COMPUTATIONAL METHODS FOR PRICING AMERICAN OPTIONS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
ATILIM UNIVERSITY

BY

BURCU AYDOĞAN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
MATHEMATICS

JULY 2014

Approval of the Graduate School of Natural and Applied Sciences, Atılım University.

Prof. Dr. İbrahim Akman
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of **Master of Science in Mathematics Department, Atılım University**.

Prof. Dr. Tanıl Ergenç
Head of Department

This is to certify that we have read the thesis COMPUTATIONAL METHODS FOR PRICING AMERICAN OPTIONS submitted by BURCU AYDOĞAN and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

Assoc. Prof. Dr. Ömür Uğur
Co-Supervisor

Assist. Prof. Dr. Ümit Aksoy
Supervisor

Examining Committee Members:

Prof. Dr. Sofiya Ostrovska
Mathematics Department, Atılım University

Assoc. Prof. Dr. Ömür Uğur
Institute of Applied Mathematics, METU

Assoc. Prof. Dr. İnci Erhan
Mathematics Department, Atılım University

Asst. Prof. Dr. Yeliz Yolcu Okur
Institute of Applied Mathematics, METU

Asst. Prof. Dr. Ümit Aksoy
Mathematics Department, Atılım University

Date: July 18, 2014

I declare and guarantee that all data, knowledge and information in this document has been obtained, processed and presented in accordance with academic rules and ethical conduct. Based on these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name : BURCU AYDOĞAN

Signature :

ABSTRACT

COMPUTATIONAL METHODS FOR PRICING AMERICAN OPTIONS

Aydoğan, Burcu

M.S., Department of Mathematics

Supervisor : Assist. Prof. Dr. Ümit Aksoy

Co-Supervisor : Assoc. Prof. Dr. Ömür Uğur

July 2014, 111 pages

In financial mathematics, option pricing is a popular problem in theory of finance and mathematics. In option pricing theory, the valuation of American options is one of the most important problems. American options are the most traded option styles in all financial markets. In spite of the recent developments, the valuation of American options continues to exist as a challenging problem. There are no closed-form analytical solutions of American options, so that a usual way to deal with this problem is to develop numerical and approximation techniques.

In this thesis, we analyze binomial, finite difference and approximation methods, for pricing American options. We first consider the binomial approximation which is very easy to implement and assumes that the asset prices follow from geometric Brownian motion. Then, we present American options as a free boundary value problem based on Black-Scholes partial differential equation, which leads to a very famous model in finance theory, and formalize it as a linear complementarity problem. We refer to the projected successive over relaxation (PSOR) method to solve this problem. Although there are no closed-form solutions for American options, we deal with some analytical approximation methods to approach the option values. We demonstrate the applications of the each method and compare their solutions.

Keywords: American Options, Black-Scholes Equation, Binomial Method, Finite Difference Methods, Approximation Methods

ÖZ

AMERİKAN OPSİYONLARININ HESAPLAMALI YÖNTEMLERLE FİYATLANDIRILMASI

Aydoğan, Burcu

Yüksek Lisans, Matematik Bölümü

Tez Yöneticisi : Yrd. Doç. Dr. Ümit Aksoy

Ortak Tez Yöneticisi : Doç. Dr. Ömür Uğur

Temmuz 2014, 111 sayfa

Finansal matematikte, opsiyon fiyatlama finansal teori ve matematiksel olarak düşünüldüğünde çok popüler bir problemdir. Opsiyon fiyatlama teorisinde, Amerikan opsiyonlarının fiyatlandırılması en önemli problemlerden biridir. Amerikan opsiyonları, finansal piyasalarda en çok işlem gören opsiyon türüdür. Son zamanlardaki birçok gelişmeye rağmen, Amerikan opsiyon fiyatlandırması hala en zor problemlerden biri olmaya devam etmektedir. Amerikan opsiyonlarının kapalı analitik çözümleri yoktur, bu sebeple bu problemle ugraşmanın en yaygın yollarından biri sayısal ve yaklaşım teknikleri geliştirmektir.

Bu tezde, Amerikan opsiyonlarını fiyatlandırmak için hesaplamalı metotlardan; binom, sonlu fark ve yaklaşım metotları analiz edilmiştir. İlk olarak, uygulaması çok kolay olan ve varlık fiyatlarının geometrik Brownian hareketinden geldiğini varsayan binom yaklaşımı ele alınmıştır. Daha sonra, Amerikan opsiyonları için Black-Scholes kısmi diferansiyel denklemine dayanarak serbest sınır değer problemi verilmiştir. Bu problemi çözmek için PSOR metodu kullanılmıştır. Amerikan opsiyonlarının kapalı çözümleri olmamasına rağmen, opsiyonun değerine çok yaklaşan bazı analitik yaklaşım metotları üzerinde çalışılmıştır. Her bir

metodun uygulamaları yapılmış ve çözümler karşılaştırılmıştır.

Anahtar Kelimeler: Amerikan Opsiyonları, Black-Scholes Denklemi, Binom Yöntemi, Sonlu Fark Yöntemi, Yaklaşım Yöntemi

To My Family

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my supervisor, Asst. Prof. Dr. Ümit Aksoy, for her excellent guidance, support and encouragement throughout my thesis study.

My deepest gratitude also goes to my co-supervisor, Assoc. Prof. Dr. Ömür Uğur, for his advices, guidance and kindness during the completion of this thesis.

I am so grateful to Asst. Prof. Dr. Yeliz Yolcu Okur for her valuable comments and help she provided at each step of my study.

My thanks also go to Head of the Department Prof. Dr. Tanıl Ergenç and all staff of the Department of Mathematics for their support.

Furthermore, I thank the members of my thesis committee, Prof. Dr. Sofiya Ostrovska and Assoc. Prof. Dr. İnci Erhan for their insightful comments.

My special thanks go to Asst. Prof. Dr. Mehmet Turan for his all kind of help and motivation.

Many thanks to my friends Nazlı Çiloğlu and İpek Çağla Yıldız supported me during my study.

Finally, I wish to express my warmest thanks to my mother Güler, my father Nihat, my brothers Aytaç and Ömer for their love and support throughout my life.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

CHAPTER 1

INTRODUCTION

Options are widely used financial derivatives, contracts derived from underlying assets, which are traded in the world financial markets and, bought and sold more than the underlying stocks themselves. Today, option pricing is a very appreciated theory of the mathematical finance. An option is a financial instrument that gives the holder the right, but not the obligation, to buy or sell a prescribed asset, for a prescribed amount, by a prescribed time or expiration date. The option to buy an asset is called a call option, an option to sell the asset is known as a put option.

A European option gives the holder of the option to buy or sell the financial instrument for a certain price at the expiration date of the contract. Fischer Black and Myron Scholes first published a paper [9] in 1973, providing a formula for pricing European style options where the stock price follows a geometric Brownian motion. They derived a second order linear partial differential equation, which is called Black-Scholes partial differential equation, for pricing European options with the corresponding boundary conditions and gave a closed-form solution for the problem. The Black-Scholes partial differential equation is given as

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0.$$

This paper [9] led to a great activity in option markets and scientific research in financial mathematics. Fischer Black and Myron Scholes won Nobel Prize in 1994 for their excellent contributions to the theory finance. In the sequel we will consider the Black-Scholes model in details.

An American option is an option that gives the option holder the right to exercise at any time before the expiration date. Therefore, the holder may choose whether or not to exercise the option up to that time. Thus, an American option has the additional feature than that of a European option; it can be exercised at any time during the life of the option. For the value of American options, there is no closed-form formula and it is much more harder to find the price comparing with European options; since the holder of American options has the right to early exercise the option at any time before the expiration date. Thus, Black-Scholes equation should be solved with a free boundary simultaneously. A natural way to study this problem is to apply computational methods of whose progress is a hot and productive research area in recent years.

There have been developed different numerical methods and techniques to solve the American option problem. The Binomial model which was developed by Cox, Ross and Rubenstein [28] is a simple approach that uses lattices to represent the future possibilities of the underlying asset of the option. Also, Boyle [11] extended the basic model to two assets case. Finite difference method gives efficient solutions to the linear complementarity problem for American options based on Black-Scholes partial differential equation. There are three well-known methods of the finite difference approximation: the explicit method which was introduced by Brennan and Schwartz [12], and then advanced by Courtadon [16] is obtained by using backward and central difference approximations. Schwartz and Brennan [40, 12] and Courtadon [16] provided the implicit method formed by forward difference approximation. The approach improved by Crank and Nicolson [17] suggest to combine the explicit and implicit methods. After obtaining the finite difference formulation for linear complementarity problem, the projected successive over relaxation (PSOR) method which was developed by Cryer [18] is used to solve the problem for American options iteratively.

Despite the fact that there is no closed-form formula for American options, some researches attempt to search some analytical methods to approximate the value of the option. Specifically, Roll-Geske-Whaley [39, 22, 23, 47], Barone-Adesi Whaley [3] and Bjerksund-Stensland [7, 8] gave approximated closed-form solutions under suitable assumptions. Moreover, Longstaff and Schwartz [36] implemented a path-wise approximation by comparing the exercise value of the option with the expected payoff from the continuation at any possible exercise time.

This thesis is devoted to study of numerical methods for pricing American style options. We focus on Binomial, Finite Difference and Approximation methods with their theoretical backgrounds and with MATLAB implementations including graphical representations. Our aim is to compare all these methods. The thesis is structured as follows: In Chapter 2, we give a complete discussion about American options and a continuous model for underlying stock price. Chapter 3 first presents the Binomial method for European options on one-step and multi-step binomial trees, then modifies the idea to obtain American option values on non-dividend, discrete and continuously compounded dividend paying assets. In Chapter 4, Finite Difference method is introduced under the Explicit, Implicit and Crank-Nicolson approximations for the Black-Scholes partial differential equation, then the θ-averaged method is discussed as a generalization of the Crank-Nicolson method in detail. For the American free boundary value problem, we formalize the θ-averaged method and use it in connection with the PSOR iteration technique. For particular examples, prices obtained from Binomial and Finite Difference methods are compared. Chapter 5 includes the theory and applications of the analytical approximation models widely used in literature, and at the end of this chapter, the methods considered thus far are compared for some specific examples. Finally, Chapter 6 is devoted to discussion of the efficiency and applicability of the methods investigated to conclude the thesis. Basic definitions and preliminary theorems of financial mathematics and MATLAB programs are compiled in two respective appendices at the end.

CHAPTER 2

AMERICAN OPTIONS AND CONTINUOUS MODEL FOR STOCK PRICES

In this chapter, we give some important definitions and results that are necessary for the concept of pricing options. For more detailed discussion, we refer to Shreve [42], Lamberton and Lapeyre [35], Hull [24], Seydel [41] and Willmott [38].

2.1 Options

An option gives the buyer of the contract the right, but not the obligation, to buy or sell an asset, by a certain date, for a certain strike price. The option to buy an asset is known as a call option while the option to sell is a put. In an option contract, the followings should be specified:

- the underlying asset: a stock, a bond or a currency,
- the amount of an underlying asset to be bought or sold,
- the expiration date or expiry date which is a prescribed time in the future: if the option can be exercised at any time before maturity, it is called an *American* option but, if it can only be exercised at maturity, it is called a *European* option,
- the exercise price (strike price) which is the prescribed amount.

The value of an option is denoted by V. The value of V depends on the price per share of the underlying, which is denoted by S. The letter S symbolizes stocks, which are the most prominent examples of underlying assets. The variation of the asset price S with time t is expressed by S_t or $S(t)$. The value of the option also depends on the remaining time to expiry $T-t$. That is, V depends on time t. The dependence of V on S and t is written as $V(S,t)$. The underlying assets may be dividend or non-dividend paid where *dividend* is a cash payment made to the owner of a stock. A dividend paying stock receives additional shares of stock when a company declares a stock dividend, while a non-dividend paying stock means that the stock is currently paying no dividends.

Let T be the expiration date and K be the strike price. If $K > S_T$, the holder of the option will not exercise the option since the asset can be purchased on the market for the cheaper price. But, if $S_T > K$, the holder makes a profit of $S_T - K$ by exercising the option. Therefore, the value of the European call at maturity is given by

$$C(S_T, T) = \max(S_T - K, 0) = (S_T - K)^+$$

which is called the *payoff* function where C denotes the value of the call option.

For a European put option, the payoff function is given by

$$P(S_T, T) = \max(K - S_T, 0) = (K - S_T)^+$$

where P denotes the value of the put option. A payoff diagram is a graph of the payoff function at expiration $t = T$ as a function of the underlying stock S.

American options and early exercise feature

Let the price of the American call option be denoted by C_A and the American put option by P_A. Then, the payoff of American call at maturity time T is

$$C_A(S_T, T) = \max(S_T - K, 0).$$

The payoff of an American put option is

$$P_A(S_T, T) = \max(K - S_T, 0).$$

An American option has an additional feature when compared to a European option that the holder can exercise the option at any time during the life of the

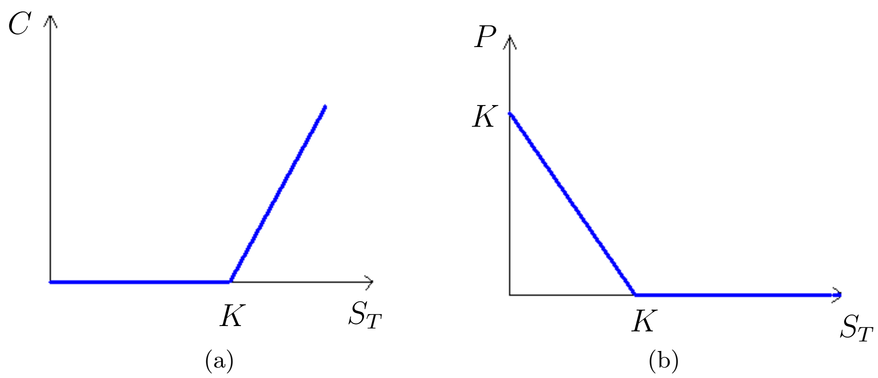

Figure 2.1: Payoff diagram for a (a) call option and (b) put option.

option. Since American options give the holder early exercise right, it is expected to have a higher value than the value of a European option.

Before the expiry date, the value of a European put option is less than its payoff such that $P(S,t) < \max\{K - S, 0\}$. When considering the effect of exercising the option, we can buy the asset in the market for price S and if we immediately exercise the option by selling the asset for price K buying the option for P, then we have a risk-free profit of $K - P - S$. Therefore, when the early exercise is permitted, we should have the constraint

$$V(S,t) \geq \max\{S - K, 0\}$$

so that European and American options will indeed have different values.

On the other hand, for a call option on a dividend paying asset, we should have the constraint

$$C(S,t) \geq \max\{S - K, 0\}.$$

In both cases, there is a value of S for which it is optimal to exercise the American options.

Consequently, the valuation of American options is more complicated, since we should check both the option value and whether it is optimal or not for each value of S at any time t. This is known as a free boundary problem. We will consider in details the free boundary value problem for American options and methods for solving it in Chapter 4. Here are two definitions that might be helpful in further references.

Definition 1. A *portfolio* is the combination of assets, options and bonds. The value of portfolio is denoted by Π_t at time t.

Definition 2 (Time value of money)**.** The value at time $t = T$ of investing an amount P is found as

$$M = Pe^{-r(T-t)}$$

under continuously compounded interest rate r.

No-Arbitrage Principle

When a zero initial investment Π_0 is identified that guarantees non-negative profit in the future such that Π_T, then an arbitrage opportunity, that is, riskless profit arises. We assume from now on that there exists no arbitrage opportunity such that all risk-free portfolios must have the same return which is the risk-free interest rate. Simply, no-arbitrage principle can be stated as there are never opportunities to make risk-free profit.

Let us denote by C_t and P_t, respectively, the prices of the European call and put options at time t. The absence of arbitrage opportunity gives the following equation, the so-called *put-call parity*, which is true for all $t < T$:

$$C_t - P_t = S_t - Ke^{-r(T-t)}.$$

The put-call parity inequality for the American option, therefore, is given by

$$S - K \leq C_A - P_A \leq S - Ke^{-r(T-t)},$$

where C_A is the value of the American call and P_A is the value of the American put option.

Since the American style exercise contains that of a European at time $t = T$ because of the early exercise feature, the value of an American option is always greater than the value of a European option. That is,

$$V_A \geq V \tag{2.1}$$

under the same contract conditions, which can be proved using no-arbitrage principle. For the case of underlying non-dividend paying assets, it can be seen again by no-arbitrage principle that the American and European call options must have the same value $V_A = V$. Nevertheless, the following lemma states the situation of an American call on dividend paying underlying asset.

Lemma 2.1.1 ([24]). *It is optimal to exercise an American call option with underlying dividend paying stock is just before the ex-dividend date, which is the time on which a security has the right to the most recently announced dividend anymore.*

Remark 2.1.2. For the case of more than one dividend payment, it can be shown that it may be optimal to exercise the American call option just before the ex-dividend date. See [24] for the details.

See [35] and [41] for some other properties of American options.

2.2 Continuous model for stock prices

In this section, we consider a model for stock price, therefore, give some essential theory that we will refer to in the following chapters. In Appendix A, we give the necessary background material to trace this section.

Definition 3. A *stochastic process* is a family of random variables X_t, which are defined for a set of parameters t.

Definition 4 (Brownian motion, Wiener process)**.** A Wiener process (or Brownian motion; notation W_t or W) is a time-continuous process for $t \geq 0$ with the following properties.

(a) $W_0 = 0$,

(b) $W_t \sim N(0, t)$ for all $t \geq 0$. That is, for each t the random variable W_t is distributed normally, with mean $\mathbb{E}[W_t] = 0$ and variance $\mathrm{Var}[W_t] = \mathbb{E}[W_t^2] = t$.

(c) All increments $\Delta W_t := W_{t+\Delta t} - W_t$ on non overlapping time intervals are independent: That is, the displacements $W_{t_2} - W_{t_1}$ and $W_{t_4} - W_{t_3}$ are independent for all $0 \leq t_1 < t_2 \leq t_3 < t_4$.

(d) W_t depends continuously on t.

The increment ΔW_t, sometimes, is written as

$$\Delta W_t = X(\Delta t)^{\frac{1}{2}}$$

where $X \sim N(0,1)$ so that

$$\mathbb{E}[\Delta W] = 0 \quad \text{and} \quad \mathbb{E}[(\Delta W)^2] = \Delta t.$$

The probability density function for W_t is

$$f(y,t) = \frac{1}{\sqrt{2\pi t}} e^{\frac{-y^2}{2t}}.$$

Algorithm 1 Simulation of a Brownian motion

```
Given t_0 = 0, W_0 = 0, Δt.
for j = 1, 2, ... do
    set t_j = t_{j-1} + Δt
    draw Z ~ N(0,1), Z is a random variable
    W_j = W_{j-1} + Z √Δt
end for
```

Definition 5. An Itô stochastic differential equation (SDE) is

$$dX_t = a(X_t,t)dt + b(X_t,t)dW_t; \tag{2.2}$$

this together with $X_{t_0} = X_0$ is a symbolic short form of the integral equation

$$X_t = X_{t_0} + \int_{t_0}^{t} a(X_s,s)ds + \int_{t_0}^{t} b(X_s,s)dW_s.$$

The term $a(X_t,t)$ is called the drift term or the drift coefficient, while the term $b(X_t,t)$ is called the diffusion term.

For the detailed information about stochastic integrals, see [42], [35] and [41].

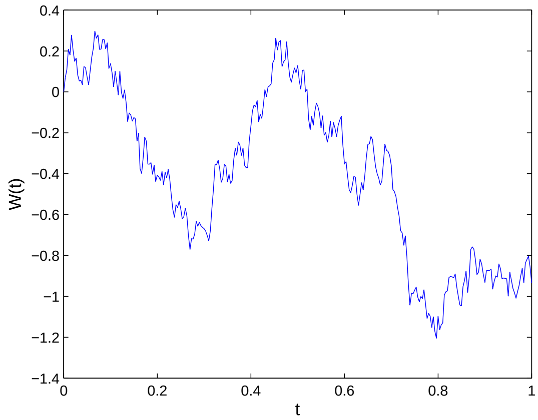

Figure 2.2: An example of paths of a Brownian motion

Note that the Brownian motion is a special case of an Itô process, since from $X_t = W_t$ the trivial SDE $dX_t = dW_t$ follows, hence the drift vanishes, $a = 0, b = 1$ in (2.2). Simulation of a Brownian motion is shown in Algorithm 2.2, and a simple path is depicted in Figure 2.2. Finally, note that $b \equiv 0$ and X_0 is constant, then the SDE becomes deterministic.

Definition 6 (Geometric Brownian motion)**.** A stochastic differential equation of the form

$$dS_t = \mu S_t dt + \sigma S_t dW_t \tag{2.3}$$

is called the *geometric Brownian motion* which is linear in $X_t = S_t$, and $a(S_t, t) = \mu S_t$ is the drift rate with the expected rate of return μ, $b(S_t, t) = \sigma S_t$, where σ is the volatility.

Lemma 2.2.1. *The closed-form solution of the geometric Brownian motion is given by*

$$S_t = S_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t}$$

where $S_{t_0} = S_0$ is the initial value of the asset price.

Proof. See [41]. □

Expectation and Variance of Geometric Brownian Motion

As is important for further investigations, expectation of a geometric Brownian motion S_t is

$$\begin{aligned}\mathbb{E}[S_t] &= \mathbb{E}[S_0 e^{(\mu-\frac{1}{2}\sigma^2)t+\sigma W_t}] \\ &= S_0 e^{(\mu-\frac{1}{2}\sigma^2)t}\mathbb{E}[\sigma W_t]\end{aligned}$$

Since $\mathbb{E}[\sigma W_t] = e^{\frac{\sigma^2}{2}t}$, then we have

$$\mathbb{E}[S_t] = S_0 e^{(\mu-\frac{1}{2}\sigma^2)t} e^{\frac{\sigma^2}{2}t} = S_0 e^{\mu t}.$$

Furthermore, since

$$\mathrm{Var}(S_t) = \mathbb{E}[S_t^2] - (\mathbb{E}[S_t])^2,$$

we first need to calculate $\mathbb{E}[S_t^2]$:

$$\begin{aligned}\mathbb{E}[S_t^2] &= \mathbb{E}[S_0^2(e^{(\mu-\frac{1}{2}\sigma^2)t+\sigma W_t})^2] = S_0^2\mathbb{E}[e^{2\mu t-\sigma^2 t+2\sigma W_t}] \\ &= S_0^2 e^{2\mu t-\sigma^2 t}\mathbb{E}[e^{2\sigma W_t}] = S_0^2 e^{2\mu t-\sigma^2 t}e^{2\sigma W_t}. \\ &= S_0^2 e^{(2\mu+\sigma^2)t}\end{aligned}$$

Thence,

$$\begin{aligned}\mathrm{Var}(S_t) &= S_0^2 e^{(2\mu+\sigma^2)t} - (S_0 e^{\mu t})^2 \\ &= S_0^2 e^{2\mu t}(e^{\sigma^2 t} - 1).\end{aligned}$$

Figure 2.3 depicts several paths of a geometric Brownian motion. Note that each path starts from $S_0 = 1$.

Finally, we conclude this chapter by stating an important lemma due to Itô, which is widely used in stochastic calculus.

Lemma 2.2.2 (Itô lemma)**.** *Let X_t for $t \geq 0$ be a stochastic Itô process defined by*

$$dX_t = a(X_t, t)dt + b(X_t, t)dW_t,$$

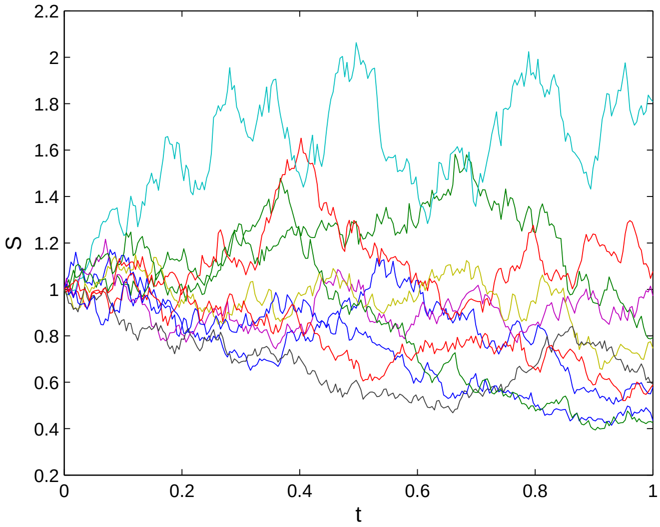

Figure 2.3: Paths of geometric Brownian motion

and let $g : (x,t) \mapsto g(x,t)$ *be a function for which the partial derivatives*

$$\frac{\partial g}{\partial x}, \quad \frac{\partial^2 g}{\partial x^2}, \quad \frac{\partial g}{\partial t}$$

are defined and continuous.

Then, $Y_t := g(X_t, t)$ *for every* $t \geq 0$ *is an Itô process, and*

$$dY_t = \left(\frac{\partial g}{\partial x} a + \frac{\partial g}{\partial t} + \frac{1}{2} \frac{\partial^2 g}{\partial x^2} \right) dt + \frac{\partial g}{\partial x} b \, dW_t$$

where the derivatives of g *as well as the coefficient functions* a *and* b *in general depend on the arguments* (X_t, t)*.*

CHAPTER 3

THE BINOMIAL METHOD

The binomial model is one of the ways for describing the random asset price dynamics. The model is a very useful technique to price stock options with a binomial tree because of its easy implementation. The method was first developed by Cox, Ross and Rubenstein [28]. The model assumes that the asset prices follow from the geometric Brownian motion. The tree which is used in the method represents the possibilities of the underlying asset price over the life of the option.

We will first describe the model for the simplest case which is European options and then, modify it for more complex versions, in particular, for American options.

3.1 Binomial Model for European Options

3.1.1 One-Step Binomial Tree

The idea of the binomial method is that at each period of time to the maturity, δ_t, the underlying non-dividend paying asset price with the initial value S_0 moves either up ($S_0 u$) with probability q by a fixed factor u, or down ($S_0 d$) with probability $1 - q$ by a factor d where the interest rate is constant. This movement can be represented as in Figure 3.1.

