Accelerating the calibration of stochastic volatility models

Fiodar Kilin

Frankfurt School of Finance & Management,
Quanteam AG

November 16, 2007

Models

- Heston model
- Bates model
- Barndorff-Nielsen-Shephard model
- Variance-Gamma with the Cox-Ingersoll-Ross stochastic clock
- Variance-Gamma with the Gamma-Ornstein-Uhlenbeck stochastic clock
- Normal Inverse Gaussian with the Cox-Ingersoll-Ross stochastic clock
- Normal Inverse Gaussian with the Gamma-Ornstein-Uhlenbeck stochastic clock

Calibration setup

- Optimization method
- Regularization
- Weights assigned to each calibration instrument
- Market data
- Pricing method

Calibration setup

- Optimization method
- Regularization
- Weights assigned to each calibration instrument
- Market data
- Pricing method : this presentation

Challenges

- Accuracy
- Numerical stability
- Stability of the calibrated parameter time series
- Speed (without sacrificing previous three points)

Challenges

- Accuracy
- Numerical stability
- Stability of the calibrated parameter time series
- Speed (without sacrificing previous three points) : this presentation

Pricing methods

- Direct integration
- Fast Fourier Transform (FFT)
- Fractional FFT

Pricing methods

- Irrelevant: one option
- Relevant: the whole set of liquid vanilla options

How to choose?

Strategy A (fortunately we did not use this strategy)

- Compare unoptimized implementations
- Choose a pricing method
- Optimize this method (mathematical modifications as well as implementation techniques)

Strategy B (we used it)

- Optimize all three methods
- Compare optimized implementations
- Choose a pricing method

How to choose?

- Strategy A: fast implementation, wrong choice (fractional FFT)
- Strategy B: right choice (modified direct integration method)

Direct integration

One-dimensional numerical quadrature, for example Gaussian quadrature.

Direct integration

Formulas

- Attari (2004) : one integral, quadratic term in the denominator - optimal for the calibration
- Lewis (2001) : one integral, quadratic term in the denominator - also optimal for the calibration
- Heston (1993) : two integrals, linear term in the denominator - not optimal for the calibration, good for calculation of delta

Direct integration

Attari (2004) :

$$C(S_0,T,K) = S_0 - \frac{1}{2}e^{-rT}K - e^{-rT}K \times \tag{1}$$

$$\times \left(\frac{1}{\pi}\int_0^{+\infty} \frac{(\text{Re}(\phi(\omega)) + \frac{\text{Im}(\phi(\omega))}{\omega})\cos(\omega l(K)) + (\text{Im}(\phi(\omega)) - \frac{\text{Re}(\phi(\omega))}{\omega})\sin(\omega l(K))}{1+\omega^2} d\omega \right)$$

where

$$l(K) = \ln\left(\frac{Ke^{-rT}}{S_0}\right),$$

S_t is the underlying price, K is the strike price, T is the maturity of the option, r is the risk-free interest rate, the dividend yield is assumed to be zero, Q is the risk-neutral measure and

$$\phi(\omega) = E_Q(e^{i\omega x})$$

is the characteristic function of

$$x = \ln\left(\frac{S_T}{S_0}\right) - rT.$$

FFT

Carr and Madan (1999) :

$$C(S_0, T, k) = \frac{e^{-\gamma k}}{\pi} \int_0^{+\infty} e^{-iku} \psi(u) du \tag{2}$$

where k denotes the log of the strike price, γ is a dampening parameter and

$$\psi(u) = \frac{e^{-rT} \widehat{\phi}(u - (\gamma + 1)i)}{\gamma^2 + \gamma - u^2 + (2\gamma + 1)ui}, \tag{3}$$

where

$$\widehat{\phi}(\omega) = E_Q(e^{i\omega\widehat{x}}) \tag{4}$$

is the characteristic function of the log price $\widehat{x} = \ln(S_T)$.

FFT

The integral in (2) is approximated using an integration rule

$$\int_0^{+\infty} e^{-iku}\psi(u)du \approx \sum_{j=0}^{N_{FFT}-1} e^{-iku_j}\psi(u_j)w_j\delta, \tag{5}$$

$$u_j = j\delta, \tag{6}$$

where N_{FFT} is the number of grid points and the weights w_j implement the integration rule.

