Computing Greeks using multilevel path simulation

S. Burgos, M. Giles

University of Oxford, Mathematical Institute

August 17, 2010

Introduction

- Financial assets evolution modelled as an SDE.
- Option's price is a functional of the underlying asset's price.
- Greeks: sensitivities of option's price to market parameters.
    - Underlying asset's price S_0, volatility σ, interest rate r...
    - Measure exposure to different sources of risk.
- Estimation of option's price and Greeks with Monte Carlo.
- Computing Greeks is more challenging than pricing derivatives.
- Problems arise with discontinuous/nonsmooth payoffs.
- Multilevel path simulation to reduce computational complexity.

Plan

Monte Carlo Greeks

Setting

Notation:

- Underlying asset S, value S_t at time t
- Time interval $[0, T]$ split in N timesteps of size h.
- Interest rate r, volatility σ.
- European option: payoff $P(S_T)$.

Evolution SDE for S:

- $\mathrm{d}S(t) = a(S,t)\mathrm{d}t + b(S,t)\mathrm{d}W_t$
- Euler discretization: $S_{n+1} = S_n + a(S_n, t_n)\, h + b(S_n, t_n)\, \Delta W_n$.
- Milstein discretization: $S_{n+1} = S_n + a(S_n, t_n)\, h + b(S_n, t_n)\, \Delta W_n + \frac{1}{2} b(S_n, t_n) \frac{db}{dS_n} \cdot (\Delta W_n^2 - h)$.

A naive method: finite differences

Pricing:

- The option's value for a certain parameter θ is $V(\theta) = \mathbb{E}\left(P(S_T)\right)$.
- Simulate paths for the underlying asset, get values of $\hat{P}(S_T)$
- Compute MC estimate: $\hat{V}(\theta)$.

Finite differences:

- Take parameter $\theta + \delta\theta$, compute $\hat{V}(\theta + \delta\theta)$
- $\frac{\partial V}{\partial \theta} \approx \frac{\hat{V}(\theta+\delta\theta) - \hat{V}(\theta)}{\delta\theta}$

Limitations:

- Requires two sets of calculations.
- Which $\delta\theta$?
- Discretisation bias/Variance tradeoff.

Pathwise Sensitivities

Principle:

- $\frac{\partial V}{\partial \theta} = \frac{\partial \mathbb{E}(P(S_T))}{\partial \theta} = \mathbb{E}\left(\frac{\partial P(S_T)}{\partial \theta}\right) = \mathbb{E}\left(\frac{\partial P}{\partial S}(S_T) \cdot \frac{\partial S_T}{\partial \theta}\right)$
- $\frac{\partial P}{\partial S}$ is a property of the payoff function.
- $\frac{\partial S_0}{\partial \theta}$ is known.
- $S_{n+1} = S_n + a(S_n, t_n)\, h + b(S_n, t_n)\, \Delta W_n$
  $\Rightarrow \frac{\partial S_{n+1}}{\partial \theta} = \frac{\partial S_n}{\partial \theta} + \frac{\partial a(S_n, t_n)}{\partial \theta}\, h + \frac{\partial b(S_n, t_n)}{\partial \theta}\, \Delta W_n$

Limitations:

- $\frac{\partial P}{\partial S}$ must be defined almost everywhere.
- For the first line to be true, $\frac{P(x+\mathrm{d}x)-P(x)}{\mathrm{d}x}$, must be uniformly integrable.
- Practically, P Lipschitz is a sufficient condition.

Likelihood Ratio Method

Principle:

- Let $p(S)$ the p.d.f. of S_T.
- $$V = \mathbb{E}\left(P(S_T)\right) = \int P(S)\, p(S) \mathrm{d}S$$
$$\Rightarrow \frac{\partial V}{\partial \theta} = \int P \frac{\partial p}{\partial \theta} \mathrm{d}S = \int P \frac{\partial \log p}{\partial \theta}\, p \mathrm{d}S = \mathbb{E}\left(P \frac{\partial \log p}{\partial \theta}\right)$$

Limitations:

- In most cases, $p(S)$ is not known.
- Have to discretize $[0, T]$ in time steps of size h.
- LRM not well suited for path simulations, $\mathbb{V}$ then explodes as $\frac{1}{h}$.

Multilevel Monte Carlo

Complexity Evaluation

We want a Monte Carlo option price with an accuracy ϵ.

- 2 sources of error: discretisation (N steps), finite number of samples (M).

Mean Square Error: $(1/M)\,\mathbb{V}(\hat{P}) + (\mathbb{E}(\hat{P}) - \mathbb{E}(P))^2$.

