Option pricing models
Black-Scholes & Wavelet-based

Harsh Pruthi

A dissertation submitted for the partial fulfillment of BS-MS dual degree in Science

Indian Institute of Science Education and Research Mohali

May 2020

Certificate of Examination

This is to certify that the dissertation titled **"Option pricing models: Black-Scholes and Wavelet-based"** submitted by **Mr. Harsh Pruthi** (Reg. No. MS14003) for the partial fulfilment of BS-MS dual degree programme of the Institute, has been examined by the thesis committee duly appointed by the Institute. The committee finds the work done by the candidate satisfactory and recommends that the report be accepted.

Dr. Abhik Ganguli | Dr. Soma Maity | Dr. Varadharaj R. Srinivasan (Supervisor)

Dated: May 4, 2020

Declaration

The work presented in this dissertation has been carried out by me under the guidance of Dr. Varadharaj R. Srinivasan at the Indian Institute of Science Education and Research Mohali.

This work has not been submitted in part or in full for a degree, a diploma, or a fellowship to any other university or institute. Whenever contributions of others are involved, every effort is made to indicate this clearly, with due acknowledgement of collaborative research and discussions. This thesis is a bonafide record of original work done by me and all sources listed within have been detailed in the bibliography.

Harsh Pruthi

(Candidate)

Dated: May 4, 2020

In my capacity as the supervisor of the candidate's project work, I certify that the above statements by the candidate are true to the best of my knowledge.

Dr. Varadharaj R. Srinivasan
(Supervisor)

Acknowledgement

With great pleasure, I express my deepest gratitude towards people and institution that have helped me all along the way to secure, initiate and successfully complete my MS thesis. The work done during this thesis was carried out by the inspiration received from the external guide, Dr. Pooja Soni, Assistant Professor, Department of Statistics, University Business School, Panjab University. Her comprehensive knowledge and experience helped me understand the essentials of this project and provided me an opportunity to learn an interesting topic.

I also acknowledge the esteemed committee members Dr. Abhik Ganguli, Dr. Soma Maity and Dr. Varadharaj R. Srinivasan (supervisor).

I would like to thank my colleague and friend Mr. Adeetya V. Tantia who introduced me to the subject of 'Option Pricing'.

Harsh Pruthi

List of Figures

Notation

S :	stock price
S_0 :	current stock price
P :	price of put option
C :	priceof call option
K :	strike price
T :	expiration time(or maturity date)
r :	riskfree rate
$\mathcal{P}$:	partition
$\mathbb{P}$:	probability measure
X :	random variable
ε,μ :	expected value or expectation
σ^2 :	variance
$\{a_1, \ldots, a_n\}$:	assets
$\mathbb{F}$:	filtration
$\mathbb{X}$:	stochastic process
Θ_i :	portfolio
$S_{i,j}$:	price random variable
$\mathcal{I}$:	pricing functional
Φ :	trading strategy
$\langle \cdot, \cdot \rangle$:	inner product
$\Vert \cdot \Vert_p$:	L^p-norm
ψ :	mother wavelet
$\psi_{a,b}$:	child wavelets
Wf :	wavelet transform of f

Contents

Abstract

A stock option is a financial contract which gives its owner the right to buy (or sell) a stock for a fixed value in the future. Option pricing models aim to determine a fair price for a stock option. The starting point of option pricing theory is considered to be the Black and Scholes published paper of 1973 providing a model for valuing European options.

This thesis aims at studying the discrete-time Binomial model for pricing options, which in the limit goes to the continuous-time Black-Scholes model.

Since then, large number of parametric and non-parametric methods have been developed to relax one or more restrictions of the original Black–Scholes model.

One amongst them are the Fourier inversion methods, which depend on the availability of an expression for the characteristic function of the stochastic processes modelling the underlying assets.

Wavelet theory, viewed as an extension of Fourier analysis, aims to represent complicated functions using sums of simple ones. In wavelets, the building blocks, instead of sinusoidal, are wavelets, which are functions that can be arbitrarily translated and dilated in order to generate basis of $L^2(\mathbb{R})$. The wavelet-based methods are based on the approximation of functions by projecting on the wavelets basis such that the coefficients of the expansion are expressed by means of the Fourier transform of the function to approximate. Two such methods, one each for European and Asian options, are studied and presented.

Chapter 1

Introduction

1.1 Classification of financial instruments

There are many possible classifications of financial instruments.
As per [CZ04], for a division see Figure 1.1. A **security** is a document that confers upon its owner a financial claim. In contrast, a general **financial contract** links two parties nominally and not through the ownership of a document.

Fixed-income securities pay fixed amounts of money to their owners. These include bonds, regular savings accounts, money-market accounts, etc.
A **bond** is a security that gives its owner the right to a fixed, predetermined payment, at a future, predetermined date.
A **stock** is a security that gives its owner the right to a proportion of any profits that

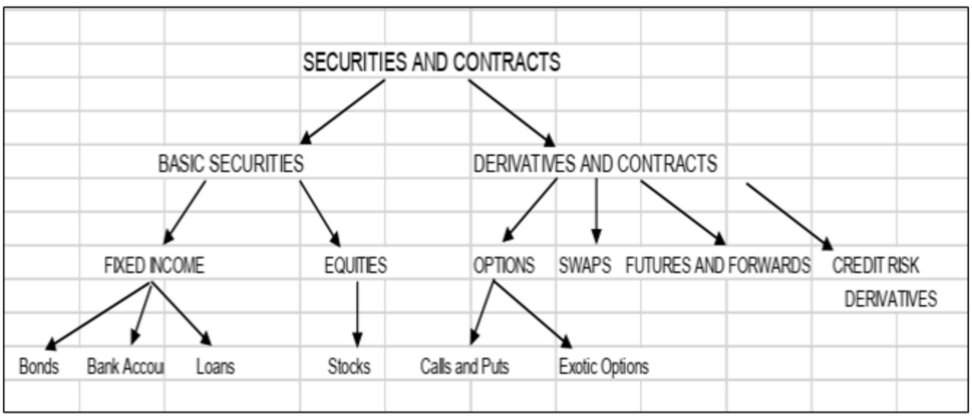

Figure 1.1: A classification of financial instruments

might be distributed (rather than reinvested) by the firm that issues the stock and to the corresponding part of the firm in case it decides to close down and liquidate.

Derivatives are financial instruments whose payoff depends on the value of another financial variable (price of a stock, price of a bond, exchange rate, and so on), called underlying.

Futures and forwards are contracts by which one party agrees to buy the underlying asset at a future, predetermined date at a predetermined price. The other party agrees to deliver the underlying at the predetermined date for the agreed price.

A **swap** is a contract by which two parties agree to exchange two cash flows with different features.

1.2 Background on Options

Definition. A **stock option** is a contract between the writer(seller) and the buyer of the option. The writer has a short position and the buyer has a long position. Every option has an underlying stock, an expiration date and a stock price, also called striking price or exercise price.

1. In a call option the buyer has the right to buy the underlying stock from the writer at the strike price K per share.
   - In a European call, the right to buy can only be exercised on the expiration date of the call.
   - In an American call, the right to buy can be exercised at any time on or before the expiration date of the call.
2. In a put option, the buyer has the right to sell the underlying stock to the writer at the strike price K per share.
   - In a European put, the right to sell can only be exercised on the expiration date of the call.
   - In an American put, the right to sell can be exercised at any time on or before the expiration date of the call.

Most of the option pricing literature considers mainly stock options and so does this thesis.

Exchanges. The options exchange, through which options on major stocks are traded, determines the terms of an option, such as the expiration date and strike price.

Purpose of Options. Primarily, options are used for hedging and for speculation. A hedge is an investment that reduces the risk in an existing position. Options can also be used for implicit leverage, that is, as a tool for borrowing money.

1.2.1 Payoff and Profit Curves

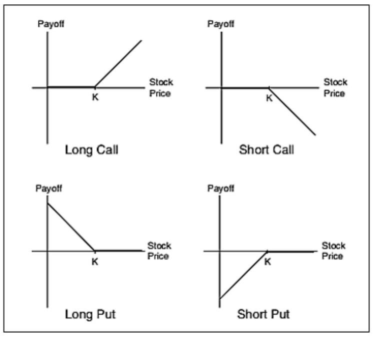

Figure 1.2: Payoff curves

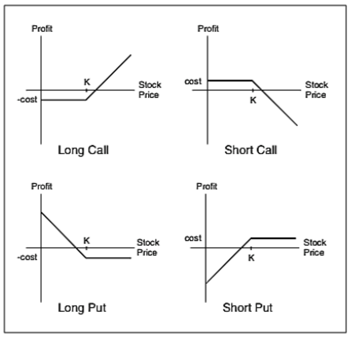

Figure 1.3: Profit curves

1.2.2 Types of Options

There are various types of options, depending on the expiration time and the way to claculate the payoff. Some of these are mentioned in [Buc12].

- **Vanilla** options
  1. **European** - an option that may only be exercised on expiry
  2. **American** - an option that may be exercised on any trading day on or before expiration
- **Exotic** Options
  1. **Lookback** - payoffs depend not only on the underlying asset price at expiry, but also on the maximum or minimum asset price over some predefined monitoring window

2. **Binary** - options that pay out at one or more future dates if and only if some exercise condition is met
3. **Barrier** - options have payoffs which depend, at least in part, on some aspect of the actual asset path traced out, and not just on the terminal value of the path. Barrier options have payoffs which depend on whether a given barrier level x = b is crossed or otherwise during the life of the option
4. **Asian** - an option whose payoff is determined by the average underlying price over some preset time period

and others.

1.2.3 Put-Call option parity

Arbitrage opportunity is an investment opportunity that is guaranteed not to result in a loss and may (with positive probability) result in a gain. If an arbitrage opportunity exists, then prices will be adjusted to eliminate that opportunity.
No-arbitrage Pricing Principle. As a consequence of the tendency to an arbitrage-free market equilibrium, it only makes sense to price assets under the assumption that there is no arbitrage.

Theorem 1.2.1 (European Options with Dividends)**.** *Suppose that a stock is currently selling at a price of S_0 per share. An European put on this stock sells for P dollars and an European call for C dollars, both having the same strike price K and expiration time T. Suppose that the present value of any dividends paid by the stock during the period in question is d_0. Then the no-arbitrage pricing principle implies that*

$$C - P \;=\; S_0 - Ke^{-rT} - d_0$$

where r is the risk-free interest rate.

Theorem 1.2.2 (American Options with Dividends)**.** *Suppose that a stock is currently selling at a price of S_0 per share. An American put on this stock sells for P dollars and an American call for C dollars, both having the same strike price K and expiration time T. Suppose that the present value of any dividends paid by the stock* *during the period in question is d_0. Then the no-arbitrage pricing principle implies that*

$$S_9 - K - d_0 \ \leq \ C - P \ \leq \ S_0 - Ke^{-rT}$$

where r is the risk-free interest rate.

Refer to [Rom04] for proofs.

Chapter 2

Mathematical prerequisites

2.1 Discrete Probability

Definition. Let Ω be a nonempty set. Then a **partition** of Ω is a collection $\mathcal{P} = \{B_1, \ldots, B_n\}$ of nonempty subsets of Ω, called the **blocks** of the partition, with the following properties:

1. $B_i \cap B_j = \emptyset$ for all $i \neq j$
2. $B_1 \cup \cdot \cup B_n = \Omega$

Definition. Let $\mathcal{P} = \{B_1, \ldots, B_n\}$ be a partition of a set Ω. Then a partition $\mathcal{Q} = \{C_1, \ldots, C_n\}$ is called a **refinement** of P, written $P \succ Q$, if each block C_i of Q is completely contained in some block B_j of P or, equivalently, if each block of P is a union of blocks of Q.

Definition. A collection $\mathcal{A}$ of subsets of Ω is called an **algebra** of sets (or algebra) if it satisfies the following properties:

1. $\emptyset \in \mathcal{A}$
2. $A \in \mathcal{A} \Rightarrow A^C \in \mathcal{A}$
3. $A, B \in \mathcal{A} \Rightarrow A \cup B \in \mathcal{A}$

Definition. An **atom** of $\mathcal{A}$ is a nonempty set $S \in \mathcal{A}$ with the property that no nonempty proper subset of S is also in $\mathcal{A}$.

Definition. A **finite probability space** is a pair $(\Omega, \mathbb{P})$ consisting of a finite non-empty set Ω, called the **sample space** and a real-valued function $\mathbb{P}$ defined on the set of all subsets of Ω, called a **probability measure** on Ω. The function $\mathbb{P}$ must satisfy the following properties:

1. For all $A \subseteq \Omega$, $0 \leq \mathbb{P}(A) \leq 1$

2. $\mathbb{P}(\Omega) = 1$

3. If A and B are disjoint, $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$

All subsets of Ω are called **events**.

Definition. We assign to each of the elements $w \in \Omega$ a number p_w satisfying $0 \leq p_w \leq 1$ and for which

$$\sum_{w \in \Omega} p_w = 1$$

Then we can define a **probability measure** $\mathbb{P}$ by setting $\mathbb{P}(\{w\}) = p_w$. Extending this to all events, we get

$$\mathbb{P}(A) = \sum_{w \in \Omega} \mathbb{P}(\{w\})$$

The set $\{p_w : w \in \Omega\}$ is called a **probability mass** or **probability distribution** and the function $f : \Omega \to \mathbb{R}$ defined by $f(w) = p_w$ is called a **probability mass function**.

Definition. A real-valued function $X : \Omega \to \mathbb{R}$ defined on a finite sample space Ω is called a **random variable** on Ω.

Definition. Let X be a random variable on Ω with

$$im(X) = \{x_1, \ldots, x_n\}.$$

Then the partition

$$\mathcal{P}_X = \{\{X = x_1\}, ..., \{X = x_n\}\}$$

is called the **partition defined by random variable** X.

Definition. If $\mathbb{P}$ is a probability measure on Ω, then we denote $\mathbb{P}(X = x) = \mathbb{P}(\{X = x\})$. The partition $\mathcal{P}_X$ defined by X then defines a probability measure $\mathbb{P}_X$ on

$im(X) = \{x_1, \ldots, x_n\}$ by

$$\mathbb{P}_X(x_i) = \mathbb{P}(X = x_i)$$

for all $x_i \in im(X)$. This probability distribution on $im(X)$ is called the **probability distribution of random variable** X and the corresponding probability mass function $f : A \to \mathbb{R}$ defined by $f(x_i) = \mathbb{P}(X = x_i)$ is called the **probability mass function of random variable** X.

