Student Research Group ’Stochastic Volatility Models’, Project ’Heston−2’

Methods of Simulation of the Heston Model: A Review

Artemy Sazonov, Danil Legenky, Kirill Korban

Lomonosov Moscow State Univesity, Faculty of Mechanics and Mathematics

November 5, 2022

Heston Model Definition

Assume that the spot asset at time t follows the diffusion

$$dS(t) = \mu S(t)dt + \sqrt{v(t)}S(t)dZ_1(t), \tag{1}$$

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

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

Table of Contents

Random and Pseudo-Random Numbers

A. N. Kolmogorov – «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.

Random and Pseudo-Random Numbers

Collins Dictionary Definitions

Definition 1 (Random numbers)

a sequence of numbers that do not form any progression, used to facilitate unbiased sampling of a population.

Definition 2 (Pseudorandom numbers)

a sequence of numbers that satisfies statistical tests for randomness but is produced by a reproducible mathematical procedure.

Law of Large Numbers

Theorem 3 (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

$$\mathbb{P}\text{-}\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} X_i = \mu. \tag{3}$$

Theorem 4 (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

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} X_i \overset{a.s.}{=} \mu. \tag{4}$$

Central Limit Theorem

Theorem 5 (Lindeberg-Lévy)

Let $X_1, \ldots, X_n$ be a sequence of i.i.d. random variables with $\mathbb{E}[X_i] = \mu$ and $\mathrm{var}\left[X_i\right] = \sigma^2$. 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.

$$\sqrt{n}\left(\bar{X}_n - \mu\right) \xrightarrow{d} \mathcal{N}\left(0, \sigma^2\right). \tag{5}$$

NB

Law of large numbers is an informal corollary of the central limit theorem.

Monte Carlo Simulation

Statistical Estimation

Lemma 6

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

$$\widehat{\mathbb{E}\left[h(X)\right]} = \frac{1}{n} \sum_{i=1}^{n} h(X_i). \tag{6}$$

Monte Carlo Simulation

Local Truncation Error

Definition 7

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 $\hat{\mu} = \widehat{\mathbb{E}[X]}$ at the confidence level α:

$$\mu \in \left(\hat{\mu} - z_{\alpha/2}\sqrt{\frac{\sigma^2}{n}}, \hat{\mu} + z_{\alpha/2}\sqrt{\frac{\sigma^2}{n}} \right). \tag{7}$$

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

Table of Contents

Forward Euler Scheme for ODEs

Definition

Suppose that we have an ODE of the form

$$dX(t) = f(X(t), t)dt, \quad X(0) = X_0. \tag{8}$$

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

$$X_{n+1} = X_n + f(t_n, X_n)h_n, \tag{9}$$

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

Forward Euler Scheme for ODEs

Global truncation error

Lemma 8

Let the solution of (8) have an M-bounded second derivative and let f be L-Lipshitz continious in its second argument. Then the global truncation error of the mesh solution (9) is

$$|X(T) - X_N| \leq \frac{hM}{2L}\left(e^{LT} - 1\right) = O(h). \tag{10}$$

Euler-Maruyama Scheme for SDEs

Definition

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:

$$X_{n+1} = X_n + f(t_n, X_n)h_n + \sigma(t_n, X_n)\sqrt{h_n}Z_n, \tag{11}$$

where $(Z_n)_{n=1,2,\dots}$ 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 $(t_i)_{i=0,1,\dots}$ is a uniform grid with $t_i = ih$.

Euler-Maruyama Scheme for SDEs

Strong and weak convergence as a global truncation error analogue

Definition 9

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

$$\mathbb{E}\left[\left|\hat{X}^n(T) - X(T)\right|\right] \leq Ch^p, \quad n \to \infty. \tag{12}$$

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

$$\left|\mathbb{E}\left[f(\hat{X}^n(T))\right] - \mathbb{E}\left[f(X(T))\right]\right| \leq Ch^p, \quad n \to \infty. \tag{13}$$

Euler-Maruyama Scheme for SDEs

Strong and weak convergence as a global truncation error analogue

Theorem 10

Under some technical assumptions the Euler-Maruyama scheme (11) has a strong convergence of order $1/2$ and a weak convergence of order 1.

NB

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.

Euler Scheme for the Heston Model

Classical Euler-Maruyama Scheme

Suppose we have the Heston model (1) – (2). Then it could be numerically solved by the following finite difference scheme:

$$S_{n+1} = S_n + \mu S_n h_n + \sqrt{v_n} S_n \sqrt{h_n} Z_{1,n}, \tag{14}$$

$$v_{n+1} = v_n + \left(\delta^2 - 2\beta v_n\right) h_n + \sigma \sqrt{v_n} \sqrt{h_n} Z_{2,n}, \tag{15}$$

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 grid.

