Heston

Nodari Alessandro & Proserpio Lorenzo

Let $(\Omega, \{(F_t)_{t\geq 0}\}, \mathbb{P})$ a complete filtered probability space and let call $\mathbb{P}$ the *physical-measure*. Let $T < \infty$ be the right limit of our time horizon from now on.

1 The Heston Model

Given a stock price process $S = (S_t)_{t\geq 0}$ the Heston model, under $\mathbb{P}$, is the following :

$$\begin{cases} \mathrm{d}S_t = \mu S_t \mathrm{d}t + S_t\sqrt{v_t}\mathrm{d}W_t \\ \mathrm{d}v_t = \kappa(\eta - v_t)\mathrm{d}t + \theta\sqrt{v_t}\mathrm{d}\tilde{W}_t \\ v_0 = \sigma_0^2 \end{cases}$$

where:

- μ is the drift of the stock returns;
- $W = (W_t)_{t\geq 0}$ and $\tilde{W} = (\tilde{W}_t)_{t\geq 0}$ are two correlated Brownian motions with $\mathrm{d}\langle W, \tilde{W}\rangle_t = \rho \mathrm{d}t$ and $\rho \in [-1, 1]$;
- $\sigma_0 > 0$ is the initial volatility;
- $\eta > 0$ is the long run variance;
- $\kappa > 0$ is the mean reversion rate;
- $\theta > 0$ is the volatility of the volatility.

The variance process is strictly positive if $2\kappa\eta > \theta^2$ (condition only sufficient, not necessary). This is known as *Feller condition*.

2 The Valuation Equation

Consider two independent claims which prices, at time t, are given as $U_t = u(t, S_t, v_t)$ and $\tilde{U}_t = \tilde{u}(t, S_t, v_t)$. Suppose that u and $\tilde{u}$ are $\mathcal{C}^1(\mathbb{R}^+)$ w.r.t. the first variable and $\mathcal{C}^2(\mathbb{R}^+ \times \mathbb{R}^+)$ w.r.t. the last two variables. Consider a portfolio consisting of a long contract U, short Δ shares of the stock and short Δ_1 contracts of $\tilde{U}$. So the value (denoted with Π_t) of our portfolio at time t is equal to:

$$\Pi_t = U_t - \Delta S_t - \Delta_1 \tilde{U}_t$$

then we can apply Itô's lemma and write the dynamics, omitting some subscripts, of the value of the portfolio as:

$$\begin{aligned}
\mathrm{d}\Pi_t = &\left\{\frac{\partial U}{\partial t} + \frac{1}{2}vS^2\frac{\partial^2 U}{\partial S^2} + \rho\theta v S\frac{\partial^2 U}{\partial S \partial v} + \frac{1}{2}\theta^2 v\frac{\partial^2 U}{\partial v^2}\right\}\mathrm{d}t \\
&+ \left\{\frac{\partial \tilde{U}}{\partial t} + \frac{1}{2}vS^2\frac{\partial^2 \tilde{U}}{\partial S^2} + \rho\theta v S\frac{\partial^2 \tilde{U}}{\partial S \partial v} + \frac{1}{2}\theta^2 v\frac{\partial^2 \tilde{U}}{\partial v^2}\right\}\mathrm{d}t \\
&+ \left\{\frac{\partial U}{\partial S} - \Delta_1\frac{\partial \tilde{U}}{\partial S} - \Delta\right\}\mathrm{d}S \\
&+ \left\{\frac{\partial U}{\partial v} - \Delta_1\frac{\partial \tilde{U}}{\partial v}\right\}\mathrm{d}v
\end{aligned}$$

to make our portfolio instantaneously risk-free we must impose the following equations to eliminate the $\mathrm{d}S$ and $\mathrm{d}v$ terms:

$$\begin{cases} \dfrac{\partial U}{\partial S} - \Delta_1\dfrac{\partial \tilde{U}}{\partial S} - \Delta = 0 \\ \dfrac{\partial U}{\partial v} - \Delta_1\dfrac{\partial \tilde{U}}{\partial v} = 0 \end{cases} \tag{1}$$

since our portfolio is now risk-free its value must be equal to the value of a portfolio with the same initial value invested in the cash market at risk-free rate r, so:

$$\mathrm{d}\Pi_t = r\Pi_t\mathrm{d}t = r(U_t - \Delta S_t - \Delta_1\tilde{U}_t)\mathrm{d}t$$

then choosing Δ and Δ_1 as in (1) and collecting all the U terms on the LHS and all the $\tilde{U}$ terms on the RHS we obtain (omitting once again some subscripts):

$$\begin{aligned}
&\frac{\frac{\partial U}{\partial t} + \frac{1}{2}vS^2\frac{\partial^2 U}{\partial S^2} + \rho\theta v S\frac{\partial^2 U}{\partial S\partial v} + \frac{1}{2}\theta^2 v\frac{\partial^2 U}{\partial v^2} + rS\frac{\partial U}{\partial S} - rU}{\frac{\partial U}{\partial v}} \\
&\qquad = \frac{\frac{\partial \tilde{U}}{\partial t} + \frac{1}{2}vS^2\frac{\partial^2 \tilde{U}}{\partial S^2} + \rho\theta v S\frac{\partial^2 \tilde{U}}{\partial S\partial v} + \frac{1}{2}\theta^2 v\frac{\partial^2 \tilde{U}}{\partial v^2} + rS\frac{\partial \tilde{U}}{\partial S} - r\tilde{U}}{\frac{\partial \tilde{U}}{\partial v}}
\end{aligned}$$

now since the LHS depends on U and the RHS depends on $\tilde{U}$ then the only way for which this is possible is if they are equal to some function $f = f(t, S_t, v_t)$ which is independent from U and $\tilde{U}$. So we obtain:

$$\frac{\partial U}{\partial t} + \frac{1}{2}vS^2\frac{\partial^2 U}{\partial S^2} + \rho\theta v S\frac{\partial^2 U}{\partial S \partial v} + \frac{1}{2}\theta^2 v\frac{\partial^2 U}{\partial v^2} + rS\frac{\partial U}{\partial S} - rU = -f\frac{\partial U}{\partial v} \tag{2}$$

Equation (2) is called the valuation equation for U. WLOG we can rewrite f as:

$$f(t, S_t, v_t) = \kappa(\eta - v_t) - \sqrt{v_t}\phi(t, S_t, v_t)$$

where ϕ_t is some arbitrary $\mathbb{F}$-adapted function and it is called the *market price of volatility risk*.

NOTE: the stock S is assumed to have yield equals to zero. If the stock in consideration has yield q we can simply switch r with $(r-q)$ in the calculations before.

