MLMC tricks for discontinuous functionals

Mike Giles

Mathematical Institute, University of Oxford

MCM 2021

Sept 18, 2021

Outline

- MLMC and the problem with discontinuities
- 9 approaches:
  - explicit smoothing
  - integration/differentiation
  - Malliavin calculus
  - conditional expectation
  - change-of-measure
  - splitting
  - conditional integration with root-finding
    (Christian Bayer talk tomorrow at 15:30)
  - adaptive sampling (Al Haji-Ali talk yesterday)
  - path branching (next talk by Al Haji-Ali)

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 to output P on level ℓ.

If $\widehat{Y}_\ell$ has expected value $\mathbb{E}[\widehat{P}_\ell - \widehat{P}_{\ell-1}]$, with variance V_ℓ and cost C_ℓ, then for a given target RMS error ε, the number of independent samples on each level can be optimised to give overall cost

$$\varepsilon^{-2} \left(\sum_{\ell=0}^{L} \sqrt{C_\ell V_\ell} \right)^2 \sim \begin{cases} \varepsilon^{-2} C_0 V_0, & C_\ell V_\ell \to 0, \\ \varepsilon^{-2} L^2 C_L V_L, & C_\ell V_\ell \to \text{const}, \quad \ell \to \infty \\ \varepsilon^{-2} C_L V_L, & C_\ell V_\ell \to \infty. \end{cases}$$

Multilevel Monte Carlo

In the case of an SDE

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

with an output quantity of interest $P \equiv f(S_T)$, the standard estimator is

$$\widehat{Y}_\ell = \widehat{P}_\ell - \widehat{P}_{\ell-1}$$

where the same Brownian motion W_t is used for both $\widehat{P}_\ell$ and $\widehat{P}_{\ell-1}$, but with different uniform timesteps h.

If f is Lipschitz, with constant L_f, then

$$V_\ell \;\leq\; \mathbb{E}\left[(\Delta P)^2\right] \;\leq\; L_f^2\,\mathbb{E}\left[(\widehat{S}_\ell - \widehat{S}_{\ell-1})^2\right]$$

so we have $V_\ell = O(h_\ell)$ for Euler-Maruyama discretisation, $V_\ell = O(h_\ell^2)$ for Milstein, and cost $C_\ell = O(h_\ell^{-1})$ in both cases.

Digital options

In mathematical finance, a digital put option payoff is 0 or 1, depending on whether S_T is above or below the strike K. The problem is that a small change in the path can give a big change in the payoff.

Using the Euler-Maruyama approximation the strong error is $O(h^{1/2})$,

$$\Longrightarrow \quad \widehat{S}_\ell - \widehat{S}_{\ell-1} = O(h_\ell^{1/2}).$$

An $O(h_\ell^{1/2})$ fraction of fine/coarse pairs straddles the strike,

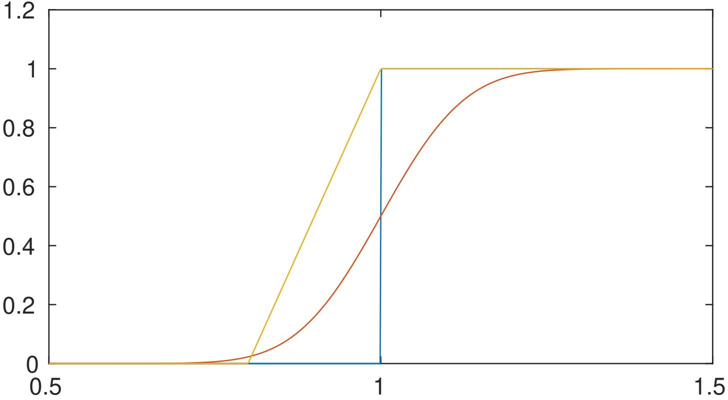
$$\Longrightarrow \quad V_\ell = O(h_\ell^{1/2}), \text{ and hence the complexity is } O(\varepsilon^{-5/2}).$$

Using the Milstein approximation the strong error is $O(h)$ so $V_\ell{=}O(h_\ell)$. This is better, but the kurtosis is $O(h_\ell^{-1})$ which causes problems in estimating V_ℓ.

