Numerical Methods II

Prof. Mike Giles

mike.giles@maths.ox.ac.uk

Oxford University Mathematical Institute

Greeks

In Monte Carlo applications we don't just want to know the expected discounted value of some payoff

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

We also want to know a whole range of "Greeks" corresponding to first and second derivatives of V with respect to various parameters:

$$\Delta = \frac{\partial V}{\partial S_0}, \qquad \Gamma = \frac{\partial^2 V}{\partial S_0^2},$$

$$\rho = \frac{\partial V}{\partial r}, \qquad \text{Vega} = \frac{\partial V}{\partial \sigma}.$$

Greeks

The Greeks are needed for hedging and risk analysis.

Whereas prices can be obtained to some extent from market prices, simulation is the only way to determine the Greeks.

In the next two lectures we will explore 3 approaches:

- finite differences
- likelihood ratio method (lecture 8)
- pathwise sensitivities (lecture 8)

Finite difference sensitivities

If $V(\theta) = \mathbb{E}[f(S(T))]$ for an input parameter θ is sufficiently differentiable, then the sensitivity $\dfrac{\partial V}{\partial \theta}$ can be approximated by one-sided finite difference

$$\frac{\partial V}{\partial \theta} = \frac{V(\theta+\Delta\theta) - V(\theta)}{\Delta\theta} \ + \ O(\Delta\theta)$$

or by central finite difference

$$\frac{\partial V}{\partial \theta} = \frac{V(\theta+\Delta\theta) - V(\theta-\Delta\theta)}{2\Delta\theta} \ + \ O((\Delta\theta)^2)$$

(This approach is referred to as getting Greeks by "bumping" the input parameters.)

Finite difference sensitivities

The clear advantage of this approach is that it is very simple to implement (hence the most popular in practice?)

However, the disadvantages are:

- expensive (2 extra sets of calculations for central differences)
- significant bias error if $\Delta\theta$ too large
- machine roundoff errors if $\Delta\theta$ too small
- large variance if $f(S(T))$ discontinuous and $\Delta\theta$ small

Finite difference sensitivities

Let $X^{(i)}(\theta{+}\Delta\theta)$ and $X^{(i)}(\theta{-}\Delta\theta)$ be the values of $f(S(T))$ obtained for different MC samples, so the central difference estimate for $\dfrac{\partial V}{\partial \theta}$ is given by

$$
\begin{aligned}
\widehat{Y} &= \frac{1}{2\Delta\theta}\left(N^{-1}\sum_{i=1}^{N} X^{(i)}(\theta{+}\Delta\theta) - N^{-1}\sum_{i=1}^{N} X^{(i)}(\theta{-}\Delta\theta) \right) \\
&= \frac{1}{2N\Delta\theta}\sum_{i=1}^{N} \left(X^{(i)}(\theta{+}\Delta\theta) - X^{(i)}(\theta{-}\Delta\theta) \right)
\end{aligned}
$$

Finite difference sensitivities

If independent samples are taken for both $X^{(i)}(\theta+\Delta\theta)$ and $X^{(i)}(\theta-\Delta\theta)$ then

$$
\begin{aligned}
\mathbb{V}[\widehat{Y}] &\approx \left(\frac{1}{2N\Delta\theta}\right)^2 \sum_j \Big(\mathbb{V}[X(\theta+\Delta\theta)] + \mathbb{V}[X(\theta-\Delta\theta)]\Big) \\
&\approx \left(\frac{1}{2N\Delta\theta}\right)^2 2\,N\,\mathbb{V}[f] \\
&= \frac{\mathbb{V}[f]}{2N(\Delta\theta)^2}
\end{aligned}
$$

which is very large for $\Delta\theta \ll 1$.

Finite difference sensitivities

It is much better for $X^{(i)}(\theta+\Delta\theta)$ and $X^{(i)}(\theta-\Delta\theta)$ to use the same set of random inputs.

