SDE Numerical Analysis for Multilevel Function Approximation

Filippo De Angelis, Mike Giles, Christoph Reisinger

Mathematical Institute, University of Oxford

Stochastic Numerics and Statistical Learning, KAUST

May 30, 2024

Outline

- Stefan Heinrich's MLMC for parametric integration
- MLFA for SDEs – extension of Heinrich's approach
  - numerical analysis for integrable SDEs
  - randomised MLMC for SDE approximations
  - MIMC decomposition for SDE approximations
  - numerical analysis for smooth and non-smooth "payoffs"
  - strong convergence for pathwise sensitivities
- conclusions and references

MLMC for parametric integration

Stefan Heinrich's original MLMC research (2001) concerned the approximation of

$$f(\theta) = \mathbb{E}[P(\theta;\omega)],$$

given exact sampling of $P(\theta;\omega)$ at unit cost (finite-dimensional ω).

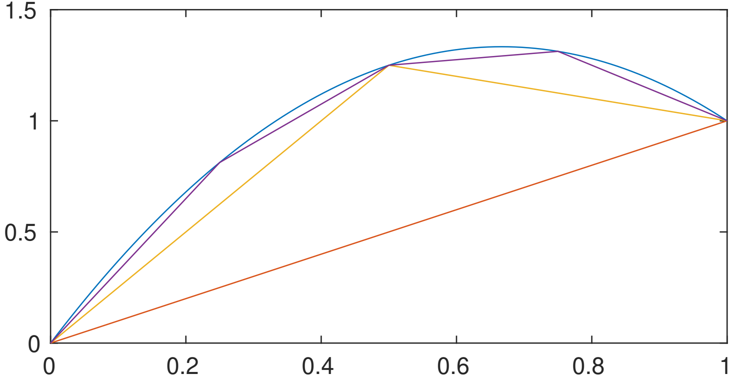
For simplicity can think of $\theta \in [0,1]^d$.

In his formulation, the MLMC telescoping sum is

$$f \approx I_L[f] = I_0[f] + \sum_{\ell=1}^{L} I_\ell[f] - I_{\ell-1}[f]$$

where $I_\ell[f]$ represents a level ℓ interpolation, e.g. piecewise linear interpolation in 1D with spacing $2^{-\ell}$, and tensor product multilinear interpolation in higher dimensions.

MLMC for parametric integration



Here we see 3 levels of approximation, with the difference $I_\ell[f] - I_{\ell-1}[f]$ getting progressively smaller as ℓ increases.

MLMC for parametric integration

Heinrich then approximates $(I_\ell - I_{\ell-1})[f]$ through Monte Carlo sampling at required values of θ:

$$(I_\ell - I_{\ell-1})[f] \approx \frac{1}{M_\ell} \sum_{m=1}^{M_\ell} (I_\ell - I_{\ell-1})[P(\,\cdot\,;\omega^{\ell,m})]$$

As $\ell \to \infty$, $(I_\ell - I_{\ell-1})[f] \to 0$ and $\mathbb{V}\,[\,(I_\ell - I_{\ell-1})[P]\,] \to 0$, so fewer MC samples needed on finer levels.

Analysis assumes the number of θ points increases exponentially with dimension so the resulting complexity for linear interpolation is of order

$$\begin{array}{rl} \varepsilon^{-2}, & d < 2r \\ \varepsilon^{-2}\,|\log\varepsilon|^2, & d = 2r \\ \varepsilon^{-d/r}, & d > 2r \end{array}$$

where d is the dimension and $r \in \{1,2\}$ is the degree of smoothness of f and P with respect to θ.

MLMC for parametric integration

Heinrich's work is the key starting point for our current work which extends it in several directions:

- PDEs with appropriate numerical approximation (not today)
- sparse grid interpolation to address curse of dimensionality (not today)
- weaker assumptions on smoothness of $P(\theta;\omega)$
- numerical approximation of $f(\theta) \equiv \mathbb{E}[P(\theta;\omega)]$ in cases without a finite variance, finite expected cost unbiased estimator
- numerical analysis for PDEs (not today) and SDEs

