Numerical Methods for Pricing Options with Transaction Costs

Juan David Aguilar Renzo Sosa Gonzalez Federico Gaspar Vega *

December 2020

Abstract

Black-Scholes option pricing theory rests on an arbitrage argument where it should be possible to adjust the portfolio continuously. In presence of transaction costs (as it happens in real markets) this assumption leads to a problematic situation. Indeed, diffusion process have infinite variation and a continuous trading would be ruinously expensive. Formally, the arbitrage argument will no longer be valid because replicating the portfolio will be infinitely costly so a discrete revision of Black-Scholes will be needed. In this work we examine Leland and Amster models where the costs behave as a non-increasing linear function through the implementation of numerical scheme using implicit and explicit finite differences methods.

Contents

*juandavid@keysoftware.com.co rensogonsales@gmail.com federicogaspar@gmail.com

1 Introduction

The interest in pricing financial derivatives arises from the fact that them can be used to minimize losses caused by price fluctuations of the underlying assets. Option pricing theory has made a great leap forward since the development of the Black–Scholes option pricing model by Fischer Black and Myron Scholes in [1] in 1973 and previously by Robert Merton in [2]. The solution of the famous (linear) Black–Scholes equation provides both the price for a European option and a hedging portfolio that replicates the option under the following assumptions:

- The price of the underlying derivative follows a Geometric Brownian motion.
- The drift, the volatility and the riskless interest rate are constant and no dividends are paid in that time period.
- The market is frictionless, thus there are no transaction costs (fees or taxes), the interest rates for borrowing and lending money are equal, all parties have immediate access to any information, and all securities and credits are available at any time and any size.
- There are no arbitrage opportunities.

Under these assumptions the market is complete, which means that any asset can be replicated with a portfolio of other assets in the market. Then, the linear Black–Scholes equation can be transformed into the heat equation and analytically solved to price the option. However, these restrictive assumptions never occur in reality. Transaction costs (see [3, 4, 5]), large investor preferences (see [6, 7, 8]) and incomplete markets [9] makes the classical model strongly or fully nonlinear, where both the volatility and the drift can depend on the time, the stock price or the derivatives of the option price itself.

In presence of transaction costs the arbitrage argument in the Black-Scholes theory to price options can no longer be used. Replicating the option by a dynamic strategy would be infinitely costly and no effective option price bounds emerge. Even when the trading takes place at discrete times it may be very costly to assure a given degree of accuracy in the replicating strategy and it could be found that the total cost of it exceeds the value of the stock itself. Additionally, transactions costs themselves are random and path-dependent: they depend not only on the initial and final stock prices, but also on the entire sequence of stock prices in between. Also, considering that computation of the maximum transactions costs is a nontrivial problem and because the maximum transactions costs will substantially exceed the average, the bounds on option prices will not be very tight.

These considerations lead us to study two transaction cost models where the costs behave as a non-increasing linear function. We implemented a numerical

scheme using implicit and explicit finite differences methods to price an European option with (and without) transaction costs modeled with Leland model. For the Amster case we studied it using the explicit scheme.

2 Finance Basics

Definition 2.1 (European Derivative Contract)**.** A financial contract F is a derivative or a contingent claim, if its value at expiration date T is determined exactly by the market price of the underlying cash instrument at time T.

Therefore we are restricting to those contracts whose value are determined by their payout at maturity[1]. Hence, at the expiration time T, the price F_T of the derivative asset is completely determined by S_T, the value of the underlying. After that date, the security ceases to exist.

Definition 2.2 (European Options)**.** An European option is a contract that allows a buyer the right to buy (call) or sell (put) an underlying asset or financial instrument at a specified strike price K on a specified expiration date T. The option can be purchased for a price called the premium, at time $t < T$.

Therefore, the seller has the corresponding obligation to fulfill the transaction (i.e., to sell or buy) if the buyer (owner) "exercises" the option. A call option would normally be exercised only when the strike price is below the market value of the underlying asset, while a put option would normally be exercised only when the strike price is above the market value. When an option is exercised, the cost to the buyer of the asset acquired is the strike price plus the premium, if any. The premium is income to the seller, and normally a capital loss to the buyer.

The most desirable way of pricing a call option is to find a closed-form formula for C_t that expresses the latter as a function of the underlying asset's price and the relevant parameters. At time t, the only known "formula" concerning C_t is the one that determines its value at the time of expiration. In fact, if the option is expiring out of money, the underlying can be purchased in the market for less than the strike and the option will have no value. From the other hand, if the option expires in the money, the holder will clearly exercise the option buying the underlying at the strike price K and then sell it in the market at a higher price S_T with a net profit of $S_T - K$. Market participants, being aware of this, will place a value of $S_T - K$ on the option. We can use a shorthand notation to express both of these possibilities as:

$$C_T = max(S_T - K, 0)$$

A similar analysis can be made for put options to arrive to the following formula:

$$P_T = max(K - S_T, 0)$$

[1] There are other types of derivatives whose payout depend on the path.

Also there exist a similar instrument that allows the writer to exercise the option before maturity:

Definition 2.3 (American Option)**.** An American option is a type of contract with similar properties to the European option but with the difference that it can be exercised any time between the writing and the expiration date.

Before we get into the details of the dynamics of the call options, and find a closed-form formula for C_t that expresses the latter as a function of the underlying asset's price, lets establish why options are so important. This will give us an insight about the relevance of the Black-Scholes work.

Options as a base of the European derivatives space

Let be $D_T = g(S_T)$ the payout at maturity T of any European derivative D with an underlying S. Then g can be replicated as a sum of call options (with same maturity T) plus some position on the stock[2] and a zero coupon bond.

If g is a step-wise function then the prove is straightforward, for any general g we prove it in the following theorem.