- 1st term $\searrow$ with the number of paths M: $\mathcal{O}(\frac{1}{M})$.
- 2nd term $\searrow$ with the number of steps N: $\mathcal{O}(\frac{1}{N^2})$.

We want to have these terms $\mathcal{O}(\epsilon^2)$.

- We set $M = \mathcal{O}(\epsilon^{-2})$ and $N = \mathcal{O}(\epsilon^{-1})$.

Total complexity $\mathcal{O}(N\,M) = \mathcal{O}(\epsilon^{-3})$ for option value - often worse for Greeks.

Multilevel path simulation idea

We simulate paths at different levels of fineness:

- At level l, $l = 0 \ldots L$, 2^l timesteps of width $h_l = T/2^l$.
- Let $\hat{P}_l$ be the payoff with level l's discretisation.

We have $\mathbb{E}(\hat{P}_L) = \mathbb{E}(\hat{P}_0) + \sum_{l=1}^{L} \mathbb{E}\left(\hat{P}_l - \hat{P}_{l-1}\right)$.

- With N_l samples, we estimate

$$\mathbb{E}(\hat{P}_l - \hat{P}_{l-1}) \simeq \hat{Y}_l = \frac{1}{N_l} \sum_{i=1}^{N_l} (\hat{P}_l^{(i)} - \hat{P}_{l-1}^{(i)})$$

- We estimate the different $\hat{Y}_l$ independently.
- We reuse the leading brownian motion from $\hat{P}_l$ in $\hat{P}_{l-1}$.

Complexity Improvements

Variance of our combined estimator:

- $\mathbb{V}(\sum_{l=1}^{L}\hat{Y}_l) = \sum_{l=1}^{L}\Big(\mathbb{V}(\hat{Y}_l)\Big) = \sum_{l=1}^{L}\Big(\frac{1}{N_l}\mathbb{V}(\hat{P}_l - \hat{P}_{l-1})\Big)$

Computational cost:

- Cost $\sum_{l=1}^{L}\Big(N_l\, h_l^{-1}\Big)$

We target an accuracy ϵ:

- Choose L to make the discretisation bias small enough.
- Take $N_l \sim \kappa\sqrt{\mathbb{V}(\hat{P}_l - \hat{P}_{l-1})h_l}$ to minimise the variance at a fixed computational cost.
- Choose κ big enough to have a $\mathcal{O}(\epsilon^2)$ variance overall.

MLMC complexity theorem

Theorem

P is a functional of the solution of an SDE. $\hat{P}_l$ is the approximation with a timestep h_l. If there are independent estimators $\hat{Y}_l$ based on N_l samples of cost C_l and constants $\alpha \geq \frac{1}{2}, \beta, c_1, c_2, c_3$ such that

1. $\mathbb{E}(\hat{Y}_0) = \mathbb{E}(\hat{P}_0), \quad \forall l \geq 1 \quad \mathbb{E}(\hat{Y}_l) = \mathbb{E}(\hat{P}_l - \hat{P}_{l-1})$
2. $|\mathbb{E}\left(\hat{P}_l - P\right)| \leq c_1 h_l^{\alpha}$
3. $\mathbb{V}(\hat{Y}_l) \leq c_2 N_l^{-1} h_l^{\beta}$
4. $C_l \leq c_3 N_l h_l^{-1}$

MLMC complexity theorem

Theorem

Then there exists a constant c_4 such that for any $\epsilon < e^{-1}$ there are values of L and N_l for which the estimator $\hat{Y} = \sum_{l=0}^{L} \hat{Y}_l$

1. *Has an MSE* $\mathbb{E}\left[\left(\hat{Y} - \mathbb{E}(P)\right)^2\right] < \epsilon^2$
2. *With a complexity*

$$C \leq \begin{cases} c_4 \epsilon^{-2} & \text{if } \beta > 1 \\ c_4 \epsilon^{-2} (\log \epsilon)^2 & \text{if } \beta = 1 \\ c_4 \epsilon^{-2-(1-\beta)/\alpha} & \text{if } 0 < \beta < 1 \end{cases}$$

Comments on the complexity theorem

The parameter α is known thanks to literature on weak convergence of discretisation schemes.

- Euler and Milstein discretisation: $\alpha = 1$, even with discontinuous payoffs (Bally and Talay, 1995).

The parameter β is related to strong convergence, it determines the efficiency of the multilevel approach.

- For a Lipschitz payoff, Euler: $\beta = 1$, Milstein: $\beta = 2$,
- Not as good for discontinuous payoffs.
- Generally we do not know β *a priori*.

