Monte Carlo approximation of Greeks

Quantitative Finance

Tilburg University

These slides are partly based on earlier versions by Nikolaus Schweizer

Ramon van den Akker

the Greeks

Previous lecture we met the Greeks:

delta	$\frac{\partial C}{\partial S}$	gamma	$\frac{\partial^2 C}{\partial S^2}$
vega	$\frac{\partial C}{\partial \sigma}$	vomma	$\frac{\partial^2 C}{\partial \sigma^2}$
rho	$\frac{\partial C}{\partial r}$	vanna	$\frac{\partial^2 C}{\partial S \partial \sigma}$
theta	$\frac{\partial C}{\partial t}$		

Monte Carlo computation of Greeks

- we have seen that Greeks are important to risk measurement and risk management (hedging)
- for standard options in the Black-Scholes market, closed-form expressions for the Greeks are available
  - for example, the Delta of European call option is given by $\Phi(d_1)$
- this lecture: how to compute an approximation, using Monte Carlo simulation, to the Greeks if we cannot obtain a closed-form solution?
- we discuss three methods:
  - bump and reprice
  - pathwise method
  - likelihood ratio method

Section 1

Bump and Reprice

Goal:

Compute, using Monte Carlo techniques, approximation to:

$$\frac{d}{d\theta}\mathbb{E}_{\theta,\eta}[h(X;\theta,\gamma)]$$

where

- θ, η, and γ are (deterministic) parameters
- X is random variable whose distribution might depend on (θ,η)
- h is a real-valued function of X and potentially the parameters (θ,γ)
- if X has density (pdf) $f(\cdot;\theta,\eta)$:

$$\mathbb{E}_{\theta,\eta}[h(X;\theta,\gamma)] = \int h(u;\theta,\gamma)f(u;\theta,\eta)\,\mathrm{d}u$$

Example (vega call option)

- assume Black-Scholes market
- vega of European call option:

$$\frac{\partial}{\partial \sigma} F(t, S_t; r, \sigma, T, K)$$

where $F(t, S_t; r, \sigma, T, K)$ is price, at time t, of European option with payoff $\max\{S_T - K, 0\}$ at maturity T

- by FFT:

$$\begin{aligned} F(t, S_t; r, \sigma, T, K) &= \mathrm{e}^{-r(T-t)} \, \mathbb{E}_{\mathbb{Q}} \left[C_T | \mathcal{F}_t \right] \\ &= \mathrm{e}^{-r(T-t)} \, \mathbb{E}_{\mathbb{Q}} \left[\max\{S_T - K, 0\} | \mathcal{F}_t \right] \end{aligned}$$

- remark: from $\mathrm{d}S_t = rS_t \, \mathrm{d}t + \sigma S_t \, \mathrm{d}W_t^{\mathbb{Q}}$ we see that conditional pdf of S_T, under $\mathbb{Q}$, (indeed) depends on σ

Core idea

- we want to obtain MC approximation to

$$\frac{d}{d\theta}\mathbb{E}_{\theta,\eta}[h(X;\theta,\gamma)]$$

- please recall, for $f : \mathbb{R} \to \mathbb{R}$,

$$f'(x) = \lim_{\epsilon \to 0} \frac{f(x+\epsilon) - f(x)}{\epsilon} \approx \frac{f(x+h) - f(x)}{h}$$

  for small number h

- this suggests to approximate $(d/d\theta)\mathbb{E}_{\theta,\eta}[h(X;\theta,\gamma)]$ by *one-sided finite-difference estimate*:

$$\frac{\mathbb{E}_{\theta+h,\eta}[h(X;\theta+h,\gamma)] - \mathbb{E}_{\theta,\eta}[h(X;\theta,\gamma)]}{h}$$

  where h is a small step

- however, we cannot (or do not want to) compute $\mathbb{E}_{\theta+h,\eta}[h(X;\theta+h,\gamma)]$ and $\mathbb{E}_{\theta,\eta}[h(X;\theta,\gamma)]$ analytically

Core idea

- if we simulate $X_1, \ldots, X_n$ i.i.d. from same distribution as $X \sim P_{\theta,\eta}$, then we have (LLN), for "large" n,

$$f(\theta) := \mathbb{E}_{\theta,\eta}[h(X;\theta,\gamma)] \approx \frac{1}{n}\sum_{i=1}^{n} h(X_i;\theta,\gamma) =: \hat{f}_n(\theta)$$