Theorem 2.1. *Let be D_t any European derivative of an underlying S_t with payout $D_T = g(S_T)$ at maturity $t = T$ and price D_t at time $t \leq T$. The set $\{C_t(K)\}_{K\in\mathbb{R}}$ of all call prices with the same maturity T determines D_t for any g such that its derivative is bounded.*

Proof. Let be $g(S_T)$ any function with bounded derivative. Then:

$$g(S_T) = g(0) + g'(0)S_T + \int_0^\infty g''(K)C_T(K)dK$$

Indeed just using integration by parts it can be checked that:

$$\int_0^\infty g''(K)C_T(K)dK = \int_0^\infty g''(K)max(S_T - K, 0)dK =$$

$$= \int_0^{S_T} g''(K)(S_T - K)dK = g'(S_T)0 - g'(0)S_T + g(S_T) - g(0)$$

where the first term is zero because of the bounded condition. □

Of course, in real markets the strike values are not continuous and this result present some limitations.

[2]Note that a position on the stock is equivalent to a position in a call option with zero strike price, $S_T = C_T(0)$.

Dynamics of an European Call Option

There are different approaches that can be used to calculate the fair market value of a derivative asset price. The one we are following here obtain the price of the derivative instrument by forming risk-free portfolios. Infinitesimal adjustments in portfolio weights and changes in the option price are going to be used to replicate unexpected movements in the underlying asset, S_t. This eliminates all the risk from the portfolio, at the same time it impose restrictions on the way $F_t(S_t)$, S_t, and the risk-free asset could jointly move over time. The assumption that we could make infinitesimal changes in positions will play an important role here and will show the advantage of continuous-time asset pricing models. We will follow this path because it is the closer one to our goal of describe options with transactions costs.

An alternative but very elegant method for pricing a derivative asset rest on the claim that we could find a probability measure $\mathbb{Q}$, such that under this probability, $e^{-rt}F_t(S_t)$ becomes a martingale with respect to the filtration I_t:

$$e^{-rt}F_t(S_t) = \mathbb{E}_{\mathbb{Q}}\left[e^{-rt}F_t(S_t)\middle| I_t\right], \quad t < T \tag{1}$$

For further details we refer the reader to [10]

Forming a risk free portfolio and Black-Scholes formula

By definition, at expiration any European derivative contract should depend solely on the value of the underlying security S_T. At that time, we know the exact form of the function $F_T = F_T(S_T)$. We assume that the same relationship is true for $t < T$ and that the price of the derivative product can be written as $F_t = F_t(S_t)$. If we would have a law of motion for the S_t process then we can use Ito's Lemma to obtain dF_t. We assume that the stochastic differential dS_t obeys is:

$$dS_t = a(S_t, t)dt + \sigma(S_t, t)dW_t\,,$$

where W_t is a Wiener process. Applying Ito's Lemma we find:

$$dF_t = \left[a_t\frac{\partial F}{\partial S} + \frac{1}{2}\sigma_t^2\frac{\partial^2 F}{\partial S^2} + \frac{\partial F}{\partial t}\right]dt + \sigma_t\frac{\partial F}{\partial S}dW_t\,. \tag{2}$$

From the above equation, we see that the dynamics of dF_t, is driven by the same Wiener increment dW_t that drives the S_t. In principle it could be possible to use one of these *SDEs* to eliminate the randomness in the other.

Let P_t dollars be invested in a combination of F_t and S_t:

$$P_t = \theta_1 F_t + \theta_2 S_t$$

where θ_1 and θ_2 are *quantities* that represent portfolio weights. Taking them as constant, allows us to write:

$$dP_t = \theta_1 dF_t + \theta_2 dS_t \tag{3}$$

In general, θ_1 and θ_2 will vary over time but at this point we ignore such dependence. If we substitute (3) in (2) we find:

$$dP_t = \theta_1 \left[\frac{\partial F}{\partial S} dS_t + \frac{1}{2}\sigma_t^2 \frac{\partial^2 F}{\partial S^2} dt + \frac{\partial F}{\partial t} dt \right] + \theta_2 dS_t \tag{4}$$

If we ignore for a minute that $\frac{\partial F}{\partial S}$ depends on S_t and without any loss of generality we select $\theta_1 = 1$ and $\theta_2 = -\frac{\partial F}{\partial S}$, then dS_t vanishes and the dP_t process becomes deterministic due to the cancellation of the random term. This means that the portfolio is risk-free and in consequence its appreciation must equal the earnings of a risk-free investment during an interval dt in order to avoid arbitrage.

Assuming that the (constant) risk-free interest rate is given by r, the expected capital gains must equal $rP_t dt$ in the case where S_t pays no "dividends" and replacing this in (4) with the proper values for θ_1 and θ_2 yields to the famous Black-Scholes equation:

$$\boxed{-rF + r\frac{\partial F}{\partial S} S_t + \frac{1}{2}\sigma_t^2 \frac{\partial^2 F}{\partial S^2} + \frac{\partial F}{\partial t} = 0\,.} \tag{5}$$

The selection for θ_1 and θ_2 "works" in the sense that it eliminates the "unpredictable" random component and makes the portfolio risk-free, but it also violates the assumption that θ_1 and θ_2 are constants. Although, formally speaking, the risk-free portfolio method is not satisfactory the method still gives us the correct *pde*. For a detailed explanation on this topic, see [10].

From now on we are going to consider the special case where $a(S_t, t) = \mu S_t$ and $\sigma(S_t, t) = \sigma S_t$. Therefore the model for the price of the stock is given by:

$$dS_t = \mu S_t dt + \sigma S_t dW_t \tag{6}$$

where W_t in a Wiener process. It is not hard, using Ito-lemma, to prove that (6) models stocks prices that varies with a log-normal distribution. In the following we are going to establish a connection between Black-Scholes and the heat equation, and then we are going to use this connection to finally price an European call option.

Connection with the heat equation

To solve the Black-Scholes *pde* we recognize that it is a Cauchy–Euler equation which can be transformed into a diffusion equation by introducing the following change-of-variable transformation:

$$x = \ln\left(\tfrac{S}{K}\right) + \left(r - \tfrac{\sigma^2}{2}\right)\tau\,, \quad \tau = T - t\,, \quad u(x,\tau) = e^{-r\tau} C_\tau(x)\,. \tag{7}$$

After some computations we find that Black-Scholes equation (5) in $\mathbb{R}^+ \times [0,T]$ is transformed into the heat equation:

$$\frac{\partial u}{\partial \tau} = \frac{\sigma^2}{2}\frac{\partial^2 u}{\partial x^2}\,, \quad x \in \mathbb{R}\,, \quad \tau \in [0,T] \tag{8}$$

Of course, if we re-scale the function $u(x,\tau) \to u(x, \frac{\sigma^2}{2}\tau)$ or, alternatively, by $u(x,\tau) \to u(\frac{\sqrt{2}}{\sigma}x, \tau)$ we obtain the a-dimensional diffusion equation. Neverless it worth to note that if σ^2 is imaginary, then the solutions in the x-direction becomes oscillatory.

