Universitat Politècnica de Catalunya

Facultat de Matemàtiques i Estadística

Master in Advanced Mathematics and Mathematical Engineering

Master's thesis

This is the long title
with a line skip

Eudald Romo Grau

Supervised by (name of the supervisor/s of the master's thesis)

Month, year

Thanks to...

Abstract

This should be an abstract in English, up to 1000 characters.

Keywords

Derivatives Trading, Heston, SWIFT, calibration

Contents

1. Motivation

Option pricing has an important role in contract trading, both as a form of derivative trading in itself, as a way to hedge other stock or derivative portfolios.

Pricing these options can have a broad range of difficulty, mainly due to two factors:

(1) `Contract type:` These options can take a variety of forms, from the vanilla European Options to more complex American or Asian options (TODO: citations).

(2) Underlying Model: The price of the derivative depends on the future price of its underlying. The evolution of this price can be modeled using simpler models like Black & Scholes, or using other models that take into account more theoretical results related to the underlying price evolution, like the Heston model. TODO: citations, add more models.

There are only closed analytic solutions for the Option pricing problem for the simpler contracts and models: European options under BS and can have no analytic solution when one either trades a more complicated contract (like American options) or uses a more complex model (like Heston) to model its dynamics.

The advantages of simpler models tend to be faster computations and possibly analytic functions to compute the price. On the other hand, they fail to capture more complex behavior of the market.

For example, BS is probably the most used model in real option trading (TODO: back this up) but fails to capture the well-known high kurtosis and negative skewness of the log returns volatility, which motivated the development of the Heston model, which will be the center of this study.

Most of these models can be used to price different strikes using a set of intrinsic parameters. Usually this parameters are calibrated either using historical data or using data from options with the same underlying at different strikes and, sometimes, different maturities. (TODO: Cite) In the second calibration mode, which is usually used in the trading world, particularly in high frequency trading, (TODO: cite) fast option calculation methods and model calibration techniques are required to be able to update the model in real time.

In particular, solving the Heston model calibration is a difficult optimization problem to solve, partly due to it being a multidimensional optimization problem with a mainly unknown structure: it is not known whether it is a convex problem or not nor whether it generally has a single solution or not. Because of that, several different branches of research have appeared, but they are either deterministic but slow, or use heuristics or assumptions about the model parameters to reduce the dimension of the problem.

A branch of numerical option pricing that, given a density function (and the value of its parameters) to describe the stochastic evolution of the underlying over time, use its characteristic function to valuate the price of the option with different strikes and maturities.

Traditionally, using this pricing method combined with a gradient-descent based method for calibration was not possible due to the characteristic function derivatives with respect of the model parameters being

too intractable.

Recently, there has been both work in fast option pricing numerical techniques using characteristic functions (TODO: cite) and giving simple expressions for the characteristic function (TODO: cite) of the Heston model (its previous expressions either had complex variable discontinuities that translated in numeric problems (TODO: cite) or complicated derivative expressions, which made it difficult to use in gradient descent based methods).

This work is centered around the calibration of the Heston model for European Options (choosing proper values of the model parameters to minimize the option valuation error) by using a gradient descent method (TODO: name properly) valuating the option and its partial derivatives using the new characteristic function approach in the SWIFT setup.

2. Previous work

3. Option Valuation

Definition 3.1. An European Option is a derivative contract on an underlying asset that fixes a transaction price (called **strike price**), an expiration time (also called **maturity**), and an **underlying** asset quantity and gives the holder of the contract the right (but not the obligation) of buying (Call Option) or selling (Put Option) the specified quantity of the underlying at the Strike price on expiration.

In the following sections we will use the following nomenclature for Option parameters: t will refer to the time at which the option is being valuated and T will be the expiration time. Given a time-dependent variable V, V_t will refer to the value of that variable at time t (respectively with T). S, K will refer to the underlying and the strike prices respectively.

Slight variations of these terms can give rise to other families of options by facilitating the exercising of the contracts, like the Bermudan and American options (which allow to exercise your buying/selling right at specified instants of time or at any point prior to expiry, respectively).