- to stress that we simulated from $P_{\theta,\eta}$ we will write $X_i = X_i^{\theta,\eta}$
- please note:

$$\mathbb{E}_{\theta,\eta}[\hat{f}_n(\theta)] = f(\theta),$$

$$\mathrm{var}_{\theta,\eta}[\hat{f}_n(\theta)] = \frac{\mathrm{var}_{\theta,\eta}\left[h\left(X_1^{\theta,\eta}\right)\right]}{n}.$$

One-sided estimate

Combining the observations on the previous slide leads to **bump and reprice** method a.k.a. **finite-difference** method:

- choose small number h
- simulate $X_1^{\theta,\eta},\dots,X_n^{\theta,\eta}$ i.i.d. from $P_{\theta,\eta}$
- simulate $X_1^{\theta+h,\eta},\dots,X_n^{\theta+h,\eta}$ i.i.d. from $P_{\theta+h,\eta}$
- approximate
$$\frac{d}{d\theta}\mathbb{E}_{\theta,\eta}[h(X;\theta,\gamma)]$$
by *one-sided finite-difference estimate*:
$$\frac{\frac{1}{n}\sum_{i=1}^n h(X_i^{\theta+h,\eta};\theta+h,\gamma)-\frac{1}{n}\sum_{i=1}^n h(X_i^{\theta,\eta};\theta,\gamma)}{h}$$
where h is a small step

One-sided estimate - quality

- denote

$$f(\theta) = \mathbb{E}_{\theta,\eta}[h(X;\theta,\gamma)] \text{ and } \hat{f}_n(\theta) = \frac{1}{n}\sum_{i=1}^{n} h(X_i^{\theta,\eta};\theta,\gamma)$$

- we approximate $f'(\theta)$ by one-sided f.d.

$$\hat{f}_n'(\theta) = \frac{\hat{f}_n(\theta+h) - \hat{f}_n(\theta)}{h}$$

- estimation error $\hat{f}_n'(\theta) - f'(\theta)$ can be written as:

$$\begin{aligned}\hat{f}_n'(\theta) - f'(\theta) = &\left[\frac{f(\theta+h) - f(\theta)}{h} - f'(\theta)\right] \\ &+ \left[\hat{f}_n'(\theta) - \frac{f(\theta+h) - f(\theta)}{h}\right]\end{aligned}$$

One-sided estimate - quality

- from

$$\hat{f}_n'(\theta) - f'(\theta) = \left[\frac{f(\theta+h) - f(\theta)}{h} - f'(\theta) \right] + \left[\hat{f}_n'(\theta) - \frac{f(\theta+h) - f(\theta)}{h} \right],$$

with

$$\hat{f}_n'(\theta) = \frac{\hat{f}_n(\theta+h) - \hat{f}_n(\theta)}{h},$$

we obtain

$$\mathrm{MSE} = \left(\frac{f(\theta+h) - f(\theta)}{h} - f'(\theta) \right)^2 + \frac{1}{h^2} \operatorname{var}\left(\hat{f}_n(\theta+h) - \hat{f}_n(\theta) \right)$$

One-sided estimate - quality

- recall (if f is sufficiently smooth)

$$f(\theta + h) = f(\theta) + f'(\theta)h + \frac{f''(\theta)}{2}h^2 + O\left(h^3\right)$$

- so

$$\left(\frac{f(\theta + h) - f(\theta)}{h} - f'(\theta)\right)^2 = O(h^2)$$

- to determine *variance*, we first consider case that $(X_1^{\theta,\eta}, \ldots, X_n^{\theta,\eta})$ are independent of $(X_1^{\theta+h,\eta}, \ldots, X_n^{\theta+h,\eta})$

$$\mathrm{var}\left[\hat{f}_n(\theta + h) - \hat{f}_n(\theta)\right] = \frac{1}{n}\left(\mathrm{var}_{\theta+h,\eta}(h(X_1^{\theta+h,\eta}; \theta + h, \gamma)) + \mathrm{var}_{\theta,\eta}(h(X_1^{\theta,\eta}; \theta, \gamma))\right]$$

- we assume that $\mathrm{var}_{\theta+h,\eta}(h(X_1^{\theta+h,\eta}; \theta + h, \gamma)$ is a (locally) bounded function of θ
- this implies

$$\mathrm{MSE} = \mathrm{bias}^2 + \mathrm{var} = c_1 h^2 + c_2 \frac{1}{nh^2}$$

One-sided estimate - quality

- we have
$$\text{MSE} = \text{bias}^2 + \text{var} = c_1 h^2 + c_2 \frac{1}{nh^2}$$
- letting h tend to 0 while keeping n constant will make the MSE tend to ∞
- choosing h such that MSE is minimal leads to $h \propto n^{-1/4}$ and
$$\text{MSE} = O\left(n^{-1/2}\right)$$
- recall that (standard) MC approximation of option price itself (i.e. $(f\theta)$) satifies $\text{MSE} = O(n^{-1})$
- can we improve our rate?