Definition. Let $\mathcal{Q} = \{B_1, ..., B_n\}$ be a partition of Ω. A random variable X on Ω is **$\mathcal{Q}$-measurable** if X is constant on each block of $\mathcal{Q}$, that is, if it has the form

$$X = \sum_{i=1}^{n} b_i 1_{B_i}$$

for (not necessarily distinct) constants $b_i \in \mathbb{R}$.

Definition. Random variables $X_1, ..., X_n$ are **independent random variables** if

$$\mathbb{P}(X_1 = x_1, \ldots, X_n = x_n) = \prod_{i=1}^{n} \mathbb{P}(X_i = x_i)$$

for all $x_i \in im(X_i)$.

Definition. Let X be a random variable on a finite probability space $(\Omega, \mathbb{P})$. The **expected value** or **expectation of random variable** X is given by

$$\varepsilon_{\mathbb{P}}(X) = \sum_{w \in \Omega} X(w)\mathbb{P}(w)$$

Definition. Let X be a random variable with finite expected value μ. The **variance of random variable** X is

$${\sigma_X}^2 = \varepsilon((X - \mu)^2)$$

Standardizing a random variable. If X is a random variable with expected value μ and variance σ^2, then we can define a new random variable Y by

$$Y = \frac{X - \mu}{\sigma}$$

Then Y has expected value 0 and variance 1.

Definition. Let $(\Omega, \mathbb{P})$ be a finite probability space and A be an event for which $\mathbb{P}(A) > 0$. The **conditional expectation of random variable** X with respect to the event A is

$$\varepsilon(X \mid A) = \sum_{w \in \Omega} X(w)\mathbb{P}(w \mid A)$$

Definition. Let $\mathcal{P} = \{B_1, ..., B_n\}$ be a partition of Ω for which $\mathbb{P}(B_i) > 0$ for all i. The **conditional expectation of random variable** X **with respect to partition** $\mathcal{P}$ is

$$\varepsilon(X \mid \mathcal{P}) = \varepsilon(X \mid B_1)1_{B_1} + ... + \varepsilon(X \mid B_n)1_{B_n}$$

2.2 Stochastic processes

A finite collection of nodes or vertices, combined with a finite collection of edges connecting certain nodes forms a **state tree**. The vertical columns in which nodes of

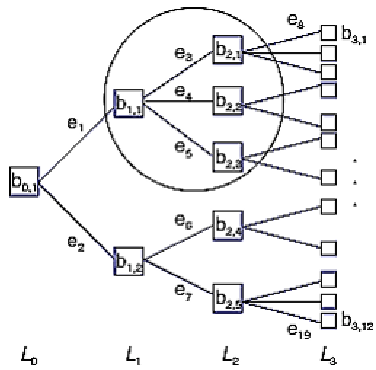

Figure 2.1: A State Tree

a state tree are organized are called **levels**, often thought of as times.
Then, the **time-t_k intermediate states** of the tree are the nodes at time t_k and the time-t_k state space is the set of all time-t_k nodes. The state space for the final time is called the final state space.
A node b, its children and the edges that connect these children to the parent becomes the **child sub-tree** of b.

Definition. A sequence $\mathbb{F} = \{\mathcal{P}_0, \ldots, \mathcal{P}_N\}$ of partitions of a set $\Omega = \{w_1, \ldots, w_m\}$

for which

$$\mathcal{P}_0 \succ \ldots \succ \mathcal{P}_N$$

is called a **filtration**.

A filtration is called an **information structure** if $\mathcal{P}_0 = \{\Omega\}$ and $\mathcal{P}_N = \{\{w_1\}, \ldots, \{w_m\}\}$

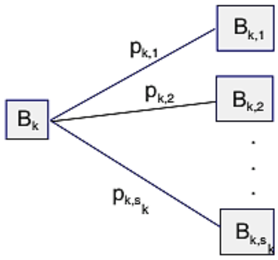

Figure 2.2: The child subtree of B_k

The **child sub-tree number** for B_k is defined as the sum of the edge labels in the child sub-tree for B_k,

$$C(B_k) = \sum_{i=1}^{s_k} p_{k,i}.$$

The product of the edge labels of the path is called the **path number** of B_k, denoted by H_k.

Theorem 2.2.1. *Let $\Omega = \{w_1, ..., w_m\}$ be a finite set with information structure $\mathbb{F} = \{\mathcal{P}_0, \ldots, \mathcal{P}_N\}$. Suppose that we label the edges of the state tree of $\mathbb{F}$ with positive real numbers such that $C(B_k) = 1$ for all $B_k \in \mathcal{P}_k$ and all $k = 0, \ldots, N-1$. Then the path number function defines a strongly positive probability distribution on Ω, with associated probability measure $\mathbb{P}(\{w\}) = H(\{w\})$ and more generally $\mathbb{P}(B_k) = H(B_k)$ for all states $B_k \in \mathcal{P}_k$ and all $k = 0, \ldots, N$. Also, $p_k = \mathbb{P}(B_{k,i} \mid B_k)$.*

Definition. A (finite) **stochastic process** on a sample space Ω is a sequence $\mathbb{X} = (X_0, X_1, \ldots, X_N)$ of random variables defined on Ω. If $k \leq m$, the **change** in $\mathbb{X}$ from k to m is the difference

$$\Delta_{k,m}(\mathbb{X}) = X_m - X_k$$

Definition. Let Ω be a finite set, with filtration $\mathbb{F} = \{\mathcal{P}_0, \ldots, \mathcal{P}_N\}$. A stochastic process $\mathbb{X} = (X_0, X_1, \ldots, X_N)$ on Ω is **adapted to the filtration** $\mathbb{F}$, or is $\mathbb{F}$**-adapted** if X_k is $\mathcal{P}_k$-measurable for all $k = 0, \ldots, N$.

$\mathbb{X}$ is **predictable** or previewable with respect to $\mathbb{F}$ if X_k is $\mathcal{P}_{k-1}$-measurable for all $k = 1, \dots, N$.

Definition. Let $\mathbb{X} = (X_0, X_1, \dots, X_N)$ be a stochastic process adapted to filtration $\mathbb{F}$. Then $\mathbb{X}$ is a **martingale** with respect to the triple $(\Omega, \mathbb{P}, \mathbb{F})$, if

$$\varepsilon_{\mathbb{P}}(X_{k+1} \mid \mathcal{P}_k) = X_k$$

or equivalently,

$$\varepsilon_{\mathbb{P}}(\Delta_{k,k+1}(X) \mid \mathcal{P}_k) = 0 \tag{2.1}$$

for all $k = 0, \dots, N-1$. This expresses the idea that $\mathbb{X}$ is "fair" over every one-period time interval $[t_k, t_{k+1}]$.

The martingale condition 2.1 is equivalent to

$$\varepsilon_{\mathbb{P}}(\triangle_{k,k+1}(X) \mid \mathcal{B}_k) = 0 \tag{2.2}$$

for all $B_k \in \mathcal{P}_k$. Since X_k is constant on B_k, by denoting this constant by $X_k(B_k)$, the martingale condition is

$$\varepsilon(X_{k+1} \mid \mathcal{B}_k) = X_k(B_k) \tag{2.3}$$

Either of 2.2 and 2.3 can be referred to as the **local martingale condition** at B_k in P_k.

Characterizing martingales

Theorem 2.2.2. *Let $\mathbb{F} = \{\mathcal{P}_0, \dots, \mathcal{P}_N\}$ be a filtration on Ω and let $\mathbb{X} = (X_0, X_1, \dots, X_N)$ be a stochastic process adapted to $\mathbb{F}$. The following are equivalent:*

1. *$\mathbb{X$ is a martingale*

$$\varepsilon(X_{k+1} \mid \mathcal{P}_k) = X_k$$

or in terms of change,

$$\varepsilon(\triangle_{k,k+1}(X) \mid \mathcal{P}_k) = 0$$

for all $k = 0, \dots, N-1$; that is, X is "fair" over any $[t_{k+1}, t_k]$.

2. *X is "fair" over any $[t_k, t_{k+i}]$; that is,*

$$\varepsilon(X_{k+i} \mid \mathcal{P}_k) = X_k \quad i.e., \quad \varepsilon(\triangle_{k,k+i}(X) \mid \mathcal{P}_k) = 0$$

for all $i \geq 0$ and $k \geq 0$ for which $k + i \leq N$.

3. *X is "fair" over any time interval of the form $[t_k, t_N]$; that is,*

$$\varepsilon(X_N \mid \mathcal{P}_k) = X_k \quad i.e, \quad \varepsilon(\triangle_{k,N}(X) \mid \mathcal{P}_k) = 0$$

for all $k = 0, \ldots, N - 1$.

4. *$\mathbb{X}$ is fair at every $\mathcal{B}_k \in \mathcal{P}_k$; that is, 2.3 holds for all $k = 0, \ldots, N - 1$ and for all states $\mathcal{B}_k \in \mathcal{P}_k$.*

Moreover, if X is a martingale, then

$$\varepsilon(X_k) = \varepsilon(X_0) = X_0(\Omega) \quad i.e. \quad \varepsilon(\Delta_{0,k}(X)) = 0$$

for all $0 \leq k \leq N$.

2.3 Continuous Probability

Definition. Let Ω be a non-empty set. A non-empty collection Σ of subsets of Ω is a **σ-algebra** if

1. $\Omega \in \Sigma$

2. If $A_1, A_2, \ldots$ is a sequence of elements of Σ, then

$$\bigcup_{i=1}^{\infty} A_i \in \Sigma$$

3. If $A \in \Sigma$, then $A^C \in \Sigma$

Definition. A **probability space** is a triple $(\Omega, \Sigma, \mathbb{P})$ comprising of a non-empty set Ω, called a **sample space**, a σ-algebra Σ of subsets of Ω whose elements are called **events** and a real-valued function $\mathbb{P}$ defined on Σ called a **probability measure**.

The function $\mathbb{P}$ must satisfy the following properties:

1. For all $A \in \Sigma$, $0 \leq \mathbb{P}(A) \leq 1$
2. $\mathbb{P}(\Omega) = 1$
3. If $A_1, A_2, ...$ Is a sequence of pairwise mutually exclusive events, then

$$\mathbb{P}(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$$

Definition. A **distribution function** is a function $F : \mathbb{R} \to \mathbb{R}$ satisfying:

1. F is non-decreasing; i.e.

$$s < t \Rightarrow F(s) \leq F(t)$$

2. F is right-continuous; i.e. the right-hand limit exists everywhere and

$$\lim_{t \to a^+} F(t) = F(a)$$

3. F satisfies

$$\lim_{t \to -\infty} F(t) = 0 \quad \mathit{and} \quad \lim_{t \to \infty} F(t) = 1$$

Theorem 2.3.1. *1. Let $\mathbb{P}$ be a probability measure on $\mathbb{R}$. The function $F_{\mathbb{P}} : \mathbb{R} \to \mathbb{R}$ defined by*

$$F_{\mathbb{P}}(t) = \mathbb{P}((-\infty, t])$$

*is a probability distribution function, called the **distribution function** of $\mathbb{P}$.*

2. Let $F : \mathbb{R} \to \mathbb{R}$ be a distribution function. Then there is a unique probability measure $\mathbb{P}_F$ on $\mathbb{R}$ whose distribution function is F; that is, for which

$$\mathbb{P}_F((-\infty, t]) = F(t)$$

Definition. 1. A **density function** $f : \mathbb{R} \to \mathbb{R}$ is a non-negative function for which

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

2. A probability measure $\mathbb{P}$ on $\mathbb{R}$ or equivalently a distribution function $F_{\mathbb{P}}$ is **absolutely continuous** if there is a density function $f : \mathbb{R} \to \mathbb{R}$ for which

$$F_{\mathbb{P}}(t) = \mathbb{P}((-\infty, t]) = \int_{-\infty}^{t} f(x)dx$$

From this definition, it follows that

$$\mathbb{P}((a, b]) = \int_{a}^{b} f(x)dx$$

Definition. A function $X : \Omega \to \mathbb{R}$ is **Σ-measurable** if the inverse image of every open interval is in Σ, i.e.,

$$X^{-1}((a, b)) \in \Sigma$$

A measurable function on (Ω, Σ) is also called a **random variable**.

If $(\Omega, \Sigma, \mathbb{P})$ is a probability space and X is a random variable on (Ω, Σ), then X defines a **distribution function** F_X and corresponding **probability measure** $\mathbb{P}_X$ on $\mathbb{R}$ by

$$F_X(t) = \mathbb{P}_X((-\infty, t]) = \mathbb{P}(X \leq t)$$

Definition. A collection $X_1, \ldots, X_n$ of random variables is **independent** if

$$\mathbb{P}(X_1 \leq t_1, \ldots, X_n \leq t_n) = \prod_{i=1}^{n} \mathbb{P}(X_i \leq t_i)$$

Definition. Let X be an absolutely continuous random variable, with density function f. The **expected value** or **mean** of X is the improper integral

$$\varepsilon(X) = \int_{-\infty}^{\infty} x f(x)dx$$

which exists provided that

$$\int_{-\infty}^{\infty} \mid x \mid f(x)dx < \infty$$

Definition. The **variance** of X is

$$Var(X) = \varepsilon((X - \mu)^2)$$

Normal Distribution. Consider the normal distribution whose density function is

$$N_{\mu,\sigma}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Theorem 2.3.2. *If $N_{\mu,\sigma}$ is a random variable with mean μ and variance σ^2 then $Z = \frac{N_{\mu,\sigma}-\mu}{\sigma}$ is a standard normal random variable. Conversely, if Z is a standard normal random variable, then $N_{\mu,\sigma} = \sigma N_{0,1} + \mu$ is a random variable with mean μ and variance σ^2.*

Theorem 2.3.3. *If X is lognormally distributed, that is, if $Y = \log X$ is normal with mean a and variance b^2, then*

$$\varepsilon(X) = \varepsilon(e^Y) = e^{a+\frac{1}{2}b^2}$$

$$Var(X) = Var(e^Y) = e^{2a+b^2}(e^{b^2} - 1)$$

Definition. A sequence of functions from $\mathbb{R}$ to $\mathbb{R}$, (f_n), **converges pointwise** to $f : \mathbb{R} \to \mathbb{R}$ if for each real x, the sequence of real numbers $(f_n(x))$ converges to the real number $f(x)$.

