UiO : Department of Mathematics
University of Oslo

Pricing Perpetual American Options with Linear Programming

Jan Aleksander Olsen Bakke

Master's Thesis, Spring 2018

This master's thesis is submitted under the master's programme *Lektorprogrammet*, with programme option *Mathematics*, at the Department of Mathematics, University of Oslo. The scope of the thesis is 30 credits.

The front page depicts a section of the root system of the exceptional Lie group E_8, projected into the plane. Lie groups were invented by the Norwegian mathematician Sophus Lie (1842–1899) to express symmetries in differential equations and today they play a central role in various parts of mathematics.

Abstract

This master thesis will demonstrate how to price perpetual American options with linear programming. American options are used both for hedging and speculation, and being able to price derivatives, without creating arbitrage opportunities, are of importance. First we introduce a deterministic security market model and exploit the mathematical structure. Then European and American put and call options are presented. With dynamic programming we show how to price American options. Dynamic programming is based on an idea that an investor would act optimally at all trading dates and the objective is yielding the maximum profit, despite the risk of not knowing the true future value of the option. With this technique, we investigate perpetual American options on a ternary Markov chain model. Perpetual options are without an expiration date. Markov chain models are only dependent of the current state when determining the future value, thus simplifying the computations. The solution, based on dynamic programming, is the smallest payoff that is greater than the discounted expected value of the option at the next trading date. The value and the payoff must not be confused, as an investor may be willing to pay more than the payoff today, if the value of the option might rise in the future. The solution is obtained by formulating the problem as an optimization problem and then using linear programming theory.

Acknowledgements

I would like to express my gratitude to Professor Kenneth Hvistendahl Karlsen, my supervisor, for his useful critiques and guidance and our regular meetings. I would also like to thank Karoline Moe, for her LaTeX expertise. The advise and assistance has truly been helpful. I am particularly grateful for the support and encouragement from my mom and dad, not just this master thesis, but throughout my studies. Finally, I wish to thank my girlfriend Christin for her support, both home and at the University. She has made this process interesting and fun.

Contents

List of Figures

List of Tables

1 Introduction

In this thesis, we will price perpetual American options based on a ternary Markov chain model. In chapter 1, we introduce a general security market model, and focus on its mathematical structure. We will discuss topics such as martingales, viability and arbitrage.

In chapter 2 we introduce European and American put and call options. Our main concern is to find their fair prices. These options will be a natural part of the models introduced in chapter 1. Then, in chapter 3, we present the binomial model, and in chapter 4 the Markov models with the Markov chain property. These chapters shows that models with different properties have different outcomes when computing fair prices of options.

In chapter 5 we introduce theory on linear programming as this will be needed to solve the perpetual American options introduced in chapter 6. Chapter 5 include the strong and the weak duality theorems and the complementary slackness conditions.

First, in chapter 6, we present the result of Vanderbei & Pınar [7] when pricing perpetual American options based on a random walk model. We then proceed with a ternary Markov chain model. This part is the main result in this master thesis and is carefully demonstrated.

In chapter 7 there is a small discussion with suggestions for further studies and about what is not included in this master thesis.

1.1 Creating a Security Market Model

A *financial market* is a broad term describing a market in which people trade financial securities and commodities. They can be found in nearly every nation in the world, some are smaller and some trade billions of dollars daily.

Our security market models will be used to determine a fair price of plain vanilla options in discrete time. Once we have the necessary machinery, we'll look at the well-known binomial and Markov models.

The models we make in our text will not look much like the real world financial markets as many details are not included. Different borrowing and lending rates, transaction cost, transaction time, brokerage, and many other details are important and should be included when making a trading strategy in the real world. We will ignore these details in our models.

Even though we ignore certain details found in the real world, we want to identify the mathematical structure in the best possible way. This is especially noted when we talk about viable markets and arbitrage opportunities in later chapters.

1.2 One-period Single-stock Binary Model

Perhaps the simplest security market model is the one-period single-stock binary model. As the name indicates, it consist only of one risky security S and two moments in time. The risky security has a known value «today», mathematically written as S_0, and an unknown value S_1 «tomorrow». The name *binary* tells us that there are two possible scenarios ω for the development of the risky security, denoted u and d. We think of scenario u as when the value of the risky security goes up at time 1, and d as the scenario where the value goes down. But this does not need to be the case. Both $S_1(u)$ and $S_1(d)$ may be greater (or less) than S_0. Remark, $S_t(\omega)$ should be read as "the price of the risky security S at time t, given scenario ω". We will later see that some requirements on the relationship between the elements in the model are necessary in order to make sence with a real market model. There is also a risk-free asset B involved. It has an initial price B_0 and a fixed interest rate $r \in \mathbb{R}$, resulting in $B_1 = (1 + r)B_0$ for all ω.

Example 1.1 (One-period single-stock binary model)**.** Let the risk-free asset B be a bond with initial value of \$1000 and 6% interest rate. Further, let the risky security S be a stock with initial price of \$15 and terminal value of \$20 or \$10, depending on which of the scenarios u and d which turns out to be the real state of the world, respectively. This can be illustrated with a *tree-representation*, see Figure 1.1. Note that we have included a real-world probability $\mathbb{P}$. We will describe this in chapter 1.3.

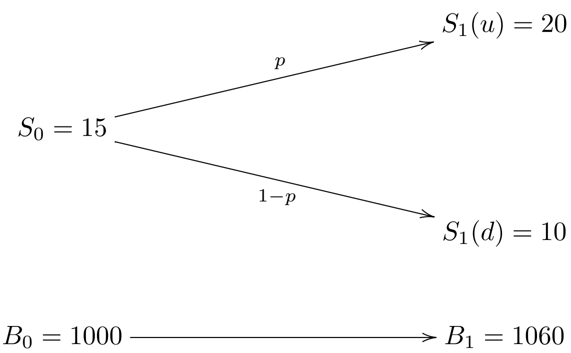

Figure 1.1: One-period single-stock binary model.

Although simple, the one-period single-stock binary model contains many desired properties that remain in more complex models. To avoid being lengthy, we skip these simple models and move straight to multi-period models. The theory developed for the multi-period models works perfectly with single period models.

1.3 Multi-period Security Models

The following multi-period models of securities markets are discrete. They are more realistic than the previous single period models and a step closer to a generalization to continuous time models. Based on the work of Pliska [5] and Cutland & Roux [3], the following elements are basic assumptions for the multi-period models:

- A fixed finite set of trading dates $t = 0, 1, \ldots, T$.
- A fixed finite sample space $\Omega = \{\omega_1, \ldots, \omega_K\}$. Each ω represent a possible scenario for the evolution of stock and bond prices, from time 0 to time T.
- A probability measure $\mathbb{P}$ on Ω with $\mathbb{P}(\omega) > 0$ for all $\omega \in \Omega$. This is the so-called *real-world probability.*
- A bank account process $\mathbb{B} = \{B_t\}_{t=0}^{T}$, where $\mathbb{B}$ is a stochastic process with $B_0 = 1$. The value $B_t(\omega)$ should be thought of as the time t value of a bond when one unit of the current medium of exchange is invested at time 0, in the case of scenario ω.
- A finite number N of risky security processes $S^n = \{S_t^n(\omega) : \omega \in \Omega, t$ is a trading date$\}$ for $n = 1, \ldots, N$. We think of $S_t^n(\omega)$ as the value of the risky security n at time t in scenario ω.
- A trading strategy process $\phi = \{\phi_t : t = 1, \ldots, T\}$. For any time intervall $(t-1, t]$ we have a trading strategy $\phi_t = \{H_t^0, H_t^1, \ldots, H_t^N\}$ that describes the investors portfolio as carried forward from time $t-1$ to time t. Here H^0 denotes the holding in the bank account process, and $\{H^n\}_{n=1}^{N}$ denotes the holdings in the N risky securities.
- A filtration $\mathbb{F} = \{\mathscr{F}_t : t = 0, 1, \ldots, T\}$, which is a submodel describing how the information about the prices of the risky securities are revealed for the investors as time progresses.

Remarks The probability measure $\mathbb{P}$ is not directly needed in the theory that follows, except we will need it for the modelling at the end. Nevertheless, the assumed property is an important one. It states that the probability of any scenario in Ω, $\mathbb{P}(\omega)$, is always greater than zero, together with

$$\sum_{\omega \in \Omega} \mathbb{P}(\omega) = 1.$$

This implies that all the impossible scenarios (zero-probability scenarios) are excluded from Ω, together with Ω including all possible scenarios.

The term *stochastic process* was used about the bank account process. This is a term we will use frequently, and as in Cutland & Roux [3, p. 142], we think of it as “a family of random variables indexed by time”.

Different assumptions about the bank account process can be made in a multi-period model. Consider $\mathbb{B}$ representing a savings account, you may expect to always have positive interest rate, i.e. $B_0 = 1$ and $B_t(\omega) > B_{t-1}(\omega)$ for all $\omega \in \Omega$ and for all $t = 1, \ldots, T$. An investment in bonds may return a higher yield, but you can also end up with loosing some of your investment, i.e. $B_t(\omega) \geq 0$. To simplify the examples that follows we assume the bank account process to be fixed.[1]

The trading strategy ϕ looks limited, it only contains $\mathbb{B}$ and $\mathbb{S}$. Later we introduce derivatives that will be an additional element in the portfolio.

The final basic assumption for the multi-period model, a filtration $\mathbb{F}$, needs a deeper explanation. The motivation is that an investor initially do not know which scenario $\omega \in \Omega$ will turn out to be the real state of the world at the final time T. The investor may want to adjust the trading strategy as time progresses and certain scenarios may not be available. At time t, only the market trend from time 0 and up to time t will be known. This idea, as stated by Cutland & Roux [3, p. 136], is "captured mathematically by means of the collection $\mathscr{F}_t$ of subsets of Ω for which membership depends only on the scenario ω up to time t".

Definition 1.2 (Filtration, [5, p. 73])**.** For any t, let $\mathscr{F}_t$ be the collection of all sets $A \subseteq \Omega$ with the property that, if $\omega \in A$ and there is another $\omega' \in \Omega$ such that the scenario ω' up to time t is equal to the scenario ω, then also $\omega' \in A$. The family $(\mathscr{F}_t)_{t=0}^T$ is called a filtration and is denoted $\mathbb{F}$.

From Definition 1.2 several properties, which are in line with what we expect to be reality, follows. The chosen theorem below is a desired property that shows that a filtration is a nested sequence in the way that as more time passes, more information is known to the investor.

Theorem 1.3 (Nested property, [3, p. 138])**.** *A filtration $\mathbb{F}$ satisfies the property $\mathscr{F}_t \subseteq \mathscr{F}_s$ whenever $t \leq s$.*

The proof of Theorem 1.3 is given by Springer [3, p. 138], and the following is a modified version.

Proof. Fix any $A \in \mathscr{F}_t$ and $\omega \in A$, and assume $t \leq s$. If the scenario ω up to time s is equal to another scenario ω' up to time s for some $\omega' \in \Omega$, then ω and ω' equals up to time t. By Definition 1.2 it follows that $\omega' \in A$. Thus $A \in \mathscr{F}_s$. □

Value and gain processes What we need next is to be able to convey the information the models provides. The assumed trading strategy ϕ gives rise to a value process $V = \{V_t : t = 0, 1, \ldots, T\}$, defined as in both the work of Pliska [5, p. 81] and Cutland & Roux [3, p. 98],

[1]Being fixed in this case does not mean that it can't change value over time, but that it uses the same value for all scenarios at each time step.

$$V_t = \begin{cases} H_1^0 B_0 + \sum_{n=1}^{N} H_1^n S_0^n, & t = 0, \\ H_t^0 B_t + \sum_{n=1}^{N} H_t^n S_t^n, & t = 1, \ldots, T. \end{cases} \tag{1.1}$$

The effect of different scenarios ω in (1.1) is not included, but clearly they matter in an actual calculation as can be seen in Example 1.5. Since the risky securities are random variables at each time step $\{(t-1, t] : t = 1, \ldots, T\}$, the value process V is a stochastic process.

Sometimes when we mention trading strategies, we want to add that they should be self-financing. This is to prevent meaningless counterarguments when developing the theory. For a trading strategy ϕ to be self-financing, it simply means that the value of ϕ before and after an adjustment, at time t, must be equal. Formally, as given by Cutland & Roux [3, p. 99], we write that ϕ is self-financing if and only if

$$V_{t+}^{\phi} = V_t^{\phi}, \tag{1.2}$$

where we define

$$V_{t+}^{\phi} \equiv H_{t+1} B_t + \sum_{n=1}^{N} H_{t+1}^n S_t^n.$$

The next concept is a gain process G. Let

$$G_{s \to t} \equiv \sum_{u=s}^{t} H_u^0 \Delta B_u + \sum_{n=1}^{N} \sum_{u=s}^{t} H_u^n \Delta S_u^n, \quad 1 \leq s < t. \tag{1.3}$$

Then $G_{s \to t}$ defines the cumulative gain from time s up to time t of a chosen portfolio.[2] The notation ΔS_t^n is defined as $S_t^n - S_{t-1}^n$, and similar for ΔB_t. We will later learn about adapted processes and $G = \{G_{(t-1) \to t} : t = 1, \ldots, T\}$ is in fact such a process. A more concise way of expressing the gain is seen in equation (1.3). It is similar without the discount as shown by Pliska [5, p. 82].

Discounting the effect of inflation and interest rates elucidates the return on risk and is commonly used in mathematical finance. For us, the movement of the security prices relative to each other are of interest, so it is convenient to normalize the prices in such a way that the bank account becomes the *numeraire*[3]. In fact, the security prices absolute value are of no interest, only their

[2]This is an example of a so-called *discrete time stochastic integral* [5, p. 81].