Before introducing the method for American options, we first consider the case of European options. Let V_0 be the current value of the European option and the

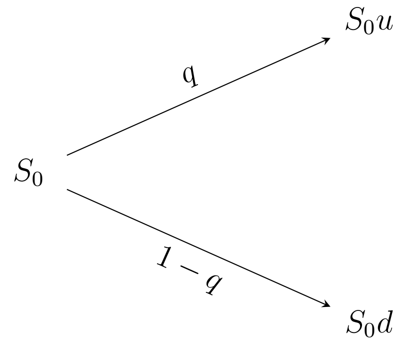

Figure 3.1: One-Step Binomial Tree

option yields V_u and V_d if the underlying asset moves up or down, respectively. Now, consider a portfolio Π_t with the initial value

$$\Pi_0 = V_0 - \Delta S_0$$

where Δ is the number of shares of the asset. Then, at time of maturity T, the possible outcomes of the portfolio are

$$\Pi_u = V_u - u\Delta S_0 \quad \text{ and } \Pi_d = V_d - d\Delta S_0.$$

We choose Δ according to the strategy to construct a portfolio which is riskless, that is,

$$\Pi_u = \Pi_d.$$

Consequently, we find

$$\Delta = \frac{V_u - V_d}{S_0(u - d)}. \tag{3.1}$$

If there were no riskless arbitrage opportunities, assuming continuously compounding risk-free interest rate r, then we should have

$$\Pi_{\delta_t} = e^{r\delta_t}\Pi_0 = e^{r\delta_t}(V_0 - \Delta S_0) = \Pi_u = \Pi_d,$$

which yields

$$\Pi_u = e^{r\delta_t}(V_0 - \Delta S_0) = V_u - u\Delta S_0$$

and hence,

$$V_0 e^{r\delta_t} = e^{r\delta_t}\Delta S_0 + V_u - u\Delta S_0. \tag{3.2}$$

Substituting (3.1) into (3.2), we obtain the value of the option as follows:

$$V_0 = e^{-r\delta_t}\{qV_u + (1-q)V_d\}, \tag{3.3}$$

where

$$q = \frac{e^{r\delta_t} - d}{u - d}.$$

Here, we may view q as the risk-neutral probability, say $\mathbb{Q}$. Under this probability measure,

$$\begin{aligned}\mathbb{E}_{\mathbb{Q}}[S_{\delta_t}] &= qS_0u + (1-q)S_0d \\ &= (e^{r\delta_t} - d)S_0 + S_0d \\ &= e^{r\delta_t}S_0. \end{aligned} \tag{3.4}$$

Therefore, the option value V_0 reduces to an expectation formula

$$V_0 = \mathbb{E}_{\mathbb{Q}}[e^{-r\delta_t}V_{\delta_t}] = e^{-r\delta_t}\mathbb{E}_{\mathbb{Q}}[V_{\delta_t}],$$

where

$$\mathbb{E}_{\mathbb{Q}}[V_{\delta_t}] = qV_u + (1-q)V_d$$

under the risk-neutral measure $\mathbb{Q}$.

Now, we start to determine the values of the parameters u and d in order to calculate the values V_u and V_d. From (3.4), we see that

$$qu + (1-q)d = e^{r\delta_t}$$

holds, and so does

$$\begin{aligned}\mathbb{E}_{\mathbb{Q}}[S_{\delta_t}^2] &= q(S_0u)^2 + (1-q)(S_0d)^2 \\ &= {S_0}^2 e^{(2r+\sigma^2)\delta_t}, \end{aligned} \tag{3.5}$$

as we assume that stock price S follows a geometric Brownian motion (6). Then the variance for the continuous model is

$$\begin{aligned}\mathrm{Var}(S_{\delta_t}) &= \mathbb{E}_{\mathbb{Q}}[(S_{\delta_t})^2] - (\mathbb{E}_{\mathbb{Q}}[S_{\delta_t}])^2 \\ &= {S_0}^2(e^{(2r+\sigma^2)\delta_t} - (e^{r\delta_t})^2) \\ &= {S_0}^2 e^{2r\delta_t}(e^{\sigma^2\delta_t} - 1),\end{aligned}$$

where σ is the volatility.

On the other hand, for one period binomial model, the variance is

$$\mathrm{Var}(S_{\delta_t}) = q{S_0}^2u^2 + (1-q){S_0}^2d^2 - {S_0}^2[qu + (1-q)d]^2.$$

By equating the variances of the asset price in both continuous and discrete models, we obtain

$$e^{2r\delta_t}(e^{\sigma^2\delta_t} - 1) = qu^2 + (1-q)d^2 - (e^{r\delta_t})^2,$$

from which it easily follows that

$$qu^2 + (1-q)d^2 = e^{(2r+\sigma^2)\delta_t}.$$

However, in order to determine the parameters u and d, another equation is needed. A convenient choice proposed by Cox, Ross and Rubenstein [28] for the other equation is

$$ud = 1,$$

so that the lattice nodes associated with the binomial tree are symmetric.

Now, we have two equations and two unknowns, u and d, to solve from the system:

$$\begin{aligned}qu^2 + (1-q)d^2 &= e^{(2r+\sigma^2)\delta_t} \\ ud &= 1\end{aligned}$$

where

$$q = \frac{e^{r\delta_t} - d}{u - d}.$$

Solving the system of nonlinear equations, we obtain the equation

$$u^2 - 2Au + 1 = 0, \tag{3.6}$$

where

$$A = \frac{1}{2}(e^{-r\delta_t} + e^{(2r+\sigma^2)\delta_t}).$$

Clearly, from (3.6), we have

$$u = A \mp \sqrt{A^2 - 1}.$$

Since $u > 1$ and $d < 1$ for up and down movements, we get

$$u = A + \sqrt{A^2 - 1}$$
$$d = A - \sqrt{A^2 - 1}.$$

Remark 3.1.1. The another choice for u and d was improved by Jarrow and Rudd [29] which differs from the model developed by [28] in that the mean and variance of the binomial tree matches that of the underlying process over any time step. The parameters are then

$$u = e^{(r-\frac{\sigma^2}{2})\delta_t + \sigma\sqrt{\delta_t}}$$
$$d = e^{(r-\frac{\sigma^2}{2})\delta_t - \sigma\sqrt{\delta_t}}$$
$$q = \frac{1}{2}.$$

3.1.2 Multi-Step Binomial Trees

We now extend the model to a multi-step binomial tree which gives more accurate results than one-step binomial model, since more payoff values will now be considered.

Let the stock price S_0 at time zero be known, and let us denote

$$\delta_t := \frac{T - t_0}{M},$$
$$t_i := t_0 + i\delta_t, \quad i = 0, 1, \ldots, M,$$
$$S_i := S_{t_i},$$

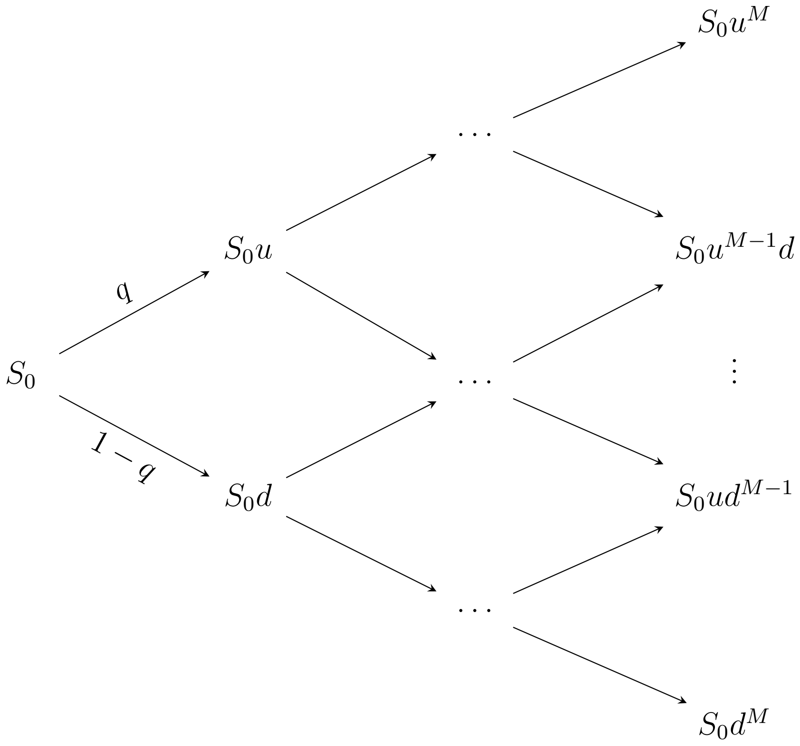

Figure 3.2: Multi-Step Binomial Trees

where M is the number of time steps, T is time to maturity and S_i is the asset price at time t_i.

Now that S will go up or down at each time steps, at time $t = t_i$, there are $i+1$ nodes in the tree and we shall denote the values of the stock prices as follows:

$$S_{ji} := S_0 u^j d^{i-j},$$

for each $i = 0, 1, \ldots, M$ and $j = 0, 1, \ldots, i$. The movements can be represented as in Figure 3.2.

The value of the European option, therefore, at each node is given by

$$V_{ji} := V(S_{ji}, t_i).$$

Now, we trace the lattice backward starting from (the payoff, $i = M$, and) $i = M - 1$ until $i = 0$ to find the value of $V_0 = V(S_0, t_0)$, the option price today. In the lattice, for every $i = M-1, M-2, \ldots, 0$ and $j = 0, 1, \ldots, i$, we have the

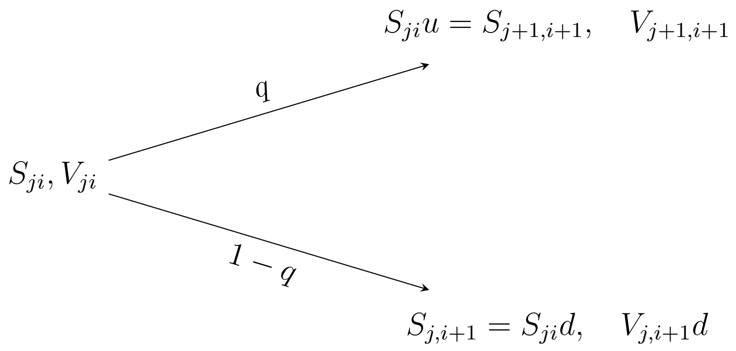

Figure 3.3: A branch of the tree of a Multi-Step Model

obvious relations

$$S_{ji}u = S_{j+1,i+1} \quad \text{and } S_{ji}d = S_{j,i+1}.$$

Fortunately, at the expiry date $T = M\delta t$, the payoff of the option is known, and thus

$$V_{jM} = V(S_{jM}, t_M) = V(S_{jM}, T)$$

for every $j = 0, 1, \ldots, M$. Therefore, the values V_{ji} associated to the nodes (S_{ji}, t_i) can be calculated for $i = M-1, M-2, \ldots, 0$ and $j = 0, 1, \ldots, i$ by the backward phase, depicted in Figure 3.3.

It follows that, for each time step the pricing formula (3.3) is generalized to

$$V_{ji} = e^{-r\delta_t}\{qV_{j+1,i+1} + (1-q)V_{j,i+1}\} \tag{3.7}$$

for $i = M-1, M-2, \ldots, 0$ and $j = 0, 1, \ldots, i$, where

$$q = \frac{e^{r\delta_t} - d}{u - d}.$$

Here is an example to illustrate the the multi-step Binomial model.

Example 1. Consider a European call option on a particular non-dividend paying asset. Today's asset price is \$20. For each month, stock will rise either by 1.25 or drop by 0.8 and the strike price is \$20. The time to maturity is $T = 3/12$ months, the risk-free interest rate is 20% and the volatility is 10%.

First, the parameters of the binomial tree are calculated to form the values of the asset. Then, the asset prices at the end of each month can easily be calculated to form the binomial tree. Therefore, we obtain approximately that the initial stock price of the call option is \$1.0561.

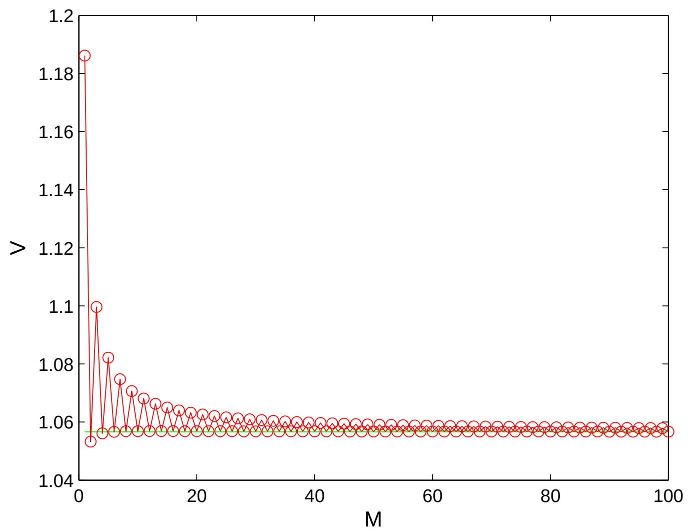

Figure 3.4: European call option by Binomial Tree

3.2 Binomial Model for American Options

European options can be exercised only at the maturity time, but American options can be exercised at any time before the maturity. Since the difference of American options from European options is that the American option gives the early exercise right to its holder, then in the binomial model, we have to check at every node if it is optimal to exercise early or not.

3.2.1 American Options on Non-Dividend Paying Assets

If we assume that the underlying asset pays no-dividend, the value of the American call option is the same as the value of the European call option. Thus, we only consider the case of the American put option with the non-dividend paying underlying asset.

Because of the early exercise feature of American options, equation (3.7) must be modified by adding an examine whether early exercise is to be preferred. The payoff of the American option at every node (t_i, S_{ji}) can be calculated as

$$\Lambda(S) := \Lambda(S_{ji}) := \max\{0, S_0 u^j d^{M-j} - K\} \tag{3.8}$$

for a call option, and

$$\Lambda(S) := \Lambda(S_{ji}) := \max\{0, K - S_0 u^j d^{M-j}\} \tag{3.9}$$

for a put option where $i = M-1, M-2, \ldots, 0$ and $j = 0, 1, \ldots, i$.

Now, at each node, the value of an American option at the node (t_i, S_{ji}) must be chosen as follows:

$$C_{ji}^A = \max\{e^{-r\delta_t}[qP_{j+1,i+1}^A + (1-q)P_{j,i+1}^A], \max\{0, S_0 u^j d^{M-j} - K\}\}$$

for a call option, and

$$P_{ji}^A = \max\{e^{-r\delta_t}[qP_{j+1,i+1}^A + (1-q)P_{j,i+1}^A], \max\{0, K - S_0 u^j d^{M-j}\}\}$$

for a put option for each $i = M-1, M-2, \ldots, 0$ and $j = 0, 1, \ldots, i$.

Example 2. The price of a non-dividend paying stock is currently for an American call option \$80. For each month, the stock price will go up 1.5 or down by 0.7, and the time to maturity is 3/12. The strike price is \$80 and the risk free interest rate is 20%. The volatility is 40%. The value of the American call option can be found as \$8.8351.

The value of the European option on the same conditions is also found as \$8.8351. This result confirms that the fact that American call options on non-dividend paying assets have the same value with the European call option under the same underlying asset and maturity date.

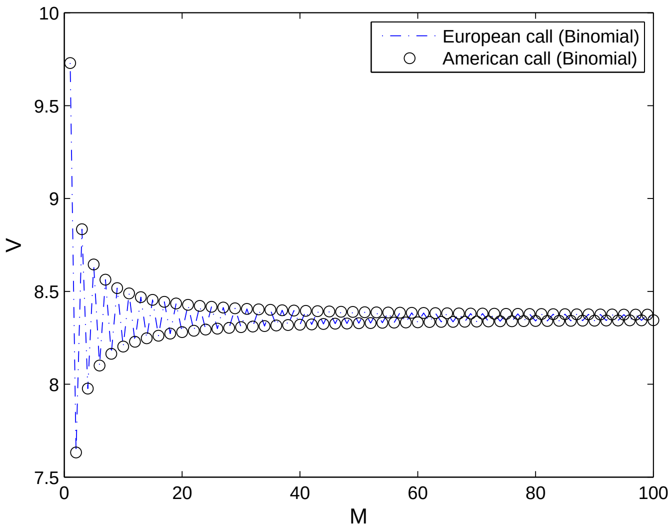

Figure 3.5: European and American call values by binomial method

For the case of American put option on the same stock with the same parameters, using binomial method, the value is found as \$5.1996.

Under the same situation, the European put option has the value \$4.9335.

Thus, the above figure illustrates the fact that the value of an American is always greater than the value of a European option which is stated in (2.1).

3.2.2 American Options on Discrete Dividend Paying Assets

Assume that the underlying asset pays one dividend D at time R. In this case, the values of the stock prices up to time t are expressed as

$$S_{ji} := S_0 u^j d^{i-j}$$

for each $i = 0, 1, \ldots, M$ and $j = 0, 1, \ldots, i$ and at time $R = kdt$, the price is

$$S_{ki} = S_0 u^{k-1} d^{k-1-j} - D,$$

and for the later times, stock prices will be found in usual way. The payoff of the American option at the node (t_i, S_{ji}) can be found as

$$\Lambda(S_{ji}) = \max\{0, S_{ji} - K\} \tag{3.10}$$

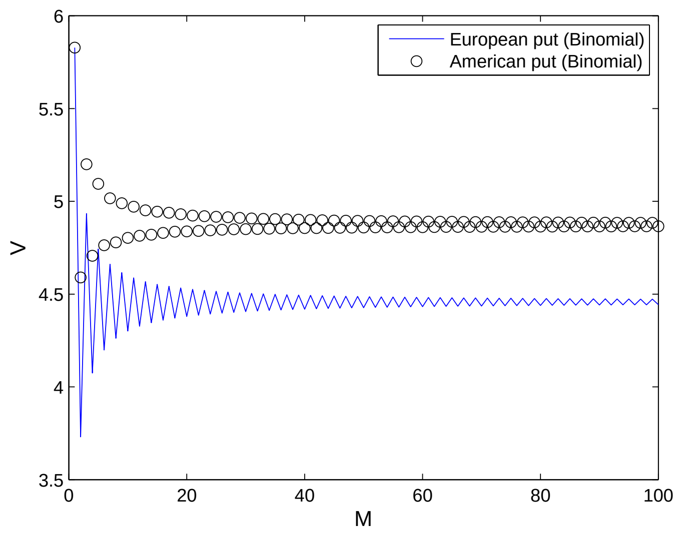

Figure 3.6: European and American put values by binomial method

for a call option, and

$$\Lambda(S_{ji}) = \max\{0, K - S_{ji}\} \tag{3.11}$$

for a put option.

Then, the value V_{ji}^A of an American option at the node (t_i, S_{ji}) is given as

$$V_{ji}^A = \max\{e^{-r\delta_t}[qV_{j+1,i+1} + (1-q)V_{j,i+1}], \Lambda(S_{ji})\}$$

for a call option, and

$$V_{ji}^A = \max\{e^{-r\delta_t}[qV_{j+1,i+1} + (1-q)V_{j,i+1}], \Lambda(S_{ji})\}$$

for a put option for each $i = M-1, M-2, \ldots, 0$ and $j = 0, 1, \ldots, i$, and

$$q = \frac{e^{r\delta_t} - d}{u - d}.$$

Example 3. Consider an American call option with the stock price \$60 and the discrete dividend payment $D = 0.2$ at time $t = 2/12$ where the time to maturity is $T = 4/12$. Over each of the next month, it will go up 1.1 or down by 0.9. The strike price is \$60 and the risk free interest rate is 10%. The volatility is 30%.

Using the binomial model with $ud = 1$, the current option price can be found as \$4.8114.

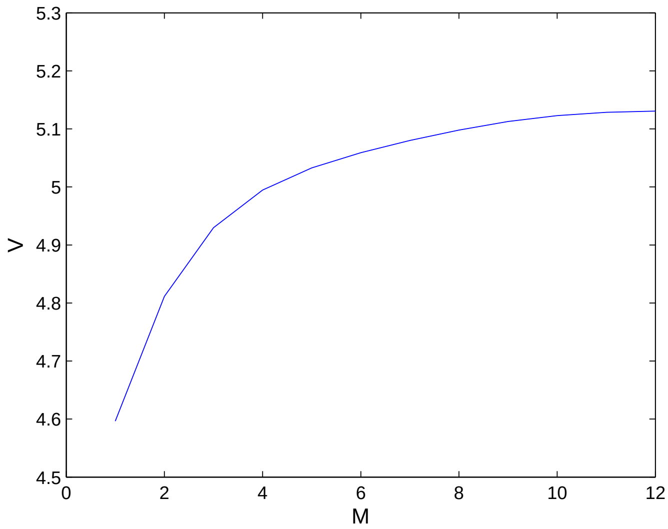

Figure 3.7: American call values where the underlying asset pays one dividend

3.2.3 American Options on Continuous Dividend Paying Assets

Consider a stock paying a continuous dividend yield D. Since the dividends provide a return of D, the capital gain return has to be $r - D$ which means that an arbitrage opportunity exists. In this case, if the initial value of the stock price for a European option is S_0, its expected value after on time step of length δ_t must be $S_0\, e^{(r-D)\delta_t}$, and the system which we solved for the parameters u and d changes.

It follows that, formula (3.7) changes to

$$V_{ji} = e^{-r\delta_t}\{qV_{j+1,i+1} + (1-q)V_{j,i+1}\},$$

where

$$q = \frac{e^{(r-D)\delta_t} - d}{u - d}.$$

The payoffs (3.10)-(3.11) would be the same and the value of the option would be given as

$$V_{ji}^{A} = \max\{e^{-r\delta_t}[qV_{j+1,i+1} + (1-q)V_{j,i+1}], \Lambda(S_{ji})\}$$

for each $i = M-1, M-2, \ldots, 0$ and $j = 0, 1, \ldots, i$, where

$$q = \frac{e^{(r-D)\delta_t} - d}{u - d}.$$

Example 4. Take the same parameters as in the example (2) differently considering a continuous dividend payment with $D = 0.2$. Then, the value of the American call option will be 6.6048.

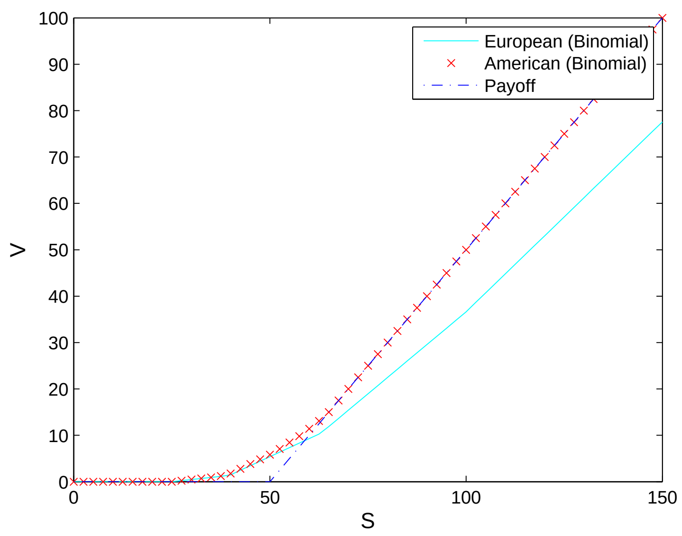

Figure 3.8: American call values where the underlying asset pays continuous dividends

As we have mentioned, the binomial method is a very popular technique to price the options contains constructing a binomial tree. However, the model works in a relatively slow speed requiring more calculations at a time, especially when large number of steps is considered. It may not be the most practical way to calculate the price of the options, especially when barrier options are under investigation; therefore, we might resort to another approach to price the options: next chapter is devoted to the finite difference method for solving the associated partial differential equation.

CHAPTER 4

AMERICAN OPTION PRICING BY FINITE DIFFERENCE METHOD

In this section, we deal with finite difference method for pricing American options. First, we introduce the Black-Scholes equation with boundary and final conditions for European options, give the closed-form solution of the partial differential equation and represent the finite difference method to solve the equation. Then, we follow to the pricing American options by partial differential equations and writing the problem as a free boundary value problem, and finally we will give the projected SOR algorithm in oder to price American type options.

4.1 The Black-Scholes Partial Differential Equation

The Black-Scholes Model is a model for option pricing which was first developed by Fischer Black and Myron Scholes [9] and then further developed by Robert Merton. The model is one of the basic buildings for the derivatives theory. Before representing the model we need the following assumptions for the derivation:

- There are no dividends on the underlying asset.
- There are no transaction costs.
- Trading takes place continuously.
- One can borrow and lend cash at a constant risk-free interest rate.
- One can buy any fraction of a share of stock.

- There are no restrictions on short selling.

Further we assume that the stock price follows a geometric Brownian motion (6) with constant expected return μ (drift) and volatility σ which is given by

$$dS = \mu S dt + \sigma S dW \tag{4.1}$$

where W is a Brownian motion.

Let $V = V(S,t)$ denote the value of an option, its second-order derivatives with respect to t be continuous in the domain

$$\mathcal{D}_V = \{(S,t) : S \geq 0, \quad 0 \leq t \leq T\}.$$

Then, from the Itô Lemma (2.2.2), we get

$$dV = \left(\frac{\partial V}{\partial S} \mu S + \frac{\partial V}{\partial t} + \frac{1}{2} \frac{\partial^2 V}{\partial S^2} \sigma^2 S^2 \right) dt + \frac{\partial V}{\partial S} \sigma S \, dW. \tag{4.2}$$

Now, let us set up a portfolio consisting of a long position in one option and a short position in Δ shares of the underlying asset with the price S and denote this portfolio by Π. Suppose that the initial value of the portfolio is Π_0, and that the value of the portfolio at time t can be determined from

$$\Pi = V - \Delta S.$$

In order to find the number of shares Δ which makes this portfolio riskless, we write the change in the value of this portfolio in the time interval dt as

$$d\Pi = dV - \Delta dS, \tag{4.3}$$

where $dS = \mu S dt + \sigma S dW$. Then, substituting the equation (4.2) into (4.3), we obtain

$$\begin{aligned} d\Pi &= dV - \Delta dS \\ &= \left(\frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \mu S \frac{\partial V}{\partial S} - \Delta \mu S \right) dt + \left(\sigma S \frac{\partial V}{\partial S} - \Delta \sigma S \right) dW. \end{aligned} \tag{4.4}$$

Observe that the randomness is eliminated from this portfolio by choosing

$$\Delta = \frac{\partial V}{\partial S}.$$

Thence, we have

$$d\Pi = \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right) dt, \tag{4.5}$$

which is completely riskless from uncertainty.

The return on an amount Π invested in a riskless asset would grow to $r\Pi dt$ within a time interval dt. If $d\Pi > r\Pi dt$, then an arbitrager could make a guaranteed riskless profit by borrowing an amount Π to invest in the portfolio. Conversely, if $d\Pi < r\Pi dt$, then the arbitrager would short the portfolio and invest Π in the bank. Consequently, we must have

$$d\Pi = r\Pi dt, \tag{4.6}$$

which is equivalent to

$$d\Pi = r(V - \Delta S)dt = \left(rV - rS\frac{\partial V}{\partial S} \right) dt. \tag{4.7}$$

Now, equating the equations (4.5) and (4.7), we obtain

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0, \tag{4.8}$$

the so-called *Black-Scholes partial differential equation.*

It is of great importance to note that drift parameter μ of underlying assets in the Black-Scholes equation has disappeared, so the price of the options will be independent of how rapidly or slowly on asset price changes.

