Finance Without Price Dynamics

by

Natasha Bushueva

B.S., Moscow Institute of Physics and Technology, 1997
M.S., Moscow Institute of Physics and Technology, 1998

Submitted to the Department of Mathematics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2003

© Natasha Bushueva, MMIII. All rights reserved.

The author hereby grants to MIT permission to reproduce and
distribute publicly paper and electronic copies of this thesis document
in whole or in part.

Author

Department of Mathematics

April 3, 2003

Certified by

Dimitris Bertsimas
Boeing Professor of Operations Research
Sloan School of Management
Thesis Supervisor

Accepted by

Rodolfo Ruben Rosales
Chairman, Applied Mathematics Committee

Accepted by

Pavel Etingof
Chairman, Department Committee on Graduate Students
Finance Without Price Dynamics

by

Natasha Bushueva

Submitted to the Department of Mathematics
on April 3, 2003, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

Traditional finance theory has addressed portfolio optimization and option pricing problems assuming specific price dynamics for securities. In this thesis we address these problems without assuming specific price dynamics.

We consider the problem of determining bounds on the price of an option based on observable prices of other options and the no-arbitrage assumption. This problem has been solved for the case of European options of a single maturity. We find exact bounds on the price of European options, when options of multiple maturities are given. We also find exact bounds if the payoff function on the option is a piecewise linear function of the price of the underlying security. We solve a similar problem in a two dimensional case. The methods developed for two-dimensional case are applicable to the multiple dimensional case, however the number of variables grows exponentially with the dimension. We also investigate how the optimal bounds in the one-dimensional case change, if, in addition to the set of observable prices, there is a condition on the variance of the underlying asset. We derive closed form expressions for the tight upper bound on the variance and propose a polynomial time algorithm for obtaining the lower bound.

In the portfolio allocation problem we prove that for large time horizons, there exists a myopic policy which leads to a distribution of the terminal wealth with the property that the probability of underperforming any other policy tends to zero as the horizon tends to infinity. Finally, we address the problem of maximizing a mean-variance function of the terminal wealth in the multiperiod case. For general price dynamics of securities we propose a monte-carlo based method for the solution which is polynomial in the number of securities.

Thesis Supervisor: Dimitris Bertsimas
Title: Boeing Professor of Operations Research
Sloan School of Management
Acknowledgments

I would like to thank my thesis advisor, Prof. Dimitris Bertsimas, for suggesting the problems of the thesis and his useful comments. I would also like to thank Prof. Richard Dudley for his careful comments on parts of the thesis.
## Contents

1 Introduction .................................................. 5
   1.1 A short history of option pricing .......................... 5
   1.2 Contributions on option pricing .......................... 9
   1.3 An outline of Chapters 2 and 3 ........................... 13
   1.4 On portfolio optimization ................................... 18
   1.5 Contributions on portfolio optimization ................... 19

2 Option Pricing Without Price Dynamics. One Dimensional Case. 21
   2.1 Structural theorems ........................................... 23
   2.2 Upper and lower bounds on option prices .................... 33
       2.2.1 Derivation ............................................. 33
       2.2.2 Algorithm ............................................ 40
       2.2.3 Computational example ................................ 42
   2.3 Bounds on option if the variance of the underlying asset is known 43
   2.4 Bounds on variance ......................................... 45

3 Option Pricing Without Price Dynamics, the Two Dimensional Case. 65
   3.1 Structural theorems .......................................... 67
       3.1.1 Restriction to Markov martingales .................... 67
       3.1.2 Treatment of future conditions ........................ 68
       3.1.3 Restriction to discrete distributions .................. 71
   3.2 Algorithm .................................................. 76
   3.3 Computational example ...................................... 78
   3.4 An approximation approach .................................. 78
       3.4.1 Algorithm ............................................ 82
       3.4.2 Approximation approach for the special case of a payoff function 83
3.5 Application of the two-dimensional approach to the one-dimensional case ........................................ 84
3.6 Generalization to a multiple dimensional case ................................................................. 86
3.7 Static arbitrage bounds on basket option prices ............................................................... 87

4 Portfolio Optimization With General Price Dynamics .......................................................... 91
4.1 Introduction ....................................................................................................................... 91
4.2 On the consistency of optimal policies ........................................................................... 92
4.3 On the asymptotic optimality of the myopic policy ....................................................... 94
4.4 Multi-period mean-variance optimization ........................................................................ 99
  4.4.1 Algorithm ................................................................................................................ 100
  4.4.2 Example ................................................................................................................ 101
  4.4.3 Analysis .................................................................................................................. 101
  4.4.4 Numerical experiments .......................................................................................... 111
Chapter 1

Introduction

Traditional finance theory has addressed portfolio optimization and option pricing problems assuming specific price dynamics for securities, for example, geometric Brownian motion for the stock price. In this thesis we address these problems without assuming specific price dynamics.

In Sections 1.1-1.3 we discuss the history of option pricing, the problems that we address and our contributions. In Sections 1.4-1.5 we discuss the portfolio optimization problem that we address and our contributions.

1.1 A short history of option pricing

A call (put) option on a stock is a contract that gives its holder the right (but not the obligation) to buy (sell) the stock in the future at price $k$, called the strike. A European call (put) option allows to exercise the contract at a particular date $T$, called the maturity date. An American call (put) option allows to exercise the contract at any time before or at time $T$. These options are among the most common financial instruments in the market over the last thirty years, and thus, drew a lot of attention in both the business and the academic literature. While the problem of pricing an American call option seems to be harder, it has been shown that prices of American and European call options are the same. This is not the case for put options.

The classical theory of option pricing initiated by Black and Scholes [5] is based on two assumptions: (a) the assumption of no-arbitrage, i.e., it is not possible to construct a zero investment portfolio that will yield a sure profit, (b) the price dynamics
of the stock price is a geometric Brownian motion. Based on these assumptions Black and Scholes [5] derive their celebrated option pricing formula.