Definition. Let (X_n) be a sequence of random variables, defined on $(\Omega_n, \mathbb{P}_n)$. Let X be a random variable on $(\Omega, \mathbb{P})$. Then (X_n) **converges in distribution** to X, denoted

$$X_n \xrightarrow{dist} X$$

if the distribution functions (F_{X_n}) converge pointwise to the distribution function F_X at all points where F_X is continuous. Thus, if F_X is continuous at s, then

$$\lim_{n\to\infty} F_{X_n}(s) = F_X(s)$$

that is,

$$\lim_{n\to\infty} \mathbb{P}_n(X_n \le s) = \mathbb{P}(X \le s)$$

Theorem 2.3.4. *For* (X_n) *and* X *as above,*

1.
$$X_n \xrightarrow{dist} X \quad \Leftrightarrow \quad \varepsilon_{\mathbb{P}_n}(g(X_n)) \to \varepsilon_{\mathbb{P}}(g(X))$$

for all bounded continuous functions $g : \mathbb{R} \to \mathbb{R}$. *In particular,*

$$X_n \xrightarrow{dist} X \quad \Rightarrow \quad \varepsilon_{\mathbb{P}_n}(X_n) \to \varepsilon_{\mathbb{P}}(X)$$

2. *For all continuous functions* $f : \mathbb{R} \to \mathbb{R}$,

$$X_n \xrightarrow{dist} X \quad \Rightarrow \quad f(X_n) \xrightarrow{dist} f(X)$$

Definition. Two (or more) random variables are said to be **identically distributed** if they have the same distribution function.

Theorem 2.3.5 (Central Limit Theorem). *Let* $X_1, X_2, ...$ *be a sequence of independent, identically distributed random variables with finite mean* μ *and variance* $\sigma^2 > 0$. *If*

$$S_n = \sum_{i=1}^{n} X_i$$

then the standardized random variables ${S_n}^* = \frac{S_n - n\mu}{\sqrt{n}\sigma}$ *converge in distribution to a standard normal random variable* $N_{0,1}$*; that is,*

$$\lim_{n\to\infty} F_{{S_n}^*}(t) = \phi_{0,1}(t)$$

and so

$$\mathbb{P}({S_n}^* < t) \approx \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{x^2}{2}} dx$$

where the error in the approximation tends to 0 *as* n *tends to* ∞.

Theorem 2.3.6 (Central Limit Theorem). *Consider a triangular array of random variables*

$$\begin{array}{ccc} B_{1,1} & & \\ B_{2,1} & B_{2,2} & \\ B_{3,1} & B_{3,2} & B_{3,3} \\ \vdots & \vdots & \vdots \end{array}$$

where for each row, the random variables $B_{n,i}$ are independent, identically distributed standard Bernoulli random variables with

$$\mathbb{P}(B_{n,i} = \tfrac{q_n}{\sqrt{p_n q_n}}) = p_n \ \text{ and } \ \mathbb{P}(B_{n,i} = \tfrac{-p_n}{\sqrt{p_n q_n}}) = q_n$$

However, the random variables in different rows need not be independent or identically distributed, or even defined on the same probability space. Suppose also that $p_n \to p$, where $0 < p < 1$. Then the random variables

$$S_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} B_{n,i}$$

converge in distribution to a standard normal random variable.

For proofs of the theorems in this chapter and other properties in probability, refer [Rom04].

Chapter 3

Some basic models

3.1 Discrete-time pricing model

The problem of determining a fair initial value of any derivative is the derivative pricing problem (DPP). At time $t = 0$, the final value of the derivative is unknown, since it generally depends on final value of underlying asset, which inturn depends on the state of the economy at the final time. So, on the space of all final states of the economy, we assume that the final value of the underlying is a known random variable. This section provides a model for the DPP.

3.1.1 Assumptions:

1. All prices are given in terms of an unspecified **unit of accounting**.

2. There is always available a **risk-free asset**, which cannot decrease in value and whose amount of increase is known in advance for each time interval.

3. Additional assumptions

    - **Infinitely divisible market**- We can speak of, for example, $-\pi$ worth of a stock or bond.

    - **Frictionless market**- All transactions take place immediately and without any external delays.

    - **Perfect market**- No transaction fees or commissions; no restrictions on

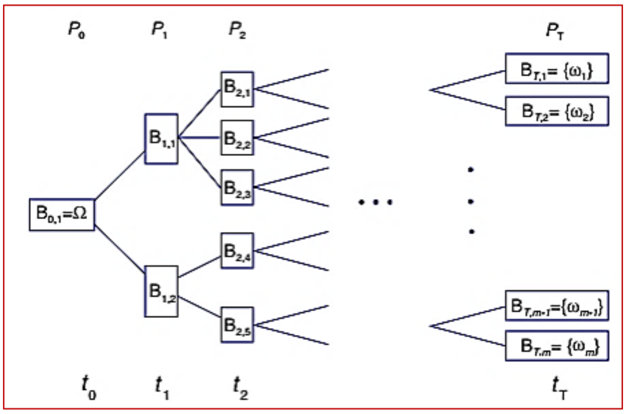

Figure 3.1: The state tree $\mathbb{T}$

short selling; borrowing rate is same as lending rate.

- **Buy-sell parity**- Any asset's buying price is equal to its selling price.
- **Prices determined under no-arbitrage assumption**- If an arbitrage opportunity exists in the market, the prices will be adjusted to eliminate that opportunity. Therefore, it makes sense to price securities under the assumption that there is no arbitrage.

3.1.2 The basic model

The basic ingredients of the discrete-time model $\mathbb{M}$ are:

- **Times**- $\mathbb{M}$ has $T+1$ times $t_0 < t_1 < \ldots < t_T$.
- **Assets**- $\mathbb{M}$ has a finite(n) number of basic assets $A = \{a_1, \ldots, a_n\}$. a_1 is assumed to be risk-free, while others are risky.
- **States of the economy**- At the final state t_T, we assume that the economy is in one of m possible final states, given by the state space $\Omega = w_1, \ldots, w_m$. At t_0, we know nothing about the final state except that it lies in Ω. As time passes, we may gain some information (but never lose information) about the possible final state of the economy. So, we use an information structure $\mathbb{F} = \{\mathcal{P}_0, \ldots, \mathcal{P}_T\}$ on Ω, called the **state information structure** for $\mathbb{M}$.
  The partition $\mathcal{P} = \{B_{i,1}, \ldots, B_{i,m_i}\}$ of Ω is called the **time-t_i state partition**. For $i < T$, the blocks of P_i are called the **time-t_i intermediate states**.
- **Natural probabilities**- We assume that there exists a probability measure on

Ω which reflects the likelihood that each final state in Ω will be the actual final state. These are called natural probabilities.

- **Asset price**- Each asset has a price at each time that depends on the state of the economy at that time.

Definition. 1. For each time t_i and each asset a_j, the **price random variable** $S_{i,j} : \Omega \to \mathbb{R}$ is a nonnegative $\mathcal{P}$-measurable random variable for which $S_{i,j}(w)$ is the time-t_i price of asset a_j under the final state w belonging to Ω.

2. For a fixed time t_k, the **price vector** for t_k is the vector of asset prices $(S_{k,1}, \ldots, S_{k,n})$.

3. For a fixed asset a_j, the sequence $\mathbb{S}_j = (S_{0,j}, \ldots, S_{T,j})$ is a stochastic process, called the **price process** for asset a_j.

4. For the risk-free asset, the price random variables are constant; that is, they do not depend upon the state of the economy. In particular $S_{0,1} = 1$, and for all times $i > 0$,

$$S_{i,1} = e^{r_0(t_1 - t_0)} \ldots e^{r_{i-1}(t_i t_{i-1})}$$

where r_k is the risk-free rate for the time interval $[t_k, t_{k+1}]$.

Definition. The **discounted asset prices** are given by $\overline{S}_{i,j} = \frac{S_{i,j}}{S_{i,1}}$, also called the **(time-t_0) present value** of $S_{i,j}$.

In particular, $\overline{S}_{i,1} = 1$.

Definition.
- The **asset holding** for asset a_j during the time interval $[t_{i-1}, t_i]$ is a $\mathcal{P}_{i-1}$-measurable random variable $\theta_{i,j}$, where $\theta_{i,j}(w)$ is the quantity of asset a_j held during this time interval under state $w \in \Omega$.
- The stochastic process $\Phi_j = (\theta_{1,j}, \ldots, \theta_{T,j})$ is the **asset holding process** for a_j.
- A **portfolio** for the time interval $[t_{i-1}, t_i]$ is a random vector $\Theta_i = (\theta_{i,1}, \ldots, \theta_{i,n})$ on Ω where $\theta_{i,j}$ is the asset quantity for asset a_j over this time interval.

The process of liquidating the portfolio θ_i and acquiring the portfolio θ_{i+1} at time t_i is termed **portfolio rebalancing**.

Definition. A **trading strategy** for $\mathbb{M}$ is a sequence of portfolios $\Phi = (\Theta_1, \ldots, \Theta_T)$

where $\Theta_i = (\theta_{i,1}, \ldots, \theta_{i,n})$ is a portfolio for the time interval $[t_{i-1}, t_i]$.

Since a portfolio Θ_i exists only during the time interval $[t_{i-1}, t_i]$, it makes sense to assign a value to Θ_i only at the **acquisition time** t_{i-1} and the **liquidation time** t_i. The **acquisition value** or **price** of the portfolio Θ_i is defined by the inner product

$$\nu_{i-1}(\Theta_i) = \langle \Theta_i, S_{i-1} \rangle = \sum_{j=1}^{n} \theta_{i,j} S_{i-1,j}$$

And the **liquidation value** or **price** of Θ_i is defined by

$$\nu_i(\Theta_i) = \langle \Theta_i, S_i \rangle = \sum_{j=1}^{n} \theta_{i,j} S_{i,j}$$

Definition. A trading strategy $\Phi = (\Theta_1, \ldots, \Theta_T)$ is **self-financing trading strategy**(SFTS) if for any time t_i where $i \neq 0, T$, the acquisition price of Θ_{i+1} is equal to the liquidation price of Θ_i; that is,

$$\nu_i(\Theta_{i+1}) = \nu_i(\Theta_i)$$

Thus, a SFTS is initially purchased for the acquisition value $\nu_0(\Theta_1)$ of the first portfolio and is liquidated at time t_T, producing a final payoff of $\nu_T(\Theta_T)$. No other money is added to or removed from the model during its lifetime.
The term **gain** refers to the change in value of a portfolio over a period of time.

1. For $j < k$, the **discounted gain** from t_j to t_k, denoted $\overline{G}_{j,k}$, is given by

$$\overline{G}_{j,k}(\Phi) = \overline{\nu}_k(\Phi) - \overline{\nu}_j(\Phi) = \frac{1}{S_{k,1}} \nu_k(\Phi) - \frac{1}{S_{j,1}} \nu_j(\Phi)$$

2. For any time t_k, the (**cumulative**) **discounted gain** $\overline{G}_k$ is

$$\overline{G}_k(\Phi) = \overline{\nu}_k(\Phi) - \overline{\nu}_0(\Phi)$$

. $\overline{G}_T(\Phi)$ is the discounted final gain in Φ.

Definition. We say that a SFTS $\Phi' = (\Theta'_1, \ldots, \Theta'_T)$ **locks in the gain** in Φ up to time t_k if

$$\overline{G}_T(\Phi') = \overline{G}_k(\Phi)$$

.

Theorem 3.1.1. *Let* $\Phi = (\Theta_1, \ldots, \Theta_T)$ *be a SFTS,* $[t_k, t_{k+m}]$ *be an interval,* $\mathcal{A} = \{B_{k_1}, \ldots, B_{k_s}\}$ *be a collection of time-*t_k *states in* $\mathcal{P}_k$ *and let* $B = B_{k_1} \cup \cdots \cup B_{k_s}$*. Then the SFTS defined by*

$$\Theta_i{}^{(k,k+m)} 1_B = \begin{cases} 0 & if 1 \le i \le k \\ \Phi_i 1_B - (\overline{\nu}_k(\Phi_{k+1}), 0, \ldots, 0) 1_B & if k+1 \le i \le k+m \\ (\overline{G}_{k,k+m}(\Phi), 0, \ldots, 0) 1_B & if k+m < i \le T \end{cases}$$

locks in the discounted gain in Φ *over* $[t_k, t_{k+m}]$ *for the states in* $\mathcal{A}$ *only; i.e.*

$$\overline{G}_T(\Phi^{(k,k+m)} 1_B) = (\overline{G}_{k,k+m}(\Phi) 1_B$$

The particular example given by

$$\Theta_i = \begin{cases} 0 & if 1 \le i \le k \\ (-\overline{S}_{k,j}, \ldots, 0, 1, 0, \ldots, 0) 1_B & if i = k+1 \\ (-\overline{S}_{k,j} + \overline{S}_{k+1,j}, 0, \ldots, 0) 1_B & if k+1 < i \le T \end{cases}$$

is denoted by $\Phi[a_j, t_k, B]$ and called **atomic trading strategy**. The final gain

$$\overline{G}_T(\Phi[a_j, t_k, B]) = \overline{G}_{k,k+1}(\Phi[a_j]) 1_B = \Delta_{k,k+1}(\mathbb{S}_j) 1_B$$

is the change in price from t_k to t_{k+1} for asset a_j in state $B \in \mathcal{P}_k$.

Definition.
- A SFTS Φ is an **arbitrage trading strategy** or **arbitrage opportunity** if $\overline{G}_T(\Phi) > 0$.
- A probability measure $\mathbb{P}$ on Ω is a **martingale measure** (or **risk-neutral probability measure**) for $\mathbb{M}$ if
  1. $\mathbb{P}$ is strongly positive; that is, $\mathbb{P}(w) > 0$ for all $w \in \Omega$.
  2. For each asset a_j, the discounted price process $(\overline{S}_{0,j}, \ldots, \overline{S}_{T,j})$ is a martin-

gale; that is, for all $k \geq 0$,

$$\varepsilon_{\mathbb{P}}((\overline{S}_{k+1,j} \mid \mathcal{P}_k) = (\overline{S}_{k,j}$$

Theorem 3.1.2. *A strongly positive probability measure $\mathbb{P}$ on Ω is a martingale measure for $\mathbb{M}$ if and only if the expected gain of every atomic trading strategy is 0, i.e.*

$$\varepsilon(\overline{G}_T(\Phi[a_j, t_k, B])) = 0$$

for all assets a_j, all times t_k and all states $B \in \mathcal{P}_k$, where $0 \leq k < T$.