On the other hand, the Black-Scholes equation needs a final condition and boundary conditions to find the unique solution of the equation. Now, we consider a European call option with value denoted by $C(S,t)$ at time t, with strike price K and expiry date T. The final condition at time $t = T$ is the value of the call option known with certainty to be the payoff

$$C(S,T) = \max\{S - K, 0\}. \tag{4.9}$$

From (4.1), we can see that when $S = 0$, the payoff is zero at expiry date. Therefore, we have

$$C(0,t) = 0. \tag{4.10}$$

When $S \to \infty$, the option will be exercised and the magnitude of the exercise price becomes less important. Hence, the option value becomes asymptotically the asset value and so, we get

$$C(S,t) \sim S - Ke^{-r(T-t)} \quad \text{as} \quad S \to \infty. \tag{4.11}$$

As for a European call option, there is no possibility of early exercise, (4.8)–(4.11) can be solved to give the Black-Scholes solution for the option.

For a put option, meanwhile, the final condition with value $P(S,t)$ is the payoff

$$P(S,T) = \max\{K - S, 0\}. \tag{4.12}$$

In this case, when $S = 0$, the final payoff of a European put option is certainly the strike price K. The present value of K at time T is the discounted one, $P(0,t)$, which is,

$$P(0,t) = Ke^{-r(T-t)} \tag{4.13}$$

where the interest rate r is assumed to be constant.

On the other hand, when $S \to \infty$, a put option is unlikely to be exercised and so that

$$P(S,t) \to 0 \quad \text{as} \quad S \to \infty. \tag{4.14}$$

To sum up, closed-form solutions of Black-Scholes equation with the corresponding boundary and final conditions for European call and put options are given in the following theorem.

Theorem 4.1.1. *The Black-Scholes formula for a European call option is*

$$C(S,t) = SN(d_1) - Ke^{-r(T-t)}N(d_2),$$

for a European put option is

$$P(S,t) = Ke^{-r(T-t)}N(-d_2) - SN(d_1),$$

where

$$d_1 = \frac{\ln\left(\frac{S}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}},$$

$$d_2 = d_1 - \sigma\sqrt{T-t},$$

and $N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt$ *is the cumulative distribution function for the standard normal distribution.*

Proof. See [42]. □

In the case of dividend paying underlying asset with yield rate q, Black-Scholes equation (4.8) can be modified as

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + (r-q)S\frac{\partial V}{\partial S} - rV = 0.$$

Closed-form solutions of modified Black-Scholes partial differential equation with the same boundary and final conditions (4.9), (4.10), and (4.11) for a call option is given by

$$C(S,t) = Se^{-q(T-t)}N(d_1) - Ke^{-r(T-t)}N(d_2),$$

and, for a put option,

$$P(S,t) = Ke^{-r(T-t)}N(-d_2) - Se^{-q(T-t)}N(d_1),$$

with the corresponding boundary and final conditions (4.12), (4.13), and (4.14).

4.2 Finite Difference Method

This part of the thesis represents the finite difference method which is based on the natural idea of approximating the partial differential equations over the area of integration by a set of algebraic equations. The main idea of this method is to replace differentials by difference quotients. The most common finite difference methods to solve the Black-Scholes equation are the explicit, implicit and the Crank-Nicolson methods.

The idea underlying finite-difference methods is to replace the partial derivatives in partial differential equations by approximations based on Taylor series expansions of derivatives of (necessarily smooth) functions. The Taylor's theorem states that a function $f(x)$ may be expressed as

$$f(x+h) = f(x) + hf'(x) + \frac{1}{2}h^2 f''(x) + \frac{1}{6}h^3 f'''(x) + \mathcal{O}(h^4).$$

If the terms of order h^2 and higher are neglected, then we get

$$f'(x) = \frac{f(x+h) - f(x)}{h} + \mathcal{O}(h).$$

From here, $f'(x)$ would be written as

$$f'(x) \approx \frac{1}{h}\left[f(x+h) - f(x)\right].$$

This particular finite difference approximation is called a forward difference since the differencing is in the forward ($h > 0$) direction.

The other way to approximate first-order derivatives could be as

$$f(x-h) = f(x) - hf'(x) + \frac{1}{2}h^2 f''(x) - \frac{1}{6}h^3 f'''(x) + \mathcal{O}(h^4).$$

From this formula, $f'(x)$ can also be written as

$$f'(x) = \frac{f(x) - f(x-h)}{h} + \mathcal{O}(h),$$

and so that we have

$$f'(x) \approx \frac{1}{h}\left[f(x) - f(x-h)\right]$$

which is called the backward difference approximation.

Furthermore, $f'(x)$ can be written by subtracting the Taylor series for $f(x+h)$ and $f(x-h)$ as the following:

$$f'(x) = \frac{1}{2h}\left[f(x+h) - f(x-h)\right] + \mathcal{O}(h^2).$$

Then, the formula

$$f'(x) \approx \frac{1}{2h}\left[f(x+h) - f(x-h)\right]$$

is called the central difference formula for $f'(x)$. As we have seen that there may be many other finite difference formulas to approximate $f'(x)$.

Now, let us see the derivation of the finite difference formulas for higher-order derivatives. Adding the Taylor's expansions of $f(x+h)$ and $f(x-h)$ as

$$f(x+h) + f(x-h) = 2f(x) + h^2 f''(x) + \mathcal{O}(h^4)$$

and making some rearrangements we obtain

$$f''(x) = \frac{1}{h^2}\left[f(x+h) - 2f(x) + f(x-h)\right] + \mathcal{O}(h^2),$$

and then

$$f''(x) \approx \frac{1}{h^2}\left[f(x+h) - 2f(x) + f(x-h)\right]$$

which is called the central difference approximation for the second-derivative f'' at x.

Likewise, if $u = u(x, y)$ is a function of two variables, then the approximation for the partial derivative $u_x(x, y)$ with respect to x is

$$u_x(x, y) = \frac{1}{h}\left[u(x+h, y) - u(x, y)\right] + \mathcal{O}(h),$$

and it is called a forward difference formula for $u = u(x, y)$, where $h = \Delta x$.

Similarly, a central difference formula for the second derivative of u with respect to x can be written as

$$u_{xx}(x, y) = \frac{1}{h^2}\left[u(x+h, y) - 2u(x, y) + u(x-h, y)\right] + \mathcal{O}(h^2).$$

Also, an approximation to u_{xy} where $u = u(x, y)$ derived from the Taylor's series can be given by

$$u_{xy}(x, y) \approx \frac{u(x+h, y+k) - u(x, y+k) - u(x+h, y) + u(x, y)}{hk}$$

with the order $\mathcal{O}(h) + \mathcal{O}(k)$, where $h = \Delta x$ and $k = \Delta y$.

4.3 Pricing European Options by Finite Difference Approximations

As we have stated, the value V of a European option at time t on an underlying asset with price S satisfies the Black-Scholes equation

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + (r - q)S\frac{\partial V}{\partial S} - rV = 0. \tag{4.15}$$

where the final condition is given by the payoff

$$C(S, T) = \max\{S - K, 0\}$$

for a European call option and

$$P(S,T) = \max\{K - S, 0\}$$

for a European put option with a strike price K. Also, the boundary conditions can be given by [46]

$$C(S_{min}, t) = 0, \quad C(S_{max}, t) = S_{max} - Ke^{-r(T-t)},$$

for a European call option and

$$P(S_{max}, t) = 0, \quad P(S_{min}, t) = Ke^{-r(T-t)} - S_{min},$$

for a European put option. Here S_{min} and S_{max} represents respectively small and large values of stock prices for the vanilla options.

To work with finite difference quotients, we now discretize the domain of S and t,

$$S_{min} \leq S \leq S_{max} \quad \text{and} \quad t_0 \leq t \leq T,$$

as follows: we divide the intervals into M and N parts such that

$$\Delta t = \frac{T - t_0}{M} \quad \text{and} \quad \Delta S = \frac{S_{max} - S_{min}}{N}.$$

Here,

$$t_i = t_0 + i\Delta t \quad \text{and} \quad S_j = S_{min} + j\Delta S,$$

where $i = 0, 1, \ldots, M$ and $j = 0, 1, \ldots, N$. Now, we denote the approximation of the value $V_{ji} = V(S_j, t_i) \approx w_{ji}$, for all grid points (S_j, t_i).

4.3.1 Explicit Method

In this part, we derive the explicit method for the Black-Scholes partial differential equation which was first developed by Brennan and Schwartz [12, 40] and then improved by Courtadon [16]. This method uses the backward difference approximation for the partial derivative $\frac{\partial V}{\partial t}$, and the central difference approximation for the partial derivatives $\frac{\partial V}{\partial S}$ and $\frac{\partial^2 V}{\partial S^2}$ such that

$$\frac{\partial V}{\partial t} \approx \frac{w_{ji} - w_{j,i-1}}{\Delta t}, \tag{4.16}$$

$$\frac{\partial V}{\partial S} \approx \frac{w_{j+1,i} - w_{j-1,i}}{2\Delta S}, \tag{4.17}$$

$$\frac{\partial^2 V}{\partial S^2} \approx \frac{w_{j+1,i} - 2w_{ji} + w_{j-1,i}}{(\Delta S)^2}. \tag{4.18}$$

Substituting the equations (4.16), (4.17) and (4.18) into the Black-Scholes PDE, we obtain the explicit method which is accurate to $\mathcal{O}(\Delta t)+\mathcal{O}((\Delta S)^2)$ as follows:

$$\frac{w_{ji} - w_{j,i-1}}{\Delta t} = rw_{ji} - (r-q)S_j\frac{w_{j+1,i} - w_{j-1,i}}{2\Delta S} - \frac{1}{2}\sigma^2 S_j^2 \frac{w_{j+1,i} - 2w_{ji} + w_{j-1,i}}{(\Delta S)^2}.$$

Here, the method can be written as

$$w_{j,i-1} = \alpha_j w_{j-1,i} + \beta_j w_{ji} + \gamma_j w_{j+1,i}, \tag{4.19}$$

where

$$\alpha_j = \frac{1}{2}\Delta t\left\{\sigma^2(\frac{S_j}{\Delta S})^2 - (r-q)\frac{S_j}{\Delta S}\right\},$$

$$\beta_j = 1 - \Delta t\left\{\sigma^2(\frac{S_j}{\Delta S})^2 + r\right\},$$

$$\gamma_j = \frac{1}{2}\Delta t\left\{(r-q)\frac{S_j}{\Delta S} + \sigma^2(\frac{S_j}{\Delta S})^2\right\}$$

for every $i = 0, 1, \ldots, M$ and $j = 0, 1, \ldots, N$.

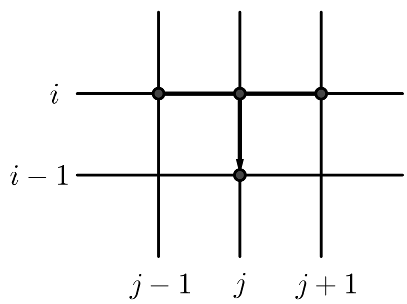

Figure 4.1: Molecules of the Explicit method

The values $w_{1,i}, w_{2,i}, \dots, w_{N-1,i}$ can be collected in the vectors as

$$w^{(i)} = \begin{pmatrix} w_{1,i} \\ \vdots \\ w_{N-1,i} \end{pmatrix}.$$

Then, the system can be written in the following matrix-vector form:

$$w^{(i-1)} = Aw^{(i)} + y^{(i)}, \quad i = M, M-1, \dots, 1,$$

where

$$A = \begin{pmatrix} \beta_1 & \gamma_1 & 0 & \dots & 0 & 0 \\ \alpha_2 & \beta_2 & \gamma_2 & \dots & 0 & 0 \\ 0 & \alpha_3 & \beta_3 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \beta_{N-2} & \gamma_{N-2} \\ 0 & 0 & 0 & \dots & \alpha_{N-1} & \beta_{N-1} \end{pmatrix}, \quad y^{(i)} = \begin{pmatrix} \alpha_1 w_{0,i} \\ 0 \\ \vdots \\ 0 \\ \gamma_{N-1} w_{N,i} \end{pmatrix}.$$

Since there is no need to take the inverse of any matrix and compute some boundaries [25] for the solution, the explicit method for the Black-Scholes PDE is easy to implement.

Example 5. Consider a European call option with the stock price \$60 on a non-dividend paying asset where the time to maturity is $T = 4/12$. The strike price is \$60 and the risk free interest rate is 10%. The volatility is 40%.

We found the value of this call option using explicit finite difference method as \$6.4526 and as \$6.4649 by using the exact solution for the Black-Scholes equation for European call option.

These figures, Figure 4.2 and Figure 4.3, show that explicit method gives a satisfactory result compared to the exact solution of the Black-Scholes equation.

Due to the stability condition that will be given later we can not reduce ΔS for given Δt as we wish.

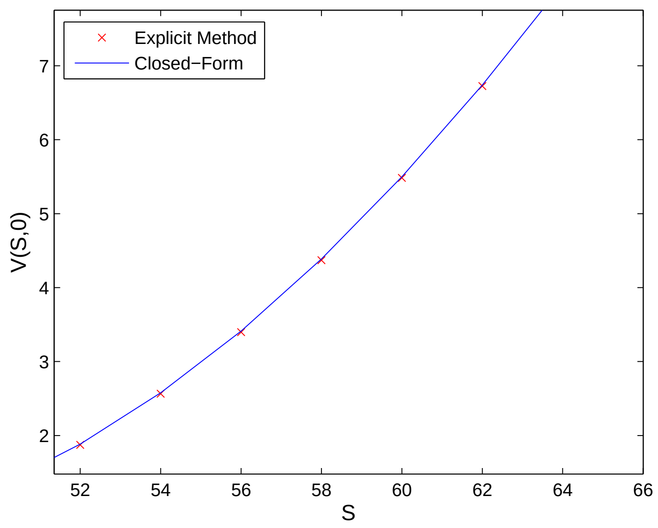

Figure 4.2: Exact and approximate solutions for a European call option by explicit method

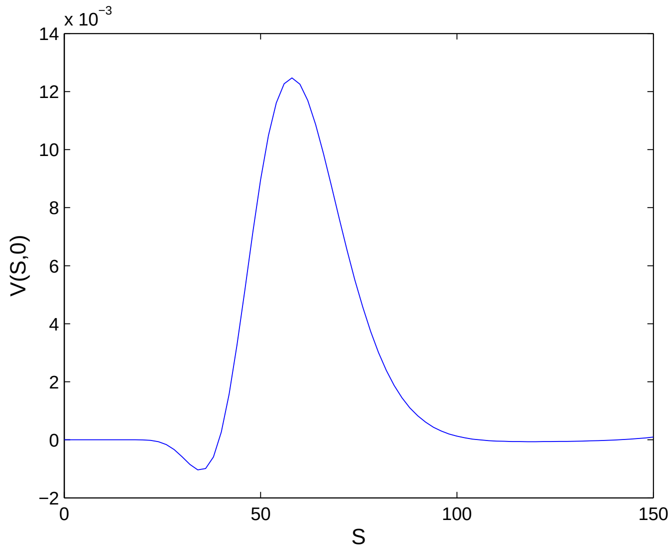

Figure 4.3: The difference between explicit method and closed form solutions for above European call option

4.3.2 Implicit Method

The forward difference approximation is used in this case, contrary to the explicit method. This scheme was first introduced by Schwartz [40], and Brennan and Schwartz [12]. It was then developed by Courtadon [16].

The partial derivatives $\frac{\partial V}{\partial t}$, $\frac{\partial V}{\partial S}$ and $\frac{\partial^2 V}{\partial S^2}$ are approximated as the following:

$$\frac{\partial V}{\partial t} \approx \frac{w_{j,i+1} - w_{ji}}{\Delta t}, \tag{4.20}$$
$$\frac{\partial V}{\partial S} \approx \frac{w_{j+1,i} - w_{j-1,i}}{2\Delta S}, \tag{4.21}$$
$$\frac{\partial^2 V}{\partial S^2} \approx \frac{w_{j+1,i} - 2w_{ji} + w_{j-1,i}}{(\Delta S)^2}. \tag{4.22}$$

Now, again substituting the equations (4.20), (4.21) and (4.22) into the Black-Scholes PDE, we get

$$\frac{w_{j,i+1} - w_{ji}}{\Delta t} = rw_{ji} - (r-q)S_j\frac{w_{j+1,i} - w_{j-1,i}}{2\Delta S} - \frac{1}{2}\sigma^2 S_j^2 \frac{w_{j+1,i} - 2w_{ji} + w_{j-1,i}}{(\Delta S)^2}.$$

Making some rearrangements, we have

$$\alpha_j w_{j-1,i} + \beta_j w_{ji} + \gamma_j w_{j+1,i} = w_{j,i+1}, \tag{4.23}$$

where

$$\alpha_j = \frac{1}{2}\Delta t \left\{ (r-q)\frac{S_j}{\Delta S} - \sigma^2 (\frac{S_j}{\Delta S})^2 \right\},$$
$$\beta_j = 1 + \Delta t \left\{ \sigma^2 (\frac{S_j}{\Delta S})^2 + r \right\},$$
$$\gamma_j = -\frac{1}{2}\Delta t \left\{ (r-q)\frac{S_j}{\Delta S} + \sigma^2 (\frac{S_j}{\Delta S})^2 \right\}$$

for all $i = 0, 1, \ldots, M$ and $j = 0, 1, \ldots, N$ which is called the implicit method which is accurate to $\mathcal{O}(\Delta t) + \mathcal{O}((\Delta S)^2)$.

Consequently, the system will be

$$Aw^{(i)} = w^{(i+1)} + y^{(i+1)}, \quad i = M-1, M-2, \ldots, 1,$$

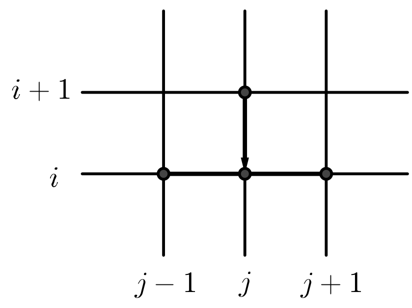

Figure 4.4: Molecules of the Implicit method

where

$$A = \begin{pmatrix} \beta_1 & \gamma_1 & 0 & \dots & 0 & 0 \\ \alpha_2 & \beta_2 & \gamma_2 & \dots & 0 & 0 \\ 0 & \alpha_3 & \beta_3 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \beta_{N-2} & \gamma_{N-2} \\ 0 & 0 & 0 & \dots & \alpha_{N-1} & \beta_{N-1} \end{pmatrix}, \quad y^{(i+1)} = \begin{pmatrix} -\alpha_1 w_{0,i} \\ 0 \\ \vdots \\ 0 \\ -\gamma_{N-1} w_{N,i} \end{pmatrix}.$$

LU-decomposition

To solve the system (4.23), in principle, we need to find the inverse of the tridiagonal matrix A, however, this is not practical computationally. Thus, we look for another way to solve this system: the easiest one is the so-called LU-decomposition [46].

An LU-decomposition of the tridiagonal matrix A has the form

$$A = LU,$$

where L is a lower and U is an upper triangular matrices of the same size as of A. We first introduce the temporary value

$$v^{(i)} := Uw^{(i)},$$

then the system

$$Lv^{(i)} = w^{(i+1)} + y^{(i+1)},$$

should be solved by backward substitution. We denote this substitution as

$$v^{(i)} = L^{-1}\left\{w^{(i+1)} + y^{(i+1)}\right\},$$

but the inverse L^{-1} will never be computed numerically.

Then, the solution $w^{(i)}$ is obtained by backward substitution represented by

$$w^{(i)} = U^{-1}v^{(i)}.$$

Decomposing these substitutions, the solution to the system of linear equations can be written as

$$w^{(i)} = U^{-1}\left\{L^{-1}(w^{(i+1)} + y^{(i+1)})\right\}$$

in terms of U and L.

LU-decomposition is a direct method to solve the system (4.23). Another way to handle the same system is by using the so-called successive over relaxation (SOR) method which is an iterative method. We will discuss this in Section 4.3.9.

4.3.3 Crank-Nicolson Method

The Crank-Nicolson method is obtained by combining the explicit and implicit methods which are given by the equations (4.19) and (4.23). The scheme was introduced by Crank and Nicolson [17].

Taking the arithmetic average of these equations, we derive the Crank-Nicolson method which is now accurate to $\mathcal{O}((\Delta t)^2)+\mathcal{O}((\Delta S)^2)$, better than the previous methods. For the Black-Scholes PDE, we have

$$\begin{aligned}-\alpha_j w_{j-1,i-1} + (1-\beta_j)w_{j,i-1} - \gamma_j w_{j+1,i-1}& \\ = \alpha_j w_{j-1,i} + (1+\beta_j)w_{ji} + \gamma_j w_{j+1,i},& \quad (4.24)\end{aligned}$$

where

$$\alpha_j = \frac{1}{4}\Delta t \left\{ \sigma^2 (\frac{S_j}{\Delta S})^2 - (r-q)\frac{S_j}{\Delta S} \right\},$$
$$\beta_j = -\frac{1}{2}\Delta t \left\{ \sigma^2 (\frac{S_j}{\Delta S})^2 + r \right\},$$
$$\gamma_j = \frac{1}{4}\Delta t \left\{ \sigma^2 (\frac{S_j}{\Delta S})^2 + (r-q)\frac{S_j}{\Delta S} \right\}$$

for every $i = M-1, M-2, \ldots, 1$ and $j = 1, 2, \ldots, N-1$.

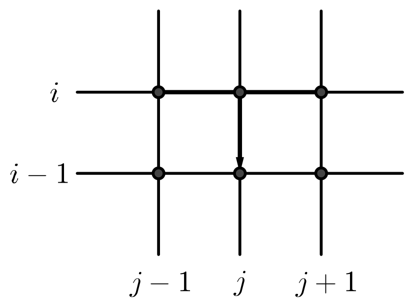

Figure 4.5: Molecules of the Crank-Nicolson method

Then, the method can be written in the matrix-vector notation as

$$Aw^{(i-1)} = Bw^{(i)} + y^{(i)}, \quad i = M-1, M-2, \ldots, 1,$$

where

$$w^{(i)} = \begin{pmatrix} w_{1,i} \\ \vdots \\ w_{N-1,i} \end{pmatrix}, \quad y^{(i)} = \begin{pmatrix} \alpha_1(w_{0,i-1} + w_{0,i}) \\ 0 \\ \vdots \\ 0 \\ \gamma_{N-1}(w_{N,i-1} + w_{N,i}) \end{pmatrix},$$

$$A = \begin{pmatrix} 1-\beta_1 & -\gamma_1 & 0 & \dots & 0 & 0 \\ -\alpha_2 & 1-\beta_2 & -\gamma_2 & \dots & 0 & 0 \\ 0 & -\alpha_3 & 1-\beta_3 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1-\beta_{N-2} & -\gamma_{N-2} \\ 0 & 0 & 0 & \dots & -\alpha_{N-1} & 1-\beta_{N-1} \end{pmatrix},$$

$$B = \begin{pmatrix} 1+\beta_1 & \gamma_1 & 0 & \dots & 0 & 0 \\ \alpha_2 & 1+\beta_2 & \gamma_2 & \dots & 0 & 0 \\ 0 & \alpha_3 & 1+\beta_3 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1+\beta_{N-2} & \gamma_{N-2} \\ 0 & 0 & 0 & \dots & \alpha_{N-1} & 1+\beta_{N-1} \end{pmatrix}.$$

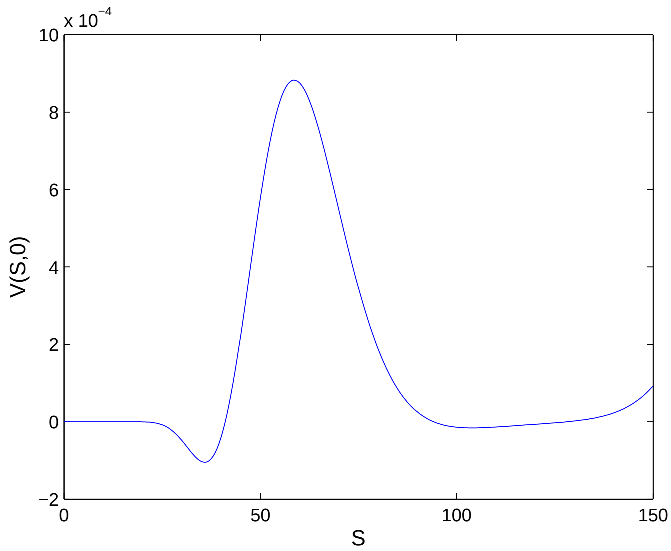

Figure 4.6: The difference between Crank-Nicolson and exact solutions with parameters $S = 50, K = 50, D = 0, \sigma = 0.4, r = 0.1, T = 4/12, S_{min} = 0, S_{max} = 150, dS = 0.5, dt = 1/1200$

Figure 4.6 depicts the error between the Crank-Nicolson approximation and exact solution. It shows that Crank-Nicolson solution oscillates around the exact solution when S is close to the strike price. This situation is due to the non-smoothness of the final condition at the strike price $S = K$. To overcome with this difficulty, in practice, one can use implicit method for the first few time steps and then continue with the Crank-Nicolson method to benefit from the accuracy of the latter.

4.3.4 θ-Averaged Method

The θ-Averaged method is a convex combination of the explicit and the implicit methods with the weights θ and $1-\theta$, [37, 46, 32]. The method can be recarded as a generalized version of the Crank-Nicolson method.

For $0 \leq \theta \leq 1$, the method is defined as

$$\begin{aligned} \frac{w_{ji} - w_{j,i-1}}{\Delta t} &= \theta \left(r w_{ji} - (r-q) S_j \frac{w_{j+1,i} - w_{j-1,i}}{2\Delta S} \right) \\ &- \frac{1}{2}\theta \left(\sigma^2 S_j^2 \frac{w_{j+1,i} - 2w_{ji} + w_{j-1,i}}{(\Delta S)^2} \right) \\ &+ (1-\theta) \left(r w_{j,i-1} - (r-q) S_j \frac{w_{j+1,i-1} - w_{j-1,i-1}}{2\Delta S} \right) \\ &- \frac{1}{2}(1-\theta) \left(\sigma^2 S_j^2 \frac{w_{j+1,i-1} - 2w_{j,i-1} + w_{j-1,i-1}}{(\Delta S)^2} \right). \end{aligned} \tag{4.25}$$

From this formula, we obtain

- the explicit method by taking $\theta = 1$,
- the implicit method by taking $\theta = 0$,
- the Crank-Nicolson method by taking $\theta = \frac{1}{2}$.