We must create estimators $\hat{Y}_l$ with β as large as possible.

- Pathwise sensitivities reduce the smoothness by one order.
- Multilevel Greeks of nonsmooth payoffs are challenging.

Multilevel Computation of Greeks

Multilevel Greeks estimators

We can write

$$\frac{\partial V}{\partial \theta} = \frac{\partial \mathbb{E}(P)}{\partial \theta} \approx \frac{\partial \mathbb{E}(\hat{P}_L)}{\partial \theta} = \frac{\partial \mathbb{E}(\hat{P}_0)}{\partial \theta} + \sum_{l=1}^{L} \frac{\partial \mathbb{E}(\hat{P}_l - \hat{P}_{l-1})}{\partial \theta}$$

We estimate this value with

$$\begin{cases} \hat{Y}_0 = \frac{1}{N_0} \sum_{i=1}^{M} \frac{\partial \widehat{P_0^{(i)}}}{\partial \theta} \\ \hat{Y}_l = (1/N_l) \sum_{i=1}^{N_l} \left(\frac{\partial \widehat{P_l^{(i)}}}{\partial \theta} - \frac{\partial \widehat{P_{l-1}^{(i)}}}{\partial \theta} \right), \; 1 \le l \le L \end{cases}$$

We compute $\frac{\mathrm{d}\widehat{P_0^{(i)}}}{\mathrm{d}\theta}$, $\frac{\mathrm{d}\widehat{P_{l-1}^{(i)}}}{\mathrm{d}\theta}$, $\frac{\mathrm{d}\widehat{P_l^{(i)}}}{\mathrm{d}\theta}$ as normal MC Greeks with LRM, PwS...

Numerical Experiments

We consider one underlying asset S:

- Black & Scholes $\mathrm{d}S_t = r\, S_t\, \mathrm{d}t + \sigma\, S_t\, \mathrm{d}W_t$
- Milstein discretization:
  $S_{n+1} = S_n \cdot (1 + r\, h + \sigma\, \Delta W_n + \frac{\sigma^2}{2}(\Delta W_n^2 - h)) := S_n \cdot D_n$

We consider European options, illustration with:

- Call: $P(S) = \max(S_T - K, 0)$ (Lipschitz).
- Digital call: $P(S) = \mathbf{1}_{S_T > K}$ (Discontinuous).

We illustrate our methods with two sensitivities:

- Δ: sensitivity to S_0.
- ν: sensitivity to σ.

Multilevel Pathwise Sensitivities

Conditions:

- P Lipschitz.
- Only works for the European call.

Implementation (call):

- $\hat{Y}_l = \frac{1}{N_l} \sum \left[\left(\frac{\partial S}{\partial \theta} \frac{\partial P}{\partial S} \right)^{(l)} - \left(\frac{\partial S}{\partial \theta} \frac{\partial P}{\partial S} \right)^{(l-1)} \right]$
- $\frac{\partial S}{\partial \theta}$ via recurrence:
  - $\frac{\partial S_{n+1}}{\partial S_0} = \frac{\partial S_n}{\partial S_0} \cdot D_n$
  - $\frac{\partial S_{n+1}}{\partial \sigma} = \frac{\partial S_n}{\partial \sigma} \cdot D_n + S_n(\Delta W_n + \sigma(\Delta W_n^2 - h))$
- Sum pairs of fine brownian increments to get coarse increments.
  - Saves the cost of generating new increments.
  - The coarse and rough paths are close $\Rightarrow$ $\mathbb{V}(\hat{Y}_l)$ is smaller.

Multilevel PwS Results

Numerical estimate of β + theorem $\Rightarrow$ Estimated complexity.

European Call:

- Value's estimator: $\beta \approx 2.0 \;\Rightarrow\;$ Complexity $\mathcal{O}(\epsilon^{-2})$.
- Delta's estimator: $\beta \approx 0.8 \;\Rightarrow\;$ Complexity $\mathcal{O}(\epsilon^{-2.2})$.
- Vega's estimator: $\beta \approx 1 \;\Rightarrow\;$ Complexity $\mathcal{O}(\epsilon^{-2}(\log \epsilon)^2)$.

PwS and Conditional Expectations

Reasons:

- Extending the PwS to discontinuous payoffs.
- Payoff smoothing unsatisfactory: tradeoff bias/variance.
- Improving Greeks' convergence rates with nonsmooth payoffs.