Theorem 3.1.3. *Let $\mathbb{M}$ be a discrete-time model with state information structure $\mathbb{F} = \{\mathcal{P}_0, \ldots, \mathcal{P}_T\}$. The following are equivalent for a strongly positive probability measure $\mathbb{P}$.*

1. *$\mathbb{P}$ is a martingale measure; i.e. for all $j = 1, \ldots, n$,*

$$\varepsilon_{\mathbb{P}}(\triangle_{k,k+1}(\mathbb{S}_j) \mid \mathcal{P}_k) = 0.$$

2. *For all $k = 0, \ldots, T-1$,*

$$\varepsilon_{\mathbb{P}}(\overline{\nu}_{k+1}(\Phi) - \overline{\nu}_k(\Phi) \mid \mathcal{P}_k) = 0 \quad i.e., \quad \varepsilon_{\mathbb{P}}(\overline{G}_{k,k+1}(\Phi) \mid \mathcal{P}_k) = 0.$$

3. *For all Φ and all t_k,*

$$\varepsilon_{\mathbb{P}}(\overline{G}_k(\Phi)) = 0 \quad i.e., \quad \varepsilon_{\mathbb{P}}(\overline{\nu}_k(\Phi)) = \overline{\nu}_0(\Phi).$$

4. *For all assets a_j, all times t_k and all states $B \in \mathcal{P}_k$, where $0 \leq k < T$,*

$$\varepsilon(\overline{G}_T(\Phi[a_j, t_k, B])) = 0.$$

Characterizing arbitrage

Theorem 3.1.4 (The First Fundamental Theorem of Asset Pricing). *A discrete-time pricing model $\mathbb{M}$ has no arbitrage opportunities iff it has a martingale measure.*

3.1.3 Computing martingale measures

Figure 2.2 shows a parent B_k and children $\{B_{k+1,1}, \ldots, B_{k+1,s_k}\}$. Edge label from B_k to $B_{k+1,i}$ is $p_{k+1,i}$.

The path numbers for the final states form a probability distribution $\mathbb{P}$ iff

$$1 = \sum_{i=1}^{s_k} p_{k+1,i} \tag{3.1}$$

The edge labels are the conditional probabilities

$$p_{k+1,i} = \mathbb{P}(B_{k+1,i} \mid B_k)$$

Therefore, the local martingale condition at B_k, $\varepsilon(\overline{S}_{k+1,j} \mid B_k) = \overline{S}_{k,j}(B_k)$ can be written as

$$\overline{S}_{k,j}(B_k) = \sum_{i=1}^{s_k} \overline{S}_{k+1,j}(B_{k+1,i})p_{k+1,i} \tag{3.2}$$

for each $j = 1, \ldots, m$.

Equations 3.1 and 3.2 are called the **local martingale equations** for a martingale measure.

Theorem 3.1.5. *If the edges of the state tree are labelled with positive real numbers $p_{k+1,i}$ as described above, then the path numbers define a martingale measure $\mathbb{P}$ on Ω iff the local martingale equations* 3.1 *and* 3.2 *hold.*

Definition. A random variable $X : \Omega \to \mathbb{R}$ is called an **alternative**, or **contingent claim**.

In a way, an alternative $X : \Omega \to \mathbb{R}$ defines an option with final payoff X. We assume that for any random variable $X : \Omega \to \mathbb{R}$, some investor will be willing to buy and some investor will be willing to sell an option whose final payoff is X.

Thus, the problem is of pricing an alternative X, the procedure for which is to find a SFTS within the model with final payoff is X; that is, for which

$$\nu_T(\Phi) = X$$

. Such a SFTS is called **replicating trading strategy**.

We set time-t_k price of X = time-t_k price of $\nu_T(\Phi)$. Any other choice will lead to arbitrage (when the alternative is added to the model).

Definition.

- An alternative X is said to be **attainable** if it has at least one replicating strategy. If every alternative is attainable, $\mathbb{M}$ is **complete**.
- **Law of One Price** states that

$$\nu_T(\Phi_1) = \nu_T(\Phi_2) \Rightarrow \nu_k(\Phi_1) = \nu_k(\Phi_2)$$

for all $0 \leq k \leq T$ and for all trading strategies Φ_1 and Φ_2.

The absence of arbitrage implies that the Law of One Price must hold and the Law of One Price ensures that the following pricing functionals are well-defined.

Definition. Let $\mathbb{M}$ be a model with no arbitrage. For any time t_k, if $\mathcal{M}$ is the vector space of all attainable alternatives, the **time-t_k pricing functional** $\mathcal{I}_k : \mathcal{M} \to \mathbb{R}$ is defined as: If $X \in \mathcal{M}$, then

$$\mathcal{I}_k(X) = \nu_k(\Phi) \tag{3.3}$$

for any replicating trading strategy Φ for X. $\mathcal{I}_0$ is called the **initial pricing functional**.

Pricing an alternative X at time t_k involves first finding a replicating trading strategy Φ and then setting 3.3.

Theorem 3.1.6. *Let $\mathbb{M}$ be a model with no arbitrage and $\mathbb{P}$ its martingale measure. Let X be an attainable alternative.*

1. *The discounted time-t_k price of X is*

$$\overline{\mathcal{I}}_k(X) = \varepsilon_{\mathbb{P}}(\overline{X} \mid \mathcal{P}_k)$$

where $\overline{X} = X/S_{T,i}$

2. *In particular,*

$$\mathcal{I}_0(X) = \varepsilon_{\mathbb{P}}(\overline{X})$$

Theorem 3.1.7 (The Second Fundamental Theorem of Asset Pricing). *Let* $\mathbb{M}$

be a model with no arbitrage opportunities, and hence at least one martingale measure. Then there is a unique martingale measure on $\mathbb{M}$ *iff* $\mathbb{M}$ *is complete.*

3.2 The Binomial model for options

3.2.1 The General Binomial Model

- **Times**- The **lifetime** L is divided into T time intervals $t_0 < t_1 < \cdots < t_T$ of equal length Δt.

- **Assets**- Two assets: a risk-free asset a_1 and a risky asset a_2.

- **States of the economy** The model assumes that during each time interval $[t_k, t_{k+1}]$, the economy either goes up, called an **up-tick** U in the economy or it goes down, called a **down-tick** D in the economy. Each movement is independent of previous movements. Thus, the state space is the set

$$\Omega = \Omega_T = \{U, D\}^T$$

of all strings of U's and D's of length T. These are the final states of the economy.
$\Omega_k = \{U, D\}^k$ is the set of all strings of U's and D's of length k. $[w]_k$ denotes the prefix of any $w \in \Omega$ of length k. For each $\delta = e_1 \dots e_k \in \Omega_k$, the intermediate state $B_\delta \in \mathcal{P}_k$ is the set of all final states having prefix δ; that is,

$$B_\delta = \{w \in \Omega | [w]_k = \delta\}$$

- **Natural probabilities**- For each interval $[t_k, t_{k+1}]$, there is a natural probability p_k of an up-tick and $1 - p_k$ of a down-tick in the economy.

- **The price functions**- The time-t_k price of the stock is denoted by S_k, which is a random variable on Ω. An up-tick during $[t_k, t_{k+1}]$ takes the stock price up by a factor of $u_k \geq 1$ to $S_{k+1} = S_k u_k$ and a down-tick takes the stock price down by a factor of $0 < d_k \leq 1$ to $S_{k+1} = S_k d_k$. u_k is called the time-t_k **up-tick**

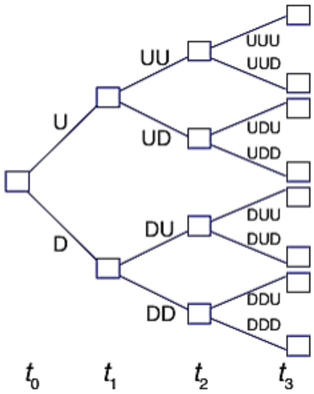

Figure 3.2: State tree for a Binomial Model

factor and d_k the time-t_k **down-tick factor**. u_k and d_k are called the **tick parameters** of the model. Note $0 < d_k \leq 1 \leq u_k$.

If $w \in \Omega$, we let $\delta_k(w)$ be the product of the up-tick and down-tick factors that determine the time-t_k price S_k. In general, we have

$$S_k(w) = S_0 \delta_k(w)$$

for any $w \in \Omega$. The price of the risk-free asset is given by the risk-free rates r_k for the intervals $[t_k, t_{k+1}]$.

Martingale measures in the Binomial Model

The binomial model is free of arbitrage iff $d_k < e^{r_k \Delta t} < u_k$ for all $k = 0, \ldots, T-1$. Then $\mathbb{M}$ is complete and the unique martingale measure $\mathbb{P}$ on $\mathbb{M}$ is defined by the path numbers in the state tree when the up-tick edges of the tree are labelled with the **martingale up-tick probabilities**

$$\pi_k = \frac{e^{r_k \Delta t} - d_k}{u_k - d_k}$$

and the down-tick edges are labelled with the **martingale down-tick probabilities** $1 - \pi_k$.

Pricing in the Model

The pricing functionals $\mathcal{I}_k$ are well defined, since $\mathbb{M}$ is arbitrage-free and complete. In particular, for any random variable X on Ω,

$$\mathcal{I}_0(X) = e^{-r\Delta t}\varepsilon(X)$$

where $r = \sum r_k$ is the sum of the risk-free rates and the expected value is taken under the martingale measure.

3.2.2 Standard Binomial Model

Definition. A binomial model is SBM if the following hold:

1. the up-tick probabilities $u = u_k$ are the same for all times.
2. the down-tick probabilities $d = d_k$ are the same for all times.
3. the risk-free rate $r = r_k$ is the same for all times.
4. the natural probabilities are the same for all times.

Theorem 3.2.1. *Let $N_U(w)$ = number of U's in w, $N_D(w)$ = number of D's in w. The standard binomial model is free of arbitrage iff $d < e^{r\Delta t} < u$. In this case, the time-t_k stock price function S_k is given by*

$$S_k(w) = S_0 u^{N_U([w]_k)} d^{N_D([w]_k)}$$

for any $w \in \Omega$. In particular, the final price is

$$S_T(w) = S_0 u^{N_U(w)} d^{N_D(w)}$$

Moreover, the model is complete and the unique martingale measure $\mathbb{P}$ on $\mathbb{M}$ is defined by

$$\mathbb{P}(w) = \pi^{N_U(w)}(1-\pi)^{N_D(w)}$$

for any $w \in \Omega_T$, where $\pi = \frac{e^{r\Delta t}-d}{u-d}$ is the martingale up-tick probability.

Pricing in a Standard Binomial Model

Theorem 3.2.2. *Let $\mathbb{M}$ be a complete SBM with no arbitrage. Then a European call option with strike price K expiring at the end of the model has initial value*

$$\mathcal{I}_0(Call) = e^{-rL} \sum_{k=0}^{T} \binom{T}{k} (S_0 u^k d^{T-k} - K)^+ \pi^k (1-\pi)^{T-k}$$

and a European put option has initial value

$$\mathcal{I}_0(Put) = e^{-rL} \sum_{k=0}^{T} \binom{T}{k} (K - S_0 u^k d^{T-k})^+ \pi^k (1-\pi)^{T-k}$$

where

$$\pi = \frac{e^{r\Delta t} - d}{u - d}$$

is the martingale up-tick probability.

Choosing the Tick parameters

Let p be the natural (not the martingale) probability of an up-tick in the market. During each time interval, the stock price takes either an up-tick or a down-tick. Hence, we can define independent Bernoulli random variables $E_1, \ldots, E_T$ to track these growth factors by

$$\mathbb{P}(E_k = u) = p$$

$$\mathbb{P}(E_k = d) = 1 - p$$

Then the stock price at time-t_k is given by

$$S_k = S_0 E_1 \cdots E_K$$

and the final time-t_T price is $S_T = S_0 E_1 \cdots E_T$ Since

$$\frac{S_{k+1}}{S_k} = E_{k+1}$$

E_{k+1} is referred to as the simple return of the stock price over $[t_k, t_{k+1}]$. We define the (annualized) instantaneous return of the stock price to be

$$s_{k+1} = \frac{1}{\Delta t} \log E_{k+1}$$

To make the stock price look like exponential growth, we write

$$S_T = S_0 E_1 \cdots E_T = S_0 e^{\sum \log E_i} = S_0 e^{H_T}$$

where

$$H_T = \log\left(\frac{S_T}{S_0}\right) = \sum_{i=1}^{T} \log E_i$$

is called the **logarithmic growth** of the stock price.

Now, the expected value μ and variance s^2 of the instantaneous return are given by

$$\mu = \frac{1}{\Delta t}\varepsilon(\log E_i) = \frac{1}{\Delta t}(p \log u + q \log d)$$

$$s^2 = \frac{1}{(\Delta t)^2} Var(\log E_i) = \frac{1}{(\Delta t)^2} pq(\log u - \log d)^2$$

μ is called the drift and the constant $\sigma = s\sqrt{\Delta t}$ is called the drift volatility of the stock price. Thus, we have

$$\varepsilon(\log E_i) = \mu \Delta t \quad \text{and} \quad Var(\log E_i) = \sigma^2 \Delta t$$

We standardize $\log E_i$ to get

$$X_i = \frac{\log E_i - \mu \Delta t}{\sigma \sqrt{\Delta t}}$$

which are independent standard Bernoulli random variables with

$$X_i = \begin{cases} \frac{q}{\sqrt{pq}} & \text{with probability } p \\ \frac{-p}{\sqrt{pq}} & \text{with probability } q \end{cases}$$

To write the logarithmic growth in terms of the random variables X_i, we have

$$H_T = \mu L + \sigma\sqrt{\Delta t}\sum_{i=1}^{T} X_i$$

The stock price is given by

$$S_T = S_0 e^{H_T} = S_0 e^{\mu L + \sigma\sqrt{\Delta t}\sum_{i=1}^{T} X_i}$$

3.3 The Black-Scholes model for options

The Black–Scholes option pricing formula gives the price of a European put or call based on five quantities:

- the initial price of the underlying stock,
- the strike price of the option,
- the time to expiration,
- the risk-free rate during the lifetime of the option, which is assumed to be constant,
- the volatility of the stock price, a constant that provides a measure of the fluctuation in the stock's price and thus is a measure of the risk involved in the stock.

The first three in the above list are known while the last two can only be estimated.

3.3.1 Stock prices and Brownian motion

Definition. A **continuous stochastic process** on an interval $I \subseteq \mathbb{R}$ of the real line is a collection $\{X_t \mid t \in I\}$ of random variables on Ω indexed by a variable t that ranges over the interval I.

Definition. A continuous stochastic process $\mathcal{W} = \{W_t \mid t \geq 0\}$ is a **Brownian motion process with volatility** σ if

1. $W_0 = 0$

2. Each increment $W_t - W_s$ is normally distributed with mean 0 and variance $\sigma^2(t-s)$. In particular, each W_t is normally distributed with mean 0 and variance $\sigma^2 t$.