MLFA for parametric integration

If $Y(\theta;\omega)$ is an unbiased estimator for $f(\theta) \equiv \mathbb{E}[P(\theta;\omega)]$, with

$$\begin{aligned} \|(I_\ell - I)\,[f]\,\| &< c_1\, 2^{-r\ell} \\ \mathbb{V}\,[(I_\ell - I_{\ell-1})\,[Y]\,] &< c_2\, 2^{-s\ell} \end{aligned}$$

and the total expected cost is bounded by $c_3 \displaystyle\sum_0^L 2^{d\ell} M_\ell$, for M_ℓ samples per level, then ε r.m.s. accuracy can be achieved with cost of order

$$\begin{array}{ll} \varepsilon^{-2}, & d < s \\ \varepsilon^{-2} |\log \varepsilon|^2, & d = s \\ \varepsilon^{-2-(d-s)/r}, & d > s \end{array}$$

This is a slight generalisation of Heinrich's original result which corresponds to $s = 2r$.

Numerical analysis

In 1D, using piecewise linear interpolation, the maximum value of $(I_\ell - I_{\ell-1})\,[Y]$ is at a midpoint of a coarse θ interval, so the numerical analysis involves bounding

$$\mathbb{E}\left[\left(\delta^2 Y(\theta;\omega)\right)^2\right]$$

where $\delta^2 Y(\theta_0;\omega) = Y(\theta_0{+}\Delta\theta;\omega) - 2\,Y(\theta_0;\omega) + Y(\theta_0{-}\Delta\theta;\omega)$

We are concerned with applications in mathematical finance for which

$$P(\theta;\omega) = g(S_T(\theta;\omega))$$

with S_T being the final value for an SDE solution with ω representing the driving Brownian motion.

Numerical analysis for integrable SDEs

For an integrable SDE, we use $Y(\theta;\omega) = P(\theta;\omega) = g(S_T(\theta;\omega))$.

We assume the SDE satisfies the usual conditions and therefore for each $p>0$ there exists $c^{(p)}$ such that

$$\mathbb{E}\left[\|S_T\|^p\right] \leq c^{(p)}$$

Furthermore, we assume the drift and diffusion coefficients are smooth w.r.t. θ and therefore for integer $q>0$, and any $p>0$, there exists $c^{(p,q)}$ such that

$$\mathbb{E}\left[\left\|\frac{\partial^q S_T}{\partial\theta^q}\right\|^p\right] \leq c^{(p,q)}$$

This can be proved given bounded derivatives for the drift and diffusion coefficients, but I haven't yet found a reference for it.

Numerical analysis for integrable SDEs

For twice-differentiable payoff functions,

$$\begin{aligned}\dot{Y}(\theta,\omega) &= g'(S_T(\theta,\omega))\ \dot{S}_T(\theta,\omega),\\ \ddot{Y}(\theta,\omega) &= g''(S_T(\theta,\omega))\ (\dot{S}_T(\theta,\omega))^2 + g'(S_T(\theta,\omega))\ \ddot{S}_T(\theta,\omega).\end{aligned}$$

where $\dot{Y} \equiv \partial Y/\partial\theta$, and $g' \equiv \mathrm{d}g/\mathrm{d}S$. We then have

$$\delta^2 Y = \int_{\theta_0-\Delta\theta}^{\theta_0+\Delta\theta} (\Delta\theta - |\theta-\theta_0|)\ \ddot{Y}(\theta,\omega)\,\mathrm{d}\theta,$$

and hence $\delta^2 Y = O(\Delta\theta^2)$ and $\mathbb{E}[(\delta^2 Y)^2] = O(\Delta\theta^4)$, giving $s{=}4$ as well as $r{=}2$ in the meta-theorem.

This corresponds to the smooth case analysed by Stefan Heinrich. However, in mathematical finance the payoff function is rarely twice-differentiable.

Numerical analysis for non-smooth payoffs

At the other extreme, consider a digital option for which the payoff is an indicator function $g(S_T) = \mathbb{1}_{S_T \in K}$

For this, we follow previous research in assuming that there exists a constant c such that for all θ, and all $\delta > 0$,

$$\mathbb{P}[d(S_T, \partial K) < \delta] < c\,\delta$$

where $d(S_T, \partial K)$ is the distance of S_T from the boundary ∂K.

Heuristically, this corresponds to S_T having a bounded density, but it also requires the set K to not be pathological.

G., Haji-Ali (2024) give conditions under which this assumption is satisfied, and also examples of pathological K for which it is not.