FFT

The crucial limitation of the FFT method is that the grid points u_j must be chosen equidistantly. This limitation prohibits the use of the most effective integration rules such as the Gaussian quadrature.

FFT

The FFT pricing method simultaneously computes the values of the integral approximations (5) for the set of log-strikes $\{k_m = -(\frac{N_{FFT}\lambda}{2}) + m\lambda, m = 0, \ldots, N_{FFT} - 1\}$. The simultaneous calculation for all strikes is not an exclusive advantage of the FFT-based methods, because a slightly modified direct integration method also has this advantage. This simple modification is described in the section "Caching technique".

FFT

The second important restriction is that the grid spacings must satisfy the condition

$$\lambda\delta = \frac{2\pi}{N_{FFT}}. \qquad (7)$$

If this condition is satisfied, the sums in (5) can be expressed in the form

$$\sum_{j=0}^{N_{FFT}-1} e^{-iku_j}\psi(u_j)w_j\delta = \sum_{j=0}^{N_{FFT}-1} e^{-i\lambda\delta jm}h_j = \sum_{j=0}^{N_{FFT}-1} e^{-i(\frac{2\pi}{N_{FFT}})jm}h_j, \qquad (8)$$

which allows the application of the FFT procedure invoked on the vector $h = \{h_j = e^{i(\frac{N\lambda}{2}j\delta)}\psi(u_j)w_j\delta, j = 0, \ldots, N_{FFT} - 1\}$.

Fractional FFT

Chourdakis (2005) has shown how the method of Carr and Madan (1999) can be accelerated using the fractional FFT algorithm. This algorithm rapidly computes sums of the form

$$D_k(h,\alpha) = \sum_{j=0}^{N-1} e^{-i2\pi kj\alpha} h_j \tag{9}$$

for any value of α.

Fractional FFT

The fractional FFT method can be applied without the need to impose the restriction

$$\lambda\delta = \frac{2\pi}{N_{FFT}}. \quad (10)$$

Fractional FFT

However, the fractional FFT method does not overcome the crucial limitation of the FFT method because the grid points u_i still must be chosen equidistantly.

Fractional FFT

Fractional FFT is implemented by invoking three FFT procedures, i.e.,

$$D_k(h,\alpha) = (e^{-i\pi k^2 \alpha})_{k=0}^{N-1} \odot D_k^{-1}(D_j(y) \odot D_j(z)), \tag{11}$$

where

$$y = ((h_j e^{-i\pi j^2 \alpha})_{j=0}^{N-1}, (0)_{j=0}^{N-1}), \tag{12}$$

$$z = ((e^{i\pi j^2 \alpha})_{j=0}^{N-1}, (e^{i\pi (N-j)^2 \alpha})_{j=0}^{N-1}), \tag{13}$$

$D_k(h)$ denotes the FFT sum

$$D_k(h) = \sum_{j=0}^{N-1} e^{-i\frac{2\pi}{N}kj} h_j, \tag{14}$$

$D_k^{-1}(h)$ is the inverse FFT sum

$$D_k^{-1}(H) = \frac{1}{N} \sum_{j=0}^{N-1} e^{i\frac{2\pi}{N}kj} H_j, \tag{15}$$

and $\odot$ denotes element-by-element vector multiplication.

Fractional FFT

The fractional FFT pricing method is faster than the FFT pricing method, because the absence of the restriction

$$\lambda\delta = \frac{2\pi}{N_{FFT}} \tag{16}$$

allows the use of sparser grids. This effect is more important in terms of computing time than the disadvantage of using three FFT routines instead of one.

Caching technique

The most time-consuming part of the computation is the evaluation of the characteristic function. For example the characteristic function of the Heston model

$$\phi(\omega) = \exp\{\eta\kappa\theta^{-2}((\kappa - \rho\theta\omega i - d)T - 2\ln(\frac{1 - ge^{-dT}}{1-g})) + \sigma_0^2\theta^{-2}(\kappa - \rho\theta\omega i - d)\frac{1-e^{-dT}}{1-ge^{-dT}}\}, \tag{17}$$

$$d = ((\rho\theta\omega i - \kappa)^2 - \theta^2(-i\omega - \omega^2))^{1/2}, \tag{18}$$

$$g = \frac{\kappa - \rho\theta\omega i - d}{\kappa - \rho\theta\omega i + d}, \tag{19}$$

contains two complex exponents (we do not count identical repeated terms), one complex logarithm and one complex square root.