3. $\mathcal{W}$ has independent increments; that is, for any times $t_1 \leq t_2 \leq \cdots \leq t_n$, the non-overlapping increments

$$W_{t_2} - W_{t_1}, W_{t_3} - W_{t_2}, \ldots, W_{t_n} - W_{T_{n-1}}$$

are independent.

Definition. A stochastic process of the form $\mathcal{W} = \{\mu t + W_t \mid t \geq 0\}$ where μ is a constant and $\{W_\}$ is Brownian motion with volatility σ is called **Brownian motion with drift μ and volatility σ**.

Definition. A brownian motion process $\{Z_t \mid t \geq 0\}$ with drift $\mu = 0$ and volatility $\sigma = 1$ is called **standard Brownian motion**. Z_t has mean 0 and variance t.

If $\{W_t \mid t \geq 0\}$ is Brownian motion with drift μ and variance σ^2, then we can write

$$W_t = \mu t + \sigma Z_t$$

where $\{Z_t \mid t \geq 0\}$ is standard Brownian motion.

Definition. We assume that the stock price S_t at time t is given by

$$S_t = S_0 e^{H_t}$$

where S_0 is the initial price and H_t is a Brownian motion process. The exponent H_t represents a continuously compounded rate of return of the stock price over the period of time $[0, t]$.

$$H_t = \log(\frac{S_t}{S_0})$$

is referred to as the **logarithmic growth** of the stock price.

Definition. A stochastic process of the form $\{e^{W_t} \mid t \geq 0\}$ where $\{W_t \mid t \geq 0\}$ is Brownian motion is called **geometric Brownian motion**.

If we assume that H_t follows a Brownian motion process with positive drift, then we can write

$$H_t = \log\left(\frac{S_t}{S_0}\right) = \mu t + \sigma W_t$$

where $\{W_t\}$ is standard Brownian motion. Therefore, H_t has a normal distribution with

$$\varepsilon(H_t) = \mu t \quad \text{and} \quad Var(H_t) = \sigma^2 t$$

3.3.2 Binomial model in the limit $\Delta t \to 0$

Let $M_{t,T}$ be a binomial model with the following parameters:

1. Lifetime $t \geq 0$
2. T time increments, each of length $\Delta t = t/T$
3. Up-tick factor u_T and down-tick factor d_T, where $0 < d_T \leq 1 \leq u_T$.
4. A probability of up-tick $p_T \in (0,1)$ and down-tick $q_T = 1 - p_T$.
5. Drift and volatility

$$\mu_T = \frac{1}{\Delta t}(p_T \log u_T + q_T \log d_T)$$
$$\sigma_T]^2 = \frac{1}{\Delta t} p_T q_T (\log u_T - \log d_T)^2$$

Let $S_{t,T}$ be the final stock price and let

$$H_{t,T} = \log(\frac{S_{t,T}}{S_0})$$

be the logarithmic price growth. If the family $\{\mathbb{M}_{t,T} \mid t \geq 0\}$ is **stable** (that is, if $p_T \to p$ for some $p \in (0,1)$ and $\mu_t \to \mu, \sigma_T \to \sigma$ for some real numbers μ and $\sigma \neq 0$), then

$$H_{t,T} \xrightarrow{dist} H_t = \mu t + \sigma\sqrt{t} Z_t \quad \text{and} \quad S_{t,T} \xrightarrow{dist} S_t = S_0 e^{\mu t + \sigma\sqrt{t} Z_t}$$

Moreover, the process $\{\sqrt{t} Z_t \mid t \geq 0\}$ is standard Brownian motion. Hence, the logarithmic growth and stock price processes $\{H_t \mid t \geq 0\}$ and $\{S_t \mid t \geq 0\}$ are Brownian and geometric Brownian, respectively, with drift μ and volatility σ. Also, the stock price growth S_t/S_0 is lognormally distributed with

$$\varepsilon(S_t) = S - 0e^{(\mu+\frac{1}{2}\sigma^2)t}$$

$$Var(S_t) = (S - 0e^{(\mu+\frac{1}{2}\sigma^2)t})^2(e^{\sigma^2 t} - 1)$$

3.3.3 The Natural Binomial Model

In the general binomial model, the up-tick probability p is arbitrary and may or may not be related to the martingale measure up-tick π. If p is determined by empirical means based on economic data, it is referred to as the **natural up-tick probability** and denoted ν. The binomial model with $p = \nu$ is the **natural binomial model**, under which we want to price alternatives.

Assumption 1: The natural probability ν of an up-tick in the stock price satisfies $0 < \nu < 1$ and does not depend on the number T of intervals.
Assumption 2: The resulting natural drift and volatility

$$\mu = \frac{1}{\Delta t}(\nu \log u_T + (1-\nu)\log d_T)$$

$$\sigma^2 = \frac{1}{\Delta t}\nu(1-\nu)(\log u_T - \log d_T)^2$$

called the **natural**(or instantaneous) **drift** and **natural**(or instantaneous) **volatility**, respectively, do not depend on T (or Δt).

Theorem 3.3.1. *Let $\mathbb{M}_T$ be the natural binomial model with natural drift μ and volatility σ. Then,*

1. *The martingale measure up-tick probability is given by*

$$\pi = \frac{e^{(r-\mu)\Delta t + \frac{\nu}{\sqrt{\nu(1-\nu)}}\sigma\sqrt{\Delta t}} - 1}{e^{\frac{1}{\sqrt{\nu(1-\nu)}}\sigma\sqrt{\Delta t}} - 1}$$

2. *The martingale probability $\pi = \pi_T$ approaches the natural probability ν as $T \to \infty$ ($\Delta t \to 0$); that is,*

$$\lim_{T\to\infty} \pi_T = \nu$$

3.3.4 The Martingale Measure Binomial Model

The payoff for a European put with strike price K is

$$X = (K - S_T)^+ = (K - S_0 e^{H_T})^+$$

and the absence of arbitrage implies that

$$\mathcal{I}(Put) = e^{-rt}\varepsilon((K - S_0 e^{H_T})^+)$$

where the expected value is taken under the martingale measure of the natural binomial model.
Taking limits as $T \to \infty$ gives

$$P_\infty = \lim_{T\to\infty} \mathcal{I}(Put) = e^{-rT} \lim_{T\to\infty} \varepsilon((K - S_0 e^{H_T})^+)$$

where P_∞ denotes the limiting price random variable.
Setting $g(x) = (K - S_0 e^x)^+$ which is bounded and continuous on $\mathbb{R}$ gives

$$P_\infty = e^{-rT} \lim_{T\to\infty} \varepsilon(g(H_t))$$

If H_T converges in distribution to a random variable X under the martingale measure, then

$$P_\infty = e^{-rT} \lim_{T\to\infty} \varepsilon(g(H_t)) = e^{-rt}\varepsilon(g(X))$$

To find the limit in distribution of H_T under the martingale measure, we consider the binomial model $\mathbb{M}_{(\pi,T)}$ formed by taking p_T to be the martingale measure up-tick probability from the natural model with natural up-tick probability ν; that is,

$$p_T = \pi_T = \frac{e^{rT} - d_T}{u_T - d_T}$$

We refer to this as the martingale measure binomial model.

Theorem 3.3.2. *The martingale measure model drift $\mu_{\pi,T}$ and volatility $\sigma_{\pi,T}$ are* *related to the natural drift μ_ν and volatility σ_ν by*

$$\mu_{\pi,T} = \mu_\nu + \sigma_\nu \frac{1}{\sqrt{\Delta t}} \frac{(\pi_T - \nu)}{\sqrt{\nu(1-\nu)}}$$

$${\sigma_{\pi,T}}^2 = {\sigma_\nu}^2 \frac{(\pi_T(1-\pi_T)}{\nu(1-\nu)}$$

Theorem 3.3.3. *The following limits hold:*

$$\lim_{T\to\infty} \mu_{\pi,T} = r - \frac{{\sigma_\nu}^2}{2} \text{and} \lim_{T\to\infty} \sigma_{\pi,T} = \sigma_\nu$$

where r is the riskfree rate.

Theorem 3.3.4. *Let $M_{t,T}$ be a binomial model with the following parameters:*

1. *Lifetime $t \geq 0$*

2. *T time increments, each of length $\Delta t = t/T$*

3. *Up-tick factor u_T and down-tick factor d_T, where $0 < d_T \leq 1 \leq u_T$.*

4. *Probability of up-tick equal to the martingale measure up-tick probability from the natural model with up-tick probability ν; that is,*

$$p_T = \pi_T = \frac{e^{rT} - d_T}{u_T - d_T}$$

5. *Drift and volatility*

$$\mu_{\pi,T} = \frac{1}{\Delta t}(\pi_T \log u_T + (1-\pi_T)\log d_T)$$

$${\sigma_{\pi,T}}^2 = \frac{1}{\Delta t}\pi_T(1-\pi_T)(\log u_T - \log d_T)^2$$

Then

$$\pi_T \to \nu, \quad \mu_{\pi,T} \to r - \frac{{\sigma_\nu}^2}{2} \quad \text{and} \quad \sigma_{\pi,T} \to \sigma_\nu$$

where r is the riskfree rate and so the model is stable. Hence the logarithmic growth satisfies

$$H_{t,T} \xrightarrow{dist} H_t = (r - \frac{{\sigma_\nu}^2}{2})t + \sigma_\nu \sqrt{t} Z_t$$

where Z_t is standard normal and

$$S_{t,T} \xrightarrow{dist} S_t = S_0 e^{(r-\frac{\sigma_\nu{}^2}{2})t+\sigma_\nu\sqrt{t}Z_t}$$

where $\{\sqrt{t}Z_t \mid t \geq 0\}$ is standard Brownian motion. The logarithmic growth and stock price processes $\{H_t \mid t \geq 0\}$ and $\{S_t \mid t \geq 0\}$ are Brownian and geometric Brownian, respectively, with drift μ and volatility σ. Also, the stock price growth S_t/S_0 is lognormally distributed with

$$\varepsilon(S_t) = S - 0e^{rt}$$

$$Var(S_t) = (S - 0e^{rt})^2(e^{\sigma_\nu{}^2 t} - 1)$$

3.3.5 The Black-Scholes Option Pricing Formula

With all the above information in this section, we can derive the following:

Theorem 3.3.5 (The Black-Scholes Option Pricing Formulas). *For European options with strike price K and expiration time t, we have*

$$C = S_0\phi_{0,1}(d_1) - Ke^{-rT}\phi_{0,1}(d_2)$$

$$P = Ke^{-rT}\phi_{0,1}(-d_2) - S_0\phi_{0,1}(-d_1)$$

where S_0 is the initial price of the underlying stock, σ is the natural volatility, $\phi_{0,1}$ is the standard normal distribution function and

$$d_1 = \frac{1}{\sigma\sqrt{t}}[\log(\frac{S_0}{K}) + t(r + \frac{1}{2}\sigma^2)]$$

$$d_2 = \frac{1}{\sigma\sqrt{t}}[\log(\frac{S_0}{K}) + t(r - \frac{1}{2}\sigma^2)] = d_1 - \sigma\sqrt{t}$$

where r is the risk-free rate.

For proofs of the theorems and other details, refer [Rom04].

Chapter 4

Wavelet-based option pricing

4.1 Basic wavelet theory

The theory in this section is from [PO18], [Dau92] and [Bla03].

4.1.1 Mathematical background

Definition. The **characteristic function** (Chf) of a continuous random variable X with probability density function f is defined as

$$\mathbb{E}[e^{-iwX}] = \int_{-\infty}^{\infty} f(x)e^{-iwx}dx. \tag{4.1}$$

Let V be a vector space over a field F ($= \mathbb{C}$ or $\mathbb{R}$).

Definition. A **norm** on V is a function $p : V \to \mathbb{R}$ that satisfies the following: For all $a \in F$ and all $u, v \in V$,

1. (Homogeneity) $p(av) = |a|\ p(v)$.

2. (Triangle inequality) $p(u + v) \leq p(u) + p(v)$.

3. (Positivity) $p(v) \geq 0$. $\quad p(v) = 0$ only for $v = 0$

The norm of a vector $v \in V$ is denoted by $\|v\|_V$.
A **normed vector space** is a vector space endowed with a norm.

Definition. A **Banach space** is a vector space over $\mathbb{R}$ or $\mathbb{C}$ equipped with a norm which is complete with respect to that norm, where completeness means that for all Cauchy sequences $\{v_n\}$ in V, there exist $v \in V$ such that $\|v_n - v\|_V \to 0$ as $n \to \infty$.

Definition. For each p, $1 \leq p < \infty$, $L^p(S)$ denotes the class of measurable functions f on a measure space S such that

$$\int_S | f(x)|^p dx < \infty$$

These are Banach spaces with the $L^p(S)$ norm defined by

$$\|f\|_p := (\int_{-\infty}^{\infty} |f(x)|^p \, dx)^{\frac{1}{p}}, \text{ for } 1 \leq p < \infty.$$

Definition. A **separable vector space** V is a vector space such that there exist a countable dense subset $\{f_n\}_{n=0}^{\infty}$, $f_n \in V$, which means that for all $g \in V$ and $\epsilon > 0$ $\exists$ n such that $\|f_n - g\| \leq \epsilon$.

For all $p < \infty$, the L^p spaces are separable.

Definition. An **inner product** on V is defined as a map $\langle \cdot, \cdot \rangle : V \times V \to F$ such that for $x, y \in V$ and $a, b \in \mathbb{C}$,

1. (Symmetry) $\langle y, x \rangle = \overline{\langle x, y \rangle}$.
2. (Bilinearity) $\langle ax_1 + bx_2, y \rangle = a\langle x_1, y \rangle + b\langle x_2, y \rangle$.
3. (Positivity) $\langle x, x \rangle \geq 0$; equality holds only for $x = 0$.

An **inner vector space** is a vector space endowed with an inner product. For $f \in V$, we define $\|f\|_V = \sqrt{\langle f, f \rangle}$

Definition. A **Hilbert space** H is an inner product space that is also a complete metric space with respect to the norm induced by the inner product.

We define an inner product on $L^2(S)$ by

$$\langle f, g \rangle = \int_S f(x)\, \overline{g(x)}\, dx, \quad f, g \in L^2(S)$$

$$\langle f, f \rangle = \|f\|_2^{\ 2}, \quad f \in L^2(S)$$

The standard norm of $L^2(S)$ is derived from this inner product. Endowed with this inner product, $L^2(S)$ becomes a Hilbert space.