Pricing an European Call Option

Here we want to find the solution to (5) with the following boundary and initial conditions:

$$\begin{cases} C_T = max(S_T - K, 0) \;\text{ for }\; 0 \le S_T < \infty \\ C_t(0) = 0 \;\text{ for }\; 0 \le t \le T \\ C_t(S_t) \sim S_t \;\text{ as }\; S_t \to \infty \;\text{ for }\; 0 \le t \le T \end{cases}$$

Using the transformation of the previous section, the initial and boundary conditions are mapped into:

$$\begin{cases} u_0(x) = u(x,0) = K\left(e^{max\{x,0\}} - 1\right) \\ u(x,\tau) \to 0 \;\text{ as }\; x \to -\infty \\ u(x,\tau) \to Ke^{x+r\tau} \;\text{ as }\; x \to \infty \end{cases}.$$

The solution to (8) is very well documented in the literature and can be written as the convolution of the initial condition with the Poisson Kernel:

$$u(x,\tau) = \frac{1}{\sigma\sqrt{2\pi\tau}}\int_{-\infty}^{\infty} u_0(y) exp\left(-\frac{x-y}{2\sigma^2\tau}\right) dy = Se^{r\tau}N(d_1) - KN(d_2)\,,$$

where N is the cumulative density function of the normal distribution:

$$N(d) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d} e^{-\frac{s^2}{2}} ds$$

and $\sigma\sqrt{\tau}d_1 = x + \sigma^2\tau$ and $\sigma\sqrt{\tau}d_2 = x$. Reversing the transformations in (7) we finally find the price of a call option:

$$C_t(S) = SN(d_1) - Ke^{-r(T-t)}N(d_2) \tag{9}$$

3 Options with transaction costs

The Black–Scholes model requires a continuous portfolio adjustment in order to hedge the position without any risk. If the cost associated with the rehedging are independent of the time scale of rehedging then the infinite number of transactions needed to maintain a hedged position until expiration leads to infinite

total transaction costs. Thus, the hedger needs to find the balance between the transaction costs that are required to rebalance the portfolio and the implied costs of hedging errors. As a result to this "imperfect" hedging, the option might be over or under priced up to the extent where the riskless profit obtained by the arbitrageur is offset by the transaction costs, so that there is no single equilibrium price but a range of feasible prices.

Leland model

In 1985 Leland [3] has proposed a very simple modification to Black-Scholes model for vanilla calls and puts where the portfolio and transaction costs are revised every finite sized time-step δt. He introduced a model for the transaction cost proportional to the monetary value of the assets bought or sold, $\frac{\kappa}{2}|\nu|S$. Here, κ denotes the round trip transaction cost per unit dollar of the transaction and the number of assets bought ($\nu > 0$) or sold ($\nu < 0$) at price S. Finally it is assumed that the hedge of the portfolio has an expected return equal to that from a riskless zero coupon bond. We can follow the Black-Scholes analysis as in the previous section but in equation (3) we should subtract the transaction costs. Indeed, equation (3) should be re-written as:

$$\delta P_t = \delta P_t^0 - k|\nu|S \tag{10}$$

where δP_0^t as in equation (4) (with the model for the stock distribution as in (6)) is defined as:

$$\delta P_t^0 = \sigma S \left(\frac{\partial F}{\partial S} - \theta \right) dW + \left(\frac{1}{2}\sigma^2 S^2 \frac{\partial^2 F}{\partial S^2} + \mu S \frac{\partial F}{\partial S} + \frac{\partial F}{\partial t} - \mu \theta S \right) dt$$

If we set up as before $\theta = \partial F/\partial S$, then the number of assets traded during the hedging are:

$$\nu = \frac{\partial F}{\partial S}(S + \delta S, t + \delta t) - \frac{\partial F}{\partial S}(S, t) \sim \sigma S \frac{\partial^2 F}{\partial S^2} dW \tag{11}$$

where we ignore $\mathcal{O}(dt)$ terms. From (11) it follows that the expected value of the transaction cost term is:

$$E\Big(k|\nu|S\Big) = \sqrt{\frac{2}{\pi}}\; k\sigma S^2 \left|\frac{\partial^2 F}{\partial S^2}\right| \sqrt{\delta t} \tag{12}$$

where $\sqrt{2\delta t/\pi} = E(|dW|)$.

If we assume that this portfolio is hedged, the holder should expect to make as much as if he had put the money in the bank, and then we can write:

$$E\Big(\delta P\Big) = E\Big(\delta P^0 - k|\nu|S\Big) = r\left(F - S\frac{\partial F}{\partial S}\right)\delta t \tag{13}$$

which leads to the nonlinear Black-Scholes equation with a modified volatility:

$$\frac{\partial F}{\partial t} + \frac{1}{2}\widetilde{\sigma}^2 S^2 \frac{\partial^2 F}{\partial S^2} + rS\frac{\partial F}{\partial S} - rF = 0 \tag{14}$$

where we defined:

$$\frac{\widetilde{\sigma}^2}{\sigma^2} = 1 - \frac{2k}{\sigma}\sqrt{\frac{2}{\pi\,\delta t}}\; sign\left(\frac{\partial^2 F}{\partial S^2}\right). \tag{15}$$

A financial interpretation can be made: the hedge eliminates the leading order component of randomness $\mathcal{O}(dt)$ leaving a small component of order $\mathcal{O}(\sqrt{dt})$ proportional to $\Gamma = \partial^2 F/\partial S^2$ (which is a measure of the degree of mishedging of hedged portfolio). Thus, the gamma is related to the amount of rehedging that takes place in the next time interval and hence to the cost expected.