Then, rearranging the equation (4.25), we have

$$\begin{aligned} &\alpha_j w_{j-1,i-1} + \beta_j w_{j,i-1} + \gamma_j w_{j+1,i-1} \\ &\qquad = \hat{\alpha}_j w_{j-1,i} + \hat{\beta}_j w_{ji} + \hat{\gamma}_j w_{j+1,i}, \end{aligned} \tag{4.26}$$

where

$$
\begin{aligned}
\alpha_j &= -\frac{1}{2}(1-\theta)\Delta t\left\{\sigma^2(\frac{S_j}{\Delta S})^2 - (r-q)\frac{S_j}{\Delta S}\right\}, \\
\beta_j &= 1 + (1-\theta)\Delta t\left\{\sigma^2(\frac{S_j}{\Delta S})^2 + r\right\}, \\
\gamma_j &= -\frac{1}{2}(1-\theta)\Delta t\left\{\sigma^2(\frac{S_j}{\Delta S})^2 + (r-q)\frac{S_j}{\Delta S}\right\}, \\
\hat{\alpha}_j &= \frac{1}{2}\theta\Delta t\left\{\sigma^2(\frac{S_j}{\Delta S})^2 - (r-q)\frac{S_j}{\Delta S}\right\}, \\
\hat{\beta}_j &= 1 - \theta\Delta t\left\{\sigma^2(\frac{S_j}{\Delta S})^2 + r\right\}, \\
\hat{\gamma}_j &= \frac{1}{2}\theta\Delta t\left\{\sigma^2(\frac{S_j}{\Delta S})^2 + (r-q)\frac{S_j}{\Delta S}\right\}
\end{aligned}
\tag{4.27}
$$

for all $i = M-1, M-2, \ldots, 1$ and $j = 1, \ldots, N-1$. Here, the matrix-vector notation of the method is

$$
Aw^{(i-1)} = Bw^{(i)} + y^{(i)}, \quad i = M-1, M-2, \ldots, 1, \tag{4.28}
$$

where

$$
w^{(i-1)} = \begin{pmatrix} w_{1,i-1} \\ \vdots \\ w_{N-1,i-1} \end{pmatrix}, \quad y^{(i)} = \begin{pmatrix} \alpha_1 w_{0,i-1} \\ 0 \\ \vdots \\ 0 \\ \gamma_{N-1} w_{N,i-1} \end{pmatrix},
$$

$$
A = \begin{pmatrix}
\beta_1 & \gamma_1 & 0 & \ldots & 0 & 0 \\
\alpha_2 & \beta_2 & \gamma_2 & \ldots & 0 & 0 \\
0 & \alpha_3 & \beta_3 & \ldots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \ldots & \beta_{N-2} & \gamma_{N-2} \\
0 & 0 & 0 & \ldots & \alpha_{N-1} & \beta_{N-1}
\end{pmatrix},
$$

$$
B = \begin{pmatrix}
\hat{\beta}_1 & \hat{\gamma}_1 & 0 & \ldots & 0 & 0 \\
\hat{\alpha}_2 & \hat{\beta}_2 & \hat{\gamma}_2 & \ldots & 0 & 0 \\
0 & \hat{\alpha}_3 & \hat{\beta}_3 & \ldots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \ldots & \hat{\beta}_{N-2} & \hat{\gamma}_{N-2} \\
0 & 0 & 0 & \ldots & \hat{\alpha}_{N-1} & \hat{\beta}_{N-1}
\end{pmatrix}.
$$

We illustrate, in Figure 4.7, that the θ-Averaged method for $\theta = 1$ and the explicit method for a European put option with the parameters $S_0 = 60, K = 60, D = 0, \sigma = 0.2, r = 0.3, T = 3/12, S_{min} = 0, S_{max} = 150, dS = 2, dt = 1/1000$. To solve the system, LU method is used.

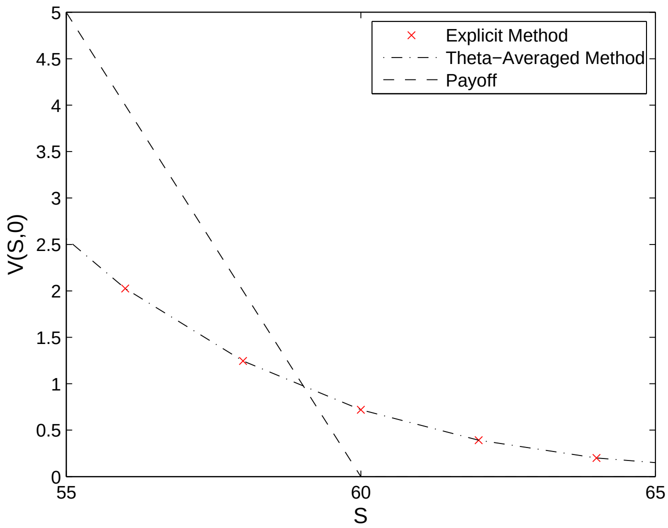

Figure 4.7: The θ-Averaged and Explicit Methods

One can see that from the figure, the values coming from the θ-averaged and explicit methods are the same as we expected.

4.3.5 Consistency, stability and convergence of methods

Truncation errors are the measures of the error by which the exact solution of a differential equation does not satisfy the numerical scheme at the grid points.

Let $u(x, t)$ be any solution of a partial differential equation and u_m^n be any solutions of a finite difference scheme. The scheme is said to be convergent if u_m^0 approaches to $u_0(x)$ as $mh \to x$, then u_m^n approaches to $u(x, t)$ as (mh, nk) approaches to (x, t) as $h, k \to 0$. Therefore, a necessary condition for the convergence of the numerical solutions to the continuous solution is that the truncation error tends to zero as the mesh size goes to zero. In this case the scheme is said to be *consistent*, that is, a finite difference numerical scheme $P_{h,k}u = f$ is consistent with a partial differential equation $Pu = f$ if for any smooth function

$\varphi(x,t)$,

$$P\varphi - P_{h,k}\varphi \to 0 \quad \text{as} \quad h,k \to 0.$$

The explicit and implicit numerical schemes have truncation errors $\mathcal{O}(\Delta t) + \mathcal{O}((\Delta S)^2)$, and the Crank-Nicolson scheme has $\mathcal{O}((\Delta t)^2) + \mathcal{O}((\Delta S)^2)$ truncation error. Thus all these methods are consistent.

As we have mentioned, consistency is only a necessary condition for the convergence of the numerical methods. While performing the calculations roundoff errors may become arbitrarily large which causes error in the whole computation. If the roundoff errors are not grown in the calculations, that is if they become bounded, then the method is called *stable*. If the roundoff error is unbounded, then the method is said to be *unstable*. In other words, a finite difference scheme $P_{h,k}u_m^n = 0$ is stable if for some positive numbers h_0, k_0, there is a constant C such that

$$\|u_m^n\| \leq C \left\|u_m^0\right\|_h,$$

for $h \leq h_0$, $k \leq k_0$.

For the stability analysis of the finite differences methods which we have investigated above, we will discuss the Fourier stability analysis which was developed by von Neumann. The Fourier method is based on the assumption that the numerical scheme (4.26) takes a solution of the form

$$w_{ji} = \lambda^i e^{-zNj\Delta S}, \tag{4.29}$$

where $z^2 = -1$, $\mathrm{Im}(z) = 1$ and N is an arbitrary constant.

Substituting (4.29) into (4.26), we obtain

$$\begin{aligned}\alpha_j\lambda^{i-1}e^{-zN(j-1)\Delta S} + \beta_j\lambda^{i-1}e^{-zNj\Delta S} + \gamma_j\lambda^{i-1}e^{-zN(j+1)\Delta S}\\ = \hat{\alpha}_j\lambda^i e^{-zN(j-1)\Delta S} + \hat{\beta}_j\lambda^i e^{-zNj\Delta S} + \hat{\gamma}_j\lambda^i e^{-zN(j+1)\Delta S}.\end{aligned} \tag{4.30}$$

Then, rearranging the equation (4.30) and removing the term $\lambda^i e^{-zNj\Delta S}$, we have

$$\frac{1}{\lambda}\left[\alpha_j e^{zN\Delta S} + \beta_j + \gamma_j e^{-zN\Delta S}\right] = \hat{\alpha}_j e^{zN\Delta S} + \hat{\beta}_j + \hat{\gamma}_j e^{-zN\Delta S}.$$

Then,

$$\lambda = \left| \frac{\alpha_j e^{zN\Delta S} + \beta_j + \gamma_j e^{-zN\Delta S}}{\hat{\alpha}_j e^{zN\Delta S} + \hat{\beta}_j + \hat{\gamma}_j e^{-zN\Delta S}} \right|.$$

The necessary and sufficient condition for the scheme to be stable is $|\lambda| \leq 1$ because of the discrete von Neumann criterion for stability, see [43]. Thus, we obtain the stability condition for the scheme (4.26) such as in [20]

$$\left| \frac{\alpha_j e^{zN\Delta S} + \beta_j + \gamma_j e^{-zN\Delta S}}{\hat{\alpha}_j e^{zN\Delta S} + \hat{\beta}_j + \hat{\gamma}_j e^{-zN\Delta S}} \right| \leq 1.$$

Now, we will investigate the stability conditions for each explicit, implicit and Crank-Nicolson schemes giving the θ values in (4.26).

For $\theta = 1$; we get

$$\begin{aligned}
\alpha_j &= 0, \\
\beta_j &= 1, \\
\gamma_j &= 0, \\
\hat{\alpha}_j &= \frac{1}{2}\Delta t \left\{ \sigma^2 \left(\frac{S_j}{\Delta S} \right)^2 - (r-q)\frac{S_j}{\Delta S} \right\}, \\
\hat{\beta}_j &= 1 - \Delta t \left\{ \sigma^2 \left(\frac{S_j}{\Delta S} \right)^2 + r \right\}, \\
\hat{\gamma}_j &= \frac{1}{2}\Delta t \left\{ \sigma^2 \left(\frac{S_j}{\Delta S} \right)^2 + (r-q)\frac{S_j}{\Delta S} \right\},
\end{aligned}$$

and then the stability condition for the explicit method turns out to be

$$\begin{aligned}
&\Bigg| \Delta t \sigma^2 \left(\frac{S_j}{\Delta S} \right)^2 \cos(N\Delta S) - \Delta t(r-q) \left(\frac{S_j}{\Delta S} \right) \sin(N\Delta S) \\
&\qquad\qquad + 1 - \Delta t \left(\sigma^2 (\frac{S_j}{\Delta S})^2 + r \right) \Bigg| \geq 1.
\end{aligned}$$

For $\theta = 0$; we get

$$\alpha_j = \frac{1}{2}\Delta t \left\{ \sigma^2 \left(\frac{S_j}{\Delta S}\right)^2 - (r-q)\frac{S_j}{\Delta S} \right\},$$

$$\beta_j = 1 + \Delta t \left\{ \sigma^2 \left(\frac{S_j}{\Delta S}\right)^2 + r \right\},$$

$$\gamma_j = -\frac{1}{2}\Delta t \left\{ \sigma^2 \left(\frac{S_j}{\Delta S}\right)^2 + (r-q)\frac{S_j}{\Delta S} \right\},$$

$$\hat{\alpha}_j = 0,$$

$$\hat{\beta}_j = 1,$$

$$\hat{\gamma}_j = 0,$$

then the condition for the stability of the implicit method will be

$$\left| -\Delta t \sigma^2 \left(\frac{S_j}{\Delta S}\right)^2 \cos(N\Delta S) + \Delta t(r-q)\left(\frac{S_j}{\Delta S}\right) \sin(N\Delta S) + 1 + \Delta t \left(\sigma^2 \left(\frac{S_j}{\Delta S}\right)^2 + r \right) \right| \leq 1.$$

Lastly, for $\theta = \frac{1}{2}$; we have

$$\alpha_j = -\frac{1}{4}\Delta t \left\{ \sigma^2 \left(\frac{S_j}{\Delta S}\right)^2 - (r-q)\frac{S_j}{\Delta S} \right\},$$

$$\beta_j = 1 + \frac{1}{2}\Delta t \left\{ \sigma^2 \left(\frac{S_j}{\Delta S}\right)^2 + r \right\},$$

$$\gamma_j = -\frac{1}{4}\Delta t \left\{ \sigma^2 \left(\frac{S_j}{\Delta S}\right)^2 + (r-q)\frac{S_j}{\Delta S} \right\},$$

$$\hat{\alpha}_j = \frac{1}{4}\Delta t \left\{ \sigma^2 \left(\frac{S_j}{\Delta S}\right)^2 - (r-q)\frac{S_j}{\Delta S} \right\},$$

$$\hat{\beta}_j = 1 - \frac{1}{2}\Delta t \left\{ \sigma^2 \left(\frac{S_j}{\Delta S}\right)^2 + r \right\},$$

$$\hat{\gamma}_j = \frac{1}{4}\Delta t \left\{ \sigma^2 \left(\frac{S_j}{\Delta S}\right)^2 + (r-q)\frac{S_j}{\Delta S} \right\},$$

then the condition for the stability of the Crank-Nicolson method turns out to be

$$\left| -\frac{1}{2}\Delta t\sigma^2 \left(\frac{S_j}{\Delta S}\right)^2 \cos(N\Delta S) + \frac{1}{2}\Delta t(r-q)\left(\frac{S_j}{\Delta S}\right)\sin(N\Delta S) \right.$$

$$\left. + 1 + \frac{1}{2}\Delta t\Delta t\left(\sigma^2\left(\frac{S_j}{\Delta S}\right)^2 + r\right)\right|$$

$$\leq \left| \frac{1}{2}\Delta t\sigma^2 \left(\frac{S_j}{\Delta S}\right)^2 \cos(N\Delta S) - \frac{1}{2}\Delta t(r-q)\left(\frac{S_j}{\Delta S}\right)\sin(N\Delta S) \right.$$

$$\left. + 1 - \frac{1}{2}\Delta t\left(\sigma^2\left(\frac{S_j}{\Delta S}\right)^2 + r\right)\right|$$

Remark 4.3.1. One can find further generalized θ-schemes in [20].

Finally, we remark that the *Lax Equivalence Theorem* [43] states that stability is a necessary and sufficient condition for a consistent finite difference scheme. Therefore, the stability of a finite difference method implies the convergence of the consistent methods by the Lax-Equivalence Theorem.

4.3.6 American Put Options as a Free Boundary Problem

In this part, we will be concerned only with American put options. We know that an American option has the additional feature, unlike a European option, that it can be exercised at any time during the life of the option. Therefore, the valuation of an American option is more complicated than the valuation of a European option since we have to determine whether or not the American option should be exercised for each value of S. To avoid arbitrage opportunities at each grid point in the (S,t)-plane, the value of an American option should never be less than the immediate payoff in the case the option is exercised [46]. This is what is known as a free boundary problem. To be more specific, there is a contact point $S_f(t)$ which tells whether it is worth holding or exercising the option:

i) if $S < S_f(t)$, one should exercise the put option and gain $P_A(S,t) = \max\{K-S,0\} = K-S$,

ii) if $S > S_f(t)$, one should hold the option, and hence $P_A(S,t)$ satisfies the Black-Scholes equation.

Consequently, $S_f(t)$ will be called the optimal exercise price.

Unfortunately, we do not know $S_f(t)$, however, we can treat $S_f(t)$ as a new unknown which is called free boundary. To be able to calculate $S_f(t)$, an additional condition is needed and therefore we consider the slope $\frac{\partial P_A}{\partial S}$ more closely with which $P_A(S,t)$ touches to the straight line $K - S$ at the point $S_f(t)$: it turns out that the slope is -1. But, it should be realized that the condition does not follow from the fact that $P_A(S_f(t),t) = K - S_f(t)$. We need this extra condition since we do not know *a priori* where $S_f(t)$ is, and by the arbitrage arguments the gradient of P should be continuous, which is the condition we need [38]. See Figure 4.8 and Figure 4.9 for visualization.

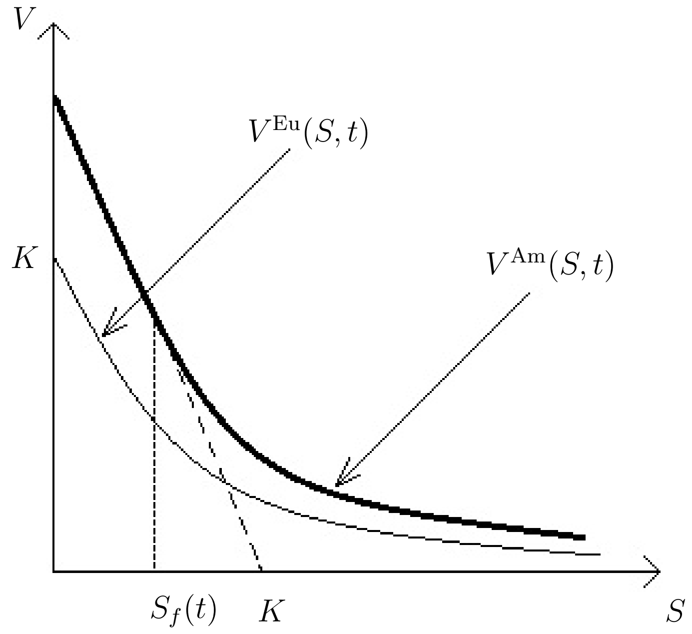

Figure 4.8: European and American option values for an American put option

Now, we have two boundary conditions at the contact point $S_f(t)$ as follows:

$$P_A(S_f(t),t) = K - S_f(t), \quad \frac{\partial P_A}{\partial S} = -1.$$

One can see that $P_A(S,t)$ touches the payoff function tangentially at $S_f(t)$ and this tangent point has an effect on the Black-Scholes inequality for an American

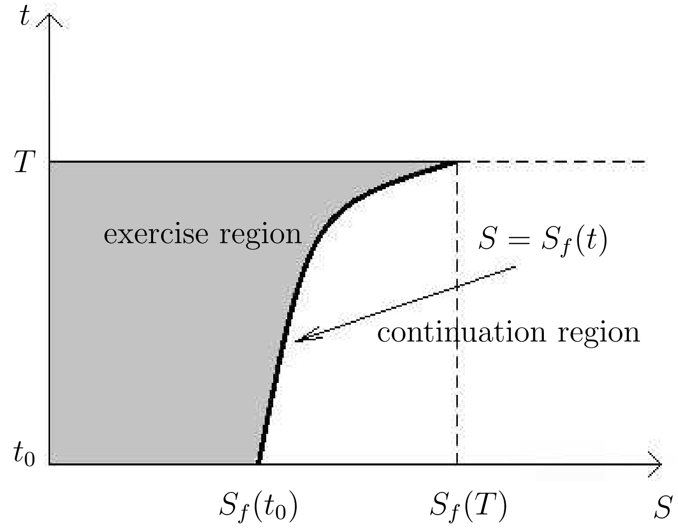

Figure 4.9: Exercise and continuation regions for American option

put option

$$\frac{\partial P_A}{\partial t} + \frac{1}{2}\sigma^2 \frac{\partial^2 P_A}{\partial S^2} + (r-q)S\frac{\partial P_A}{\partial S} - rP_A \leq 0 \tag{4.31}$$

which comes from the fact that the return from the portfolio cannot be greater than the return from a riskless asset as we have discussed previously in Section 4.1.

When it is optimal to hold the option, the Black-Scholes equation is valid and $P_A(S,t) \geq \max\{S-K, 0\}$. If not, it is optimal to exercise the option and only (4.31) holds. Also, $P_A(S,t) \geq \max\{S-K, 0\}$ is satisfied. Hence, we conclude that when $P_A = K - S$, for $S < K$, and (4.31) gives

$$\frac{\partial P_A}{\partial t} + \frac{1}{2}\sigma^2 \frac{\partial^2 P_A}{\partial S^2} + (r-q)S\frac{\partial P_A}{\partial S} - rP_A = -rK < 0.$$

Summarizing all these facts, in Figure 4.9, the American put problem can be written as a free boundary problem: for each time t, the S-axis is divided into two distinct regions. The first region is for $0 \leq S < S_f(t)$ where the early exercise is optimal and

$$P_A = K - S, \quad \frac{\partial P_A}{\partial t} + \frac{1}{2}\sigma^2 \frac{\partial^2 P_A}{\partial S^2} + (r-q)S\frac{\partial P_A}{\partial S} - rP_A < 0. \tag{4.32}$$

The second region is for $S_f(t) < S < \infty$, and the early exercise is not optimal there:

$$P_A > K - S, \quad \frac{\partial P_A}{\partial t} + \frac{1}{2}\sigma^2 \frac{\partial^2 P_A}{\partial S^2} + (r-q)S\frac{\partial P_A}{\partial S} - rP_A = 0. \tag{4.33}$$

Furthermore, we have the boundary conditions at $S = S_f(t)$, and the slope $\frac{\partial P_A}{\partial S}$ is continuous such that

$$P_A(S_f(t), t) = \max\{K - S_f(t), 0\}, \quad \frac{\partial P_A}{\partial S}(S_f(t), t) = -1.$$

The free boundary problem (4.32) and (4.33) for the American put option can be formulated as a linear complementarity problem [38]

$$\begin{aligned} \left(\frac{\partial P_A}{\partial t} + \frac{1}{2}\sigma^2 \frac{\partial^2 P_A}{\partial S^2} + (r-q)S\frac{\partial P_A}{\partial S} - rP_A\right) \cdot (P_A(S,t) - \Lambda(S)) = 0, \\ \left(\frac{\partial P_A}{\partial t} + \frac{1}{2}\sigma^2 \frac{\partial^2 P_A}{\partial S^2} + (r-q)S\frac{\partial P_A}{\partial S} - rP_A\right) \leq 0, \\ (P_A(S,t) - \Lambda(S)) \geq 0, \end{aligned} \tag{4.34}$$

which does not explicitly include the free boundary and where $\Lambda(S) = \max\{K - S, 0\}$. The final condition is $P_A(S,T) = \Lambda(S)$ and the boundary conditions are

$$P_A(0,t) = K, \quad \lim_{S\to\infty} P_A(S,t) = 0,$$

as stated previously.

4.3.7 The American Call with Dividends

We now consider the model for an American call option on a dividend-paying asset. The value of the call option $C_A(S,t)$ satisfies

$$\frac{\partial C_A}{\partial t} + \frac{1}{2}\sigma^2 \frac{\partial^2 C_A}{\partial S^2} + (r-q)S\frac{\partial C_A}{\partial S} - rC_A = 0, \tag{4.35}$$

and the payoff condition at maturity T is $C_A(S,T) = \max\{S-K, 0\}$. Since the option can be exercised at any time, we have

$$C_A(S,t) \geq \max\{S - K, 0\}.$$

At the point $S = S_f(t)$ which is the optimal exercise boundary, we have

$$C_A(S_f(t), t) = S_f(t) - K, \quad \frac{\partial C_A}{\partial S}(S_f(t), t) = 1.$$

If there is no optimal exercise boundary, then only (4.35) is valid when $C_A(S,t) > \max\{S-K,0\}$. Then, (4.35) can be again written as an inequality

$$\frac{\partial C_A}{\partial t} + \frac{1}{2}\sigma^2 \frac{\partial^2 C_A}{\partial S^2} + (r-q)S\frac{\partial C_A}{\partial S} - rC_A \leq 0.$$

Therefore, the option value is

$$C_A(S,t) = \begin{cases} S-K, & 0 \leq S < S_f(t), \\ \max\{S-K,0\}, & S_f(t) < S < \infty, \end{cases}$$

with boundary and final conditions

$$\begin{aligned} C_A(0,t) &= 0, \\ C_A(S,T) &= \max\{S-K,0\}, \\ C_A(S_f(t),t) &= S_f(t) - K. \end{aligned}$$

For rewriting the American call with underlying dividend paying assets in the form of free boundary problem and its linear complementarity problem, see [38]. Also, it is possible to obtain the finite difference formulation for the corresponding linear complementarity problem for an American call option, similar to the American put option.

4.3.8 Finite Difference Discretization of the Linear Complementarity Problem for an American Put Option

In this section, we will formulate the θ-averaged scheme for the linear complementarity problem for American put options given by (4.34). The finite difference approximation for the Black-Scholes equation using θ-method is derived to be

$$\alpha_j w_{j-1,i-1} + \beta_j w_{j,i-1} + \gamma_j w_{j+1,i-1} = \hat{\alpha}_j w_{j-1,i} + \hat{\beta}_j w_{ji} + \hat{\gamma}_j w_{j+1,i}, \qquad (4.36)$$

where $\alpha_j, \beta_j, \gamma_j, \hat{\alpha}_j, \hat{\beta}_j, \hat{\gamma}_j$ are given in (4.27) for all $i = M-1, M-2, \ldots, 1$ and $j = 1, \ldots, N-1$.

The boundary condition (4.14) gives

$$w_{N-1}^{(i-1)} = w_{N-1}^{(i)} = 0.$$

Therefore, the linear complementarity problem (4.34) can be written in the form:

$$\begin{aligned} Aw^{(i-1)} - C^{(i)} &\geq 0, \\ w^{(i-1)} &\geq \Lambda^{(i)}, \\ (Aw^{(i-1)} - Bw^{(i)}) \cdot (w^{(i-1)} - \Lambda^{(i)}) &= 0, \end{aligned} \tag{4.37}$$

where

$$A = \begin{pmatrix} 1-\beta_1 & \gamma_1 & 0 & \dots & 0 & 0 \\ \alpha_2 & 1-\beta_2 & \gamma_2 & \dots & 0 & 0 \\ 0 & \alpha_3 & 1-\beta_3 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1-\beta_{N-2} & \gamma_{N-2} \\ 0 & 0 & 0 & \dots & \alpha_{N-1} & 1-\beta_{N-1} \end{pmatrix}, \tag{4.38}$$

$C^{(i)} = Bw^{(i)} + y^{(i-1)}$, and

$$\Lambda^{(i)} = \begin{pmatrix} \Lambda_{1,i} \\ \vdots \\ \Lambda_{N-1,i} \end{pmatrix},$$

and $\Lambda_{j,i} = \Lambda(j\Delta S, i\Delta t)$.

Remark 4.3.2. Here, the expression $a \geq b$, where a and b are vectors, means that each component of a is greater than or equal to the corresponding component of b, that is, $a_n \geq b_n$ for all n.

After this point, we now give a widely used iterative algorithm for pricing American options in literature. The problem (4.37) is in fact equivalent to

$$\min \left\{ Aw^{(i-1)} - C^{(i)} w^{(i)}, w^{(i-1)} - \Lambda^{(i)} \right\} = 0,$$

so that

$$\min \left\{ w^{(i-1)} - A^{-1} C^{(i)}, w^{(i-1)} - \Lambda^{(i)} \right\} = 0$$

holds. Therefore, we obtain

$$w^{(i-1)} = \max \left\{ A^{-1} C^{(i)}, \Lambda^{(i)} \right\}.$$

4.3.9 The Projected (SOR) Method

We will discuss the iterative solvers for linear systems only if certain conditions for convergence hold. The Gauss-Seidel algorithm solves the linear system iteratively and it converges to the true solution of the system. However, a method, called the successive overrelaxation (SOR) to accelerate convergence will be used. Then, we will use the modified SOR method which is called the projected SOR.