If $X^{(i)}(\theta)$ is differentiable with respect to θ, then

$$X^{(i)}(\theta+\Delta\theta) - X^{(i)}(\theta-\Delta\theta) \approx 2\,\Delta\theta\,\frac{\partial X^{(i)}}{\partial\theta}$$

and hence

$$\mathbb{V}[\widehat{Y}] \approx N^{-1}\,\mathbb{V}\left[\frac{\partial X}{\partial\theta}\right],$$

which behaves well for $\Delta\theta \ll 1$, so one should choose a small (but not ridiculously small) value for $\Delta\theta$ to minimise the bias due to the finite differencing.

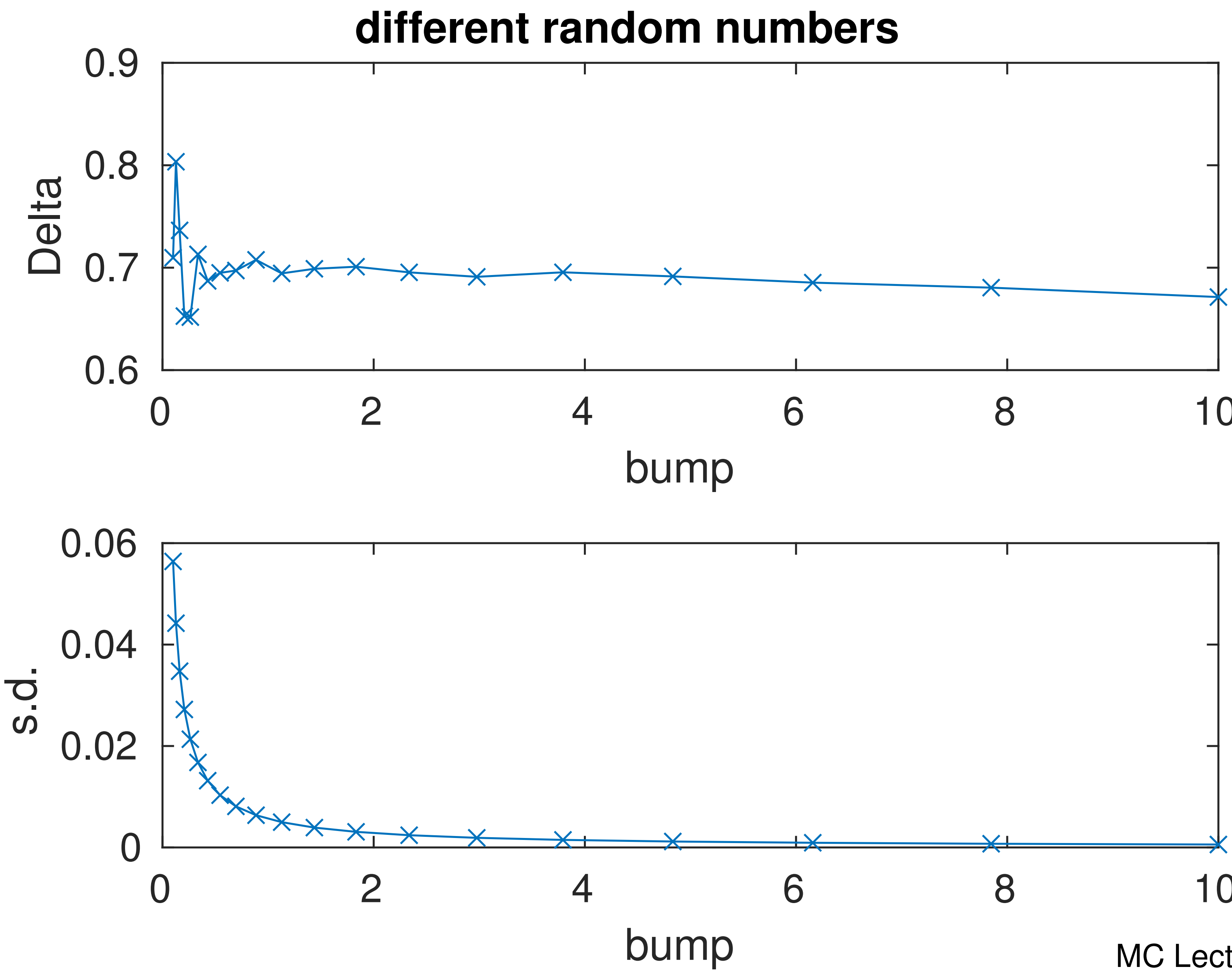
Basket call option

- 5 underlying assets starting at $S_0 = 100$, with call option on arithmetic mean with strike $K = 100$
- Geometric Brownian Motion model, $r = 0.05, T = 1$
- volatility $\sigma = 0.2$ and correlation matrix