2.1 The market price of volatility risk

Without assuming that there are two traded independent claims with one of which is dependent from the variance process we cannot a priori fix an $\mathbb{F}$-adapted ϕ. Indeed, we can rewrite $\tilde{W} = \rho W + \sqrt{1-\rho^2} W^{\perp}$ where $W^{\perp}$ is a continuous Brownian motion on the filtration $\mathcal{F}$ such that $\mathrm{d}\langle W, W^{\perp}\rangle = 0$. Then we can rewrite the Heston model under $\mathbb{P}$ as:

$$\begin{cases} \mathrm{d}S_t = \mu S_t \mathrm{d}t + S_t\sqrt{v_t}\mathrm{d}W_t \\ \mathrm{d}v_t = \kappa(\eta - v_t)\mathrm{d}t + \theta\rho\sqrt{v_t}\mathrm{d}W_t + \theta\sqrt{1-\rho^2}\sqrt{v_t}\mathrm{d}W_t^{\perp} \end{cases}$$

now take $\lambda_t = \lambda(t, S_t, v_t)$ defined as:

$$\lambda_t := \frac{\mu - r}{\sqrt{v_t}}$$

and take ϕ_t $\mathbb{F}$-adapted such that the Novikov's condition is satisfied, i.e.;

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^T [\lambda_t^2 + \phi_t^2]\mathrm{d}t\right)\right] < \infty$$

then we have that the process $Z = (Z_t)_{t\in[0,T]}$defined as:

$$Z_t = \mathcal{E}(-\lambda_t W - \phi_t W^{\perp})_t$$

where $\mathcal{E}$ is the Doléans-Dade exponential is a $\mathbb{P}$-martingale, with $E[Z_T] = 1$ and strictly positive. So, using the Radon-Nikodym's theorem we can define a measure $\mathbb{Q}$ on $\mathbb{F}$ as:

$$\left.\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\right|_{F_t} := Z_t$$

and we have that $\mathbb{Q}$ is equivalent to $\mathbb{P}$ on F_T and moreover is martingale equivalent measure. By Girsanov's theorem we can define a 2-dimensional $\mathbb{Q}$-Brownian motion $(W^{\mathbb{Q}}, W^{\mathbb{Q},\perp})$ as:

$$W_t^{\mathbb{Q}} = W_t + \int_0^t \lambda_s \mathrm{d}s \quad \text{and} \quad W_t^{\mathbb{Q},\perp} = W_t + \int_0^t \phi_s \mathrm{d}s$$

and under such a $\mathbb{Q}$ we have that the Heston model has the following dynamics:

$$\begin{cases} \mathrm{d}S_t = rS_t\mathrm{d}t + S_t\sqrt{v_t}\mathrm{d}W_t^{\mathbb{Q}} \\ \mathrm{d}v_t = [\kappa(\eta - v_t) - \theta\sqrt{v_t}(\rho\lambda_t + \sqrt{1-\rho^2}\phi_t)]\mathrm{d}t + \theta\sqrt{v_t}\mathrm{d}\tilde{W}_t^{\mathbb{Q}} \\ v_0 = \sigma_0^2 \end{cases}$$

and under this we have that the discounted stock price is a $\mathbb{Q}$ local martingale. However, we can choose any arbitrary ϕ and obtain another $\tilde{\mathbb{Q}}$ with the same properties. But, as in our case, both the claims are actively traded we can infer the ϕ from the market prices of the options, which will determine a $\mathbb{Q}$ and then fix the ϕ. Among all the possible choices from now on we will choose $\phi = 0$. The resulting $\mathbb{Q}^M$ is called the *minimal martingale measure* and Föllmer and Schweizer have proven that, for models with continuous price trajectories, solves the minimization problem:

$$\mathbb{Q}^M = \arg\min_{\mathbb{Q}\in\mathcal{M}} \mathbb{H}(\mathbb{Q}|\mathbb{P})$$

where $\mathcal{M}$ is the set of equivalent martingale measures and $\mathbb{H}$ is the reverse relative entropy. Indeed this is exactly what we obtain if we can completely infer the ϕ from the market prices of the options (because we think that $\mathbb{P}$ contains all the information to reconstruct $\mathbb{Q}$).

NOTE: From now on, to lighten the notation, our Heston model under $\mathbb{Q}$ is written in the following way:

$$\begin{cases} \mathrm{d}S_t = (r-q)S_t\mathrm{d}t + S_t\sqrt{v_t}\mathrm{d}W_t \\ \mathrm{d}v_t = \kappa(\eta - v_t)\mathrm{d}t + \theta\sqrt{v_t}\mathrm{d}\tilde{W}_t \\ v_0 = \sigma_0^2 \end{cases}$$

with $W = (W_t)_{t\geq 0}$ and $\tilde{W} = (\tilde{W}_t)_{t\geq 0}$ are two correlated $\mathbb{Q}$ Brownian motions with $\mathrm{d}\langle W, \tilde{W}\rangle_t = \rho\mathrm{d}t$.

3 The Characteristic Formula

3.1 Derivation of the pseudo-probabilities

Setting the $\phi = 0$ we have that the valuation equation for our model, to price a call option C, becomes:

$$\frac{\partial C}{\partial t} + \frac{1}{2}vS^2\frac{\partial^2 C}{\partial S^2} + \rho\theta vS\frac{\partial^2 C}{\partial S\partial v} + \frac{1}{2}\theta^2 v\frac{\partial^2 C}{\partial v^2} + rS\frac{\partial C}{\partial S} - rC + [\kappa(\eta - v)]\frac{\partial C}{\partial v} = 0$$

if we now substitute with $\tau = T - t$ and $x = \log\left(\frac{S_t e^{(r-q)\tau}}{K}\right)$ in the previous equation we obtain:

$$-\frac{\partial C}{\partial \tau} + \frac{1}{2}v\frac{\partial^2 C}{\partial x^2} + \rho\theta v\frac{\partial^2 C}{\partial x\partial v} + \frac{1}{2}\theta^2 v\frac{\partial^2 C}{\partial v^2} - \frac{1}{2}v\frac{\partial C}{\partial x} + [\kappa(\eta - v)]\frac{\partial C}{\partial v} = 0 \quad (3)$$

according to Duffie, Pan and Singleton (2000), the solution to this equation has the form of:

$$C(x, v, \tau) = K\left\{e^x P_1(x, v, \tau) - P_0(x, v, \tau)\right\} \quad (4)$$

notice how this remind us to the solution of the Black&Scholes' Equation. Moreover, the P_j are absolutely continuous, differentiable and both P_j and P_j' are absolutely integrable on $\mathbb{R}$. Putting (4) into (3) we obtain that P_j with $j = 0, 1$ must satisfy the following PDE:

$$-\frac{\partial P_j}{\partial \tau} + \frac{1}{2} v \frac{\partial^2 P_j}{\partial x^2} - \left(\frac{1}{2} - j\right) v \frac{\partial P_j}{\partial x} + \frac{1}{2}\theta^2 v \frac{\partial^2 P_j}{\partial v^2} + \rho\theta v \frac{\partial^2 P_j}{\partial x \partial v} + [\kappa(\eta+1) - j\rho\theta v] \frac{\partial P_j}{\partial v} = 0$$

with boundary conditions:

$$\lim_{\tau \to 0} P_j(x, v, \tau) = \begin{cases} 1 & x > 0 \\ 0 & x \leq 0 \end{cases}$$