Conditional expectation technique:

- $\hat{S}_N(W,Z) = \hat{S}_{N-1}(1+r\,h) + \sigma\,\hat{S}_{N-1}\sqrt{h}\,Z \;:= \mu_W + \sigma_W\,Z$
- $p(\hat{S}_N|W) = \frac{1}{\sigma_W\sqrt{2\pi}}\exp\left(-\frac{(\hat{S}_N-\mu_W)^2}{2\sigma_W^2}\right)$
- Tower property : $\mathbb{E}(P(\hat{S}_N)) = \mathbb{E}_W\left[\mathbb{E}_Z(P(\hat{S}_N)|W)\right]$.
- $\mathbb{E}(P(\hat{S}_N)|W) = \Phi\left(\frac{\mu_W-K}{\sigma_W}\right)$ for the digital call.
- Apply the Multilevel PwS method to this differentiable function.

Multilevel PwS Cond Exp results

Implementation

- Use final step's first half to reduce variance.
- $\Delta W^{(c)}_{N-1} = \Delta W^{(f)}_{2N-2} + \Delta W^{(f)}_{2N-1}$
- Tower prop: $\mathbb{E}\left[P(\hat{S}_N)|W\right] = \mathbb{E}\left[\mathbb{E}(P(\hat{S}_N)|W, \Delta W^{(f)}_{2N-2})|W\right]$.
- $\mathbb{E}(P(\hat{S}_N)|W, \mathrm{d}W^{(f)}_{2N-2}) = \Phi\left(\frac{\mu_W + \sigma_W/\sqrt{h_f}\cdot\Delta W^{(f)}_{2N-2} - K}{\sigma_W}\right)$.

Digital Call:

- Value's estimator: $\beta = 1.5 \;\Rightarrow\;$ Complexity $\mathcal{O}(\epsilon^{-2})$.
- Delta's estimator: $\beta = 0.5 \;\Rightarrow\;$ Complexity $\mathcal{O}(\epsilon^{-2.5})$.
- Vega's estimator: $\beta = 0.6 \;\Rightarrow\;$ Complexity $\mathcal{O}(\epsilon^{-2.4})$.

European Call:

- Value's estimator: $\beta = 2 \;\Rightarrow\;$ Complexity $\mathcal{O}(\epsilon^{-2})$.
- Delta's estimator: $\beta = 1.5 \;\Rightarrow\;$ Complexity $\mathcal{O}(\epsilon^{-2})$.
- Vega's estimator: $\beta = 2 \;\Rightarrow\;$ Complexity $\mathcal{O}(\epsilon^{-2})$.

Split Pathwise Sensitivities

Limitations of conditional expectations:

- Need an analytical expression for $\mathbb{E}(P(\hat{S}_N)|\hat{S}_{N-1})$ and its derivative.

Principle:

- Numerical approximation of a conditional expectation.
- Split each path, take d samples for the final increment.
- $\mathbb{E}(P(\hat{S}_N)|\hat{S}_{N-1}) \approx \frac{1}{d} \sum_{m=1}^{d} P(\hat{S}_{N-1}^{(m)} \cdot D_{N-1}^{(m)})$.
- $\frac{\partial \mathbb{E}(P(\hat{S}_N)|\hat{S}_{N-1})}{\partial \hat{S}_{N-1}} \approx \frac{1}{d} \sum_{m=1}^{d} \left(\frac{\partial P}{\partial \hat{S}_N} \frac{\partial \hat{S}_N}{\partial \hat{S}_{N-1}}\right)^{(m)}$.
- Use these expressions as before in Multilevel PwS with Conditional Expectations.

Warning:

- Conditions apply for Greeks.

Multilevel Split PwS results

European Call, $d = 1$ (Same as PwS) :

- Value's estimator: $\beta \approx 2.0 \Rightarrow$ Complexity $\mathcal{O}(\epsilon^{-2})$.
- Delta's estimator: $\beta \approx 0.9 \Rightarrow$ Complexity $\mathcal{O}(\epsilon^{-2.1})$.
- Vega's estimator: $\beta \approx 1.3 \Rightarrow$ Complexity $\mathcal{O}(\epsilon^{-2})$.

European Call, $d = 20$:

- Value's estimator: $\beta \approx 2.0 \Rightarrow$ Complexity $\mathcal{O}(\epsilon^{-2})$.
- Delta's estimator: $\beta \approx 1.1 \Rightarrow$ Complexity $\mathcal{O}(\epsilon^{-2})$.
- Vega's estimator: $\beta \approx 1.8 \Rightarrow$ Complexity $\mathcal{O}(\epsilon^{-2})$.