We investigate in this thesis results that can be derived by only assuming no-arbitrage but without making an assumption on the underlying price dynamics.

Harrison and Kreps [20] found the necessary and sufficient condition on securities price processes for the existence of no-arbitrage. To state the result, we introduce some notations and the mathematical model first.

Denote the price of a security (stock) at time \( t \) by \( S(t) \). Then \( \{S(t)\}_{t \in [0,T]} \) can be interpreted as a stochastic process defined on a probability space \((\Omega, \mathcal{F}, P)\).

Let \( r(s) \) designate the instantaneous riskless rate of return. Then

\[
\{e^{-\int_0^t r(s)ds} S(t)\}_{t \in [0,T]}
\]

is called the discounted security price process.

Harrison and Kreps [20] show that the condition of no-arbitrage is equivalent to the existence of a probability measure \( Q \) equivalent to \( P \), such that price processes of all discounted securities are martingales under \( Q \). (Two measures are called equivalent if they have the same null sets.) This measure is also often referred to as the risk-neutral measure.

Now we can return to the question of option pricing. An option on a stock has its value at maturity specified by some function \( f(S(T)) \), called a payoff function. In the case of a European call option \( f(S(T)) = (S(T) - k)^+ = \max(S(T) - k, 0) \). If \( C(t) \) is the price of the option at time \( t \), then from the no-arbitrage condition it follows that

\[
C(t) = e^{-\int_0^t r(s)ds} E_Q[f(S(T))]. \tag{1.1}
\]

As we have already mentioned, Black and Scholes [5] make the additional assumption that the stock price follows a geometric Brownian motion process:

\[
dS(t) = \mu S(t) dt + \sigma S(t) dB.
\]

From the Markov property of \( S(t) \) and (1.1) it follows that the price of the option at time \( t \) is a function of \( S(t) \) and \( t \), that is \( C = C(S, t) \). Applying Itô's lemma and taking into account the no-arbitrage condition, Black and Scholes derive the partial
differential equation

$$\frac{\partial C}{\partial t} + rS\frac{\partial C}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} = rC. \tag{1.2}$$

The key boundary condition is $C(S, T) = f(S(T))$.

For $f(S(T)) = (S(T) - k)^+$ (and $r(s) \equiv r$) the celebrated Black-Scholes formula is

$$C(t) = S(t)N(d_1) - ke^{-(T-t)}N(d_2), \tag{1.3}$$

where

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

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

and $N(x)$ is a cumulative probability distribution function for a variable that is normally distributed with a mean of zero and a standard deviation of 1.

There is empirical evidence, however, that the underlying price dynamics is not indeed a geometric Brownian motion. In practical terms, this manifests itself in that for Formula (1.3) to match observed option prices, the volatility parameter need to be a function of the strike (what is called the volatility smile.)

Thus, a new approach emerged which considers market prices of options as given and view them as a source of information on the market.

Ross [41] introduces the idea that it is possible in principle to infer the probability distribution of the underlying asset at time $T$ corresponding to the martingale measure $Q$ from the complete set of European option prices with maturities $T$.

For the limited availability of option prices, various computational methods were proposed for recovering the risk-neutral probability distributions at time $T$. A summary of them can be found in Jackwerth [22] and Cont [9].

Jackwerth [22] classifies these methods as parametric and non-parametric. Parametric methods are divided into three groups. First, expansion methods start with a basic distribution, and add correction terms in order to make it more flexible. Second, generalized distribution methods introduce more flexible distributions with additional parameters beyond the two parameters of the normal or lognormal distributions. Third, mixture methods create new distribution from mixture of simple distributions.
The non-parametric methods are also divided into three groups. First, kernel methods are conceptually related to regressions in that they try to fit a function to observed data. The main difference is that kernel regression does not specify the parametric form of the function. Second, maximum-entropy methods find a nonparametric probability distribution, which is as close as possible in terms of information content to prior distribution while satisfying certain constraints such as pricing observed options correctly. Third, curve fitting methods are a loosely associated group of methods that try to fit the implied volatilities or the risk-neutral probability as well as possible with some flexible function. A commonly used measure of fit is the sum of squared differences between observed option prices and option prices suggested by the curve fitting method.

To cover the whole stochastic process of asset price across all times the method of implied binomial trees has been developed by Cox, Ross and Rubinstein [12]. The binomial tree model can be used to extract from option prices an "implied tree", the parameters of which are conditioned to reproduce correctly a set of observed option prices. Rubinstein [42] proposes an algorithm which, starting from a set of option prices at a given maturity, constructs an implied binomial tree which reproduces them exactly. The tree then can be used to price other options.

However, the uniqueness of the distribution obtained by the above methods is due to the constraints imposed or implied by the model. In general the observed option prices can be supported by a whole set of risk-neutral distributions (martingale measures.) Thus, with no additional assumptions, only bounds on option prices can be guaranteed.

Bertsimas and Popescu [4] investigate the relation between options based only on the no-arbitrage assumption. Given observable option prices on a single underlying asset, they provide best possible bounds on option prices as well as on moments of the prices of the asset. The solution is based on convex optimization methods and duality theory. This paper has inspired the problems considered in the option pricing part of this thesis.

d'Aspremont and Ghaoui [2] address the problem of computing bounds on the price of a European basket call option, given prices on other similar baskets. In mathematical terms, given \( C \in \mathbb{R}^{n+}, \ k \in \mathbb{R}^{n+}, \ w_i \in \mathbb{R}^n, \ i = 0, \ldots, m \) and \( k_0 > 0 \), find upper and lower bounds on

\[
E_Q \left( w_0^T X_t - k_0 \right)^+, \]

8
with respect to all probability distributions $Q$ with nonnegative support of the asset price vector $X_t = (X_t^1, X_t^2, \ldots, X_t^n)^1$ under the condition

$$E_Q \left( w_i^T X_t - k_i \right)^+ = C_i, \quad i = 1, \ldots, m.$$ 

Their bounds are not necessarily tight except in some special cases. For the simpler version of the problem of determining the upper and lower bounds on

$$E_Q \left( w_0^T X_t - k_0 \right)^+,$$

given $2n$ constraints

$$E_Q(X_i^i - k_i) = C_i, \quad E_Q X_i^i = q_i, \quad i = 1, \ldots, n,$$

where $k_0 > 0$ and $w_0, k, C, q$ are given vectors of $\mathbb{R}^n^+,$ they find a tight upper bound and a lower bound that they conjecture to be tight. The idea of the solution is based on the duality theory.