Explicit smoothing

Digital options are also a problem for pathwise sensitivity analysis, estimating the sensitivity of the expectation to parameter change

One common solution to this is to explicitly smooth the payoff, and that can be used also for MLMC – involves a tradeoff between bias and variance



Explicit smoothing

G, Nagapetyan, Ritter (2015) used explicit smoothing for estimating CDFs.

For a scalar S_T, to estimate $C(x) = \mathbb{P}(S_T < x) = \mathbb{E}[H(x - S_T)]$ where $H(x)$ is the Heaviside step function, the approach was

- use MLMC to estimate $C(x_j)$ for a set of spline points x_j
- interpolate with a cubic spline

To improve the MLMC variance, $H(x)$ was replaced by $H_\delta(x)$ which smoothed H over a width of δ. Overall, had to balance four errors:

- SDE discretisation bias on finest level
- MLMC sampling error
- smoothing error
- interpolation error

Integration/differentiation

Krumscheid, Nobile (2018) came up with a slightly different approach for estimating CDFs, based on

$$\frac{\mathrm{d}}{\mathrm{d}x}\,\mathbb{E}\left[\max(0, x - S_T)\right] = \mathbb{E}\left[H(x - S_T)\right]$$

- use MLMC to estimate $\mathbb{E}[\max(0, x_j - S_T)]$ for spline points x_j
- interpolate with a cubic spline
- differentiate to obtain desired CDF $C(x)$

This avoids the smoothing error, but differentiating the cubic spline amplifies the noise in the spline data.

Malliavin calculus

On a similar note, Altmayer & Neuenkirch (2015) used Malliavin calculus integration by parts to handle discontinuous payoffs with the Heston model

Used on its own it improves the asymptotic behaviour, but makes the variance on coarse levels worse.

To address this, they split the payoff into a smooth part (handled by standard MLMC) and a compact-support discontinuous part (handled using Malliavin MLMC)

Again one key lesson is that techniques which help with computing sensitivities can also help with MLMC.

Conditional expectation

For the Milstein discretisation, one "fix" for digital options is to use E-M approximation for the final timestep, then take conditional expectation over final fine path Brownian increment ΔW_N.

For fine path

$$\widehat{S}_T = \widehat{S}_{T-h} + a_{T-h}\, h_\ell + b_{T-h}\, \Delta W_N,$$

$$\Longrightarrow \quad \widehat{P}_\ell = \Phi\left(\frac{\widehat{S}_{T-h} + a_{T-h}\, h_\ell - K}{b_{T-h}\sqrt{h_\ell}} \right)$$

while for the coarse path,

$$\widehat{S}_T = \widehat{S}_{T-2h} + 2\, a_{T-2h}\, h_\ell + b_{T-2h}\, (\Delta W_{N-1} + \Delta W_N),$$

$$\Longrightarrow \quad \widehat{P}_{\ell-1} = \Phi\left(\frac{\widehat{S}_{T-2h} + 2\, a_{T-2h}\, h_\ell + b_{T-2h}\, \Delta W_{N-1} - K}{b_{T-2h}\sqrt{h_\ell}} \right)$$

Conditional expectation

Analysis (G, Debrabant, Roessler, 2019) proves $V_\ell = O(h_\ell^{3/2})$ and the kurtosis is $O(h_\ell^{-1/2})$, so much better.

