Stability of finite difference sensitivities

Sensitivities of the option price with respect to input parameters, so-called *Greeks*, are crucial trading (hedge) parameters. It is key that a computational method does not only give accurate prices, but can also produce accurate and stable sensitivities.

In- and out-of-model sensitivities

Important examples are the derivatives with respect to the underlying asset price, the *delta* and *gamma* of the option,

$$\Delta = \frac{\partial V}{\partial S}, \quad \Gamma = \frac{\partial^2 V}{\partial S^2},$$

and its time derivative, *theta*,

$$\Theta = \frac{\partial V}{\partial t}.$$

These are often linked by the pricing equation, e.g. in the Black-Scholes model

$$\Theta + \frac{1}{2}\sigma^2 S^2 \Gamma + rS\Delta - rV = 0,$$

so if we have the price and two sensitivities, the other can be computed directly from the pricing equation. When solvig the pricing PDE numerically, the value function is computed over a whole range of underlying prices and times, which lends itself to a sensitivity analysis in these variables.

A different type of sensitivities are those to model parameters, such as the *vega*

$$\mathcal{V}(S,t) = \frac{\partial V}{\partial \sigma}(S,t).$$

This is an "out of model" sensitivity, because the assumption underlying the Black-Scholes model is that the volatility σ is constant. Solving the pricing equation numerically only gives an approximation $\widehat{V}(\sigma)$ to the option value $V(\sigma)$ for a single input parameter σ. To obtain an approximation to the vega, one can compute

$$\widehat{\mathcal{V}} = \frac{\widehat{V}(\sigma + \Delta\sigma) - \widehat{V}(\sigma)}{\Delta\sigma}, \tag{6.65}$$

which involves pricing the option numerically for a second input parameter shifted by a small amount $\Delta\sigma$. This is a procedure referred to in practice as "bumping", and is really independent of the pricing method (e.g. finite differences) used to determine the option value for fixed input parameters. The total error incurred is a combination of the truncation error of the finite difference in the σ direction, and the error of the numerical method used to determine the prices, say ϵ,

$$\widehat{\mathcal{V}} = \frac{V(\sigma+\Delta\sigma)-V(\sigma)}{\Delta\sigma} + \frac{(\widehat{V}(\sigma+\Delta\sigma)-V(\sigma+\Delta\sigma))-(\widehat{V}(\sigma)-V(\sigma))}{\Delta\sigma} \quad (6.66)$$

$$= \mathcal{V} + O(\Delta\sigma) + O(\epsilon/\Delta\sigma). \quad (6.67)$$

The pricing error ϵ is divided by the small step $\Delta\sigma$ and hence magnified, so for *fixed* ϵ, choosing the optimal step size $\Delta\sigma$ is a trade-off between the two terms. This is slightly pessimistic because the numerical errors for parameters σ and $\sigma+\Delta\sigma$ will normally be related and will cancel out to some extent in the difference in the second term of (6.66). The first term can be reduced e.g. by the use of a central difference instead of the one-sided difference in σ direction, giving better accuracy $O(\Delta\sigma^2)$, which allows larger step sizes hence reducing the magnification of the pricing error in the second term.

We note in passing that as an alternative one could differentiate the PDE with respect to σ, and solve the resulting PDE for $\mathcal{V}$. This removes some of the aforementioned problems, but requires implementation of a separate PDE to be solved in conjunction with the original one. A largely equivalent approach is *algorithmic differentiation* of the pricing algorithm with respect to the input parameters [Griewank and Walther, 2008]. Tools for *automatic differentiation* are available, see e.g. [Research website on automatic differentiation,]. See [Giles and Glasserman, 2006] for the application of these techniques to Monte Carlo sensitivities. Although the set-up of the pricing method is entirely different between Monte Carlo and finite differences, several issues concerning the computation of these "out of model" sensitivities stand outside the particular pricing method.