European Call, $d = 500$ (Similar to PwS with Cond. Expectations):

- Value's estimator: $\beta \approx 2.0 \Rightarrow$ Complexity $\mathcal{O}(\epsilon^{-2})$.
- Delta's estimator: $\beta \approx 1.5 \Rightarrow$ Complexity $\mathcal{O}(\epsilon^{-2})$.
- Vega's estimator: $\beta \approx 2.0 \Rightarrow$ Complexity $\mathcal{O}(\epsilon^{-2})$.

Vibrato Monte Carlo

Goal and Principle:

- Approximate Conditional Expectations for discontinuous P.
- $\hat{S}_N = \mu_W + \sigma_W Z$.
- Use PwS on W and LRM on last step Z.

Principle

- $V = \mathbb{E}_W \left[\mathbb{E}_Z (P(\hat{S}_N) | W) \right]$
- $\frac{\partial V}{\partial \theta} = \mathbb{E}_W \left[\frac{\partial}{\partial \theta} \mathbb{E}_Z (P(\hat{S}_N) | W) \right] = \mathbb{E}_W \left[\mathbb{E}_Z (P(\hat{S}_N) \frac{\partial (\log p_S)}{\partial \theta} | W) \right]$
- $\frac{\partial}{\partial \theta} (\log p_S) = \frac{\partial}{\partial \mu_W} (\log p_S) \cdot \frac{\partial \mu_W}{\partial \theta} + \frac{\partial}{\partial \sigma_W} (\log p_S) \cdot \frac{\partial \sigma_W}{\partial \theta}$
  - $\frac{\partial}{\partial \mu_W} (\log p_S)$, $\frac{\partial}{\partial \sigma_W} (\log p_S)$ simple functions of Z.
  - $\frac{\partial \mu_W}{\partial \theta}$, $\frac{\partial \sigma_W}{\partial \theta}$ known by PwS.

Vibrato Monte Carlo - Multilevel

Multilevel Implementation:

- Treatment similar to Split PwS.
- Final step at coarse level:
  - $\Delta W^{(c)}_{N-1} = \Delta W^{(f)}_{2N-2} + \Delta W^{(f)}_{2N-1}$
  - $V = \mathbb{E}_W \left[\mathbb{E} \left[P(\hat{S}_N) | W \right] \right]$
    $= \mathbb{E}_W \left[\mathbb{E} \left[\mathbb{E}(P(\hat{S}_N) | W, \Delta W^{(f)}_{2N-2}) | W \right] \right].$
  - Apply LRM to $\frac{\partial}{\partial \theta} \mathbb{E}(P(\hat{S}_N) | W, \Delta W^{(f)}_{2N-2})$.
  - $\frac{\partial V}{\partial \theta} =$
    $\mathbb{E}_W \left[\mathbb{E}_{\Delta W^{(f)}_{2N-2}} \left[\mathbb{E}_{\Delta W^{(f)}_{2N-1}} (P(\hat{S}_N) \frac{\partial (\log p_S)}{\partial \theta} | W, \Delta W^{(f)}_{2N-2}) | W \right] \right]$

Vibrato Monte Carlo results

European Call, $d = 100$:

- Value's estimator: $\beta \approx 2.0 \;\Rightarrow\;$ Complexity $\mathcal{O}(\epsilon^{-2})$.
- Delta's estimator: $\beta \approx 1.5 \;\Rightarrow\;$ Complexity $\mathcal{O}(\epsilon^{-2})$.
- Vega's estimator: $\beta \approx 2.0 \;\Rightarrow\;$ Complexity $\mathcal{O}(\epsilon^{-2})$.

Digital Call, $d = 100$:

- Value's estimator: $\beta \approx 1.2 \;\Rightarrow\;$ Complexity $\mathcal{O}(\epsilon^{-2})$.
- Delta's estimator: $\beta \approx 0.3 \;\Rightarrow\;$ Complexity $\mathcal{O}(\epsilon^{-2.7})$.
- Vega's estimator: $\beta \approx 0.5 \;\Rightarrow\;$ Complexity $\mathcal{O}(\epsilon^{-2.5})$.

Remarks:

- $d \gg 1$ is important to get good approximates of good approximates of Cond. Exp.
- $d = 500\ldots$ β similar to Cond. Exp. $(1.5, 0.5, 0.6)$

Conclusion

Final Words

MLMC provides computational savings for the computation of Greeks:

- Benefits dependent on β, convergence rate of $\mathbb{V}(\hat{Y}_l)$.
- Discontinuous/nonsmooth payoffs are challenging:
  - Special treatment (Cond. Exp., Vibrato).
  - Smaller β.

Unexpected result:

- ν converges slightly faster than Δ.
- Consistent feature across all simulations.

Research will now focus on a rigorous numerical analysis of MLMC for Greeks.