Now take the Fourier transform of P_j as:

$$\hat{P}_j(u, v, \tau) = \int_{\mathbb{R}} e^{-iux} P_j(x, v, \tau) \mathrm{d}x \quad \text{then} \quad \hat{P}_j(u, v, 0) = \frac{1}{iu}$$

and the inverse transform given by:

$$P_j(x, v, \tau) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-iux} \hat{P}_j(u, v, \tau) \mathrm{d}u$$

thanks to the regularity of P_j we have that $[\widehat{P_j(u)'}] = iu\hat{P}_j(u)$. So, using this property, we can rewrite our equation as:

$$-\left[\frac{1}{2}u^2 - \left(\frac{1}{2} - j\right) iu\right] v\hat{P}_j + \left\{\rho\theta iuv + [\kappa(\eta+1) - j\rho\theta v]\right\} \frac{\partial \hat{P}_j}{\partial v} + \frac{1}{2}\theta^2 v \frac{\partial^2 \hat{P}_j}{\partial v^2} = \frac{\partial \hat{P}_j}{\partial \tau}$$

now take two arbitrary functions $f(u, \tau)$, $g(u, \tau)$. The solution should be of the form:

$$\hat{P}_j(u, v, \tau) = \frac{1}{iu} \exp\{f(u, \tau)\eta + g(u, \tau)v\}$$

it follows that:

$$\begin{aligned} \frac{\partial \hat{P}_j}{\partial \tau} &= \left\{\eta \frac{\partial f}{\partial \tau} + v \frac{\partial g}{\partial \tau}\right\} \hat{P}_j \\ \frac{\partial \hat{P}_j}{\partial v} &= g\hat{P}_j \\ \frac{\partial^2 \hat{P}_j}{\partial v^2} &= g^2 \hat{P}_j \end{aligned}$$

now define:

$$\begin{aligned} \alpha &= -\frac{u^2}{2} - \frac{iu}{2} + iju \\ \beta &= \kappa - \rho\theta(j + iu) \\ \gamma &= \frac{\theta^2}{2} \end{aligned}$$

then it must be:

$$\begin{aligned}\frac{\partial f}{\partial \tau} &= \kappa g \\ \frac{\partial g}{\partial \tau} &= \alpha - \beta g - \gamma g^2 \\ &= \gamma(g - r_+)(g - r_-)\end{aligned}$$

where $r_\pm = \frac{\beta \pm \sqrt{\beta^2 - 4\alpha\gamma}}{2\gamma} =: \frac{\beta \pm d}{\theta^2}$. Integrating and putting the terminal condition $f(u,0) = 0$ and $g(u,0) = 0$, we obtain:

$$\begin{aligned}f(u,\tau) &= \kappa\left\{r_-\tau - \frac{2}{\theta^2}\log\left(\frac{1 - \frac{r_-}{r_+}e^{-d\tau}}{1 - \frac{r_-}{r_+}}\right)\right\} \\ g(u,\tau) &= r_-\frac{1 - e^{-d\tau}}{1 - \frac{r_-}{r_+}e^{-d\tau}}\end{aligned} \tag{5}$$

Taking the inverse transform on $\hat{P}_j$ we obtain finally the form of the pseudo probabilities as:

$$P_j(x, v, \tau) = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \Re\left\{\frac{\exp\{f_j(u,\tau)\eta + g_j(u,\tau)v + iux\}}{iu}\right\}\mathrm{d}u \tag{6}$$

3.2 Numerical integration of the complex logarithm

In (5) we decided to take r_- to define f, but we could have taken r_+ and everything would have been "almost" the same (this definition coincides with our previous one only if the imaginary part of the complex logarithm is chosen so that $f(u,\tau)$ is continuous with respect to u). However, we are interested in integrating numerically the characteristic function of the Heston model. We have to keep in mind that the complex logarithm has a branch (we decided to take as branch the semiaxis of negative real numbers in the complex plane), so in order to avoid any kind of discontinuity we want that the characteristic function never crosses the negative real axis on $(0,\infty)$. Albrecher, Mayer, Schoutens, Tistaert (2006) proved the following result when using the FFT-like approach:

Proposition 1. *Whenever the parameters of the Heston model are such that* $\Im\{d(u)\} := \Im\{\sqrt{(\rho\theta ui - \kappa)^2 + \theta^2(iu + u^2)}\} \neq 0$ *and* $2\kappa\eta \neq \theta^2 n$ *(where* $n \in \mathbb{N}$*), then defining (5) using* r_+ *leads to numerical instabilities for sufficiently large maturities.*

In order to use methods which leverage the frequency domain we usually have to evaluate the characteristic function in $u - (\alpha + 1)i$ for positive u. So they also proved the following proposition:

Proposition 2. *Denote with* $\phi(u)$ *the characteristic function of the Heston model obtained using* r_- *in (5). Then* $\forall\alpha > 0$ *and* $\forall u \in (0,\infty)$ *the function* $\phi(u - (\alpha + 1)i)$ *does not cross the negative real axis.*

In conclusion if we use r_- in (5) then the characteristic function that we will obtain in the next paragraph is suitable for numerical integration.

3.3 Derivation of the characteristic function

By definition the characteristic function under $\mathbb{Q}$ is defined as:

$$\phi(u) = \mathbb{E}_{\mathbb{Q}}\left[e^{iux_T}\middle| x_t = \log\left(\frac{S_t e^{(r-q)\tau}}{K}\right); v_t\right]$$

now thanks to (6) we know that:

$$\mathbf{Pr}_{\mathbb{Q}}(x_T > x_t) = P_0(x_t, v_t, \tau)$$

so, if we define $k = -x_t$ the density is:

$$\begin{aligned} p(k) &= -\frac{\partial P_0}{\partial k} \\ &= \frac{1}{2\pi}\int_{\mathbb{R}} \exp\left\{f_0(s,\tau)\eta + g_0(s,\tau)v_t - isk\right\}\mathrm{d}s \end{aligned}$$

then we have that, using Fubini theorem:

$$\begin{aligned} \phi_H(u;\tau,x_t,v_t) &= \int_{\mathbb{R}} e^{iuk+iux_t} p(k)\mathrm{d}k \\ &= \frac{1}{2\pi}\int_{\mathbb{R}}\int_{\mathbb{R}} \exp\left\{f_0(s,\tau)\eta + g_0(s,\tau)v_t\right\} e^{iuk} e^{-isk} e^{iux_t}\mathrm{d}s\mathrm{d}k \\ &= \frac{e^{iux_t}}{2\pi}\int_{\mathbb{R}} \exp\left\{f_0(s,\tau)\eta + g_0(s,\tau)v_t\right\}\int_{\mathbb{R}} e^{ik(u-s)}\mathrm{d}k\mathrm{d}s \\ &= e^{iux_t}\int_{\mathbb{R}} \exp\left\{f_0(s,\tau)\eta + g_0(s,\tau)v_t\right\}\delta(u-s)\mathrm{d}s \\ &= \exp\left\{f_0(u,\tau)\eta + g_0(u,\tau)v_t + iu\log\left[S_t e^{(r-q)\tau}\right] - iu\log(K)\right\} \end{aligned}$$

4 Fourier Cosine Expansion for Vanilla Options

In order to calibrate our model we need a way to price vanilla options (calls & puts) which usually require integrating the probability density function. However, since we have its Fourier transform (the characteristic function) we can leverage the *Fourier Cosine Expansion* method developed by F. Fang & C.W. Osterlee. The computational speed, especially for plain vanilla options, makes this integration method state-of-the-art for calibration at financial institutions. From now on, in this section, $t_0 = 0$, the final time is T, we will call the Heston transition density function f_H and K is always used for the strike.