This is different for sensitivities to the underlying of the option, and time. We focus in the following on this first type of sensitivities, the derivatives of the value function with respect to its arguments S and t.

6.3.2 Finite difference sensitivities

In the finite difference method, one solves the pricing equation by approximating the derivatives by finite differences. An obvious route to Δ, Γ, and Θ is therefore to use these finite difference approximations, say V_n^m at point $S_n = n\Delta S$ and $t_m = m\Delta t$, applied to the computed finite difference solution,

$$\Delta(S_n,t_m) = \tfrac{\partial V}{\partial S}(S_n,t_m) \approx \Delta_n^m = \frac{V_{n+1}^m - V_{n-1}^m}{2\Delta S},$$

$$\Gamma(S_n,t_m) = \tfrac{\partial^2 V}{\partial S^2}(S_n,t_m) \approx \Gamma_n^m = \frac{V_{n+1}^m - 2V_n^m + V_{n-1}^m}{\Delta S^2},$$

$$\Theta(S_n,t_m) = \tfrac{\partial V}{\partial t}(S_n,t_m) \approx \Theta_n^m = \frac{V_n^{m+1} - V_n^{m-1}}{2\Delta t}.$$

We investigate the approximation "≈" more quantitatively below. The boundary points $n = 0$ and $n = N$ would need a different treatment, e.g. by one-sided differences, but it is unlikely these are required in practice. Note also that for Θ we have used a central difference and that this is somewhat decoupled from the question whether a central time difference is a stable time marching scheme. If we are interested in the Θ at $t = 0$, we can perform one more timestep backward in time to $t = -\Delta t$ and still use a central difference.

The key empirical observation is that finite differences taken from the numerical solution do not necessarily have the same accuracy as the numerical solution V_n^m itself. Far from it, the error is potentially amplified through numerical differentiation. Specifically, if V_{n-1}^m, V_n^m and V_{n-1}^m are all accurate to say $O(\Delta S^2)$ at every point, we may find

$$\begin{aligned} \Delta_n^m &= \frac{V_{n+1}^m - V_{n-1}^m}{2\Delta S} = \frac{V(S_n + \Delta S, t_m) + O(\Delta S^2) - V(S_n - \Delta S, t_m) + O(\Delta S^2)}{2\Delta S} \\ &= \frac{\partial V}{\partial S}(S_n, t_m) + O(\Delta S), \end{aligned} \tag{6.68}$$

instead of $O(\Delta S^2)$ as expected from a central difference, and similarly

$$\Gamma_n^m = \frac{\partial^2 V}{\partial S^2}(S_n, t_m) + O(1).$$

That is to say the error is not reduced by refining the grid size, hence the finite difference gamma does not necessarily approximate the true gamma at all. This is already highlighted in [Shaw, 1998].

The way to get around this is to ensure that the error terms "$O(\Delta S^2)$" in (6.68) cancel out at neighbouring points S_{n-1}, S_n, S_{n+1} to some higher order in ΔS. This does happen if the (truncation) error is sufficiently smooth and is the case for smooth data. As we will see next, this is not the case in most relevant applications from derivative pricing.

Consider a European put. The discretised payoff/terminal condition is

$$V_n^M = (K - S_n)^+,$$

and we assume for simplicity that the grid has been chosen such that a grid point, say S_k, coincides with the strike, $S_k = K$. Then[1]

$$\Delta_n^M = \begin{cases} -1 & n < k \\ -\frac{1}{2} & n = k \\ 0 & n > k \end{cases} \tag{6.69}$$

and

$$\Gamma_n^M = \begin{cases} \frac{1}{\Delta S} & n = k \\ 0 & n \neq k \end{cases}. \tag{6.70}$$

This falls square into the numerical analysis of smoothing differences for non-smooth data in 5.3. In fact, differentiating the pricing PDE, e.g. in the Black-Scholes model

