"Vibrato" Monte Carlo evaluation of Greeks
(Smoking Adjoints: part 3)

Mike Giles
mike.giles@maths.ox.ac.uk

Oxford University Mathematical Institute
Oxford-Man Institute of Quantitative Finance

MCQMC 2008, July 7-11
“Smoking Adjoint”

Paper with Paul Glasserman in *Risk* in 2006 showed how adjoints can be used in computing pathwise sensitivities – gives lots of first order sensitivities for negligible cost

This attracted a lot of interest, and questions:

- what is involved in practice in creating an adjoint code, and can it be simplified? (see HERCMA paper, available from website)
- do we really have to differentiate the payoff?
- what about discontinuous payoffs?
- what about American options? (not addressed yet!)
Outline

- different approaches to computing Greeks
  - finite differences
  - likelihood ratio method
  - pathwise sensitivity
- use of conditional expectation for a digital option
- “vibrato” extension for scalar SDE
- generalisation to multidimensional SDEs
Stochastic differential equation with general drift and volatility terms:

$$dS_t = a(S_t, t) \, dt + b(S_t, t) \, dW_t$$

For a simple European option we want to compute the expected discounted payoff value dependent on the terminal state:

$$V = \mathbb{E}[f(S_T)]$$

Note: the drift and volatility functions are almost always differentiable, but the payoff $f(S)$ is often not.
Euler discretisation with timestep $h$:

$$\hat{S}_{n+1} = \hat{S}_n + a(\hat{S}_n, t_n) \cdot h + b(\hat{S}_n, t_n) \cdot \Delta W_n$$

Simplest Monte Carlo estimator for expected payoff is an average of $M$ independent path simulations:

$$M^{-1} \sum_{i=1}^{M} f(\hat{S}^{(i)}_N)$$

Greeks: for hedging and risk management we also want to estimate derivatives of expected payoff $V$
Simple Problem

For Geometric Brownian motion

$$dS_t = r S_t \, dt + \sigma S_t \, dW_t$$

the SDE can be solved analytically to give

$$S_T = S_0 \exp \left( (r - \frac{1}{2} \sigma^2) T + \sigma W_T \right)$$

In this case, we can directly sample $W_T$ to get

$$V \equiv \mathbb{E} \left[ f(S_T) \right] \approx M^{-1} \sum_{i=1}^{M} f(S_T^{(i)})$$

– will use this to illustrate approaches to calculating sensitivities
Finite Differences

Simplest approach is to use a finite difference approximation,

\[ \frac{\partial V}{\partial \theta} \approx \frac{V(\theta + \Delta \theta) - V(\theta - \Delta \theta)}{2 \Delta \theta} \]

\[ \frac{\partial^2 V}{\partial \theta^2} \approx \frac{V(\theta + \Delta \theta) - 2V(\theta) + V(\theta - \Delta \theta)}{(\Delta \theta)^2} \]

– very simple, but expensive and inaccurate if \( \Delta \theta \) is too big, or too small in the case of discontinuous payoffs
Likelihood Ratio Method

For simple cases where we know the terminal probability distribution

\[ V \equiv \mathbb{E}[f(S_T)] = \int f(S) \, p_S(\theta; S) \, dS \]

we can differentiate this to get

\[ \frac{\partial V}{\partial \theta} = \int f \frac{\partial p_S}{\partial \theta} \, dS = \int f \frac{\partial (\log p_S)}{\partial \theta} \, p_S \, dS = \mathbb{E} \left[ f \frac{\partial (\log p_S)}{\partial \theta} \right] \]

This is the Likelihood Ratio Method (Broadie & Glasserman, 1996) – its great strength is that it can handle discontinuous payoffs

“Vibrato” Monte Carlo Greeks – p. 8/30
Likelihood Ratio Method

The LRM weakness is in its generalisation to full path simulations for which we get the multi-dimensional integral

\[ \hat{V} = \mathbb{E}[f(\hat{S})] = \int f(\hat{S}) p(\hat{S}) \, d\hat{S}, \]

where

\[ d\hat{S} \equiv d\hat{S}_1 \, d\hat{S}_2 \, d\hat{S}_3 \, \ldots \, d\hat{S}_N \]

and the joint probability density function \( p(\hat{S}) \) is the product of the p.d.f.s for each timestep

\[
p(\hat{S}) = \prod_{n} p_n(\hat{S}_{n+1} | \hat{S}_n)
\]

\[
\log p(\hat{S}) = \sum_{n} \log p_n(\hat{S}_{n+1} | \hat{S}_n)
\]
Likelihood Ratio Method