[3]A common example of a numeraire is setting oil prices in U.S. Dollars, allowing different countries to compare the value of oil prices in its own currency. Norway, a country that is a net exporter of oil, will for instance earn more (in local currency terms) than it did in the past, if its currency is weakening against the U.S. dollar.

relative behavior, and especially in relation to the bank account process. *Why take risk if you can be without!*

Definition 1.4 (Discounted price process, [5, p. 83])**.** The discounted price process $\bar{S}^n = \{\bar{S}^n_t : t = 0, 1, \ldots, T\}$ is defined by

$$\bar{S}^n_t \equiv S^n_t / B_t.$$

From Definition 1.4 both a discounted value process and a discounted gain process can be defined:

- The discounted value process: $\bar{V} = \{\bar{V}_t : t = 0, 1, \ldots, T\}$, as in Pliska [5, p. 83], where

$$\bar{V}_t \equiv V_t / B_t. \tag{1.4}$$

- The discounted gain process from time s up to time t is

$$\bar{G}_{s\to t} \equiv \sum_{n=1}^{N} \sum_{u=s}^{t} H^n_u \Delta \bar{S}^n_u,$$

where $\Delta \bar{S}^n_u := \bar{S}^n_u - \bar{S}^n_{u-1}$. Additionally, we let

$$\bar{G}_{s\to t} := \bar{V}_t - \bar{V}_s. \tag{1.5}$$

Equation (1.5) is an important result which is not necessarily true, but we define it to be so. Then any trading strategy is self-financing. To see this, recall (1.1), (1.3) and (1.4), and let

$$\begin{aligned}
\bar{G}_{s\to t} &= \bar{V}_t - \bar{V}_s \\
\bar{G}_{s\to t+1} - \bar{G}_{s\to t} &= \bar{V}_{t+1} - \bar{V}_t \\
\bar{V}_t &= \bar{V}_{t+1} + \bar{G}_{s\to t} - \bar{G}_{s\to t+1} \\
\bar{V}_t &= H^0_{t+1} + \sum_{n=1}^{N} H^n_{t+1} \bar{S}^n_{t+1} + \sum_{n=1}^{N} \sum_{u=s}^{t} H^n_u \Delta \bar{S}^n_u - \sum_{n=1}^{N} \sum_{u=s}^{t+1} H^n_u \Delta \bar{S}^n_u \\
\bar{V}_t &= H^0_{t+1} + \sum_{n=1}^{N} H^n_{t+1} \bar{S}^n_{t+1} - \sum_{n=1}^{N} H^n_{t+1} \Delta \bar{S}^n_{t+1} \\
\bar{V}_t &= H^0_{t+1} + \sum_{n=1}^{N} H^n_{t+1} \bar{S}^n_t.
\end{aligned}$$

Multiplying both sides of the last equation with B_t gives (1.2), the desired property of a self-financing trading strategy. We finalize this section with an example.

Example 1.5 (Multi-period multi-stock model)**.** Figure 1.2 is an example of a multi-period security market model. It consists of $N = 2$ risky securities, in this case two stocks, a fixed interest rate r of 4%, $T = 2$ time steps, and $|\Omega| = 9$, nine different possible scenarios. A key observation is that no real world probability measure $\mathbb{P}$ is mentioned. We will later see that this is not needed in our finite and deterministic models.

Table 1.3 lists the discounted price processes, which we will make use of later. A table is another way of representing a security market model, serving the same purpose as the tree-representation. Notice that at time 1, the scenario u is equivalent with either scenario ω_1, ω_2 or ω_3 being the true state of the world, and similar for scenario m and d.

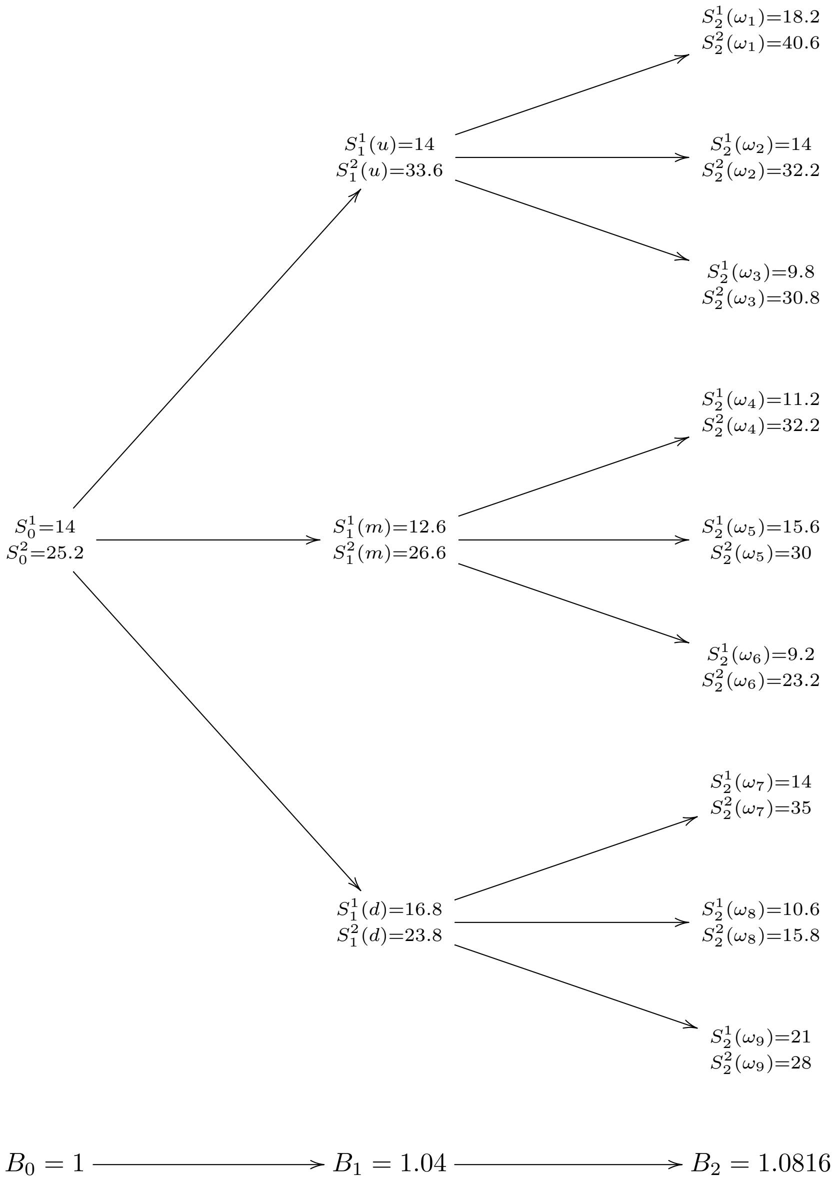

Figure 1.2: Two-step ternary branching model with two stocks.

ω	$\bar{S}_0^1$	$\bar{S}_0^2$	$\bar{S}_1^1(\omega)$	$\bar{S}_1^2(\omega)$	$\bar{S}_2^1(\omega)$	$\bar{S}_2^2(\omega)$
ω_1	14	25.2	$\frac{175}{13}$	$\frac{420}{13}$	$\frac{875}{52}$	$\frac{25375}{676}$
ω_2	14	25.2	$\frac{175}{13}$	$\frac{420}{13}$	$\frac{4375}{338}$	$\frac{20125}{676}$
ω_3	14	25.2	$\frac{175}{13}$	$\frac{420}{13}$	$\frac{6125}{676}$	$\frac{9625}{338}$
ω_4	14	25.2	$\frac{315}{26}$	$\frac{665}{26}$	$\frac{1750}{169}$	$\frac{20125}{676}$
ω_5	14	25.2	$\frac{315}{26}$	$\frac{665}{26}$	$\frac{375}{26}$	$\frac{9375}{338}$
ω_6	14	25.2	$\frac{315}{26}$	$\frac{665}{26}$	$\frac{2875}{338}$	$\frac{3625}{169}$
ω_7	14	25.2	$\frac{210}{13}$	$\frac{595}{26}$	$\frac{4375}{338}$	$\frac{21875}{676}$
ω_8	14	25.2	$\frac{210}{13}$	$\frac{595}{26}$	$\frac{6625}{676}$	$\frac{9875}{676}$
ω_9	14	25.2	$\frac{210}{13}$	$\frac{595}{26}$	$\frac{13125}{676}$	$\frac{4375}{169}$

Table 1.3: Discounted price processes of Example 1.2.

1.4 Arbitrage Opportunities

Our security market model aims to represent the mathematical structure of a real world market model. An *arbitrage opportunity* is a risk free way of making money and should therefore not be allowed. Definition 1.6 is a modified version of what given by Pliska [5, p. 92]. Note the importance of discounting. We are not thinking of money in the bank as a risk free way of gaining profit.[4]

Definition 1.6 (Arbitrage opportunity, [5, p. 92])**.** An arbitrage opportunity in a multi-period security market model is a self-financing trading strategy ϕ such that:

1. $\bar{V}_0^\phi = 0$,
2. $\bar{V}_T^\phi(\omega) \geq 0$ for all $\omega \in \Omega$,
3. $\bar{V}_T^\phi(\omega) > 0$ for at least one $\omega \in \Omega$.

Directly from Definition 1.6, the following can be said about arbitrage opportunities:

Theorem 1.7. *The existence of an arbitrage opportunity implies that an investor can create a portfolio which guarantees either a zero or a positive discounted gain.*

Proof. The proof given by Pliska [5, p. 9] is only for single-period models, but because of (1.5) Pliska show in [5, p. 92] that it's also true for multi-period models. Again, showing the importance of self-financing portfolios. □

[4] Definition 1.6 without discounting would be that any investment in the bank account is an arbitrage opportunity.

Arbitrage opportunities do exist in real world security markets.[5] Once they are founded, the balance between supply and demand are changed, as most investors would not let go of a *free lunch*[6]. Therefore, they are short-lived and will not be included in our models. It is worth mentioning that if you are lucky to find an arbitrage opportunity, friction will make it harder to exploit the opportunity as the margins often are relative small for private investors.

Example 1.8 (Example 1.5 continued)**.** Looking at Figure 1.2, it is not easy to spot an arbitrage opportunity - if it even exists. And how to prove arbitrage opportunities are non-existent? Assume for now that the value of S_0^1 is changed to $\mu \geq \max\{S_1^1(\omega) : \omega \in \Omega\}$. Then the portfolio $\phi_1 = (-x, 0, x\mu)$, for some constant x, has initial value 0, and time 1 value

$$V_1^{\phi_1} = \begin{cases} -xS_1^1(u) + x\mu \cdot B_1 > 0, & \omega = \omega_1, \omega_2 \text{ or } \omega_3, \\ -xS_1^1(m) + x\mu \cdot B_1 > 0, & \omega = \omega_4, \omega_5 \text{ or } \omega_6, \\ -xS_1^1(d) + x\mu \cdot B_1 \geq 0, & \omega = \omega_7, \omega_8 \text{ or } \omega_9. \end{cases} \tag{1.6}$$

Liquidating this portfolio at time 1, and then investing all in the bank account process until time 2 is clearly an arbitrage opportunity.

The negative holding, $-x$, in portfolio ϕ_1 is understood as short selling x units of the risky security S^1. We say that a model is *viable* when no arbitrage opportunities exist. The model in Example 1.8 is therefore not viable. To be able to classify and create viable market models are obviously of interest, and we will look into this in the next chapter.

1.5 Martingales

Martingales are important in several ways in multi-period security market models. The final result in this chapter shows that the existence of an equivalent martingale measure excludes the existence of an arbitrage opportunity. Later we will see that martingales are an important tool when pricing derivatives.

Before the main theorem for this chapter, we must extend the foundation we are working with. The first definition below associate a random variable X with the filtration $\mathbb{F}$ generated by the risky securities in the model.

Definition 1.9 (Measurability, [3, p. 140])**.** For any trading date t, a random variable X is said to be measurable with respect to $\mathscr{F}_t$ if $\{X \leq x\} := \{\omega \in \Omega | X(\omega) \leq x\} \in \mathscr{F}_t$ for every $x \in \mathbb{R}$. In short we say X is $\mathscr{F}_t$-measurable.

The next definition extend the concept of dependency on available information from single random variables to stochastic processes. This will ensure an investor the full knowledge about the past and present prices. The investor

[5] Topics that will not be discussed here: does there really exist completely risk-free investments? What is the limit between a risk-free and a risky investment?

[6] Investment slang term referring to arbitrage opportunities.

know at time t that the true state ω is contained in some subset $A \in \mathscr{F}_t$. From the definition of measurability the price of the risky security at time t, S_t, will be constant on this subset. In addition, since the filtration $\mathbb{F}$ consists of a nested sequence of subsets, the investor can at time t infer the observed subsets in earlier subsets $\mathscr{F}_s$, $s < t$, and thereby deduce the earlier security prices.

Definition 1.10 (Adaptedness, [3, p. 142])**.** A stochastic process $X = \{X_t\}_{t=0}^T$ is said to be adapted to the filtration $\mathbb{F}$ if the random variable X_t is $\mathscr{F}_t$-measurable for every t.

Definition 1.11 tell us that since we have assumed that the real-world probability $\mathbb{P}$ is greater than zero for all $\omega \in \Omega$, it follows that $\mathbb{Q}$ is equivalent to $\mathbb{P}$ if and only if $\mathbb{Q}(\omega) > 0$, for all $\omega \in \Omega$.

Definition 1.11 (Equivalent probability measures, [3, p. 143])**.** Two probability measures $\mathbb{Q}$ and $\mathbb{Q}'$ on Ω are called equivalent if, for any $A \subseteq \Omega$, we have $\mathbb{Q}(A) = 0$ if and only if $\mathbb{Q}'(A) = 0$.

We are ready to define a martingale process. Conditional expectation is then a key notation. For Definition 1.12 the conditional expectation under $\mathbb{Q}$ is such that for any time t scenario ω, the probability of $q_\omega \in \mathbb{Q}$ is the probability that a process (e.g. a risky security) reaches scenario ω at time t, conditional on the fact that it has reached an associated scenario of ω at an earlier time s.

Definition 1.12 (Martingale, [3, p. 152])**.** A process M is a martingale with respect to $\mathbb{Q}$ if it is adapted and $\mathbb{E}_{\mathbb{Q}}(M_t|\mathscr{F}_s) = M_s$ for every $s \leq t$.

We accumulates the different definitions above to define an equivalent martingale measure which we will need for the main theorem in this section[7].

Definition 1.13 (Equivalent martingale measure, [3, p. 154])**.** A probability measure $\mathbb{Q}$ on the set Ω is called an equivalent martingale measure if it is equivalent to the real-world probability $\mathbb{P}$ and the discounted stock price process $\bar{S}$ is a martingale under $\mathbb{Q}$.