The prior related work includes the paper by Lo [28], who derives best-possible closed-form bounds on the price of a European call option, given the mean and the variance of the underlying stock price; by Grundy [17], who extends Lo's work for the case when the first and the $k$th moments of the stock price are known; by Boyle and Lin [6], who use semidefinite optimization to find an upper bound on the price of a European call option on the maximum of a number of assets, given the means, variances, and covariances of these assets.

All the above papers consider the static problem, that is they only consider options of the same maturity $T$ and moments of the underlying asset at time $T.$

### 1.2 Contributions on option pricing

In this thesis we address the following problems

(a) The first problem that we consider is determining tight upper and lower bounds on the price of a European call option with given strike price and maturity, based

\footnote{We will use $X_t$ instead of $S(t),$ the price of the stock, to underline that we operate in the environment with zero interest rates, so that \( \{ X_t \}_{t \geq 0} \) is a martingale under $Q.$ We will show later that in all problems that we consider interest rates can be assumed zero without loss of generality.}
on the set of options of different maturities extending the work of Bertsimas and Popescu [4]. As in the case of a static problem, the interest rate is not essential here. (Notice, that this is not true for all options. For example, in formulating an analogous problem for American put options, interest rates cannot be assumed zero, as the problem would become qualitatively different.) In mathematical terms: given an ordered set

\[ S := \{(k_{ij}, C_{ij}) \mid (k_{ij}, C_{ij}) \in \mathbb{R}^{2+}, \ t \in \{1, 2, \ldots, n\}, \ j \in \{1, \ldots, U(t)\}\} \] (1.4)

and \( k^* \in \mathbb{R}^{+} \) and \( t^* \in \{1, 2, \ldots, n\} \), we are interested in determining the domain of possible values of

\[ E \left[ (X_{t^*} - k^*)^+ \right], \] (1.5)

under the condition, that \( \{X_1, X_2, \ldots, X_{t^*}, \ldots, X_n\} \) is a nonnegative martingale such that

\[ E \left[ (X_t - k_{ij})^+ \right] = C_{ij} \ \text{for all} \ (k_{ij}, C_{ij}) \in S. \] (1.6)

(b) Given the same set of conditions as in (a), we want to determine the domain of

\[ E[g(X_{t^*})], \]

if \( g : \mathbb{R}^{+} \to \mathbb{R} \) is a given continuous piecewise linear function, i.e., we are interested in option prices with more general payoff.

(c) Given the same set of conditions as in (a), determine tight lower and upper bounds on the variance of the underlying asset at time \( t \in \{1, 2, \ldots, n\} \), that is \( Var(X_t) \). Similarly, we wish to extend problem (a) when \( Var(X_t) \) is given.

(d) We address a similar problem in the two dimensional case: given two sets of options on two individual assets, find tight bounds on an option with the payoff which is a function of both assets. The problem can be formulated as follows.

Given the set \( \{(k_{ij}^h, C_{ij}^h) \mid (k_{ij}^h, C_{ij}^h) \in \mathbb{R}^{2+}\}, (h, t, j) \in \mathcal{T}, \) where

\[ \mathcal{T} := \{(h, t, j) \mid h \in \{1, 2\}, \ t \in \{1, 2, \ldots, n\}, \ j = 1, 2, \ldots, U(t, h)\}, \]

10
and a continuous piecewise linear function $g : \mathbb{R}^2^+ \rightarrow \mathbb{R}$, find the maximum and minimum of $E[g(X_t^*)]$ for a given $t^* \in \{1, 2, \ldots, n\}$, under the condition that $X_1, X_2, \ldots, X_n$ is a two-dimensional nonnegative martingale, such that

$$E \left[ (X_t^h - k_{tj})^+ \right] = C_{tj}^h, \ (h, t, j) \in \mathcal{T}. \quad (1.7)$$

Here $h$ denotes coordinates of the two-dimensional martingale.

(e) Does the solution to problem (d) work for higher dimensions?

(g) We also address the problem of computing upper and lower bounds on the price of a European basket call option, given prices of other similar baskets, that is the problem considered by A. d’Aspremont and L. Ghaoui: given $C \in \mathbb{R}^{m^+}$, $k \in \mathbb{R}^{m^+}$, $w_i \in \mathbb{R}^n$, $i = 0, \ldots, m$ and $k_0 > 0$, find the upper and lower bounds on

$$E_Q \left( w_i^T X - k_0 \right)^+, \quad \text{with respect to all probability distributions } Q \text{ of an } n\text{-dimensional random variable } X \text{ with a nonnegative support under the condition}$$

$$E_Q \left( w_i^T X - k_i \right)^+ = C_i, \quad i = 1, \ldots, m.$$ 

**Contributions**

We state our contributions in the same order as the problems above are listed.

(a) We state a sufficient and necessary condition that a set of options of different maturities and the same underlying asset should jointly satisfy so that there exists no arbitrage. This is based on a theorem of Kertz-Rosler-Strassen [24, 43]. We give a simpler proof to that result and propose an algorithm for determining tight upper and lower bounds on the price of another option based on this set of options. The complexity of the algorithm in the worst case is at most $O(N^2)$, where $N$ is the number of options ($N = \sum_{t=1}^n U(t)$.)