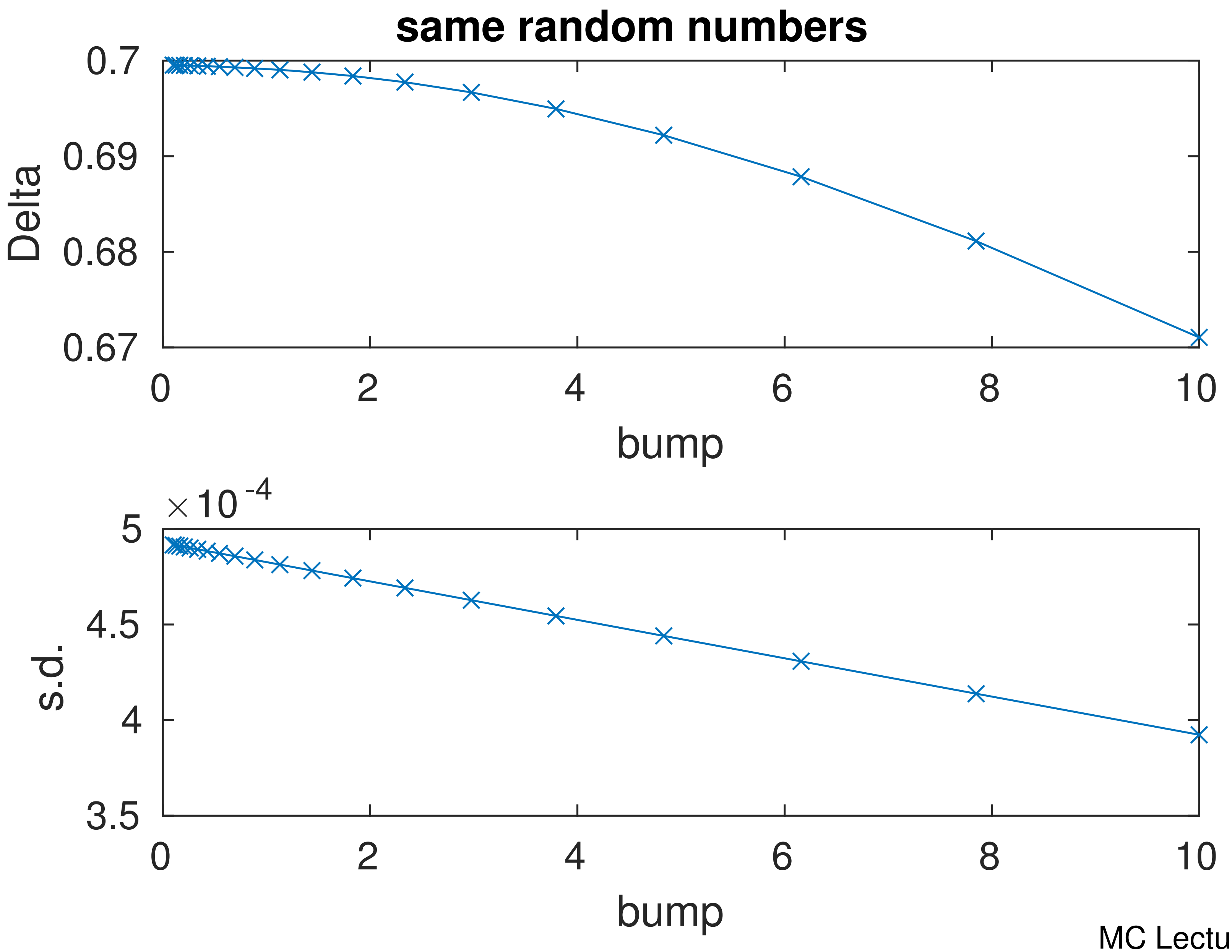
$$\Omega = \begin{pmatrix} 1 & 0.1 & 0.1 & 0.1 & 0.1 \\ 0.1 & 1 & 0.1 & 0.1 & 0.1 \\ 0.1 & 0.1 & 1 & 0.1 & 0.1 \\ 0.1 & 0.1 & 0.1 & 1 & 0.1 \\ 0.1 & 0.1 & 0.1 & 0.1 & 1 \end{pmatrix}$$

The aim is to estimate $\dfrac{\partial}{\partial S_0}\mathbb{E}[\exp(-rT)\, f(S_T)]$ where $f(S_T)$ is the basket call option payoff, using central differences.