4.1 Inverse Fourier integral via cosine expansion

Remember that:

$$\phi_H(u) = \int_{\mathbb{R}} e^{iux} f_H(x)\mathrm{d}x$$

we are interested in approximating:

$$f_H(x) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-iux}\phi_H(u)\mathrm{d}u$$

Since $f_H \in \mathcal{L}^1(\mathbb{R})$ and it is ≥ 0 we are able to choose a finite interval $[a, b]$ such that we have:

$$\tilde{\phi}_H(u) = \int_{\mathbb{R}} e^{iux} f_H(x)\mathbb{1}_{[a,b]}(x)\mathrm{d}x = \int_a^b e^{iux} f_H(x)\mathrm{d}x \approx \int_{\mathbb{R}} e^{iux} f_H(x)\mathrm{d}x = \phi_H(u)$$

Now the function $\tilde{f}_H \coloneqq f_H\mathbb{1}_{[a,b]}$ is obviously supported on a finite interval $[a, b]$ and it is continuously differentiable on that interval (look at how it is defined ϕ_H and remember that we choose the root in order to have that continuous), so it admits a cosine expansion:

$$\tilde{f}_H(x) = \frac{\tilde{A}_0}{2} + \sum_{k=1}^{\infty} \tilde{A}_k \cos\left(k\pi\frac{x-a}{b-a}\right)$$

where:

$$\tilde{A}_k = \frac{2}{b-a}\int_a^b f_H(s)\cos\left(k\pi\frac{s-a}{b-a}\right)\mathrm{d}s$$

comparing the two equations before we obtain:

$$\tilde{A}_k = \frac{2}{b-a}\Re\left\{\tilde{\phi}_H\left(\frac{k\pi}{b-a}\right)\cdot e^{-\frac{ik\pi a}{b-a}}\right\} \approx \frac{2}{b-a}\Re\left\{\phi_H\left(\frac{k\pi}{b-a}\right)\cdot e^{-\frac{ik\pi a}{b-a}}\right\} =: A_k$$

so if we replace $\tilde{A}_k$ with A_k we have that:

$$\tilde{f}_H(x) \approx \frac{A_0}{2} + \sum_{k=1}^{\infty} A_k \cos\left(k\pi\frac{x-a}{b-a}\right) \approx f_H(x)$$

now since f_H is an *entire function* (function without any singularities anywhere in the complex plane, except at ∞) the cosine Fourier expansion converges with exponential speed then we can choose a suitable $N \in \mathbb{N}$ such that:

$$\frac{A_0}{2} + \sum_{k=1}^{\infty} A_k \cos\left(k\pi\frac{x-a}{b-a}\right) \approx \frac{A_0}{2} + \sum_{k=1}^{N-1} A_k \cos\left(k\pi\frac{x-a}{b-a}\right)$$

so in conclusion we obtain:

$$f_H(x) \approx \frac{A_0}{2} + \sum_{k=1}^{N-1} A_k \cos\left(k\pi\frac{x-a}{b-a}\right)$$

4.2 Pricing european options

Let $x \coloneqq \log\left(\frac{S_0}{K}\right)$ and $y \coloneqq \log\left(\frac{S_T}{K}\right)$ then consider a european option and call its value at expiration as $v(y,T)$. Given the transition density $f_H(y|x)$ we have that the value of the claim at time 0 should be:

$$v(x,0) = e^{-rT}\int_{\mathbb{R}} v(y,T)f(y|x)\mathrm{d}y$$

Since the density rapidly decays to zero as $y \to \pm\infty$ we truncate the infinite integration range without loosing significant accuracy to $[a,b]$ and we obtain approximation:

$$v(x,0) \approx e^{-rT}\int_a^b v(y,T)f(y|x)\mathrm{d}y$$

no we replace $f_H(y|x)$ with the approximation obtained in the previous paragraph:

$$v(x,0) \approx e^{-rT}\int_a^b v(y,T)\left[\frac{A_0(x)}{2} + \sum_{k=1}^{N-1} A_k(x)\cos\left(k\pi\frac{y-a}{b-a}\right)\right]\mathrm{d}y$$

now we define:

$$V_k \coloneqq \frac{2}{b-a}\int_a^b v(y,T)\cos\left(k\pi\frac{y-a}{b-a}\right)\mathrm{d}y$$

and obtain:

$$v(x,0) \approx \frac{b-a}{2}e^{-rT}\left[\frac{A_0(x)V_0}{2} + \sum_{k=1}^{N-1} A_k(x)V_k\right]$$

substituting the $A_k(x)$ with the formulation in the previous paragraph we finally obtain:

$$v(x,0) \approx e^{-rT}\left[\frac{1}{2}\Re\big\{\phi_H(0;x)\big\}V_0 + \sum_{k=1}^{N-1}\Re\left\{\phi_H\left(\frac{k\pi}{b-a};x\right)\cdot e^{-\frac{ik\pi a}{b-a}}\right\}V_k\right]$$

4.3 Coefficients V_k for european put options

The value of a put option at maturity, keeping the notation of the previous paragraph, is:

$$v(y,T) = \big[K(1-e^y)\big]^+$$

so we obtain that V_k for a put can be expressed as:

$$V_k = \frac{2K}{b-a}\int_a^0 (1-e^y)\cos\left(k\pi\frac{y-a}{b-a}\right)\mathrm{d}y$$

this is the formulation given by Fang & Osterlee. However, Le Floc'h underlined that the V_k are computed relatively to the strike price, but the truncation range is relative to the spot price. The consequence is that very deep out-the-money & in-the-money puts are severely mispriced, especially for short maturities. Le Floc'h proposed an improved version to calculate the V_k, we will present here a reformulation of that method. We will compute the V_k this time with respect to the spot price. Let's define $z := \log\left(\frac{S_T}{S_0}\right)$, then we have that:

$$v(z,T) = S_0\left[\frac{K}{S_0} - e^z\right]^+ = S_0\left[e^{-x} - e^z\right]^+$$

so we have that (notice that it is possible to price puts with $-x \in (a,b)$):

$$\begin{aligned}
V_k &= \frac{2S_0}{b-a}\int_a^b \left[e^{-x} - e^z\right]^+ \cos\left(k\pi\frac{z-a}{b-a}\right)\mathrm{d}z \\
&= \frac{2S_0}{b-a}\int_a^{-x} \left[e^{-x} - e^z\right] \cos\left(k\pi\frac{z-a}{b-a}\right)\mathrm{d}z \\
&= \frac{2S_0e^{-x}}{b-a}\int_a^{-x} \left[1 - e^{z+x}\right] \cos\left(k\pi\frac{z-a}{b-a}\right)\mathrm{d}z \\
&= \frac{2S_0e^{-x}}{b-a}\int_a^{-x} \left[1 - e^{z+x}\right] \cos\left(k\pi\frac{z-a}{b-a}\right)\mathrm{d}z \\
&= \frac{2}{b-a}\left[K\psi_k(a,-x) - S_0\chi_k(a,-x)\right]
\end{aligned}$$

where:

$$\begin{aligned}
\psi_k(a,-x) &:= \int_a^{-x} \cos\left(k\pi\frac{z-a}{b-a}\right)\mathrm{d}z \\
\chi_k(a,-x) &:= \int_a^{-x} e^z\cos\left(k\pi\frac{z-a}{b-a}\right)\mathrm{d}z
\end{aligned}$$

after integration we obtain:

$$\psi_k(a,-x) = \begin{cases} \frac{a-b}{k\pi}\sin\left(k\pi\frac{x+a}{b-a}\right) & k \neq 0 \\ -x-a & k = 0 \end{cases}$$

$$\chi_k(a,-x) = \frac{1}{1+\left(\frac{k\pi}{b-a}\right)^2}\left[e^{-x}\cos\left(k\pi\frac{x+a}{b-a}\right) - e^a - \frac{k\pi}{b-a}e^{-x}\sin\left(k\pi\frac{x+a}{b-a}\right)\right]$$

Now the only thing which is missing is how to compute the extremes of the truncation interval.