Euler Scheme for the Heston Model

Modified Euler-Maruyama Scheme

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

$$X(t) := \log \frac{S(t)}{S(0)}. \tag{16}$$

Euler Scheme for the Heston Model

Modified Euler-Maruyama Scheme

We take the positive part of the variance:

$$X_{n+1} = X_n + (\mu - 0.5v_n^+)h_n + \sqrt{v_n^+}X_n\sqrt{h_n}Z_{1,n}, \tag{17}$$

$$v_{n+1} = v_n + \left(\delta^2 - 2\beta v_n^+\right)h_n + \sigma\sqrt{v_n^+}\sqrt{h_n}Z_{2,n}, \tag{18}$$

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

$$S_n = S_0 e^{X_n}. \tag{19}$$

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

Table of Contents

Broadie-Kaya Simulatation Method

Definition

It folows from Heston model, that for $t > u$

$$S_t = S_u e^{\left(r(t-u) - \frac{1}{2}\int_u^t v_s\, ds + \rho \int_u^t \sqrt{v_s}\, dZ_1(s) + (1-\rho)\int_u^t \sqrt{v_s}\, dZ_2(s)\right)}, \tag{20}$$

$$v_t = v_u + \kappa\theta(t-u) - \kappa \int_u^t v_s\, ds + \sigma \int_u^t \sqrt{v_s}\, dZ_2(s), \tag{21}$$

Exact Simulation Algorithm for the SV Model

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

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

Step 3 Recover $\int_u^t \sqrt{v_s} dZ_1(s)$ given v_t, v_u, and $\int_t^u v_s ds$

Step 4 Generate a sample from the distribution of S_t given $\int_u^t \sqrt{v_s} dZ_1(s)$ and $\int_u^t v_s ds$

Broadie-Kaya Simulatation Method

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

As shown in Cox et al. (1985) 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:

$$v_t = \frac{\sigma^2(1 - e^{-\kappa(t-u)})}{4\kappa} \chi_d'^2 \left(\frac{4\kappa e^{-\kappa(t-u)}}{\sigma^2(1 - e^{-\kappa()t-u})} v_u \right), t > u, \tag{22}$$

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

$$d = \frac{4\theta\kappa}{\sigma^2} \tag{23}$$

Broadie-Kaya Simulatation Method

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

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) show that for $d > 1$, the following representation is valid:

$$\chi_d'^2(\lambda) = \chi_1'^2(\lambda) + \chi_{d-1}'^2 = N(\lambda, 1)^2 + \chi_{d-1}^2 \tag{24}$$

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:

$$\chi_d'^2(\lambda) \sim \chi_{d+2N}^2 \tag{25}$$

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

Broadie-Kaya Simulatation Method

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

$$\phi(a) = \mathbb{E}\left[\exp ia \int_u^t V_s ds \middle| v_u, v_t\right] = \frac{\gamma(a)e^{-(1/2)(\gamma(a)-\kappa)(t-u)}}{\kappa(1-e^{-\gamma(a)(t-u)})} \tag{26}$$

$$\exp \frac{v_u + v_t}{\sigma^2}\left[\frac{\kappa(1+e^{-\kappa(t-u)})}{1-e^{-\kappa(1-u)}}\right] \tag{27}$$

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:

$$\chi_d'^2(\lambda) \sim \chi^2_{d+2N} \tag{28}$$

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

Table of Contents

Quadratic-Exponential Discretization Scheme

Motivation

We denote

$$m = \mathbb{E}\left[\hat{V}(t+\Delta)\middle|\hat{V}(t)\right], \tag{29}$$

$$s^2 = \mathbb{E}\left[\left(\hat{V}(t+\Delta) - m\right)^2\middle|\hat{V}(t)\right], \tag{30}$$

$$\psi = \frac{s^2}{m^2}. \tag{31}$$

Quadratic-Exponential Discretization Scheme

Motivation

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 implicates that there exist some 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-Exponential Discretization Scheme

The Problem with Variance

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:

$$\mathrm{Law}\left(\hat{V}(t+\Delta)\,\middle|\,\hat{V}(t)\right) = a(\Delta, \hat{V}(t), VP)\chi_1'^2(b(\Delta, \hat{V}(t), VP)), \tag{32}$$

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

Quadratic-Exponential Discretization Scheme

The Problem with Variance

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 $\hat{V}(t)$)

$$\hat{V}(t+\Delta) = \xi \cdot \eta, \tag{33}$$

what gives us the following distribution density:

$$p_{\hat{V}(t+\Delta)|\hat{V}(t)} = p \cdot \delta(x) + (1-p) \cdot \beta e^{-\beta x}, \tag{34}$$

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.

Quadratic-Exponential Discretization Scheme

Finding the constants

Lemma 11

We have

$$b^2 = \frac{2}{\psi} - 1 + \sqrt{\frac{2}{\psi}\left(\frac{2}{\psi} - 1\right)}, \tag{35}$$

$$a = \frac{m}{1 + b^2}. \tag{36}$$

Proof.

Plain equating of the theoretical and real moments. □

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

Quadratic-Exponential Discretization Scheme

Finding the constants

Lemma 12

We have

$$p = \frac{\psi - 1}{\psi + 1}, \qquad \beta = \frac{1 - p}{m} = \frac{2}{m(\psi + 1)}. \tag{37}$$

Proof.

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

$$\frac{1 - p}{\beta} = m, \qquad \frac{1 - p^2}{\beta^2} = s^2. \tag{38}$$

□

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

Table of Contents

Table of Contents

Table of Contents

Conclusion

We introduced the three most common simulation methods for dynamics of the two-factor Gaussian diffusion model:

1. Euler scheme;
2. Broadie-Kaya scheme;
3. Andersen scheme.