Basket call option



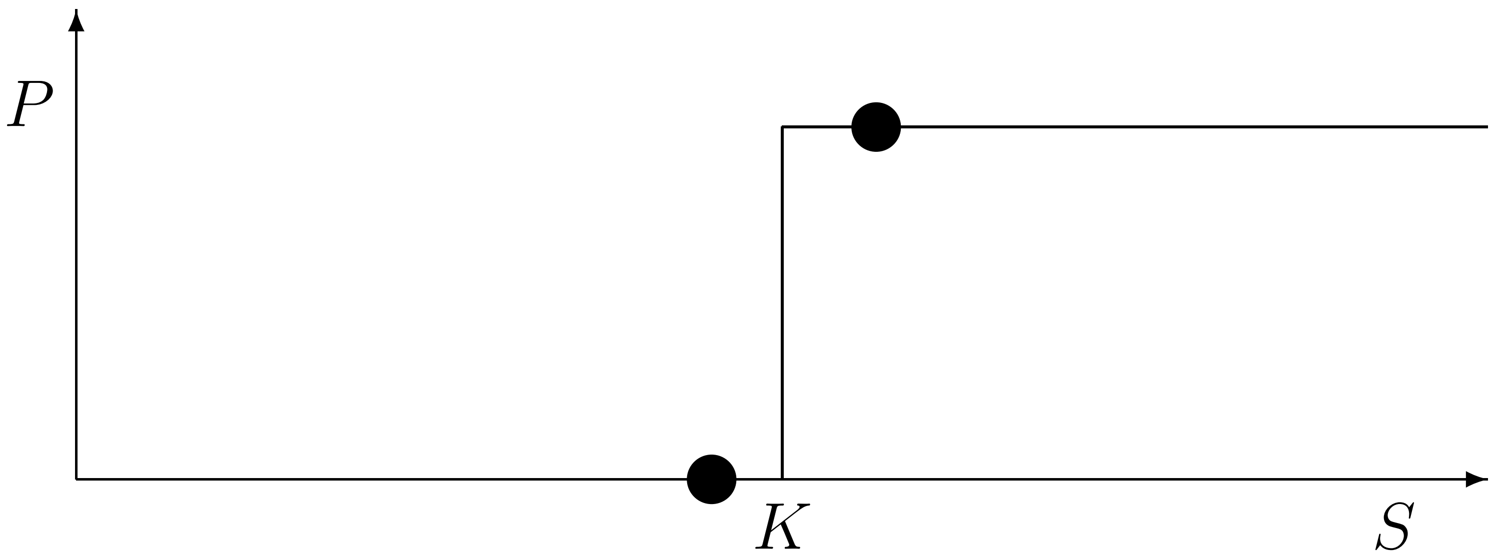
Basket call option



Finite difference sensitivities

Next, we analyse the variance of the finite difference estimator when the payoff is discontinuous.

The problem is that a small bump in the asset S can produce a big bump in the payoff – not differentiable



Finite difference sensitivities

What is the probability that $S(\theta \pm \Delta\theta)$ will be on different sides of the discontinuity?

Separation of $S(\theta \pm \Delta\theta)$ is $O(\Delta\theta)$

$\mathbb{P}(\,|S(\theta) - K| < c\,\Delta\theta) = O(\Delta\theta)$

Hence, $O(\Delta\theta)$ probability of straddling the strike.