Non-linearity has a profound impact from theoretical point of view, it breaks down the "superposition principle". Indeed, the value of the portfolio is no longer expected to be the sum of the value of each individual option that compose it. Neverless, differentiating the solution of the usual Black-Scholes (9), it follows that $\Gamma > 0$ for one call (the same is true for a put option) held long and $\Gamma < 0$ for the short position. Assuming the same behaviour in the presence of transaction costs, the equation (14) becomes linear with an adjusted constant variance $\widetilde{\sigma}^2 = \sigma^2 - Le$ for the long position and $\widetilde{\sigma}^2 = \sigma^2 + Le$ for the short one, where Le is some positive constant. Therefore a long position has a lower apparent volatility than in the classical model. This is because when the asset price rises the owner of the option must sell some of the asset to remain delta hedged, however the effect of the bid-offer spread on the underlying is to reduce the price at which the asset is sold and so the effective increase in the asset price is less than the actual increase. The converse is true for the short position.

If $\delta t > 1$ then the transaction cost term surpass the basic variance which implies that costs are too high. If $\delta t > 1$ then $\widetilde{\sigma} \sim \sigma$ which implies very low transaction costs and hence δt is too large and it should be decreased to minimize the risk.

Amster model

In the case of Amster model they consider a portfolio where the transaction costs behave as a nonincreasing linear function:

$$\delta P_t = \delta P_t^0 - (a + b|\nu|)|\nu|S \tag{16}$$

The modified Black Scholes equation when ones consider this extra term is:

$$\frac{\partial F}{\partial t} + \frac{1}{2}\widetilde{\sigma}^2 S^2 \frac{\partial^2 F}{\partial S^2} + b\sigma^2 S^3 \left(\frac{\partial^2 F}{\partial S^2}\right)^2 + rS\frac{\partial F}{\partial S} - rF = 0 \tag{17}$$

This is a manifestly non-linear differential equation which is really hard to tackle analytically, therefore we will explore the possibility to find a solution

numerically. Neverless, in [11] they obtain solutions for the stationary problem. Moreover, they give conditions for the existence of solutions of the general evolution equation.

4 Numerical methods

The models presented above do not admit simple closed form solutions and a numerical approach is the only mean of obtaining quantitative results. In this section we discuss numerical techniques for solving partial differential equation and we construct algorithms to solve these models. The idea underlying finite-difference methods is to replace the partial derivatives occurring in the partial differential equations by finite difference approximations. The numerical considerations we have to care about can be summarized as:

- the discretisation problem: how to construct a finite approximation to the partial differential equation that can be implemented as an algorithm.
- the stability problem: consider whether or not the discretised model is sensitive to small errors that arise because of the finite precision of the computer.
- the convergence problem: the degree to which the solutions of the discretised equations approximate the solutions of the *pde*.

We are not going to discuss one fundamental consideration: the efficiency problem which can be measure by two indicators of the solution algorithm:

- the amount of computer memory required
- the number of arithmetical calculations performed.

In the following we are going to consider implicit and explicit finite-difference methods.

The finite-difference mesh and derivatives approximations

If we choose a constant x-spacing δx, we cannot solve the problem for all $-\infty < x < \infty$ with a finite number of x-steps. We get around this problem taking a finite, but suitable large, number of x-steps. Therefore, we define the bounded region for the finite difference method as $[-\delta x N_-, \delta x N_+] \times [0, \delta\tau M]$, where N_- and N_+ are large enough constants and $M = \frac{1}{2\delta\tau}\sigma^2 T$. In the next place, we discretise this region into equally spaced nodes in each direction, where the mesh points have the form $(n\delta x, m\delta\tau)$, and where $u(n\delta x, m\delta\tau) = u_n^m$ is the exact solution of the *pde* at the point $(n\delta x, m\delta\tau)$.

The partial derivative is defined as a limiting process, if we substitute dx for some finite but small δx, we obtain a finite difference approximation. Indeed, lets define:

- the forward difference:

$$\frac{\partial u}{\partial x}(x,\tau) \approx \frac{u(x+\delta x,\tau)-u(x,\tau)}{\delta x} \tag{18}$$

- the backward difference:

$$\frac{\partial u}{\partial x}(x,\tau) \approx \frac{u(x,\tau)-u(x-\delta x,\tau)}{\delta x} \tag{19}$$

- the central difference:

$$\frac{\partial u}{\partial x}(x,\tau) \approx \frac{u(x+\delta x,\tau)-u(x-\delta x,\tau)}{2\delta x} \tag{20}$$

making a Taylor series expansion, is easy to check that the forward and backward approximation is $\mathcal{O}(\delta x)$ and the central is $\mathcal{O}((\delta x)^2)$. For second partial derivatives, such as $\partial^2 u/\partial x^2$ we can define the symmetric central difference:

$$\frac{\partial^2 u}{\partial x^2}(x,\tau) \approx \frac{u(x+\delta x,\tau)-2u(x,\tau)+u(x-\delta x,\tau)}{(\delta x)^2}\,. \tag{21}$$

This approximation, of order $\mathcal{O}((\delta x)^2)$, is preferred among all of the other possibilities, because it preserves the the reflection symmetry $x \to -x$ of the partial second derivative.

The explicit finite difference method

Consider the transformed Black-Scholes equation for the value of a European option,

$$\frac{\partial u}{\partial \tau} = \frac{\partial^2 u}{\partial x^2}$$

with the boundary and initial condition:

$$\lim_{x\to-\infty} u(x,\tau) = f(x,\tau)\,, \quad \lim_{x\to\infty} u(x,\tau) = g(x,\tau)\,, \quad u(x,0) = u_0(x)\,.$$

Using the forward difference for $\partial u/\partial\tau$ and the symmetric central difference for $\partial^2 u/\partial x^2$, restricting our attention to the values of $u(x,\tau)$ on the regular mesh u_n^m and neglecting $\mathcal{O}(\delta\tau)$ terms for the first derivative and $\mathcal{O}((\delta x)^2) = \mathcal{O}(\delta\tau)$ for the second, we can approximate the diffusion equation by:

$$v_n^{m+1} = \alpha\, v_{n+1}^m + (1-2\alpha)\, v_n^m + \alpha\, v_{n-1}^m\,, \quad \text{where } \alpha := \frac{\delta\tau}{(\delta x)^2} \tag{22}$$

with the following boundary conditions:

$$\begin{cases} v_{-N_-}^m = f(-N_-\delta x, m\delta\tau)\,, \quad v_{N_+}^m = g(N_+\delta x, m\delta\tau)\,, \\ v_n^0 = u_0(n,\delta x)\,, \quad \text{with} \qquad -N_- \le n \le N_+\,. \end{cases} \tag{23}$$

It is clear why this is an explicit method: if at time-step m, we know v_n^m for all values of n we can explicitly calculate v_n^{m+1}.

