Student Research Group Report

Monte-Carlo Methods for the Heston Model

Artemy Sazonov, Danil Legenky, Kirill Korban

Contents

List of Figures

List of Tables

Introduction

One of the first diffusion-based models in mathematical finance was introduced in 1973 in the paper by Fisher Black and Myron Sholes [BS73]. However, the model was not very realistic, as it did not take into account the variability of the volatility process, which was proven not to be a constant in the real stock market. The implied volatility of the stock options was not the same for different maturities and strikes.

Later, the class of so-called local volatility models was developed (Dupire et. al.). They fixed the problem of the spot implied volatility: now we could get a perfect fin into the spot prices of the options. However, the local volatility models give us the wrong dynamics, which is crucial to valuate the price of different derivatives.

In 1993, Steven Heston introduced a new diffusion-based model [Hes93], but he made a vital assumption: the variance process is not a constant, not a determenistic function of time and stock price, but follows a diffusion process, called the Cox-Ingersol-Ross (CIR) process. The stochastic volatility models cannot be perfectly calibrated to fit the volatility smile, but they give us a realistic dynamics of the implied volatility surface.

In this paper we revise the Heston model and its most popular simulation methods. We remind the reader of some basic facts abou the Monte-Carlo methods in finance. We also study the empirical speed of convergence of the simulation methods and the accuracy of the option greeks. Futhermore, we implement a multi-threaded versions of the desired simulation teqniques and optimize them for the best possible performance in **Python**.

We provide the reader with the code for the simulation methods and the greeks computation for the results to be reproductible.

Part I

Monte-Carlo Methods for the Heston Model: A Theoretical Review

Chapter 1

A review of the original Heston model

1.1 Basic facts

We shall use the following resources in this chapter: [Hes93] and [Gat12]. Assume that the spot asset's price S at time t follows the diffusion (1.1) – (1.2):

$$dS(t) = \mu S(t)dt + \sqrt{v(t)}S(t)dZ_1(t), \tag{1.1}$$

$$dv(t) = \left(\delta^2 - 2\beta v(t)\right) dt + 2\delta\sqrt{v(t)}dZ_2(t), \tag{1.2}$$

where Z_1, Z_2 are the correlated Wiener processes with $dZ_1 dZ_2 = \rho dt$.

1.2 PDEs

1.3 A closed-form solution for the European call option

Chapter 2

A review of the Monte-Carlo methods for diffusions

2.1 Randomness in Probability Theory

A. N. Kolmogorov in «On Logical Foundations of Probability Theory»: *In everyday language we call random these phenomena where we cannot find a regularity allowing us to predict precisely their results. Generally speaking there is no ground to believe that a random phenomenon should possess any definite probability. Therefore, we should have distinguished between randomness proper (as absence of any regularity) and stochastic randomness (which is the subject of the probability theory). Since randomness is defined as absence of regularity, we should primarily specify the concept of regularity. The natural means of such a specification is the theory of algorithms and recursive functions...*

Check out the lecture by A. N. Shiryaev for more details.

2.2 Laws of large numbers and central limit theorems

Theorem 1 (Khinchin). *Let $X_1, X_2, \ldots, X_n$ be a sequence of independent and identically distributed random variables with $\mathbb{E}X_i = \mu$. Then*

$$\mathbb{P}\text{-}\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} X_i = \mu. \tag{2.1}$$

Theorem 2 (Kolmogorov). *Let $X_1, X_2, \ldots, X_n$ be a sequence of independent and identically distributed random variables. Then $\exists \mathbb{E}X_i = \mu$, if and only if*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} X_i \overset{a.s.}{=} \mu. \tag{2.2}$$

Theorem 3 (Lindeberg-Lévy). *Let $X_1, \ldots, X_n$ be a sequence of i.i.d. random variables with $\mathbb{E}[X_i] = \mu$ and $\operatorname{var} X_i = \sigma^2$. Then as n approaches infinity, the random variables $\sqrt{n}(\bar{X}_n - \mu)$ converge in law to a normal distribution $\mathcal{N}(0, \sigma^2)$, i.e.*

$$\sqrt{n}\left(\bar{X}_n - \mu\right) \xrightarrow{d} \mathcal{N}\left(0, \sigma^2\right). \tag{2.3}$$