Finite difference sensitivities

If we are interested in $\mathbb{E}[f(\omega)]$, and samples ω are either in set A, or its complement A^c, then

$$\begin{aligned} \mathbb{E}[f(\omega)] &= \mathbb{E}[f(\omega)\mathbf{1}_A] + \mathbb{E}[f(\omega)\mathbf{1}_{A^c}] \\ &= \mathbb{P}(\omega \in A)\ \mathbb{E}[f(\omega) \mid \omega \in A] \\ &+ \mathbb{P}(\omega \notin A)\ \mathbb{E}[f(\omega) \mid \omega \notin A] \end{aligned}$$

and similarly

$$\begin{aligned} \mathbb{E}[f^2(\omega)] &= \mathbb{P}(\omega \in A)\ \mathbb{E}[f^2(\omega) \mid \omega \in A] \\ &+ \mathbb{P}(\omega \notin A)\ \mathbb{E}[f^2(\omega) \mid \omega \notin A] \end{aligned}$$

Finite difference sensitivities

In this case of a discontinuous payoff

- For most samples, $X^{(i)}(\theta{+}\Delta\theta) - X^{(i)}(\theta{-}\Delta\theta) = O(\Delta\theta)$
- For an $O(\Delta\theta)$ fraction, $X^{(i)}(\theta{+}\Delta\theta) - X^{(i)}(\theta{-}\Delta\theta) = O(1)$

$$
\begin{aligned}
\Longrightarrow \quad \mathbb{E}\left[\frac{X^{(i)}(\theta{+}\Delta\theta) - X^{(i)}(\theta{-}\Delta\theta)}{2\Delta\theta}\right] &= O(1) \\
\mathbb{E}\left[\left(\frac{X^{(i)}(\theta{+}\Delta\theta) - X^{(i)}(\theta{-}\Delta\theta)}{2\Delta\theta}\right)^2\right] &= O(\Delta\theta^{-1})
\end{aligned}
$$

This gives $\mathbb{E}[\widehat{Y}] = O(1)$, but $\mathbb{V}[\widehat{Y}] = O(N^{-1}\Delta\theta^{-1})$.

Finite difference sensitivities

So, small $\Delta\theta$ gives a large variance, while a large $\Delta\theta$ gives a large finite difference discretisation error.

To determine the optimum choice we use the following result: if $\widehat{Y}$ is an estimator for $\mathbb{E}[Y]$ then

$$
\begin{aligned}
\mathbb{E}\left[\left(\widehat{Y} - \mathbb{E}[Y]\right)^2\right] &= \mathbb{E}\left[\left(\widehat{Y}-\mathbb{E}[\widehat{Y}] \ + \ \mathbb{E}[\widehat{Y}]-\mathbb{E}[Y]\right)^2\right] \\
&= \mathbb{E}\left[(\widehat{Y}-\mathbb{E}[\widehat{Y}])^2\right] + (\mathbb{E}[\widehat{Y}]-\mathbb{E}[Y])^2 \\
&= \mathbb{V}[\widehat{Y}] + \left(\mathbb{E}[\widehat{Y}]-\mathbb{E}[Y]\right)^2
\end{aligned}
$$

$$
\text{Mean Square Error} \;=\; \text{variance} + (\text{bias})^2
$$

Finite difference sensitivities

In our case, the MSE (mean-square-error) is

$$\mathbb{V}[\widehat{Y}] + \text{bias}^2 \sim \frac{a}{N\,\Delta\theta} + b\,\Delta\theta^4.$$

This is minimised by choosing $\Delta\theta \propto N^{-1/5}$, giving

$$\sqrt{\text{MSE}} \;\propto\; N^{-2/5}$$

in contrast to the usual MC result in which

$$\sqrt{\text{MSE}} \;\propto\; N^{-1/2}$$

Finite difference sensitivities

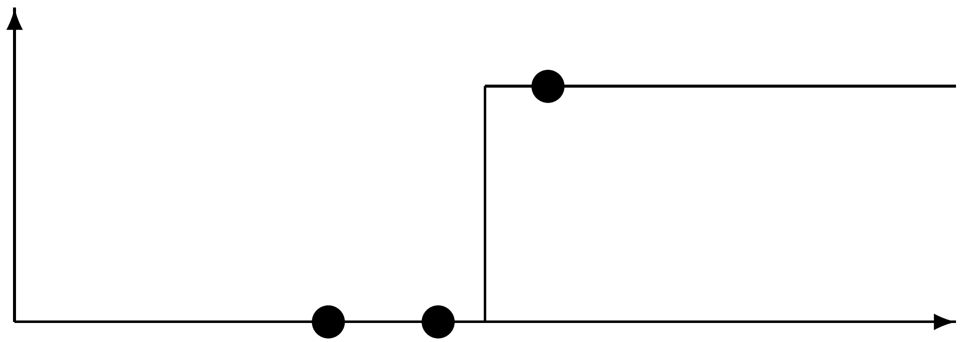
Second derivatives such as Γ can also be approximated by central differences:

$$\frac{\partial^2 V}{\partial \theta^2} = \frac{V(\theta+\Delta\theta) - 2\,V(\theta) + V(\theta-\Delta\theta)}{\Delta\theta^2} \; + \; O(\Delta\theta^2)$$

This will again have a larger variance if either the payoff or its derivative is discontinuous.

Finite difference sensitivities

Discontinuous payoff:



For an $O(\Delta\theta)$ fraction of samples

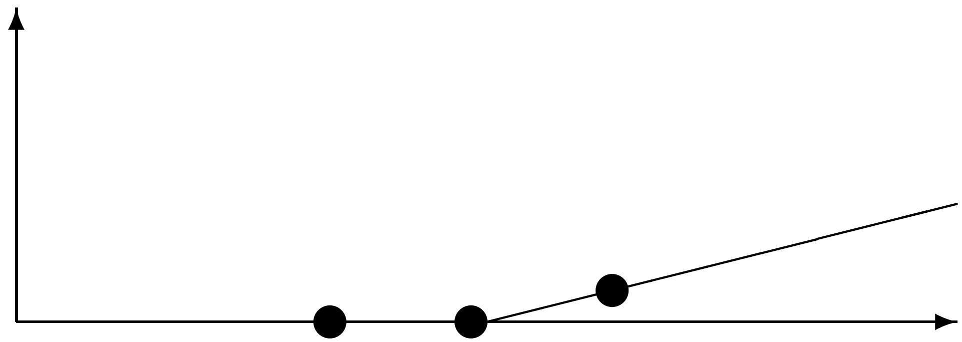
$$X^{(i)}(\theta{+}\Delta\theta) - 2X^{(i)}(\theta) + X^{(i)}(\theta{-}\Delta\theta) = O(1)$$

$$\Longrightarrow \quad \mathbb{E}\left[\left(\frac{X^{(i)}(\theta{+}\Delta\theta) - 2X^{(i)}(\theta) + X^{(i)}(\theta{-}\Delta\theta)}{\Delta\theta^2}\right)^2\right] = O(\Delta\theta^{-3})$$

This gives $\mathbb{V}[\widehat{Y}] = O(N^{-1}\Delta\theta^{-3})$.

Finite difference sensitivities

Discontinuous derivative:



For an $O(\Delta\theta)$ fraction of samples

$$X^{(i)}(\theta+\Delta\theta) - 2X^{(i)}(\theta) + X^{(i)}(\theta-\Delta\theta) = O(\Delta\theta)$$

$$\Longrightarrow \quad \mathbb{E}\left[\left(\frac{X^{(i)}(\theta+\Delta\theta) - 2X^{(i)}(\theta) + X^{(i)}(\theta-\Delta\theta)}{\Delta\theta^2}\right)^2\right] = O(\Delta\theta^{-1})$$

This gives $\mathbb{V}[\widehat{Y}] = O(N^{-1}\Delta\theta^{-1})$.

Finite difference sensitivities

Hence, for second derivatives the variance of the finite difference estimator is

- $O(N^{-1})$ if the payoff is twice differentiable
- $O(N^{-1}\Delta\theta^{-1})$ if the payoff has a discontinuous derivative
- $O(N^{-1}\Delta\theta^{-3})$ if the payoff is discontinuous

These can be used to determine the optimum $\Delta\theta$ in each case to minimise the Mean Square Error.

Final Words

- estimating Greeks is an important task, often more important than estimating the prices
- finite differences are simplest approach, but least accurate and most expensive
- always use the same random numbers for both sets of simulations
- in some cases, the optimum step size is a tradeoff between variance and bias (due to finite difference discretisation error)