Theorem 1.14 (Fundamental theorem of asset pricing)**.** *A multi-period security model with a finite number of trading dates, scenarios and risky securities is viable if and only if it admits an equivalent martingale measure.*

Proof. See Cutland & Roux [3, p. 159] for a proof. □

The theorem above is powerful as it provides a tool to show the existence (or non-existence) of arbitrage opportunities. Now we make use of another result that appear in the work of Cutland & Roux [3, p. 147]. It simply states that $\mathbb{Q}$ can be constructed by first calculating the conditional probabilities at each time steps at each scenarios. Then, the equivalent martingale measure is given for each scenario by multiplying all the conditional probabilities that belong to the same scenario, for all time steps. This provides a tool for calculating an equivalent martingale measure $\mathbb{Q}$ on Ω. Lets see this in an example:

[7] This theorem is sometimes called the *fundamental theorem of asset pricing.*

Example 1.15 (Example 1.5 continued)**.** We use Definition 1.13 together with Definition 1.12, and get that $\mathbb{Q}$ should satisfy $\mathbb{E}_{\mathbb{Q}}[\bar{S}_1|\mathscr{F}_0] = \bar{S}_0$ at time 0. That is, we need the conditional probabilities q_u, q_m and q_d to satisfy the system

$$\begin{aligned} q_u\bar{S}_1^1(u) + q_m\bar{S}_1^1(m) + q_d\bar{S}_1^1(d) &= \bar{S}_0^1, \\ q_u\bar{S}_1^2(u) + q_m\bar{S}_1^2(m) + q_d\bar{S}_1^2(d) &= \bar{S}_0^2, \\ q_u + q_m + q_d &= 1. \end{aligned} \tag{1.7}$$

Using Table 1.3 we get

$$\begin{aligned} \frac{175}{13}q_u + \frac{315}{26}q_m + \frac{210}{13}q_d &= 14, \\ \frac{420}{13}q_u + \frac{665}{26}q_m + \frac{595}{26}q_d &= 25.2, \\ q_u + q_m + q_d &= 1. \end{aligned}$$

This is routine to solve, and the solution is

$$(q_u, q_m, q_d) \approx (0.1154, 0.4564, 0.4282).$$

At the next trading date, $t = 1$, $\mathbb{Q}$ should satisfy $\mathbb{E}_{\mathbb{Q}}[\bar{S}_2|\mathscr{F}_1] = \bar{S}_1$. There are three separated scenarios, and we must solve for each one of them. For scenario u, we need a solution $(\omega_1, \omega_2, \omega_3)$ to the system

$$\begin{aligned} q_{\omega_1}\bar{S}_2^1(\omega_1) + q_{\omega_2}\bar{S}_2^1(\omega_2) + q_{\omega_3}\bar{S}_2^1(\omega_3) &= \bar{S}_1^1(u), \\ q_{\omega_1}\bar{S}_2^2(\omega_1) + q_{\omega_2}\bar{S}_2^2(\omega_2) + q_{\omega_3}\bar{S}_2^2(\omega_3) &= \bar{S}_1^2(u), \\ q_{\omega_1} + q_{\omega_2} + q_{\omega_3} &= 1. \end{aligned}$$

We get similar systems for scenarios m and d. Solving these yields

$$\begin{aligned} (q_{\omega_1}, q_{\omega_2}, q_{\omega_3}) &\approx (0.3653, 0.4027, 0.2320), \\ (q_{\omega_4}, q_{\omega_5}, q_{\omega_6}) &\approx (0.0460, 0.5956, 0.3584), \\ (q_{\omega_7}, q_{\omega_8}, q_{\omega_9}) &\approx (0.0575, 0.3005, 0.6420). \end{aligned}$$

Finally, we can use the conditional probabilities above to create an equivalent martingale measure $\mathbb{Q}$:

$$\begin{aligned}
\mathbb{Q}(\omega_1) &= q_u q_{\omega_1} \approx 0.0422,\\
\mathbb{Q}(\omega_2) &= q_u q_{\omega_2} \approx 0.0464,\\
\mathbb{Q}(\omega_3) &= q_u q_{\omega_3} \approx 0.0268,\\
\mathbb{Q}(\omega_4) &= q_m q_{\omega_4} \approx 0.0210,\\
\mathbb{Q}(\omega_5) &= q_m q_{\omega_5} \approx 0.2718,\\
\mathbb{Q}(\omega_6) &= q_m q_{\omega_6} \approx 0.1636,\\
\mathbb{Q}(\omega_7) &= q_d q_{\omega_7} \approx 0.0246,\\
\mathbb{Q}(\omega_8) &= q_d q_{\omega_8} \approx 0.1287,\\
\mathbb{Q}(\omega_9) &= q_d q_{\omega_9} \approx 0.2749.
\end{aligned} \tag{1.8}$$

From Theorem 1.14 above, since our model admits an equivalent martingale measure, no arbitrage opportunities exists. Lets do (1.7) again, but now for Example 1.8, which we remember contains arbitrage opportunities:

$$\begin{aligned}
\frac{175}{13}q_u + \frac{315}{26}q_m + \frac{210}{13}q_d &= \mu,\\
\frac{420}{13}q_u + \frac{665}{26}q_m + \frac{595}{26}q_d &= 25.2,\\
q_u + q_m + q_d &= 1.
\end{aligned}$$

Solving this system yields

$$(q_u, q_m, q_d) \approx (0.0874\mu - 1.1082, -0.3059\mu + 4.7388, 0.2185\mu - 2.6306).$$

Clearly, there does not exists any value of μ such that both $\mathbb{Q}(\omega) > 0$ and $\mu \geq \max\{S_1^1(\omega) : \omega \in \Omega\}$, for all $\omega \in \Omega$. Again, showing that the modified model in Example 1.8 contains arbitrage opportunities.

2 Introducing Derivatives

We are ready to introduce another element to our multi-period security market model - derivative securities. The theory already developed is consistent with this new financial asset, given that it may be included in any trading strategy ϕ. A derivative is a random variable that can be taken to be a function of one or more underlying security prices[8] [5, p. 112]. Derivatives describe a broad

[8]From an investors point of view a derivative is a contract between two parties. One party is the seller (or issuer) of the derivative, which promise to pay the other party, the buyer (or holder) of the derivative the amount X when the derivative is exercised.

range of securities and include options, futures, swaps, and forward contracts [5]. An option gives the holder the right, but not the obligation, to exercise the derivative[9].

The so-called *vanilla options* that we will look into have always non-negative payoff. The fundamental question to be addressed is: what is the appropriate value for this agreement? That is, if the buyer acquires this asset at time t, and is expecting a payoff $X \geq 0$ at a later time T, then the investor would be expecting to pay something at time t. The first derivatives we will investigate is the simple European call and put options. As explained by Cutland & Roux [3, p. 21], a European call option C^E on a security S takes form of a binding agreement that entitles its owner to buy 1 unit of S at a fixed strike price K at a fixed time T in the future. The holder is not obliged to exercise the option. Thus,

- if $K > S_T$ the investor will buy the security in the market rather than exercise the derivative.
- If $K < S_T$ it makes sense to exercise the option as selling the acquired security will yield a net profit of $S_T - K > 0$.
- If $K = S_T$ it doesn't matter if the investor exercise or not. In a real world market it would depend on which alternative that have less friction, in order to acquire the security.[10]

Summarizing, a European call option has payoff

$$(S_T - K)^+ \equiv \max\{S_T - K, 0\},$$

which at time $t < T$ is well defined, but unknown. A European put option P^E is similar to the call option, except it gives the right to sell and not to buy one unit of the underlying asset. Thus it has payoff

$$(K - S_T)^+ \equiv \max\{K - S_T, 0\}.$$

We assume our markets to be viable, so that none combinations of the risk free and the risky securities can create an arbitrage opportunity. However, the challenge of derivatives is that we, in one way, must decide the price. The chief topic of the rest of the theory presented in this text is to find what we call a fair price of a given derivative. A price is called *fair* if it guarantees that neither the buyer nor the seller can gain an arbitrage opportunity, even though the future is unknown. A price is called *unfair* if it creates arbitrage opportunities. The remarkable theory of derivative pricing shows that in many situations there exists prices for large classes of derivatives that are fair [3, p. 3].

[9] To be clear, an option is a derivative, but a derivative is not necessary an option.
[10] Friction, as in transaction cost, transaction time, etc.

2.1 Fair Pricing of European Options

When we are to find fair prices of European options, there are two types to consider: the Europen call option and the European put option. Although these two options are defined in two different ways, our results will apply to both, plus a wide range of other options, see Definition 3.2. But, the derivatives we will look at afterwards, American options, need a new mindset.

Theorem 2.1 (Fair pricing of European options, [3, p. 168])**.** *Let D be any European option in a viable multi-period model. Then $F_D = \{\mathbb{E}_{\mathbb{Q}}(\bar{D}) : \mathbb{Q}$ is an equivalent martingale measure$\}$ is a non-empty interval of fair prices of D.*

Proof. Suppose that $\pi = \mathbb{E}_{\mathbb{Q}}(\bar{D})$ is a price for D at time 0. Take the pricing structure $(\pi_t)_{t=0}^{T}$ for D given by $\pi_t = (1+r)^{t-T}\mathbb{E}_{\mathbb{Q}}[D|\mathscr{F}_t]$ for all t. Clearly, from the just given definition of π, we have $\pi_0 = \pi$ and $\pi_T = D$. If we regard D as an additional asset in the model, then $\mathbb{Q}$ is an equivalent martingale measure for this extended model, so that it is viable. Thus, there can be no arbitrage opportunity involving $(\pi)_{t=0}^{T}$, and so π is a fair price for D at time 0. For the converse, suppose that π is a fair price for D. Thus there exists some pricing structure $(\pi)_{t=0}^{T}$ with $\pi_0 = \pi$ for which there is no arbitrage opportunity. Thus the extended model with the new asset D with pricing structure $(\pi)_{t=0}^{T}$ is viable, and admits an equivalent martingale measure $\mathbb{Q}$. The discounted stock and bond price processes are martingales with respect to $\mathbb{Q}$, so it is also an equivalent martingale measure for the basic model without the derivative D. At the same time we have $\mathbb{E}_{\mathbb{Q}}[\bar{D}] = \mathbb{E}_{\mathbb{Q}}[\bar{\pi}_T] = \bar{\pi}_0 = \pi_0 = \pi$, which concludes the proof of the first part of the theorem. A similar version of whats written above is done by Cutland & Roux [3, p. 168].

For the second part, assume both $\mathbb{Q}_1$ and $\mathbb{Q}_2$ belong to the set of all equivalent martingales measures. Then $\lambda\mathbb{Q}_1(\omega) + (1-\lambda)\mathbb{Q}_2(\omega) > 0$ for all $\omega \in \Omega$, and

$$\lambda \sum_{\omega\in\Omega} \mathbb{Q}_1(\omega) + (1-\lambda) \sum_{\omega\in\Omega} \mathbb{Q}_2(\omega) = 1.$$

Also, for any scenario ω at every trading date $t < T$ we have

$$\begin{aligned}
\mathbb{E}_{\lambda\mathbb{Q}_1+(1-\lambda)\mathbb{Q}_2}[\bar{S}_{t+1}|\mathscr{F}_t] &= \mathbb{E}_{\lambda\mathbb{Q}_1}[\bar{S}_{t+1}|\mathscr{F}_t] + \mathbb{E}_{(1-\lambda)\mathbb{Q}_2}[\bar{S}_{t+1}|\mathscr{F}_t] \\
&= \lambda\mathbb{E}_{\mathbb{Q}_1}[\bar{S}_{t+1}|\mathscr{F}_t] + (1-\lambda)\mathbb{E}_{\mathbb{Q}_2}[\bar{S}_{t+1}|\mathscr{F}_t] \\
&= \lambda\bar{S}_t + (1-\lambda)\bar{S}_t \\
&= \bar{S}_t.
\end{aligned}$$

Thus showing that the set of all equivalent martingale measures is convex and the expectation $\mathbb{E}_{\mathbb{Q}}$ depends linearly on $\mathbb{Q}$. We conclude that F_D is a convex set of $\mathbb{R}$, and must be an interval. □

Let's see how this result works out in an example.

Example 2.2 (European put option)**.** Consider an European put option P^E with exercise date $T = 2$ and strike price $K = 12$ in the model in Example 1.5, with S^1 as the underlying asset. Its payoff $P_2^E = (K - S_2^1)^+$ satisfies

$$\begin{aligned}
&P_2^E(\omega_1) = P_2^E(\omega_2) = P_2^E(\omega_5) = P_2^E(\omega_7) = P_2^E(\omega_9) = 0,\\
&P_2^E(\omega_3) = 2.2,\\
&P_2^E(\omega_4) = 0.8,\\
&P_2^E(\omega_6) = 2.8,\\
&P_2^E(\omega_8) = 1.4.
\end{aligned}$$

For the unique equivalent martingale measure $\mathbb{Q}$ found in (1.8) we have

$$\begin{aligned}
\mathbb{E}_{\mathbb{Q}}[P_2^E] &= \tfrac{1}{1.0816}\big(2.2\mathbb{Q}(\omega_3) + 0.8\mathbb{Q}(\omega_4) + 2.8\mathbb{Q}(\omega_6) + 1.4\mathbb{Q}(\omega_8)\big)\\
&\approx \tfrac{1}{1.0816}\big(2.2 \cdot 0.0268 + 0.8 \cdot 0.0210 + 2.8 \cdot 0.1636 + 1.4 \cdot 0.1287\big)\\
&\approx 0.6602.
\end{aligned}$$

The fair price to pay for P^E is ≈ 0.6602 at time 0. To find the fair price of P^E at time 1 is similar as the computation above, only now we solve the expectation conditioned that the history up to time 1 is known:

$$\mathbb{E}_{\mathbb{Q}}[P_2^E|\mathscr{F}_1] = \begin{cases} P_2^E(\omega_3)q(\omega_3) \approx 0.5104, & \omega = u,\\ P_2^E(\omega_4)q(\omega_4) + P_2^E(\omega_6)q(\omega_6) \approx 1.0403, & \omega = m,\\ P_2^E(\omega_8)q(\omega_8) \approx 0.4207, & \omega = d. \end{cases}$$

Any other prices will give an arbitrage opportunity. We can see this with an example. Suppose the price of P^E at time 1 and scenario m is decreased from 1.0403 to 0.7. Then in case of scenario m, buying 3 derivatives and 1 unit of security S^1 together with shorting ≈ 27.5962 units of B_1, all in time 1, yields a portfolio with initial value 0 and final value

$$V_2 \approx \begin{cases} 3 \cdot P_2^E(\omega_4) + S_2^1 - 27.5962B_2 > 0, & \omega = \omega_4,\\ S_2^1 - 27.5962B_2 > 0, & \omega = \omega_5,\\ 3 \cdot P_2^E(\omega_6) + S_2^1 - 27.5962B_2 > 0, & \omega = \omega_6. \end{cases}$$

This is clearly an arbitrage opportunity. Therefore, 0.7 is not a fair price at time 1 in case of scenario m for the derivative P^E.

2.2 Fair Pricing of American Options

American options, studied in this chapter, are like the European options except the owner is allowed to exercise[11] at any time up to and included the fixed expiry time T. Thus, the payoff Y may depend not only on the scenario $\omega \in \Omega$, but also on the exercise time $t \leq T$. As in Cutland & Roux [3, p. 211], we can represent this with a non-negative adapted stochastic process $Y = \{Y_t : t = 0, 1, \ldots, T\}$. The payoff is non-negative as the holder of the option can always choose to never exercise the option, yielding a zero payoff. It is assumed to be adapted because when deciding whether or not to exercise the option, in scenario ω, at a given time t, the payoff is assumed to be known for the investor. But, we do not assume it is possible to know the true future payoff after time t. Therefor, Y_t depend only on the history of scenario ω up to time t.