2.3 The statistical foundations of the Monte-Carlo methods

Lemma 4. *Let $X_1, X_2, \dots, X_n$ be a series of independent and identically distributed random variables, and $h : \mathbb{R} \to \mathbb{R}$ be a borel function. Then $h(X_1), h(X_2), \dots, h(X_n)$ is a series of independent and identically distributed random variables.*

Thus, we could write an unbiased consistent estimator of $\mathbb{E}\left[h(X)\right]$ as follows:

$$\widehat{\mathbb{E}\left[h(X)\right]} = \frac{1}{n}\sum_{i=1}^{n} h(X_i). \tag{2.4}$$

Definition 1. *Monte Carlo simulation is a set of techniques that use pseudorandom number generators to solve problems that might be too complicated to be solved analytically. It is based on the central limit theorem.*

Asymptotic confidence interval for $\hat{\mu} = \widehat{\mathbb{E}\left[X\right]}$ at the confidence level α:

$$\mu \in \left(\hat{\mu} - z_{\alpha/2}\sqrt{\frac{\sigma^2}{n}}, \hat{\mu} + z_{\alpha/2}\sqrt{\frac{\sigma^2}{n}}\right). \tag{2.5}$$

That means that the estimation error is equal to $2z_{\alpha/2}\sqrt{\frac{\sigma^2}{n}}$.

2.4 General Monte-Carlo methods for Gaussian diffusions

2.5 The Main Methods

2.5.1 Euler-Maruyama

Forward Euler Scheme for ODEs

Suppose that we have an ODE of the form

$$dX(t) = f(X(t), t)dt, \quad X(0) = X_0. \tag{2.6}$$

Then it could be numerically solved by the following finite difference scheme:

$$X_{n+1} = X_n + f(t_n, X_n)h_n, \tag{2.7}$$

where $t_n = \sum_{k=1}^{n} h_n, t_0 = 0$ is a grid.

Backward Euler Scheme for ODEs

Suppose that we have an ODE of the form

$$dX(t) = f(X(t), t)dt, \quad X(0) = X_0. \tag{2.8}$$

Then it could be numerically solved by the following finite difference scheme:

$$X_{n+1} = X_n + f(t_{n+1}, X_{n+1})h_n, \tag{2.9}$$

where $t_n = \sum_{k=1}^{n} h_n, t_0 = 0$ is a grid.

Euler-Maruyama Scheme for SDEs

Suppose we have a diffusion of the form

$$dX(t) = f(X(t), t)dt + \sigma(X(t), t)dW(t), \quad X_0 = X_0.$$

Then it could be numerically solved by the following finite difference scheme:

$$X_{n+1} = X_n + f(t_n, X_n)h_n + \sigma(t_n, X_n)\sqrt{h_n}Z_n, \tag{2.10}$$

where $(Z_n)_{n=1,2,\ldots}$ is a sample of standard normal random variables, and $t_n = \sum_{k=1}^n h_n, t_0 = 0$ is a grid. The same method could be generalized for the two-factor Gaussian diffusions. Further we assume that $(t_i)_{i=0,1,\ldots}$ is a uniform grid with $t_i = ih$.

Definition 2. *Let $\hat{X}^n(t)$ be a piecewise mesh approximation of an SDE solution $X(t)$ (we assume that there exists a unique strong solution). Then a scheme is said to have a strong convergence of order p if*

$$\mathbb{E}\left[\left|\hat{X}^n(T) - X(T)\right|\right] \leq Ch^p, \quad n \to \infty. \tag{2.11}$$

A scheme is said to have a weak convergence of order p if for any polynomial $f : \mathbb{R} \to \mathbb{R}$ we have

$$\left|\mathbb{E}\left[f(\hat{X}^n(T))\right] - \mathbb{E}\left[f(X(T))\right]\right| \leq Ch^p, \quad n \to \infty. \tag{2.12}$$

Theorem 5. *Under some technical assumptions the Euler-Maruyama scheme (2.10) has a strong convergence of order $1/2$ and a weak convergence of order 1.*

Since our goal is to approximate $\mathbb{E}\left[h(X)\right]$ with a given accuracy and the least possible number of simulations, we need to compare the weak convergence rate between the methods.