(b) We find tight upper and lower bounds on the price of an option with a piecewise linear continuous payoff function based on the set of priced European call options of different maturities. Here an additional assumption is made, that $P(X_t \leq L) = 1$, $t = 1, 2, \ldots, n$, for some value $L > 0$, that is, there is an
upper bound on the price of the underlying asset. The solution requires solving a linear programming problem with the number of variables being at most $O\left((N_k + N_g)^2\right)$, where $N_k$ is the number of different strike prices $k_{ij}$ and $N_g$ is the number of points of $\mathbb{R}^+$, where the payoff function $g(X_{t^*})$ changes its derivative.

(c) We derive a tight upper bound on the variance of the underlying asset at time $t \in \{1, 2, \ldots, n\}$ and propose an algorithm for determining the tight lower bound. When there are no options of maturity smaller than $t$, the complexity of the algorithm is $O\left(N_t^2\right)$, where $N_t$ is the number of options of maturity $t$. If there are options of maturity smaller than $t$, then in the worst case the algorithm is exponential in $N_{<t}$, where $N_{<t}$ is the number of options of maturity smaller than $t$. We show that the problem of determining bounds on the price of an option, based on the set of other options and the variance of the underlying asset at a time $t$, is equivalent to the problem of determining bounds on the variance of the underlying asset at time $t$.

(d) We show how to reduce the problem to a linear programming problem. The number of variables in that problem grows exponentially with the number of different maturities smaller than or equal to $t^*$. The solution leads to the tight upper and lower bounds. The bound on the price of each stock is required, as in (b).

We also propose an approximation algorithm with the solution converging to the exact solution as the precision increases. The number of variables in the linear programming problem, that has to be solved in this case, grows linearly with $t^*$.

(e) All the algorithms proposed for the two-dimensional case are applicable in the multiple dimensional case. However the number of variables in linear programming problems, which are key to the solution, grows exponentially with the dimension.

(f) We notice, that the idea of the method proposed for the exact solution of problem (d) also works for problem (f). Exact upper and lower bounds can be found. Again, a solution of the linear programming problem is required. The number of variables is bounded from above by $O\left(C_{2n+m}^m\right)$, where $C_{2n+m}^m = \frac{m!2n!}{(2n+m)!}$.
and \( n \) and \( m \) are as defined in problem (f).

1.3 An outline of Chapters 2 and 3

Chapter 2 and Chapter 3 are devoted to problems on option pricing listed above. In Chapter 2 we consider the one-dimensional case and address problems (a), and (c), while in Chapter 3 we consider the two-dimensional case with extensions to multiple dimensions and address problems (d)-(f) and (b). The problem (b) is a one-dimensional problem, however the solution comes from the two-dimensional approach, so it is included in Chapter 3.

Chapter 2 starts with characterization theorems, which establish the necessary and sufficient condition for the set \( S \) (See Eq. 1.4) of options to satisfy the no-arbitrage condition.

For each probability law \( \Pi \) with a finite expectation and support in \( \mathbb{R}^+ \) we introduce a function (transform): \( \mathbb{R}^+ \to \mathbb{R}^+ \)

\[
\Psi_\Pi(t) = E_\Pi [(X - t)^+].
\]

We show (Theorem 1 and Proposition 1) that \( \Psi_\Pi(t) \)

- uniquely determines \( \Pi \), that is the inverse transform is uniquely defined;
- is a convex nonincreasing function of \( t \) which is above the line \( E[X] - t \) and approaches 0 as \( t \to \infty \).

If one-dimensional random variables \( \{X_1, X_2, \ldots, X_n\} \) represent stock prices at times corresponding to a given set of maturities, then \( \Psi \)-transforms of these random variables uniquely determine prices of options with corresponding maturities. We want to find a necessary and sufficient condition for the sequence of \( \Psi \)-transforms to satisfy in order to ensure the existence of a martingale with marginal laws defined by the given sequence of \( \Psi \)-transforms. Based on that it must be possible to find a necessary and sufficient condition that the given set of options needs to satisfy in order for the no-arbitrage condition to hold.

Theorem 3 establishes the relationship between the sequence of \( \Psi \)-transforms and the existence of a martingale:

if \( \Pi_1, \Pi_2, \ldots, \Pi_n \) are probability laws with finite expectations and support in \( \mathbb{R}^+ \), then there exists a martingale \( \{X_1, X_2, \ldots, X_n\} \), such that \( X_i \) has marginal law \( \Pi_i \),
for each \( i = 1, \ldots, n \) if and only if the sequence of functions \( \Psi_{\Pi_1}, \Psi_{\Pi_2}, \ldots, \Psi_{\Pi_n} \) is nondecreasing and \( \Psi_{\Pi_1}(0) = \Psi_{\Pi_2}(0) = \ldots = \Psi_{\Pi_n}(0) \).

That had been shown by Kertz and Rosler\[24\]. In particular they prove that for two measures \( \mu_1 \) and \( \mu_2 \) on \( \mathbb{R} \) the following two conditions are equivalent:

(i) \( \int (x-t)^+ d\mu_1(x) \leq \int (x-t)^+ d\mu_2(x) \) for all \( t \in \mathbb{R} \);

(ii) \( \int \phi d\mu_1 \leq \int \phi d\mu_2 \) for all nondecreasing convex functions \( \phi \) for which the integral exists.

Strassen \[43\] had previously shown that if \( \mu_1 \) and \( \mu_2 \) are two measures on \( \mathbb{R} \), such that they satisfy condition (ii) and have equal expectations, then there exists a martingale \( X_1, X_2 \) with corresponding marginals \( \mu_1 \) and \( \mu_2 \).

We discovered the above result after stating and proving the theorem necessary for the solution of our problem. Since our proof is different, uses the nonnegativity of the martingale and has a simple geometric argument, we include our original proof.