We used v_n^m instead of u_n^m to emphasise that the solution of (22) is only an approximation to the solution of the *pde*, since we ignore higher order terms. This allows us to consider the following convergence question: Does $v_n^m \to u_n^m$ as $\delta\tau \to 0$ and $\delta x \to 0$? It can be proven that if we choose δx and $\delta\tau$ in such a way that $0 < \alpha < 1/2$ the finite-difference method approximation converges to the actual solution. It worth to note that any errors in $v_{N_+}^m$ can only influence only the values of v_j^i such that $i + j \geq m + N_+$. Since $m \leq M$, this means that the errors are confined to a finite region.

The equation (22) has a natural interpretation as a discrete random walk. If the random walk is taking place on every node of the mesh and we interpret v_n^m as the probability that the walker is at the node $(n\delta x, m\delta\tau)$ at time $m\delta\tau$ then the walker has a probability of α to move to right or to the left and a probability of $(1-2\alpha)$ to stay in the same place than the previous step. In particular, if $\alpha = 1/2$ (which marks the boundary between stability and instability) corresponds to a binomial walk where the walker cannot stay put. In the unstable regime, $\alpha < 0$ or $\alpha < 1/2$, there is a negative probability.

The implicit finite difference method

Implicit finite difference methods are used to overcome the convergence limitation imposed by the explicit method $0 < \alpha < 1/2$. Implicit finite difference method allow us to use a large number of x-mesh points without having to take ridiculously small time-steps.

The fully implicit method

Here we use the backward difference for $\partial u/\partial\tau$ and the symmetric central difference for $\partial^2 u/\partial x^2$, neglecting $\mathcal{O}(\delta\tau)$ terms for both derivatives and restricting our attention to the values of $u(x,\tau)$ on the regular mesh u_n^m, we can approximate the diffusion equation by:

$$v_n^{m-1} = -\alpha\, v_{n+1}^m + (1+2\alpha)\, v_n^m - \alpha\, v_{n-1}^m\,, \tag{24}$$

where the boundary conditions are the same as in equation (23). Again, we use v_n^m to denote the approximation to the exact value u_n^m. Here, v_i^m depend on v_n^{m-1} in an implicit manner. Indeed, the new values cannot be separated and solved in terms of the old ones.

Writing equation (24) in a matricial form:

$$\mathbf{M}\mathbf{v}^m = \mathbf{v}^{m-1} + \mathbf{b}^m\,, \tag{25}$$

where:

$$\mathbf{v}^m = \begin{bmatrix} v^m_{N_+-1} \\ \vdots \\ v^m_0 \\ \vdots \\ v^m_{-N_-+1} \end{bmatrix} \quad \mathbf{b}^m = \alpha \begin{bmatrix} v^m_{N_+} \\ 0 \\ \vdots \\ 0 \\ v^m_{-N_-} \end{bmatrix} \quad \mathbf{M} = \begin{bmatrix} 1+2\alpha & -\alpha & & 0 \\ -\alpha & 1+2\alpha & \ddots & \\ & \ddots & \ddots & -\alpha \\ 0 & & -\alpha & 1+2\alpha \end{bmatrix}$$

Here, is a straightforward task to conclude that if $\mathbf{M}$ is not invertible then (25) either will not have a solution or it will not be unique. But this is a triangular matrix, so it will be invertible if and only if its principal diagonal has no zero element. $\mathbf{M}$ satisfies this condition, so it is invertible and the solution can be found as:

$$\mathbf{v}^m = \mathbf{M}^{-1}\left(\mathbf{v}^{m-1} + \mathbf{b}^m\right) .$$

In practice the tri-diagonal structure means that there is a highly efficient algorithm called LU decomposition for solving (25) in $\mathcal{O}(N)$ arithmetic operations. The details of the implementation of this algorithm can be found in [12, 13]. Also, as in the explicit method, a convergence criterion of the implicit finite-difference approximation to the solution of the partial differential equation can be proved. Indeed, the implicit method converges if and only if $\alpha > 0$.

The Crank-Nicolson method

The Crank-Nicolson method is essentially an average of the implicit and explicit method and it is used to overcome the convergence restrictions imposed by the explicit method and to have $\mathcal{O}((\delta\tau)^2)$ rates of convergence to the solution of the diffusion equation, instead of $\mathcal{O}(\delta\tau)$ as it was for both implicit and explicit methods. This method is obtained by averaging the discretization proposed by the implicit and explicit method leading to:

$$v^{m+1}_n - \frac{\alpha}{2}\left(v^{m+1}_{n-1} - 2v^{m+1}_n + v^{m+1}_{n+1}\right) = v^m_n + \frac{\alpha}{2}\left(v^m_{n-1} - 2v^m_n + v^m_{n+1}\right) . \quad (26)$$

It can be proved (see [12]) that this approximation is $\mathcal{O}((\delta\tau)^2)$. As in the previous case, it is convenient to write equation (26) in a matricial form:

$$(\mathbb{1} + \mathbf{M})\mathbf{v}^{m+1} = (3\mathbb{1} - \mathbf{M})\mathbf{v}^m + 2\mathbf{b}^m , \quad (27)$$

where $\mathbf{M}$ and $\mathbf{v}^m$ are defined as in equation (25) and:

$$\mathbf{b}^m = \alpha\left[v^m_{N_+} + v^{m+1}_{N_+}, \quad 0, \quad \ldots, \quad 0, \quad v^m_{-N_-} + v^{m+1}_{-N_-}\right]^T .$$

As previously, it can be show that $(\mathbb{1} + \mathbf{M})$ and $(3\mathbb{1} - \mathbf{M})$ are invertible matrices (since they are tri-diagonal matrices with nonzero diagonal elements) and they are inverted efficiently using the tri-diagonal property using LU decomposition algorithm.