Chapter 3

Methods of simulation of the Heston stochastic volatility model

3.1 Euler Scheme

Suppose we have the Heston model (1.1) – (1.2). Then it could be numerically solved by the following finite difference scheme:

$$S_{n+1} = S_n + \mu S_n h_n + \sqrt{v_n} S_n \sqrt{h_n} Z_{1,n}, \tag{3.1}$$

$$v_{n+1} = v_n + \left(\delta^2 - 2\beta v_n\right) h_n + \sigma \sqrt{v_n} \sqrt{h_n} Z_{2,n}, \tag{3.2}$$

where $(Z_{1,n})_{n=1,2,\dots}$ and $(Z_{2,n})_{n=1,2,\dots}$ are ρ-correlated samples of standard normal random variables, and $t_n = \sum_{k=1}^{n} h_n$ is a mesh grid. But we have a problem: during simulation of the Heston model using Euler method S_{t_n} and v_{t_n} could be negative. How do we deal with this inconvenience? Let us introduce the log-prices

$$X(t) := \log \frac{S(t)}{S(0)}. \tag{3.3}$$

We take the positive part of the variance:

$$X_{n+1} = X_n + (\mu - 0.5 v_n^+) h_n + \sqrt{v_n^+} X_n \sqrt{h_n} Z_{1,n}, \tag{3.4}$$

$$v_{n+1} = v_n + \left(\delta^2 - 2\beta v_n^+\right) h_n + \sigma \sqrt{v_n^+} \sqrt{h_n} Z_{2,n}, \tag{3.5}$$

and then we take the exponential of the log-prices:

$$S_n = S_0 e^{X_n}. \tag{3.6}$$

However, the scheme is not accurate, since we ignore the $dZ_i dZ_j$ terms in the Itô-Taylor series approximation.

3.2 Broadie-Kaya Scheme

It follows from Heston model that for $t > u$

$$S_t = S_u e^{\left(r(t-u) - \frac{1}{2}\int_u^t v_s\,ds + \rho\int_u^t \sqrt{v_s}\,dZ_1(s) + (1-\rho)\int_u^t \sqrt{v_s}\,dZ_2(s)\right)}, \tag{3.7}$$

$$v_t = v_u + \kappa\theta(t-u) - \kappa\int_u^t v_s\,ds + \sigma\int_u^t \sqrt{v_s}\,dZ_2(s), \tag{3.8}$$

Exact simulation algorithm for the Heston model:

1. Generate a sample from the distribution of v_t given v_u;
2. Generate a sample from the distribution of $\int_u^t V_s ds$ given v_t and v_u;
3. Recover $\int_u^t \sqrt{v_s} dZ_1(s)$ given v_t, v_u, and $\int_t^u v_s ds$;
4. Generate a sample from the distribution of S_t given $\int_u^t \sqrt{v_s} dZ_1(s)$, $\int_u^t \sqrt{v_s} dZ_2(s)$, $\int_u^t v_s ds$.

Step 1: Generate a sample from the distribution of v_t given v_u

As shown in Cox et al. (1985) ADD CITING TO BIB FILE the distribution of v_t given v_u for some $u < t$ is, up to a scale factor, a noncentral chi-squared distribution. The transition law of v_t can be expressed as:

$$v_t = \frac{\sigma^2(1-e^{-\kappa(t-u)})}{4\kappa}\chi_d'^2\left(\frac{4\kappa e^{-\kappa(t-u)}}{\sigma^2(1-e^{-\kappa(t-u)})}v_u\right), \quad t > u, \tag{3.9}$$

where $\chi_d'^2(\lambda)$ denotes the noncentral chi-squared random variable with d degrees of freedom, and noncentrality parameter λ, and

$$d = \frac{4\theta\kappa}{\sigma^2}. \tag{3.10}$$

Thus, we can sample from the distribution of v_t exactly, provided that we can sample from the noncentral chisquared distribution. Johnson et al. (1994) ADD CITING TO BIB FILE show that for $d > 1$, the following representation is valid:

$$\chi_d'^2(\lambda) = \chi_1'^2(\lambda) + \chi_{d-1}'^2 = N(\lambda, 1)^2 + \chi_{d-1}^2. \tag{3.11}$$

Therefore, when $d > 1$, sampling from a noncentral chi-squared distribution is reduced to sampling from an ordinary chi-squared and an independent normal. When $d < 1$ we can use the the fact that

$$\chi_d'^2(\lambda) \sim \chi_{d+2N}^2, \tag{3.12}$$

where N is a Poisson random variable with mean $\frac{\lambda}{2}$.