Definition. For an inner product space V, orthogonailty means

- Vectors X and Y are **orthogonal vectors** if $\langle X, Y \rangle = 0$.
- Subspaces V_1 and V_2 of V are **orthogonal subspaces** if each vector in V_1 is orthogonal to each vector in V_2.
- Functions f and g are **orthogonal functions**, denoted $f \perp g$, when $\langle f, g \rangle = 0$.
- A set of functions is an **orthogonal set of functions** if and only if every distinct pair in the set is orthogonal.

Definition. Suppose V_0 is a finite dimensional subspace of an inner product space V. For any vector $v \in V$, the **orthogonal projection** of v onto V_0 is the unique vector $v_0 \in V_0$ that is closest to v; i.e.,

$$\|v - v_0\| = \min_{w \in V_0} \|v - w\|$$

Definition.

- The collection of vectors $e_i, i = 1, \ldots, N$, is an **orthonormal vector collection** if each e_i has unit length, $\|e_i\| = 1$, and e_i and e_j are orthogonal for $i \neq j$.
- A sequence of functions $\{f_n\}_{n \in Z}$ is said to be an **orthonormal sequence of functions** if $\langle f_m, f_n \rangle = \delta_{m,n}$, where $\delta_{j,k}$ is the Kronecker delta defined by

$$\delta_{j,k} := \begin{cases} 1, & \text{for } j = k, \\ 0, & \text{for } j \neq k. \end{cases}$$

  If $\{\phi_1, \phi_2, \phi_3, \ldots\}$ is any orthogonal set of non-zero functions, then a corresponding orthonormal set $\{\psi_1, \psi_2, \psi_3, \ldots\}$ can be constructed by "normalizing" each ϕ_k, that is,

$$\psi_k(x) = \frac{\phi_k(x)}{\|\phi_k\|}$$

Definition. Given a Hilbert space V, a **Hilbert basis**(or simply basis) for V is an orthonormal set of vectors, H, with the property that every vector in V can be written as an infinite linear combination of the vectors in the basis.

Theorem 4.1.1. *Let* $\{f_n\}_{n=1}^{\infty}$ *be an orthonormal set on* $L^2(S)$. *Then the following* *conditions also characterize an orthonormal basis.*

1. *For eacg $g \in L^p(S)$,*

$$g = \sum_{n=0}^{\infty} \langle g, f_n \rangle f_n.$$

2. *For each $g \in L^p(S)$,*

$$\|g\|^2 = \sum_n |\langle g, f_n \rangle|^2 .$$

4.1.2 Fourier analysis

Fourier series aims at decomposing a periodic signal into its frequency components which are represented by the sine and cosine. The Fourier transform, viewed as an extension of Fourier series to general, non-periodic functions.

Fourier series

A Fourier series decomposes any periodic function of period $2a$ for $a \in \mathbb{R}$ into the sum of a (possibly infinite) set of simple oscillating functions, specifically complex exponentials (sines and cosines). For each p, $L^p(-a, a)$ denotes the Banach space of functions f satisfying $f(x + 2a) = f(x)$ almost everywhere (a.e.) in $\mathbb{R}$ and $\|f\|_{L^p(-a,a)} < \infty$.

Theorem 4.1.2. *For $a \in \mathbb{R}$, the set of functions,*

$$\left\{ \frac{1}{\sqrt{2a}} e^{\frac{in\pi x}{a}}; \quad n \in \mathbb{Z} \right\},$$

is an orthonormal basis for $L^2(-a, a)$.

Definition. The **complex Fourier series** of a periodic function $f(x) \in L^2(-a, a)$ is given by,

$$f(x) = \sum_{n=-\infty}^{\infty} \alpha_n e^{\frac{in\pi x}{a}},$$

where the coefficients of the complex Fourier series are,

$$\alpha_n = \frac{1}{2a} \int_{-a}^{a} f(x) e^{\frac{-in\pi x}{a}} dx.$$

Definition. The **Fourier series** of a periodic function $f(x) \in L^2(-a, a)$ of period $2a$ is given by,

$$f(x) = a_0 + \sum_{n=1}^{\infty} a_n \cos(\frac{\pi n}{a}x) + \sum_{n=1}^{\infty} b_n \sin(\frac{\pi n}{a}x),$$

where the coefficients of the Fourier series are,

$$a_0 = \frac{1}{2a} \int_{-a}^{a} f(x)dx,$$

$$a_n = \frac{1}{a} \int_{-a}^{a} f(x) \cos(\frac{\pi n}{a}x)dx,$$

$$b_n = \frac{1}{2a} \int_{-a}^{a} f(x) \sin(\frac{\pi n}{a}x)dx.$$

Fourier and inverse Fourier Transform

Definition. The **Fourier transform** of a function $f \in L^1(\mathbb{R})$ is defined by

$$\widehat{f}(w) := \int_{-\infty}^{\infty} e^{-iwx} f(x)dx.$$

The above integral converges and is bounded.

Theorem 4.1.3. *If $f \in L^1(\mathbb{R})$ and $\widehat{f} \in L^1(\mathbb{R})$, then the* ***inverse Fourier transform*** *of f is*

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \widehat{f}(w)e^{iwx}dw.$$

Discrete Fourier Transform

Definition. The **discrete Fourier transform** (DFT) of f is

$$\hat{f}[k] = \sum_{n=0}^{N-1} f[n]exp(\frac{-i2\pi kn}{N})$$

and the **inverse discrete Fourier transform** (IDFT) formula,

$$f[n] = \frac{1}{N} \sum_{k=0}^{N-1} \hat{f}[k]exp(\frac{i2\pi kn}{N})$$

Windowed Fourier transform

Choose a window function $g : \mathbb{R} \to \mathbb{R}_{\geq 0}$, which has "total mass" 1 and is more or less concentrated around $t = 0$, which means that it has, e.g., a compact support containing 0 or at least a maximum at $t = 0$ and fast decay when $|t| \to \infty$.
For a given $s \in \mathbb{R}$, the function

$$g_s : t \mapsto g(t-s)$$

represents the window g, translated by the amount s(to the right, if $s > 0$). We define the **windowed Fourier transform** by

$$Gf : \mathbb{R} \times \mathbb{R} \to \mathbb{C}, \quad (\alpha, s) \mapsto Gf(\alpha, s)$$

of a function f by

$$Gf(\alpha, s) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) g(t-s) e^{-i\alpha t} dt.$$

4.1.3 Wavelets

A limitation of Fourier series is that its building blocks are periodic.
So, we have a different set of building blocks, called wavelets, which, roughly speaking, are waves that travel for one or more periods and are non-zero only over a finite interval. A wavelet can be translated in time, stretched or compressed by scaling to obtain low and high frequency wavelets.
Wavelets are a family of functions constructed from dilation and translation of a single function in $L^2(\mathbb{R})$, ψ, with $\|\psi\| = 1$ and

$$C_\psi := 2\pi \int_{\mathbb{R}*} \frac{|\hat{\psi}(a)|^2}{|a|} \, da < \infty.$$

This is called the **mother wavelet**. When the dilation parameter a and the translation parameter b vary continuously, we have the following family of **continuous wavelets**,

$$\psi_{a,b}(x) = |a|^{-\frac{1}{2}} \, \psi\left(\frac{x-b}{a}\right), \quad a, b \in \mathbb{R}, \, a \neq 0.$$

These are also called **child wavelets**.

Definition. The parameters a and b restricted to the discrete values: $a = {a_0}^{-m}$ and $b = nb_0{a_0}^{-m}$, $a_0 > 1$, $b_0 > 0$ and $n, m \in \mathbb{Z}$, give rise to the following family of **discrete wavelets**

$$\psi_{m,n}(x) = |a_0|^{m/2}\ \psi({a_0}^m x - nb_0).$$

Definition. The **wavelet series** of $f \in L^2(\mathbb{R})$ is defined as

$$f(x) = \sum_{m\in\mathbb{Z}}\sum_{n\in\mathbb{Z}} \langle f, \psi_{m,n}\rangle \psi_{m,n}(x),$$

provided that the functions $\{\psi_{m,n} : m, n \in \mathbb{Z}\}$ form an orthonormal basis of $L^2(\mathbb{R})$.

Continuous Wavelet Transform

Assume that a certain wavelet ψ has been chosen and is held fixed. Then the function

$$Wf(a,b) := \frac{1}{|\ a\ |^{1/2}} \int f(t)\, \overline{\psi\left(\frac{t-b}{a}\right)}\, dt \quad (a \neq 0)$$

is called the **continuous wavelet transform** of $f \in L^2$ with respect to ψ.
After dilation by a and translation by b, we get

$$\psi_{a,b}(t) := \frac{1}{|\ a\ |^{1/2}} \psi(\frac{t-b}{a}).$$

Clearly, $\|\psi_{a,b}\|= 1$. Then, $Wf(a,b) = \langle f, \psi_{a,b}\rangle$.
This implies that

1. At each point $(a,b) \in \mathbb{R}^* \times \mathbb{R}$ the wavelet transform Wf has a well determined value $Wf(a,b)$
2. By Schwarz' inequality, Wf is uniformly bounded on ${\mathbb{R}_-}^2$:

$$|Wf(a,b)| \leq \|f\| \quad \forall\, (a,b) \in \mathbb{R}^2_-.$$

The Fourier transforms of the functions $\psi_{a,b}$ are

$$\hat{\psi}_{a,b}(\xi) = |a|^{1/2}\ e^{-ib\xi}\ \hat{\psi}(a\xi).$$

We therefore can write $Wf(a,b)$ in the following form

$$Wf(a,b) = \langle \hat{f}, \hat{\psi}_{a,b}\rangle = |a|^{1/2} \int \hat{f}(\xi)\, e^{ib\xi}\, \overline{\hat{\psi}(a\xi)}\, d\xi.$$

Theorem 4.1.4. *For fixed $a \neq 0$ the function*

$$Wf(a, \cdot) : b \mapsto Wf(a, b)$$

can be regarded as the Fourier transform of the function F_a, given by

$$F_a(\xi) := \sqrt{2\pi}\,|a|^{1/2}\;\hat{f}(\xi)\,\overline{\hat{\psi}(a\xi)}$$

A Plancherel formula. For the definition of a scalar product for functions $u : \mathbb{R}^2_- \to \mathbb{C}$ we need a measure on $\mathbb{R}_-{}^2$. We see that a point $(a, b) \in \mathbb{R}^2_-$ is used implicitly to characterize the affine transformation

$$S_{a,b} : \mathbb{R} \to \mathbb{R}, \quad \tau \mapsto t := a\tau + b$$

of the time axis. The totality

$$\text{Aff}(\mathbb{R}) := \{S_{a,b} \mid (a, b) \in \mathbb{R}^2_-\}$$

of these affine transformations is a topological group with respect to $\circ$ (i.e. composition) and as such it carries a "natural" measure $d\mu$, called left invariant Haar measure. Formula above defines a parametrization of the group Aff($\mathbb{R}$) by the set $\mathbb{R}^2_-$, so the measure $d\mu$ becomes manifest as a measure in the (a, b)-plane. The resulting expression for $d\mu = d\mu(a, b)$ can be computed explicitly; one finds

$$d\mu = d\mu(a, b) := \frac{1}{|\,a\,|^2} da db.$$

We define the Hilbert space as

$$H := L^2(\mathbb{R}^2_-, d\mu) = L^2\left(\mathbb{R}^* \times \mathbb{R}, \frac{dadb}{|\,a\,|^2}\right)$$

whose scalar product is defined by

$$\langle u, v\rangle_H := \int_{\mathbb{R}^2_-} u(a, b)\overline{v(a, b)}\frac{dadb}{|\,a\,|^2}.$$

Theorem 4.1.5. *Let ψ be an arbitrary wavelet and let W denote the corresponding* *wavelet transform. Then* $\forall\ f, g \in L^2$ *the following is true:*

$$\langle Wf, Wg\rangle_H \;=\; C_\psi \langle f, g\rangle.$$

Theorem 4.1.6. *Let* ψ *and* χ *be two wavelets and assume that the integral*

$$2\pi \int_{\mathbb{R}^*} \frac{\overline{\hat{\psi}(a)}\hat{\chi}(a)}{\mid a \mid} da =: C_{\psi\chi}$$

is defined, i.e., finite. If W_ψ *and* W_χ *denote the wavelet transform with respect to* ψ *and* χ*, then the following is true for arbitrary* $f, g \in L^2$*:*

$$\langle W_\psi f, W_\chi g\rangle_H = C_{\psi\chi}\langle f, g\rangle.$$

Inversion Formula. For deatils and proofs, refer [Dau92].

Theorem 4.1.7. *Let* x *be a point of continuity of* f*. Under suitable assumptions about* f *and* ψ*, one has*

$$f(x) = \frac{1}{C_\psi} \int_{\mathbb{R}^2_-} Wf(a,b)\psi_{a,b}(x)\frac{dadb}{\mid a\mid^2})$$

Theorem 4.1.8.

$$f(x) = \frac{1}{C_{\psi\chi}} \int_{\mathbb{R}_-{}^2} W_\psi f(a,b)\chi_{a,b}(x)\frac{dadb}{\mid a\mid^2})$$

if the quantity $C_{\psi\chi}$ *is defined.*

Decay of the Wavelet Transform

Theorem 4.1.9. *Assume that a wavelet* ψ *with* $t\psi \in L^1$ *has been chosen. Let the time signal* $f \in L^2$ *be globally bounded and assume that* f *is Hoelder continuous at the point* b*, i.e., there is* $\alpha \in (0, 1]$ *such that in a neighbourhood of* b *an estimate of the form*

$$|f(t) - f(b)| \leq C\ |t - b|^\alpha$$

holds. Then

$$|Wf(a,b)| \leq C' \, |a|^{\alpha+\frac{1}{2}}$$

Multiresolution Analysis

Definition. A **multiresolution analysis** (MRA) consists of the following:

(a) A bilateral sequence $\{V_j | j \in \mathbb{Z}\}$ of closed subspaces of L^2. V_j's are ordered by inclusion

$$\ldots \subset V_1 \subset V_0 \subset V_{-1} \subset \ldots \subset V_{-(j-1)} \subset V_{-j} \subset \ldots \subset L^2 \qquad (4.2)$$

and one has

$$\bigcap_j V_j = \{0\} \qquad \text{(separation axiom)}, \qquad (4.3)$$

$$\overline{\bigcup_j V_j} = L^2 \qquad \text{(completeness axiom)}. \qquad (4.4)$$

In the limit, any $f \in L^2$ can be obtained from functions $f_j \in V_j$.