4.4 Computation of a and b

We used the formula proposed by Fang & Osterlee to compute the truncation range as:

$$[a,b] = [c_1 - 12\sqrt{|c_2|}, c_1 + 12\sqrt{|c_2|}]$$

where c_1 and c_2 are the two first cumulants. Indeed, defining the cumulant generating function as

$$g(u) \coloneqq \log \phi_H(-iu)$$

we have that $c_1 = g'(0)$ and $c_2 = g''(0)$. However, those are easier to compute numerically than analytically. Here below there are the formulae obtained trough a Taylor expansion of the characteristic function:

$$c_1 = (r-q)\tau + (1-e^{-\kappa\tau})\frac{\eta-\sigma_0}{2\kappa} - \frac{\eta\tau}{2}$$

$$\begin{aligned} c_2 = \frac{\sigma_0}{4\kappa^3}\{4\kappa^2[1+e^{-\kappa\tau}(\rho\theta\tau-1)]+\kappa[4\rho\theta(e^{-\kappa\tau}-1)-2\theta^2\tau e^{-\kappa\tau}]+\theta^2(1-e^{-2\kappa\tau})]\} \\ + \frac{\eta}{8\kappa^3}\{8\kappa^3\tau - 8\kappa^2[1+\rho\theta\tau+e^{-\kappa\tau}(\rho\theta\tau-1)] \\ + 2\kappa[(1+2e^{-\kappa\tau})\theta^2\tau + 8\rho\theta(1-e^{-\kappa\tau})] + \theta^2(e^{-2\kappa\tau}+4e^{-\kappa\tau}-5)\} \end{aligned}$$

5 Calibration

In this section we present the Yiran Cui's method for a fast calibration of the Heston model. The calibration problem is simply to choose a set of parameters $\Xi \coloneqq [\sigma_0^2, \eta, \rho, \kappa, \theta]^\top$ such that the difference between the observed price of the calls and the puts and the price given by the Heston model with that parameters is minimum. We denote with $C^*(K_i, T_i)$ the market prices for calls with strike K_i and maturity T_i and with $C(\Xi; K_i, T_i)$ the prices for calls under Heston model with parameters Ξ. From now on we will focus only on calls, but the discussion is analogue on puts (indeed we will derive the gradient w.r.t. the parameters for calls, but using the put-call parity is the same for puts). Given n call options to be calibrated we define:

$$r_i(\Xi) \coloneqq C(\Xi; K_i, T_i) - C^*(K_i, T_i) \qquad i = 1, ..., n$$

and the residual vector $r(\Xi) = [r_1(\Xi), ..., r_n(\Xi)]^\top$. With this notation the calibration of the Heston model is an inverse problem in the nonlinear least square form as:

$$(\star) \qquad \min_{\Xi} \frac{1}{2}\|r(\Xi)\|^2$$

since we suppose to have $n \gg 5$ (where 5 is the number of parameters that we have to determine) it is an overdetermined problem. To tackle this kind of problem we will use the Levenberg-Marquardt method, which requires to find the analytical gradient of the call price w.r.t. the parameters.

5.1 Analytical gradient

In order to obtain the analytical gradient Cui uses an equivalent form for the Heston characteristic function which is easier to derive and it is still numerically continuous. First of all define:

$$\begin{aligned}
\xi &:= \kappa - \theta\rho iu \\
d &:= \sqrt{\xi^2 + \theta^2(u^2 + iu)} \\
A &:= \frac{A_1}{A_2} = \frac{(u^2 + iu)\sinh\frac{Td}{2}}{d\cosh\frac{Td}{2} + \xi\sinh\frac{Td}{2}} \\
D &:= \log d + \frac{(k-d)T}{2} - \log\left(\frac{d+\xi}{2} + \frac{d-\xi}{2}e^{-dT}\right) =: \log B
\end{aligned}$$

then the new form of the ϕ_H is:

$$\begin{aligned}
\phi_H(\Xi; u, T) = \exp&\left\{iu[\log S_0 + (r-q)T] - \frac{T\kappa\eta\rho iu}{\theta}\right\} \\
\times \exp&\left\{-\sigma_0^2 A + \frac{2\kappa\eta}{\theta^2}D\right\}
\end{aligned}$$

Since it is only an algebrical manipulation from the previous characteristic function we will omit how it is obtained. Deriving we obtain that $\nabla\phi_H(\Xi; u, T) = \phi_H(\Xi; u, T)\mathbf{h}(u)$ where $\mathbf{h}(u) := [h_1(u), ..., h_5(u)]^\top$ with elements:

$$\begin{aligned}
h_1(u) &= -A \\
h_2(u) &= \frac{2\kappa}{\theta^2}D - \frac{T\kappa\rho iu}{\theta} \\
h_3(u) &= -\sigma_0^2\frac{\partial A}{\partial\rho} + \frac{2\kappa\eta}{\theta^2 d}\left(\frac{\partial d}{\partial\rho} - \frac{d}{A_2}\frac{\partial A_2}{\partial\rho}\right) - \frac{T\kappa\eta iu}{\theta} \\
h_4(u) &= \frac{\sigma_0^2}{\theta iu}\frac{\partial A}{\partial\rho} + \frac{2\eta}{\theta^2}D + \frac{2\kappa\eta}{\theta^2 B}\frac{\partial B}{\partial\kappa} - \frac{T\eta\rho iu}{\theta} \\
h_5(u) &= -\sigma_0^2\frac{\partial A}{\partial\theta} - \frac{4\kappa\eta}{\theta^3}D + \frac{2\kappa\eta}{\theta^2 d}\left(\frac{\partial d}{\partial\theta} - \frac{d}{A_2}\frac{\partial A_2}{\partial\theta}\right) + \frac{T\kappa\eta\rho iu}{\theta^2}
\end{aligned}$$

computing also the partial derivatives in the previous formulae we have that:

$$\begin{aligned}
\frac{\partial d}{\partial\rho} &= -\frac{\xi\theta iu}{d} \\
\frac{\partial A_2}{\partial\rho} &= -\frac{\theta iu(2 + T\xi)}{2d}\left(\xi\cosh\frac{Td}{2} + d\sinh\frac{Td}{2}\right) \\
\frac{\partial A_1}{\partial\rho} &= -\frac{iu(u^2 + iu)T\xi\theta}{2d}\cosh\frac{Td}{2} \\
\frac{\partial A}{\partial\rho} &= \frac{1}{A_2}\frac{\partial A_1}{\partial\rho} - \frac{A}{A_2}\frac{\partial A_2}{\partial\rho} \\
\frac{\partial B}{\partial\kappa} &= \frac{ie^{\kappa T/2}}{\theta u}\left(\frac{1}{A_2}\frac{\partial d}{\partial\rho} - \frac{d}{A_2^2}\frac{\partial A_2}{\partial\rho}\right) + \frac{TB}{2}
\end{aligned}$$