Other groups of option families are obtained by changing the definition of the option payoff, that is, the amount of money (be it in cash or in the value of the underlying obtained) the option gives on expiration.

Definition 3.2. Payoff of a Call European option:

$$v_E C(y, T) = \begin{cases} S_T - K & \text{if} \quad S_T > K \\ 0 & \text{if} \quad S_T \leq K \end{cases}$$

So, on maturity, a European Call option is only profitable if it allows the holder of the contract to buy the underlying asset at a cheaper price than the current asset spot price. In most models any option has a certain positive value if it's not expired, as the underlying price can always theoretically move enough so that a currently non-profitable option ends up with a positive payoff on expiry. Another

Definition 3.3. A Call European option at time t is called:

- Out of the money (OTM): If $K > S_t$ (that is, it has no value apart from time value).

- At the money (ATM): If $K = S_t$ (the price of a real option contract will rarely match perfectly the underlying asset price, but one can refer to a theoretic ATM option, for example to describe a volatility surface).

- In the money (ITM): If $K < S_t$ (that is, even if it were to expire right now, it would still hold some implicit value).

Some examples of other families of functions that have different payoffs are Binary Options (only have two possible payoffs dependent on K and S_T) and Asian Options (the payoff depends on the mean price of the underlying since the start of the contract and until maturity).

In order to price an option that expires at some point in the future, one usually first models the temporal evolution of the price of its underlying asset by using some stochastic process. This model will typically be expressed in terms of parameters that will be calibrated to fit the underlying price evolution of each product one wants to valuate.

3.1 Risk neutrality

Put-Call parity (that's why we only talk about Calls in this article) Model Log-Returns

3.2 European Option Valuation

European options only depend on the final price of the underlying asset on the moment of expiration of the contract. Thus, if we model the underlying asset's price with an stochastic process, the discounted expectation (TODO: Explain in 3.1) (under risk neutrality assumptions) of the gain one will obtain by holding the contract can be computed by integrating the density function of the price of the asset at expiry.

$$v(x, t) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}(v(y, T)|x) = e^{-r(t-T)} \int_{\mathbb{R}} v(y, T)f(y|x)dy \tag{1}$$

, where v denotes the option value, r the risk-neutral interest rate, $\mathbb{E}^Q$ the expectation under the risk-neutral measure, and x and y denote state variables that fully describe the underlying asset process state at times t and T respectively.

There is a broad literature in methods to solve this integration problem for basic underlying asset models and some simpler models, as Black and Scholes even have closed solutions (TODO: ref BS formula) but it is common that, for more complex models, the probability function at expiry doesn't have a known expression but it's characteristic function does, so several methods have been developed using numerical integration and the Fourier transform.

3.3 Volatility smile and surface

4. Black and Scholes

5. Lévy Processes in Option Pricing

Some well known underlying asset log returns theoretical stochastic processes, as the Geometric Brownian Motion (GBM) model (known as the Black-Scholes-Merton[2, 13] model), can be considered inside a more general concept called Lévy processes, which have a known Fourier transform.

Definition 5.1. A stochastic process $X = \{X_t : t \geq 0\}$ is considered a Lévy process if:

- $X_0 = 0$ almost surely.

- For any $0 \leq t_1 \leq t_2 \leq \cdots \leq t_n \leq \infty$, $X_{t_2} - X_{t_1}, X_{t_3} - X t_2, \cdots X t_n - X_{t_{n-1}}$ are independent.

- $\forall s < t$, $X_t - X_s$ is equal in distribution to X_{t-s}

A famous property of Lévy processes is the Lévy-Khintchine formula:

Lemma 5.2. Let $X = (X_t)_{t \geq 0}$ be a Lévy process. Then its characteristic function $Char_X(\omega)$ is:

$$Char_X(\omega)(t) := \mathbb{E}\left[e^{i\theta X(t)}\right] = e^{t\left(ai\omega - \frac{\sigma^2 \omega^2}{2} + \int_{\{0\}^c}(e^{i\omega x}\mathbb{I}_{|x|<1})\Pi(dx)\right)} = e^{\psi_L(w)}$$

, where $\psi_L(\omega)$ is called the characteristic Lévy exponent, and $a \in \mathbb{R}$, $\sigma \geq 0$, and Π is the Lévy measure of X, a σ-finite measure satisfying $\int_{\{0\}^c}(1 \wedge x^2)\Pi(dx) < \infty$

Note that any Lévy process can be completely characterized by the triplet (a, σ, Π), which represent a linear drift, a Brownian motion, and a L'evy jump process respectively.