Numerical analysis for non-smooth payoffs

Heuristic analysis:

- $O(\Delta\theta)$ probability of $S_T(\theta_0;\omega)$ being within $O(\Delta\theta)$ of ∂K
- $\Longrightarrow$ $O(\Delta\theta)$ probability of $S_T(\theta;\omega)$ for $\theta_0 - \Delta\theta < \theta < \theta_0 + \Delta\theta$ crossing ∂K, giving $\delta^2 Y = O(1)$
- otherwise, $\delta^2 Y = 0$
- hence, $\mathbb{E}[(\delta^2 Y)^2] = O(\Delta\theta)$

The rigorous version of this gives

$$\mathbb{E}[(\delta^2 Y)^2] = o(\Delta\theta^{1-\delta})$$

for any $\delta > 0$, so $s \approx 1$, but $r = 2$.

Numerical analysis for non-smooth payoffs

Similarly, for Lipschitz functions with a bounded second derivative except on ∂K (e.g. European put/call functions), the heuristic analysis is:

- $O(\Delta\theta)$ probability of $S_T(\theta_0;\omega)$ being within $O(\Delta\theta)$ of ∂K
- $\implies$ $O(\Delta\theta)$ probability of $S_T(\theta;\omega)$ for $\theta_0-\Delta\theta < \theta < \theta_0+\Delta\theta$ crossing ∂K, giving $\delta^2 Y = O(\Delta\theta)$
- otherwise, $\delta^2 Y = O(\Delta\theta^2)$
- hence, $\mathbb{E}[(\delta^2 Y)^2] = O(\Delta\theta^3)$

The rigorous version of this gives

$$\mathbb{E}[(\delta^2 Y)^2] = o(\Delta\theta^{3-\delta})$$

for any $\delta > 0$, so $s \approx 3$, but $r = 2$.

MLMC for SDE approximations

Almost all SDEs in mathematical finance are not integrable, and instead need to be approximated, e.g. using the Euler-Maruyama discretisation.

The standard MLMC method for path approximations uses

$$\mathbb{E}[\widehat{P}_L] = \sum_{\ell=0}^{L} \mathbb{E}[\Delta \widehat{P}_\ell], \qquad \Delta \widehat{P}_\ell \equiv \widehat{P}_\ell - \widehat{P}_{\ell-1}, \quad \widehat{P}_{-1} \equiv 0$$

where $\widehat{P}_\ell$ approximates P on level ℓ using timestep $h_\ell = 2^{-\gamma\ell} h_0$.

If there are constants α, β such that

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

then we get the optimal complexity when $\beta > \gamma$.

Randomised MLMC for SDE approximations

If $\beta > \gamma$ then we can use the randomised MLMC of Rhee & Glynn (2015) in which

$$Y = p_{\ell'}^{-1} \Delta P_{\ell'}$$

with ℓ' being a random level chosen with probability $p_\ell \propto 2^{-(\beta+\gamma)\ell/2}$. This works because

$$\mathbb{E}[Y] \;=\; \sum_\ell \mathbb{P}[\ell'{=}\ell]\, p_\ell^{-1}\, \mathbb{E}[\Delta P_\ell] \;=\; \sum_\ell \mathbb{E}[\Delta P_\ell] \;=\; \mathbb{E}[P]$$

and it can be proved that Y has finite variance and finite expected cost because $\beta > \gamma$.

We then have to work out the corresponding s value for the interpolation meta-theorem.

MIMC for SDE approximations

Alternatively, we can use MIMC approach of Haji-Ali, Nobile & Tempone (2016). The starting point is the interpolation decomposition:

$$f \approx \sum_{\ell=0}^{L} (I_\ell - I_{\ell-1})[f], \qquad I_{-1}[f] \equiv 0,$$

where I_ℓ uses a dense interpolation with spacing proportional to $2^{-\ell}$.

We then replace f with a timestep approximation expansion

$$f \approx \sum_{\ell=0}^{L} \sum_{\ell'=0}^{L'_\ell} \Delta I_\ell[\Delta f_{\ell'}], \qquad \Delta I_\ell[\Delta f_{\ell'}] \equiv (I_\ell - I_{\ell-1})[f_{\ell'} - f_{\ell'-1}]$$

in which L'_ℓ is a decreasing function of ℓ, since $\Delta I_\ell[f]$ becomes smaller as ℓ increases and so less relative accuracy is required in its approximation.