The Gauss-Seidel Method

Now, we will consider the Gauss-Seidel method to solve the linear systems iteratively. For the details, see [13].

Consider a linear system of equations

$$Ax = b,$$

where

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}.$$

Then, the Gauss-Seidel iteration is defined as

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left\{ b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=1+1}^{n} a_{ij} x_j^{(k)} \right\}, \tag{4.39}$$

for a given initial guess $x^{(0)} = [x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)}]^T$, for all $i = 1, 2, \dots, n$ and $k = 1, 2, \dots$. The iteration should stop for some positive integer k.

If the matrix A is a strictly diagonally dominant matrix, that is

$$|a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}|,$$

for every $i = 1, 2, \dots, n$, then the Gauss-Seidel method for $Ax = b$ is convergent.

Recall that θ-method is writen as $Aw^{i-1} = Bw^i + y^i = \psi^{(i)}$, then the iteration formula for our problem takes a form such as

$$\hat{w}_{ji}^{(k+1)} = \frac{1}{\alpha_{ii}} \left\{ \psi_i - \sum_{r=1}^{j-1} \alpha_{jr} w_{ri}^{(k+1)} - \sum_{r=j+1}^{N-1} \alpha_{jr} w_{ri}^{(k)} \right\},$$

for each $j = 1, \ldots, N-1$ and $k = 1, \ldots$.

Successive Over Relaxation (SOR) Method

For an introduced relaxation parameter $\omega \in (0, 2)$, (4.39) turns out to be

$$x^{(k+1)} = x^{(k)} + \omega(x^{(k+1)} - x^{(k)}) = (1-\omega)x^{(k)} + \omega x^{(k+1)}, \tag{4.40}$$

which is called successive overrelaxation. Note that

- if $\omega < 1$, then the scheme is called underrelaxation;
- if $\omega > 1$, the scheme is called overrelaxation;
- if $\omega = 1$, it is exactly the Gauss-Seidel method.

The equation (4.40) can be rewritten by using the components of each iterate $x^{(k+1)}$ as

$$x_i^{(k+1)} = x_i^{(k)} + \frac{\omega}{a_{ii}} \left\{ (b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=1}^{n} a_{ij} x_j^{(k)}) \right\}$$

for every $i = 1, 2, \ldots, n$.

The iteration for our problem will be

$$\hat{w}_{ji}^{(k+1)} = \hat{w}_{ji}^{(k)} + \frac{\omega}{\alpha_{jj}} \left\{ (\psi_i - \sum_{r=1}^{j-1} \alpha_{jr} w_{ri}^{(k+1)} - \sum_{r=j+1}^{N-1} \alpha_{jr} w_{ri}^{(k)}) \right\},$$

where

$$Bw^{(i)} + y^{(i)} := \psi_i$$

for each $j = 1, \ldots, N-1$ and $k = 1, \ldots$.

The Projected SOR Method

The projected SOR method which was developed by Cryer [18] includes the iterative solutions of linear system of equations. We aim to modify the SOR method to value American options by projected SOR method. We apply the method to the finite difference formulation (4.36).

We approximate the solution of the linear system of equations by a modified iterative method such as SOR. Let us denote $A = (\alpha_{ji})$ and the desired solution $w^{(i)} = (w_{1,i}, \ldots, w_{N-1,i})^T$. We start with an initial guess $w_i^{(0)} = (w_{1,i}^0, \ldots, w_{N-1,i}^0)^T$ and we choose a relaxation parameter, $\omega \in (0, 2)$, to converge to the desired solution of the iteration, that is proved by Kahan [30]. Then, we define an iteration such that

$$\hat{w}_{ji}^{(k)} = \hat{w}_{ji}^{(k-1)} + \frac{\omega}{\alpha_{jj}} \left\{ \psi_i - \sum_{r=1}^{j-1} \alpha_{jr} w_{ri}^{(k)} - \sum_{r=j+1}^{N-1} \alpha_{jr} w_{ri}^{(k-1)} \right\},$$

where

$$Bw^{(i)} + y^{(i)} := \psi_i$$

for each $j = 1, \ldots, N-1$ and $k = 1, \ldots$.

Here, the iteration continues for a given tolerance ϵ, by the condition

$$\left\| \hat{w}_{ji}^{(k+1)} - \hat{w}_{ji}^{(k)} \right\| \leq \epsilon.$$

To construct the algorithm for an American option, we should compare the SOR iteration with the payoff Λ_{ji} at the (S_j, t_i) nodes. This modified algorithm is called the projected SOR method. Therefore, the approximate value of the American option is given by

$$\hat{w}_{ji}^{(PSOR)} = \max \left\{ \hat{w}_{ji}^{(SOR)}, \Lambda_{ji} \right\},$$

where $\hat{w}_{ji}^{(SOR)}$ is called as the SOR iteration under the no free boundary assumption.

As a summary, we may give compose the projected SOR method in Algorithm 2 for American put options.

Algorithm 2 The projected SOR Method for American Put Options

Given S_0, K, r, D, σ, T.

Grid variables: S_{max}, S_{min}.

Define $\Delta S = \frac{S_{max} - S_{min}}{N}$ where N is the number of the nodes, and $\Delta t = \frac{T}{M}$ where M is the number of the time intervals.

Choose the algorithm variables such that $\theta \in [0, 1]$ and $\omega \in (0, 2)$.

Choose a PSOR convergence tolerance ϵ.

Set up matrices A and B with boundary conditions.

Start with an initial guess $w_i^{(0)} = (w_{1,i}^0, \ldots, w_{N-1,i}^0)^T$.

for $j = 1, \ldots, N-1$ **do**

set $Bw^{(i)} + y^{(i)} = \psi_i$

for $k = 1, \ldots, K$ **do**

calculate $\hat{w}_{ji}^{(k)} = \hat{w}_{ji}^{(k-1)} + \frac{\omega}{\alpha_{jj}} \left\{ \psi_i - \sum_{r=1}^{j-1} \alpha_{jr} w_{ri}^{(k)} - \sum_{r=j+1}^{N-1} \alpha_{jr} w_{ri}^{(k-1)} \right\}$

set $\hat{w}_{ji}^{(PSOR)} = \max\{\hat{w}_{ji}^{(K)}, \Lambda_{ji}\}$

while $\left\| \hat{w}_{ji}^{(k+1)} - \hat{w}_{ji}^{(k)} \right\| > \epsilon$ **do**

$w_{j,i-1} = w_{j,i-1}^{(0)} = \max\{w_{ji}^M, \Lambda_{j,i-1}\}$. {% PSOR loop}

end while

end for

end for

Return: (w_{ji}).

Example 6. In Figure 4.10, Crank-Nicolson method for European and American call options are compared with the parameters $S_0 = 60, K = 60, D = 0, \sigma = 0.2, r = 0.3, T = 3/12, S_{min} = 0, S_{max} = 150, dS = 0.5, dt = 1/1000$.

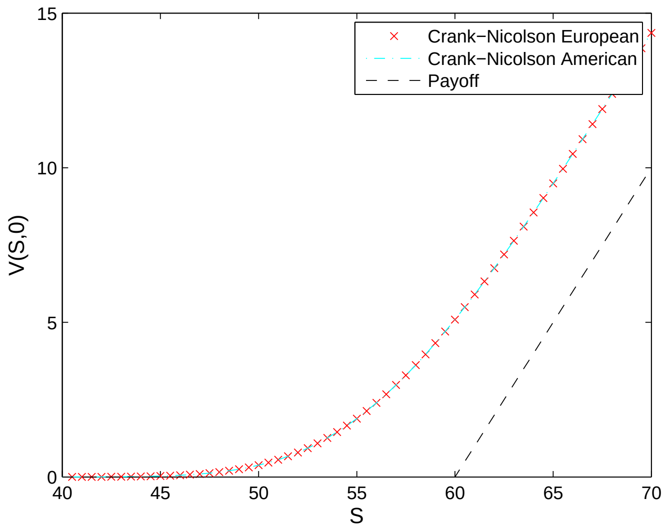

Figure 4.10: The comparison of the Crank-Nicolson method for European and American call options on a non-dividend paying asset

It can be seen that the prices of European and American options are the same for non-dividend paying underlying asset which confirms the theoretical fact. In the case of a dividend payment on a call option with $D = 0.5$, the European option value will be less valuable than the American option as shown in Figure 4.11. Furthermore, Crank-Nicolson results for European and American put options with the same parameters and $D = 0$ are illustrated in Figure 4.12.

Figures so far demonstrate that the value of the American call options on dividend and non-dividend paying asset and put options are always greater than the European option values.

Example 7. In Figure 4.13 and Figure 4.14, we illustrate and compare the Crank Nicolson method under LU and Projected SOR methods for an American put option with the parameters $S_0 = 60, K = 60, D = 0, \sigma = 0.4, r = 0.1, T = 3/12, S_{min} = 0, S_{max} = 150, dS = 0.5, dt = 1/1200, \omega = 0$.

Note that when the stock price S is getting near to the strike price, LU and

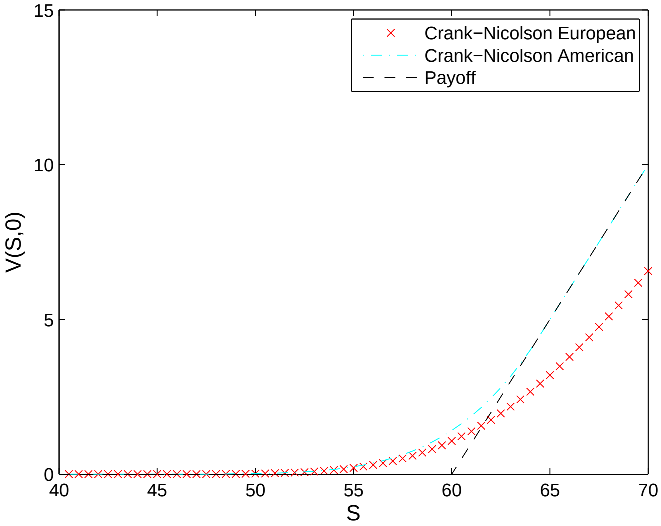

Figure 4.11: The comparison of the Crank-Nicolson method for European and American call options on a dividend paying asset

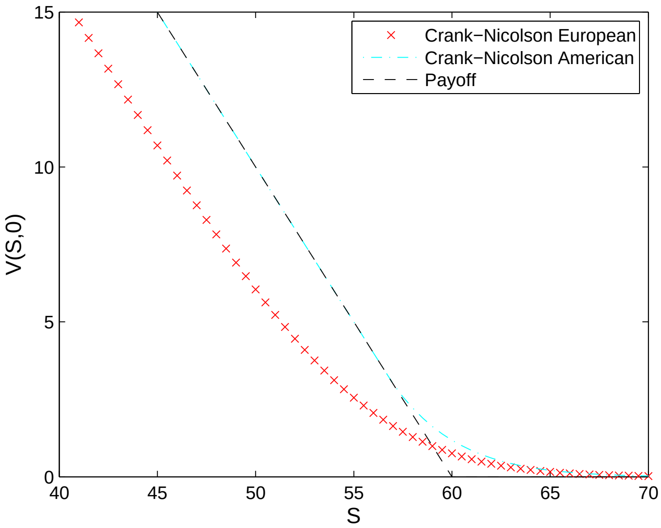

Figure 4.12: The comparison of the Crank-Nicolson method for European and American put options on a non-dividend paying asset

PSOR solutions differ from each other.

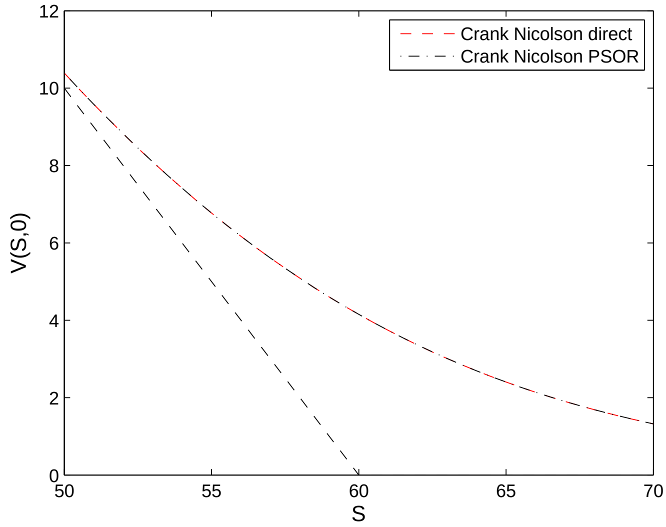

Figure 4.13: Crank Nicolson *LU* and PSOR solutions for an American put option on non-dividend paying asset

Because of the non-differentiable final condition, Crank-Nicolson method oscillates around the strike price. To eliminate this behaviour of the Crank-Nicolson method, one can use Implicit method in first few steps and then go on with the Crank-Nicolson after these steps. In Figure 4.18, this behaviour is demonstrated.

MATLAB codes of the Explicit, Implicit and Crank-Nicolson methods can be found in [46].

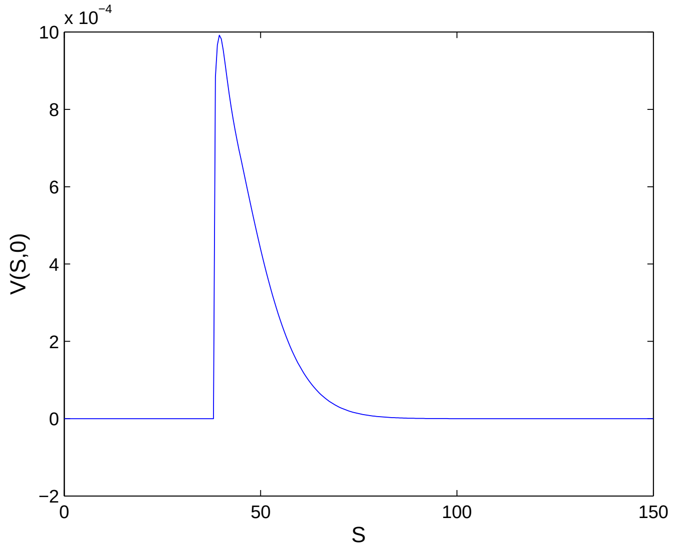

Figure 4.14: The difference between Crank Nicolson LU and PSOR solutions for an American put option on non-dividend paying asset

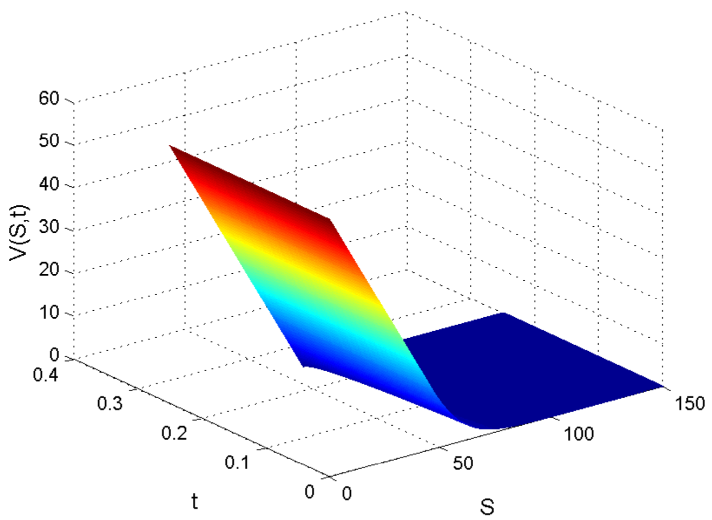

Figure 4.15: The surface $V = V(S, t)$ for an American put option on non-dividend paying asset

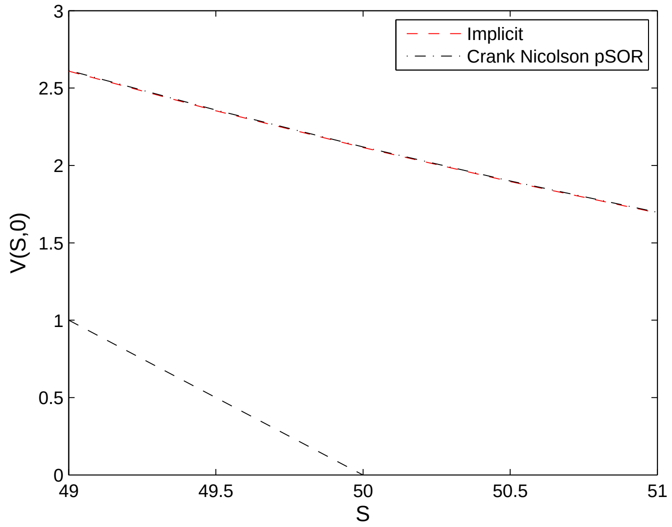

Figure 4.16: The comparison of Implicit and Crank Nicolson PSOR for a put option with parameters $S = 50, K = 50, D = 0, \sigma = 0.4, r = 0.1, T = 1/12, S_{min} = 0, S_{max} = 150, dS = 0.05, dt = 1/1200$

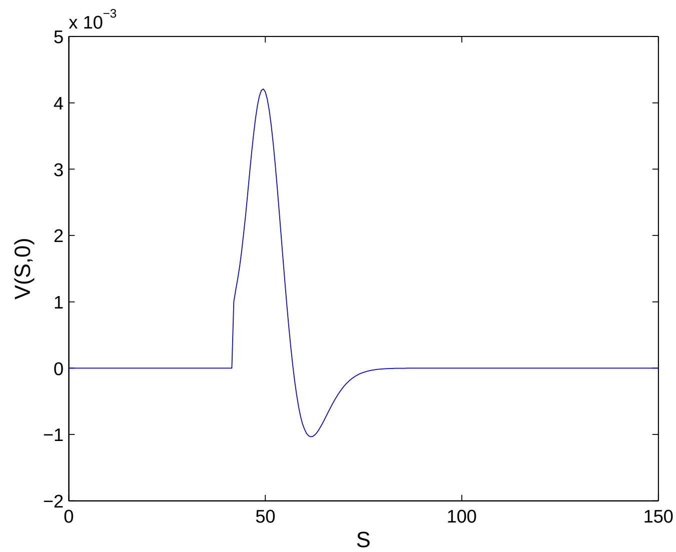

Figure 4.17: The error between Implicit and Crank Nicolson PSOR for a put option with the same parameters

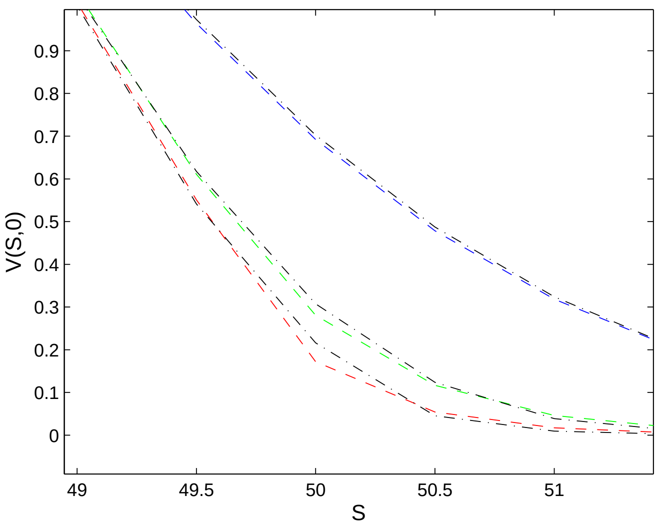

Figure 4.18: Implicit and Crank Nicolson solutions

CHAPTER 5

APPROXIMATION METHODS IN PRICING AMERICAN OPTIONS

In this chapter, we study the analytical approximations for the valuation of American options. Typical methods in this category are the Roll [39], Geske [22, 23], Whaley [47] Approximation, the Bjerksund and Stensland [7, 8] Approximation, and the Quadratic Approximation (Barone-Adesi Whaley) [3] and also Least-Squares Monte Carlo (LSM) [36] method. After introducing the techniques numerical examples will be given and the results coming from the approximation methods with other methods such as binomial and finite differences will be compared.

5.1 Roll-Geske-Whaley Approximation

The Roll-Geske-Whaley (RGW) Approximation is based on a Black-Scholes model which is the case of only one dividend during the life of the American call option. The RGW model was constructed by Roll [39], Geske [22, 23], and Whaley [47].

In this model, the stock price S is given by using the discounted stock price; $S = S_0 - De^{-rt}$, where the dividend D is discounted to the ex-dividend date t. Then, the option is priced as if it is a European option. The RGW model uses the approximated price which corresponds to the risky period $(T - t)$ of the life of the option.

Theorem below formalizes the Roll-Geske-Whaley model, and then we analyze the proof of the method.

Theorem 5.1.1 (Roll-Geske-Whaley Formula)**.** *The price of an American call on a stock with a single dividend paid during the life of the option is given by*

$$\begin{aligned} C(S_0, K, T, 0) =& (S_0 - De^{-rt})N(b_1) - (K - D)e^{-rt}N(b_2) \\ & + (S_0 - De^{-rt})M\left(a_1, -b_1; -\sqrt{\frac{t}{T}}\right) - Ke^{-rt}M\left(a_2, -b_2; -\sqrt{\frac{t}{T}}\right) \end{aligned}$$

where

$$a_1 = \frac{\ln\left[(S_0 - De^{-rt})/K\right] + (r + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}, \quad a_2 = a_1 - \sigma\sqrt{T},$$

$$b_1 = \frac{\ln\left[(S_0 - De^{-rt})/S^*\right] + (r + \frac{1}{2}\sigma^2)t}{\sigma\sqrt{t}}, \quad b_2 = b_1 - \sigma\sqrt{t}. \tag{5.1}$$

Here, $N(\cdot)$ is the distribution function of the standard normal distribution and $M(a, b; \rho)$ is the distribution function of the bivariate standard normal distribution with correlation ρ. The term S^* is the fixed point of the equation

$$C_{BS}(S^*, T, K) = S^* + D - K \tag{5.2}$$

where C_{BS} is the Black-Scholes price of a European call.

Proof. We know by Lemma 2.1.1 that, it may be only optimal to exercise an American call just before the dividend payment. So,

$$C_A(S_t, K, T, t) = \mathbb{E}[e^{-rt}\max\{(S_t - K)^+, C(S_t, K, T)\}].$$

Since

$$C(S^*, K, T) = S^* + D - K,$$

then we have

$$\begin{aligned} C_A(S_t, K, T, t) &= \mathbb{E}[e^{-rt}(S_t + D - K)\mathbf{1}_{\{S_t \geq S^*\}}] + \mathbb{E}[e^{-rt}C(S_t, K, T)\mathbf{1}_{\{S_t \leq S^*\}}] \\ &:= A + B \end{aligned}$$

where the indicator function $\mathbf{1}_{\{A\}}(x)$ which is defined as

$$\mathbf{1}_{\{A\}}(x) = \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases}.$$

Now, we first express S_t in terms of $(S_0 - De^{-rt})$. By the Itô formula, we know that the solution of the stochastic differential equation

$$dS = rS_t dt + \sigma S_t dW_t \tag{5.3}$$

is

$$S_t = (S_0 - De^{-rt})e^{(r-\frac{1}{2}\sigma^2)t+\sigma W_t},$$

where $W_t = X\sqrt{t}$ is the Brownian motion and X is a standard normal random variable.

On the other hand, evaluation of A can be carried out as follows:

$$\begin{aligned} A &= \mathbb{E}\left[e^{-rt}(S_t + D - K)\mathbf{1}_{\{S_t \geq S^*\}}\right] \\ &= \int_{-\infty}^{\infty} e^{-rt}(S_t + D - K)\mathbf{1}_{\{S_t \geq S^*\}} f_X(x)dx \\ &= \int_{-\infty}^{\infty} e^{-rt}((S_0 - De^{-rt})e^{(r-\frac{1}{2}\sigma^2)t+\sigma x\sqrt{t}} + D - K)\mathbf{1}_{\{S_t \geq S^*\}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx. \end{aligned}$$

Since $S_t \geq S^*$, we get

$$\begin{aligned} (S_0 - De^{-rt})e^{(r-\frac{1}{2}\sigma^2)t+\sigma x\sqrt{t}} &\geq S^* \\ e^{(r-\frac{1}{2}\sigma^2)t+\sigma x\sqrt{t}} &\geq \frac{S^*}{S_0 - De^{-rt}} \\ (r - \frac{1}{2}\sigma^2)t + \sigma x\sqrt{t} &\geq \ln\left(\frac{S^*}{S_0 - De^{-rt}}\right) \\ x &\geq \frac{\ln\left(\frac{S^*}{S_0 - De^{-rt}}\right) - (r - \frac{1}{2}\sigma^2)t}{\sigma\sqrt{t}} := C. \end{aligned}$$

Hence,

$$\begin{aligned} A &= \int_C^{\infty} (S_0 - De^{-rt}) \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\sigma\sqrt{t})^2}{2}} dx - \int_C^{\infty} e^{-rt}(K - D)\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\ &= (S_0 - De^{-rt})(1 - N(C - \sigma\sqrt{t})) - (K - D)e^{-rt}(1 - N(C)) \\ &= (S_0 - De^{-rt})N(b_1) - (K - D)e^{-rt}N(b_2), \end{aligned}$$

where b_1 and b_2 are given in (5.1), and N is defined in (20).