5.1 Underlying Process Modeling

In option pricing, one can denote by $\{S_t\}_{t=0}^T$ the underlying price process and model it with an exponential Lévy process. For a given strike K, one can consider the log-transformed process $X_t := log(U_t/K)$, which will follow another Lévy process with an extra drift. Let the state variables be $x := log(S_t/K)$ and $y := log(U_T/K)$. In order to solve equation 1 one needs to evaluate $f(y|x)$, which is not always available in analytic form, but formula 5.2 can be used to obtain the form of its Fourier transform at any strike:

$$\hat{f}(\theta; x) := \mathbb{E}\left[e^{i\theta X_T}\right] = e^{-i\theta x}e^{-i\theta\mu T + T\psi_L(-\theta)} =: e^{-i\theta x}\hat{f}(\theta)$$

Model	$\psi_L(w)$	Param. Restrictions
GBM	$-\dfrac{\sigma^2}{2}\omega^2$	$\sigma > 0$

Table 1: Test

, where $\mu := r - \psi_L(-i)$ is a drift correction term and $\psi_L(\theta)$ is the characteristic L'evy exponent of the underlying log-returns process.

Given this price transformation, we can express the payoff of a put option as

$$v(y, T) := K \cdot max(1 - e^y, 0)$$

It is usually better for numerical reasons to always calculate the price of the call via the price of the put and the put-call parity formula, as e^y can have arbitrarily large values inside the integration domain.

6. Heston Model

The Black and Scholes model discussed in section 4 fails to capture essential well known properties of the real world market dynamics of the log returns distributions, as its high kurtosis, its negative skew, and the correlation between the underlying price and its volatility.

Several variations of this model have been proposed since it was introduced. Some of them, called stochastic volatility (SV) models, considered all the real world dynamics mentioned above and treated both the underlying price and its volatility as (potentially correlated) stochastic processes.

One of the first and most well-known SV models is the Heston model, defined by the following system of stochastic differential equations.

Definition 6.1. Heston model price-volatility equations:

$$dS_t = \mu S_t dt + \sqrt{\nu_t} S_t W_t^{(1)}$$

$$d\nu_t = \kappa(\overline{\nu} - \nu_t)dt + \sigma\sqrt{\nu_t}dW_t^{(2)}$$

$$dW_t^{(1)}dW_t^{(2)} = \rho dt$$

, where ν_t is the variance of the underlying asset price at time t. The parameters $\kappa, \overline{\nu}, \sigma, \rho$ are respectively called: mean-reversion rate, long-term variance, volatility of volatility, and correlation between the Brownian processes $W_t^{(1)}$ and $W_t^{(2)}$. From now on, $\theta := (\nu_0, \overline{\nu}, \rho, \kappa, \sigma)$ will refer to the set of model parameters.

Several studies have shown the relations between the Heston model parameters and the shape of the implied volatility surface [4, 7, 9, 10] which can be summarized as follows:

- ν_0 controls the position of the volatility surface.

- ρ controls its skewness.

- κ and σ control the convexity of the surface.

- $\kappa(\nu_0 - \overline{\nu}$ controls the term structure of implied volatility.