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0, \tag{6.71}$$

gives equations for the sensitivities. For the Δ, after S differentiation,

$$\frac{\partial \Delta}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 \Delta}{\partial S^2} + (r+\sigma^2)S\frac{\partial \Delta}{\partial S} = 0, \tag{6.72}$$

with terminal condition for the put

$$\Delta(S,T) = \begin{cases} -1 & \text{if } S < K \\ 0 & \text{else} \end{cases}.$$

By a further differentiation, for Γ,

$$\frac{\partial \Gamma}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 \Gamma}{\partial S^2} + (r+2\sigma^2)S\frac{\partial \Gamma}{\partial S} + (r+\sigma^2)\Gamma = 0, \tag{6.73}$$

where

$$\Gamma(S,T) = \delta(S-K),$$

the Dirac delta.

So, if we solve the orginal pricing equation and then apply finite differences to the finite difference solution to compute the Greeks, in essence what we are doing is numerically solve a PDE with more irregular final condition.

6.3.3 Stability analysis

We know from 5.3 that irregular data are characterised by slowly decaying Fourier transforms, and we need strong stability of timestepping schemes to get realistic solutions. A key indicator for stability were the eigenvalues of the iteration matrix $K = {K_1}^{-1}K_2$, denoted say by λ_j. For constant coefficient problems these are known in closed form, with their corresponding eigenvectors. This is no longer the case for the Black-Scholes PDE or other models, but we can still investigate the behaviour *numerically*. Fig. 6.5 shows the eigenvalues of K in descending order ($\lambda_{j+1} \leq \lambda_j$), and corresponding eigenvectors W_j.

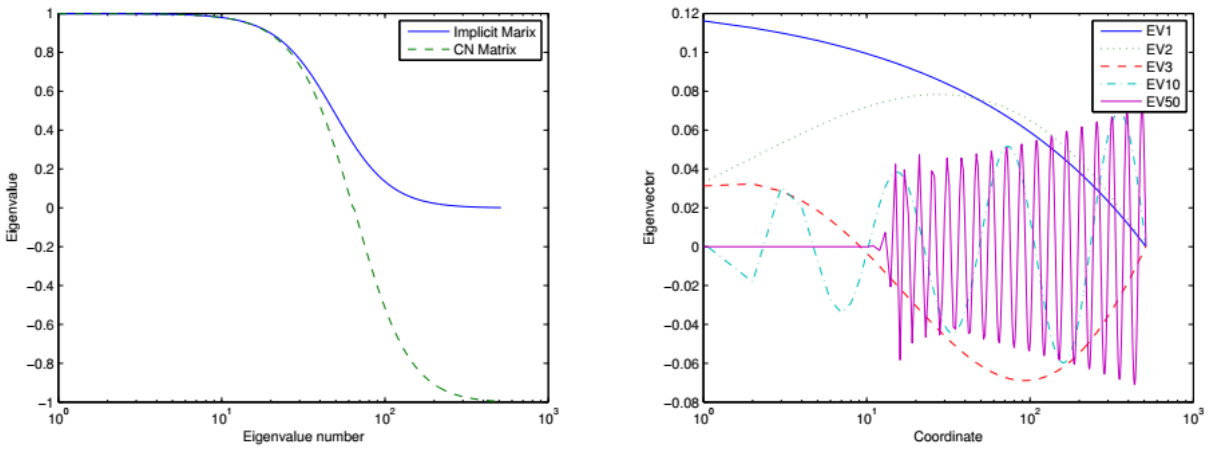

Figure 6.5: Left: Eigenvalues λ_k of the implicit and Crank-Nicolson matrix. Right: Eigenvectors to a few different eigenvalues λ_k, $k = 0, 1, 2, 9, 49$.