Heuristically, this is because there is an $O(h_\ell^{1/2})$ probability of paths being within $O(h_\ell^{1/2})$ of the strike, and for these

$$\widehat{S}_\ell - \widehat{S}_{\ell-1} = O(h_\ell) \quad \Longrightarrow \quad \Delta P = O(h_\ell^{1/2})$$

Unfortunately, the conditional expectation approach does not help with the E-M discretisation where

$$\widehat{S}_\ell - \widehat{S}_{\ell-1} = O(h_\ell^{1/2}) \quad \Longrightarrow \quad \Delta P = O(1)$$

Conditional expectation

This conditional expectation is a standard technique for smoothing the payoff to enable pathwise sensitivity calculations. (L'Ecuyer, Glasserman)

Another example is a down-and-out barrier options, where the option is knocked out if the path drops below a certain value.

Payoff can be smoothed by computing probability of this happening, conditional on computed path at discrete timesteps.

Again, this works well for both pathwise sensitivity analysis and MLMC.

(G, 2008, Burgos, G, 2012)

Change of measure

Another approach with the Milstein discretisation is to use a change of measure – similar to Likelihood Ratio Method for sensitivity analysis

For both the fine and coarse paths, we have conditional Gaussian distributions for $\widehat{S}_T$, with different means and variances.

Can perform a change of measure to the same Gaussian distribution, and then pick the same sample for both paths.

$$\widehat{P}_\ell - \widehat{P}_{\ell-1} = \widehat{P}(\widehat{S}_T)\,(R_\ell - R_{\ell-1})$$

where $R_\ell, R_{\ell-1}$ are the respective Radon-Nikodym derivatives. Works well in multiple dimensions where often cannot evaluate the analytic conditional expectation. (Burgos, 2014)

Problem: still doesn't help with Euler-Maruyama discretisation because $R_\ell - R_{\ell-1} = O(1)$.

Change of measure

An earlier example of its use was for a Merton-style jump-diffusion SDE with path-dependent jump rate $\lambda(S,t)$. (Xia, G, 2012, Xia, 2014)

Problem is that coarse and fine paths will jump at different times; one might jump just before T, the other just after $\implies$ large $\widehat{P}_\ell - \widehat{P}_{\ell-1}$

Solution: use Glasserman & Merener thinning technique, over-sampling possible jump times using rate $\lambda_{sup} > \lambda(S,t)$, and combine with change of measure for identical acceptance/rejection decision for fine/coarse paths.

Leads to an estimator which looks like

$$\widehat{P}_\ell \, R_\ell - \widehat{P}_{\ell-1} \, R_{\ell-1}$$

and gives $V_\ell = O(h_\ell^2)$ when combined with Milstein discretisation.

Similar ideas have also been used for importance sampling and chaotic SDEs (Fang, G, 2019)

Splitting

Back again to the multi-dimensional digital option.

The conditional expectation can be estimated numerically by averaging over a number of independent samples for the final Brownian increment.

$O(h_\ell^{-1})$ samples can be used without increasing the path cost significantly.

This is sufficient to reduce V_ℓ to about the same level as using the analytic conditional expectation.

Bonus: can use more accurate Milstein method for final timestep.

Burgos, G (2012) and Burgos (2014) have also used splitting for MLMC for pathwise sensitivity analysis for put/call options.

Splitting

Bernal, G (2019) also used splitting for Feynman-Kac functionals arising for stopped diffusions – SDE calculations which terminate when the path leaves the domain.

The issue here is that when a fine path exits, there is an $O(h_\ell^{1/2})$ probability that the corresponding coarse path does not leave until much later.

This is solved by estimating a conditional expectation by splitting the coarse path into $O(h_\ell^{-1/2})$ independent sub-simulations.

V_ℓ is improved from $O(h_\ell^{1/2})$ to approximately $O(h_\ell)$.

Conditional integration with root-finding

Bayer, Ben Hammouda, Tempone (2020 – talk on Thurs at 15:30) split the random inputs W into a scalar Z and the remainder W_r.