The payoff function Y may fail to represent the value Z of the American option. If the expected future time τ discounted payoff where higher than the payoff today,

$$\mathbb{E}[\bar{Y}_\tau] > Y_t,$$

then you may be willing to pay more than the actual payoff today, Y_t. We will define the value process $Z = \{Z_t : t = 0, 1, \ldots, T\}$ for the American option, but first some definitions are needed.

Definition 2.3 (Supermartingales, [5, p. 127])**.** An adapted stochastic process $X = \{X_t : t = 0, 1, \ldots, T\}$ is said to be a supermartingale if $\mathbb{E}[X_t | \mathscr{F}_t] \leq X_s$, for $0 \leq s \leq t \leq T$.

Obviously, all martingales are supermartingales, but not vice versa. We know from Theorem 2.1 that the discounted value of an European option is a martingale under a risk neutral probability measure. Theorem 2.5 states that the discounted value process Z of an American option is a supermartingale under the same measure.

Definition 2.4 (Stopping times, [5, p. 127])**.** A stopping time is a random variable τ taking values in the set $\{0, 1, \ldots, T, \infty\}$ such that $\{\tau = t\} \in \mathscr{F}_t$, for all $t = 0, 1, \ldots, T$. If an event never occurs, stopping times are allowed to take the value ∞.

Stopping times can be hard to grasp, but think of them as random variables whose value is interpreted as the time at which a given stochastic process exhibits a certain behavior of interest. Taking the example of Pliska [5, p. 127], for a risky security with $S_0 = 10$, $\tau_1 \equiv \min\{t : S_t \geq 20\}$ is a stopping time, because you learn the event $\{\tau_1 = t\}$ by time t. However, the random variable $\tau_2 \equiv \max\{t : S_t \geq 20\}$ is not a stopping time, because you may not learn whether $\{\tau_2 = t\}$ until time T. There are many stopping times associated with

[11] That is, to take the payoff.

our security model's filtration. Let the set of random variables which are stopping times taking finite values in the closed interval $[s, t]$, to be denoted by the set $\zeta(s, t)$.

Theorem 2.5 (Value process for America options, [5, p. 127])**.** *The value process Z for the American option Y is given by*

$$Z_t = \max_{\tau \in \zeta[t,T]} \mathbb{E}_{\mathbb{Q}}[Y_\tau B_t / B_\tau | \mathscr{F}_t]. \tag{2.1}$$

Also, the discounted process Z/B is the smallest supermartingale under a risk neutral probability measure satisfying $Z_t \geq Y_t$, for all t, ω. Moreover, the optimal stopping time that maximizes (2.1) *is $\tau(t) \equiv \{s \geq t : Z_s = Y_s\}$ for $t = 0, 1, \dots, T$.*

Proof. See Pliska [5, pp. 127–131] for a proof. □

With Theorem 2.5 we are able to define a fair price of the American option, but we don't have any methods to compute it. Springer [3, pp. 258, 263] shows that in a viable multi-period model, for the American option X,

$$F_X = \{\sup_\tau \mathbb{E}_{\mathbb{Q}}(\bar{X}_\tau) : \mathbb{Q} \text{ is an equivalent martingale measure}\} \tag{2.2}$$

is a non-empty interval of fair prices of X. Later we will look at options without an expiration date in models with an infinite number of trading dates. Then (2.2) will no longer work. We therefore introduce dynamic programming, a topic of importance in its own right. At first we will use it to compute the value process Z of the American options presented above.

The idea, as explained in Pliska [5, p.128], is to work backwards in time, and we will justify our choices on the way. Clearly at time T, $Z_T = Y_T$. If the option is not exercised, it expires and have value 0. We remember that $Y_T \geq 0$ and in order to act optimally we set $Z_T = Y_T$.

At time $T - 1$ we do the first iteration:

$$Z_{T-1} \equiv \max\{Y_{T-1}, \mathbb{E}_{\mathbb{Q}}[Z_T B_{T-1} / B_T | \mathscr{F}_{T-1}]\}. \tag{2.3}$$

That is, our choice depends on which value is greater: exercising now, or the expected value of the option at the next trading time. Since we have already calculated Z_T and assumed that there exists an equivalent martingale measure $\mathbb{Q}$, this equation is well defined.

We continue these iterations until time 0. Then we will know the value process Z for the American option at all times and scenarios. There is also a bonus. Let $\tau = t$ at each time $Z_t = Y_t$. This gives the optimal stopping time in Theorem 2.5. Of course, we must take into account the different possible scenarios. Lets see this in an example.

Example 2.6 (Pricing American options with dynammic programming)**.** Consider the American call option C^A with exercise date $T = 2$, strike price $K = 18$, and underlying security S^1 in the model in Example 1.5. This options has payoff $Y_t = (S^1_t(\omega) - K)^+$, for $t = 0, 1, 2$ and for all $\omega \in \Omega$.

We want to determine what an investor should expect to pay for this option. That is, to determine the value process Z for C^A. As explained above, at time $T = 2$ we set

$$Z_2 = Y_2 = \begin{cases} 0.2, & \omega_1, \\ 0, & \omega_2, \omega_3, \omega_4, \omega_5, \omega_6, \omega_7, \omega_8, \\ 3, & \omega_9. \end{cases}$$

The value at time 2 is now known, and at time 1 we compute

$$\begin{aligned} Z_1 &= \max\left\{Y_1, \mathbb{E}_{\mathbb{Q}}[Z_2 B_1/B_2 | \mathscr{F}_1]\right\} \\ &= \begin{cases} \max\{0, \frac{\omega_1}{q_{\omega_u}} \cdot \frac{B_1}{B_2} \cdot Z_2(\omega_1)\} \approx 0.1022, & \omega = u, \\ \max\{0, 0\} = 0, & \omega = m, \\ \max\left\{0, \frac{\omega_9}{q_{\omega_d}} \cdot \frac{B_1}{B_2} \cdot Z_2(\omega_9)\right\} \approx 1.7764, & \omega = d. \end{cases} \end{aligned}$$

Similar, at time 0 we compute

$$\begin{aligned} Z_0 &= \max\{Y_0, \mathbb{E}_{\mathbb{Q}}[Z_1 B_0/B_1 | \mathscr{F}_0]\} \\ &= \max\left\{0, q_u Z_1(u) \cdot \frac{1}{B_1} + q_d \cdot Z_1(d) \cdot \frac{1}{B_1}\right\} \\ &\approx 0.7427. \end{aligned}$$

We have computed the fair prices of C^A. An investor should, given the principles of dynamic programming, only exercise the option at time 1 if scenario m occurs. At any other scenarios, the investor should treat the asset as an European option.[12]

3 Binomial Model

In the next two chapters we will look at two models with some special features. The first one is the binomial model, which we recognize as a model that handles frequent (discrete) attempts with a fixed probability. The binomial model is of interest as it allows a simplified calculation of pricing and replication of a large

[12] Of course, exercising at time 1 in case of scenario m yields zero payoff, so it does not matter if the investor exercise or not.

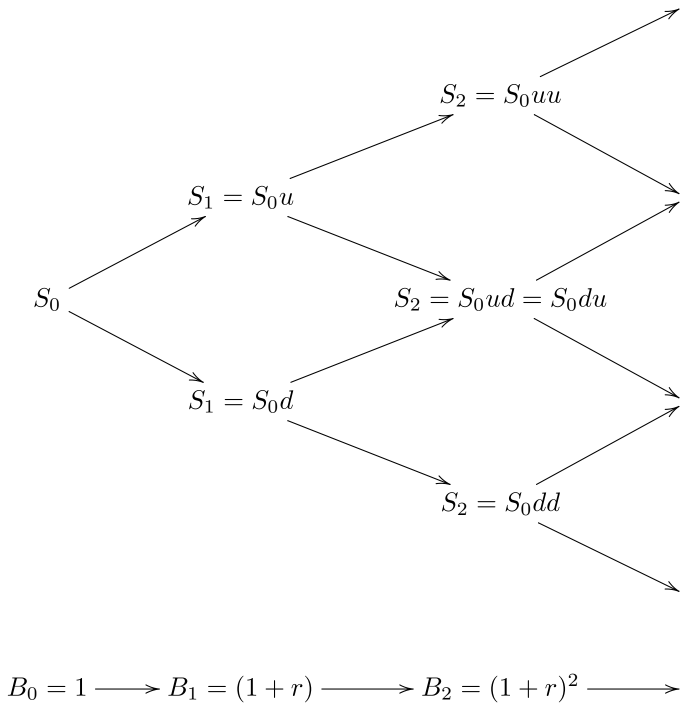

Figure 3.1: Showing the three first steps of a generalization of the binomial model.

class of derivatives that are path-independent. As we can see from Definition 3.2, given by Cutland & Roux [3, p. 184], the European call and put options are examples of derivatives that are path-independent.

Definition 3.2 (Path-independent derivative)**.** A derivative D is called path-independent if there exists a payoff function $\hat{D}$ such that

$$D = \hat{D}(S_T).$$

The binomial model, as in Pliska [5, pp. 100–106], is a multi-period binary model with one risky security and one risk-free security. It is binary since the stock price evolution of the risky security is completely determined of two parameters, say u and d. These are fixed and known at time 0, with $0 < d < u$. At each period, either the risky security moves with a factor of u or with a factor of d, where we assume $S_0 > 0$. We will for simplicity also assume $B_0 = 1$ and that the risk-free security has a fixed and known interest rate $r > 0 \in \mathbb{R}$. See Figure 3.1 for a general example.

The binomial model is, for our purposes so far, completely determined by the parameters T, r, S_0, u and d. There is also a real world probability p for the probability of the risky security to move «up» with a factor of u, and the probability $1-p$ for a «down»-move with a factor of d. These are constant on all periods.

From Figure 3.1 we see that at time t there will be t^2, for $t > 1$, different possible scenarios and $t-1$ different possible values of the risky security. The number of scenarios with the same final risky security price is given, as in Pliska [5, p. 102], by the so-called *binomial coefficient*

$$\binom{T}{t} := \frac{t!}{n!(t-n)!}.$$

This follows from the fact that the order of the «ups» and «downs» is not important, only the total number of «ups» (or «downs»). With this, Pliska [5, p. 104] shows that the binomial model has the following probability distribution of S_t under an equivalent martingale measure $\mathbb{Q}$:

$$\mathbb{Q}(S_t = S_0 u^n d^{t-n}) = \binom{t}{n} q^n (1-q)^{t-n}, \quad n = 0, 1, \ldots, t. \tag{3.1}$$

We want our model to be viable, which we know is the case if it admits an equivalent martingale measure $\mathbb{Q}$. The following theorem by Cutland & Roux [3, p. 108] tells us that this is the case if and only if each single-step submodel is viable.

Theorem 3.3. *A finite multi-period model is viable if and only if each single-step submodel is viable.*

Proof. The result and the underlying idea is easy to grasp, but the detailed proof is somewhat lenghty. The proof connects the definitions of arbitrage and viability. For a detailed proof see Cutland & Roux [3, p. 108]. □

Since each single-step submodel looks the same, we only need to solve the first one to see the whole picture. The result is simply a restriction on the parameters u and d.

Theorem 3.4 (Viable binomial model, [3, p. 181])**.** *A finite multi-period binomial model with parameters T, r, S_0, u and d, is viable if and only if $d < (r+1) < u$.*

Proof. This is one of several results we could get from investigating the single-period single-stock model we introduced at the very beginning. The results depend on the assumption $S_0 > 0$. See Cutland & Roux [3, p. 181] for a proof. □

We now have a method to check that any given binomial model is viable. We want to find an equivalent martingale measure. In the work of Pliska [5,

p. 103] a detailed presentation of this is found. First he shows that any one-step equivalent martingale measure, normally called a risk-neutral probability measure in single-period models, must satisfy

$$q\frac{u-1-r}{1+r} + (1-q)\frac{d-1-r}{1+r} = 0,$$

where q is the conditional probability that the next move is «up». Solving for q yields $q = \frac{r+1-d}{u-d}$. From this we deduce that $d < (r+1) < u$, in order for q to satisfy the conditions of being a risk neutral probability measure. This is the result we saw in Theorem 3.4.

We have a risk-neutral probability measure for all the single-period sub-models. But what about the complete model? This is given by the following result.

Theorem 3.5. *In a viable binomial model, if $S_T(\omega) = S_0 u^s d^{T-s}$ for some $s \leq T$ and $\omega \in \Omega$, then it admits a unique equivalent martingale measure $\mathbb{Q}$ given by $\mathbb{Q}(\omega) = q^s(1-q)^{T-s}$, where s represent the number of «up» moves.*

Proof. See Cutland & Roux [3, p. 182] for a proof. □

The last result we need before we can begin pricing path-independent derivatives is the probability distribution of S_t under an equivalent martingale measure.

Notice that the martingale $\mathbb{Q}$ depends only on $S_T(\omega)$ and not the full price history. This is one of several properties of the binomial model. From Theorem 2.1 we deduce that a fair price at time 0 of a path-independent derivative D with exercise date T in a viable binomial model is given by

$$\begin{aligned} D_0 &= (1+r)^{-T}\mathbb{E}_{\mathbb{Q}}[D] \\ &= (1+r)^{t-T}\mathbb{E}_{\mathbb{Q}}[\hat{D}(S_T)] \\ &= (1+r)^{-T}\sum_{\omega\in\Omega}\mathbb{Q}(\omega)\hat{D}(S_T(\omega)). \end{aligned} \tag{3.2}$$

Knowing that S_T must take one of the values $S_0 d^T, S_0 d^{T-1}u, \ldots, S_0 u^T$, together with (3.1), Cutland & Roux [3, p. 185] shows that this can be rewritten as

$$D_0 = (1+r)^{-T}\sum_{s=0}^{T}\binom{T}{s}q^s(1-q)^{T-s}\hat{D}(S_0 u^s d^{T-s}). \tag{3.3}$$

They recognize this as the general *Cox-Ross-Rubinstein formula*[13] for the fair price of a path-independent derivative at time 0. While equation (3.2) contains 2^T terms, as we noticed from Figure 3.1, equation (3.3) contains $T+1$ terms, which is significantly less for models with more trading times.

4 Markov Models

We will here introduce the *Markov chain*, as in Pliska [5, p. 106], which is a simple Markov model where the system state is fully observable and autonomous. It consists of a filtration $\mathbb{F} = \{\mathscr{F}_t\}_{t=0}^T$ generated by a stochastic process $X = \{X_t\}_{t=0}^T$. This process takes values in a finite state space E. If $X_t = j \in E$, we think of this as the process X beeing in state j at time t. As before, there is a sample space Ω and a probability measure P.

For a stochastic process X to be a *Markov chain* it need to satisfy the so-called Markov property:

$$P\{X_{t+1} = j | \mathscr{F}_t\} = P\{X_{t+1} = j | X_t\}$$

Stated in words: Given the history of the process, only the current state is of importance to determine the future value.

