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 advantage 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).

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.

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.

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 Srike 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:

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

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).

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, T the time of expiration (or maturity) of the expiration, t the initial time, 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 broad literature in methods to solve this integration problem for basic underlying asset models. Some simpler models, as Black and Scholes even have closed solutions (TODO: ref BS formula), but they typically use models for the natural logarithm of the contract returns (Given a starting initial price S_0 and a final price S_T the returns are defined as $\dfrac{S_T}{S_0}$), as they appear to follow simpler stochastic processes. For some 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. the probability distributions do not have known expressions (TODO: cite).

4. Black and Scholes

5. Lévy Processes in Option Pricing

Some well known underlying asset log returns models, as the Geometric Brownian Motion (GBM) model (known as the Black-Scholes-Merton[2, 4] model), use a 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:*

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

, 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). Also, the

The evaluation of an European option can be obtained through equation 1.

5.1 Black-Scholes-Merton Model

6. Heston Model

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 [3].

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 $L^2(\mathbb{R})$ [1] 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 [3].

7.2 Shannon Wavelets

Claude Shannon introduced the usage of the cardinal sine function for information modeling [6]:

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

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 [5] 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 [9]

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 [5]

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 $|\epsilon_m(x)| \leq H(2^m\pi)$ [5]*

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 [7, 8] 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. Goal

Body

Numerical Results

Conclusions

10. Bibliography

References

[1] I. Daubechies. Ten lectures on wavelets. *CBMS-NSF Reg. Conf. Ser. Appl. Math.*, 1992.

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

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

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

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

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

[7] B.M. Quine S.M. Abramov. A rational approximation for efficient computation of the voigt function in quantitative spectroscopy. *Journal of Math. Res.*, 7(2):163–174, 2015.

[8] B.M. Quine S.M. Abramov. Sampling by incomplete cosine expansion of the sinc function: application to the voigt/complex error function. *Appl. Math. Comput.*, 258:425–435, 2015.

[9] 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.