Heston also provided an analytic equation to obtain the price of European options in terms of the characteristic function of the stochastic distribution of the underlying asset. Below is its expression for a Call option:

Lemma 6.2. *Heston's pricing method.*

$$
\begin{aligned}
C(\theta; K, \tau) = &\frac{1}{2} \left(S_t e^{-q\tau} - K e^{-r\tau} \right) \\
&+ \frac{e^{-r\tau}}{\pi} \left[S_0 \int_0^\infty \text{Re} \left(\frac{e^{-iu \log \frac{K}{S_0}}}{iu} \phi(\theta; u - i, \tau) \right) du \right. \\
&\left. -K \int_0^\infty \text{Re} \left(\frac{e^{-iu \log \frac{K}{S_0}}}{iu} \phi(\theta; u, \tau) \right) du \right]
\end{aligned}
$$
(2)

. The proposed expression of the characteristic function is

$$
\begin{aligned}
\phi(\boldsymbol{\theta}; u, \tau) := \text{E} \left[\exp \left(iu \log \frac{S_t}{S_0} \right) \right] = \exp \Big\{ iu \log \frac{F}{S_0} \\
+ \frac{\kappa \overline{\nu}}{\sigma^2} \left[(\xi + d)\tau - 2 \log \frac{1 - g_1 e^{d\tau}}{1 - g_1} \right] + \frac{v_0}{\sigma^2} (\xi + d) \frac{1 - e^{d\tau}}{1 - g_1 e^{d\tau}} \Big\}
\end{aligned}
$$
(3)

, where

$$
\xi := \kappa - \sigma \rho iu \quad (4) \qquad d := \sqrt{\xi^2 + \sigma^2 \left(u^2 + iu \right)} \quad (5) \qquad g_1 := \frac{\xi + d}{\xi - d} \quad (6)
$$

From this expression, one can compute the analytic gradients of the option price in terms of the derivatives of the characteristic function. Then they can be used to optimize an appropriate objective function using gradient-descent based methods (See chapter 9) but, as will be explained in the next section, the analytic derivatives have not been widely used traditionally to calibrate Heston models prior to Cui et al. due to the complexity of the expressions obtained by derivating the available characteristic function expressions as eq 3 [5].

6.1 Calibration Difficulties

As opposed to simpler 1-dimensional models, Heston model calibrations is a multidimensional optimization problem with 5 degrees of freedom. Furthermore, the structure of this optimization problem is not known.

According to [5] no consensus exists among researcher regarding whether the objective function of this optimization problem is convex or not. Some results point to a non-convex function, as the calibration

methods proposed in [3, 14] (which yielded different results for different initial points) and one must use long or short term approximations and rules to provide a sufficient initial guess. Recent research claims to provide methods that reach a unique solution independently of the initial point [8] which, according to that study this indicates some structure that, even if not necessarily convex, tends to lead an initial guess to a stationary result. There's also no consensus on whether the problem always has a single optimum. In particular, it is known that there exists dependencies between the parameters that yield to similar results. For example, $lim_{t\to\infty} Var(\nu_t) = \dfrac{\sigma^2 \bar{\nu}}{\kappa}$, so large values of κ and σ can provide a model that prices options similarly to one with proportionally smaller values of these two parameters. The study by Cui et al. [5] defends that this yields to the the objective function of the optimization problem flat close to the optimum.

As said above, there's no guarantee that a gradient-based method converges to the global optimum of the model parameters, but even obtaining a local optimum has been traditionally difficult. A lot of literature uses numerical gradients[8] for these methods when trying to solve the Heston calibration problem (which are less accurate and more computationally consuming), because no simple analytic gradients were available and the ones obtained with symbolic algebra packages from the expressions of the characteristic function were intractable.

Prior to the work of Cui et al. [5], the existing methods could be summarized as:

- **Heuristic based models:** Using the relationships stated in 6, some studies reduce the dimension of the optimization problem by assuming some values or relationships between the parameters from the observation of a specific volatility surface. For example, Gatheral sets ν_0 to the short-term ATM implied variance obtained by using a BS model [7], an heuristic further justified by Chen [3], where the linearity between ν_0 and the BS implied volatility was verified for short maturities (less than 2 months). Other heuristics used in the industry are $\kappa = \dfrac{2.75}{\tau}$ and setting $\bar{\nu}$ to the BS short term volatility [4].

  These assumptions may restrict the optimization problem domain and exclude the optimum. 2.2.2.