Now, we continue by calculating B as follows:

$$\begin{aligned}
B =& \mathbb{E}\left[e^{-rt}C(S_t, K, T)\mathbf{1}_{\{S_t \leq S^*\}}\right] \\
=& \int_{-\infty}^{C} e^{-rt} C_{BS}(S_t, K, T) \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx
\end{aligned}$$

since $S_t \leq S^*$ implies $x \leq C$. Furthermore,

$$\begin{aligned}
B =& \int_{-\infty}^{C} e^{-rt} \left[S_t N\left(\frac{\ln(\frac{S_t}{K}) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T}} \right) \right. \\
& \left. - KN\left(\frac{\ln(\frac{S_t}{K}) + (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T}} \right) \right] \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\
=& \int_{-\infty}^{C} e^{-rt}(S_0 - De^{-rt}) e^{(r-\frac{1}{2}\sigma^2)t + \sigma x\sqrt{t}} \\
& \times N\left(\frac{\ln(\frac{S_0 - De^{-rt}}{K}) + (r - \frac{1}{2}\sigma^2)t + \sigma x\sqrt{t} + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T}} \right) \\
& \times \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\
& - \int_{-\infty}^{C} e^{-rt} KN \\
& \times \left(\frac{\ln(\frac{S_0 - De^{-rt}}{K}) + (r - \frac{1}{2}\sigma^2)t + \sigma x\sqrt{t} + (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T}} \right) \\
& \times \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx
\end{aligned}$$

Since

$$e^{-rt} e^{(r-\frac{1}{2}\sigma^2)t + \sigma x\sqrt{t}} e^{-\frac{x^2}{2}} = e^{-\frac{(x-\sigma\sqrt{t})^2}{2}},$$

we have

$$\begin{aligned}
B =& \int_{-\infty}^{C} (S_0 - De^{-rt}) N\left(\frac{\ln(\frac{S_0 - De^{-rt}}{K}) + (r + \frac{1}{2}\sigma^2)T - \sigma^2 t + \sigma x\sqrt{t}}{\sigma\sqrt{T}} \right) \\
& \times \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\sigma\sqrt{t})^2}{2}} dx \\
& - \int_{-\infty}^{C} e^{-rt} KN\left(\frac{\ln(\frac{S_0 - De^{-rt}}{K}) + (r - \frac{1}{2}\sigma^2)T + \sigma x\sqrt{t}}{\sigma\sqrt{T}} \right) \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\
:=& B_1 - B_2.
\end{aligned}$$

We define

$$\beta := \frac{\ln(\frac{S_0 - De^{-rt}}{K}) + (r + \frac{1}{2}\sigma^2)T - \sigma^2 t}{\sigma\sqrt{T}}.$$

Then,

$$B_1 = \int_{-\infty}^{C} (S_0 - De^{-rt}) N\left(\beta + x\sqrt{\frac{t}{T}}\right) \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\sigma\sqrt{t})^2}{2}} dx$$

$$= \int_{-\infty}^{C} (S_0 - De^{-rt}) \left(\int_{-\infty}^{\beta + x\sqrt{\frac{t}{T}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz\right) \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\sigma\sqrt{t})^2}{2}} dx.$$

Then, by Lemma A.1.3, we have

$$B_1 = (S_0 - De^{-rt}) \mathbb{P}(X \leq C, Z \leq \beta),$$

where

$$(X, Z) \sim \mathcal{M}\left(\left(\begin{array}{c} \sigma\sqrt{t} \\ \sigma\frac{t}{\sqrt{T}} \end{array}\right), \left(\begin{array}{cc} 1 & -\sqrt{\frac{t}{T}} \\ -\sqrt{\frac{t}{T}} & \frac{T+t}{T} \end{array}\right)\right).$$

We have $\mu_1 = \sigma\sqrt{t}$ and it is known that $Z_1 = \frac{X-\mu_1}{\sigma_1}$. Hence,

$$X = \sigma_1 Z_1 + \mu_1 = Z_1 + \sigma\sqrt{t} \leq C.$$

Here, $Z_1 \sim N(0,1)$ and $Z_1 \leq C - \sigma\sqrt{t} = b_1$.

Therefore,

$$B_1 = (S_0 - De^{-rt}) \mathbb{P}(Z_1 \leq -b_1, Z \leq a_1)$$

with

$$(Z_1, Z) \sim \mathcal{M}\left(\left(\begin{array}{c} 0 \\ 0 \end{array}\right), \left(\begin{array}{cc} 1 & -\sqrt{\frac{t}{T}} \\ -\sqrt{\frac{t}{T}} & 1 \end{array}\right)\right),$$

and, hence,

$$B_1 = (S_0 - De^{-rt}) M\left(a_1, -b_1; -\sqrt{\frac{t}{T}}\right).$$

Similarly, we obtain easily calculate B_2 to yield

$$B_2 = \int_{-\infty}^{C} e^{-rt} K N\left(\frac{\ln(\frac{S_0 - De^{-rt}}{K}) + (r - \frac{1}{2}\sigma^2)T + \sigma x\sqrt{t}}{\sigma\sqrt{T}}\right) \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx.$$

Further, we define

$$\delta := \frac{\ln(\frac{S_0 - De^{-rt}}{K}) + (r - \frac{1}{2}\sigma^2)T - \sigma^2 t}{\sigma\sqrt{T}}$$

so that

$$B_2 = \int_{-\infty}^{C} e^{-rt} K \cdot N\left(\delta + x\sqrt{\frac{t}{T}}\right) \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$
$$= \int_{-\infty}^{C} e^{-rt} K \left(\int_{-\infty}^{\delta + x\sqrt{\frac{t}{T}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz\right) \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx.$$

By Lemma A.1.3, it follows that

$$B_2 = e^{-rt} K \mathbb{P}(X \leq C, Z \leq \delta)$$

with

$$(X, Z) \sim \mathcal{M}\left(\begin{pmatrix} \sigma\sqrt{t} \\ \sigma\frac{t}{\sqrt{T}} \end{pmatrix}, \begin{pmatrix} 1 & -\sqrt{\frac{t}{T}} \\ -\sqrt{\frac{t}{T}} & 1 \end{pmatrix}\right).$$

Again, we have $\mu_2 = \sigma\sqrt{t}$ and it is already known that $Z_2 = \frac{X - \mu_2}{\sigma_2}$. Thus we get,

$$X = \sigma_2 Z_2 + \mu_2 = Z_2 + \sigma\sqrt{t} \leq C.$$

Here, $Z_2 \sim N(0, 1)$ and $Z_2 \leq C - \sigma\sqrt{t} = b_2$.

Therefore, it follows that

$$B_2 = e^{-rt} K \mathbb{P}(Z_2 \leq -b_2, Z \leq a_2)$$

with

$$(Z_2, Z) \sim \mathcal{M}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & -\sqrt{\frac{t}{T}} \\ -\sqrt{\frac{t}{T}} & 1 \end{pmatrix}\right).$$

Finally, we obtain

$$B_2 = e^{-rt} K M\left(a_2, -b_2; -\sqrt{\frac{t}{T}}\right)$$

which completes the proof. □

Remark 5.1.2. For the price at time t and more than one dividend, see [34].

In [33], Roll-Geske-Whaley formula is derived for different dividend dates, and the proof is given for any time $t < T$.

Table 5.1: The comparison of the method Roll-Geske-Whaley with the Binomial method for an American call option

Stock Price (S_0)	Roll-Geske-Whaley	Binomial
40	0.9732	0.5372
45	2.2883	2.2477
50	4.3769	3.9582
55	7.2278	7.0739
60	10.7323	10.9900
65	14.7416	14.9062
70	19.1110	18.8447
75	23.7224	23.7089
80	28.4890	28.5730
85	33.3516	33.4372

Example 8. Consider an American call option on stock with a single dividend payment $D = 0.2$ at time $t = 3/12$ where the time to maturity is $T = 9/12$. The strike price $K = \$60$, the risk free interest rate is 20% and the volatility is 30%. We compare the Roll-Geske-Whaley method with the Binomial method for these parameters as given in Table 5.1.

In our application, the critical price S^* is determined by using the Bisection method [13]. MATLAB implementation of the method can be found in Appendix B.

5.2 Quadratic Approximation (Barone-Adesi Whaley Formula)

We now discuss the Quadratic Approximation which was developed by Barone-Adesi and Whaley [3]. Quadratic approximation is an accurate and inexpensive method which can be used to value stock indices, currencies, futures and stocks carrying a constant dividend yield [3]. The method is based on the Black-Scholes model and the Merton model, and provides an analytic approximation solution to the values of the American options on stocks which pay continuous dividends.

5.2.1 Quadratic Approximation for an American Call Option

The basic key for the quadratic approximation approach is that, if the modified Black-Scholes differential equation

$$\frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + (r-q)S\frac{\partial V}{\partial S} - rV + \frac{\partial V}{\partial t} = 0$$

applies to American options as well as European options, then the early exercise premium which is $w(S,T)$ for an American call option written on a commodity, is defined as

$$w(S,T) = C_A(S,T) - C(S,T),$$

where $C_A(S,T)$ is the American call option and $C(S,T)$ is the European call option. Thus, the partial differential equation for $w(S,T)$ is defined as

$$\frac{\partial w}{\partial t} + (r-q)S\frac{\partial w}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 w}{\partial S^2} - rw = 0 \tag{5.4}$$

where r is the interest rate, σ is the volatility and q is the dividend yield. Let us define the following variables:

$$\begin{aligned}
\tau &= T - t, \\
h(\tau) &= 1 - e^{-r\tau}, \\
\alpha &= \frac{2r}{\sigma^2}, \\
\beta &= \frac{2(r-q)}{\sigma^2}, \\
w &= h(\tau)g(S,h).
\end{aligned} \tag{5.5}$$

Since $\tau = T - t$, we have $\frac{\partial w}{\partial \tau} = -\frac{\partial w}{\partial t}$. If the equation (5.4) is multiplied by $\frac{2}{\sigma^2}$, and the substitutions $\alpha = \frac{2r}{\sigma^2}$ and $\beta = \frac{2(r-q)}{\sigma^2}$ are used, we have rewritten (5.4) as follows:

$$S^2\frac{\partial^2 w}{\partial S^2} - \alpha w + \beta S\frac{\partial w}{\partial S} - \frac{\alpha}{r}\frac{\partial w}{\partial T} = 0. \tag{5.6}$$

Using the substitutions (5.5), we get

$$\begin{aligned}
\frac{\partial w}{\partial \tau} &= g\frac{\partial h}{\partial \tau} + h\frac{\partial g}{\partial h}\frac{\partial h}{\partial \tau}, \\
\frac{\partial w}{\partial S} &= h\frac{\partial g}{\partial S}, \\
\frac{\partial^2 w}{\partial S^2} &= h\frac{\partial^2 g}{\partial S^2}.
\end{aligned}$$

Substituting these partial derivatives into (5.6), we have

$$S^2\frac{\partial^2 g}{\partial S^2}+\beta S\frac{\partial g}{\partial S}\alpha w-\frac{\alpha}{h}g-(1-h)\alpha\frac{\partial g}{\partial h}=0. \tag{5.7}$$

To make an approximation, the last term on the left-hand side of the equation (5.7) will be assumed to be equal to 0. For the options with very short (long) times to expiration, this assumption is reasonable, since as T goes to 0 (∞), $\frac{\partial g}{\partial h}$ goes to 0 (h goes to 1), and the term $(1-h)\alpha\frac{\partial g}{\partial h}$ disappears. Therefore, the last term is dropped and the partial differential equation is

$$S^2\frac{\partial^2 g}{\partial S^2}+\beta S\frac{\partial g}{\partial S}\alpha w-\frac{\alpha}{h}g=0. \tag{5.8}$$

There are two linearly independent solutions of the form $g=aS^{\lambda}$ of the ordinary differential equation (5.8) and substituting $g=aS^{\lambda}$ into (5.8), we have

$$aS^{\lambda}\left[\lambda^2+(\beta-1)\lambda-\frac{\alpha}{h}\right]=0. \tag{5.9}$$

The roots of the indicial equation (5.9) are

$$\lambda_1=\frac{1-\beta-\sqrt{(\beta-1)^2+\frac{4\alpha}{h}}}{2}$$

and

$$\lambda_2=\frac{1-\beta+\sqrt{(\beta-1)^2+\frac{4\alpha}{h}}}{2}. \tag{5.10}$$

Thus, the general solution of (5.8) is given by

$$g(S)=a_1S^{\lambda_1}+a_2S^{\lambda_2}. \tag{5.11}$$

Here, λ_1 and λ_2 are known, so a_1 and a_2 should be determined. Since $\lambda_1<0$ and $a_1\neq 0$, the function g goes to ∞ as S goes to 0 and the early exercise premium of the American call becomes worthless when S goes to 0, so it is not reasonable. Hence, we have $a_1=0$, and the approximate value of the American call is

$$C_A(S,T)=C(S,T)+ha_2S^{\lambda_2}. \tag{5.12}$$

We note that as $S=0$, $C_A(S,T)$=0. However, if $C(S,T)$ and $ha_2S^{\lambda_2}$ rise, then the value of $C_A(S,T)$ rises as S rises where $a_2>0$. The value of the American call is equal to its exercisable proceeds, $S-K$ called S^* which is the point of tangency.

To find the critical value S^*, we set the value of the American call equal to the value of $C_A(S^*, T)$ as denoted by (5.12). Therefore,

$$S^* - K = C(S^*, T) + ha_2 S^{*\lambda_2}. \tag{5.13}$$

Since the value of European option is

$$C(S^*, T) = S^* N[d_1(S^*)] - KN[d_2(S^*)]$$

by the Black-Scholes formula, we can easily see that

$$\frac{\partial C(S^*, T)}{\partial S^*} = e^{(b-r)T} N[d_1(S^*)], \tag{5.14}$$

where

$$d_1(S^*) = \frac{\ln\left(\frac{S^*}{K}\right) + \left(b + \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}.$$

Now, by considering the derivative of both sides of (5.13) with respect to S^*, we get

$$e^{(b-r)T} N[d_1(S^*)] + h\lambda_2 a_2 S^{*\lambda_2 - 1} = 1. \tag{5.15}$$

From (5.15), we have

$$a_2 = \frac{1 - e^{(b-r)T} N[d_1(S^*)]}{h\lambda_2 S^{*\lambda_2 - 1}}. \tag{5.16}$$

we have two equations, (5.13) and (5.15), two unknowns, a_2 and S^* from equation (5.15). Writing (5.16) into (5.13), we get

$$\begin{aligned} S^* - K &= C(S^*, T) + ha_2 S^{*\lambda_2} \\ &= C(S^*, T) + \frac{(1 - e^{(b-r)T} N[d_1(S^*)])S^*}{\lambda_2}. \end{aligned} \tag{5.17}$$

In (5.17), the only unknown is S^*, so this is the classical problem of finding a root of the equation

$$f(S^*) := S^* - K - C(S^*, T) - \frac{(1 - e^{(b-r)T} N[d_1(S^*)])S^*}{\lambda_2} = 0.$$

This can be solved using Newton-Raphson method given in Appendix A.1.4, for finding the root.

We start with a first seed value S_0 suggested in [3]. At each step we need to evaluate $f'(S^*)$:

$$f'(S^*) = 1 - \frac{\partial C}{\partial S^*} + \frac{e^{(b-r)T}\frac{\partial N[d_1(S^*)]}{\partial S^*}S^* - (1 - e^{(b-r)T}N[d_1(S^*)])}{\lambda_2},$$

and

$$\frac{\partial C}{\partial S^*} = e^{(b-r)T}N(d_1) + S^* e^{(b-r)T}\frac{\partial N(d_1)}{\partial S^*} - Ke^{-rT}\frac{\partial N(d_2)}{\partial S^*},$$

$$\frac{\partial N(d_1)}{\partial S^*} = N'(d_1)\frac{\partial d_1}{\partial S^*} = \frac{1}{\sqrt{2\pi}}e^{-\frac{d_1{}^2}{2}}\frac{1}{S^*}\frac{1}{\sigma\sqrt{T}},$$

$$\frac{\partial N(d_2)}{\partial S^*} = N'(d_2)\frac{\partial d_2}{\partial S^*} = \frac{1}{\sqrt{2\pi}}e^{-\frac{d_2{}^2}{2}}\frac{1}{S^*}\frac{1}{\sigma\sqrt{T}}.$$

Putting all these values into Newton-Raphson method, the value S^* can be found. With known S^*, we can find the value of a_2 from (5.16). Lastly, substituting (5.16) into (5.12) and making some simplifications, we obtain that for an American call option the approximation is

$$C_A(S,T) = C(S,T) + A_2\left(\frac{S}{S^*}\right)^{\lambda_2} \quad \text{if} \quad S < S^*,$$

$$C_A(S,T) = S - K \quad \text{if} \quad S \geq S^* \tag{5.18}$$

where

$$A_2 = \frac{(1 - e^{(b-r)T}N[d_1(S^*)])S^*}{\lambda_2}.$$

Since S^*, λ_2 and $1 - e^{(b-r)T}N[d_1(S^*)]$ are positive when $b < r$, then $A_2 > 0$. Therefore, the analytic approximation is efficient for the value of an American call option when the cost of carry is less than the riskless rate of interest.

In equation (5.18), the early exercise premium of the American call option goes to 0 as the time to expiration of the option goes to 0. However, as T gets small, $N[d_1(S^*)]$ goes to 1, $1 - e^{(b-r)T}N[d_1(S^*)]$ goes to 0, A_2 goes to 0, and so that $A_2\left(\frac{S}{S^*}\right)^{\lambda_2}$ goes to 0.

Example 9. Consider an American call option with a stock price \$90, and a volatility 15%. The strike price is \$100 and the risk free rate is 10%, and annual yield from the stock is 10%. The option has 1.2 months to expiry.

Applying the Quadratic Approximation, the price of the option is obtained as \$0.0382. To implement the method, we use MATLAB code found in Appendix B.

5.2.2 Quadratic Approximation for an American Put Option

Since the modified Black-Scholes partial differential equation

$$\frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + (r-q)S\frac{\partial V}{\partial S} - rV + V_t = 0$$

applies to the early exercise premium for an American put option

$$w(S,T) = P_A(S,T) - P(S,T),$$

the equations (5.6)–(5.10) are all the same for the American put.

Following from (5.11), the approximate value of the American put option becomes

$$P_A(S,T) = P(S,T) + ha_1 S^{\lambda_1}$$

since the term $a_1 S^{\lambda_1}$ is of interest when the early exercise premium of the American put should go to 0 as S goes to $+\infty$.

Similarly, the coefficient a_1 and the critical price S^{**} should be determined in the same way as it was done before for the American call.

After making the same steps for the American call, we get

$$a_1 = \frac{-(1 - e^{(b-r)T} N[-d_1(S^{**})])}{h\lambda_1 {S^{**}}^{\lambda_1 - 1}},$$

where

$$-e^{(b-r)T} N[-d_1(S^{**})] = \frac{\partial P(S^{**},T)}{\partial S^{**}}.$$

Also, $a_1 > 0$ since $\lambda < 0$, and the all other terms are positive.

Therefore, the critical price is found by solving the equation

$$K - S^{**} = P(S^{**},T) - \frac{(1 - e^{(b-r)T} N[-d_1(S^{**})])S^{**}}{\lambda_1}$$

by iteration again. With known S^{**}, the approximation for an American put can be written as

$$P_A(S,T) = P(S,T) + A_1\left(\frac{S}{S^{**}}\right)^{\lambda_1} \quad \text{if} \quad S > S^{**},$$

$$P_A(S,T) = K - S \quad \text{if} \quad S \le S^{**},$$

Table 5.2: The comparison of Barone-Adesi Whaley (BAW) with the Binomial and finite differences methods for an American put option

Stock Price (S_0)	BAW	Binomial	Finite Differences
70	20.6553	20.6756	20.5808
75	16.1234	16.1720	16.0812
80	12.0242	12.0822	11.9900
85	8.5264	8.5811	8.5105
90	5.7357	5.7817	5.7251
95	3.6600	3.6965	3.6527
100	2.2191	2.2427	2.2140
105	1.2820	1.2784	1.2978
110	0.7082	0.7172	0.7057
115	0.3755	0.3810	0.3738

where

$$A_1 = \frac{-(1 - e^{(b-r)T} N[-d_1(S^{**})])S^{**}}{\lambda_1}.$$

Here, $A_1 > 0$ since $\lambda_1 < 0$ and $S^{**} > 0$, also $N[-d_1(S^{**})] < e^{-bT}$.

Example 10. Consider an American call option on stock with parameters $K = 90, r = 0.06, T = 0.25, \sigma = 0.3, q = 0.1, b = -0.04$. We now compare the Barone-Adesi Whaley method with the Binomial and finite differences methods for these parameters as given in Table 5.2.

MATLAB implementation of the Barone-Adesi Whaley method for an American put option can be found in Appendix B.

5.3 Bjerksund and Stensland Approximation

The Bjerksund and Stensland Approximation [7, 8] is a useful method to compute the values of the American options. They obtain an accurate and computer efficient approximation to the value of an American option by imposing feasible but non-optimal exercise strategy. They assume a flat early exercise boundary.

5.3.1 Assumptions of the model

First, we consider a Black-Scholes model with a risk free rate r on an underlying asset price which is a geometric Brownian motion with respect to the equivalent martingale measure. So, we define the underlying asset price S_t at time t as

$$S_t = Se^{(b-\frac{1}{2}\sigma^2)t+\sigma W_t},$$

where S is the stock price, $b = r - q$, q is the dividend yield which is called *cost of carry*, is the drift rate with respect to the equivalent martingale measure, σ is volatility, and W_t is a Brownian motion.

Secondly, we consider an American call option with maturity time T and strike price K. For a given feasible exercise strategy which is represented by a stopping time $\tau \in [0, T]$, the option value can be written according to the this strategy as

$$C_A = \mathbb{E}_{\mathbb{Q}}\left[e^{-r\tau}(S_\tau - K)^+\right]. \tag{5.19}$$

It is known that the value of a contingent claim can be found as the expected discounted payoff where the expectation is taken with respect to the equivalent martingale measure $\mathbb{Q}$ and the risk free rate r is used to discount. Then, the value of the American call option is given by

$$\begin{aligned} C_{A\mathbb{Q}} &\equiv C_{\mathbb{Q}}(S, K, T, r, b, \sigma^2) \\ &= \sup_{\tau\in[0,T]} \mathbb{E}_{\mathbb{Q}}[e^{-r\tau}(S_\tau - K)^+]. \end{aligned}$$

5.3.2 Approximating the American call

Now, define the stopping time as

$$\tau_{\mathbb{Q}}(S^*) \equiv \inf\{\{\inf_{t\in[0,\infty)} S_t > S^*\}, T\} \tag{5.20}$$

where $S^* > K$ is the trigger price, which is defined as early exercise boundary. Then, applying (5.20) to (5.19) and making some rearrangements, we obtain

$$\begin{aligned} \bar{C}_A \equiv& \bar{C}_A(S, K, T, r, b, \sigma^2, S^*) \\ =& \mathbb{E}_{\mathbb{Q}}[e^{-r\tau_{\mathbb{Q}}(S^*)}(S^* - K)\mathbf{1}_{\{\tau_{\mathbb{Q}}(S^*)<T\}}] \\ &+ \mathbb{E}_{\mathbb{Q}}[e^{-rT}(S_T - K)^+\mathbf{1}_{\{\tau_{\mathbb{Q}}(S^*)=T\}}]. \end{aligned} \tag{5.21}$$

In Equation (5.21), the first term represents the value of the early exercise of the option when $S_T > S^*$ and the second term defines the value of the option at the maturity date T.

A strategy following from a flat early exercise boundary is clearly feasible, but not optimal. This means that the value coming from this strategy shows a lower bound to the real option value. By the numerical researches, this lower bound can be thought as an accurate approximation [7, 8, 6].

The following theorem gives an approximation for the American call price under the above exercise strategy. This formula is given in [7].

Theorem 5.3.1 (Bjerksund and Stensland Formula)**.** *Given the stopping rule, the approximation for the value of the American call option is given by*

$$\begin{aligned} \bar{C}_A =& \alpha S^\beta - \alpha\phi(S,T,\beta,S^*,S^*) + \phi(S,T,1,S^*,S^*) \\ & - \phi(S,T,1,K,S^*) - X\phi(S,T,0,S^*,S^*) + K\phi(S,T,0,K,S^*) \end{aligned} \tag{5.22}$$

with

$$\begin{aligned} \alpha &= (S^* - K)(S^*)^{-\beta}, \\ \beta &= \left(\frac{1}{2} - \frac{b}{\sigma^2}\right) + \sqrt{\left(\frac{b}{\sigma^2} - \frac{1}{2}\right)^2 + 2\frac{r}{\sigma^2}}. \end{aligned} \tag{5.23}$$

The function ϕ is given by

$$\phi(S,T,\gamma,H,S^*) = e^\lambda S^\gamma \left[N(d) - \left(\frac{S^*}{S}\right)^k N\left(d - \frac{2\ln(S^*/S)}{\sigma\sqrt{T}}\right)\right] \tag{5.24}$$

where $N(\cdot)$ is the cumulative normal distribution function and

$$\begin{aligned} \lambda &\equiv \left(-r + \gamma b + \frac{1}{2}\gamma(\gamma-1)\sigma^2\right)T, \\ d &\equiv -\frac{\ln(S/H) + (b + (\gamma - \frac{1}{2})\sigma^2)T}{\sigma\sqrt{T}}, \\ k &\equiv \frac{2b}{\sigma^2} + (2\gamma - 1) \end{aligned} \tag{5.25}$$

where the trigger price S^* to be equal to S_T^* which is defined as

$$S_T^* = B_0 + (B_\infty - B_0)(1 - e^{h(T)}),$$

for which

$$
\begin{aligned}
h(T) &\equiv -\left(bT + 2\sigma\sqrt{T}\right)\left(\frac{B_0}{B_\infty - B_0}\right),\\
B_\infty &= \frac{\beta}{\beta-1}X,\\
B_0 &= \max\left\{K, \left(\frac{r}{r-b}\right)K\right\}.
\end{aligned}
$$

For the complete proof of the above approximation formula, see the appendices of [6, 7]. We will only give a sketch of the proof. First a contingent claim

$$
\phi(S,T,\gamma,H,S^*) = \mathbb{E}_{\mathbb{Q}}\left[e^{-rT}S_T^{\gamma}\mathbf{1}_{\{S_t\le H\}}\mathbf{1}_{\{\tau_{\mathbb{Q}}(S^*)=T\}}\right] \tag{5.26}
$$

is defined, where S_t is the risk-adjusted price process given as

$$
S_t = Se^{(r-\frac{1}{2}\sigma^2)t+\sigma W_t}
$$

and $H \le S^*$. Using the fact that for $\gamma > 0$

$$
\mathbf{1}_{\{S_t\le H\}} = \mathbf{1}_{\{S_T^{\gamma}\le H^{\gamma}\}}, \tag{5.27}
$$

it is true that

$$
\mathbf{1}_{\{\tau_{\mathbb{Q}}(S^*)=T\}} = \mathbf{1}_{\{\sup_{\tau\in[0,T]}S_T^{\gamma}\le S^{*\gamma}\}}. \tag{5.28}
$$