5 Application

The aim of this section is to apply the discretization methods described previously to price European options with transaction cost. We either explore Leland model through implicit and explicit technique and Amster model using the explicit one.

Explicit Method

Black Scholes and Leland model

Using backward difference (19) for the time derivative and the central difference (20) for first and second Stock derivative in (5) we obtain:

$$F_n^{m-1} = \alpha_n F_{n-1}^m + (1-\gamma_n) F_n^m + \beta_n F_{n+1}^m \,, \tag{28}$$

where we have defined:

$$\alpha_n = \frac{1}{2}\left(\sigma^2 n^2 - rn\right) dt \,, \quad \beta_n = \frac{1}{2}\left(\sigma^2 n^2 + rn\right) dt \,, \quad \gamma_n = rdt + \alpha_n + \beta_n \,.$$

Since $\alpha_n + \beta_n + 1 - \gamma_n = 1 - rdt$, eqn. (28) has a probabilistic interpretation, as expected. Boundary conditions on the European option for large S and for $S = 0$ makes the F_N^m and F_0^m coefficients known for every time step m. Additionally, the final time condition (the payoff of the option) makes the F_n^M coefficients known for all n. Therefore the algorithm implemented computes F_n^{M-1} from (28) for all n and then repeats it process for every time step m up to $m = 0$.

This process can be efficiently written as:

$$\mathbf{M}\mathbf{F}^m = \mathbf{F}^{m-1} \equiv \tag{29}$$

$$\begin{bmatrix} \alpha_1 & 1-\gamma_1 & \beta_1 & 0 \\ \ddots & \ddots & \ddots & \\ 0 & \alpha_{N-1} & 1-\gamma_{N-1} & \beta_{N-1} \end{bmatrix} \begin{bmatrix} F_0^m \\ \vdots \\ F_N^m \end{bmatrix} = \begin{bmatrix} F_1^{m-1} \\ \vdots \\ F_{N-1}^{m-1} \end{bmatrix}$$

where $\mathbf{M}$ is a rectangular matrix of $(N-1)\times(N+1)$. Hence for every time step $m+1$, the value of the portfolio can be obtained by the multiplication of the matrix $\mathbf{M}$ and the price of the portfolio at time step m. It worth to note that for the standard Black-Scholes model $\mathbf{M}$ does not depend on m. This result does not hold when transaction costs are taken into account through Leland model. Even though equation (28) remains, $\sigma \to \tilde{\sigma}$ and $\mathbf{M}$ will depend on m index.

Amster model

Using backward difference (19) for the time derivative and the central difference (20) for first and second Stock derivative in (17) we obtained:

$$F_n^{m-1} = \alpha_n F_{n-1}^m + (1-\gamma_n) F_n^m + \beta_n F_{n+1}^m + bn^3\sigma^2 \left(F_{n-1}^m - 2F_n^m + F_{n+1}^m\right)^2 \frac{dt}{dS} \,,$$

where $\alpha_n, \beta_n, \gamma_n$ defined as previously with $k = 2a$.

This process can be written as:

$$\mathbf{F}^{m-1} = \mathbf{M}\mathbf{F}^m + b\sigma^2 \frac{dt}{dS} \quad \mathbf{\Gamma}^{\dagger m}\mathbf{B}\mathbf{\Gamma}^m$$

with $\mathbf{B} = i^3\delta_{ij}$ and $\mathbf{\Gamma}_n^m = F_{n-1}^m - 2F_n^m + F_{n+1}^m$[3]

Implicit Method

Black Scholes and Leland model

Using forward difference (18) for the time derivative and the central difference (20) for first and second Stocks derivative in (14) we obtain the recurrence relation:

$$F_n^{m+1} = -\alpha_n F_{n-1}^m + (1+\gamma_n) F_n^m - \beta_n F_{n+1}^m \,, \tag{30}$$

or written in a matricial form:

$$\mathbf{M}\mathbf{F}^{m-1} = \mathbf{F}^m + \mathbf{b}^{m-1} \equiv \tag{31}$$

$$\begin{bmatrix} 1+\gamma_1 & -\beta_1 & & 0 \\ -\alpha_2 & 1+\gamma_2 & \ddots & \\ & \ddots & \ddots & -\beta_{N-2} \\ 0 & & -\alpha_{N-1} & 1+\gamma_{N-1} \end{bmatrix} \begin{bmatrix} F_1^{m-1} \\ \vdots \\ F_{N-1}^{m-1} \end{bmatrix} = \begin{bmatrix} F_1^m \\ \vdots \\ F_{N-1}^m \end{bmatrix} + \begin{bmatrix} \alpha_1 F_0^m \\ 0 \\ \vdots \\ 0 \\ \beta_{N-1} F_N^m \end{bmatrix}$$

A couple of observations should be made here. Vectors are of length $N-1$, while $\mathbf{M}$ matrix is of size $(N-1)\times(N-1)$. Final condition is imposed when $m = M$, while the boundary condition is incorporated in the vector $\mathbf{b}^{m-1}$.

Numerical Accuracy

As with any form of numerical analysis, it is important to consider the accuracy of the algorithms we have coded and the data we have produced.

In particular we should be careful to ensure that our algorithms are stable, that error is kept to a minimum and by extension that our output is converging.

Stability

Standard matrix algebra dictates that for a difference equation system $y_t = b + Ay_{t-1}$ the fundamental necessary and sufficient stability condition for the system is $\mu(A) < 1$, where $\mu(A)$ is the spectral radius of A. There exists a class of scalar real-valued functions of matrices called matrix norms; and any matrix A

[3] In fact, the last term in python can be easily written as a product of two column vectors.

has the property $\mu(A) < f(A)$. Hence, a sufficient stability condition is $f(A) < 1$, where $f(A)$ is any matrix norm function. The advantage of demonstrating existence, uniqueness, and stability through a norm $f(A)$ rather than through the spectral radius itself is that many norms are far easier to compute but it is good to remember that the bound may be far stronger than necessary. By definition A matrix norm is a real-valued scalar function of a square matrix such that satisfies: (i) $f(A) > 0$ with equality holding if and only if $A = 0$; (ii)$f(cA) = |c|f(A)$ for any scalar c; (iii) $f(AB) < f(A)f(B)$ and (iv) $f(A+B) < f(A) + f(B)$.