MIMC for SDE approximations

The final step is to replace $\Delta I_\ell[\Delta f_{\ell'}]$ by a Monte Carlo estimate, giving the MIMC estimator

$$\widetilde{f} = \sum_{\ell=0}^{L} \sum_{\ell'=0}^{L'_\ell} \left(\frac{1}{M_{\ell,\ell'}} \sum_{m=1}^{M_{\ell,\ell'}} \Delta I_\ell[\Delta g_{\ell'}(\cdot;\omega^{\ell,\ell',m})] \right)$$

We now need to choose $L, L'_\ell, M_{\ell,\ell'}$ to achieve the desired accuracy at the minimum cost.

$$\begin{aligned} \mathbb{E}[\widetilde{f}-f] &= (I_L-I)[f] + \sum_{\ell=0}^{L} (I_\ell-I_{\ell-1})[f_{L'(\ell)}-f] \\ \Longrightarrow \quad \left\| \mathbb{E}[\widetilde{f}-f] \right\| &\leq \|(I_L-I)[f]\| + \sum_{\ell=0}^{L} \|(I_\ell-I_{\ell-1})[f_{L'(\ell)}-f]\| \end{aligned}$$

and

$$\mathbb{V}[\widetilde{f}] = \sum_{\ell=0}^{L} \sum_{\ell'=0}^{L'_\ell} \left(\frac{1}{M_{\ell,\ell'}} \mathbb{V}\left[\Delta I_\ell[\Delta g_{\ell'}(\cdot;\omega)]\right] \right)$$

MIMC for SDE approximations

If we have

$$\begin{aligned} \| \Delta I_\ell [\Delta f_{\ell'}] \| &< c_1 \, 2^{-r\ell - \alpha\ell'} \\ \mathbb{V}\,[\Delta I_\ell [\Delta g_{\ell'}]\,] &< c_2 \, 2^{-s\ell - \beta\ell'} \end{aligned}$$

and the total cost is bounded by

$$c_3 \sum_{\ell=0}^{L} \sum_{\ell'=0}^{L'_\ell} 2^{d\ell+\gamma\ell'} M_{\ell,\ell},$$

then ε RMS accuracy can be achieved at a computational cost of order

$$\begin{aligned} &\varepsilon^{-2}, && \eta < 0 \\ &\varepsilon^{-2-\eta}\, |\log \varepsilon|^p, && \eta \geq 0 \end{aligned}$$

for some p (see MIMC analysis by Haji-Ali *et al* (2016)), where

$$\eta = \max\left(\frac{\gamma - \beta}{\alpha}, \frac{d - s}{r} \right).$$

Numerical analysis

The challenge now for both the randomised MLMC and the MIMC approaches is to bound

$$\mathbb{V}\left[\Delta I_\ell[\Delta g_{\ell'}]\right]$$

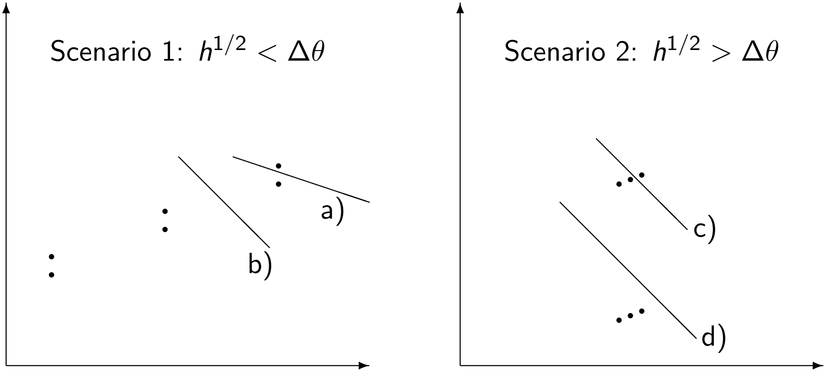
On a level ℓ with spacing $\Delta\theta$, and level ℓ' with timestep h, this involves bounding

$$\begin{aligned}\mathbb{V}\Big[\ &\Big(g(\widehat{S}(\theta_0{-}\Delta\theta, h, \omega)) \ - \ 2\,g(\widehat{S}(\theta_0, h, \omega)) \ + \ g(\widehat{S}(\theta_0{+}\Delta\theta, h, \omega))\Big) \\ &- \Big(g(\widehat{S}(\theta_0{-}\Delta\theta, 2h, \omega)) - 2\,g(\widehat{S}(\theta_0, 2h, \omega)) + g(\widehat{S}(\theta_0{+}\Delta\theta, 2h, \omega))\Big)\ \Big]\end{aligned}$$

In the smooth case, this variance is $O(\Delta\theta^4 h)$ for the E-M discretisation, and $O(\Delta\theta^4 h^2)$ for Milstein.