Substituting the equations (5.27) and (5.28) into equation (5.26) gives

$$
\begin{aligned}
\phi =&\mathbb{E}_{\mathbb{Q}}[e^{-rT}S_T^{\gamma}\mathbf{1}_{\{S_T^{\gamma}\le H^{\gamma}\}}\mathbf{1}_{\{\sup_{\tau\in[0,T]}S_T^{\gamma}<S^{*\gamma}\}}]\\
=&\mathbb{E}_{\mathbb{Q}}\left[e^{-rT}S^{*\gamma}\left(\frac{S_T}{S^*}\right)^{\gamma}\mathbf{1}_{\left\{\left(\frac{S_T}{S^*}\right)^{\gamma}\le\left(\frac{H}{S^*}\right)^{\gamma}\right\}}\mathbf{1}_{\left\{\sup_{\tau\in[0,T]}\left(\frac{S_\tau}{S^*}\right)^{\gamma}<1\right\}}\right]\\
=&\mathbb{E}_{\mathbb{Q}}\left[e^{-rT}S^{*\gamma}\left(\frac{S_T}{S^*}\right)^{\gamma}\mathbf{1}_{\left\{\gamma\ln\left(\frac{S_T}{S^*}\right)\le\gamma\ln\left(\frac{H}{S^*}\right)\right\}}\mathbf{1}_{\left\{\sup_{\tau\in[0,T]}\gamma\ln\left(\frac{S_\tau}{S^*}\right)<0\right\}}\right]\\
=&\mathbb{E}_{\mathbb{Q}}\left[e^{-rT}S^{*\gamma}e^{Y_T}\mathbf{1}_{\{Y_T\le\hat{Y}\}}\mathbf{1}_{\{\sup_{\tau\in[0,T]}Y_\tau<0\}}\right]
\end{aligned} \tag{5.29}
$$

where

$$
\begin{aligned}
\hat{Y} &\equiv \gamma\ln\left(\frac{H}{S^*}\right), \quad \hat{Y}\le 0,\\
Y_t &\equiv \gamma\ln\left(\frac{S_t}{S^*}\right)\\
&= \gamma\ln\left(\frac{S_0e^{(b-\frac{1}{2}\sigma^2)t+\sigma W_t}}{S^*}\right)\\
&= \gamma\ln\left(\frac{S_0}{S^*}\right)+\gamma\left(b-\frac{1}{2}\sigma^2\right)t+\gamma\sigma W_t\\
&= Y_0+\hat{\mu}t+\hat{\sigma}W_t
\end{aligned}
$$

and

$$\hat{\mu} = \gamma\left(b - \frac{1}{2}\sigma^2\right),$$
$$\hat{\sigma} = \gamma\sigma.$$

Now, to evaluate ϕ, the probability density function as follows:

$$\begin{aligned}
f_0 \equiv f_0 &\left(\{Y_T\} \quad , \quad \{ \sup_{\tau\in[0,T]} Y_\tau < 0\} \right) \\
= f_0 &\left(\{-Y_T\} \quad , \quad \{ \inf_{\tau\in[0,T]} -Y_\tau > 0\} \right) \\
= n &\left[\frac{Y_T - Y_0 - \hat{\mu}T}{\hat{\sigma}\sqrt{T}} \right] - e^{\frac{-2\hat{\mu}Y_0}{\hat{\sigma}^2}} n \left[\frac{Y_T + Y_0 - \hat{\mu}T}{\hat{\sigma}\sqrt{T}} \right] \qquad (5.30)
\end{aligned}$$

is defined where $n(\cdot)$ is the normal density function, which follows from Ingersoll [27]. Rearranging the equation (5.29) and using the probability density function from (5.30),

$$\begin{aligned}
\phi =& e^{-rT} S^{*\gamma} \int_{-\infty}^{\hat{Y}} e^{Y_T} n\left(\frac{Y_T - Y_0 - \hat{\mu}T}{\hat{\sigma}\sqrt{T}} \right) dY_T \\
&- e^{-rT} S^{*\gamma} e^{\frac{-2\hat{\mu}Y_0}{\hat{\sigma}^2}} \int_{-\infty}^{\hat{Y}} e^{Y_T} n\left(\frac{Y_T + Y_0 - \hat{\mu}T}{\hat{\sigma}\sqrt{T}} \right) dY_T \\
=& S^{*\gamma} e^{Y_0} e^{(-r+\hat{\mu}+\frac{1}{2}\hat{\sigma}^2)T} N\left(\frac{\hat{Y} - Y_0 - (\hat{\mu}+\hat{\sigma}^2)T}{\hat{\sigma}\sqrt{T}} \right) \\
&- S^{*\gamma} e^{Y_0} e^{(-r+\hat{\mu}+\frac{1}{2}\hat{\sigma}^2)T} e^{-Y_0\left(\frac{2\hat{\mu}}{\hat{\sigma}^2}+1\right)} N\left(\frac{\hat{Y} + Y_0 - (\hat{\mu}+\hat{\sigma}^2)T}{\hat{\sigma}\sqrt{T}} \right)
\end{aligned}$$

is obtained, where the relation

$$\int_{-\infty}^{\hat{Z}} e^{Z} n\left(\frac{Z - \mu_Z}{\sigma_Z} \right) dZ = e^{\left(\mu_Z + \frac{1}{2}\sigma_Z{}^2\right)} N\left(\frac{\hat{Z} - (\mu_Z + \sigma_Z{}^2)}{\sigma_Z} \right)$$

is used. Inserting the definitions of $Y_0, \hat{\mu}, \hat{\sigma}$ and $\hat{Y}$ and making some calculations, the equations (5.24)–(5.25) are obtained.

Remark 5.3.2. The approximation for S^* is similar to Equation (31) given by [3]. The first term of (5.22) represents the value of an infinite-lived American call option with exercise price K and the trigger price S^*. The other terms of (5.22) are given by the function ϕ. The second term of (5.22) represents an obligation to return the option at date T if it is still alive at that date. When $\gamma = 0$ or $\gamma = 1$

in (5.26), the function returns the conditional claims on the riskless asset and the underlying asset. Thus, the remaining terms of the equation (5.22) represent the value of a European call option with time to maturity T and exercise price K with condition $S_t < S^*$ for all $t \in [0, T)$.

Here, B_T is defined as the optimal exercise price for an American call option with exercise price K and time to maturity T. This approximation is similar to Quadratic Approximation (Barone-Adesi Whaley Formula) except the term B_0.

The value of the American put option is, therefore, given by

$$\begin{aligned} P_{\mathbb{Q}} &\equiv P_{\mathbb{Q}}(S, K, T, r, b, \sigma^2) \\ &= \sup_{\tau \in [0,T]} \mathbb{E}_{\mathbb{Q}}[e^{-r\tau}(K - S_\tau)^+], \end{aligned}$$

that is, the optimal solution to a stopping problem.

From the following put-call transformation in Bjerksund and Stensland [6],

$$P_{\mathbb{Q}}(S, K, T, r, b, \sigma^2) = C_{\mathbb{Q}}(S, K, T, r - b, -b, \sigma^2),$$

where $C_{\mathbb{Q}}$ is the value of the American call option, we approximate the American option by

$$P(S, K, T, r, b, \sigma^2) = C(S, K, T, r - b, -b, \sigma^2; X),$$

where C is given by equations (5.22)–(5.23) and K is the exercise boundary.

Table 5.3 shows a simple comparison of the method with others in pricing American options.

The MATLAB code of the Bjerksund and Stensland, and Barone-Adesi Whaley methods can be found in Appendix B.

5.4 Least Squares Monte Carlo

This section introduces Least-Squares Monte Carlo Method (LSMC) which was introduced by Longstaff and Schwartz [36]. They present a new approach to approximate the values of American options by a path-wise approximation of

Table 5.3: The comparison of Barone-Adesi Whaley (BAW), Bjerksund and Stensland (BS), Binomial and Finite Differences methods for an American call option with parameters $K = 80, r = 0.06, T = 0.25, \sigma = 0.4, q = 0.1, b = -0.04$.

Stock Price (S_0)	BAW	BS	Binomial	Finite Differences
60	0.4149	0.4078	0.4097	0.4101
65	1.0041	0.9922	0.9980	0.9971
70	2.0542	2.0367	2.0492	2.0475
75	3.6802	3.6586	3.6787	3.6780
80	5.9361	5.9157	5.9456	5.9447
85	8.8107	8.8007	8.8394	8.8378
90	12.2433	12.2546	12.2985	12.2965
95	16.1479	16.1872	16.2342	16.2324
100	20.4354	20.4956	20.5524	20.5511
105	25.0297	25.0776	25.1679	25.1668

the optimal exercise rule. The key to this approach is to compare the immediate exercise value at any exercise time with the expected payoff from continuation. Therefore, the conditional expected value of continuation should be calculated at any exercise time. They use the cross-sectional information to estimate the conditional expectation function for in-the-money paths. Then, the option can be valued by discounting each cash flow in all paths to time zero, and averaging over all paths.

Reviewing the recent articles which address the pricing of American options by simulation, Bossaerts [10], Tilley [44], Barraquand and Martineau [4], Averbukh [2], Carr [14], Ibanez and Zapatero [26], and Garcia [21] are the important examples as a contribution to the Least-Squares Monte Carlo literature. These articles are fundamentally different approaches by focusing directly on the conditional expectation function. The articles Carriere [15] and Tsitsiklis and Van Roy [45] are similar approaches to LSM algorithm.

In the sequel, we describe the LSM approach for the valuation of American options.

First, we review the LSM approach discussed in Longstaff and Schwartz [36]. The algorithm is one of the most popular methods, in particular, for pricing American options on more than one underlying asset.

Assume an underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and finite time interval $[0, T]$. We define a filtration generated by the price processes of the underlying securities $F = \{\mathcal{F}_t; t \in [0, T]\}$. By the no-arbitrage principle, we assume the existence of an equivalent martingale measure $\mathbb{Q}$ for this model. The method is suitable for options with payoffs which are the elements of the space of square-integrable functions of $\mathcal{L}^2$. By a standard result from Bensoussan [5] and Karatzas [31], the value of an American option can be represented by the *Snell-envelope* (26) which means that the value of an American option equals the maximum value of the discounted payoffs of the option, where the maximum is taken over all stopping times with respect to the filtration F.

Let $CF(\omega, s; t, T)$ denote the different ω paths of cash flows generated by the option, conditional of not having exercised the option before or at time t and on following the optimal exercise strategy $s, t < s \leq T$. We assume that the American option can only be exercised at the N discrete times $0 = t_0 \leq t_1 \leq t_2 \leq \ldots \leq t_N = T$, and consider the optimal stopping strategy at each exercise date.

At the final exercise date T of the option, the optionholder exercises the option depending on it being in the money or not. However, at any prior time t_k, with the expected value of continuation, the holder must decide whether to exercise immediately or to continue the life of the option. The option value is maximized pathwise if the holder exercises as soon as the immediate exercise value is greater than or equal to the continuation value. By no-arbitrage principle, the continuation value is given by the expectation of all future discounted payoffs $CF(\omega, s; t_k, T)$ with respect to the risk neutral pricing measure $\mathbb{Q}$. The value of continuation at time t_k is given by

$$F(\omega; t_k) = \mathbb{E}_{\mathbb{Q}}\left[\sum_{j=k+1}^{N} \exp\left(-\int_{t_k}^{t_j} r(\omega, s)ds\right) CF(\omega, t_j; t_k, T)|\mathcal{F}_{t_k}\right], \qquad (5.31)$$

where $r(\omega, t)$ is the risk-free interest rate, and the expectation is taken conditionally on the set of information, $\mathcal{F}_{t_k}$, obtained until time t_k with respect to the risk-neutral measure. Then, the price of the American option is found by averaging $F(\omega, 0)$ over all ω paths.

The LSM approach uses least squares regression to approximate the conditional expectation function at every exercise times $t_{N-1}, t_{N-2}, \ldots, t_1$. It uses a backward induction method since the path of payoffs $CF(\omega, s; t, T)$ is defined recursively. The unknown functional form of the equation (5.31) can be approximated by a linear combination of a countable set of $F_{t_{k-1}}$-measurable basis functions at time t_{k-1}.

The conditional expectation should be an element of the Hilbert space $\mathcal{L}^2$ to justify the above assumption. Then, it has a countable orthonormal basis and the conditional expectation can be expressed by a linear function of the elements of the basis.

Longstaff and Schwartz [36] suggested to use as basis functions the weighted Laguerre polynomials:

$$\begin{aligned} L_0(X) &= \exp(-X/2), \\ L_1(X) &= \exp(-X/2)(1 - X), \\ L_2(X) &= \exp(-X/2)(1 - 2X + X^2/2), \\ L_n(X) &= \exp(-X/2)\frac{e^X}{n!}\frac{d^n}{dX^n}(X^n e^{-X}). \end{aligned}$$

With this choice, $F_M(\omega; t_{k-1})$ can be approximated by a set of $M < \infty$ basis functions;

$$F_M(\omega; t_{k-1}) \approx \sum_{j=0}^{M} a_j L_j(X_{t_{k-1}})$$

where the a_j coefficients are constants. The other types of basis functions include the Hermite, Legendre, Chebyshev, Gegenbauer, and Jacobi polynomials. Also, simple set of polynomials gives accurate solutions.

After this subset of basis functions has been specified , $F_M(\omega; t_{k-1})$ is estimated by regressing the discounted payoffs of $CF(\omega, s; t_{k-1}, T)$ on the set of basis functions where the option is in-the-money at time t_{k-1}. The LSM algorithm only uses in-the-money paths for the regression since for out-of-the money paths it is never optimal to exercise the option. Therefore, this limits the region over which the conditional expectation function must be estimated yielding more accurate approximation with far fewer number of basis functions. Since the values of basis functions are identically and independently distributed across all paths, White [48] shows that the fitted value of the regression, a particular value that fits the line of best fit, $\hat{F}_M(\omega; t_k)$ converges in mean square and in probability to $F_M(\omega; t_k)$ as the number of in-the-money paths goes to infinity. Also, Amemiya [1] states that $\hat{F}_M(\omega; t_k)$ is the best linear unbiased estimator of $F_M(\omega; t_k)$ based on the mean-squared metric.

The objective of the algorithm is to approximate the optimal path–wise stopping rule that maximizes the value of the American option.

Once the conditional expectation function is estimated at time t_{N-1}, the exercise decision is determined by comparing the immediate exercise value of all of the in-the-money paths with the expected value of continuation. Now, the payoffs for t_{N-1} can be approximated. After discounting the payoffs, these values can be regressed on a set of basis functions of state variables of time t_{N-2}. In this way, we obtain an accurate estimation of the continuation function at time t_{N-2}, and repeat this procedure until the stopping rule for every exercise times over all paths are determined. Then, the value of the American option is calculated by starting at time zero, moving forward along each path until the exercise time occurs, discounting the indicated payoff to time zero, and taking the average over all paths ω.

The algorithm can be summarized as follows:

- Given S_0, K, r, σ, T, M.
- Generate stock prices S_{t_i} for each M paths and times $0 = t_0 \leq t_1 \leq t_2 \ldots t_N = T$.
- Find the payoffs $P = \max\{K - S_j\}$, $j = 1, \ldots, N$ and take only in-the-money paths.
- Find the discounted payoffs $e^{-rt} CF(\omega; t_k)$ for in-the-money paths.
- Choose a set of basis functions.

Table 5.4: The comparison of the Least Squares Monte Carlo (LSMC), Binomial and Finite Differences methods with parameters $K = 80, r = 0.08, T = 0.25, \sigma = 0.4$.

Stock Price (S_0)	LSMC	Binomial	Finite Differences
60	20.0392	20.0047	20.0039
65	15.3810	15.3856	15.3839
70	11.3670	11.4352	11.4326
75	8.0361	8.2050	8.2030
80	5.4013	5.6878	5.6866
85	3.5616	3.8182	3.8155
90	2.2999	2.4864	2.4839
95	1.3484	1.5755	1.5732
100	0.7917	0.9737	0.9721
105	0.4415	0.5889	0.5878

- Calculate $\hat{F}_M(\omega; t_k)$ by regressing the discounted payoffs and determine the continuation value.
- If the continuation value $F(\omega; t_k)$ is greater than exercise value, hold the option.
- Repeat this procedure for each paths until the decision rule is made for every time $t_1, \ldots, t_N$.
- Find the option value $P_A = \frac{1}{M} \sum_{j=1}^{M} \hat{F}_M(\omega; t_k)$.

Remark 5.4.1. For the convergency of the Least-Squares Monte Carlo method, Longstaff and Schwartz [36] gives an objective criteria such as

$$V(x) \geq \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} LSM(\omega_i; M, N)$$

where $V(x)$ is the value of the American option and ω_i is the ith path. As a result of this criteria, the number of basis functions can be determined. Also, [36] states that the algorithm converges to any desired accuracy.

Table 5.4 presents a comparison of the Least Squares Monte Carlo with other methods, previously investigated.

Now, we will study on a numerical example to understand the algorithm of the approach in a best way.

Example 11. Consider an American put option on a non-dividend paying stock with a stock price $S_0 = 2$ and strike price $K = 2.1$. The option is exercisable at time $1, 2$, and 3, where time 3 is the final expiration date of the option and the risk free rate is 6%. The other parameters are $\sigma = 0.15$ and $T = 1$.

We use simple set of polynomials to make linear regression and take only ten simple paths which are obtained by the geometric Brownian motion for the price of the stock and the paths are shown in the following table.

Stock price paths

Path	$t = 0$	$t = 1$	$t = 2$	$t = 3$
1	2	2.21	2.35	2.71
2	2	1.92	2.21	1.97
3	2	1.84	1.89	1.66
4	2	1.91	1.93	1.87
5	2	2.27	2.25	2.11
6	2	1.88	2.08	2.30
7	2	2.13	2.41	2.29
8	2	2.05	1.84	1.85
9	2	2.14	2.15	2.01
10	2	2.09	1.70	1.77

Now, we start with computing the option's payoff at time 3, which is $\max\{K - S_3, 0\}$, as in the following table.

Payoff vector at time 3

Path	$t = 3$
1	0
2	0.13
3	0.44
4	0.23
5	0
6	0
7	0
8	0.25
9	0.09
10	0.33

If the put option is in the money, that is the strike price is greater than the stock price, at time 2, then the option holder should decide whether it is optimal to exercise the option or not.

The next table shows the in-the-money paths at time 2.

Paths in-the-money at time 2

Path	$t = 2$
1	0
2	0
3	0.21
4	0.17
5	0
6	0.02
7	0
8	0.26
9	0
10	0.4

Only the paths $3, 4, 6, 8$, and 10 are in-the-money at time 2, so the holder should decide whether it is optimal to early exercise for these paths or not. We use a least square regression to estimate the conditional expectation function which gives the continuation values.

Let X be a vector of stock prices at time 2 for these 5 paths and Y be the vector of corresponding discounted payoffs at time 3 as shown in the following table.

Regression at time 2

Path	$Y(\max\{K-S_3\}e^{-0.06})$	X
1	-	-
2	-	-
3	0.44×0.94176	1.89
4	0.23×0.94176	1.93
5	-	-
6	0	2.08
7	-	-
8	0.25×0.94176	1.84
9	-	-
10	0.33×0.94176	1.70

To decide whether or not it is optimal to early exercise for these paths, we regress Y onto $1, X$ and X^2.

It is found that the regression coefficients are $a_0 = -12.361, a_1 = 14.171$ and $a_2 = -3.955$. Then, the conditional expectation function is $\mathbb{E}[Y|X] = -12.361 + 14.171X - 3.955X^2$.

The next table shows the exercise values, the profit that an optionholder would get by exercising an in-the-money option, and continuation values, the value that the optionholder does not exercise the option will receive, at time 2.

Path	Exercise	Continuation
1	-	-
2	-	-
3	0.21	0.2945
4	0.17	0.2570
5	-	-
6	0.02	0.0037
7	-	-
8	0.26	0.3235
9	-	-
10	0.4	0.2997

For the paths 6, and 10, since the continuation values are less than exercise values, it is optimal to exercise the option at time 2 for these paths.

The next table illustrates the corresponding payoffs.

Payoff matrix at time 2

Path	$t = 2$	$t = 3$
1	0	0
2	0	0.13
3	0	0.44
4	0	0.23
5	0	0
6	0.02	0
7	0	0
8	0	0.25
9	0	0.09
10	0.4	0

Now, we decide whether the option is exercised at time 1.

Paths in-the-money at time 1

Path	$t = 1$
1	0
2	0.18
3	0.26
4	0.19
5	0
6	0.22
7	0
8	0.05
9	0
10	0.01

It can be easily seen that, there are six paths where the option is in-the-money at time 1.

Now, we again use a regression to estimate the conditional expectation function. Let X be a vector of stock prices at time 1 for these five paths and Y be the vector of corresponding discounted payoffs at time 2.

Regression at time 1

Path	$Y(\max\{K - S_3\}e^{-0.06})$	X
1	-	-
2	0	1.92
3	0	1.84
4	0	1.91
5	-	-
6	0.02×0.94176	1.88
7	-	-
8	0	2.05
9	-	-
10	0.4×0.94176	2.09

We regress Y onto $1, X$ and X^2 and the regression coefficients are found as $a_0 = 44.8953, a_1 = -46.6325$, and $a_2 = 12.0988$.

The following table shows the exercise values and continuation values at time 1.

Path	Exercise	Continuation
1	-	-
2	0.18	0
3	0.26	0.0531
4	0.19	0
5	-	-
6	0.22	0
7	-	-
8	0.05	0.1438
9	-	-
10	0.01	0.2821

It is seen that it is optimal to exercise paths $2, 3, 4$, and 6 at time 1.

Now, the next table shows the stopping rule for the option where ones denote the exercise dates.

Stopping rule

Path	$t = 1$	$t = 2$	$t = 3$
1	0	0	0
2	1	0	0
3	1	0	0
4	1	0	0
5	0	0	0
6	1	0	0
7	0	0	0
8	0	0	0
9	0	0	1
10	0	1	0

This stopping rule table leads to the following payoff matrix.

Path	$t = 1$	$t = 2$	$t = 3$
1	0	0	0
2	0.18	0	0
3	0.26	0	0
4	0.19	0	0
5	0	0	0
6	0.22	0	0
7	0	0	0
8	0	0	0
9	0	0	0.09
10	0	1.4	0

Finally, the option can be valued by discounting each payoff to time zero, and taking the average over all paths. Therefore, the value for the American put option can be found as in the following:

$$\begin{aligned}\text{Option value} &\approx \frac{1}{10}[e^{-0.06}(0.18 + 0.26 + 0.19 + 0.22) + e^{-2\times 0.06}1.4 + e^{-3\times 0.06}0.09] \\ &= 0.2117.\end{aligned}$$

CHAPTER 6

CONCLUSION

In this thesis, we analyzed the valuation of American options by computational methods. We first investigated the binomial model which assumes the stock price is following geometric Brownian motion and gives possibilities of underlying asset price at each time. Comparing European and American call options on non-dividend paying assets by simulation, we noticed that these two types of options give the same value on the same contract conditions. For the case of put option, we are verified by the test results that the American option value is higher than the European option value. For the dividend paying underlying asset, we observed the same theoretical information.

Next, we considered the Black-Scholes model for pricing American options by replacing partial differentials by difference quotients in this model. We made the implementations on different scenarious for each finite difference method. For the explicit difference method, we compared the results coming from this method and closed-form solution of the Black-Scholes partial differential equation and we were satisfied that the results are very close by a setting of higher values of ΔS for given Δt. When we test explicit method with the θ-averaged method, the solution curve for each method fits properly.

The finite difference method gives a higher value of an American call option with dividend payment asset rather than European type, and this is also current for put option. We performed the Crank-Nicolson method under LU-decomposition and projected SOR (PSOR) and we detected that there were slightly larger errors around the strike price. Implicit and Crank-Nicolson methods were also investigated between themselves by a simulation on the theoretical concepts. As a result of this test, we noted that the Crank-Nicolson method does not give very proper results around the strike price because of the non-differentiable final condition. Then, we contrasted these two methods for a put option and confirmed by the result.

The American option valuation problem does not have a closed-form solution. Therefore, we then studied some approximation methods to approach the exact value of American options. We analyzed the Roll-Geske-Whaley approximation with Binomial method on a single discrete dividend paying asset for different stock prices and we saw that when the stock price is getting higher, the error between these two methods is getting smaller. Also, the quadratic approximation gives satisfactory results by comparing with binomial and finite difference methods. Then, we discussed the quadratic approximation with another approximation method which is Bjerksund and Stensland approximation with the numerical experiments and observed that they give the results that we already expected, also near to the strike price. The Least-Squares Monte Carlo method is an easy and accurate technique which needs a regression to evaluate the continuation value. As a simulation experiment, Least-Squares Monte Carlo method also gives fitting values , but comparing this method with the other approximation methods, the error occuring with this method is higher.

As a continuation of this work, it would be good to analyze different numerical methods and approximation methods existing in the literature for the valuation of American options for real data values. Also, for the case of more than one discrete dividend payment case can be handled through a recent study [34] about generalization of Roll-Geske-Whaley approximation and comparison of their results with the other methods.

REFERENCES

[1] T. Amemiya. *Advanced Econometrics*. Basil Blackwell, London, UK, 1985.

[2] V. Averbukh. Pricing american options using monte carlo simulation. *Ph.D. dissertation, Cornell University*, 1997.

[3] G. Barone-Adesi and R. E. Whaley. Efficient analytic approximation of american options values. *The Journal of Finance*, 42(2):301–320, 1987.

[4] J. Barraquand and D. Martineau. Numerical valuation of high dimensional multivariate american securities. *Journal of Financial and Quantitative Analysis*, 30:383–405, 1995.

[5] A. Bensoussan. On the theory of option pricing. *Acta Applicandae Mathematicae*, 2(2):139–158, 1984.

[6] P. Bjerksund and G. Stensland. American exchange options and a put-call transformation. *Working paper no. 4/92, Institute of Finance and Management Science, Norwegian School of Economics and Business Administration*, 1992.

[7] P. Bjerksund and G. Stensland. Closed-form approximation of american options. *Scandinavian Journal of Management*, 9:87–99, 1993.

[8] P. Bjerksund and G. Stensland. Closed-form valuation of american options. *Working Paper NHH*, 2002.

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

[10] P. Bossaerts. Simulation estimators of optimal early exercise. *Working paper, Carnegie-Mellon University*, 1989.

[11] P. P. Boyle. A lattice framework for option pricing with two state variables. *Journal of Financial and Quantitative Analysis*, 23:1–12, 1988.

[12] M. Brennan and E. Schwartz. Finite difference methods and jump processes arising in the pricing of contingent claims: A synthesis. *Journal of Financial and Quantitative Analysis*, 13(3):461–474, 1978.

[13] R. L. Burden and J. D. Faires. *Numerical Analysis*. Thomson Brooks/Cole, 8th edition edition, 2005.

[14] P. Carr. Randomization and the american put. *Review of Financial Studies*, 11:597–626, 1998.

[15] J. Carriere. Valuation of early-exercise price of options using simulations and nonparametric regression. *Insurance: Mathematics and Economics*, 19:19–30, 1996.

[16] G. Courtadon. A more accurate finite difference approximation for the valuation of options. *Journal of Financial and Quantitative Analysis*, 17:697–705, 1982.

[17] J. Crank and P. Nicolson. A practical method for numerical evaluation of solutions of partial differential equations of the heat conduction type. *Proceedings of the Cambridge Philosophical Society*, 43(1):50–67, 1947.

[18] C. W. Cryer. The solution of a quadratic programming problem using systematic overrelaxation. *SIAM J. Control and Optimization*, 9(3):385–392, 1971.

[19] M. H. DeGroot and M. J. Schervish. *Probability and Statistics*. Pearson, 2011.

[20] R. England. The use of numerical methods in solving pricing problems for exotic financial derivatives with a stochastic volatility. *M.Sc. Thesis, University of Reading*, 2006.