Step 2: Generate a sample from the distribution of $\int_u^t V_s ds$ given v_t and v_u

The folowing formula can be derived. The derivation could be found in in the original paper. DERIVE HERE

$$\phi(a) = \mathbb{E}\left[\exp\left(ia\int_u^t V_s ds\right)\middle| v_u, v_t\right] = \frac{\gamma(a)e^{-(1/2)(\gamma(a)-\kappa)(t-u)}}{\kappa(1-e^{-\gamma(a)(t-u)})}$$

$$\exp\left(\frac{v_u+v_t}{\sigma^2}\left[\frac{\kappa(1+e^{-\kappa(t-u)})}{1-e^{-\kappa(1-u)}}\right]\right)\frac{I_{0.5d-1}\left(\sqrt{v_u v_t}\frac{4\gamma(a)e^{-0.5\gamma(a)(t-u)}}{\sigma^2(1-e^{-\kappa(a)(t-u)})}\right)}{I_{0.5d-1}\left(\sqrt{v_u v_t}\frac{4\kappa e^{-0.5\kappa(t-u)}}{\sigma^2(1-e^{-\kappa(t-u)})}\right)}, \tag{3.13}$$

where $\gamma(a) = \sqrt{\kappa^2 - 2\sigma i a}$ and $I_{0.5d-1}$ is a modified Bessel function of the first kind.

Let $V(u,t)$ denote the random variable that has the conditional distribution of the integral $\int_u^t V_s ds$ given v_u and v_t. Then we need to invert the characteristic function to get the cumulative distribution function

$$F(x) = \mathbb{P}(V(u,t) \le x) = E\left[e^{iaV(u,t)} \middle| v_u, v_t\right] = \\ = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\sin ux}{u}\Phi(u)du = \frac{2}{\pi}\int_0^{\infty} \frac{\sin ux}{u}\Phi(u)du. \quad (3.14)$$

To calculate the integral the trapezoidal rule is being used, ??????????? because the errors tend to annihilate for periodic and other oscillating integrands

$$\mathbb{P}(V(u,t) \le X) = \frac{hx}{\pi} + \frac{2}{\pi}\sum_{j=1}^{\infty} \frac{\sin hjx}{j}\Re[\Phi(hj)] - e_d(h), \quad (3.15)$$

where h is a grid scale and $e_d(h)$ is the discretization error e_d. It can be bounded above by using a Poisson summation formula:

$$0 \le e_d(h) = \sum_{k=1}^{\infty}\left[F\left(\frac{2k\pi}{h} + x\right) - F\left(\frac{2k\pi}{h} - x\right)\right] \le 1 - F\left(\frac{2\pi}{h} - x\right). \quad (3.16)$$

If we want to achieve a discretization error α, then the step size should be

$$h = 2\frac{2\pi}{x + u_\alpha} \ge \frac{\pi}{u_\alpha}, \quad (3.17)$$

where $1 - F(u_\alpha) = \alpha$ and $0 \le x \le u_\alpha$. To be able to calculate $P(V(u,t) < x)$ using (3.15), we need to determine a point at which the summation can be terminated. Let N represent the last term to be calculated so that the approximation becomes

$$F(x) = \mathbb{P}(V(u,t) \le X) = \frac{hx}{\pi} + \frac{2}{\pi}\sum_{j=1}^{N} \frac{\sin hjx}{j}\Re[\Phi(hj)] - e_d(h) - e_T(N). \quad (3.18)$$