Further, we assume the Markov chain to be stationary if the probability $P\{X_{t+1} = j | \mathscr{F}_t\}$ is independent of the time t. Thus the binomial model is a stationary Markov chain.

These properties simplify computations and yet the models lead to realistic representations of true prices [5, p. 106]. In the coming parts we will use models with the Markov property. But, first a demonstration of the strength of Markov chains when computing conditional probability distributions. The computation that follows is a generalization of an example given by Pliska [5, p. 110]. Assume $t < T$ and $j \in E$. Then

[13] The binomial model was first introduced by Cox, Ross and Rubinstein in 1979 for valuation of options, and therefor many authors use the name *Cox-Ross-Rubinstein model* instead of *Binomial model* [1].

$$
\begin{aligned}
P\{X_T = j|\mathscr{F}_t\} &= P\{X_T = j|X_t = i\} \\
&= \sum_{i=1}^{T-t+1} \sum_{k_i \in E} P\{X_T = j, X_{T-1} = k_1, \ldots, X_{t+1} = k_{T-t+1}|X_t = i\} \\
&= \sum_{i=1}^{T-t+1} \sum_{k_i \in E} P\{X_T = j|X_{T-1} = k_1, \ldots, X_t = i\} \\
&\qquad \cdot P\{X_{T-1} = k_1|X_{T-2} = k_2, \ldots, X_t = i\} \\
&\qquad \cdot \ldots \\
&\qquad \cdot P\{X_{t+1} = k_{T-t+1}|X_t = i\} \\
&= \sum_{i=1}^{T-t+1} \sum_{k_i \in E} P\{X_T = j|X_{T-1} = k_1\} \\
&\qquad \cdot P\{X_{T-1} = k_1|X_{T-2} = k_2\} \\
&\qquad \cdot \ldots \\
&\qquad \cdot P\{X_{t+1} = k_{T-t+1}|X_t = i\}
\end{aligned}
$$

Both in the second top and in the second bottom equation we make use of the strong Markov property. If the Markov chain is stationary this simplifies further to

$$
= \sum_{i=1}^{T-t+1} \sum_{k_i \in E} P\{X = j|X = k_1\} P\{X = k_1|X = k_2\} \cdot \ldots \cdot P\{X = k_{T-t+1}|X = i\}.
$$

5 Linear Programming

This chapter will be about constrained linear optimization, one of the basic pillars of mathematical optimization. For instance, in the work of Pliska [5], Alex & Cutland [3] and many more, we see how strong linear programming is as a proof technique. First, some basic vocabulary is needed, and then an example of a linear programming problem (LP) will be introduced. Further, we show how we can solve a LP with dual theory. The strong complimentary property will also be given attention, as we will need it for the work in the next section.

The following is a summary of Vanderbei [6] to accommodate some notation. A LP comes in many forms, but what we recognize is that it is some function we want to maximize (or minimize). This is called the *objective function.* More precise, its the variables of the objective function we want to decide, called the *decision variables.* The objective function comes with a set of *constraints*, which the solution need to satisfy. A *solution* is said to be *feasible* with respect to the LP if it satisfy all the given constraints. The set of all feasible solutions is called the *feasible region*. A feasible solution is *optimal* if its objective function value is maximized (or minimized) with respect to the feasible region. Of course, the problem may be *unbounded*, that is, the decision variables can be increased (or decreased) indefinitely, and still satisfy the set of constraints. Then no optimal solution will exists.

Associated with every LP is another called its *dual*. The dual of this dual linear program is the original LP, therefor referred to as the *primal linear program*. Many textbooks will restrict maximization problems to always be defined as the primal problem, while minimization problems are always defined as the dual problem. Which of the two cohesive LP problems that are defined as dual and primal does not matter and the theorems below is valid regardless of the choice. For simplicity we assume that the primal LP is defined as

$$\begin{aligned}
\text{maximize} \quad & \sum_{j=1}^{n} c_j x_j \\
\text{subject to} \quad & \sum_{j=1}^{n} a_{ij} x_j \le b_i, \quad && i = 1, 2, \ldots, m, \\
& x_j \ge 0, && j = 1, 2, \ldots, n.
\end{aligned}$$

The associated dual LP is then

$$\begin{aligned}
\text{minimize} \quad & \sum_{i=1}^{m} b_i y_i \\
\text{subject to} \quad & \sum_{i=1}^{m} y_i a_{ij} \ge c_j, \quad && j = 1, 2, \ldots, n, \\
& y_i \ge 0, && i = 1, 2, \ldots, m.
\end{aligned}$$

Above, x_j and y_i are decision variables, and b_i, c_j and a_{ij} are constants. The definitions above may seem to limit the set of allowed LP's to work with. But, a constraint can often in some way be converted to fit the definition:

$$\begin{aligned}
a_1x_1 + a_2x_2 + \ldots + a_nx_n &= b \\
&\Longleftrightarrow \\
a_1x_1 + a_2x_2 + \ldots + a_nx_n &\le b \\
a_1x_1 + a_2x_2 + \ldots + a_nx_n &\ge b \\
&\Longleftrightarrow \\
a_1x_1 + a_2x_2 + \ldots + a_nx_n &\le b \\
b - a_1x_1 - a_2x_2 - \ldots - a_nx_n &\le 0.
\end{aligned}$$

Working with non-negative variables in optimization problems is a common situation, but not a necessary criterion. Transportation, production, traveling and cost reduction are some topics which encounters non-negative variables. We will not adjust the results to suit negative variables, but mention that it is fully possible to solve such optimization problems. The most important results we encounters are the so-called weak and strong duality theorems and the complementary slackness conditions.

Theorem 5.1 (Weak duality, [7, p. 58])**.** *For any feasible solution for the primal (dual) problem, the value of the objective function provides a lower (upper) bound on the optimal value to the dual (primal) problem .*

Theorem 5.2 (Strong duality, [7, p. 60])**.** *The existence of an optimal solution is mutual for the primal/dual pair. If it exists, then the optimal value of the dual objective function equals the optimal value of the primal objective function.*

The strong duality theorem is important as it provides a confirmation of whether the suggested solution is actually the right answer. The weak duality theorem tells us that given two feasible solutions, one for the primal and one for the dual, there will be a gap between the values of the two objective functions. There will be an upper bound u for the maximization problem, and a lower bound l for the minimization problem, with $u \leq l$, see Figure 5.3. What the strong duality theorem states is that any solutions satisfying $u = l$ are optimal. That is, no gap.

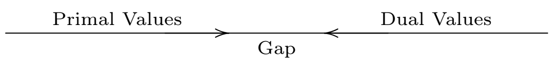

Figure 5.3: Primal objective values are all less than the dual objective values.

Theorem 5.4 (The complementary slackness conditions, [7, p. 67])**.** *Let x be a feasible solution for the primal, and y be a feasible solution for the dual. They are optimal solutions of their respective LP's if and only if the complementary slackness conditions holds:*

$$y_i(b_i - \sum_{j=1}^{n} a_{ij}x_j) = 0, \quad 1 \leq i \leq m,$$

$$x_j(\sum_{i=1}^{m} a_{ij}y_i - c_j) = 0, \quad 1 \leq j \leq n.$$

Proof. See [6, pp. 58–67] for a proof of the weak and strong duality theorems and the complementary slackness conditions. □

Theorem 5.4 is a very strong property as it allows us to choose the simplest model to address (from an algorithmic point of view). Either way you will get the results of the associated equivalence model (may it be the primal or dual problem). If you know the optimal solution of the primal, then you can find the solution of the dual problem (and vice versa) by solving a system of equations formed by the decision variables and the set of constraints [2].

Example 5.5 (A linear programming problem)**.** There are loads of intuitive and well formulated examples of LP problems, but we will simply make up a typical example (not completely random, as we will see later). Let the function

$$\zeta = \sum_{i=1}^{3} y_i,$$

be the objective function we want to minimize. Each decision variable y_i is greater than or equal to a non-negative constant f_i, for $i = 1, 2, 3$. Additionally, $c_1 y_{i-1} + c_2 y_i + c_3 y_{i+1} \geq 0$, for $i = 2, 3$, where each c_j is a constant. Below is the problem reformulated:

$$\begin{aligned} &\text{minimize} && \sum_{i=1}^{3} y_i && \\ &\text{subject to} && y_i \geq f_i, && i = 1, 2, 3 \\ & && c_1 y_{i-1} + c_2 y_i + c_3 y_{i+1} \geq 0, && i = 2, 3 \\ & && y_1, y_2, y_3 \geq 0 && \end{aligned} \tag{5.1}$$

The non-negativity of the decision variables of the ζ-function follows from the fact that each f_i is non-negative. We have not said anything about y_4, that appears at $i = 3$. This is not of importance in this example, so we simply assume $y_4 = 0$ (without removing it from the constraints).

Until now we have only stated that there always exists an associated LP. Lets see this by following the logic of Vanderbei [6, pp. 55–56] and examining (5.1) closer. Obviously, any feasible solution gives an upper bound. But how good is it? How close is it to the optimal value? What we need is a lower bound. We start by multiplying the constraints with non-negative numbers x (thus the inequalities are preserved):

$$\begin{aligned} x_i y_i \geq& x_i f_i, \quad i = 1, 2, 3, \\ x_j(c_1 y_{i-1} + c_2 y_i + c_3 y_{i+1}) \geq& 0, \qquad i = 2, 3, j = i + 2. \end{aligned}$$

Adding the constraints yields

$$\begin{aligned} &y_1(x_1 + c_1 x_4) + y_2(x_2 + c_2 x_4 + c_1 x_5) \\ &\quad + y_3(x_3 + c_3 x_4 + c_2 x_5) + y_4(c_3 x_5) \geq x_1 f_1 + x_2 f_2 + x_3 f_3. \end{aligned} \tag{5.2}$$

We stipulate that each of the coefficients of the y_i's in (5.2) to be at least smaller or equal to the corresponding coefficient in the objective function:

$$\begin{aligned} x_1 + c_1x_4 &\leq 1, \\ x_2 + c_2x_4 + c_1x_5 &\leq 1, \\ x_3 + c_3x_4 + c_2x_5 &\leq 1, \\ c_3x_5 &\leq 0. \end{aligned} \tag{5.3}$$

Next we compare the inequalities in (5.3) with the objective function:

$$\begin{aligned} \zeta = \sum_{i=1}^{3} y_i &\geq y_1(x_1 + c_1x_4) + y_2(x_2 + c_2x_4 + c_1x_5) + y_3(x_3 + c_3x_4 + c_2x_5) + y_4(c_3x_5) \\ &\geq x_1f_1 + x_2f_2 + x_3f_3. \end{aligned}$$

We now have a lower bound, $x_1f_1 + x_2f_2 + x_3f_3$, which we should maximize in order to obtain the best possible lower bound. Finally, we are led to the dual problem of (5.1):

$$\begin{aligned} \text{maximize} \quad & \sum_{i=1}^{3} x_if_i \\ \text{subject to} \quad & x_1 + c_1x_4 \leq 1, \\ & x_2 + c_2x_4 + c_1x_5 \leq 1, \\ & x_3 + c_3x_4 + c_2x_5 \leq 1, \\ & c_3x_5 \leq 0, \\ & x_i \geq 0, \quad i = 1, 2, 3, 4. \end{aligned} \tag{5.4}$$

6 American Perpetual Options

6.1 Introduction

A *perpetual option* is an option that does not have an expiration date. An *American perpetual option* is similar to the American options in chapter 2.2, except that they are perpetual. Again, the problem is to determine a fair price for both the writer and the holder. We will work out the solution by writing the problem as a linear problem, and then make use of linear programming duality.

Before we can derive the solution we need a model to work with. The choice of model will have important effects on the solution. First, we present a short summary of the work of Vanderbei & Pınar [7], assuming the risky security price behaves as a random walk. After, we will in more detail look at a similar Markov chain model, but with three possible outcomes at each trading step.

6.2 Random Walk

Let $S = \{S_t : t = 0, 1, \dots\}$ be a stochastic process, where each S_t denotes the price of the risky security at time t, as before. Let us assume S is a random walk with absorption at 0 on the state space $E = \{j\Delta x : j = 0, 1, \dots\}$, where Δx is a fixed positive number. The solution to pricing perpetual American options based on this model is given by Vanderbei & Pınar [7, pp. 768–781].

With *absorption at* 0; once the risky security equals 0, it has to stay in that state. Thus it is not possible to take negative values,[14] or to reach zero value and then have a positive gain. With a *random walk* we mean that if the current price is $x \in E$, then the price at the next period will be $x \pm \Delta x$ with probabilities p for an increase in value, and probability $1 - p$ for a decrease in value, $0 < p < 1$. Thus, the random walk possess the desired Markov property shown in chapter 4.

The goal is to price perpetual American options based on the risky security above. We need to figure out the expected value, in today's dollars. That is, $\mathbb{E}_x[\alpha^\tau Y_\tau]$, for some future time τ.

The payoff function, Y_τ, is similar to the one introduced in chapter 2.1 and 2.2. The discount factor α^τ represents the value today of one dollar[15] at time τ. Since future dollars (normally) are worth less than present dollars, we assume $\alpha^\tau < 1$. We could stick to the earlier bank account process, $\mathbb{B}$, but for simplicity we express this with the number α. The expectation is under the given probabilities p and $1 - p$.

It is natural to assume that the holder of the option will follow an optimal exercising strategy. The value function $v(x)$ below tells us to exercise at time τ when the expected payoff is at its greatest:

[14] Given the assumption $S_0 > 0$.
[15] Or, one unit of the current medium of exchange.

$$v(x) = \max_{\tau} \mathbb{E}_x[\alpha^\tau Y^\tau].$$

The holder of the option can always exercise the option or choose to keep it for one more period. Therefore, $v(x)$ satisfy the principles of dynamic programming. That is, in order to act optimally, $v(x)$ must satisfy

$$\begin{aligned} &v(x) \geq Y(x), \ x \in E, \\ &v(x) \geq \alpha(pv(x+\Delta x) + (1-p)v(x-\Delta x), \ x \in E\backslash\{0\}. \end{aligned}$$

Following Vanderbeid & Pınar [7], recalling the discreteness of the model's state space, an infinite-dimensional linear programming problem is formulated

$$\begin{aligned} &\text{minimize} && \sum_{j=0}^{\infty} v_j && \\ &\text{subject to} && v_j \geq Y_j, && j = 0, 1, \dots, \\ & && v_j \geq \alpha(pv_{j+1} + (1-p)v_{j-1}, && j = 1, 2, \dots, \end{aligned} \tag{6.1}$$

where $v_j = v(j\Delta x)$ and $Y_j = Y(j\Delta x)$.

