cases} O\left(\varepsilon^{-2}\right), & \beta > \gamma, \\ O\left(\varepsilon^{-2}(\log\varepsilon)^2\right), & \beta = \gamma, \\ O\left(\varepsilon^{-2-(\gamma-\beta)/\alpha}\right), & 0 < \beta < \gamma. \end{cases}$$

Prior work – Higham

Higham *et al* (2013) developed a MLMC treatment of the exit time problem:

- Euler-Maruyama discretisation
- $O(h_\ell^{1/2})$ weak convergence $\Longrightarrow\ \alpha = 1/2$
- $\mathbb{V}[\widehat{P}_\ell - \widehat{P}_{\ell-1}] = O(h_\ell^{1/2} |\log h_\ell|^{1/2}) \Longrightarrow\ \beta \approx 1/2$
- $O(h_\ell^{-1})$ cost per path $\Longrightarrow\ \gamma = 1$

Hence, overall cost is $O(\varepsilon^{-3} |\log \varepsilon|^{1/2})$.

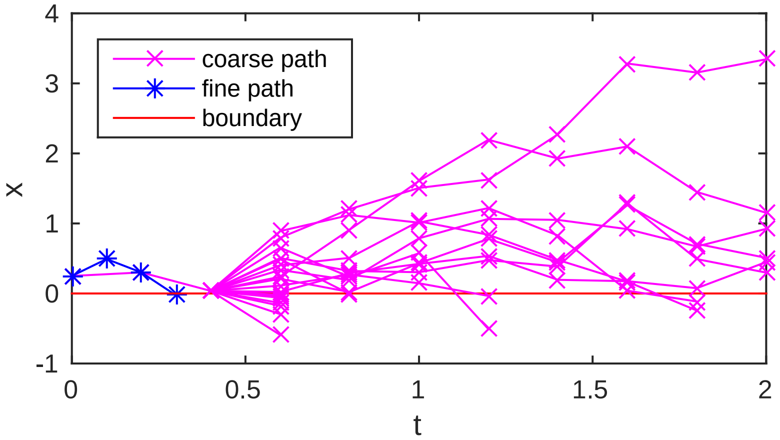
Gobet & Menozzi's boundary treatment would improve this to $O(\varepsilon^{-2.5} |\log \varepsilon|^{1/2})$.

G & Primozic (2011) developed $O(\varepsilon^{-2})$ treatment using Milstein discretisation for SDEs with special commutativity property.

MLMC challenge

When coarse or fine path exits the domain, the other is within $O(h^{1/2})$.

However, there is a $O(h^{1/2})$ probability that it will not exit the domain until much later $\implies V_\ell \approx O(h^{1/2})$.



MLMC challenge

How can we do better?

Similar to previous work on digital options (G, Burgos), and also used by Dickmann & Schweizer for stopping times, split second path into multiple copies and average their outputs to approximate the conditional expectation.

Approximately $O(h^{1/2})$ expected time to exit for second path, so can afford to use approximately $O(h^{-1/2})$ copies of second path.

This gives an approximation to the conditional expectation resulting in $\widehat{P}_\ell - \widehat{P}_{\ell-1} \approx O(h^{1/2})$, so $V_\ell \approx O(h)$.

This gives $\alpha = 1/2,\ \beta \approx 1,\ \gamma \approx 1$ and the complexity is $O(\varepsilon^{-2}\,|\log\varepsilon|^3)$.

Numerical Analysis

Assumption 1: There is a Lipschitz constant L_f such that

$$|f(x,t) - f(y,s)| \leq L_f \left(\|x-y\|_2 + |t-s| \right), \quad \forall (x,t),(y,s) \in D,$$

and there are similar Lipschitz constants L_g, L_V, L_a, L_b, L_u, L_{exit} for g, V, a, b, u, u_{exit}. In addition, $g \in C^{2,1}(D)$, with a bounded Hessian $H_g \equiv \partial^2 g / \partial x^2$.

Comment: assumption about L_u, L_{exit} may require the boundary ∂D to be smooth, or at least not have re-entrant corners.

Assumption 2: There is a unit computational cost for each timestep, and in determining whether or not $\widehat{X}_{t_{n+1}} \in D$.