When computing Vega from an Euler discretisation of Geometric Brownian motion this leads to

\[
\frac{\partial \hat{V}}{\partial \sigma} = \mathbb{E} \left[ \left( \sum_n \frac{Z_n^2 - 1}{\sigma} \right) f(\hat{S}_N) \right]
\]

where \( Z_n \) is the unit Normal used in the \( n^{th} \) timestep

\[
\hat{S}_{n+1} = \hat{S}_n (1 + r \ h) + \sigma \hat{S}_n \sqrt{h} \ Z_n
\]

Since \( \mathbb{V}[Z_n^2 - 1] = 2 \) it follows that the variance of the estimator is \( O(h^{-1}) \)

This blow-up as \( h \to 0 \) is the weakness of the LRM.
Pathwise sensitivities

Alternatively, for simple Geometric Brownian Motion

\[ V \equiv \mathbb{E} [f(S_T)] = \int f(S_T(\theta; W)) \ p_W(W) \ dW \]

and differentiating this gives

\[ \frac{\partial V}{\partial \theta} = \int \frac{\partial f}{\partial S} \frac{\partial S_T}{\partial \theta} \ p_W(W) \ dW = \mathbb{E} \left[ \frac{\partial f}{\partial S} \frac{\partial S_T}{\partial \theta} \right] \]

with \( \frac{\partial S_T}{\partial \theta} \) being evaluated at fixed \( W \).

This is the pathwise sensitivity approach – it can’t handle discontinuous payoffs, but generalises well to full path simulations
Pathwise sensitivities

The generalisation involves differentiating the Euler path discretisation,

\[ \hat{S}_{n+1} = \hat{S}_n + a(\hat{S}_n, t_n) \, h + b(\hat{S}_n, t_n) \, \Delta W_n \]

holding fixed the Brownian increments, to get

\[ \frac{\partial \hat{S}_{n+1}}{\partial \theta} = \left( 1 + \frac{\partial a}{\partial S} \, h + \frac{\partial b}{\partial S} \, \Delta W_n \right) \frac{\partial \hat{S}_n}{\partial \theta} + \frac{\partial a}{\partial \theta} \, h + \frac{\partial b}{\partial \theta} \, \Delta W_n \]

leading to

\[ \frac{\partial \hat{V}}{\partial \theta} = \mathbb{E} \left[ \frac{\partial f}{\partial S}(\hat{S}_N) \frac{\partial \hat{S}_N}{\partial \theta} \right]. \]
Pathwise sensitivities

In the case of Vega for an Euler discretisation of GBM

\[ \hat{S}_{n+1} = \hat{S}_n + r \hat{S}_n h + \sigma \hat{S}_n \Delta W_n \]

we get

\[ \frac{\partial \hat{S}_{n+1}}{\partial \sigma} = \left(1 + r h + \sigma \Delta W_n \right) \frac{\partial \hat{S}_n}{\partial \sigma} + \hat{S}_n \Delta W_n \]

and the variance

\[ \nabla \left[ \frac{\partial f}{\partial \hat{S}}(\hat{S}_N) \frac{\partial \hat{S}_N}{\partial \sigma} \right] \]

is $O(1)$ if $f(S)$ is Lipschitz.
Vibrato Monte Carlo

What is best if payoff is discontinuous?

- LRM
  - estimator variance $O(h^{-1})$
- Malliavin calculus
  - estimator variance $O(1)$
  - recent paper by Glasserman & Chen shows it can be viewed as a pathwise/LRM hybrid
  - might be good choice when few Greeks needed
- new “vibrato” Monte Carlo idea
  - also a pathwise/LRM hybrid
  - estimator variance $O(h^{-1/2})$
  - efficient adjoint implementation

“Vibrato” Monte Carlo Greeks – p. 14/30
new idea is based on use of conditional expectation for a simple digital option in Paul Glasserman’s book

output of each SDE path calculation becomes a narrow (multivariate) Normal distribution

combine pathwise sensitivity for the differentiable SDE, with LRM for the discontinuous payoff

avoiding the differentiation of the payoff also simplifies the implementation in real-world setting
Final timestep of Euler path discretisation is

\[ \hat{S}_N = \hat{S}_{N-1} + a(\hat{S}_{N-1}, t_{N-1}) \, h + b(\hat{S}_{N-1}, t_{N-1}) \, \Delta W_{N-1} \]