The spectrum is qualitatively very similar to the heat equation, see Fig. 6.5, with the eigenvalues smoothly decaying from 1 to 0 for the implicit scheme, and to -1 for the Crank-Nicolson scheme. The corresponding eigenvectors, in log coordinates, are still oscillatory for increasing eigenvalue number. Note the eigenvectors only depend on the space discretisation and are therefore identical for both timestepping schemes. The eigenvectors are no longer orthogonal, but given the remarkable similarity to the trigonometric eigenvectors of the heat equation, a transformation to the eigenvector basis is still very similar to taking the Fourier transform. So, if we introduce $W = (W^0, \ldots, W^N)$, the matrix of eigenvectors of K, $\widehat{V}^m = W^{-1}V^m$ contains the coordinates of $V^m = (V_0^m, \ldots, V_N^m)'$ in the eigenvector basis. The top plots of Fig. 6.6 show $\widehat{V}^M$ and $\widehat{V}^0$ for both the Crank-Nicolson and implicit Euler scheme.

Moving on to the Greeks, define similarly $\Delta^m = (\Delta_0^m, \ldots, \Delta_N^m)'$, $\Gamma^m = (\Gamma_0^m, \ldots, \Gamma_N^m)'$, and $\widehat{\Delta}^m = W^{-1}\Delta^m$, $\widehat{\Gamma}^m = W^{-1}\Gamma^m$. The remaining plots of Fig. 6.6 show $\widehat{\Delta}^M$ and $\widehat{\Delta}^0$, $\widehat{\Gamma}^M$ and $\widehat{\Gamma}^0$, the finite difference delta and gamma represented in the eigenvector basis. Direct computation shows that

$$\widehat{V}^0 = \Lambda^M \widehat{V}^M,$$

where $\Lambda = \mathrm{diag}(\lambda_j)$ the diagonal matrix of eigenvalues of K. Observe particularly at the bottom left plot of Fig. 6.6 the band of high frequency components which Crank-Nicolson fails to eliminate. This destroys the accuracy of the Gamma.

To obtain stable results, a smoothing timestepping scheme like the fractional-step θ-scheme or Crank-Nicolson with Rannacher start-up has to be used, as discussed in 5.3.2. The delta of a standard call or put is a step-function, its gamma a Dirac distribution. For a digital option, the delta is already a Dirac distribution and the gamma the derivative thereof.

[1] If the strike does not coincide with a point S_k, but lies between two S_k and S_{k+1}, one gets a different weighting of Δ_k^M, Δ_{k+1}^M etc. We return to the accurate grid approximation of non-smooth and singular functions more systematically in 6.4.