Caching technique

Therefore an extremely important requirement for an effective implementation of the calibration algorithm is the following: The number of evaluations of the characteristic function should be as low as possible. If the calibration algorithm uses the direct integration method to compute the values of vanilla options, a caching technique should be used to avoid unnecessary recalculations of the characteristic function.

Caching technique

If the caching technique is **not** used, the calculation of the values of vanilla options at each iteration of the optimization algorithm includes the following steps:

1. Loop over expiries of the vanilla options.
2. Loop over strikes of the vanilla options.
3. Loop over the points $\omega_i, i = 1, \ldots, U$ that are used to evaluate the integral in (1) numerically.
4. Evaluate the characteristic function in ω_i.
5. Evaluate the integrand in ω_i.
6. Calculate the value of the vanilla option.

Caching technique

However, the value of the characteristic function does not depend on the strike. If we use the same grid $\omega_i, i = 1, \ldots, U$ for all options and run the described algorithm, we recalculate the same values of the characteristic function at each step of the strike-loop. We can use the following modification of the algorithm in order to avoid these recalculations:

Caching technique

1. Loop over expiries of the vanilla options.
2. Loop over strikes of the vanilla options.
3. Loop over the points $\omega_i, i = 1, \ldots, U$ that are used to evaluate the integral in (1) numerically.
4. If we are at the first step of the strike-loop, evaluate the characteristic function in ω_i and save this value in the cache.
5. If we are not at the first step of the strike-loop, read the value of the characteristic function in ω_i from the cache.
6. Evaluate the integrand in ω_i.
7. Calculate the price of the vanilla option.

Caching technique

The numerical evaluation of the integral in (1) requires a choice of the numerical upper integration limit. Suppose a maximum tolerable truncation error is given. Then the numerical upper integration limit $\overline{\omega}$ depends on the maturity T and the strike K of the vanilla option: $\overline{\omega} = \overline{\omega}(T, K)$. We can still use the same ω-grid for all T and K - we just define the index of the last integration point as an integer-valued function $U(T, K)$ that satisfies the condition

$$\omega_{U(T,K)} \leq \overline{\omega}(T, K) < \omega_{U(T,K)+1}. \tag{20}$$

Caching technique

The grid at step 3 of the algorithm can now be defined as $\omega_i, i = 1, \ldots, U(T, K)$. In most cases the function U(T,K) is an increasing function of K. It leads to a different number of loop iteration at step 3 for different K. Therefore, we have to modify the described algorithm once more in order to take this fact into account. We can use a reverse order of strikes or we can control at each point ω_i whether the characteristic function has been already evaluated at this point. We can also combine these two solutions. In this case the algorithm is:

Caching technique

1. Loop over expiries of the vanilla options.
2. Loop over strikes of the vanilla options. Use a reverse order of strikes.
3. Loop over the points $\omega_i, i = 1, \ldots, U(T, K)$ that are used to evaluate the integral in (1) numerically.
4. If the value of the characteristic function in ω_i is still not in the cache, evaluate it and save this value in the cache.
5. If the value of the characteristic function in ω_i is already in the cache, use this precomputed value.
6. Evaluate the integrand in ω_i.
7. Calculate the value of the vanilla option.

Caching technique

There is a further possibility to accelerate this algorithm. Some terms of the characteristic function do not depend on T. These terms can be precomputed before starting the loop over expiries of the vanilla options. For example, we recommend to compute the term (18) only once and store it, because it contains a time-consuming square root operator.

Comparison

Simultaneous pricing for different strikes?

- FFT: yes (per definition)
- Fractional FFT: yes (per definition)
- Direct integration without caching technique : no
- Direct integration with caching technique : yes

Comparison

Simultaneous pricing for different strikes?

- FFT: yes (per definition) $\Rightarrow$ The most popular argument in favour of FFT
- Fractional FFT: yes (per definition)
- Direct integration without caching technique : no
- Direct integration with caching technique : yes

Comparison

Simultaneous pricing for different strikes?