- **Stochastic methods:** They are usually used in combination with deterministic search methods, as the Nelder and Mead simplex method [**?**] and would avoid the pitfalls of the gradient-based methods if the optimization problem is not convex. Some examples are Wang-Landau [3], differential evolution and particle swarm [**?**], and simulated annealing [**?**]. These methods are too computationally expensive for real-time use as of now: Fernandez et al. use GPU computations to calibrate options using a SV model called SABR, and it took 421.72 seconds to calibrate 12 instruments with tolerance of 10^{-2} using 2 NVIDIA Geforce GTX470 GPUs [**?**].

- **Deterministic methods:**

7. Multi Resolution Analysis and Shannon Wavelets

7.1 Multi Resolution Analysis

Multi Resolution Analysis (MRA) is a method that ultimately allows to express any function in $L^2(\mathbb{R})$ using a countable orthogonal family of wavelets. This family can then be truncated into a finite family and the original function can be orthogonally projected into the resulting subspace, obtaining an approximation with a certain level of resolution. Increasing the considered number of members of the wavelet family will increase the resolution of the approximation, converging to a perfect representation when all the wavelets are used [11].

Given the space $L^2(\mathbb{R}) = \left\{ f : \int_{-\infty}^{\infty} |f(x)|^2 dx < \infty \right\}$, a Multi Resolution Analysis is defined as a family of nested successive approximation closed spaces:

$$\cdots \subset V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset \cdots$$

Where the subpsaces V_i are complete (they are not redundant and cover $L^2(\mathbb{R})$):

$$\overline{\bigcup_{i \in \mathbb{Z}} V_i} = L^2(\mathbb{R}), \text{ and } \bigcap_{m \in \mathbb{Z}} = 0$$

, they have self-similarity in scale (all spaces are geometric scalings of V_0 by powers of 2):

$$f(x) \in V_i \Leftrightarrow f(2x) \in V_{i+1}$$

, they have self-similarity in time:

$$f(x) \in V_0 \Rightarrow f(x - k) \in V_0, \forall k \in \mathbb{Z}$$

(Note that self-similarity in scale implies that the self-similarity in time translates to all spaces V_i as $f(x) \in V_i \rightarrow f(x - 2^i k) \in V_i$.), and the integer shifts of a (or a finite group of) generator function ϕ form an orthogonal basis of V_0.

In summary, we can define:

Definition 7.1. (MRA): Consider $\phi \in L^2(\mathbb{R})$ a wavelet that spans the family $\{\phi_{m,k}\} m, k \in \mathbb{Z}$ defined as the normalized scaled integer shifts of ϕ. That is, $\phi_{m,k} = 2^{m/2}\phi(2^m x - k)$, and let $V_m :=$ $closure_{L^2(\mathbb{R})} \langle \{\phi_{m,k}\}_{k \in \mathbb{Z}} \rangle$

Then, if ϕ and V_m fulfill the conditions above, we say that ϕ is the scaling function or father wavelet of the MRA $\{V_m\}$ (note that the previous definition properly defines $\{V_m\}$ as a sequence of nested subspaces and that ϕ provides an orthonormal basis for each of them).

One of the important implications of obtaining a father wavelet and its MRA is that another wavelet family can be obtained from it, which will be a basis of $L^2(\mathbb{R})$. In order to do that, let's consider the set of subspaces W_m such that $V_{m+1} = V_m \oplus W_m$. Then $L^2(\mathbb{R}) = \sum_m \oplus W_m$ and there exists a function $\psi \in W_0$ (called mother wavelet) that generates an orthonormal basis of $L^2(\mathbb{R})$ [6] by defining the wavelet functions:

$$\psi_{m,k} = 2^{m/2} \psi(2^m x - k)$$

. Note that each $\{W_{m,k}\}_{k \in \mathbb{Z}}$ gives an orthonormal basis of W_m, and $\{W_{m_k}\}_{m \in [-\infty, m-1], k \in \mathbb{Z}}$ is an orthonormal basis of V_m. So for any $m \in \mathbb{Z}$ we can define a projection mapping P_m from any function $f \in L^2(\mathbb{R})$ into V_m:

$$P_m f(x) = \sum_{j=-\infty}^{m-1} \sum_{k \in \mathbb{Z}} d_{j,k} \psi_{j,k}(x) = \sum_{k \in \mathbb{Z}} c_{m,k} \phi_{m,k}(x)$$

, where $d_{j,k} = \{f, \psi_{j,k}\}$, $c_{m,k} = \{f, \phi_{m,k}\}$, and $\{f, g\} = \int_{\mathbb{R}} f(x) \overline{g(x)} dx$. Further, this projection converges in the L^2 norm as m tends to infinity [11].