Under some assumptions, the solution to the LP (6.1) can be shown to be:

$$v_j = \begin{cases} 0, & j = 0, \\ Y_{j^*} \frac{\zeta_+^j - \zeta_-^j}{\zeta_+^{j^*} - \zeta_-^{j^*}}, & 0 < j < j^*, \\ Y_j, & j^* \leq j, \end{cases}$$

where

$$j^* := \max \left\{ k : Y_k \frac{\zeta_+^{k-1} - \zeta_-^{k-1}}{\zeta_+^k - \zeta_-^k} > Y_{k-1} \right\}.$$

This solution implies that the price of the perpetual American option should be $Y_j^* = (j^*\Delta x - K)^+$ for a call option, where K is the exercise price. Unfortunately, it's not enough to let $Y_j^* = (K - j^*\Delta x)^+$ for a put option. We must take into account different boundary conditions as well. Of course, as time changes, so will the development of the underlying risky security change, and thus so will the price of the perpetual American option. We mentioned there were some assumptions needed, and we have not shown how Vanderbei & Pınar reached the results. This will not be given for the random walk model, but detailed calculations and necessary assumptions will be given for the similar ternary Markov chain model. We will also make a remark for the boundary conditions for the put option.

6.3 Markov Chain

Again, a Markov chain model with absorption at zero is used, but now the risky security will at all trading periods have three ways to develop. As with the random walk, at each step the risky security moves up by Δx with probability p, or moves down by Δx with probability q. In addition, the risky security does not move with probability $s = 1 - (p + q)$, implying $p + q < 1$.

Let $S = \{S_t : t = 0, 1, 2, \dots\}$ be a Markov chain, where each S_t denotes the price of the risky security at time t. The state space is similar to the random walk model, $E = \{j\Delta x : j = 0, 1, 2, \dots\}$, where Δx is a fixed small positive number.

The problem is to price perpetual American options based on the risky security model above. We want to find the exercise time τ that maximizes the expected payoff, with the discount factor α taken into account:

$$\max_{\tau} \mathbb{E}_x[\alpha^\tau Y_\tau].$$

We recognize this expression as the value process given by (2.1). We saw in (2.3), by the principles of dynamic programming, that the price of the option, say $v(x)$, must satisfy

$$v(x) = \max\{Y(x), \alpha(pv(x + \Delta x) + qv(x - \Delta x) + sv(x)\}, \tag{6.2}$$

for all $x \in E\backslash\{0\}$. With $Y(x)$ we mean the payoff when the price of the underlying risky security is x.[16] Because of the absorption at 0, we get the boundary condition $v(0) = 0$. The value function in (6.2) follows the principle of dynamic programming: If the payoff today is greater than the expected discounted value tomorrow, we exercise the option or sell it short if that is possible. If the opposite is the case, then we hold the option for one more period, or buy the option. This is also explained by Vanderbei & Pınar and is similar to what we used in the random walk model [7, p. 768].[17]

Next, We decompose the value function $v(x)$. Since it is equal to the optimal choice between exercising the option and the discounted expected value for the next period, we get the inequalities

$$\begin{aligned} &v(x) \geq Y(x), \quad x \in E, \\ &v(x) \geq \alpha(pv(x + \Delta x) + qv(x - \Delta x) + sv(x)), \quad x \in E\backslash\{0\}. \end{aligned}$$

[16] The difference now from the finite time options is that we are expressing the value functions based on the stock price x, instead of the time t.

[17] One of the crucial assumptions for the logic behind the dynamic programming to work, is that you can trade the option at all trading dates t, as many times needed.

Assume $W(x)$ is another value function satisfying

$$\begin{aligned} &W(x) \geq Y(x), \\ &W(x) \geq \max\{Y(x), \alpha(pv(x+\Delta x) + qv(x-\Delta x) + sv(x))\}. \end{aligned}$$

Then for all optimal exercising times τ,

$$W(x) \geq Y(x) = v(x).$$

In particular, for non-optimal exercising times t,

$$W(x) \geq \alpha(pv(x+\Delta x) + qv(x-\Delta x) + sv(x)) = v(x).$$

Therefore, $v(x)$ is the smallest function, in the L^1 sense [7, p. 769], satisfying (6.2), for $0 < t < \tau$. Combining this with the discreteness of our model's state space, we formulate an infinite-dimensional linear programming problem, which we from now will refer to as the *dual problem*,

$$\begin{aligned} &\text{minimize} && \sum_{j=0}^{\infty} v_j \\ &\text{subject to} && v_j \geq Y_j, && j = 0, 1, \dots, \\ & && v_j \geq \alpha(pv_{j+1} + qv_{j-1} + sv_j), && j = 1, 2, \dots, \end{aligned} \tag{6.3}$$

where $x_j = j\Delta x$, $v_j = v(x_j)$, and $Y_j = Y(x_j)$. We don't mention the non-negativity of the decision variables, $v_j \geq 0$, since this follows naturally from the fact that $Y_j \geq 0$, as explained in chapter 2.1 and 2.2. The idea is that no matter the values of the different constants α, p, q, s and K, as the value of the risky security increases, for some value x we will have

$$Y_j > \alpha(pv_{j+1} + qv_{j-1} + sv_j).$$

That exercise point, say j^*, is basically all we need to know, as will be shown later. Thus, the sum don't need to run to infinity. Since v_j is minimized for all $j \geq 0$, and because of the constraints making sure (6.2) is satisfied, the dual problem solves our value function $v(x)$.

Next, we notice that the dual problem is similar to the finite-dimensional problem demonstrated in Example 5.5. Adjust the indices, replace y_i with v_j, c_1 with $-\alpha q$, c_2 with $1-\alpha s$, c_3 with $-\alpha p$, and then generalise the problem to be infinite-dimensional. The first two constraints need some extra attention. For the next constraints, those defined for $j \geq 2$, a fixed system follows. Thus, the associated LP is

$$\begin{aligned}
\text{maximize} \quad & \sum_{j=0}^{\infty} Y_j y_j \\
\text{subject to} \quad & y_0 - \alpha q z_1 \le 1, \\
& y_1 + (1-\alpha s) z_1 - \alpha q z_2 \le 1, \\
& y_j + (1-\alpha s) z_j - \alpha q z_{j+1} - \alpha p z_{j-1} \le 1, \quad && j \ge 2, \\
& y_j \ge 0, && j \ge 0, \\
& z_j \ge 0, && j \ge 1.
\end{aligned}$$

We immediately note that the z_j's don't contribute to the objective function. This, together with maximizing over all the y_j's, gives a LP, from now referred to as the *primal problem*, with a more strict set of constraints.[18]

$$\begin{aligned}
\text{maximize} \quad & \sum_{j=0}^{\infty} Y_j y_j \\
\text{subject to} \quad & y_0 - \alpha q z_1 = 1, \\
& y_1 + (1-\alpha s) z_1 - \alpha q z_2 = 1, \\
& y_j + (1-\alpha s) z_j - \alpha q z_{j+1} - \alpha p z_{j-1} = 1, \quad && j \ge 2, \\
& y_j \ge 0, && j \ge 0, \\
& z_j \ge 0, && j \ge 1.
\end{aligned} \tag{6.4}$$

Statement of claim Let v_j denote the optimal dual solution and y_j, z_j the optimal primal solution.[19] Assume that there exists an optimal exercise point j^* such that the following holds:

$$\alpha(p v_{j+1} + q v_{j-1} + s v_j) > Y_j, \quad 0 < j < j^*, \tag{6.5}$$

$$\alpha(p v_{j+1} + q v_{j-1} + s v_j) < Y_j, \quad j \ge j^*. \tag{6.6}$$

Before we move on, what does it really mean to assume that there exists an optimal exercise point j^*? There is a one-to-one connection between j and the risky security price, and they both grow linearly. In addition, there is no upper boundary on either of them. We conclude that assuming there exists an optimal exercise point is similar as assuming that the exercise price is finite. An option with an infinite exercise price is not very interesting in a financial setting, as no investor with an interest in making money would invest in such an option. Still, it is important, and interesting, in a mathematical point of view.

Back to the statement of claim, we remember to include the absorption at 0. Thus

[18] Rearranging the constraints we realize that there is no reason for y_j to be smaller than $1 - (1-\alpha s) z_j + \alpha q z_{j+1} + \alpha p z_{j-1}$, when equality is a possibility.

[19] The usual «stars» denoting optimality is left out.

$$v_0 = 0, \tag{6.7}$$
$$v_j = \alpha(pv_{j+1} + qv_{j-1} + sv_j), \text{ for } 0 < j < j^*, \tag{6.8}$$
$$v_j = Y_j, \text{ for } j^* \leq j. \tag{6.9}$$

With put options we have $v_0 = K$, where K is the exercise price. The following computations would be similar with this boundary condition, but to avoid being too lengthy we focus on call options.

Invoke complementary The complimentary slackness conditions introduced in Theorem 5.4, in our particular case, states that any optimal solution v_j to the dual problem (6.3), and any optimal solutions y_j, z_j to the primal problem (6.4) satisfies

$$y_j(Y_j - v_j) = 0, \quad j \geq 0, \tag{6.10}$$
$$z_j(\alpha pv_{j+1} + \alpha qv_{j-1} - (1 - \alpha s)v_j) = 0, \quad j \geq 1. \tag{6.11}$$

To see where (6.11) comes from, it is easier to investigate (5.4) from chapter 5, linear programming. The confusion may be the fact that the x-variables are divided into y's and z's in (6.4) and that we are now working with the challenging mathematical concept of infinity. The fact that the z's is not *seen* in the objective function need be nothing but their coefficients equaling 0. (This is x_4 and x_5 in the particular case of (5.4).)

Equation (6.10) is true for $j^* \leq j$, by (6.9). For $j < j^*$ we know from (6.8) that $v_j = \alpha(pv_{j+1} + qv_{j-1} + sv_j)$. Combining this with (6.5) yields

$$y_j = 0, \quad 0 < j < j^*. \tag{6.12}$$

Similarly, (6.11) is true for $j < j^*$ by (6.8). For $j^* \leq j$ we know from (6.9) that $v_j = Y_j$. Combining this with (6.6) yields

$$z_j = 0, \quad j \geq j^*. \tag{6.13}$$

Using (6.13) combined with (6.4) we obtain

$$\begin{aligned} y_{j^*} - \alpha p z_{j^*-1} &= 1, \\ y_j &= 1, \quad j > j^*. \end{aligned} \tag{6.14}$$

Similarly, using (6.12) combined with (6.4) we obtain

$$\begin{aligned}(1-\alpha s)z_1 - \alpha q z_2 &= 1,\\ -\alpha p z_{j-1} + (1-\alpha s)z_j - \alpha q z_{j+1} &= 1, \quad 1 < j < j^*.\end{aligned} \tag{6.15}$$

Difference equations We can now formulate a pair of second-order difference equations. The difference equation for the primal problem is given by (6.7), (6.8) and (6.9):

$$\begin{aligned}(1-\alpha s)v_j - \alpha p v_{j+1} - \alpha q v_{j-1} &= 0, \quad 0 < j < j^*,\\ v_0 &= 0,\\ v_{j^*} &= Y_{j^*}.\end{aligned} \tag{6.16}$$

The difference equation for the dual problem is given by (6.15) and (6.13):

$$\begin{aligned}(1-\alpha s)z_j - \alpha q z_{j+1} - \alpha p z_{j-1} &= 1, \quad 0 < j < j^*,\\ z_0 &= 0,\\ z_{j^*} &= 0.\end{aligned} \tag{6.17}$$

Note that in (6.17) we added a new variable, z_0, which is fixed to be zero. Thus we don't need to include both equations in (6.15). We do this to consolidate the difference equations.[20]

Solving the difference equations We refer to Knut Mørken's compendium [4] for necessary background information about how to solve difference equations. First we rearrange the difference equation in (6.16):

$$v_{j+1} + \frac{\alpha s - 1}{\alpha p} v_j + \frac{q}{p} v_{j-1} = 0.$$

We set up the characteristic equation:

$$\begin{aligned} r^2 + \frac{\alpha s - 1}{\alpha p} r + \frac{q}{p} = 0 \quad &\Rightarrow\\ r = \frac{(1-\alpha s) \pm \sqrt{(\alpha s - 1)^2 - 4\alpha^2 pq}}{2\alpha p}&.\end{aligned}$$

There are two distinct roots, thus for constants C_1 and C_2 the general solution is

$$v_j = C_1 r_+^j + C_2 r_-^j.$$

[20] When we solve a difference equation for an unknown function, say u, and the value of u is given on the boundary conditions, then we say we have Dirichlet boundary conditions.

The next step is solving for the boundary conditions. From (6.16), the first one is $v_0 = 0$. Inserting into the general solution yields

$$\begin{aligned} C_1 r_+^0 + C_2 r_-^0 = 0 \Rightarrow C_1 = -C_2, \\ \Rightarrow v_j = C_1 r_+^j - C_1 r_-^j. \end{aligned}$$

The second boundary condition is $v_{j^*} = Y_{j^*}$,

$$C_1 r_+^{j^*} - C_1 r_-^{j^*} = Y_{j^*} \Rightarrow C_1 = \frac{Y_{j^*}}{r_+^{j^*} - r_-^{j^*}}.$$

Inserting for the constant C_1, we finally have the exact solution

$$v_j = Y_{j^*} \frac{r_+^j - r_-^j}{r_+^{j^*} - r_-^{j^*}}, \quad 0 < j < j^*. \tag{6.18}$$

Now we solve the difference Equation (6.17). Again we start by rearranging:

$$z_{j+1} + \frac{\alpha s - 1}{\alpha q} z_j + \frac{p}{q} z_{j-1} = -\frac{1}{\alpha q}. \tag{6.19}$$

First we solve the homogeneous version of (6.19). It has characteristic equation

$$\begin{aligned} & r^2 + \frac{\alpha s - 1}{\alpha q} r + \frac{p}{q} = 0 \quad \Rightarrow \\ & r = \frac{(1 - \alpha s) \pm \sqrt{(\alpha s - 1)^2 - 4\alpha^2 pg}}{2\alpha q}. \end{aligned}$$

There are two distinct roots, thus for constants C_1 and C_2 the general homogeneous solution is

$$z_j^h = C_1 r_+^j + C_2 r_-^j.$$

For the particular solution we guess $z_j^p \equiv c$, where c is a constant. Inserting this solution into (6.19) yields

$$\begin{aligned} & c + \frac{\alpha s - 1}{\alpha q} c + \frac{p}{q} c = -\frac{1}{\alpha q} \\ \Rightarrow \quad & \alpha c(p + q + s) - c = -1 \\ \Rightarrow \quad & c = \frac{1}{1 - \alpha}. \end{aligned}$$