- FFT: yes (per definition) ⇒ The most popular argument in favour of FFT
- Fractional FFT: yes (per definition)
- Direct integration without caching technique : no
- Direct integration with caching technique : yes ⇒ The possibility of simultaneous pricing for different strikes cannot be considered as a criterium for comparison of pricing methods

Comparison

We have to define other criteria for the comparison.

Comparison

The FFT algorithm reduces the number of multiplications in the required N_{FFT} summations from an order of N_{FFT}^2 to that of $N_{FFT} \ln_2 N_{FFT}$ (Carr and Madan (1999)). However the computing time required for these multiplications is negligible in comparison with the time required for the evaluations of the characteristic function. Therefore we concentrate on the number of the calculations of the characteristic function only.

Comparison

We have to define other criteria for the comparison.

Comparison

We have to define other criteria for the comparison:
The rate of decay of the integrand ? $\Rightarrow$ No.
The formulas of Carr and Madan (FFT, Fractional FFT) of Attari (Modified direct integration) and of Lewis(Modified direct integration) have quadratic rate of decay of the integrand.

Comparison

We have to define other criteria for the comparison.

Comparison

We have to define other criteria for the comparison:
The speed of the numerical integration method ? $\Rightarrow$ Yes.
Obviously, there are a lot of techniques of simple numerical integration in the general case that are both faster and more accurate than integration using FFT. They are designed to minimize the number of integrand evaluations. One of these techniques is the Gaussian quadrature formula. This section shows that the grid for the numerical integration (1) with six-point Gaussian quadrature is at least seven times more economical than the FFT-grid in (2).

Comparison

Numerical experiments

- As we have already pointed out, the number of the evaluations of the characteristic function is the main factor driving the calibration time. We have carried out a numerical experiment to compare the influence of this factor in each pricing method.
- Then we have conducted a second numerical experiment where we have compared the calibration time directly.

Experiment 1

Accuracy (implied volatility basis points)	FFT	Fractional FFT	Modified direct integration
2.0	4096	1024	96
1.0	4096	2048	126
0.2	8192	2048	162
0.02	16384	4096	582

Table: Grid sizes that are needed to obtain some benchmark accuracy levels

Experiment 2

Model	FFT	Fractional FFT	Direct integration
Heston	466	239	15
Bates	620	316	20
BNS	405	208	13
VG-CIR	540	281	17
VG-GOU	522	269	17
NIG-CIR	546	280	18
NIG-GOU	521	273	17

Table: Average calibration time (in seconds).

Conclusion

- An efficient implementation of the direct integration method results in a sizable speed up of the calibration of stochastic volatility models.

Conclusion

- An efficient implementation of the direct integration method results in a sizable speed up of the calibration of stochastic volatility models.
- This method even outperforms the calibration with the fractional FFT.

Conclusion

- An efficient implementation of the direct integration method results in a sizable speed up of the calibration of stochastic volatility models.
- This method even outperforms the calibration with the fractional FFT.
- The simultaneous pricing of options with different strikes is not an exclusive advantage of the FFT methods compared to the direct integration method, because an application of a cache technique leads to simultaneous pricing of options with different strikes in the framework of direct integration.

Conclusion

- The pricing methods differ in two aspects only: the numerical integration technique and the pricing formula. The combination of these factors results in higher calculation speed of the direct integration method in comparison to the FFT and fractional FFT methods.

Conclusion

- The pricing methods differ in two aspects only: the numerical integration technique and the pricing formula. The combination of these factors results in higher calculation speed of the direct integration method in comparison to the FFT and fractional FFT methods.
- Specifically: (1) Gaussian quadrature is a much faster numerical integration technique than the FFT, (2) The transformed pricing formula of Attari (2004) provides approximately the same rate of decay of the integrand in comparison with the main formula of the FFT method.

Conclusion

- The direct integration method is frequently criticized in the literature. However this critique is valid only if we consider an unoptimized implementation of the general formula.

Conclusion

- The direct integration method is frequently criticized in the literature. However this critique is valid only if we consider an unoptimized implementation of the general formula.
- The use of the modified pricing formula and the caching technique makes the direct integration method the best choice for practical applications.

Thank you!

Contact Details

Fiodar Kilin
Frankfurt School of Finance & Management,
Quanteam AG

Mobile: +49 178 8055 681
Email: Fiodar.Kilin@quanteam.de
www: http://www.quanteam.de