and the last ones:

$$\begin{aligned}
\frac{\partial d}{\partial \theta} &= \left(\frac{\rho}{\theta} - \frac{1}{\xi}\right)\frac{\partial d}{\partial \rho} + \frac{\theta u^2}{d} \\
\frac{\partial A_1}{\partial \theta} &= \frac{(u^2+ui)T}{2}\frac{\partial d}{\partial \theta}\cosh\frac{Td}{2} \\
\frac{\partial A_2}{\partial \theta} &= \frac{\rho}{\theta}\frac{\partial A_2}{\partial \rho} - \frac{2+T\xi}{iut\xi}\frac{\partial A_1}{\partial \rho} + \frac{\theta T A_1}{2} \\
\frac{\partial A}{\partial \theta} &= \frac{1}{A_2}\frac{\partial A_1}{\partial \theta} - \frac{A}{A_2}\frac{\partial A_2}{\partial \theta}
\end{aligned}$$

so now we can use the formula with the pseudo-probabilities calculated in the section 3.1 to obtain the gradient for a european call option at maturity and, as previously said, for a european put option, w.r.t. the parameters Ξ:

$$\begin{aligned}
\nabla C(\Xi; K, T) = \frac{e^{-rT}}{\pi}\Bigg[\int_0^\infty \Re\left\{\frac{K^{-iu}}{iu}\nabla\phi_H(\Xi; u-i, T)\right\}\mathrm{d}u \\
- K\int_0^\infty \Re\left\{\frac{K^{-iu}}{iu}\nabla\phi_H(\Xi; u, T)\right\}\mathrm{d}u\Bigg]
\end{aligned} \tag{7}$$

in order to evaluate the two integrals in (7) in the practical implementation we will use the Gauss-Legendre quadrature with 60 nodes and leveraging the fact that the integrands decay fast enough to justify a truncation of the integral domain to $[0, 100]$.

5.2 Levenberg-Marquardt

Let $J = \nabla r^\top \in \mathbb{R}^{5\times n}$ be the Jacobian matrix of the residual vector then for how it is defined the residual vector we have that:

$$J = [J_{ji}]_{j=1,\dots,5}^{i=1,\dots,n} = \left[\frac{\partial C(\Xi; K_i, T_i)}{\partial \Xi_j}\right]_{j=1,\dots,5}^{i=1,\dots,n}$$

now let $H(r_i) = \nabla\nabla^\top r_i \in \mathbb{R}^{5\times 5}$ be the Hessian matrix of each residual r_i. Then the gradient and Hessian of the objective function f of the problem $(\star)$ are:

$$\begin{aligned}
\nabla f &= Jr \\
\nabla\nabla^\top f &= JJ^\top + \sum_{i=1}^n r_i H(r_i)
\end{aligned}$$

The Levenberg-Marquardt algorithm is an iterative algorithm suitable for solving nonlinear least square problems. It is an hybrid between the Gauss-Newton algorithm and the steepest descent, however it is more robust than the first one. The stopping criterion for the LM algorithm is when one of the following is satisfied:

1. $\|r(\Xi_k)\| \leq \varepsilon_1$, where $\varepsilon_1 \in \mathbb{R}^+$, so if the solution found is sufficiently near the real solution;
2. $\|J_k\|_\infty \leq \varepsilon_2$, where $\varepsilon_2 \in \mathbb{R}^+$, so if the gradient is sufficiently small;
3. $\frac{\|\Xi_k - \Xi_{k-1}\|}{\|\Xi_k\|} \leq \varepsilon_3$, where $\varepsilon_3 \in \mathbb{R}^+$, so if the update in parameters is too small.

In practice works as follow:

Algorithm 1 Levenberg - Marquardt algorithm

1: Given the initial guess Ξ_0, compute $\|r(\Xi_0)\|$ and J_0.
2: Choose the initial damping factor as $\mu_0 = \tau \max\{diag(J_0)\}$ and $\nu_0 = 2$.
3: Set $k = 0$.
4: **while** TRUE **do**
5:     Compute $\Delta\Xi_k = (J_k J_k^\top + \mu_k I)^{-1} J_k r(\Xi_k)$.
6:     Compute $\Xi_{k+1} = \Xi_k + \Delta\Xi_k$ and $\|r(\Xi_{k+1})\|$.
7:     Compute $\delta_L = \Delta\Xi_k^\top[\mu\Delta\Xi_k + J_k r(\Xi_k)]$ and $\delta_F = \|r(\Xi_k)\| - \|r(\Xi_{k+1})\|$.
8:     **if** $\delta_L > 0$ and $\delta_F > 0$ **then**
9:         Accept the step: compute $J_{k+1}, \mu_{k+1} = \mu_k, \nu_{k+1} = \nu_k$.
10:     **else**
11:         Recalculate the step: set $\mu_k = \mu_k\nu_k, \nu_k = 2\nu_k$ and go to 4.
12:     **end if**
13:     **if** At least one stopping criterion is met **then**
14:         Break.
15:     **end if**
16:     $k = k + 1$.
17: **end while**

As we can see when the iteration is far from the optimum we give a large value to μ_k, which is called the *damping factor* and so the Hessian of the objective function is dominated by the scaled identity matrix, meanwhile when the iteration is closed to the optimum the Hessian matrix is dominated by the Gauss-Newton approximation. In particular we say that:

$$\nabla\nabla^\top f \approx JJ^\top$$

so we use the conjecture that near the optimum the problem is a *small residual problem* (in the sense that $\sum_{i=1}^n r_i H(r_i)$ is negligible due to having small r_i). This is sensible since we think that the Heston model is a good model to explain the smile and the skew of the volatility surface, so in a couple of words it is an appropriate model for the volatility. It can also be shown that solving the equation at step 5. of the algorithm is equivalent in finding the solution to the minimization problem:

$$\Xi_{k+1} := \arg\min_{z\in\mathbb{R}^5} \left\{ \|J_k^\top(z - \Xi_k) + r(\Xi_k)\|^2 + \mu_k\|z - \Xi_k\|^2 \right\}$$

so whenever the μ_k is sufficiently small we treat the problem as quasi-linear.

6 Simulation

Suppose now that we have calibrated the model and we wanted to sample paths from it, then we need an integration scheme. Broadie and Kaya developed an exact and bias-free scheme to sample Heston's paths, however it has a major drawback: it is incredibly slow for practical applications. To circumvent this problem many practitioners use variation of the Euler scheme or the Milstein scheme, which are classical scheme for numerically integrating SDEs. These schemes need a correction in order to keep the variance process positive, the correction can be that if the variance at a certain time step becomes negative then it is put equal to zero. They require small timesteps and have worse convergence than the Broadie-Kaya's one, moreover whenever the Feller's condition is violated the results are not reliable. We chose to present the Gamma Approximation scheme proposed by Jean-François Bégin, Mylène Bédard and Patrice Gaillardetz, which has the advantages to be almost as accurate as the Broadie-Kaya's one, to be computationally faster and to be low-bias. Before starting, in the following sections we will use the Heston model w.r.t. the log-prices $X_t = \log(S_t)$. Using Itô's formula we can rewrite, under risk-neutral measure, the model as:

$$\begin{cases} \mathrm{d}X_t = \left(r - q - \frac{1}{2}v_t\right)\mathrm{d}t + \sqrt{v_t}\mathrm{d}W_t \\ \mathrm{d}v_t = \kappa(\eta - v_t)\mathrm{d}t + \theta\sqrt{v_t}\mathrm{d}\tilde{W}_t \\ v_0 = \sigma_0^2 \end{cases}$$

6.1 Cumulative distribution function of $v_T|v_t$

We state here an important result (proven in Dufresne, "The integrated square-root process") that we will use later about the cumulative distribution function of v_T given v_t.