Numerical analysis

In the non-smooth case, there are a number of scenarios to consider regarding the position of ∂K



Numerical analysis

Eventually, the conclusion is that the variance for the Euler-Maruyama discretisation is approximately

$$O(\min(h^{1/2}, \Delta\theta))$$

for the digital case, and

$$O(\min(\Delta\theta\, h, \Delta\theta^3))$$

for the Lipschitz case.

These are not of the form required by the meta-theorem, so the analysis has to be extended.

Strong convergence for pathwise sensitivities

The numerical analysis requires the following strong convergence result for the Euler-Maruyama discretisation.

For any $p > 0$ there exists $c^{(p)}$ such that

$$
\begin{aligned}
\mathbb{E}\left[\sup_{0<t<T} \|\widehat{\dot{S}}_t - \dot{S}_t\|^p\right] &\leq c^{(p)} h^{p/2} \\
\mathbb{E}\left[\sup_{0<t<T} \|\widehat{\ddot{S}}_t - \ddot{S}_t\|^p\right] &\leq c^{(p)} h^{p/2}
\end{aligned}
$$

This can also be proved given bounded derivatives for the drift and diffusion coefficients, but I haven't yet found a reference for it, so derived it from first principles

Usual analysis of SDEs

When considering, for simplicity, the autonomous SDE

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

the "usual conditions" assume that $a(S)$ and $b(S)$ are globally Lipschitz, i.e. there exists L such that

$$\|a(v) - a(u)\| + \|b(v) - b(u)\| < L\,\|v{-}u\|, \quad \forall u, v.$$

Under these conditions, the SDE has a unique solution given initial S_0, and for any finite time interval $[0, T]$ and $p > 0$ there exist constants $c_p^{(1)}$, $c_p^{(2)}$ such that

$$\begin{aligned}
\mathbb{E}\left[\sup_{0<t<T} \|S_t\|^p\right] \;&\leq\; c_p^{(1)}, \\
\mathbb{E}\left[\,\|S_t{-}S_{t_0}\|^p\right] \;&\leq\; c_p^{(2)}\,(t{-}t_0)^{p/2}, \quad \text{for } 0 < t_0 < t < T.
\end{aligned}$$

Usual analysis of SDE discretisations

Furthermore, for the Euler-Maruyama discretisation

$$\widehat{S}_{(n+1)h} = \widehat{S}_{nh} + a(\widehat{S}_{nh})\, h + b(\widehat{S}_{nh})\, \Delta W_n,$$

with a uniform timestep of h, we have $O(h^{1/2})$ strong convergence so that for any $p>0$ there exists $c_p^{(3)}$ such that

$$\mathbb{E}\left[\sup_{0<t<T} \|\widehat{S}_t - S_t\|^p\right] \;\leq\; c_p^{(3)}\, h^{p/2}.$$

This strong convergence is important for the effectiveness and analysis of MLMC algorithms.

Pathwise sensitivities

Suppose now that S_t is scalar, and $a(\theta; S)$ and $b(\theta; S)$ depend smoothly on a scalar parameter θ as well as S

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

and we are interested in the expected value of a "payoff" function $P(S_T)$,

$$f(\theta) = \mathbb{E}\left[\, P(S_T(\theta; \{W_t\}_{0\leq t\leq T}))\,\right]$$

and want to compute its derivative

$$\dot{f} \equiv \frac{\mathrm{d}f}{\mathrm{d}\theta}$$

Pathwise sensitivities

If P is globally Lipschitz and piecewise smooth, then

$$\dot{f} \;\equiv\; \frac{\mathrm{d}}{\mathrm{d}\theta}\,\mathbb{E}[\,P(S_T)\,] \;=\; \mathbb{E}[\,\dot{P}(S_T)\,]$$

where

$$\dot{P} = \frac{\mathrm{d}P}{\mathrm{d}S}\;\dot{S}_T$$

and $\dot{S}_t \equiv \dfrac{\mathrm{d}S_t}{\mathrm{d}\theta}$ satisfies the SDE

$$\mathrm{d}\dot{S}_t = (\dot{a}(\theta;S_t) + a'(\theta;S_t)\,\dot{S}_t)\,\mathrm{d}t + (\dot{b}(\theta;S_t) + b'(\theta;S_t)\,\dot{S}_t)\,\mathrm{d}W_t$$

subject to $\dot{S}_0 = 0$, with $\dot{a} \equiv \dfrac{\partial a}{\partial \theta}$, $a' \equiv \dfrac{\partial a}{\partial S}$, and $\dot{b}, b'$ defined similarly.