Instead of using random number generator to get a value for \( \Delta W_{N-1} \), consider the whole distribution of possible values, so \( \hat{S}_N \) has a Normal distribution with mean

\[ \mu_W = \hat{S}_{N-1} + a(\hat{S}_{N-1}, t_{N-1}) \, h \]

and standard deviation

\[ \sigma_W = b(\hat{S}_{N-1}, t_{N-1}) \, \sqrt{h} \]

where \( W \equiv (\Delta W_0, \Delta W_1, \ldots \Delta W_{N-2}) \).
For a particular path given by a particular vector $W$, the expected payoff is

$$\mathbb{E}_Z[f(\mu_W + \sigma_W Z)]$$

where $Z$ is a unit Normal random variable.

Averaging over all $W$ then gives the same overall expectation as before.

Note also that, for given $W$, $\hat{S}_N$ has a Normal distribution

$$p_S(\hat{S}) = \frac{1}{\sqrt{2\pi} \sigma_W} \exp\left(-\frac{(\hat{S} - \mu_W)^2}{2 \sigma_W^2}\right)$$
In the case of a simple digital call with strike $K$, the analytic solution is

$$\mathbb{E}_Z[f(\mu_W + \sigma_W Z)] = \exp(-rT) \Phi \left( \frac{\mu_W - K}{\sigma_W} \right).$$

- for each $W$, the payoff is now smooth, differentiable
- derivative is $O(h^{-1/2})$ near strike, near zero elsewhere
  - variance is $O(h^{-1/2})$
- analytic evaluation of conditional expectation not possible in general for multivariate cases
  - use Monte Carlo estimation!
Main novelty comes in calculating the sensitivity.

For a particular $W$, we have a Normal probability distribution for $\hat{S}_N$ and can apply the Likelihood Ratio method to get

$$\frac{\partial}{\partial \theta} \mathbb{E}_Z \left[ f(\hat{S}_N) \right] = \mathbb{E}_Z \left[ f(\hat{S}_N) \frac{\partial (\log p_S)}{\partial \theta} \right],$$

where

$$\frac{\partial (\log p_S)}{\partial \theta} = \frac{\partial (\log p_S)}{\partial \mu_W} \frac{\partial \mu_W}{\partial \theta} + \frac{\partial (\log p_S)}{\partial \sigma_W} \frac{\partial \sigma_W}{\partial \theta}$$

Averaging over all $W$ then gives the expected sensitivity.
To improve the variance, we note that

\[
\mathbb{E}_Z \left[ f(\mu_W + \sigma_W Z) \, Z \right] = \mathbb{E}_Z \left[ -f(\mu_W - \sigma_W Z) \, Z \right] = \mathbb{E}_Z \left[ \left( f(\mu_W + \sigma_W Z) - f(\mu_W - \sigma_W Z) \right) \, Z \right] = \frac{1}{2} \mathbb{E}_Z \left[ \left( f(\mu_W + \sigma_W Z) - 2f(\mu_W - f(\mu_W - \sigma_W Z)) \right) \, (Z^2 - 1) \right]
\]

This gives an estimator with \( O(1) \) variance when \( f(S) \) is Lipschitz, and \( O(h^{-1/2}) \) variance when it is discontinuous.
Vibrato Monte Carlo

Test case: Geometric Brownian motion

\[ dS_t = r \, S_t \, dt + \sigma \, S_t \, dW_t \]

with simple digital call option.

Parameters: \( r = 0.05, \ \sigma = 0.2, \ T = 1, \ S_0 = 100, \ K = 100 \)

Numerical results compare:

- LRM
- vibrato with one \( Z \) per \( W \)
- pathwise with conditional expectation
Vibrato Monte Carlo

![Graph showing variance (Var) vs. h for LRM, vibrato, and pathwise methods.]

“Vibrato” Monte Carlo Greeks – p. 22/30
These results used just one $Z$ per path. If $M_Z$ are used, the variance is

$$V_W \left[ \mathbb{E}_Z [g(W, Z)] \right] + M_Z^{-1} \mathbb{E}_W \left[ \nabla_Z g(W, Z) \right]$$

where $g(W, Z)$ is the estimator.

The limit $M_Z \to \infty$ gives the variance for the estimator based on the analytic conditional expectation.