The algorithm for computing upper and lower bounds on the price of an option, given the set of priced options of different maturities, is based on Theorem 4: Let \( k_{t_1} = 0 \) and \( C_{t_1} = C_{11} \) for all \( t \). Then the martingale such that

\[
E \left[ (X_t - k_{t_j})^+ \right] = C_{t_j}, \text{ for all } (k_{t_j}, C_{t_j}) \in \mathcal{S}
\]

exists, if and only if

(a) \( C_{t_j} \geq C_{11} - k_{t_j} \) for each \( (k_{t_j}, C_{t_j}) \in \mathcal{S} \);

(b) for each \( t \), none of the points of \( \mathcal{S}_t \) is in the interior of the convex hull of \( \{(0, +\infty)\} \cup \{(+\infty, 0)\} \cup \{(k_{t_j}, C_{t_j}) \mid (k_{t_j}, C_{t_j}) \in \mathcal{S}, l \geq t\} \) (options of maturities greater than or equal to \( t \)).

The algorithm for computing tight upper and lower bounds is given in Section 2.3. In Section 2.4 we show how the solution of the problem of determining bounds on the variance of the underlying asset at time \( t \) leads to the solution of the problem of determining bounds on the price of an option, given the set of other priced options and the variance of the underlying asset at time \( t \) (Problem (c).) We state an exact upper bound on the variance in Proposition 5 and then develop an algorithm for computing the lower bound.
Chapter 3 starts with structural theorems for the problem of determining bounds on the price of an option on two different assets, given prices of options on individual assets (Problem (d).) These theorems allow to reduce the class of arbitrary distributions over which we have to maximize (minimize) to discrete distributions with a finite number of atoms, thus making the problem finite-dimensional.

We first show that it is enough to consider martingales with Markov property only (Subsection 3.1.1.) Then in Theorem 6 we state that constraints (1.7) for \( t > t^* \) have a “simple effect” on our problem, that is, they can be reduced to two separate sets of conditions on \( X_t^1 \) and \( X_t^2 \):

If \( X_1, X_2, \ldots, X_{t^*} \) is a two-dimensional martingale satisfying

\[
E \left[ (X_t^h - k_{ij}^h)^+ \right] = C_{ij}^h, \quad \{ (h, t, j) \in T \mid t \leq t^* \},
\]

\[
E \left[ (X_t^h - k_{ij}^h)^+ \right] \leq C_{ij}^h, \quad \{ (h, t, j) \in T \mid t > t^* \},
\]

then there exist two-dimensional random variables \( X_{t^*+1}, \ldots, X_n \) such that

(a) \( X_1, X_2, \ldots, X_{t^*}, X_{t^*+1}, \ldots, X_n \) is a martingale;
(b) all the conditions \( E \left[ (X_t^h - k_{ij}^h)^+ \right] = C_{ij}^h \) are satisfied for all \( (h, t, j) \in T \).

Conditions for \( t < t^* \) can not be reduced to two separate conditions on \( X_t^1 \) and \( X_t^2 \). We give a counterexample to prove it.

In Subsection 3.1.3 we show how to reduce the class of distributions to the class of atomic distributions with a finite number of atoms. The bound on the price of the stock over the considered time period is required, that is for \( h = 1, 2 \)

\[
P \left( X_1^h, X_2^h, \ldots, X_n^h \leq L \right) = 1.
\]

Let us assume that there are no options of maturity higher than \( t^* \) (we do not assume that in our solution, but make this assumption here just to introduce the idea of obtaining a discrete distribution.) We define graph \( G \) (Figure 1-1), which is contained in \( Q = \{ (x, y) \in \mathbb{R}^2 \mid 0 \leq x, y \leq L \} \), and has vertices on the intersections of the lines:

(a) \( x = k_{ij}^1 \), for \( t \leq t^*, j \leq U(t, 1) \),
(b) \( y = k_{ij}^2 \), for \( t \leq t^*, j \leq U(t, 2) \),
(c) \( \alpha x + \beta y = k_g \),
(d) segments of \( Q \): from \((0, 0)\) to \((0, L)\), \((0, L)\) to \((L, L)\), \((L, L)\) to \((L, 0)\), \((0, 0)\) to
Each edge of the graph belongs to one of the above lines (or segments). The intersection between two edges is either empty or consists of one vertex point.

In Theorem 7 we prove that it is enough to consider martingales with the state space on edges of Graph G only to satisfy all the constraints and to achieve optimal bounds.

In Theorem 8 we prove that it is enough to consider martingale distributions \( \mathcal{L}(X_1, \ldots, X_t) \), such that conditioned on the set of edges visited up to time \( t \), \( X_t \) has an atomic measure with at most one atom per edge for all \( t = 2, 3, \ldots, t^* \), while \( X_1 \) has an atomic measure with at most one atom per edge.

In Section 4.2 we first show, how the problem can be written as an optimization problem with a quadratic objective function and quadratic and linear constraints. Then we show how the problem can be made linear both in objective function and constraints.

Section 4.3 provides a computational example.

Since, as we have mentioned before, the number of variables in the linear programming problem, that leads to an optimal solution, grows exponentially with \( t^* \), we also propose an approximation approach in Section 4.3, in which the number of variables in the linear programming problem grows linearly with \( t^* \). The idea of the approach is based on Theorem 9:

Let \( X_1, X_2, \ldots, X_t \) be a two-dimensional Markov Martingale with values in \([0, L] \times [0, L]\). Then there exists a two-dimensional Markov Martingale \( \tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_{t^*} \) such
that for all $t = 1, \ldots, t^*$, $P \left( \tilde{X}_t \in L_{t}^{1} \right) = 1$ and

$$P \left( \left| \tilde{X}_t^h - X_t^h \right| \leq t \epsilon \right) = 1, \quad h = 1, 2.$$ 

Here $L_{t}^{1}$ is an $\epsilon$-square lattice in $[0, L + (t - 1)\epsilon] \times [0, L + (t - 1)\epsilon]$.