One-sided estimate - using common random numbers

- we considered case that $(X_1^{\theta,\eta}, \dots, X_n^{\theta,\eta})$ are independent of $(X_1^{\theta+h,\eta}, \dots, X_n^{\theta+h,\eta})$
- In many cases, it is possible to use *common random numbers* in the simulation:
  - simulate $Z_1, \dots, Z_n$ i.i.d.
  - suppose there exists function g such that

$$h(X_i^{\theta,\eta}; \theta, \eta) \stackrel{d}{=} g(Z_i; \theta, \eta) \text{ and } h(X_i^{\theta+h,\eta}; \theta+h, \eta) \stackrel{d}{=} g(Z_i; \theta+h, \eta)$$

- if g is differentiable with respect to the parameter θ, then

$$\begin{aligned} \operatorname{var}\left(\hat{f}_n(\theta + h) - \hat{f}_n(\theta)\right) &= \frac{1}{n} \operatorname{var}\left(g(Z_1; \theta + h, \eta) - g(Z_1; \theta, \eta)\right) \\ &= \frac{1}{n} \operatorname{var}\left(\frac{\partial g}{\partial \theta}(Z_1, \theta) h + O\left(h^2\right)\right) = \frac{1}{n} O\left(h^2\right) \end{aligned}$$

which yields

$$\text{MSE} = c_1 h^2 + c_2 \frac{1}{n}$$

One-sided estimate - using common random numbers

- under assumption of “smoothness” (function g differentiable with respect to θ), the mean square error (of the one-sided estimate) is of the form

$$\mathrm{MSE} = c_1 h^2 + \frac{c_2}{n} .$$

  so, for fixed n, the variance does not explode as h becomes small
- bias is purely controlled by h and variance by n
- so we can take h as small as we want, and the convergence rate of MSE is $O(n^{-1})$ just as in the case of MC approximation of option value itself.

Remarks

Refinement: two-sided approach

- approximate $f'(\theta)$ by

$$\frac{f(\theta+h)-f(\theta-h)}{2h}$$

- from Taylor expansion $f(\theta+h)=f(\theta)+f'(\theta)h+.5f''(\theta)h^2+O(h^3)$ we see that bias is reduced from $O\big(h\big)$ to $O\big(h^2\big)$
- more accurate, but requires extra set of simulations

How to determine second-order Greeks?

- for second-order Greeks the following finite-difference could be used

$$f''(\theta)\approx\frac{f'(\theta+h)-f'(\theta)}{h}\approx\frac{\frac{f(\theta+h)-f(\theta)}{h}-\frac{f(\theta)-f(\theta-h)}{h}}{h}$$
$$=\frac{f(\theta+h)-2f(\theta)+f(\theta-h)}{h^2}$$

- now replace population moments by MC approximations

Section 2

Pathwise and likelihood ratio methods

The pathwise and likelihood ratio methods

- pathwise and likelihood ratio method are the most common refinements of Bump and Reprice
- strategy behind both methods is to write

$$\frac{d}{d\theta}\mathbb{E}[\ldots] = \frac{d}{d\theta}\int \cdots \mathrm{d}u \overset{?}{=} \int \frac{d}{d\theta} \cdots \mathrm{d}u = \tilde{\mathbb{E}}\left[\frac{d}{d\theta}\cdots\right]$$

  (if admissible) and to compute the right hand side by MC
- one often has a choice whether to put the dependence on θ into the payoff or into the density
  - first case leads to pathwise method, second to LRM method
- more advanced methods (like Malliavin Greeks) build on these two methods... but Bump-and-Reprice remains relevant in applications

Core idea Pathwise method

Suppose $X \sim N(\mu, \sigma^2)$ and we wish to compute the derivative of $\mathbb{E}_{\mu,\sigma}[f(X)]$ w.r.t μ. Denote by $\phi_{\mu,\sigma}$ the density of $N(\mu, \sigma)$.