Because $|\sin ux| \le 1$, the integrand in (3.16) is bounded by

$$\frac{2|\Re[\Phi(u)]|}{\pi u} \le \frac{2|\Phi(u)|}{\pi u}. \quad (3.19)$$

To simulate the value of the integral, the Smirnov's transform method is used. We generate a uniform random variable U and then find the value of x for which

$$\mathbb{P}(V(u,t) \le x) = U. \quad (3.20)$$

Step 3: Generate a sample from the distribution of $V(u,t)$ given v_u and v_t

The following formula can be used to calculate $\int_u^t \sqrt{v_s} dZ_1(s)$, as we already generated samples for $v_t, v_u, V(u,t)$

$$\int_u^t \sqrt{v_s} dZ_1(s) = \frac{1}{\sigma}(v_t v_u) - \kappa\theta(t-u) + V(u,t). \quad (3.21)$$

Step 4: Generate a sample from the distribution of $V(u,t)$ given v_u and v_t

Lastly, we need to bring everything together:

- $\int_u^t \sqrt{v_s} dZ_1(s)$ and $\int_u^t \sqrt{v_s} dZ_2(s)$ are already calculated;
- $V(u,t) = \int_u^t v_s ds$ is also calculated.

$$S_t = S_u \exp\left(r(t-u) - \frac{1}{2}V(u,t) + \rho \int_u^t \sqrt{v_s}\, dZ_1(s) + (1-\rho) \int_u^t \sqrt{v_s}\, dZ_2(s) \right) \tag{3.22}$$

3.3 Andersen Scheme

Motivation for these schemes is the following two facts:

- Euler scheme is not very accurate, but fast and easy to implement;
- Broadie-Kaya scheme is more accurate, but significantly slower and way more complicated.

3.3.1 Quadratic-Exponential Discretization Scheme

We denote

$$m = \mathbb{E}\left[\hat{V}(t+\Delta) \middle| \hat{V}(t) \right], \tag{3.23}$$

$$s^2 = \mathbb{E}\left[\left(\hat{V}(t+\Delta) - m \right)^2 \middle| \hat{V}(t) \right], \tag{3.24}$$

$$\psi = \frac{s^2}{m^2}. \tag{3.25}$$

Andersen proposes an approximation based on moment-matching techniques. His goal is then to speed up the first step of Broadie and Kaya's method. He observes that the conditional distribution of $\hat{V}(t+\Delta)$ given $\hat{V}(t)$ visually difers when $\hat{V}(t)$ is small or large (in the variation coefficient sense). The scheme is constructed from the following two subschemes:

1. Quadratic sampling scheme ($\psi \leq 2$);
2. Exponential sampling scheme ($\psi \geq 1$).

Fortunately, these two intervals cover the whole positive real line. Furthermore, these two schemes could be applied at the same time when $\psi \in [1,2]$. This implies that there exists a critical value $\psi_{\text{crit}} \in [1,2]$, which could be an indicator of which scheme is more applicable at the given value of ψ. Let us show you this.

Quadratic Sampling Scheme

For large enough $\hat{V}(t)$ (in the CV-sense) we can approximate the distribution of $\hat{V}(t+\Delta)$ by the scaled non-central chi-squared distribution with 1 degree of freedom:

$$\text{Law}\left(\hat{V}(t+\Delta) \middle| \hat{V}(t) \right) = a(\Delta, \hat{V}(t), VP)\chi_1'^2(b(\Delta, \hat{V}(t), VP)), \tag{3.26}$$

where VP is the vector of parameters of the CIR variance.

Lemma 6. *We have*

$$b^2 = \frac{2}{\psi} - 1 + \sqrt{\frac{2}{\psi}\left(\frac{2}{\psi} - 1\right)}, \tag{3.27}$$

$$a = \frac{m}{1+b^2}. \tag{3.28}$$

Proof. Plain equating of the theoretical and real moments. □

Remark. *The above lemma is not valid for $\psi \geq 2$.*

Therefore, if $\hat{V}(t)$ is close to zero, then we have a problem in finding such $a = a(\Delta, \hat{V}(t), VP)$ and $b = b(\Delta, \hat{V}(t), VP)$ such that the moments of the desired conditional distribution could be properly matched.