In Theorem 10 we prove that the solution of the problem of maximizing (minimizing) $E \left[ \tilde{g} \left( \tilde{X}_{t*} \right) \right]$ with the state space of the martingale restricted to $L_{t*}^{\epsilon}$ and constraints loosened to

$$ \left| E \left[ f_{ij}^h \left( \tilde{X}_t \right) \right] - C_{ij}^h \right| \leq \epsilon t, \quad \{(h, t, j) \in T\}, $$

converges to the optimal solution of our original problem as $\epsilon \to 0$.

Function $g : \mathbb{R}^{2+} \to \mathbb{R}$ can be any continuous function here. In Subsection 4.4.1 we formulate the problem as a linear programming problem.

In Subsection 3.4.2 we show, how the efficiency of the approximation algorithm can be improved if $g$ is a continuous piecewise linear function. The idea is to combine Theorems 10 and 7.

In Section 3.5 we apply the approach developed for the exact solution in a two dimensional case to the one dimensional case. It follows, that to maximize (minimize) $g(X_{t*})$, where $g : \mathbb{R}^+ \to \mathbb{R}$ is a continuous piecewise linear function, under the condition that $X_1, \ldots, X_{t*}, \ldots, X_n$ is a nonnegative one-dimensional martingale which satisfies the set of conditions (1.6), it is enough to consider martingale distributions which are atomic with atoms located at strike prices $k_{ij}$, points where $g$ changes its derivative and points $\{0\}$ and $\{L\}$. Here $L$ is the bound on the price of the stock, as in the two-dimensional case. This argument leads to an efficient algorithm that solves the problem.

In Section 3.6 we notice, that approaches developed for the two-dimensional case also work for higher dimensions. We estimate the complexity of algorithms in each case.

Finally in Section 3.7 we show how to find exact upper and lower bounds on the price of an option on the basket of securities, given prices of similar options of the same maturity (Problem (f).) To solve the problem we first have to find vertices of
the graph $G \subset [0, L]^n$ formed by intersecting hyperplanes

$$w_i^T X = k_i, \quad i = 1, 2, \ldots, m,$$
$$X^h = 0, \quad h = 1, 2, \ldots, n,$$
$$X^h = L, \quad h = 1, 2, \ldots, n.$$  \tag{1.8}

Vector $r \in [0, L]^n$ a vertex of $G$ if there are $n$ independent hyperplanes of (1.8) which intersect at $r$. Then we show that it is enough to optimize over atomic distributions with atoms located in the vertices of graph $G$.

### 1.4 On portfolio optimization

Portfolio optimization is the problem of finding the best allocation of wealth among a basket of securities. The mean-variance formulation by Markowitz [31] provides a fundamental basis for portfolio selection in a single period. In the mean-variance framework the optimal portfolio is derived by minimizing the variance of the portfolio return, subject to a given mean return. Tobin [47] adds unlimited borrowing and lending at the riskless rate to the mean-variance model and shows that every investor holds a combination of just two portfolios: the riskless asset and one particular efficient portfolio of risky assets. Sharpe [45], Lintner [29] and Mossin [37] consequently derive an equilibrium model for the entire financial market based on the mean-variance framework of Markowitz and the two-fund separation result of Tobin, known as the CAPM.

Research on multi-period portfolio choice problems starts in the financial economic literature of the late 60's with Mossin [38], Samuelson [44] and Hakansson [18]. They study a multi-period consumption-investment problem for an investor maximizing the expectation of a power utility function over wealth at the planning horizon, under the assumption that returns are independent across times. As a companion paper of Samuelson [44], Merton [34] formulates the consumption-investment problem in a continuous-time framework, where the timestep between consecutive trading dates decreases to zero in the limit. The asset prices are assumed to follow geometric Brownian motions. For the class of power utility functions Merton [34] concludes that myopic investment policies are optimal.

Brennan, Lagnado and Schwartz [7] numerically investigate the impact of re-
turn predictability (time-varying expected returns) on optimal portfolio choice in a continuous-time investment model based on Merton [34]. Brennan [8], Barberis [3] and Xia [49] study continuous-time investment models under parameter uncertainty. Kim and Omberg [26] derive exact analytical formula for optimal portfolio strategies when investors have power utility and expected returns are governed by a single mean-reverting state variable.

After Harrison and Kreps [20] and Harrison and Pliska [21] introduce a martingale method to price securities, the approach is applied by Karatzas [25], Cox and Huang [10] and Pliska [39] to provide a closed form solution for the optimal portfolio when the underlying security prices follow a general diffusion process. The basic idea is to use the completeness and the arbitrage-free property of the market to separate the computation of optimal consumption and that of the corresponding trading strategy.

Richardson [40] derives closed-form solutions for the mean-variance optimization of a portfolio consisting of a riskless bond and a single stock in continuous time. The stock is assumed to follow a geometric Brownian motion process. Duffie and Richardson [15] find closed-form solutions for futures hedging policies under mean-variance and quadratic objectives with the same assumptions for the processes.

Li and Ng [27] derive an analytical optimal portfolio policy to the mean-variance formulation in multiperiod selection in discrete time under the assumption that returns are independent over time.

Additional material and references can be found in a book by Markowitz [33] or in any standard text on mathematical finance, like Sharpe [46], Elton and Gruber [16], Alexander and Sharpe [1], Merton [36].

1.5 Contributions on portfolio optimization

We address a few questions in dynamic portfolio allocation, where we do not assume a specific price dynamics for the assets. While the distribution of price processes is required for the solution, the arguments that we formulate or methods that we propose do not depend on the form of the distribution. The contributions and structure are as follows.