Pathwise method: (assume f is smooth!)

$$\begin{aligned}\frac{\partial}{\partial\mu}\mathbb{E}_{\mu,\sigma}[f(X)] &= \frac{\partial}{\partial\mu}\int f(\mu+\sigma x)\phi_{0,1}(x)dx \stackrel{?}{=} \int f'(\mu+\sigma x)\phi_{0,1}(x)dx \\ &= \mathbb{E}_{\mu,\sigma}[f'(X)]\end{aligned}$$

If we are able to:

- obtain a closed-form formula for f' and
- simulate $X_1, \ldots, X_n$ i.i.d. from the same distribution as X,

then we can estimate $(\partial/\partial\mu)\mathbb{E}_{\mu,\sigma}[f(X)]$ by

$$\frac{1}{n}\sum_{i=1}^{n} f'(X_i).$$

Note that this is an unbiased estimator!

Suppose $X \sim N(\mu, \sigma^2)$ and we wish to compute the derivative of $\mathbb{E}_{\mu,\sigma}[f(X)]$ w.r.t μ. Denote by $\phi_{\mu,\sigma}$ the density of $N(\mu, \sigma)$.

LRM:

$$\frac{\partial}{\partial \mu}\mathbb{E}_{\mu,\sigma}[f(X)] = \frac{\partial}{\partial \mu}\int f(x)\phi_{\mu,\sigma}(x)dx \overset{?}{=} \int f(x)\frac{\frac{\partial \phi_{\mu,\sigma}}{d\mu}(x)}{\phi_{\mu,\sigma}(x)}\phi_{\mu,\sigma}(x)dx$$

$$= \mathbb{E}_{\mu,\sigma}\left[f(X)\frac{\frac{\partial \phi_{\mu,\sigma}}{\partial \mu}(X)}{\phi_{\mu,\sigma}(X)}\right] = \mathbb{E}_{\mu,\sigma}\left[f(X)\frac{\partial}{\partial \mu}\log\phi_{\mu,\sigma}(X)\right]$$

If we are able to:

- obtain a closed-form formula for $(\partial/\partial\mu)\log\phi_{\mu,\sigma}(X)$ and
- simulate $X_1, \ldots, X_n$ i.i.d. from the same distribution as X,

then we can estimate $(\partial/\partial\mu)\mathbb{E}_{\mu,\sigma}[f(X)]$ by

$$\frac{1}{n}\sum_{i=1}^{n}\left[f(X_i)\frac{\partial}{\partial \mu}\log\phi_{\mu,\sigma}(X_i)\right].$$

Note that this is an unbiased estimator!

Pathwise method (1)

Under sufficient smoothness, we indeed have:

$$\frac{\partial}{\partial\theta}E[f(Z,\theta)] = E\Big[\frac{\partial}{\partial\theta}f(Z,\theta)\Big].$$

Example: compute the delta, at $t = 0$, of a call option in the B-S market.

$$f(Z,S_0) = \max(S_T - K, 0)$$

$$S_T = S_0 \exp\big((r - \tfrac{1}{2}\sigma^2)T + \sigma\sqrt{T}\,Z\big), \quad Z \sim N(0,1).$$

The partial derivative with respect to S_0 is given by

$$\frac{\partial f}{\partial S_0} = \frac{\partial f}{\partial S_T}\frac{\partial S_T}{\partial S_0} = 1\{S_T > K\}\exp\Big((r - \frac{1}{2}\sigma^2)T + \sigma\sqrt{T}\,Z\Big).$$

Interchange of expectation and differentiation is valid in this case. The delta can be computed by straightforward MC.

Pathwise method (2)

What makes functions like $f(x) = \max(x, 0)$ sufficiently smooth for our purposes is that – even though they are not differentiable – there exists a (non-unique) function $f'(x)$ such that

$$f(b) - f(a) = \int_a^b f'(x)dx = \int_a^b 1_{x>0}dx.$$

Intuitively, this is what we need here: A function that behaves like the (non-existent) derivative of f when its inside an integral. Such functions are called *absolutely continuous*. You implicitly are aware of this from probability theory: random variables that have a pdf (and such random variables are said to have an absolutely continuous distribution).

Conversely, the pathwise approach fails for functions f that have jumps and thus cannot be represented as integrals. Bump & Reprice remains applicable.

Failure of pathwise method

What happens if we try the pathwise method to compute the delta, at $t = 0$, of a digital option?

$$\begin{aligned} f(Z, S_0) &= 1\{S_T > K\} \\ S_T &= S_0 \exp\big((r - \tfrac{1}{2}\sigma^2)T + \sigma\sqrt{T}\,Z\big), \quad Z \sim N(0,1). \end{aligned}$$

The partial derivative with respect to S_0 is given by

$$\frac{\partial f}{\partial S_0} = \frac{\partial f}{\partial S_T}\,\frac{\partial S_T}{\partial S_0} = 0 \qquad \text{(w.p. 1)}.$$

Interchange of expectation and differentiation is *not* valid in this case. Indeed, the expectation of the derivative is 0, but this is not the delta of the digital option.