(Note: analysis can be extended with P and S_0 depending explicitly on θ)

Pathwise sensitivities

The Euler-Maruyama discretisation of the pathwise sensitivity SDE is

$$\begin{aligned}\widehat{\dot{S}}_{(n+1)h} &= \widehat{\dot{S}}_{nh} + \left(\dot{a}(\theta;\widehat{S}_{nh}) + a'(\theta;\widehat{S}_{nh})\,\widehat{\dot{S}}_{nh}\right) h \\ &\qquad + \left(\dot{b}(\theta;\widehat{S}_{nh}) + b'(\theta;\widehat{S}_{nh})\,\widehat{\dot{S}}_{nh}\right) \Delta W_n\end{aligned}$$

This is also the equation one gets by differentiating the E-M discretisation of the original SDE.

Question: what is the order of strong convergence of $\widehat{\dot{S}}$ to $\dot{S}$?

Previous MLMC work has assumed

$$\mathbb{E}\left[\sup_{0<t<T} \|\widehat{\dot{S}}_t - \dot{S}_t\|^p\right] = O(h^{p/2})$$

but I have not found a reference for this.

Pathwise sensitivities

The pathwise sensitivity SDE can be appended to the original SDE to form a vector SDE with $\mathsf{S}_t \equiv (S_t, \dot{S}_t)^T$

$$\mathrm{d}\mathsf{S}_t = \mathsf{a}(\theta; \mathsf{S}_t)\,\mathrm{d}t + \mathsf{b}(\theta; \mathsf{S}_t)\,\mathrm{d}W_t.$$

I think past work assumed this vector SDE satisfied the "usual conditions" and hence gave 1/2-order strong convergence for both $\widehat{S}$ and $\widehat{\dot{S}}$.

However, this is not true in general.

Pathwise sensitivities

Looking at the pathwise sensitivity SDE

$$\mathrm{d}\dot{S}_t = \left(\dot{a}(\theta; S_t) + a'(\theta; S_t)\, \dot{S}_t\right) \mathrm{d}t + \left(\dot{b}(\theta; S_t) + b'(\theta; S_t)\, \dot{S}_t\right) \mathrm{d}W_t$$

even if we assume all derivatives of $a(\theta; S)$ and $b(\theta; S)$ are bounded, then

$$\begin{aligned} a'(\theta; v_1)\, v_2 - a'(\theta; u_1)\, u_2 &= \left(a'(\theta; v_1) - a'(\theta; u_1)\right) v_2 + a'(\theta; u_1)\,(v_2 - u_2) \\ &= a''(\theta; w)\, v_2\, (v_1 - u_1) + a'(\theta; u_1)\,(v_2 - u_2) \end{aligned}$$

for some $u_1 < w < v_1$.

The problem is that $|a''(\theta; w)\, v_2| \to \infty$ as $v_2 \to \infty$ unless $a''(\theta; w) = 0$, and something similar applies for $b'(\theta; S)\, \dot{S}$.

Pathwise sensitivities

If we use the shorthand $a_t \equiv a(\theta; S_t)$, $\dot{a}_t \equiv \dot{a}(\theta; S_t)$, $a'_t \equiv a'(\theta; S_t)$, and similarly for $b_t, \dot{b}_t, b'_t$ and higher derivatives, then the first order pathwise sensitivity SDE is

$$\mathrm{d}\dot{S}_t = (\dot{a}_t + a'_t\, \dot{S}_t)\, \mathrm{d}t + (\dot{b}_t + b'_t\, \dot{S}_t)\, \mathrm{d}W_t$$

The second order pathwise sensitivity SDE is then

$$\mathrm{d}\ddot{S}_t = (\ddot{a}_t + 2\dot{a}'_t \dot{S}_t + a''_t (\dot{S}_t)^2 + a'_t \ddot{S}_t)\, \mathrm{d}t + (\ddot{b}_t + 2\dot{b}'_t \dot{S}_t + b''_t (\dot{S}_t)^2 + b'_t \ddot{S}_t)\, \mathrm{d}W_t$$

and the $(\dot{S}_t)^2$ terms makes it even clearer that the "usual conditions" are not satisfied.