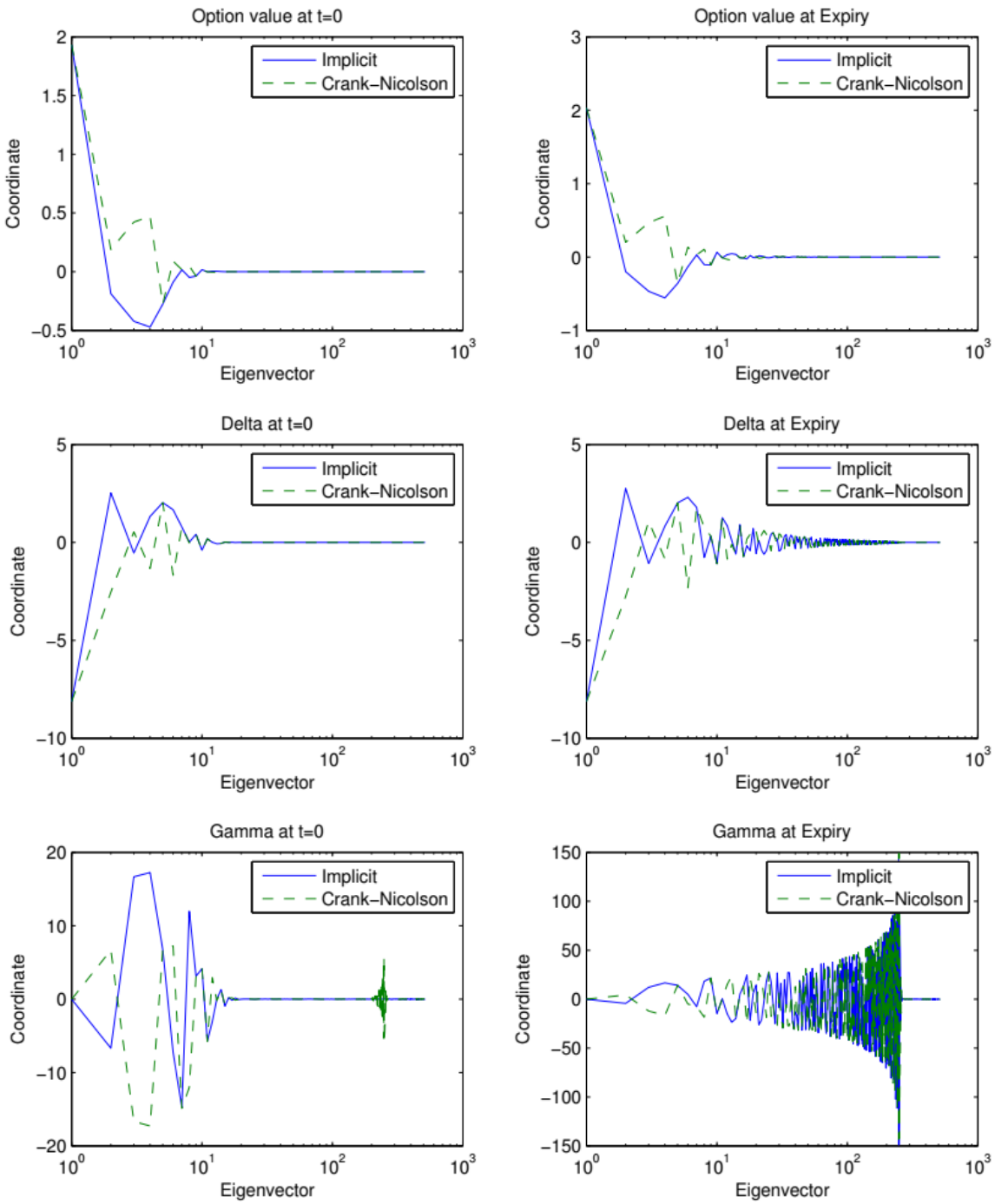

Figure 6.6: Finite difference value, delta, and gamma of a put option in the implicit and Crank-Nicolson scheme, in the eigenvector basis of the discretisation matrix.

We note without proof that strongly A-stable schemes like the fractional-step θ-scheme from 5.3.2 can still handle the situation of higher order derivatives of distributions. It is also shown in [Giles and Carter, 2006] that four fully implicit steps in the Rannacher start-up are still sufficient for a digital gamma. It is unlikely that higher derivative are needed in practice.

Remark 6.3.1. *The above PDEs (6.72) and (6.73) are Black-Scholes-type PDEs, but with slightly modified coefficients. We do not want to elaborate on this excessively as it usually causes no problems in practice, but this modification could conceivably have a bearing on the viability of numerical schemes if finite difference Greeks are required. Recall from 6.2.2 that for the Black-Scholes PDE, the central difference scheme imposed a constraint $\sigma^2 \geq r$ on the parameters for guaranteed maximum norm stability. This results from the type of degeneracy of the coefficients at the boundary, and expresses that the drift factor r has to be smaller than the diffusion factor σ^2. With r replaced by $r + \sigma^2$ in the Δ PDE, and $r + 2\sigma^2$ in the Γ case, this is* never *fulfilled. From the analysis in 6.2.2 it is clear that this problem only arises for nodes close to the $S = 0$ boundary, and if need be we could apply upwinding for these nodes only, which would not effect the second order convergence of the scheme overall. In practice this does not turn out to be necessary though.*