Applicability of pathwise method

In general, the pathwise method fails when the payoff is *not* sensitive to small changes in the simulated path, such as in the case of digital options and barrier options. These are also the cases in which the bump-and-reprice method converges slowly but may be the only method.

The pathwise method requires that you know the derivative of the payoff with respect to the parameter of interest.

Because payoff functions often have kinks, so that their derivatives are not continuous, the pathwise method frequently does not apply to the calculation of second-order sensitivities such as gamma. We can still use a method based on finite differences (analogous to bump-and-reprice).

Likelihood ratio method

- in cases in which the payoff depends on a random variable whose density function is known explicitly, the difficulties arising from discontinuous payoff functions can be avoided by making θ a parameter of the *density*
- We can then differentiate the density rather than the payoff function

$$\frac{\partial}{\partial\theta}\mathbb{E}_{\theta,\eta}[F(X)] = \frac{\partial}{\partial\theta}\int F(u)\,g(u;\theta,\eta)\,du = \int F(u)\,\frac{\partial g}{\partial\theta}(u;\theta,\eta)\,du$$

$$= \int F(u)\,\frac{(\partial g/\partial\theta)(u;\theta,\eta)}{g(u;\theta,\eta)}g(u;\theta,\eta)\,du$$

$$= \mathbb{E}_{\theta,\eta}\left[F(X)\,\frac{(\partial g/\partial\theta)(X;\theta,\eta)}{g(X;\theta,\eta)}\right] = \mathbb{E}_{\theta,\eta}\left[F(X)\,\frac{\partial\log g(X;\theta,\eta)}{\partial\theta}\right].$$

Example (1)

Consider an option written on S_T in the standard BS model, and suppose we want to compute its delta at $t = 0$ (sensitivity w.r.t. S_0). We have

$$S_T = S_0 \exp\left((r - \tfrac{1}{2}\sigma^2)T + \sigma\sqrt{T}\, Z\right).$$

The density of S_T can be found by computing $\mathbb{Q}(S_T \leq s)$ for given $s \in \mathbb{R}$ and differentiating with respect to s:

$$g(s, S_0) = \frac{1}{s\sigma\sqrt{T}}\, \phi\left(\frac{\log(s/S_0) - (r - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}\right)$$

where $\phi(z) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}z^2)$, the standard normal density. So

$$\log\, g(s, S_0) = -\frac{\left(\log(s/S_0) - (r - \frac{1}{2}\sigma^2)T\right)^2}{2\sigma^2 T} + \cdots$$

where the dots indicate terms that do not depend on S_0.

Example (2)

The score function is

$$\frac{\partial}{\partial S_0} \log g(s, S_0) = \frac{\log(s/S_0) - (r - \frac{1}{2}\sigma^2)T}{\sigma^2 T S_0}.$$

The delta, at $t = 0$, of an option with payoff $F(S_T)$ can now be computed as

$$e^{-rT}\mathbb{E}_{\mathbb{Q}}\left[F(S_T)\,\frac{\log(S_T/S_0) - (r - \frac{1}{2}\sigma^2)T}{\sigma^2 T S_0}\right].$$

Given that $S_T = S_0 \exp\left((r - \frac{1}{2}\sigma^2)T + \sigma\sqrt{T}\,Z\right)$, we can also write this as

$$e^{-rT}\mathbb{E}_{\mathbb{Q}}\left[F(S_T)\,\frac{Z}{\sigma\sqrt{T}\,S_0}\right].$$

Comments

Comments on the LR method:

- It is applicable to *any* payoff function F, continuous or not.
- It can be used analogously for other Greeks, including higher-order derivatives.
- It depends on availability of the density of the underlying at time T in analytic form.

Section 3

Concluding remarks

Summary of sensitivity estimation

- Bump and reprice is conceptually simple and easy to apply, but comes with a bias. It is essential to use common random numbers.
- For continuous (piecewise differentiable) payoffs, the pathwise method can be used as an alternative to bump and reprice. The pathwise method is not applicable to discontinuous payoffs.
- The likelihood ratio method can be used both for continuous and discontinuous payoffs. It requires that the density of the underlying at expiry is available in analytic form.
- For second-order derivatives, finite differences are most generally applicable. Convergence is often slow, however.