However, notice that $\dot{S}_t$ in the first equation, and $\ddot{S}_t$ in the second, are multiplied by a'_t and b'_t which are bounded

Pathwise sensitivities

There is a large literature on the approximation of SDEs which do not satisfy the usual conditions.

These use modified numerical approximations (e.g. tamed schemes, or adaptive timesteps) for which stability and strong convergence can be proved.

However, with these pathwise equations there is no problem using the standard Euler-Maruyama discretisation – all that is needed is a new numerical analysis to prove it has the observed $O(h^{1/2})$ strong convergence order.

Numerical analysis

The numerical analysis is not difficult – essentially retraces the steps of the standard analysis.

Focussing on the first order sensitivity equation, the key is that in the drift and diffusion terms $\dot{S}_t$ is multiplied by the bounded a'_t and b'_t.

Arbitrary moments of all other terms are bounded due to standard results for S_t and $\widehat{S}_t$.

Beyond this, the methodology is standard: use Jensen, Hölder, and Burkholder-Davis Gundy inequalities to set things up for finally using Grönwall's inequality

Numerical analysis: SDE

Lemma

For a given time interval $[0,T]$, *and any* $p \geq 2$, *there exists a constant* $c_p^{(1)}$ *such that*

$$\mathbb{E}\left[\sup_{0<t<T} |\dot{S}_t|^p\right] \leq c_p^{(1)}.$$

Lemma

For a given time interval $[0,T]$, *and any* $p \geq 2$, *there exists a constant* $c_p^{(2)}$ *such that*

$$\mathbb{E}\left[\,|\dot{S}_t - \dot{S}_{t_0}|^p\right] \leq c_p^{(2)}(t-t_0)^{p/2}$$

for any $0 \leq t_0 \leq t \leq T$.

Numerical analysis: Euler-Maruyama scheme

Lemma

For a given time interval $[0, T]$, and any $p \geq 2$, there exists a constant $c_p^{(1)}$ such that

$$\mathbb{E}\left[\sup_{0<t<T} |\widehat{\dot{S}}_t|^p\right] \leq c_p^{(1)}.$$

Theorem

Given the boundedness of all first and second derivatives, for a given time interval $[0, T]$, and any $p \geq 2$, there exists a constant $c_p^{(3)}$ such that

$$\mathbb{E}\left[\sup_{0<t<T} |\widehat{\dot{S}}_t - \dot{S}_t|^p\right] \leq c_p^{(3)} h^{p/2}.$$

Numerical analysis: extensions

- higher derivatives – no problem
- vector SDEs – no problem
- non-autonomous SDEs – no problem if a and b have bounded derivs in θ, S, t (probably OK if θ, S derivatives are $1/2$-Hölder in time)
- other discretisations – probably fine for Milstein discretisation with additional bounded derivatives

Conclusions and future work

Conclusions:

- excellent asymptotic efficiency in approximating parametric functions arising from SDEs – nearly optimal in some cases
- initial numerical results support numerical analysis

Future work:

- a lot more numerical results
- extension of analysis to sparse interpolation – challenging in non-smooth cases
- investigate path-branching and conditional expectation for improved variance for non-smooth cases

References

S. Burgos. 'The computation of Greeks with multilevel Monte Carlo'. PhD thesis, Oxford University, 2013.

F. De Angelis. PhD thesis, in preparation, Oxford University, 2024.

M.B. Giles, A.-L. Haji-Ali. 'Multilevel path branching for digital options'. Annals of Applied Probability, to appear, 2024/5.

A.-L. Haji-Ali, F. Nobile, R. Tempone. 'Multi-index Monte Carlo: when sparsity meets sampling'. Numerische Mathematik, 132:767-806, 2016.

S. Heinrich. 'Multilevel Monte Carlo methods'. Lecture Notes in Computer Science, 2179:58-67, 2001.

C.-H. Rhee, P.W. Glynn. 'Unbiased estimation with square root convergence for SDE models'. Operations Research, 63(5):1026-1043, 2015.