7.2 Shannon Wavelets

Claude Shannon introduced the usage of the cardinal sine function for information modeling [15]:

$$sinc(x) := \begin{cases} \dfrac{sin(\pi x)}{\pi x} & \text{for} \quad x \neq 0 \\ 1 & \text{for} \quad x = 0 \end{cases}$$

This function serves as the father wavelet from which we obtain the families $\phi_{m,k}$ and ψ_{m_k}:

$$\phi_{m,k}(x) = 2^{m/2} \frac{sin(\pi(2^m x - k))}{\pi(2^m x - k)}, \quad k \in \mathbb{Z}$$

$$\psi_{m,k}(x) = 2^{m/2} \frac{sin(\pi(2^m x - k - 1/2)) - sin(2\pi(2^m x - k - 1/2))}{\pi(2^m x - k - 1/2)}, \quad k \in \mathbb{Z}$$

Shannon wavelets are interesting in part because while they have a slow decay in time domain, they have a simple expression and sharp compact support in the frequency domain.

$$\hat{\phi}_{m,k}(x) = \frac{e^{-ik/2^m w}}{2^{m/2}} rect\left(\frac{w}{2^{m+1}\pi}\right), \quad k \in \mathbb{Z}$$

$$\hat{\psi}_{m,k}(x) = -\frac{e^{-i\frac{k+1/2}{2^m}w}}{2^{m/2}} \left(rect\left(\frac{w}{2^m \pi} - \frac{3}{2}\right) + rect\left(-\frac{w}{2^m \pi} - \frac{3}{2}\right) \right), \quad k \in \mathbb{Z}$$

When using Shannon Wavelets to approximate a function with a truncated wavelet expansion [12] shows a bound for the projection error into V_m, by using concepts of band-limited functions.

Definition 7.2. A function f is called band-limited if $\exists B \in \mathbb{R}^+$, with $B < \infty$ such that

$$f(x) = \frac{1}{2\pi} \int_{-B\pi}^{B\pi} \hat{f}(\omega) e^{i\omega x} d\omega$$

, that is, the support of $\hat{f}$ is contained in the interval $[-B, B]$. The parameter B is referred to as the bandwidth of f.

The nested subspaces of a Shannon MRA can be expressed in terms of band-limited functions because of the sinc Fourier transform rectangular shape, as stated in the following lemma from [16]

Lemma 7.3. *Consider an MRA generated from the Shannon scaling function as defined in 7.1, then each subspace V_m corresponds to the space of functions $f \in L^2(\mathbb{R})$ with bandwidth $B \leq 2^m$*

Combining this lemma with the L^2 convergence of the projections P_m of f into V_m yields the following corollary [12]

Corollary 7.4. *The orthogonal projection P_m of a Shannon MRA is equivalent to:*

$$P_m f(x) = \frac{1}{2\pi} \int_{-2^m\pi}^{2^m\pi} \hat{f}(\omega) e^{i\omega x} d\omega$$

Which, in turn, can be used to derive the following bound to the error of the orthogonal projection.

Definition 7.5. Given $f \in L^2(\mathbb{R})$, let $H(\xi)$ be:

$$H(\xi) := \frac{1}{2\pi} \int_{|\omega| > \xi} \left| \hat{f}(\omega) \right| d\omega$$

, the normalized mass of the two-side tails of $\hat{f}$ defined by ξ.