The mathematical formulation of the problem is as follows. Stock prices \( \{P_s\}_{s \in [0,T]} \) evolve according to some stochastic process, the distribution of which is known. The wealth at time \( T, W_T \), depends on the initial capital \( W_0 \), the path \( \{P_s\}_{s \in [0,T]} \), and the
policy \( \delta_{[0 \rightarrow T]} \). The self-financing policy \( \delta_{[0 \rightarrow T]} \) can be represented as a vector process \( \{x_t\}_{t \in [0,T]} \) with \( i \)-th coordinate being the proportion of capital invested in asset \( i \). Thus \( W_T = W_T(\{P_s\}_{s \in [0,T]}, \delta_{[0 \rightarrow T]}|W_0) \). Given the distribution of \( \{P_s\}_{s \in [0,T]} \) and \( W_0 \), each policy \( \delta_{[0 \rightarrow T]} \) leads to a certain distribution of \( W_T \).

While in the academic literature the goal is usually to maximize either \( E[f(W_T)] \), the expected utility function of the terminal wealth, or \( E[W_T] - \gamma \text{Var}[W_T] \), the mean-variance function of the terminal wealth, we notice that while the problem of maximizing the expected value of a function of the terminal wealth can always be solved by a stochastic dynamic programming, this method cannot be applied to the problem of maximizing the mean-variance function of the terminal wealth. Specifically, the principle of optimality is violated, in the sense that the optimal policy on \([t, T] \) might not coincide with the optimal policy on \([0, T] \) when restricted to \([t, T] \).

We prove that for large time horizons, there exists a myopic policy which leads to a distribution of the terminal wealth with the property that the probability of underperforming any other policy tends to zero as the horizon tends to infinity.

Finally, we define an analogue to the mean-variance function in the multiperiod case, so that the principle of optimality is not violated. We suggest a monte-carlo based method for the solution which is polynomial in a number of securities.
Chapter 2

Option Pricing Without Price Dynamics. One Dimensional Case.

This chapter is devoted to understanding the relationship between prices of options of different maturities and the same underlying asset. No assumption on the price dynamics of the underlying asset is made.

The price of a European call option with a strike price $k$ and maturity $t$ is given by

$$C = e^{-\int_0^t r(s)ds} E_Q \left[(S(t) - k)^+\right] = E_Q \left[\left(X_t - ke^{-\int_0^t r(s)ds}\right)^+\right],$$

where $r(s)$ is the instantaneous riskless rate of return, $\{S(t)\}_{t \geq 0}$ is the price process of the underlying security defined on some probability space $(\Omega, \mathcal{B}, P)$, and $Q$ is a measure equivalent to $P$, such that $\{X_t\}_{t \geq 0} := \left\{S(t)e^{-\int_0^t r(s)ds}\right\}_{t \geq 0}$ is a martingale process under $Q$.

Our first problem is to determine exact bounds on the price of a European call option given the set of priced options on the same underlying security. We allow options in the set to have different maturities.

The problem can be formulated as follows: given $k^*, t^* \in \mathbb{R}^+$, and a set

$$S := \{(k_{ij}, C_{ij}, t_i) \mid k_{ij}, C_{ij}, t_i \in \mathbb{R}^+, t_1 < t_2 \ldots < t_n, i \in \{1, 2, \ldots, n\}, j \in \{1, 2, \ldots, U(i)\}\}$$
find the domain of values of

\[ E \left[ \left( X_{t^*} - k^*e^{-\int_0^{t^*} r(s)ds} \right)^+ \right], \]

under the condition, that \( \{X_{t_1}, X_{t_2}, \ldots, X_{t^*}, \ldots, X_{t_n}\} \) is a nonnegative martingale such that

\[ E \left[ \left( X_{t_i} - k_{ij}e^{-\int_0^{t_i} r(s)ds} \right)^+ \right] = C_{ij}. \]

**Remark 1.** The rate \( r(s) \) is not essential in this problem. Indeed, we can take \( \tilde{k}_{ij} = k_{ij}e^{-\int_0^{t_i} r(s)ds} \) instead of \( k_{ij} \), and rewrite all the equalities as \( E \left[ \left( X_{t_i} - \tilde{k}_{ij} \right)^+ \right] = C_{ij} \). Here \( \tilde{k}_{ij} \) are known, once \( k_{ij} \) are known, since \( r(s) \) is assumed to be a given function of time. So without loss of generality we can assume \( r(s) \equiv 0 \).

**Remark 2.** Since \( r(s) \equiv 0 \), only the order of \( t_1, t_2, \ldots, t_n \) is important, but not the specific values that the sequence takes. Thus we can assume \( \{t_1, t_2, \ldots, t_n\} \equiv \{1, 2, \ldots, n\} \) for convenience.

**Remark 3.** The price of an option with a strike price 0 is given by \( E[X_t] = X_0 \), which is the current stock price. Thus, it is always known.

In view of Remarks 1 and 2, a simpler formulation of the problem can be used.

---

Given an ordered set

\[ S := \{(k_{tj}, C_{tj}) \mid (k_{tj}, C_{tj}) \in \mathbb{R}^{2+}, \ t \in \{1, 2, \ldots, n\}, \ j \in \{1, 2, \ldots, U(t)\}\} \quad (2.1) \]

and \( k^* \in \mathbb{R}^+ \) and \( t^* \in \{1, 2, \ldots, n\} \), find the domain of values of

\[ E \left[ (X_{t^*} - k^*)^+ \right], \quad (2.2) \]

under the condition, that \( \{X_1, \ldots, X_{t^*}, \ldots, X_n\} \) is a nonnegative martingale such that

\[ E \left[ (X_t - k_{tj})^+ \right] = C_{tj} \text{ for all } (k_{tj}, C_{tj}) \in S. \quad (2.3) \]
Remark 4. It is important that $\mathcal{S}$ be an ordered set: given $t \in \{1, \ldots, n\}$ and $j \in \{1, \ldots, U(t)\}$ one should be able to extract $(k_{ij}, C_{ij})$.