We see that it is possible to choose c such that (6.19) is satisfied. Thus, our guess was correct. By the superposition principle[21] the general solution can be found as follows

$$z_j = z_j^h + z_j^p = C_1 r_+^j + C_2 r_-^j + \frac{1}{1-\alpha}. \tag{6.20}$$

Solving for the boundary conditions, starting with $z_0 = 0$, gives

$$\begin{aligned} & C_1 + C_2 + \frac{1}{1-\alpha} = 0 \\ \Rightarrow \quad & C_1 = \frac{1}{\alpha - 1} - C_2. \end{aligned}$$

Next, we insert for C_1 in the general solution (6.20)

$$z_j = \big(\frac{1}{\alpha-1} - C_2\big) r_+^j + C_2 r_-^j + \frac{1}{1-\alpha}.$$

The final boundary condition is $z_{j^*} = 0$:

$$\begin{aligned} & \big(\frac{1}{\alpha-1} - C_2\big) r_+^{j^*} + C_2 r_-^{j^*} + \frac{1}{1-\alpha} = 0 \\ \Rightarrow \quad & C_2 = \frac{1 - r_+^{j^*}}{(\alpha-1)(r_-^{j^*} - r_+^{j^*})}. \end{aligned}$$

Inserting for C_2 in the general solution, and after some computations, the exact solution is

$$z_j = \frac{r_+^j (r_-^{j^*} - 1) + r_-^j (1 - r_+^{j^*}) - r_-^{j^*} + r_+^{j^*}}{(\alpha-1)(r_-^{j^*} - r_+^{j^*})}, \quad 0 < j < j^*.$$

Summarizing our results we have

$$v_j = \begin{cases} 0, & j = 0, \\ Y_{j^*} \frac{r_+^j - r_-^j}{r_+^{j^*} - r_-^{j^*}}, & 0 < j < j^*, \\ Y_j, & j^* \le j, \end{cases}$$

[21]The term *superposition* is not seen in the compendium [4], but it's commonly used in science. It refers in this case to the fact that the solution of an inhomogeneous difference equation is the linear combination of the associated particular- and the associated homogeneous solutions.

$$\mathrm{z}_j = \begin{cases} \frac{r_+^j(r_-^{j^*}-1)+r_-^j(1-r_+^{j^*})-r_-^{j^*}+r_+^{j^*}}{(\alpha-1)(r_-^{j^*}-r_+^{j^*})}, & 0 < j < j^*, \\ 0, & j^* \leq j, \end{cases}$$

$$\mathrm{y}_j = \begin{cases} 1 + \alpha q z_1, & j = 0, \\ 0, & 0 < j < j^*, \\ 1 + \alpha p z_{j^*-1}, & j = j^*, \\ 1, & j^* < j, \end{cases}$$

with

$$r = \frac{(1-\alpha s) \pm \sqrt{(\alpha s - 1)^2 - 4\alpha^2 pq}}{2\alpha p}$$

in the case of v_j, and

$$r = \frac{(1-\alpha s) \pm \sqrt{(\alpha s - 1)^2 - 4\alpha^2 pq}}{2\alpha q}$$

in the case of z_j.

We now have the necessary equations to be able to find the optimal value of our primal problem. We can verify that our solution is in fact optimal by solving the equations for the dual problem and making sure that it is *no gap*. But, in advance not knowing the value of j^* aggravates the problem.

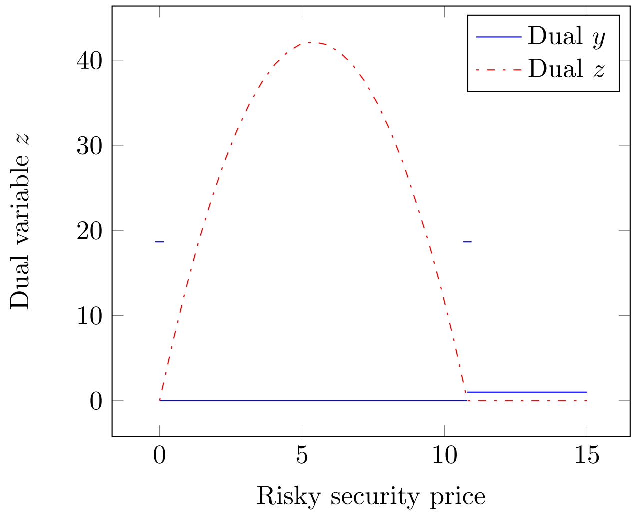

Figure 6.2: Plot of optimal values of y and z corresponding to the case where $\alpha = 0.999$, $p = q = s = 1/3$, $\Delta x = 0.1$, $K = 9$, $j^* = 108$ and $x_{j^*} = 10.8$. Note, the values for y at 0 and at j^* are enlarged to be more visible.

Check the inequalities Verifying that the various inequalities from the dual problem (6.3) and the primal problem (6.4) are satisfied is both a necessity, and gives a formula for finding the optimal exercise point.

$$y_j \geq 0, \qquad j \geq 0, \tag{6.21}$$
$$z_j \geq 0, \qquad j \geq 0, \tag{6.22}$$
$$v_j \geq Y_j, \qquad j \geq 0, \tag{6.23}$$
$$v_j \geq \alpha(pv_{j+1} + qv_{j-1} + sv_j), \qquad j \geq 1. \tag{6.24}$$

Inequality (6.22) verifies trivially for $j \geq j^*$ by (6.13). For $j < j^*$ we follow the reasoning of Vanderbei & Pınar [7], assuming there exists a k at which z_k is negative. We know that $z_0 = 0$ by (6.17) and that $z_j = 0$ for $j \geq j^*$.

There are two possibilities. Either z_k is a local minimum or it's not. If z_k is not a local minimum, either $z_{k-1} < z_k$ or $z_{k+1} < z_k$. In the first case we must have either z_{k-1} being a local minimum or $z_{k-2} < z_{k-1}$. If z_{k-1} is not a local minimum we continue in the same fashion. We are then bounded to encounter a local minimum either at $z_1 < 0$ (since $z_0 = 0$) or earlier. In the second case the reasoning is identical and yields a local minimum either at $z_{j^*} < 0$ (since $z_{j^*} = 0$) or earlier.

Now, let z_k be negative and a local minimum. Then

$$z_k \leq z_{k-1} \quad \text{and } z_k \leq z_{k+1}. \tag{6.25}$$

Combining (6.25) with (6.17) yield

$$(1 - \alpha s)z_k = 1 + \alpha q z_{k+1} + \alpha p z_{k-1} \geq 1 + \alpha q z_k + \alpha p z_k.$$

Rearranging we get

$$\begin{aligned} z_k &\geq 1 + \alpha z_k(p + q + s) \\ &\geq 1 + \alpha z_k, \end{aligned}$$

which implies $z_k \geq \frac{1}{1-\alpha} > 0$. We recall the assumption of future money being worth less than todays money, thus $\alpha < 1$. This implies z_k being positive and we have a contradiction.

Verifying inequality (6.21) is a simple matter as we now can combine the fact that $z_j \geq 0$ and the formula for y_j:

$$y_j = \begin{cases} 1 + \alpha q z_1, & j = 0, \\ 0, & 0 < j < j^*, \\ 1 + \alpha p z_{j^*-1}, & j = j^*, \\ 1, & j^* < j, \end{cases}$$

For inequality (6.24) we see from (6.8) that it holds for $j < j^*$. For $j \geq j^*$ we know that $v_j = Y_j$ by (6.9). Let $j = j^* + k$, for $k = 1, 2, \ldots$, i.e., the j's we want to examine, except for the missing $j = j^*$. We will deal with this later. Recall the notation $x_j = j\Delta x$ and $Y_j = Y(x_j)$ together with $v_{t+1} = v(x + \Delta x)$ if and only if $v_t = v(x)$. We are ready to do some computations:

$$\begin{aligned}
\alpha(pv_{j+1} + qv_{j-1} + sv_j) &= \alpha(pv_{j^*+k+1} + qv_{j^*+k-1} + sv_{j^*+k}) \\
&= \alpha(pY_{j^*+k+1} + qY_{j^*+k-1} + sY_{j^*+k}) \\
&= \alpha\big(p(Y_{j^*} + (k+1)\Delta x) + q(Y_{j^*} + (k-1)\Delta x) + s(Y_{j^*} + k\Delta x)\big) \\
&= \alpha\big(Y_{j^*}(p+q+s) + k\Delta x(p+q+s) + \Delta x(p-q)\big) \\
&= \alpha(Y_{j^*} + k\Delta x + \Delta x(p-q) \\
&= \alpha(Y_{j^*+k} + \Delta x(p-q)).
\end{aligned}$$

We want the last expression to be less than or equal to $v_j = v_{j^*+k} = Y_{j^*+k}$. There are more than one way to make this work, for instance assuming $q \geq p$. Following Vanderbei & Pınar [7] we assume

$$1/3 \geq \begin{cases} \alpha q, \\ \alpha p, \\ \alpha s. \end{cases}$$

Then

$$\begin{aligned}
\alpha(pv_{j+1} + qv_{j-1} + sv_j) &\leq \frac{1}{3}(v_{j+1} + v_{j-1} + v_j) \\
&= \frac{1}{3}(v_{j^*+k+1} + v_{j^*+k-1} + v_{j^*+k}) \\
&= \frac{1}{3}(3Y_{j^*+k}) \\
&= v_j.
\end{aligned}$$

Inequality (6.23) follows trivially for $j \geq j^*$ by (6.9). For $j < j^*$ we have

$$v_j = Y_{j^*} \frac{r_+^j - r_-^j}{r_+^{j^*} - r_-^{j^*}}$$

by (6.18), while we want $v_j \geq Y_j$. Thus we need j^* to be such that

$$Y_{j^*} \frac{r_+^j - r_-^j}{r_+^{j^*} - r_-^{j^*}} \geq Y_j, \quad j < j^*. \tag{6.26}$$

Recall that $v_{j^*} = Y_{j^*}$ and by (6.26) we have $v_{j^*-1} \geq Y_{j^*-1}$. Following the idea of Vanderbeid & Pınar [7], we assume that $v_{j'} < Y_{j'}$ for some $j' < j^*$. Then the sequence $u_j := v_j - Y_j$ must have a local maximum point, say k, strictly between j' and j^*. Even better, we can assume k to have strict inequalities on at least one side. That is, $u_k < u_{k+1}$ or $u_k < u_{k-1}$. This is a natural consequence of $u_{j'} < 0$, $u_{j^*-1} \geq 0$ and $u_{j^*} = 0$. Optionally, see the explanation for the local minimum for inequalities (6.22). Computing u_k yields

$$\begin{aligned}
u_k &= v_k - Y_k \\
&= \alpha(pv_{k+1} + qv_{j-1} + sv_j) - \frac{1}{3}(Y_{k+1} + Y_{k-1} + Y_k) \\
&\leq \frac{1}{3}(v_{k+1} + v_{k-1} + v_k) - \frac{1}{3}(Y_{k+1} + Y_{k-1} + Y_k) \\
&= \frac{1}{3}(u_{k+1} + u_{k-1} + u_k) \\
&< u_k.
\end{aligned}$$

This is a contradiction. Thus, it's not possible to have $v_j < Y_j$ for any $j < j^*$, and the inequality is verified. The splitting of Y_k follows from

$$\begin{aligned}
Y_k &= x_k - S \\
&= \frac{1}{3}\big(3(x_k - S + \Delta x - \Delta x)\big) \\
&= \frac{1}{3}(Y_{k+1} + Y_{k-1} + Y_k).
\end{aligned}$$

We are only left with inequalities (6.24), in the case of $j = j^*$. Some groundwork is needed in order to be able to verify this equation.

We are looking for the optimal exercising point j^*. By (6.26) we need

$$j^* \in K := \big\{k : Y_k \frac{r_+^j - r_-^j}{r_+^k - r_-^k} \geq Y_j\big\}, \quad j < j^*.$$

Since we are only dealing with j's less than j^*, we can assume

$$j^* \in K := \big\{k : Y_k \frac{r_+^j - r_-^j}{r_+^k - r_-^k} > Y_j\big\}, \quad j < j^*.$$

This follows from the assumption of j^*, see equation (6.5) and (6.8). Clearly, we must assume K to be non-empty. In addition,

$$\begin{aligned} & Y_{j^*}\frac{r_+^{j^*-1}-r_-^{j^*-1}}{r_+^{j^*}-r_-^{j^*}} > Y_{j^*-1} \\ \Leftrightarrow\quad & \frac{r_+^{j^*-1}-r_-^{j^*-1}}{r_+^{j^*}-r_-^{j^*}} > \frac{Y_{j^*-1}}{Y_{j^*}} \\ \Leftrightarrow\quad & \frac{r_+^{j^*-1}-r_-^{j^*-1}}{r_+^{j^*-1}r_+-r_-^{j^*}} > \frac{Y_{j^*-1}}{Y_{j^*}} \end{aligned} \tag{6.27}$$

As j^* inreases, the right hand side of the last equation of (6.27) is getting closer to 1,

$$\lim_{j^*\to\infty}\left\{\frac{Y_{j^*-1}}{Y_{j^*}}\right\} = 1.$$

For the left hand side we will need the following results on r_- and r_+, see section A.1 in the appendix.

$$\begin{aligned} 0 < r_-^j < 1, &\qquad \text{for all } j > 0, \\ r_+^j > 1, &\qquad \text{for all } j > 0. \end{aligned}$$

Then

$$\lim_{j^*\to\infty}\left\{\frac{r_+^{j^*-1}-r_-^{j^*-1}}{r_+^{j^*-1}r_+-r_-^{j^*}}\right\} = \lim_{j^*\to\infty}\left\{\frac{r_+^{j^*-1}}{r_+^{j^*-1}r_+}\right\} = \frac{1}{r_+}.$$

Since $r_+ > 1$ it follows that $\frac{1}{r_+}$ is a fixed number smaller than 1. Thus, the value of j^* must be bounded above in order for (6.27) to hold. In addition, we must assume that $j^*+1 \notin K$, making K bounded above as well. To see why, we follow Vanderbei's & Pınar's idea [7], letting w_j denote the solution to the difference equation

$$\begin{aligned} & w_j - \alpha(pw_{j+1} + qw_{j-1} + sw_j) = 0, \qquad 0 < j, \\ & w_0 = 0, \\ & w_{j^*} = Y_j^*. \end{aligned} \tag{6.28}$$

This is similar to the difference equation for the primal problem (6.16), but extended to all j. Solving (6.28) will yield

$$w_j = \begin{cases} 0, & j = 0, \\ Y_{j^*}\frac{r_+^j - r_-^j}{r_+^{j^*}-r_-^{j^*}}, & 0 < j < j^* \text{ and } j^* < j, \\ Y_{j^*}, & j = j^*. \end{cases}$$