Proposition 3. *Let $F_{\chi^2}(z;\nu,\lambda)$ be the cdf of the non-central chi-square distribution with non centrality parameter λ and ν degrees of freedom,*

$$F_{\chi^2}(z;\nu,\lambda) = e^{-\lambda/2}\sum_{j=0}^{\infty}\frac{(\lambda/2)^j}{j!2^{\nu/2+j}\Gamma(\nu/2+j)}\int_0^z x^{\nu/2+j-1}e^{-x/2}\mathrm{d}x$$

Let

$$\overline{\nu} = \frac{4\kappa\eta}{\theta^2}$$

and for $t < T$:

$$n(t,T) = \frac{4\kappa e^{-\kappa(T-t)}}{\theta^2\left[1 - e^{-\kappa(T-t)}\right]}$$

then we have that

$$\boldsymbol{Pr}_{\mathbb{Q}}(v_T < x|v_t) = F_{\chi^2}\left(\frac{x\cdot n(t,T)}{e^{-\kappa(T-t)}};\overline{\nu}, v_t\cdot n(t,T)\right)$$

6.2 Explicit solution for the log-asset price

Let $\Delta t > 0$ be our timestep and discretize time as $0, \Delta t, 2\Delta t, ..., j\Delta t, ..., T$. Integrating the log-price equation in the Heston model and using the Cholesky's decomposition in the Brownian motion we obtain:

$$X_{j\Delta t} = X_{(j-1)\Delta t} + (r-q)\Delta t - \frac{1}{2}\int_{(j-1)\Delta t}^{j\Delta t} v_t \mathrm{d}t \\ + \rho \int_{(j-1)\Delta t}^{j\Delta t} \sqrt{v_t}\mathrm{d}\tilde{W}_t + \sqrt{1-\rho^2}\int_{(j-1)\Delta t}^{j\Delta t} \sqrt{v_t}\mathrm{d}\tilde{W}_t^{\perp}$$

and in the same way integrating the variance process yields:

$$V_{j\Delta t} = V_{(j-1)\Delta t} + \int_{(j-1)\Delta t}^{j\Delta t} \kappa[\eta - v_t]\mathrm{d}t + \theta \int_{(j-1)\Delta t}^{j\Delta t} \sqrt{v_t}\mathrm{d}\tilde{W}_t$$

Since $\theta > 0$ we can isolate the integral in the last term as:

$$\int_{(j-1)\Delta t}^{j\Delta t} \sqrt{v_t}\mathrm{d}\tilde{W}_t = \theta^{-1}\left[v_{j\Delta t} - v_{(j-1)\Delta t} - \kappa\eta\Delta t + \kappa \int_{(j-1)\Delta t}^{j\Delta t} v_t \mathrm{d}t\right]$$

and substituting into the first equation of this section we obtain:

$$X_{j\Delta t} = X_{(j-1)\Delta t} + (r-q)\Delta t - \frac{1}{2}\int_{(j-1)\Delta t}^{j\Delta t} v_t \mathrm{d}t + \frac{\kappa\rho}{\theta}\int_{(j-1)\Delta t}^{j\Delta t} v_t \mathrm{d}t \\ + \frac{\rho}{\theta}[v_{j\Delta t} - v_{(j-1)\Delta t} - \kappa\eta\Delta t] + \sqrt{1-\rho^2}\int_{(j-1)\Delta t}^{j\Delta t} \sqrt{v_t}\mathrm{d}\tilde{W}_t^{\perp}$$

we define the integrated variance as:

$$IV_{(j-1)\Delta t}^{j\Delta t} := \int_{(j-1)\Delta t}^{j\Delta t} v_t \mathrm{d}t$$

using this notation and the equation above and approximating the stochastic integral, Brodie and Kaya obtained the following equation:

$$\hat{X}_{j\Delta t} = \hat{X}_{(j-1)\Delta t} + (r-q)\Delta t - \frac{1}{2}I\hat{V}_{(j-1)\Delta t}^{j\Delta t} + \frac{\kappa\rho}{\theta}I\hat{V}_{(j-1)\Delta t}^{j\Delta t} + \\ \frac{\rho}{\theta}[\hat{v}_{j\Delta t} - \hat{v}_{(j-1)\Delta t} - \kappa\eta\Delta t] + Z\sqrt{1-\rho^2}\sqrt{I\hat{V}_{(j-1)\Delta t}^{j\Delta t}} \tag{8}$$

where $Z \sim \mathcal{N}(0,1)$.

6.3 Gamma approximation scheme

The algorithm works as follow:

Algorithm 2 GA Scheme

1: Create caches for the moments of $IV_{(j-1)\Delta t}^{j\Delta t}$ as explained in section 6.4.
2: Sample $\hat{v}_{j\Delta t}$ given $\hat{v}_{(j-1)\Delta t}$ from the non-central chi-square distribution described in section 6.1.
3: Given $\hat{v}_{j\Delta t}$ and $\hat{v}_{(j-1)\Delta t}$, calculate the integrated variance over time, $I\hat{V}_{(j-1)\Delta t}^{j\Delta t}$ from a moment-matched gamma distribution using the moments available in the caches.
4: Sample $Z \sim \mathcal{N}(0,1)$ and use equation (8) to obtain $\hat{X}_{j\Delta t}$.

the fact that the caches can be precomputed is what make this algorithm so computationally efficient.

6.4 Implementation of caches

Before discussing how the caches are implemented we need two technical propositions.

Proposition 4. *The integrated variance admits the representation*

$$IV_t^T \stackrel{d}{=} X_1 + X_2 + \sum_{j=1}^{\xi} Z_j$$

where X_1, X_2, Z_j $\forall j$ and ξ are mutually independent. The random variables X_1, X_2 and Z_j have the following representations:

$$X_1 \stackrel{d}{=} \sum_{n=1}^{\infty} \frac{1}{\gamma_n} \sum_{j=1}^{N_n} A_j$$

$$X_1 \stackrel{d}{=} \sum_{n=1}^{\infty} \frac{1}{\gamma_n} B_n$$

$$Z_j \stackrel{d}{=} \sum_{n=1}^{\infty} \frac{1}{\gamma_n} C_{n,j}$$

where

$$\gamma_n = \frac{\kappa(T-t)^2 + 4\pi^2 n^2}{2\theta^2 (T-t)^2}$$

Here, the A_j are independent exponential random variables with mean 1, N_n are independent Poisson random variables with respective means $(v_t + v_T)\lambda_n$ and

$$\lambda_n = \frac{16\pi^2 n^2}{\theta^2 (T-t)[\kappa^2 (T-t)^2 + 4\pi^2 n^2]}$$

The B_n are independent gamma random variables with a shape parameter of $\overline{\nu}/2$ and a scale parameter of 1. Finally, ξ is a Bessel random variable with parameter

$$z = \frac{2\kappa}{\theta^2 \sinh\left[\kappa(T-t)/2\right]}\sqrt{v_t v_T}$$

and degrees of freedom equal to $\xi/2 - 1$.