(b) V_j's are connected through the property:

$$V_{j+1} = D_2(V_j) \quad \forall j \in \mathbb{Z}, \qquad (4.5)$$

where for ψ, $D_2\,\psi(t) = \sum_{k=0}^{n} c_k\,\psi(t-k)$. For f, this means

$$f \in V_j \;\Leftrightarrow\; f(2^j\cdot) \in V_0. \qquad (4.6)$$

(c) There is a function $\phi \in L^2 \cap L^1$ such that $(\phi(\cdot - k) \mid k \in \mathbb{Z})$ forms an orthonormal basis of V_0. This function ϕ is commonly called the **scaling function** of the MRA.

Then, the space V_0 can be described as a set of time signals f in the following way:

$$V_0 = \{f \in L^2 \mid f(t) = \textstyle\sum_k c_k \phi(t-k),\; \sum_k \mid c_k \mid^2 < \infty\}\,.$$

Using ϕ as a template we now define the functions

$$\phi_{j,k}(t) := 2^{-j/2}\phi\left(\frac{t-k.2^j}{2^j}\right) = 2^{-j/2}\phi\left(\frac{t}{2^j} - k\right) \quad (j \in \mathbb{Z}, k \in \mathbb{Z});$$

this being in obvious concordance with the formulas defining the wavelet functions $\psi_{m,n}$. Then the family $(\phi_{j,k}|k \in \mathbb{Z})$ is an orthonormal basis of V_j, two subsequent functions $\phi_{j,k}$ and $\phi_{j,k+1}$ now being translated by the amount 2^j with respect to each other.

We may interpret the orthogonal projection P_j of L^2 onto V_j as: The image $P_j f$ of a time signal $f \in L^2$ incorporates all features of f whose horizontal spread over the time axis is of size 2^j or larger. P_j is given by:

$$P_j f = \sum_{k=-\infty}^{\infty} \langle f, \phi_{j,k} \rangle \phi_{j,k}.$$

Because of the inclusions 4.2 the $\phi_{j,k}$ cannot be brought together to form a "big" orthonormal basis of L^2. So we construct a system $(W_j|j \in \mathbb{Z})$ of pairwise orthogonal subspaces $W_j \subset L^2$ in the following way: W_j is the space gained in the transition from V_j to the next larger space V_{j-1} in the chain 4.2, which means that W_j is the orthogonal complement of V_j in V_{j-1}. Then

$$V_{j-1} = V_j \oplus W_j, \quad W_j \perp V_j \qquad \forall j \in \mathbb{Z};$$

Furthermore, everything is set up in such a way that the formulas analogous to 4.5 and 4.6, namely

$$W_{j+1} = D_2(W_j) \; resp. \; f \in W_j \quad \Leftrightarrow \quad f(2^j.) \in W_0,$$

hold likewise;

Theorem 4.1.10. *If the system $(V_j|j \in \mathbb{Z})$ possesses the properties (a) of an MRA, then the corresponding subspaces W_j are pairwise orthogonal, and furthermore*

$$\overline{\bigoplus_j W_j} = L^2 \textit{(orthogonal direct sum).}$$

It can be proven that there exists a wavelet function $\psi(x) \in W_0$ such that the set $\{\psi_{j,k}(x) = 2^{j/2}\, \psi(2^j x - k),\, k \in \mathbb{Z}\}$ is an orthonormal basis of W_j.

Examples of wavelet functions

We have a wide range of wavelet functions: Daubechies, Haar, Shannon, Meyer, Gaussian, Chebyshev, Morlet, Symlets, etc.

Haar wavelet The Haar scaling function is defined as

$$\phi_H(x) = \begin{cases} 1, & \text{if } 0 \leq x < 1, \\ 0, & \text{elsewwhere.} \end{cases}$$

Its Fourier transform is

$$\hat{\phi}_H(w) = \frac{1 - e^{-iw}}{iw}$$

The Haar mother wavelet function is

$$\phi_H(x) = \begin{cases} 1, & \text{for } 0 \leq x < \frac{1}{2}, \\ -1, & \text{for } \frac{1}{2} \leq x < 1, \\ 0, & \text{otherwise.} \end{cases}$$

Shannon wavelet Shannon scaling function is

$$\phi_S(x) = sinc(x) = \begin{cases} \frac{sin(\pi x)}{\pi x}, & \text{if } x \neq 0, \\ 1, & \text{if } x = 0. \end{cases}$$

Its Fourier transform is

$$\hat{\phi_S}(w) = rect(\frac{w}{2\pi})$$

where rect is the function

$$rect(x) = \begin{cases} 1, & \text{if } \mid x \mid < \frac{1}{2}, \\ 1/2, & \text{if } \mid x \mid = \frac{1}{2}, \\ 0, & \text{if } \mid x \mid > \frac{1}{2}. \end{cases}$$

The mother wavelet is,

$$\psi_S(x) = \frac{sin(\pi(x - \frac{1}{2})) - sin(2\pi(x - \frac{1}{2}))}{\pi(x - \frac{1}{2})}$$

4.2 A wavelet-based model for European options

This section is taken from [LCMS19].

4.2.1 Introduction

The book [Ma11] describes a nonparametric option pricing model that focuses on approximating the implied risk-neutral Moment Generating Function(MGF) of the underlying asset returns using wavelets. The MGF can be used to obtain all the statistical moments of the underlying asset distributions and the preference parameter of the utility function; and out-of-sample options with different maturity dates can be directly estimated using the risk-neutral MGF.

MGF of a continuous random variable x is defined as the bilateral Laplace transform of the probability density function $\rho(x)$, i.e.

$$M(s) = \int_{-\infty}^{\infty} \rho(x)e^{-xs}dx.$$

where s is a complex number.

There are mainly three types of application of wavelet methods in finance and economics, as per [HLMS09]

1. for multi-scaling analysis;
2. to de-noise raw data;
3. to estimate unknown parameters of a model.

4.2.2 Bilateral Laplace Transform

Following [HLMS09], let $f(t)$ be a real-valued function, piecewise continuous on $(-\infty, \infty)$. Its **bilateral Laplace transformation** is a complex valued function given by

$$\mathcal{L}\{f(t)\}(s) = F(s) = \int_{-\infty}^{\infty} f(t)e^{-st}dt,$$

where s is a complex value and $\mathcal{L}$ denotes the Laplace transform operator.

The **inverse Laplace transform** can be written as:

$$\mathcal{L}^{-1}\{F(s)\}(t) = f(t) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} F(s)e^{st}ds.$$

where c is a specific real number.

Let $F(s)$ denote $\mathcal{L}\{f(x)\}(s)$ and $G(s)$ denote $\mathcal{L}\{g(x)\}(s)$. Then we have the following properties:

1. Linearity: $\mathcal{L}\{af(x) + bg(x)\}(s) = aF(s) + bG(s)$;
   $\mathcal{L}^{-1}\{aF(s) + bG(s)\}(x) = af(x) + bg(x)$.

2. Frequency shifting: $\mathcal{L}\{e^{-lx}f(x)\}(s) = F(s + l), \quad \forall l \in \mathbb{R}$;
   $\mathcal{L}^{-1}\{F(s + l)\}(x) = e^{-lt}f(x), \quad \forall l \in \mathbb{R}$.

3. Time shifting: $\mathcal{L}\{f(x - x_0)\}(s) = e^{-x_0 s}F(s), \quad \forall x_0 \in \mathbb{R}$;
   $\mathcal{L}^{-1}\{e^{-x_0 s}F(s)\}(x) = f(x - x_0), \quad \forall x_0 \in \mathbb{R}$.

4. Convolution: $\mathcal{L}\{f(x) * g(x)\} = F(s)G(s)$;
   $\mathcal{L}^{-1}\{F(s)G(s)\}(x) = f(x) * g(x)$.
   where $*$ indicates the convolution operator on f and g:
   $f * g = \int_{-\infty}^{\infty} f(\tau)g(t - \tau)d\tau = \int_{-\infty}^{\infty} g(\tau)f(t - \tau)d\tau$

4.2.3 The Model

Under fairly general assumptions including *i.i.d.* distribution for asset returns, the wavelet-based option pricing model can be expressed as follows:

$$C_t(S_t, X, T) = Xe^{-r(t_t)}\mathcal{L}^{-1}\left(\frac{\Theta_{T-t}(s)}{s(s+1)}\right)\left(ln\frac{X}{S_t}\right)$$

where $\mathcal{L}^{-1}$ denotes the bilateral inverse Laplace transform, C_t is the time-t price for a European call option written on asset whose price is S_t with strike price X and a future maturity date T. Interest rate r and the dividend yield are assumed to be constant.

The main ingredient of the model is $\frac{\Theta_{T-t}(s)}{s(s+1)}$, where s is a complex value whose real part $Re(s) < -1$ for calls and $Re(s) > 0$ for puts. The MGF $\Theta_{T-t}(s)$ of the logarithmic returns $\ln \frac{S_T}{S_t}$ captures the underlying asset dynamics and investor expectation embedded in option prices, and needs to be approximated with wavelets. A wavelet which meets the requirements (as per [Mal99]) for the given case is the Franklin hat function defined as

$$h(t) = \begin{cases} (1- \mid t \mid) & \text{if } -1 \leq t < 1 \\ 0 & \text{otherwise} \end{cases}$$

The Laplace transform of $h(t)$, denoted $m_h(s)$, is:

$$m_h(s) = (\frac{e^{s/2} - e^{-s/2}}{s})^2.$$

Then, a set of generalized functions can be generated from $h(t)$:

$$h_{l,k}(t) = 2^{\frac{1}{2}} h(2^l t - k), \quad l, k = 0, \pm 1, \pm 2, \ldots$$

Here, l (scaling parameter) determines the degree of dilation or contraction and k (shifting parameter) controls the horizontal location of the function.
Laplace transform of these $h_{l,k}(t)$, denoted $m_{l,k}(s)$, are:

$$m_{l,k}(s) = 2^{-\frac{l}{2}} e^{-\frac{ks}{2^l}} m_h(\frac{s}{2^l}), \quad l, k = 0, \pm 1, \pm 2, \ldots$$

The risk-neutral MGF of the return per unit of time $\Theta(s)$ can be expanded as:

$$\Theta(s) = \sum_{l=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} a_{lk} m_{l,k}(s).$$

where a_{lk} is a set of unknown coefficients and needs to be estimated by minimizing the sum of squared error between market option prices and theoretical prices. To estimate these, we use the procedure in [HLMS09]

1. For positive integers L and K, set $a_{lk} = 0$ for all $|l| > L$ and $|k| > K$. Let $\theta_{L,K} \equiv \{a_{lk}\}_{l=L,|k| \leq K}$

2. Given the collection $\{S_t, X_i, C_{t,i}, T, r \; ; \; i = 1, 2, \ldots, N\}$, of market data for options at time t, estimate $\theta_{L,K}$ by minimizing the sum of squared errors between

market option prices $C_{t,i}$ and theoretical prices $\hat{C}_{t,i}$:

$$\min_{\theta_{L,K}} \sum_i (C_{t,i} - \hat{C}_{t,i}(\theta_{L,K}, S_t, X_i, C_{t,i}, T, r))^2.$$

3. In each iteration, increment L by 1 and repeat steps 1 and 2 until $\sum_i (C_{t,i} - \hat{C}_{t,i})^2 < \varepsilon$ for an arbitrary $\varepsilon > 0$.

This yields:

$$\hat{\Theta}(s) = \sum_{|l|=L} \sum_{|k|\leq K} \hat{a}_{lk} m_{lk}(s).$$

In the empirical analysis, L and K are chosen by the optimisation programme so that a satisfactory estimation result can be obtained.

4.3 A wavelet-based model for Asian options

This section is taken from [CMM15].

4.3.1 Introduction

The value of Asian options depends on the average stock price. For fixed Strike Asian option, the payoff depends on the difference between the average of the underlying and a fixed strike. For floating Strike options, the payoff depends on the difference between the average of the underlying and the value of the asset at maturity. For most Asian options, the average is computed by considering the underlying asset values at prefixed dates, like the end of each day, week or month.

The pricing procedure described in [FMM11] is based on a randomization technique, according to which the expiry date of the option is modelled as a random variable distributed as geometric. The computational kernel is the solution of integral equations. The integral equations involved in the model are Fredholm integral equations of the second kind, the kernels of which fall within the class of the aforementioned operators for which projection onto wavelet bases is particularly effective. For this reason, we apply the DWT to the linear systems which arise from the discrete operators. Transforming the linear systems matrices into sparse matrices is possible owing to the wavelet localization property, resulting in a fast algorithm that preserves the accuracy of the original method.

4.3.2 The randomization pricing method

We assume that the risk-neutral process for the stock price $S(t)$ is described by

$$S(t) = S_0 e^{(r-d+g)t+L(t)},$$

where r is the short rate(continuously compounded interest rate), d is the dividend yield and g is the compensator, chosen to ensure that the discounted price process is a martingale. $L(t)$ is a Levy process, identified by its characteristic exponent $\psi(w) = \log \mathbb{E}(e^{iwL(1)})$.

Consider M equidistant monitoring dates, with amplitude of the interval Δ, such that $t_0 = 0$, $t_1 = \Delta, \cdots$, $t_n = n\Delta, \cdots$, $t_M = M\Delta = T$. The log-return on each time interval has the following characteristic function

$$\phi(w) = e^{(\psi(w)+iw(r-d+g))\Delta}. \tag{4.7}$$

The density f of the log-return is obtained computing the Fast Fourier Transform (FFT) of 4.7. Let S_n denote the price of the underlying at time $n\Delta$, i.e. $S_n = S(n\Delta)$; the pay-off of an arithmetic Asian option is given by

$$\text{Payoff} = (I_M - cS_M)^+, \text{ where } I_M = \sum_{n=0}^{M} \lambda_n S_n.$$

The following recursion holds for the option price:

$$V(S_M, I_M, M) = (I_M - cS_M)^+$$

$$V(S_n, I_n, n) = e^{-r\Delta} \int_{-\infty}^{\infty} f(s)\, V(S_n e^s, I_n + \lambda_{n+1} S_n e^s, n+1)\, ds;$$

$$n = M-1, \ldots, 0. \tag{4.8}$$

The expiry date T is modelled as a random variable distributed as geometric of parameter q; if one defines

$$H(x,q) := (1-q)\sum_{k=0}^{\infty} q^k v(x,k) \tag{4.9}$$

with $v(x,k) := V(1,x,M-k)$, the option price is given by $S_0 v(\lambda_0, M)$, where we have set $x = I_n/S_n$.

By suitable choices of values of λ_n and c, we can describe a wide class of Asian options.