Clearly, $v_{j^*} = w_{j^*}$ and $v_{j^*-1} = w_{j^*-1}$. Since w_j solves the difference equation from the primal problem (6.3) we have

$$w_{j^*} \geq \alpha(pw_{j^*+1} + qw_{j^*-1} + sw_{j^*}).$$

Inserting for v_j we get

$$v_{j^*} \geq \alpha(pw_{j+1} + qv_{j-1} + sv_{j^*}).$$

For inequality (6.24) to be true, in the case of $j = j^*$, we need to assume that $v_{j^*+1} \leq w_{j^*+1}$. Then

$$Y_{j^*+1} = v_{j^*+1} \leq w_{j^*+1} = Y_{j^*} \frac{r_+^{j^*+1} - r_-^{j^*+1}}{r_+^{j^*} - r_-^{j^*}}.$$

Rearranging, we get

$$Y_{j^*+1} \frac{r_+^{j^*} - r_-^{j^*}}{r_+^{j^*+1} - r_-^{j^*+1}} \leq Y_{j^*}.$$

That is, $j^* + 1 \notin K$. This implies

$$j^* := \max\left\{k : Y_k \frac{r_+^j - r_-^j}{r_+^k - r_-^k} > Y_j\right\}, \quad j < j^*.$$

In addition, both $Y_k \frac{r_+^j - r_-^j}{r_+^k - r_k}$ and Y_j are increasing functions, where Y_j becomes the greater one for $j \geq j^*$. See section A.2 in the appendix. That is, if

$$Y_k \frac{r_+^{k'} - r_-^{k'}}{r_+^k - r_-^k} > Y_{k'},$$

with $k' < j^* - 1$, then

$$Y_k \frac{r_+^{k'+1} - r_-^{k'+1}}{r_+^k - r_-^k} > Y_{k'+1}.$$

Thus, we only need to check for $j = k - 1$:

$$j^* := \max\left\{k : Y_k \frac{r_+^{k-1} - r_-^{k-1}}{r_+^k - r_-^k} > Y_{k-1}\right\}.$$

We have verified all the inequalities, and we have a formula for the optimal exercise point j^*. In Figure 6.2 the fair price of a perpetual American call option is computed. With $p = q = s = 1/3$, $\Delta x = 0.1$, $\alpha = 0.999$ and $K = 9$ the optimal exercise j^* is 108. That is, when the risky security price is 10.8. Thus, the fair price of the option will be 1.8.

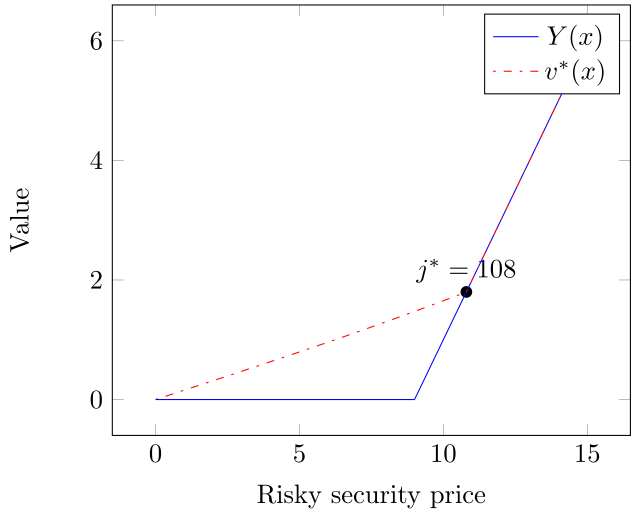

Figure 6.2: Plot of $v^*(x)$ and $Y(x)$ corresponding to the case where we have a perpetual American call option with $\alpha = 0.999$, $p = q = s = 1/3$, $\Delta x = 0.1$, $K = 9$. In this case, $j^* = 108$ and $x_{j^*} = 10.8$.

Comparing Lets compare the results from the random Walk model by Vanderbei & Pınar [7], with the ternary Markov chain model. Letting $s = 0$ in the last case yields

$$r = \frac{1 \pm \sqrt{1 - 4\alpha^2 pg}}{2\alpha p},$$

in the case of v_j. Now, with this choice of s, the solution in the Markov chain model is identical to the random walk solution [7, p. 770]. For instance, solving the example in Vanderbei & Pınar [7] with the Matlab code (see section A.3 in the appendix) adapted for the Markov chain model, yields the same solution. We conclude that the Markov chain model is a similar, but extended, model. It allows a richer and larger set of possible scenarios.

7 For Further Study

There are lots of topics not mentioned in this work, that may be considered as normal to include when talking about pricing derivatives. Attainability and complete markets are two topics of interest. The theory presented for the European and American options has managed to include both attainable and nonattainable derivatives, and apply to both complete and incomplete market models. This was to avoid the master thesis being to lengthy. A discussion about the existence of several fair prices for a single derivative is of importance if this was something to practice in a real world market.

It would also be interesting to look at other models, for instance, Vanderbei & Pınar has done a similar job with a geometric random walk model [8]. One could, in a similar way as in this thesis, modify the model to include more different type of scenarios.

We mentioned in the early beginning that certain details as transaction cost, lending rates and so forth was ignored.[22] Both Pliska [5] and Cutland & Roux [3] show how some of this can be implemented in the models, and it is only a matter of adding extra details, not actually changing any of the mathematical structure.

At equation (6.7) we made a decision and left out the put options. We could include them if it had brought new theory or new methods to arrive at the result, but the outcome would only been minor differences in the computation. Nevertheless, the put options are just as interesting and important as the call options.

7.1 Pricing perpetual American Options With a Risk Neutral Probability Measure

What was so great about pricing derivatives with an equivalent martingale measure, was that we found a fair price, without including any real world probability. That is, with one less factor to assume in the model.[23] The problem of finding an equivalent martingale measure in a perpetual model is that the number of scenarios is increasing as time progresses. Therefore, the number of scenarios would be infinite while the probability of each scenario would be zero,

$$|\Omega| = \infty \implies q_\omega = 0, \ q \in \mathbb{Q}, \ \omega \in \Omega.$$

Obviously, we are able to price American options in a finite model, as we saw in chapter 2.2. An approach to pricing perpetual American options is to first classify them into two groups:

- Those who will be exercised at some time $t < \infty$.
- Those who will never be exercised.

We are only interested in the options in the first group. Then, with whatever model we want to use, assume the number of trading dates to be T_1, where T_1

[22] Assumptions known as *friction* in mathematical finance.
[23] No one knows the true probabilities of the different movements of the risky securities in the real world.

is finite and positive. Next, price the American option in this finite model. The following step is crucial:

- If the solution, for any scenario, is to exercise at the final time T_1, start over again, but now let T_1 be greater. E.g. $T_2 := T_1 + 1$.
- If the solution, for all scenarios, is to exercise before time ${}_1T$, we are done.

We already know that the option will be exercise for some time $t < \infty$. As T increases, we will at some point be done. The model might be to complex to compute for large values of T, but it will be without any assumptions of a real world probability measure.

A Appendix

A.1 The Roots r_+ and r_-

We recall the assumption

$$1/3 \geq \begin{cases} \alpha q, \\ \alpha p, \\ \alpha s, \end{cases}$$

together with

$$p + q + s = 1 \text{ and } \alpha < 1,$$

implies it's impossible to have $\alpha p = \alpha q = \alpha s = 1/3$ at the same time. That is,

$$\alpha p + \alpha q + \alpha s < 1.$$

Next, the greatest possible value of $4\alpha^2 pq$ is $4/9$. Then, from the observation above, $\alpha s < 1/3$, implying

$$(\alpha s - 1)^2 > 4/9.$$

If

$$4\alpha^2 pq < 4/9 \quad \Rightarrow \quad (\alpha s - 1)^2 \geq 4/9 > 4\alpha^2 pq.$$

We conclude that

$$(\alpha s - 1)^2 - 4\alpha^2 pg > 0.$$

.

Starting with r_+ we compute

$$\begin{aligned} r_+ &= \frac{(1-\alpha s) + \sqrt{(\alpha s - 1)^2 - 4\alpha^2 pq}}{2\alpha p} \\ &= \frac{1-\alpha s}{2\alpha p} + \frac{\sqrt{(\alpha s - 1)^2 - 4\alpha^2 pq}}{2\alpha p} \\ &\geq \frac{1-1/3}{2/3} + \frac{\sqrt{(\alpha s - 1)^2 - 4\alpha^2 pq}}{2\alpha p} \\ &= 1 + \frac{\sqrt{(\alpha s - 1)^2 - 4\alpha^2 pq}}{2\alpha p} \\ &> 1. \end{aligned}$$

Working with r_-, we see that

$$(1-\alpha s) > \sqrt{(\alpha s - 1)^2 - 4\alpha^2 pg}. \tag{A.1}$$

Obivously, (A.1) is true if and only if both p and q are non-zero. But, assuming $p = q = 0$ is equivalent to assume that the risky security moving neither up or down, which is a situation we are not interested in in this work. Thus

$$r_- = \frac{(1-\alpha s) - \sqrt{(\alpha s - 1)^2 - 4\alpha^2 pq}}{2\alpha p} > 0.$$

In addition, plotting r_- and manually test for different values of p, q, s and α shows that $r_- < 1$. See figure A.2. This is not a valid proof, but an indication, which we will use as a valid result.

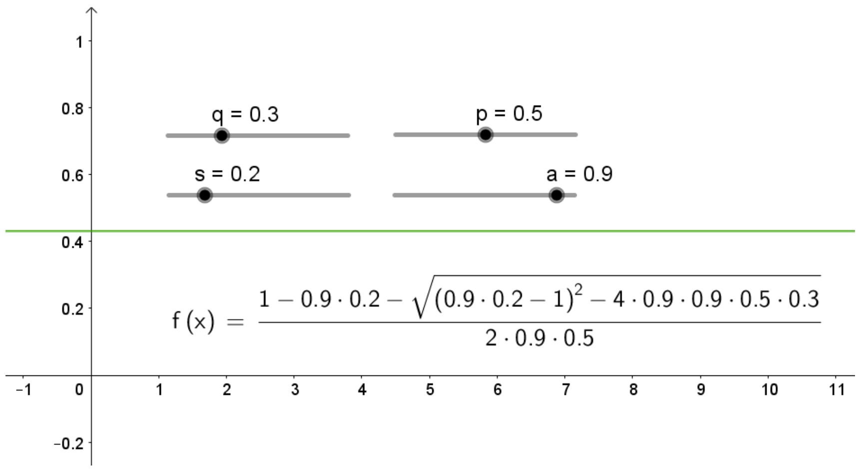

Figure A.2: A plot of r_- for some specific values of p, q, s and α, showing that r_- is less than 1.

Changing p with q in the computations above would have made no difference on the results. We conclude that

$$\begin{aligned} 0 < r_-^j &< 1, \qquad \text{for all } j > 0, \\ r_+^j &> 1, \qquad \text{for all } j > 0. \end{aligned}$$

This is valid for both

$$r = \frac{(1-\alpha s) \pm \sqrt{(\alpha s - 1)^2 - 4\alpha^2 pq}}{2\alpha p},$$

and

$$r = \frac{(1-\alpha s) \pm \sqrt{(\alpha s - 1)^2 - 4\alpha^2 pq}}{2\alpha q}.$$

Comments One could try to verify that $r_- < 1$ by assuming different restrictions on the variables, e.g. $(1-\alpha s) < 2\alpha p$, $(1-\alpha s) = 2\alpha p$ and $(1-\alpha s) > 2\alpha p$. The first two cases are easy to verify. For the last, one could assume even more restrictions. But, even with this strategy it eventually stops, without a complete verification. One solution could be to make more assumptions about the variables p, q, s and α. Instead of further restricting possible values of the variables in the model, it was decided to go for the plotting in Figure A.2.

A.2 The Increasing Value Function

For $j < j^*$ we have

$$v_j = Y_{j^*} \frac{r_+^j - r_-^j}{r_+^{j^*} - r_-^{j^*}}.$$

Note that both Y_{j^*} and $r_+^{j^*} - r_-^{j^*}$ are constants greater than zero. Let $f(x) = r_+^x - r_-^x$, for $x > 0$. Then

$$f'(x) = r_+^x \ln(r_+) - r_-^x \ln(r_-).$$

Since $f'(x) > 0$ for all $x > 0$, we conclude that v_j is increasing for all $j < j^*$.

A.3 Script Used To Compute Examples

a=0.999;
%The discount factor.

p=0.5;
%The probability of an «up» move of the risky security.

q=0.5;
%The probability of a «down» move of the risky security.

s=0;
%The probability of no change in value of the risky security.

K=9;
%The exercise price of the option.

dx=0.1;
%The change of value of the risky security in case of p or q.

rp=(1-a*s+sqrt((a*s-1)$^2 - 4 * a^2 * p * q))/(2 * a * p)$;
%Roots from the solution of the difference equation (pluss).

rm=(1-a*s-sqrt((a*s-1)$^2 - 4 * a^2 * p * q))/(2 * a * p)$;
%Roots from the solution of the difference equation (minus).

xans = fzero(@(x) ((x*dx-K)*(rp$^{(x-1)} - rm^{(x-1)})/(rp^x - rm^x) - x * dx + K + dx), [1\ K/(a * dx)]$);
%Numerical solution of the optimal exercise point. The given problem was an inequality. To solve this, we switch the inequality with an equality sign, and find the zeroes. Choose lower boundary at 1, and upper boundary such that the function values at the interval endpoints differ in sign. The fzero function finds a zero in the given interval, and @(x) assigns the variable x with the written function. Without @(x), x would be classified as 'undefined'.

joptimal = floor(xans)
%The exact solution xans is probably not always an integer. Since the term we are solving is decreasing for non-negative values of x, we use the floor function (and not the ceil or round functions) to get an integer solution.

References

[1] Binomial model. https://en.wikipedia.org/wiki/Binomial_options_pricing_model. Accessed: 2018-06-10.

[2] Complementary slackness conditions. http://www.matem.unam.mx/omar/math340/comp-slack.html. Accessed: 2018-06-12.

[3] N. J. Cutland and A. Roux. *Derivative Pricing in Discrete Time*. Springer Science & Business Media, 2012.

[4] K. Mørken. Numerical algorithms and digital representation. *Lecture Notes for course MATINF1100 Modelling and Computations*, 2016.

[5] S. Pliska. *Introduction to mathematical finance*. Blackwell publishers Oxford, 2001.

[6] R. J. Vanderbei et al. *Linear programming*. Springer, 2008.

[7] R. J. Vanderbei and M. C. Pınar. Pricing american perpetual warrants by linear programming. *SIAM review*, 51(4):767–782, 2008.

[8] R. J. Vanderbei, M. Ç. Pınar, and E. B. Bozkaya. Discrete-time pricing and optimal exercise of american perpetual warrants in the geometric random walk model. *Applied Mathematics & Optimization*, 67(1):97–122, 2013.