Assumption 3: There exist constants C_u and C_{exit} s.t. for all $(x,t) \in D$

$$\begin{aligned} |u(x,t) - \widehat{u}(x,t)| &\leq C_u \, h^{1/2} \\ |u_{exit}(x,t) - \widehat{u}_{exit}(x,t)| &\leq C_{exit} \, h^{1/2} \end{aligned}$$

Numerical Analysis

Defining the output functional

$$P_t = \int_t^\tau E(t,s)\, f(X_s,s)\, \mathrm{d}s + E(t,\tau)\, g(X_\tau,\tau)$$

we get

Lemma

Given Assumption 1, there exists C such that for any $(x,t) \in D$

$$\mathbb{V}[P_t \,|\, X_t{=}x] \leq C\ \mathbb{E}[\tau{-}t \,|\, X_t{=}x].$$

$$\mathrm{d}\left(E(t,s)\, g(X_s,s) \right) =$$

$$E(t,s) \left(\left(-\, V\, g + \dot{g} + (\nabla g)^T\, a + \tfrac{1}{2}\, \mathrm{trace}(b^T H_g\, b) \right)\, \mathrm{d}s + (\nabla g)^T\, b\, \mathrm{d}W_s \right)$$

with a, b, g, $\dot{g} \equiv \partial g / \partial t$, ∇g, H_g, all evaluated at (X_s, s).

Numerical Analysis

Hence, $P_t - g(x,t) = p^{(1)} + p^{(2)}$, where

$$p^{(1)} = \int_t^\tau E(t,s)\left(f - V\,g + \dot{g} + (\nabla g)^T a + \tfrac{1}{2}\,\mathrm{trace}(b^T H_g\, b)\right) \mathrm{d}s,$$

$$p^{(2)} = \int_t^\tau E(t,s)\,(\nabla g)^T\, b\,\mathrm{d}W_s.$$

Considering the second term, since $E(t,s) \le \exp(T\|V\|_\infty)$, we have

$$\begin{aligned} \mathbb{E}[(p^{(2)})^2] &= \mathbb{E}\left[\int_t^\tau (E(t,s))^2\ \|(\nabla g)^T b\|_2^2\ \mathrm{d}s\right] \\ &\le \exp(2T\|V\|_\infty)\ \|\nabla g\|_{2,\infty}^2 \|b\|_{2,\infty}^2\ \mathbb{E}[\tau - t \,|\, X_t = x], \end{aligned}$$

where $\|b\|_{2,\infty}$, $\|\nabla g\|_{2,\infty}$ are the maximum values of $\|b\|_2$, $\|\nabla g\|_2$ over D.

The first term is handled similarly to complete the proof.

Numerical Analysis

The following is a standard result:

Lemma

If W and Z are independent random variables, then

$$Y = M^{-1} \sum_{m=1}^{M} f(W, Z^{(m)})$$

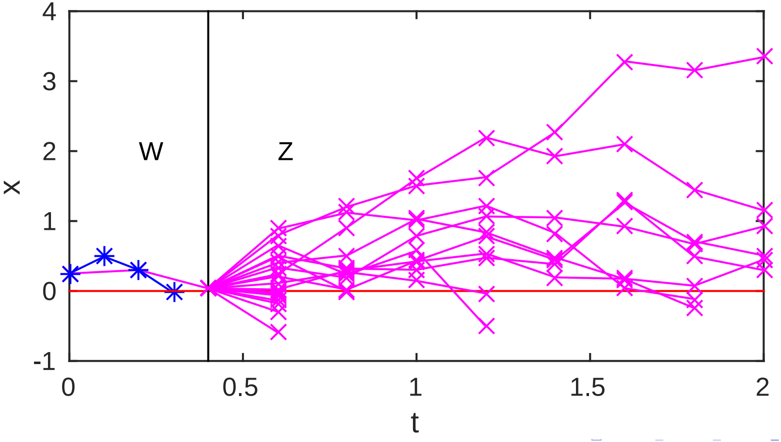
with independent samples W and $Z^{(m)}$ is an unbiased estimator for $\mathbb{E}\left[f(W,Z)\right]$ and its variance is

$$\mathbb{V}[Y] = \mathbb{V}\left[\mathbb{E}[f(W,Z)\,|\,W]\right] + M^{-1}\,\mathbb{E}\left[\mathbb{V}[f(W,Z)\,|\,W]\right].$$

Numerical Analysis

Let $\underline{\tau}$ be the exit time of the first of a pair of coarse/fine paths, and $\overline{\tau}$ be $\underline{\tau}$ rounded up to the end of a coarse timestep.