Floating Strike Call Options Standard case for Floating Strike call options:

$$c = -1,\ \lambda_0 = \frac{\gamma}{M+\gamma};\ \lambda_n = \lambda = -\frac{1}{M+\gamma},\ n = 1,\ldots,M; \tag{4.10}$$

$$\gamma = \begin{cases} 1, & \text{if } S_0 \text{is included in the average} \\ 0, & \text{otherwise} \end{cases}$$

$$H(x,q) = q\int_{-\infty}^{\lambda} K(x,y)H(y,q)dy + (1-q)\phi(x) \tag{4.11}$$

where

$$K(x,y) = -e^{-r\Delta}\, f(\log(\frac{x}{y-\lambda}))\,\frac{x}{(y-\lambda)^2} \quad \text{and} \quad \phi(x) = (x-c)^+$$

Fixed Strike Call Options Standard case for Fixed Strike call options:

$$c = 0, \lambda_0 = \frac{\gamma}{M+\gamma} - \frac{K}{S_0}; \lambda_n = \lambda = \frac{1}{M+\gamma}, n = 1,\ldots,M; \tag{4.12}$$

γ is set as in the floating strike case, For Fixed Strike options, the value of $v(x,k), k = 1,\ldots,M$, is analytically known for $x \geq 0$ and so the integral equation becomes

$$H(x,q) = q\int_{-\infty}^{0} K(x,y)H(y,q)dy + (1-q)\tilde{\phi}(x,q) \tag{4.13}$$

with

$$\tilde{\phi}(x,q) = \phi(x) + \frac{q}{1-q}\int_{0}^{\lambda} K(x,y)H(y,q)dy$$

where if $y \geq 0$,

$$H(y,q) = \frac{y(1-q)}{1-qe^{-r\Delta}} + \frac{e^{(r-d)\Delta}(1-q)}{(M+\gamma)(1-e^{(r-d)\Delta})} x(\frac{1}{1-qe^{-r\Delta}} - \frac{1}{1-qe^{-d\Delta}})$$

Applying a quadrature rule, with N nodes x_i and weight w_i, to it, we obtain the linear system

$$(I - qKD)h = b, \tag{4.14}$$

where I is the identity matrix and for $i, j = 1, \ldots, N$, the vector and matrices elements are given by

$$\begin{aligned} h(i) &= H(x_i, q) \\ K(i,j) &= K(x_i, x_j) \\ b(i) &= (1-q)\Phi(x_i, q) \\ D(i,i) &= w_i, D(i,j) = 0 if i \neq j \end{aligned}$$

with

$$\Phi(x_i, q) = \begin{cases} \phi(x_i), & \text{for Floating strike Asian options} \\ \tilde{\phi}(x_i, q), & \text{for Fixed strike Asian options.} \end{cases}$$

System 4.14 is the main computational kernel in the algorithm.

The option price is recovered de-randomizing the option maturity, that is, exploiting the complex inversion integral

$$v(\lambda_0, M) = \frac{1}{2\pi\rho^M} \int_0^{2\pi} \frac{H(\lambda_0, \rho e^{is})}{1-\rho e^{is}} e^{-iMs} ds. \tag{4.15}$$

In particular, we approximate numerically 4.15 using:

$$v_h(\lambda_0, M) = \frac{1}{2M\rho^M}(\frac{H(\lambda_0, \rho)}{1-\rho} + (-1)^M \frac{H(\lambda_0, -\rho)}{1+\rho} + 2\sum_{j=1}^{M-1}(-1^j)Re(\frac{H(\lambda_0, \rho e^{ij\pi/M})}{1-\rho e^{ij\pi/M}})), \tag{4.16}$$

where ρ is set to $10^{-4/M}$.

The procedure involves the following steps:

- solve 4.11 for $q = q_j = \rho e^{ij\pi/M}, j = 0, \ldots, M$;

Procedure Asian_floating_pricing
1 : compute K, D and b
2 : for $j = 0, ..., \min(M, n_e + m_e)$ do
    solve $(I - q_j KD)h_j = b$
  end
3 : reconstruct $v(\lambda_0, M)$
End Asian_floating_pricing

Procedure Asian_fixed_pricing
1 : compute K and D
2 : for $j = 0, ..., \min(M, n_e + m_e)$ do
    compute $b = b(q_j)$
    solve $(I - q_j KD)h_j = b$
  end
3 : reconstruct $v(\lambda_0, M)$
End Asian_fixed_pricing

Figure 4.1: Asian call randomization pricing algorithm. Up: Floating Strike options; down: Fixed Strike options.

- approximate $v(\lambda_0, M)$ by $v_h(\lambda_0, M)$ as in 4.16.

Rewrite 4.16 as

$$v_h(\lambda_0, M) = \frac{1}{\rho^M} \sum_{j=0}^{M} (-1^j) a_j H(\lambda_0, \rho e^{ij\pi/M}). \tag{4.17}$$

To compute this, we use the Euler summation technique, a convergence acceleration technique for evaluating alternating series, as folllows:
Fix two positive integers m_e, n_e. Then 4.17 can be approximated by

$$\tilde{v}(\lambda_0, M) \approx \frac{1}{2^{m_e}\rho^M} \sum_{j=0}^{m_e} \binom{m_e}{j} b_{n_e+j}(\lambda_0, M), \tag{4.18}$$

where

$$b_k(\lambda_0, M) = \sum_{j=0}^{k} (-1)^j a_j H(\lambda_0, \rho e^{ij\pi/M}).$$

When the number of monitoring dates $M > m_e + n_e$, instead of 4.16 which involves solving $M + 1$ linear systems, we use the acceleration technique and evaluate $H(\lambda_0, \rho e^{ij\pi/M})$ for $j = 0, \ldots, n_e + m_e$, thus solving $n_e + m_e + 1$ systems.
Figure 4.1 shows a sketch of the pricing algorithm.

4.3.3 Discrete Wavelet Transform

The wavelets used are from [Dau88].

Recall the portion on MRA 4.1.3. The resolution index j of each subspace V_j in the MRA must be intended as a scale: the higher the scale, the more accurate the approximation $P_j f$. An element of a MRA can be viewed as a screen with a certain resolution: the successive element in the sequence could then be a screen with twice number of pixels along each dimension. The tool for moving between resolutions is the DWT, which is introduced owing to the following relations:

$$\varphi(x) = \sum_{k \in \mathbb{Z}} h_k \varphi(2x - k)$$

$$\psi(x) = \sum_{k \in \mathbb{Z}} g_k \psi(2x - k)$$

called *refinement equation* and *wavelet equation*, respectively. The sequences $\mathbf{h} := \{h_k\}_{k \in \mathbb{Z}}$ and $\mathbf{g} := \{g_k\}_{k \in \mathbb{Z}}$ are usually referred to as **filters** of the MRA. Now, if

$$P_{l+1} f = \sum_{k \in \mathbb{Z}} c_{l+1,k} \varphi_{l+1,k}$$

$$P_l f = \sum_{k \in \mathbb{Z}} c_{l,k} \varphi_{l,k}$$

$$Q_l f = \sum_{k \in \mathbb{Z}} d_{l,k} \psi_{l,k}$$

are the projections of f on V_{l+1}, V_l and W_l, respectively, then it holds:

$$c_{l,k} = \sum_{n \in \mathbb{Z}} h_{n-2k} c_{l+1,n}, \, d_{l,k} = \sum_{n \in \mathbb{Z}} g_{n-2k} c_{l+1,n} \tag{4.19}$$

and

$$c_{l+1,k} = \sum_{n \in \mathbb{Z}} h_{2k-n} c_{l,n} + \sum_{n \in \mathbb{Z}} g_{2k-n} d_{l,n} \tag{4.20}$$

Relations 4.19 and 4.20 ensure that to move between different levels of resolution in MRA, we only need to know the filters of the MRA.

Let c_{l+1} be the set $\{c_{l+1,k}\}_{k \in \mathbb{Z}}$. Then, the operator defined by 4.19 is the DWT applied to c_{l+1}. We compute c_l convolving h with c_{l+1}: c_l contains the coefficients of

the projection of the function at a lower resolution (scaling coefficients). d_l contains the wavelet coefficients, which retain the information that is lost when moving from resolution $l+1$ to resolution l.
So, if we neglect wavelet coefficients under a fixed threshold (Hard Threshold technique (HT), [Mal99]), accuracy can be preserved with a significant gain in efficiency.

In matrix form, if

$$L = (\overline{h}_{i,j} = h_{j-2i}), H = (\overline{g}_{i,j} = g_{j-2i})$$

relations 4.19 can be written as

$$\begin{pmatrix} c_{l-1} \\ d_{l-1} \end{pmatrix} = \begin{pmatrix} \mathrm{L} \\ \mathrm{H} \end{pmatrix} \cdot c_l \quad \Leftrightarrow \quad \begin{cases} c_{l-1} = \mathrm{L}c_l \\ d_{l-1} = \mathrm{H}c_l \end{cases}$$

The DWT can be applied recursively; at each step, only the scaling coefficients resulting from the previous step are transformed.
To expand the kernel of the integral equations in wavelet bases, we introduce the bi-dimensional DWT. For this, recursively define the following matrices:

$$Q^{(1)} = \begin{pmatrix} L \\ H \end{pmatrix}, Q^{(k)} = \begin{pmatrix} Q^{(k-1)} & I \\ I & I \end{pmatrix}, k \geq 2$$

and let

$$Q_s := \prod_{i=1}^{s} Q^{(i)}$$

then, the bi-dimensional DWT in s steps of a discrete operator A is defined as

$$A_s{}^W := Q_s A Q_s{}^T.$$

4.3.4 Wavelet-based pricing algorithm

We discretize the integrals in 4.13 and 4.11 by means of a quadrature rule on a truncated domain $[\mathcal{L}, \lambda]$ for Floating Strike Asian options, on $[\mathcal{L}, 0]$ for Fixed ones. Different values for $\mathcal{L}$ can be chosen according to the optimality criterion discussed in [FMM11] and tested.
Some of the following results are from [BCR91]. Let $K(x, y)$ be the kernel of an

integral operator and P the number of wavelet vanishing moments (i.e. $\int_{\mathbb{R}} x^p \psi(x)dx = 0; \quad p = 0, \dots, P-1$). Suppose a certain level in the MRA has been fixed. Then, denote with $\varphi_J, \psi_{J'}$ respectively the scaling and the wavelet function at the fixed level, having supports J, J'. The coefficients of the expansion of $K(x,y)$ are given by:

$$\alpha_{J,J'} = \int_{\mathbb{R}} \int_{\mathbb{R}} K(x,y)\psi_J(x)\psi_{J'}(y)dxdy$$

$$\alpha_{J,J'} = \int_{\mathbb{R}} \int_{\mathbb{R}} K(x,y)\psi_J(x)\phi_{J'}(y)dxdy$$

$$\gamma_{J,J'} = \int_{\mathbb{R}} \int_{\mathbb{R}} K(x,y)\phi_J(x)\psi_{J'}(y)dxdy$$

If partial derivatives of K up to order P exist on the square JxJ', then

$$\mid \alpha_{J,J'} \mid + |\beta_{J,J'}| + |\gamma_{J,J'}| \leq C \; |J|^{P+1} \; sup_{(x,y)\in J\times J'} \sum_j |\tfrac{\partial^P}{\partial x^j \partial y^{P-j}} K(x,y)|.$$

Thus, to have small wavelet coefficients, one should have a small RHS, which holds whenever either

1. $|J|$ is small (points out the importance of having wavelet bases with narrow supports); or
2. the derivatives are small (suggests the classes of integral operators for which a wavelet-based representation can be effective).

Recall the pricing procedure in 4.1: in the solution of the linear systems (step 2), we apply to both sides the DWT operator Q_s, for a fixed number of DWT steps s; for each value of q, we thus obtain the linear system

$$Q_s(I - qKD)h = Q_s b \quad \Leftrightarrow \quad (I - qQ_s(KD){Q_s}^T)Q_s h = Q_s b$$

for the orthogonality of the operator Q_s. If we denote by KD^W, h^W, b^W the DWT of KD, h, b respectively, we have

$$(I - qKD^W)h^W = b^W. \tag{4.21}$$

We then apply a hard threshold to the coefficient matrix of 4.21, thus we actually solve the linear system

$$(I - q(KD^W)_\epsilon)y = b^W, \tag{4.22}$$

where $(KD^W)_\epsilon$ is the hard threshold of KD^W with threshold ϵ. Finally, the inverse DWT is applied to the solution y of 4.22, thus an approximation of h, ${Q_s}^T y$, is obtained.

Bibliography

[BCR91] Gregory Beylkin, Ronald Coifman, and Vladimir Rokhlin, *Fast wavelet transforms and numerical algorithms i*, Communications on pure and applied mathematics **44** (1991), no. 2, 141–183.

[Bla03] Christian Blatter, *Wavelets: a primer*, Universities Press, 2003.

[Buc12] Peter Buchen, *An introduction to exotic option pricing*, CRC Press, 2012.

[CMM15] S Corsaro, Daniele Marazzina, and Z Marino, *A parallel wavelet-based pricing procedure for asian options*, Quantitative Finance **15** (2015), no. 1, 101–113.

[CZ04] Jakša Cvitanić and Fernando Zapatero, *Introduction to the economics and mathematics of financial markets*, MIT press, 2004.

[Dau88] Ingrid Daubechies, *Orthonormal bases of compactly supported wavelets*, Communications on pure and applied mathematics **41** (1988), no. 7, 909–996.

[Dau92] Ingrid Daubechis, *Ten lectures on wavelets*, Siam, 1992.

[FMM11] Gianluca Fusai, Daniele Marazzina, and Marina Marena, *Pricing discretely monitored asian options by maturity randomization*, SIAM Journal on Financial Mathematics **2** (2011), no. 1, 383–403.

[HLMS09] Emmanuel Haven, Xiaoquan Liu, Chenghu Ma, and Liya Shen, *Revealing the implied risk-neutral mgf from options: The wavelet method*, Journal of Economic Dynamics and Control **33** (2009), no. 3, 692–709.

[LCMS19] Xiaoquan Liu, Yi Cao, Chenghu Ma, and Liya Shen, *Wavelet-based option pricing: An empirical study*, European Journal of Operational Research **272** (2019), no. 3, 1132–1142.

[Ma11] Chenghu Ma, *Advanced asset pricing theory*, vol. 2, World Scientific, 2011.

[Mal99] Stéphane Mallat, *A wavelet tour of signal processing*, Elsevier, 1999.

[PO18] Gemma Colldeforns Papiol and CW Oosterlee, *Wavelet approach in computational finance.*

[Rom04] Steven Roman, *Introduction to the mathematics of finance: from risk management to options pricing*, Springer Science & Business Media, 2004.

Index