In our case we have selected:

$$f(A) = ||A||_\infty = max\left(\sum_j |A_{0j}|, \sum_j |A_{1j}|, ...\right)$$

It worth to note here that for the explicit method $A = M$, where M have been defined in (29) while for the implicit method $A = M^{-1}$ where M have been defined in (31). Our program incorporates a function that test the stability of the method through this norm. For further details or different possible implementations of matrix norms we refer the reader to [14].

Convergence

Another element worth considering is the rate our algorithm converges, which is directly related to the truncation error we introduced in our model when we approximated the PDE. As such, our algorithm converges at a rate of $\mathcal{O}(\delta t)$ and $\mathcal{O}((\delta S)^2)$. To calculate the truncation error in our data we considered the closed form solution for a European call option as the reference value to compare.

6 Summary, Remarks and Conclusions

Summary and Remarks

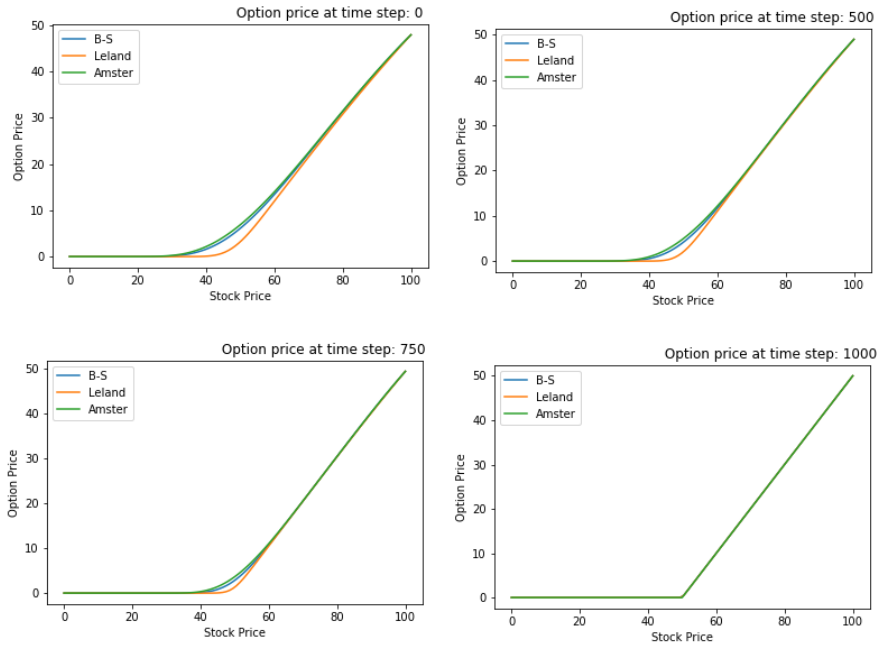
In the present work we set up numerical simulations to explore the effects of different transaction cost models when they are incorporated into the conventional Black-Scholes model. We analyze Leland and Amster model using implicit/explicit finite difference techniques. This models were programmed in python in OOP paradigm to obtain a closer approach to a production code utilized by the industry. For efficiency reasons the code is fully vectorized.

We have studied the numerical stability issues of both implicit/explicit method through the $||A||_\infty$ norm and we have compared the numerical computations of the B-S model with their analytical counterpart.

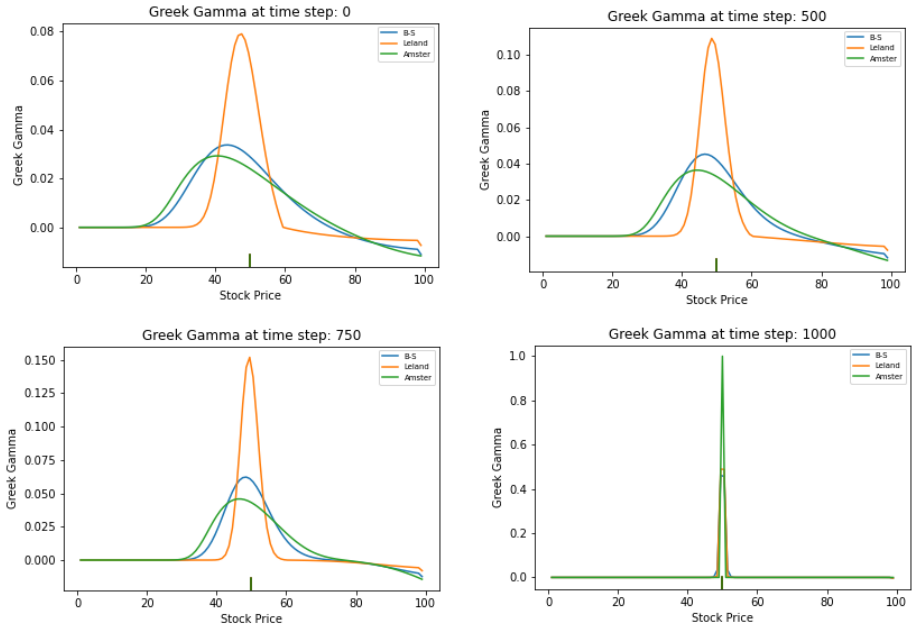
Conclusions

Both, Amster and Leland model result to be stable for a small range of parameters. In this stable region this models exhibits smalls corrections to the price

of the option computed through the B-S model.



The figures below represent the greek gamma for each of this models.



Further work

Optimization

Some open questions and further work remains. We did not have the time to test this results for a broader set of parameters and compare them with actual real data extracted from real markets. Additionally, a natural question would be if there is an optimal re-hedge time that maximize the value of the portfolio.

We have studied the numerical stability issues of both implicit/explicit method through the $||A||_\infty$ norm. Tighter bounds can be obtained studying the spectral radius. Consistency of the methods implemented is another debt.

Amster model revisited.

In spite of the mentioned limitations of the numerical methods to extract valuable information of the Amster model, we would like to briefly describe another alternative path. The ideas developed below are inspired in perturbation techniques.