In our application W represents the Brownian path up to $\overline{\tau}$, and Z is the Brownian path therafter.



Numerical Analysis

Lemma

Given Assumptions 1 and 3, we have

$$\begin{aligned}
\mathbb{E}[\sup_{[0,\underline{\tau}]} \|\widehat{X}_{\ell,t} - \widehat{X}_{\ell-1,t}\|^2\,]^{1/2} &= O(h_{\ell-1}^{1/2}\,|\log h_{\ell-1}|^{1/2}) \\
\mathbb{E}[\,\|\widehat{X}_{\ell,\underline{\tau}} - \widehat{X}_{\ell-1,\overline{\tau}}\|^2\,]^{1/2} &= O(h_{\ell-1}^{1/2}\,|\log h_{\ell-1}|^{1/2}) \\
\Longrightarrow \quad \mathbb{V}\left[\mathbb{E}[\widehat{P}_\ell - \widehat{P}_{\ell-1} \,|\, W]\right] &= O(h_{\ell-1} |\log h_{\ell-1}|)
\end{aligned}$$

The key to the proof is that if $0 \le t \le \tau$ then

$$P_0 = \int_0^t E(0,s)\, f(X_s,s)\,\mathrm{d}s + E(0,t)\left\{\int_t^\tau E(t,s)\, f(X_s,s)\,\mathrm{d}s + E(t,\tau)\, g(X_\tau,\tau)\right\}$$

$$\Rightarrow \; \mathbb{E}[P_0 \,|\, \mathcal{F}_t] - \int_0^t E(0,s)\, f(X_s,s)\,\mathrm{d}s \;=\; E(0,t)\,\mathbb{E}[P_t \,|\, \mathcal{F}_t] \;=\; E(0,t)\, u(X_t,t)$$

Something similar for the discrete approximation yields

$$\mathbb{E}[\widehat{P}_\ell - \widehat{P}_{\ell-1} \,|\, W] = O(h_{\ell-1}^{1/2}\,|\log h_{\ell-1}|^{1/2})$$

Numerical Analysis

Lemma

Given Assumptions 1 and 3,

$$\mathbb{E}\left[\, \mathbb{V}[\widehat{P}_\ell - \widehat{P}_{\ell-1} \,|\, W] \,\right] = O(h_{\ell-1}^{1/2} |\log h_{\ell-1}|^{1/2})$$

The key here is that, similar to the SDE analysis, there exists C such that

$$\begin{aligned} \mathbb{V}[\widehat{P}_\ell - \widehat{P}_{\ell-1} \,|\, W] &\leq C\, \mathbb{E}\,[\,|\widehat{\tau}_\ell - \widehat{\tau}_{\ell-1}| \,|\, W]) \\ &= O(h_{\ell-1}^{1/2} \,|\log h_{\ell-1}|^{1/2}) \end{aligned}$$

Numerical Analysis

Corollary

Under the given assumptions, an RMS error of ε can be achieved with an $O(\varepsilon^{-2}|\log \varepsilon|^3)$ expected computational cost.

The proof is slightly non-standard because of log terms.

- $h_\ell = 4^{-\ell} h_0$
- $M_\ell = \lceil 2^\ell / \ell^{1/2} \rceil$ paths in the splitting estimator
- expected cost is $O(h_\ell^{-1})$
- variance $V_\ell = O(h_\ell\, |\log h_\ell|) = O(h_\ell\, \ell)$.

This eventually gives the desired cost bound.

Conclusions

- conditional expectation / splitting is a useful technique in MLMC estimation
- in Feynmac-Kac application it improves the MLMC variance from approximately $O(h^{1/2})$ to approximately $O(h)$, reducing the complexity to $O(\varepsilon^{-2}|\log\varepsilon|^3)$
- numerical analysis is now complete but relies on key assumption of uniform $O(h^{1/2})$ weak convergence – an open problem

Webpages:

`people.maths.ox.ac.uk/gilesm/mlmc.html`

`people.maths.ox.ac.uk/gilesm/mlmc_community.html`

`people.maths.ox.ac.uk/gilesm/acta/`