Lemma 7.6. *Let $\epsilon_m(x) := f(x) - P_m f(x)$ (the pointwise approximation error due to the projection of f into V_m). Then $|\epsilon_m(x)| \leq H(2^m\pi)$ [12]*

8. SWIFT

Given a certain stochastic model for the underlying log returns price evolution over time and its density function f (equation 1) chosen to valuate the options and its correspondent characteristic function $\hat{f}(w) = \int_{\mathbb{R}} e^{-iwx} f(x) dx$, Shannon Wavelets Inverse Fourier Technique combines Shannon Wavelets and $\hat{f}$ to truncate f at a fixed approximation level m and provide a fast way to obtain the Fourier coefficients and recover the time domain ones through fast Fourier transform. (TODO: Cite)

8.1 Coefficients via Vieta's formula

$$c_{m,k} \approx c_{m,k}^* = \frac{2^{m/2}}{2^{J-1} \sum_{j=1}^{2^{J-1}}} \mathbb{R} \left[\hat{f}\left(\frac{(2j-1)\pi 2^m}{2^J} \right) e^{\frac{ik\pi(2j-1)}{2^J}} \right]$$

8.2 Sinc Integral

SWIFT works with the cardinal sinc function and it needs to evaluate it's integral $Si(t) = \int_0^t sinc(x)dx$ for which there's no closed form. This problem was addressed in [1, **?**] by using a combination of Vieta's formula and a cosine product-to-sum identity, resulting in an approximation by an incomplete cosine expansion

9. Calibration as an optimization problem

The goal of calibrating a model using market data, is to obtain model parameters that, when used for option valuation with an appropriate option valuation method, yield results similar to the real market option prices.

Completely matching the market for all strikes and maturities is an over-determined problem so, generally, the calibration consists on a least squares minimization problem, using a set of option market prices at different strikes and maturities.

In this section, the objective function of the optimization problem used by [5] is presented, as it will be the one used in this study.

9.1 Objective Function

Let $C^*(K_i, \tau_i)$ be the market price of a European call option and $C(\boldsymbol{\theta}; K_i, \tau_i)$ be the price at the same strikes and maturities obtained by using the Heston analytic formula 2. Let's also assume that we use n different options to calibrate the model, so $i \in [1, n] \subset \mathbb{Z}$. Then:

Definition 9.1. The calibration of the model is defined as the minimization problem

$$min_{\boldsymbol{\theta} \in \mathbb{R}^5}(f(\boldsymbol{\theta}))$$

$$f(\boldsymbol{\theta}) := \frac{1}{2}||\mathbf{r}(\theta)||^2 = \frac{1}{2}\mathbf{r}^T(\boldsymbol{\theta})\mathbf{r}(\boldsymbol{\theta})$$

, where $\mathbf{r}(\boldsymbol{\theta})$ is the n-dimensional vector of the residuals obtained when pricing the options considered for calibration using the model parameters. That is:

$$\mathbf{r}(\boldsymbol{\theta}) := [r_1(\boldsymbol{\theta}), \dots, r_n(\boldsymbol{\theta})]^T, \qquad r_i(\boldsymbol{\theta}) := C(\boldsymbol{\theta}; K_i, \tau_i) - C^*(K_i, \tau_i), \qquad i = 1, \dots, n \qquad (7)$$

It is worth noting that the choice of a specific objective function (and appropriate strikes and maturities to use as calibration points) will intimately characterize the model. This can be used when trading to fit to the specific needs of trader, ultimately leading to different objective functions, as relative price differences instead of absolute ones, or log-likelihood based methods [**?**] for, for example, hedging and market making [**?**].

One must be cautious, as this choice can lead to consistently giving more weight in the calibration to some set of options over the others. For example, according to [5], a commonly used minimization problem in industry for calibration consists on a least-squares problem similar to the one used in this study, using squared differences of the implied BS volatility of each calibration option, which can lead to giving more weight to OTM options in comparison with ITM ones.

The choice of calibration strikes and maturities can also have a similar impact in the pricing model. For example, as an option gets closer to the expiration date, the variance of the price change until expiration of the underlying asset gets smaller. So, if one were to always use the same strikes across different expiries, then for long-term maturities the model would end up with a lot of calibration points ATM but would lack control points in regions where the density function of the underlying asset is still relevant, which could lead to serious pricing errors for mildly OTM and ITM options. For the short-term maturities, the relevant regions of the density function would have small number of sparse strikes and there would be a lot of strikes in very OTM and ITM regions, where prices are mostly close to 0 and to the discounted payout respectively, so they are less sensitive to the change in the model parameters and do not help much in calibration (the price derivatives with respect to the model parameters will be almost flat at those expiry-maturity points, so they will not affect much the optimization problem), effectively wasting computational power.