Proposition 5. *Let $C_1 = \coth\left[\kappa(T-t)/2\right]$ and $C_2 = \operatorname{csch}^2[\kappa(T-t)/2]$. Given v_t and v_T the mean and variance of IV_t^T are expressed as*

$$\mathbb{E}[IV_t^T|v_t, v_T] = \mathbb{E}[X_1] + \mathbb{E}[X_2] + \mathbb{E}[\xi]\mathbb{E}[Z]$$

and

$$\operatorname{Var}[IV_t^T|v_t, v_T] = \operatorname{Var}[X_1] + \operatorname{Var}[X_2] + \mathbb{E}[\xi]\operatorname{Var}[Z] + \left(\mathbb{E}[\xi^2] - \mathbb{E}[\xi]^2\right)\mathbb{E}[Z]^2$$

the mean and variance of X_1 respectively satisfy

$$\mathbb{E}[X_1] = (v_t + v_T)\left[\frac{C_1}{\kappa} - (T-t)\frac{C_2}{2}\right]$$
$$\operatorname{Var}[X_1] = (v_t + v_T)\theta^2\left[\frac{C_1}{\kappa^3} + (T-t)\frac{C_2}{2\kappa^2} - (T-t)^2\frac{C_1 C_2}{2\kappa}\right]$$

the mean and variance of X_2 respectively satisfy

$$\mathbb{E}[X_2] = \overline{\nu}\theta^2\left[\frac{-2 + \kappa(T-t)C_1}{4\kappa^2}\right]$$
$$\operatorname{Var}[X_2] = \overline{\nu}\theta^4\left[\frac{-8 + 2\kappa(T-t)C_1 + \kappa^2(T-t)^2 C_2}{8\kappa^4}\right]$$

the mean and the variance of Z respectively satisfy

$$\mathbb{E}[Z] = 4\mathbb{E}[X_2]/\overline{\nu}$$
$$\operatorname{Var}[Z] = 4\operatorname{Var}[X_2]/\overline{\nu}$$

Finally,

$$\mathbb{E}[\xi] = \frac{z\mathcal{I}_{\overline{\nu}/2}(z)}{2\mathcal{I}_{\overline{\nu}/2-1}(z)}$$
$$\mathbb{E}[\xi^2] = \mathbb{E}[\xi] + \frac{z^2\mathcal{I}_{\overline{\nu}/2+1}(z)}{4\mathcal{I}_{\overline{\nu}/2-1}(z)}$$

where z is as in the previous proposition and $\mathcal{I}_\nu(\cdot)$ is a Bessel function of the first kind with ν degrees of freedom.

With these two proposition we can now proceed and define $IV^{*j\Delta t}_{(j-1)\Delta t} := IV^{j\Delta t}_{(j-1)\Delta t} - X_1$ using the same notation as in Proposition 4. Now we obtain:

$$\mathbb{E}[IV^{j\Delta t}_{(j-1)\Delta t}|v_{(j-1)\Delta t}, v_{j\Delta t}] = \mathbb{E}[IV^{*j\Delta t}_{(j-1)\Delta t}|v_{(j-1)\Delta t}, v_{j\Delta t}] + \mathbb{E}[X_1]$$
$$\operatorname{Var}[IV^{j\Delta t}_{(j-1)\Delta t}|v_{(j-1)\Delta t}, v_{j\Delta t}] = \operatorname{Var}[IV^{*j\Delta t}_{(j-1)\Delta t}|v_{(j-1)\Delta t}, v_{j\Delta t}] + \operatorname{Var}[X_1]$$

And since the first two moments of X_1 depend only on $v_{(j-1)\Delta t} + v_{j\Delta t}$ and do not require the evaluation of any Bessel function, can be computed later in an inexpensive fashion. The moments computation will be performed in the following way:

Algorithm 3 IV Moment Computation

1: Precompute in some predefined points (called *totems*) $\mathbb{E}[IV^{*j\Delta t}_{(j-1)\Delta t}|v_{(j-1)\Delta t}, v_{j\Delta t}]$ and $\mathrm{Var}[IV^{*j\Delta t}_{(j-1)\Delta t}|v_{(j-1)\Delta t}, v_{j\Delta t}]$. These values are called *caches*.

2: Compute $\mathbb{E}[X_1]$ and $\mathrm{Var}[X_1]$.

3: Use linear interpolation of caches to approximate $\mathbb{E}[IV^{*j\Delta t}_{(j-1)\Delta t}|v_{(j-1)\Delta t}, v_{j\Delta t}]$ and $\mathrm{Var}[IV^{*j\Delta t}_{(j-1)\Delta t}|v_{(j-1)\Delta t}, v_{j\Delta t}]$.

4: Add $\mathbb{E}[X_1]$ and $\mathrm{Var}[X_1]$ to the previous moments respectively to obtain $\mathbb{E}[IV^{j\Delta t}_{(j-1)\Delta t}|v_{(j-1)\Delta t}, v_{j\Delta t}]$ and $\mathrm{Var}[IV^{j\Delta t}_{(j-1)\Delta t}|v_{(j-1)\Delta t}, v_{j\Delta t}]$.

Notice that $\mathbb{E}[IV^{*j\Delta t}_{(j-1)\Delta t}|v_{(j-1)\Delta t}, v_{j\Delta t}]$ and $\mathrm{Var}[IV^{*j\Delta t}_{(j-1)\Delta t}|v_{(j-1)\Delta t}, v_{j\Delta t}]$ are functions of $v_{(j-1)\Delta t} \times v_{j\Delta t}$ and with respect to this quantity they display an exponential behaviour. So the natural grid (way to choose the totems) is to choose an exponentially spaced grid.

References

1. Gatheral J. The volatility surface. A practitioner's guide. Wiley, 2006.
2. F. Fang and C.W. Oosterlee: Cosine expansions for option pricing. AMS, 2008.
3. Albrecher, Hansjoerg and Mayer, Philipp and Schoutens, Wim and Tistaert, Jurgen. The Little Heston Trap. Wilmott, 2007.
4. M. Burzoni. Lecture notes for the course: Mathematical Finance 2. Università Statale degli Studi di Milano, 2022.
5. Le Floc'h F. More Robust Pricing of European Options Based on Fourier Cosine Series Expansions. ArXiv, 2020.
6. Y. Cui, S. del Bano Rollin, G. Germano. Full and fast calibration of the Heston stochastic volatility model. ArXiv, 2016.
7. J.F. Bégin, M. Bédard and P. Gaillardetz. Simulating from the Heston Model: A Gamma Approximation Scheme. Monte Carlo Methods and Applications, 2014.
8. M. Broadie, O. Kaya. Exact Simulation of Stochastic Volatility and Other Affine Jump Diffusion Processes. Operations Research, 2006.