[21] D. Garcia. A monte carlo method for pricing american options. *Working paper, University of California, Berkeley*, 1999.

[22] R. Geske. The valuation of compound options. *Journal of Financial Economics*, 7:63–81, 1979.

[23] R. Geske. Comments on whaley's note. *Journal of Financial Economics*, 9(2):213–215, 1981.

[24] J. Hull. *Options, Futures, and Other Derivatives*. Pearson Prentice Hall, 7th edition, 2009.

[25] J. Hull and A. White. Valuing derivative securities using the explicit finite difference method. *Journal of Financial and Quantitative Analysis*, 25(1):87–100, 1990.

[26] A. Ibanez and F. Zapatero. Monte carlo valuation of american options through computation of the optimal exercise frontier. *Working paper, University of Southern California*, 1998.

[27] J. E. Ingersoll Jr. *Theory of Financial Decision Making*. Rowman and Littlefield Publishers, 1987.

[28] S. A. Ross J. C. Cox and M. Rubinstein. On the valuation of american call options on stocks with known dividens. *Journal of Financial Economics*, 7:229–264, 1979.

[29] R. A. Jarrow and A. T. Rudd. Option pricing. *Journal of Banking and Finance*, 10:157–161, 1983.

[30] W. Kahan. *Gauss-Seidel Methods of Solving Large Systems of Linear Equations*. Ph.D. thesis, 1958.

[31] I. Karatzas. On the pricing of american options. *Applied Mathematics and Optimization*, 17:37–60, 1988.

[32] S. Kozpinar. Pricing american options under discrete and continuous time setting. *M.Sc. Thesis, Middle East Technical University*, 2013.

[33] H. Kroner. Diploma thesis. *TY Kaiserslautern*, 2004.

[34] S. Kruse and M. Muller. Pricing american call options under the assumption of stochastic dividends - an application of the korn-rogers-model. *Report, Fraunhofer (ITWM)*, 2009.

[35] D. Lamberton and B. Lapeyre. *Introduction to Stochastic Calculus Applied to Finance*. Chapman and Hall CRC, 1996.

[36] F. A. Longstaff and E. S. Schwartz. Valuing american options by simulation: A simple least-squares approach. *The Society for Financial Studies*, 14(1):113–147, 2001.

[37] K. W. Morton and D. F. Mayers. *Numerical Solutions of Partial Differential Equations*. Cambridge University Press, 1995.

[38] S. Howison P. Willmott and D. Dewynne. *The Mathematics of Financial Derivatives: A Student Introduction*. Cambridge University Press,, 1995.

[39] R. Roll. An analytic valuation formula for unprotected american call options on stocks with known dividends. *Journal of Financial Economics*, 5:251–258, 1977.

[40] E. S. Schwartz. The valuation of warrants: Implementing a new approach. *Journal of Financial Economics*, 13:79–93, 1977.

[41] R. U. Seydel. *Tools for Computational Finance*. Springer, 2002.

[42] S. E. Shreve. *Stochastic Calculus for Finance II: Continuous-Time Models*. Springer Finance, 2004.

[43] J. Thomas. *Numerical Partial Differential Equations: Finite Difference Methods*. Springer Verlag, 1995.

[44] J. A. Tilley. Valuing american options in a path simulation model. *Transactions of the society of the actuaries*, 45:83–104, 1993.

[45] J. Tsitsiklis and B. Van Roy. Optimal stopping of markov processes: hilbert space theory, approximation algorithms, and an application to pricing high-dimensional financial derivatives. *IEEE Transactions on Automatic Control*, 44:1840–1851, 1999.

[46] Ö. Uğur. *An Introduction to Computational Finance*, volume 1. Imperial College Press, 2009.

[47] R. E. Whaley. On the valuation of american call options on stocks with known dividens. *Journal of Financial Economics*, 9:207–211, 1981.

[48] H. White. *Asymptotic Theory for Econometricians*. Academic Press, New York, 1984.

Appendix A

MATHEMATICAL PRELIMINARIES

In this part we refer to Shreve [42], Lamberton and Lapeyre [35], Seydel [41] and DeGroot [19] to give some necessary definitions.

A.1 Definitions and Theorems

Definition 7. Let Ω be a nonempty set, and let $\mathcal{F}$ be a collection of subsets of Ω. We say that $\mathcal{F}$ is σ-algebra provided that:

(i) the empty set $\varnothing$ belongs to $\mathcal{F}$,

(ii) whenever a set A belongs to $\mathcal{F}$, its complement A^c also belongs to $\mathcal{F}$, and

(iii) whenever a sequence of sets $A_1, A_2, \ldots$ belongs to $\mathcal{F}$, their union $\cup_{n=1}^{\infty} A_n$ also belongs to $\mathcal{F}$.

Definition 8. Let Ω be a nonempty set, and let $\mathcal{F}$ be a σ-algebra of subsets of Ω. A probability measure $\mathbb{P}$ is a function that, to every set $A \in \mathcal{F}$, assigns a number in $[0, 1]$, called the probability of A and written $\mathbb{P}(A)$. We require

(i) $\mathbb{P}(\Omega) = 1$, and

(ii) (countable additivity) whenever $A_1, A_2, \ldots$ is a sequence of disjoint sets in $\mathcal{F}$, then

$$\mathbb{P}\left(\cup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n).$$

The triple $(\Omega, \mathcal{F}, \mathbb{P})$ is called a probability space.

Definition 9. A filtration on $(\Omega, \mathcal{F}, \mathbb{P})$ is an increasing family of σ-algebras $(\mathcal{F})_t : \forall t \geq s > 0, \mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$.

Definition 10. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A real-valued function X defined on Ω is called random variable if the sets

$$\{X \leq x\} := \{\omega \in \Omega | X(\omega) \leq x\} = X^{-1}((-\infty, x])$$

are measurable for all $x \in \mathbb{R}$. That is, $\{X \leq x\} \in \mathcal{F}$.

Definition 11. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\forall \omega \in \Omega, \mathbb{P}(\omega) > 0$, equipped with a filtration $(\mathcal{F}_n)_{0 \leq n \leq N}$. A sequence $(X_n)_{0 \leq n \leq N}$ of random variables is adapted to the filtration if for any n, X_n is $\mathcal{F}_n$-measurable.

Definition 12. For $x \in \mathbb{R}$ the distribution function $F(x)$ of a continuous random variable X is defined by the probability $\mathbb{P}$ that $X \leq x$,

$$F(x) := \mathbb{P}(X \leq x).$$

Distributions are nondecreasing, right-continuous and satisfy the limits

$$\lim_{x \to -\infty} F(x) = 0 \quad \text{and} \quad \lim_{x \to +\infty} F(x) = 1.$$

For all $x \in \mathbb{R}$ if $f(x) \geq 0$ and

$$F(x) = \int_{-\infty}^{x} f(t)dt,$$

then f is called a density function.

If X has a density f, then the expectation or mean of X is

$$\mu := \mathbb{E}[X] = \int_{-\infty}^{\infty} x f(x) dx.$$

The variance is defined as

$$\sigma^2 := \text{Var}[X] := \mathbb{E}[(X - \mu)^2] = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx.$$

A consequence is

$$\sigma^2 = \mathbb{E}[X^2] - \mu^2.$$

The expectation of a discrete random variable X on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is defined by

$$\mathbb{E}[X] = \sum_{\omega \in \Omega} X(\omega)\mathbb{P}(\omega).$$

Definition 13 (Conditional expectation)**.** Let $\mathcal{G}$ be a sub σ-algebra of $\mathcal{F}$. Then, a conditional expectation of X given $\mathcal{G}$ is

$$\int_A \mathbb{E}[X|G]d\mathbb{P} = \int_A X d\mathbb{P} \quad \text{for all} \quad A \in \mathcal{G}.$$

Theorem A.1.1 (Girsanov Theorem)**.** *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space and let* Z *be an almost surely nonnegative random variable with* $\mathbb{E}[Z] = 1$*. For* $A \in \mathcal{F}$*, define*

$$\tilde{\mathbb{P}}(A) = \int_A Z(\omega) d\mathbb{P}(\omega).$$

Then $\tilde{\mathbb{P}}$ *is a probability measure. Furthermore, if* X *is a nonnegative random variable, then*

$$\tilde{\mathbb{E}}[X] = \mathbb{E}[XZ]. \tag{A.1}$$

If Z *is almost surely strictly positive, we also have*

$$\mathbb{E}[Y] = \tilde{\mathbb{E}}\left[\frac{Y}{Z}\right]$$

for every nonnegative random variable Y*.*

The $\tilde{\mathbb{E}}$ appearing in (A.1) is expectation under the probability measure $\tilde{\mathbb{P}}$.

Definition 14 (Martingale)**.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let T be a fixed positive number, and let $(\mathcal{F}_t)_{0 \le t \le T}$ be a filtration of sub-σ-algebras of $\mathcal{F}$. Consider an adapted stochastic process $(M_t)_{0 \le t \le T}$. If

$$\mathbb{E}[M_t|\mathcal{F}_s] = M_s \quad \text{for all} \quad 0 \le s \le t \le T,$$

this process is called as a martingale.

Definition 15. Let Ω be a nonempty set and $\mathcal{F}$ a σ-algebra of subsets of Ω. Two probability measures $\mathbb{P}$ and $\tilde{\mathbb{P}}$ on $(\Omega, \mathcal{F})$ are said to be equivalent if they agree which sets in $\mathcal{F}$ have probability zero.

Definition 16 (Risk-Neutral Measure)**.** Let $\tilde{\mathbb{P}}$ be a probability measure on $(\Omega, \mathcal{F})$, equivalent to the market measure $\mathbb{P}$. If the discounted process $e^{-rt}S_t$ is a martingale under $\tilde{\mathbb{P}}$, then $\tilde{\mathbb{P}}$ is called as a risk-neutral measure.

Definition 17 (Convergence in mean)**.** A sequence X_n is said to converge in the (square) mean to X, if $\mathbb{E}[X_n^2] < \infty, \quad \mathbb{E}[X^2] < \infty$ and if

$$\lim_{n\to\infty} \mathbb{E}[(X - X_n)^2] = 0.$$

A notation for convergence in mean is

$$l.i.m_{n\to\infty} X_n = X.$$

Definition 18 (Convergence in probability)**.** A sequence $X_1, X_2, \ldots$ of random variables converges to x in probability if for every number $\epsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}(|X_n - x| < \epsilon) = 1.$$

Definition 19 (Normal Distribution)**.** The density of the normal distribution is

$$\phi_{\mu,\sigma^2}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} \quad \text{for} \quad -\infty < x < \infty.$$

$X \sim N(\mu, \sigma^2)$ means that X is normally distributed with expectation μ and variance σ^2.

Definition 20 (The Standard Normal Distribution)**.** The normal distribution with mean 0 and variance 1 is called the standard normal distribution.

The probability density function of the standard normal distribution is denoted as

$$n(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad \text{for} \quad -\infty < x < \infty,$$

and the cumulative distribution function is

$$N(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \quad \text{for} \quad -\infty < x < \infty.$$

Theorem A.1.2. *Suppose that Z_1 and Z_2 are independent random variables, each of which has the standard normal distribution. Let $\mu_1, \mu_2, \sigma_1, \sigma_2$, and ρ be constants such that $-\infty < \mu_i < \infty \quad (i = 1, 2), \quad \sigma_i > 0 \quad (i = 1, 2)$, and $-1 < \rho < 1$. Define the two new random variables X_1 and X_2 as follows:*

$$X_1 = \sigma_1 Z_1 + \mu_1,$$
$$X_2 = \sigma_2[\rho Z_1 + (1 - \rho^2)^{1/2} Z_2] + \mu_2.$$

Then, the joint probability density function of X_1 and X_2 is

$$f(x_1,x_2) = \frac{1}{2\pi(1-\rho^2)^{1/2}\sigma_1\sigma_2}\exp\left(-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 -2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right)+\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]\right) \quad \text{(A.2)}$$

Definition 21 (Bivariate Normal Distribution)**.** When the joint probability density function of two random variables X_1 and X_2 is of the form in Equation (A.2), it is said that X_1 and X_2 have the bivariate normal distribution with means μ_1 and μ_2, and variances σ_1^2 and σ_2^2, and correlation ρ. Here, X_1 is normal with mean μ_1 and variance σ_1^2, X_2 is normal with mean μ_2 and variance σ_2^2 and $\rho = \frac{\mathrm{Cov}(X_1,X_2)}{\sigma_1\sigma_2}$.

We denote

$$(X_1,X_2) \sim \mathcal{M}\left(\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix},\begin{pmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix}\right).$$

Definition 22 (Standard Bivariate Normal Distribution)**.** The random variables X_1 and X_2 have the standard bivariate normal distribution with correlation coefficient ρ if their joint probability density function is given by

$$n(x_1,x_2) = \frac{1}{2\pi(1-\rho^2)^{1/2}}\exp\left(-\frac{1}{2(1-\rho^2)}[x_1^2-2\rho x_1x_2+x_2^2]\right)$$

where $\rho \in (-1,1)$.

The distribution function of the bivariate standard normal distribution is

$$M(a,b;\rho) = \int_{-\infty}^{a}\int_{-\infty}^{b}\frac{1}{2\pi(1-\rho^2)^{1/2}}\exp\left(-\frac{1}{2(1-\rho^2)}[x_1^2-2\rho x_1x_2+x_2^2]\right).$$

Lemma A.1.3. *Let $c,\alpha,\beta\in\mathbb{R}$, $\alpha>0$, then we have*

$$\int_{-\infty}^{c}\phi_{\mu,\sigma^2}(x)N(\alpha x+\beta)dx = \mathbb{P}(X\le c, Z\le\beta),$$

with

$$(X,Z) \sim \mathcal{M}\left(\begin{pmatrix}\mu\\ -\alpha\mu\end{pmatrix},\begin{pmatrix}\sigma^2 & -\alpha\sigma^2\\ -\alpha\sigma^2 & 1+\alpha^2\sigma^2\end{pmatrix}\right)$$

where ϕ_{μ,σ^2} denotes the density of a normal distribution with expectation μ and variance σ^2, N is the standard normal distribution function.

Remark A.1.4 (The Newton-Raphson Method)**.** Assume that $f \in C^2[a,b]$. The first Taylor polynomial of $f(x)$ about the point $x = x_0 + p$ is

$$f(x) = f(x_0) + f'(x_0)p + \frac{1}{2}f''(x_0)p^2. \tag{A.3}$$

Taking $f(x_0 + p) = 0$, Equation (A.3) gives for $p = p_0$

$$p_0 = -\frac{f(x_0)}{f'(x_0)}.$$

This process goes on iteratively until it converges to a fixed point using

$$p_n = -\frac{f(x_n)}{f'(x_n)}.$$

Then, starting with an initial approximation x_0 and generating the sequence $\{x_n\}_{n=0}^{\infty}$ by

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \quad \text{for} \quad n \geq 1,$$

we obtain the Newton-Raphson method.

Definition 23. An adapted sequence $(M_n)_{0\leq n\leq N}$ of real random variables is:

(i) a martingale if $\mathbb{E}[M_{n+1}|\mathcal{F}_n] = M_n$ for all $n \leq N-1$,

(ii) a supermartingale if $\mathbb{E}[M_{n+1}|\mathcal{F}_n] \leq M_n$ for all $n \leq N-1$,

(iii) a submartingale $\mathbb{E}[M_{n+1}|\mathcal{F}_n] \geq M_n$ for all $n \leq N-1$.

Definition 24. A random variable ν taking values in $\{0, 1, \ldots, N\}$ is a stopping time if, for any $n \in \{0, 1, \ldots, N\}$,

$$\nu = n \in \mathcal{F}_n.$$

Proposition A.1.5. The random variable defined by

$$\nu_0 = \inf\{n \geq 0 | U_n = Z_n\}$$

is a stopping time and the stopped sequence $(U_{n\wedge\nu_0})_{0\leq n\leq N}$ is martingale where $(Z_n)_{0\leq n\leq N}$ is an adapted sequence.

Proof. See [35]. □

Corollary A.1.6. *The stopping time ν_0 satisfies*

$$U_0 = \mathbb{E}[Z_{\nu_0}|\mathcal{F}_0] = \sup_{\nu \in \tau_{0,N}} \mathbb{E}[Z_\nu|\mathcal{F}_0]$$

where $\tau_{n,N}$ is the set of stopping times taking values in $\{n, n+1, \ldots, N\}$.

Proof. See [35]. □

Definition 25. A stopping time ν is called optimal for the sequence $(Z_n)_{0\leq n\leq N}$ if

$$\mathbb{E}[Z_{\nu_0}|\mathcal{F}_0] = \sup_{\tau_{0,N}} \mathbb{E}[Z_\nu|\mathcal{F}_0].$$

We can see that ν is optimal. The following theorem gives a characterisation of optimal stopping times that shows that ν_0 is the smallest optimal stopping time.

Theorem A.1.7. *A stopping time ν is optimal if and only if*

$$Z_\nu = U_\nu \quad \text{and} \quad (U_{n\wedge\nu_0})_{0\leq n\leq N} \quad \text{is a martingale.}$$

Proof. See [35]. □

Proposition A.1.8. The largest optimal stopping time for (Z_n) is given by

$$\nu_{max} = \begin{cases} N, & A_N = 0 \\ \inf\{n, A_{n+1} \neq 0\}, & A_N \neq 0 \end{cases}.$$

Proof. See [35]. □

Definition 26. (Snell-envelope) Let $(Z_n)_{0\leq n\leq N}$ be an an adapted sequence on probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Define a sequence $(U_n)_{0\leq n\leq N}$ by

$$\begin{aligned} U_N &= Z_N, \\ U_n &= \max\left(Z_n, \mathbb{E}U_{n+1}|\mathcal{F}_n\right) \quad \forall n \geq N-1. \end{aligned}$$

The process U is called Snell-envelope of the sequence Z.

Appendix B

ALGORITHMS

B.1 Algorithms

Binomial American One Dividend

```
function [price, lattice, latticeS] = Binomial_American_Onedividend(S0, K, r, T,...
sigma, M, put, D, R);

if nargin < 7
    put = 0;
end
dt = T/M;
x = floor(R*M/T);
u = exp(sigma.*sqrt(dt));
d = 1./u;
q = (exp((r)*dt) - d) ./ (u-d);
lattice = zeros(M+1,M+1);
latticeS = zeros(M+1,M+1);
for i = 0:x
    for j=0:i
latticeS(j+1,i+1)=S0*u^j*d^(i-j);
    end
end
    for j=0:x
       latticeS(j+1,x+1) = latticeS(j+1,x+1)-D;
    end
for i = x+1:M
    for j=0:i
       latticeS(j+1,i+1)=latticeS(j+1,i)*d;
```

```
       latticeS(i+1,i+1)=latticeS(i,i)*u;
    end
end
   for j = 0:M
   if (put)
        lattice(j+1,M+1) = max(0, K - latticeS(j+1,M+1)); % Put Payoff
    else
        lattice(j+1,M+1) = max(0, latticeS(j+1,M+1) - K); % Call Payoff
    end
end
for i=M-1:-1:0
     for j=0:i
        if (put)
            Pji = K - latticeS(j+1,i+1);
        else
            Pji = latticeS(j+1,i+1) - K;
        end
        lattice(j+1,i+1) = max( Pji, exp(-r*dt) * (q*lattice(j+2,i+2) + ...
        (1-q)*lattice(j+1,i+2)) );
           end
end
price = lattice(1,1);
```

Roll Geske Whaley Method

```
function C = rollgeskewhaley(S0,K,r,T,sigma,D,t);

for i=1:length(D)
  a1 = (log((S0 - D(i) * exp(-r*t))/K) + (r + sigma^2/2)*T)/(sigma*sqrt(T));
  a2 = a1 - sigma*sqrt(T);

  S_star = bisect(S0,K,r,T,sigma,D(i),t);

  b1 = (log((S0 - D(i) * exp(-r*t))/S_star) + (r + sigma^2/2)*T)/(sigma*sqrt(T));
  b2 = b1 - (sigma * sqrt(t));

  C1 = (S0 - D(i) * exp(-r*t))*normcdf(b1);
```

```
  C2 = (S0 - D(i) * exp(-r*t))*bivnormcdf(a1,-b1,-sqrt(t/T));
  C3 = - K * exp(-r*T) * bivnormcdf(a2,-b2,-sqrt(t/T));
  C4 = - (K - D(i))* exp(-r*t)*normcdf(b2);
end
C(i) = C1+C2+C3+C4;
plot(D, C);
```

Bisection Method

```
function S_star = bisect(S0,K,r,T,sigma,D,t)

[blscall,p] = bsprice(S0, K, r, T, sigma);
S_high = S0;
temp = blscall - S_high -D + K;
ACCURACY = 1e-6;
S_low = 0;

while ((temp>0) && (S_high<=1e10) )
      S_high = S_high * 2;
      c = bsprice(S_high, K, r, T-t, sigma);
      temp = c-S_high-D+K;
end

if (S_high>1e10) %exercise not optimal
   c = bsprice(S_high, K, r, T-t, sigma);
end

S_star = 0.5 * S_high; % // now find S_star
   c = bsprice(S_star, K, r, T-t, sigma);
   test = c-S_star-D+X;

   while (abs(test) > 0 && (S_high-S_low)>ACCURACY  )
      if (test<0.0)
          S_high = S_star;
      else
          S_low = S_star;
      end
          S_star = 0.5 * (S_high + S_low);
```

```
      c = bsprice(S_star,K,r,T-t,sigma);
      test = c-S_star-D+K;
    end
   S_star = S_star;
```

Bjerksund and Stensland Method

```
function ret=BSCallPrice(S,K,r,b,sigma,T);


sigma2=sigma*sigma;
beta=(.5-b/sigma2)+sqrt((b/sigma2-.5)^2+2*r/sigma2);


%calculate trigger price
B_inf=(beta/(beta-1))*X;
B0=max(K,(r/(r-b))*X);
h=-(b*T+2*sigma*sqrt(T))*(K*K/((B_inf-B0)*B0));
S_star=B0+(B_inf-B0)*(1-exp(h));


alpha=(S_star-K)*S_star^-beta;


f1=alpha*S^beta;
f2=alpha*BSPhi(S,T,beta,S_star,S_star,sigma,r,b);
f3=BSPhi(S,T,1,S_star,S_star,sigma,r,b);
f4=BSPhi(S,T,1,K,S_star,sigma,r,b);
f5=X*BSPhi(S,T,0,S_star,S_star,sigma,r,b);
f6=X*BSPhi(S,T,0,K,S_star,sigma,r,b);
ret=f1-f2+f3-f4-f5+f6;
```

Function Φ for Bjerksund and Stensland Method

```
function ret=BSPhi(S,T,gamma,H,S_star,sigma,r,b)
sigma2=sigma*sigma;
k=2*b/sigma2+(2*gamma-1);
lambda=-r+gamma*b+.5*gamma*(gamma-1)*sigma2;
t1=(log(S/H)+(b+(gamma-.5)*sigma2)*T)/(sigma*sqrt(T));
t2=(log(S_star^2/(S*H))+(b+(gamma-.5)*sigma2)*T)/(sigma*sqrt(T));
ret=exp(lambda*T)*S^gamma*(normcdf(-t1)- ((S_star/S)^k)*normcdf(-t2));
```

Barone-Adesi Whaley Method for American Call

```
function C = bawhaleycall(S,K,r,T,sigma,q);
```

```
h = 1 - exp(-r*T);
a = 2*r/sigma^2;
beta = 2*(r-q)/sigma^2;

%b=-0.04;
Sx = findSx(K,r,q,T,sigma);
lambda2 = (1-beta+sqrt((beta-1)^2+4*a/h))/2;
d1 = (log(Sx/K)+((r-q)+1/2*sigma^2)*T)/(sigma*sqrt(T));
d11 = (log(S/K) + (r-q + 0.5*sigma^2)*T)/(sigma*sqrt(T));
d2 = d11 - sigma*sqrt(T);
N_d1 = 0.5*(1+erf(d1/sqrt(2)));
N_d11 = 0.5*(1+erf(d11/sqrt(2)));
N_d2 = 0.5*(1+erf(d2/sqrt(2)));
y1 = exp((-q)*T);
y2 = exp(-r*T);
K = S*exp(-q*T)*N_d11-K*exp(-r*T)*N_d2;
A2 = (1-exp(-q*T)*N_d1)*Sx/lambda2;
if S<Sx
        C = K + A2*(S/Sx)^lambda2;
else
        C = S - K;
end
```

Seed Value for Barone-Adesi Whaley Method

```
function S_star = findSx(K,r,q,T,sigma)

h = 1-exp(-r*T); alpha = 2*r/sigma^2;
beta = 2*(r-q)/sigma^2;
lambda2 = (-beta+1+sqrt((beta-1)^2+4*alpha/h))/2;
A = 1/(1-beta+sqrt((beta-1)^2+4*alpha));
B = K/(1-2/A);
h2 = -((r-q)*T+2*sigma*sqrt(T))*(K/(B-K));
Sx = K + (B-K)*(1-exp(h2));
d1 = (log(Sx/K)+((r-q)+1/2*sigma^2)*T)/(sigma*sqrt(T));
d2 = d1 - sigma*sqrt(T);
dN_d1 = 1/(sqrt(2*pi*T)*sigma*Sx)*exp(-d1^2/2);
```

```
dN_d2 = 1/(sqrt(2*pi*T)*sigma*Sx)*exp(-d2^2/2);
N_d1 = 0.5*(1+erf(d1/sqrt(2)));
N_d2 = 0.5*(1+erf(d2/sqrt(2)));
y1 = exp((-q)*T); y2 = exp(-r*T);
c = Sx*y1*N_d1-X*y2*N_d2;
iter = 0;
while abs(Sx-K-c-(1-y1*N_d1)*Sx/lambda2)/K > 10^-5 && Sx ~= 0
    dc = y1*N_d1+y1*Sx*dN_d1-X*y2*dN_d2;
    fprime = 1-dc+(y1*dN_d1*Sx-1+y1*N_d1)/lambda2;
    f = Sx-K-c-(1-y1*N_d1)*Sx/lambda2;
    Sx = Sx - f/fprime;
    d1 = (log(Sx/K)+((r-q)+1/2*sigma^2)*T)/(sigma*sqrt(T));
    d2 = d1 - sigma*sqrt(T);
    dN_d1 = 1/(sqrt(2*pi*T)*sigma*Sx)*exp(-d1^2/2);
    dN_d2 = 1/(sqrt(2*pi*T)*sigma*Sx)*exp(-d2^2/2);
    N_d1 = 0.5*(1+erf(d1/sqrt(2)));
    N_d2 = 0.5*(1+erf(d2/sqrt(2)));
    y1 = exp((-q)*T);
    y2 = exp(-r*T);
c = Sx*y1*N_d1-K*y2*N_d2;
    iter = iter + 1;
end
S_star=Sx;
```