As this study does not strive for any specific trading strategy, the choice of TODO: finish this and mention that the calibration objective function should be the same as the one used when valuating the model performance [**?**]

10. Characteristic function and Gradient Calculation

$$A := \frac{A_1}{A_2} \quad A_1 := \left(u^2 + iu\right) \sinh \frac{d\tau}{2} \quad A_2 := \frac{d}{v_0} \cosh \frac{d\tau}{2} + \frac{\xi}{v_0} \sinh \frac{d\tau}{2}$$

$$B := \frac{de^{\kappa\tau/2}}{v_0 A_2} \tag{8}$$

11. Goal

First draft, to be improved

In this study:

The model is calibrated by using a gradient descent method as in [5], even though there is no guarantee the function is convex, so we might hit a local minimum.

We will start from a set of parameters to price several options and we will try to recover the initial parameters after adding some random noise to the obtained prices.

Then we will use real market data and we will calibrate the model starting from several different initial guesses and check whether we obtain the same result.

Finally we will compare the results with the results obtained from a derivative-free optimization scheme that, even if slower, will avoid local minimums if repeated enough times.

12. Body

13. Numerical Results

Conclusions

14. Bibliography

References

[1] Sanjar M. Abrarov and Brendan M. Quine. A Rational Approximation for Efficient Computation of the Voigt Function in Quantitative Spectroscopy. *Journal of Mathematics Research*, 7(2):163, jun 2015.

[2] Fischer Black and Myron Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654, 1973.

[3] B Chen. Calibration of the Heston model with applications in derivative pricing and hedging. *Technical University of Delft(Master's thesis)*, 2007.

[4] I Clark. Foreign exchange option pricing: A practitioner's guide. *Chichester: Wiley*, 2011.

[5] Yiran Cui, Sebastian del Baño Rollin, and Guido Germano. Full and fast calibration of the Heston stochastic volatility model. *European Journal of Operational Research*, 263(2):625–638, dec 2017.

[6] I Daubechies. Ten Lectures on Wavelets. *CBMS-NSF Reg. Conf. Ser. Appl. Math.*, 1992.

[7] J Gatheral. The volatility surface: A practitioner's guide. *Finance*, 357, 2006.

[8] F Gerlich, A M Giese, J Maruhn, and E Sachs. Parameter identification in financial market models with a feasible point SQP algorithm. *Computational Optimization and Applications*, 51(3):1137–1161, 2012.

[9] M Gilli and E Schumann. Calibrating option pricing models with heuristics. *Computational Optimization and Applications*, 51(3):1137–1161, 2012.

[10] Agnieszka Janek, Tino Kluge, Rafał Weron, and Uwe Wystup. FX smile in the Heston model. In *Statistical Tools for Finance and Insurance*, pages 133–162. Springer Berlin Heidelberg, 2011.

[11] S G Mallat. *A wavelet tour to signal processing*. Academic Press, 1999.

[12] S C Maree, L Ortiz-Gracia, and C W Oosterlee. Pricing early-exercise and discrete barrier options by Shannon wavelet expansions. *Numer. Math.*, 136:1035–1070, 2017.

[13] Robert C Merton. Theory of rational option pricing. *Bell Journal of Economics and Management Science*, 4(1):141–183, 1973.

[14] S Mikhalov and U Nögel. Heston's stochastic volatility model: Implementation, calibration, and some extensions. *Wilmott Magazine*, July:74–79, 2003.

[15] C E Shannon. Communication in the presence of noise. *Proc. IRE*, 37:10–21, 1949.

[16] F Stenger. Handbook of sinc numerical methods. *CRC Press*, 2001.

A. Title of the appendix

You can include here an appendix with details that can not be included in the core of the document. You should reference the sections in this appendix in the core document.

B. Title of the appendix

Second appendix.