Lets suppose $\hat{\mathcal{H}}_0 = \frac{1}{2}\sigma^2\partial_S^2 + rS\partial_S + \partial_t - r$ is the standard BS linear operator and $\psi_0(S,t)$ is the known solution to the equation $\hat{\mathcal{H}}_0\psi_0(S,t) = 0$. Now we would like to study the problem $\hat{\mathcal{H}}\psi = (\hat{\mathcal{H}}_0 + \lambda\hat{V})\psi = 0$ where $\hat{V}$ is some perturbation to $\hat{\mathcal{H}}_0$ controlled by the parameter λ.

If $\psi(S,t)$ is an analytical function of λ then it can be written as:

$$\psi = \psi_0 + \lambda\psi_1 + ... + \lambda^n\psi_n\,. \tag{32}$$

This summation will be studied up to $n = N$. In order to be ψ compatible with the standard initial and boundary conditions, is reasonable to propose that each ψ_i satisfy the following conditions:

$$\psi_i(S,t) \to \tfrac{1}{(N+1)\lambda^i}S \quad \text{if} \quad S \to \infty$$

$$\psi_i(S,t) \to 0 \quad \text{if} \quad S \to 0$$

$$\psi_i(S,t) = \frac{1}{(N+1)\lambda^i}max(S_T - K, 0) \quad \text{if} \quad t = T$$

with $i = 0, 1, .., N$.

Using eq. (32) we can write:

$$\hat{\mathcal{H}}\psi = \hat{\mathcal{H}}_0\psi_0 + \lambda(\hat{\mathcal{H}}_0\psi_1 + \hat{V}\psi_0) + O(\lambda^2)\,.$$

From above it follows that $\hat{\mathcal{H}}\psi = 0$, $\forall\lambda$ if and only if every coefficient in the polynomial expansion vanishes.

For the Amster model (17) with $a = 0$:

$$\hat{V}\psi_0 = \sigma^2 S^3\left(\frac{\partial^2\psi_0}{\partial S^2}\right)^2 = \frac{K}{2\pi\tau}e^{\left(3r - \frac{\sigma^2}{4}\right)\tau}e^{-\frac{1}{\sigma^2\tau}\left(x + \frac{\sigma^2\tau}{2}\right)^2}$$

In resume, in order to obtain first order correction we need to solve the non-homogeneous BS equation $(\hat{\mathcal{H}}_0\psi_1 + \hat{V}\psi_0) = 0$, where the term $\hat{V}\psi_0$ is analytically known. Using the transformations described in (7) we map this non-homogeneous equation to the non-homogeneous heat equation:

$$\frac{\partial u}{\partial \tau} - \frac{\sigma^2}{2}\frac{\partial^2 u}{\partial x^2} = f$$

From an analytical point of view this problem can be tackled either by convolution of f with the corresponding Green functions or using Fourier Transformation but this margin is too narrow to contain all the computations.

At last, if we consider $a \neq 0$, then $\hat{\mathcal{H}}_0$ should be redefined and solved numerically (Leland model). The solutions obtained can be taken as the entry solutions of the perturbative model.

References

[1] Fischer Black and Myron Scholes. "The Pricing of Options and Corporate Liabilities". en. In: *Journal of Political Economy* 81.3 (May 1973), pp. 637–654. ISSN: 0022-3808, 1537-534X.

[2] Robert C. Merton. "Theory of Rational Option Pricing". In: *The Bell Journal of Economics and Management Science* 4.1 (1973), p. 141. ISSN: 00058556.

[3] Hayne E. Leland. "Option Pricing and Replication with Transactions Costs". en. In: *The Journal of Finance* 40.5 (Dec. 1985), pp. 1283–1301. ISSN: 00221082.

[4] Phelim P. Boyle and Ton Vorst. "Option Replication in Discrete Time with Transaction Costs". en. In: *The Journal of Finance* 47.1 (Mar. 1992), pp. 271–293. ISSN: 00221082.

[5] Guy Barles and Halil Mete Soner. "Option pricing with transaction costs and a nonlinear Black-Scholes equation". In: *Finance and Stochastics* 2.4 (Aug. 1998), pp. 369–397. ISSN: 0949-2984, 1432-1122.

[6] Rudiger Frey and Alexander Stremme. "Market Volatility and Feedback Effects from Dynamic Hedging". en. In: *Mathematical Finance* 7.4 (Oct. 1997), pp. 351–374. ISSN: 0960-1627, 1467-9965.

[7] Paul Wilmott and Philipp J. Schönbucher. "The Feedback Effect of Hedging in Illiquid Markets". en. In: *SIAM Journal on Applied Mathematics* 61.1 (Jan. 2000), pp. 232–272. ISSN: 0036-1399, 1095-712X.

[8] Rüdiger Frey and Pierre Patie. "Risk Management for Derivatives in Illiquid Markets: A Simulation Study". en. In: *Advances in Finance and Stochastics*. Ed. by Klaus Sandmann and Philipp J. Schönbucher. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002, pp. 137–159. ISBN: 9783642077920 9783662047903.

[9] H. Mete Soner and Nizar Touzi. "Superreplication Under Gamma Constraints". en. In: *SIAM Journal on Control and Optimization* 39.1 (Jan. 2000), pp. 73–96. ISSN: 0363-0129, 1095-7138.

[10] Ali Hirsa and Salih N. Neftci. *An introduction to the mathematics of financial derivatives*. Third edition. San Diego, CA, USA: Academic Press, 2014. ISBN: 9780123846822.

[11] P. Amster et al. "A Black–Scholes option pricing model with transaction costs". en. In: *Journal of Mathematical Analysis and Applications* 303.2 (Mar. 2005), pp. 688–695. ISSN: 0022247X.

[12] Paul Wilmott, Jeff Dewynne, and Sam Howison. *Option pricing: mathematical models and computation*. Repr. with corrections. Oxford, UK: Oxford Financial Press, 2000. ISBN: 9780952208204.

[13] Paul Wilmott, Sam Howison, and Jeff Dewynne. *The mathematics of financial derivatives: a student introduction*. Oxford ; New York: Cambridge University Press, 1995. ISBN: 9780521496995 9780521497893.

[14] John Conlisk. "Quick Stability Checks and Matrix Norms". In: *Economica* 40.160 (Nov. 1973), p. 402. ISSN: 00130427.