They then express the desired MLMC level ℓ expectation as

$$\mathbb{E}[\widehat{P}_\ell - \widehat{P}_{\ell-1}] = \mathbb{E}\left[\ \mathbb{E}[\widehat{P}_\ell - \widehat{P}_{\ell-1} \,|\, W_r]\ \right]$$

and observe that in many financial applications it is possible to perform this split in a way such that

$$\mathbb{E}[\widehat{P}_\ell \,|\, W_r], \quad \mathbb{E}[\widehat{P}_{\ell-1} \,|\, W_r]$$

are smooth functions of W_r, and can be evaluated very accurately by performing root-finding in Z to locate the one discontinuity.

For a scalar SDE, Z could be W_T, the terminal value of the driving Brownian motion, and W_r would be the other random variables required for a Brownian Bridge construction of the Brownian increments.

Conclusions

- most MLMC applications use "plain" MLMC with no need for any of these tricks – good strong convergence implies small variance
- for applications with discontinuities, there is a growing toolkit of tricks to consider, with the same tricks being used in widely differing applications.
- in many cases, ideas have been taken from sensitivity analysis which also has problems with discontinuous functionals
- next talk by Al Haji-Ali introduces another new trick, and yesterday he talked about another one based on adaptive sampling
- Christian Bayer presents another technique tomorrow at 15:30

Webpages:

http://people.maths.ox.ac.uk/gilesm/mlmc/

http://people.maths.ox.ac.uk/gilesm/mlmc.html

http://people.maths.ox.ac.uk/gilesm/mlmc_community.html

References

M. Altmayer, A. Neuenkirch. 'Multilevel Monte Carlo quadrature of discontinuous payoffs in the generalized Heston model using Malliavin integration by parts'. SIAM Journal on Financial Mathematics, 6(1):22-52, 2015.

C. Bayer, C. Ben Hammouda, R. Tempone. 'Numerical smoothing and hierarchical approximations for efficient option pricing and density estimation'. arXiv preprint 2003.05708, 2020.

S. Burgos, G. 'Computing Greeks using multilevel path simulation'. pp.281-296 in Monte Carlo and Quasi-Monte Carlo Methods 2010, Springer, 2012.

S. Burgos. 'The computation of Greeks with multilevel Monte Carlo'. PhD thesis, University of Oxford, 2014.

W. Fang, G. 'Multilevel Monte Carlo method for ergodic SDEs without contractivity'. Journal of Mathematical Analysis and Applications, 476(1):149-176, 2019

G. 'Improved multilevel Monte Carlo convergence using the Milstein scheme'. pp.343-358, in Monte Carlo and Quasi-Monte Carlo Methods 2006, Springer, 2008.

References

G. ‘Multilevel Monte Carlo methods’. Acta Numerica, 24:259-328, CUP, 2015.

G, F. Bernal. ’Multilevel estimation of expected exit times and other functionals of stopped diffusions’. SIAM/ASA Journal on Uncertainty Quantification, 6(4):1454-1474, 2018.

G, K. Debrabant, A. Roessler. ’Analysis of multilevel Monte Carlo path simulation using the Milstein discretisation’. Discrete and Continuous Dynamical Systems – series B, 24(8):3881-3903, 2019.

G, T. Nagapetyan, K. Ritter. ’Multi-Level Monte Carlo approximation of distribution functions and densities’. SIAM/ASA Journal on Uncertainty Quantification, 3:267-295, 2015.

S. Krumscheid, F. Nobile. ’Multilevel Monte Carlo approximation of functions’. SIAM/ASA Journal on Uncertainty Quantification, 6(3):1256-1293, 2018.

Y. Xia, G. ‘Multilevel path simulation for jump-diffusion SDEs’, pp.695-708 in Monte Carlo and Quasi-Monte Carlo Methods 2010, Springer, 2012.

Y. Xia. ‘Multilevel Monte Carlo for jump processes’. PhD thesis, University of Oxford, 2014.