The optimal $M_Z$ can be determined if one knows/estimates $V_W \left[ \mathbb{E}_Z [g(W, Z)] \right]$ and $\mathbb{E}_W \left[ \nabla_Z g(W, Z) \right]$, and the relative cost of the path simulation and the payoff evaluation.
Multivariate extension

In general we have

\[ \hat{S}(W, Z) = \mu_W + C_W Z \]

where \( \Sigma_W = C_W C_W^T \) is the covariance matrix, and \( Z \) is a vector of uncorrelated Normals. The joint p.d.f. is

\[ \log p_S = -\frac{1}{2} \log |\Sigma_W| - \frac{1}{2} (\hat{S} - \mu_W)^T \Sigma_W^{-1} (\hat{S} - \mu_W) - \frac{1}{2} d \log(2\pi) \]

and so

\[ \frac{\partial \log p_S}{\partial \mu_W} = C_W^{-T} Z, \]

\[ \frac{\partial \log p_S}{\partial \Sigma_W} = \frac{1}{2} C_W^{-T} \left( ZZ^T - I \right) C_W^{-1} \]
Multivariate extension

This leads to

$$\frac{\partial}{\partial \theta} \mathbb{E}_Z \left[ f(\hat{S}) \right] = \mathbb{E}_Z \left[ f(\hat{S}) \frac{\partial(\log p_S)}{\partial \theta} \right]$$

where

$$\frac{\partial(\log p_S)}{\partial \theta} = \left( \frac{\partial \log p_S}{\partial \mu_W} \right)^T \frac{\partial \mu_W}{\partial \theta} + \text{tr} \left( \frac{\partial \log p_S}{\partial \Sigma_W} \frac{\partial \Sigma_W}{\partial \theta} \right)$$

and $\frac{\partial \mu_W}{\partial \theta}, \frac{\partial \Sigma_W}{\partial \theta}$ come from pathwise sensitivity analysis.

A more efficient estimator can be obtained by similar reasoning to the scalar case.
Test case: Geometric Brownian motion

\[
\begin{align*}
\frac{dS_t^{(1)}}{S_t^{(1)}} &= r S_t^{(1)} \, dt + \sigma^{(1)} S_t^{(1)} \, dW_t^{(1)} \\
\frac{dS_t^{(1)}}{S_t^{(1)}} &= r S_t^{(2)} \, dt + \sigma^{(2)} S_t^{(2)} \, dW_t^{(2)}
\end{align*}
\]

with a simple digital call option based solely on \( S_T^{(1)} \).

Parameters: \( r = 0.05, \, \sigma^{(1)} = 0.2, \, \sigma^{(2)} = 0.3, \, T = 1, \, S_0^{(1)} = S_0^{(2)} = 100, \, K = 100, \, \rho = 0.5 \)

Numerical results again compare LRM, vibrato with one \( Z \) per \( W \), and pathwise with conditional expectation.
Vibrato Monte Carlo

![Graph showing the comparison of LRM, vibrato, and pathwise methods for different values of h.](image)

"Vibrato" Monte Carlo Greeks – p. 27/30
Multivariate extension

Can also treat payoffs dependent on \( S(\tau) \) at intermediate times, by taking
\[
t_n < \tau < t_{n+1}
\]
and using simple Brownian motion interpolation between \( \hat{S}_n \) and \( \hat{S}_{n+1} \) to get a Normal distribution for \( \hat{S}(\tau) \), with

mean:
\[
\hat{S}_n + \frac{\tau-t_n}{t_{n+1}-t_n} \left( \hat{S}_{n+1} - \hat{S}_n \right)
\]

variance:
\[
\frac{(\tau-t_n)(t_{n+1}-\tau)}{t_{n+1}-t_n} b^2(\hat{S}_n, t_n)
\]
Conclusions

“Vibrato” idea for computing Greeks offers

- $O(1)$ variance for Lipschitz payoffs, and easy implementation – no derivatives required
- $O(h^{-1/2})$ variance for discontinuous payoffs
- adjoint implementation for multiple Greeks

Future work:

- similar idea for digital options in multilevel Monte Carlo path simulation – introduces Radon-Nikodym derivative from change in measure
Acknowledgements

- Paul Glasserman for collaboration on adjoint technique and discussions on vibrato Monte Carlo
- Funding from Microsoft, EPSRC and Man Investments

Further information

- www.maths.ox.ac.uk/~gilesm/
- Email: mike.giles@maths.ox.ac.uk