Exponential Sampling Scheme

Therefore, we approximate the desired distribution with the following method. Let ξ and η be independent random variables and $\xi \sim Be(1-p)$, $\eta \sim Exp(\beta)$ for some $p \in (0,1)$ and $\beta > 0$. Then we have (given $\hat{V}(t)$)

$$\hat{V}(t+\Delta) = \xi \cdot \eta, \tag{3.29}$$

what gives us the following distribution density:

$$p_{\hat{V}(t+\Delta)|\hat{V}(t)} = p \cdot \delta(x) + (1-p) \cdot \beta e^{-\beta x}, \tag{3.30}$$

where $\delta(x)$ is a standart delta function and for some β and p. Sampling ξ and η: Smirnov's transform. Or we can use the Smirnov transform with the cdf of the desired distribution.

Lemma 7. *We have*

$$p = \frac{\psi - 1}{\psi + 1}, \qquad \beta = \frac{1-p}{m} = \frac{2}{m(\psi+1)}. \tag{3.31}$$

Proof. By direct integration of the given densities we get the following:

$$\frac{1-p}{\beta} = m, \qquad \frac{1-p^2}{\beta^2} = s^2. \tag{3.32}$$

□

Remark. *The above lemma is not valid for $\psi \leq 1$.*

3.3.2 Truncated Gaussian Discretization Scheme

The main idea of the method: in this scheme the idea is to sample from a moment-matched Gaussian density where all probability mass below zero is inserted into a delta-function at the origin. Formalization of the idea:

$$\left(\hat{V}(t+\Delta)\Big| V(t)\right) = (\mu + \sigma Z)^+, \tag{3.33}$$

where Z is a standard normal random variable and μ and σ are the 'mean' and the 'standard deviation' of the desired distribution. We find μ and σ from the moment-matching techniques (see the previous method, equations (3.23) – (3.25)).

Lemma 8. *Let $\phi(x)$ be a standart Gaussian density and define a function $r : \mathbb{R} \to \mathbb{R}$ by the following equation:*

$$r(x)\phi(r(x)) + \Phi(r(x))(1 + r(x)^2) = (1 + x)\left(\phi(r(x)) + r(x)\Phi(r(x))\right)^2. \tag{3.34}$$

Then the moment-matching parameters are

$$\mu = \frac{m}{\frac{\phi(r(\psi))}{r(\psi)} + \Phi(r(\psi))}, \tag{3.35}$$

$$\sigma = \frac{m}{\phi(r(\psi)) + r(\psi)\Phi(r(\psi))}. \tag{3.36}$$

Proof. PROOF HERE □

Problem: no closed-form solution for $r(\psi)$.
Solution: numerical solution.
Problem: no known limits to use the numerical solution.
Solution:

$$m = \frac{\delta^2}{2\beta} + \left(\hat{V}(t) - \frac{\delta^2}{2\beta}\right)e^{-2\beta\Delta}, \tag{3.37}$$

$$s^2 = \frac{\hat{V}(t)\sigma^2 e^{-2\beta\Delta}}{2\beta}\left(1 - e^{-2\beta\Delta}\right) + \frac{\delta^2\sigma^2}{8\beta^2}\left(1 - e^{-2\beta\Delta}\right)^2. \tag{3.38}$$

Then we analyze ψ wrt $\hat{V}(t)$ and obtain a finite interval as a domain for $r(\psi)$.

Proof. PROOF HERE. Redo as a lemma

□

Part II

Implementation Problems and Pricing Exotics

Chapter 4

Implementation of the Methods

4.1 Euler Scheme

4.2 E+M Scheme

4.3 Broadie-Kaya Scheme

4.4 Andersen Scheme

Chapter 5

Comparison of the Methods

5.1 Performance

5.2 Accuracy

Chapter 6

Pricing Exotics

Conclusion

Bibliography

[BS73] Fischer Black and Myron Sholes. “The Pricing of Options and Corporate Liabilities”. In: *Journal of Political Economy* 81.3 (1973), pp. 637–657.

[Hes93] Steven L. Heston. “A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options”. In: *Review of Financial Studies* 6.2 (1993), pp. 327–343.

[Gat12] Jim Gatheral. *The Volatility Surface*. John Wiley & Sons, Ltd, 2012. Chap. 1-3, pp. 1–42. ISBN: 9781119202073. DOI: https://doi.org/10.1002/9781119202073.