Another problem that we consider in this chapter is determining bounds on $E \left[ (X_t^* - k^*)^+ \right]$ if an addition to the set of options $\mathcal{S}$, the variance of $X_t$ at a certain time $t \in \{1, 2, \ldots, n\}$ is known. And, at last, we address the problem of finding lower and upper bounds on variance of $X_t$, based on the set of options $\mathcal{S}$.

Prior to solving the above problems, we have to find a necessary and sufficient condition that the set of options $\mathcal{S}$ should jointly satisfy so that there exists no-arbitrage, or, equivalently, that there exists a martingale $\{X_1, \ldots, X_n\}$ such that Eq. (2.3) holds. This condition is established in Section 2.1. In Section 2.2 we derive a closed-form upper and lower bounds on the price of an option and propose an efficient algorithm for computing them. In Section 2.3 we show how the problem of determining bounds on the price of an option, based on the set of options $\mathcal{S}$ and the variance of the stock at time $t$ can be converted to the problem of finding bounds on the variance, based on the given set of options. In Section 2.4 we derive an exact upper bound on the variance of the stock price at a given time and propose a polynomial time algorithm for determining a tight lower bound.

### 2.1 Structural theorems

**Definition 1.** (a) We will say that the martingale $\{X_1, X_2, \ldots, X_n\}$ satisfies the set of points $\mathcal{S}$ if Eq. (2.3) holds.

(b) We will say that the set $\mathcal{S}$ satisfies the no-arbitrage condition, if there exists a martingale $\{X_1, X_2, \ldots, X_n\}$, such that Eq. (2.3) holds.

The martingale $\{X_1, X_2, \ldots, X_n\}$, that we consider, is nonnegative, since the domain of its values is the same as the domain of values that the stock price can take.

Let $\mathcal{C}$ designate the class of probability laws with support in $\mathbb{R}^+$ and finite expectations. We are going to associate with each element $\Pi$ of $\mathcal{C}$ a function (transform) $\Psi_\Pi : \mathbb{R}^+ \to \mathbb{R}^+$, defined as

$$
\Psi_\Pi(t) = E_{\Pi} \left[ (x - t)^+ \right] = \int_0^\infty (x - t)^+ d\Pi(x). \tag{2.4}
$$

The inverse transform is also well-defined, that is $\Psi_\Pi(t)$ uniquely determines $\Pi$. 
Theorem 1. (a) The right derivative $\Psi'_\Pi(t^+)$ of $\Psi_\Pi(t)$ exists for all $t \geq 0$ and

$$\Psi'_\Pi(t^+) = -\Pi ( (t, +\infty) ).$$

(b) Let $\Pi_1, \Pi_2 \in \mathcal{C}$ and for each $t \in \mathbb{R}^+$, $\Psi_{\Pi_1}(t) = \Psi_{\Pi_2}(t)$. Then $\Pi_1 = \Pi_2$.

Proof. Let us take any $t \in [0, +\infty)$ and $h > 0$. We have

$$\frac{\Psi_\Pi(t+h) - \Psi_\Pi(t)}{h} = \int_0^\infty \frac{(x - (t+h))^+ - (x-t)^+}{h} \, d\Pi(x) = \int_0^\infty f_{t,h}(x) \, d\Pi(x),$$

where

$$f_{t,h}(x) := \frac{(x - (t+h))^+ - (x-t)^+}{h} = \begin{cases} 0, & x \leq t, \\ \frac{x-t}{h}, & t \leq x \leq t+h, \\ -1, & x \geq t+h. \end{cases}$$

See also Figure 2-1.

![Figure 2-1: The function $f_{t,h}(x)$.](image)

Since the sequence of functions $f_{t,h}(x)$ is monotone (in $h$), and $f_{t,h}(x)$ converges as $h \to 0^+$, then by the monotone convergence theorem [13, p.100], $\int_0^\infty f_{t,h}(x) \, d\Pi(x) \to -\Pi((t, +\infty))$ as $h \to 0^+$. Thus $\Psi'_\Pi(t^+)$ exists and equals $-\Pi((t, +\infty))$. \qed

The following proposition summarizes the properties of $\Psi-$transform. The theorem after the proposition asserts that these properties are also sufficient for the function to be a $\Psi-$transform of some distribution $\Pi$.

Proposition 1. Let $\Pi \in \mathcal{C}$. Then $\Psi_\Pi(t)$ satisfies

(a) $\Psi_\Pi(t)$ is a nonincreasing function of $t \in \mathbb{R}^+$;

(b) $\Psi_\Pi(t)$ is convex;

(c) $\lim_{t \to +\infty} \Psi_\Pi(t) = 0$. In particular, if $\Pi$ has a bounded support, then $\Psi_\Pi(t) = 0$.
for $t$ large;

(d) $\Psi_{\Pi}(t) \geq \Psi_{\Pi}(0) - t$.

**Proof.** (a) Notice, that $\Psi_{\Pi}(t)$ is a continuous function:

$$|\Psi_{\Pi}(t + h) - \Psi_{\Pi}(t)| = \left| \int_{t+h}^{t'} (x - t)d\Pi(x) \right| \leq |h|.$$

Since $\Psi_{\Pi}(t)$ is continuous and $\Psi_{\Pi}'(t+)= -\Pi\left( (t, +\infty) \right) \leq 0$, then $\Psi_{\Pi}(t)$ is nonincreasing.

(b) $f_x(t) := (x-t)^+$ is convex in $t$ for each $x$. Thus $\Psi_{\Pi}(t)$ is convex as a convex combination of convex functions.

(c)

$$\Psi_{\Pi}(t) = \int_{0}^{+\infty} (x-t)^+d\Pi(x) = \int_{t}^{+\infty} (x-t)^+d\Pi(x) \leq \int_{t}^{+\infty} xd\Pi(x).$$

Since $\int_{0}^{+\infty} xd\Pi(x) < \infty$, then $\int_{t}^{+\infty} xd\Pi(x) \to 0$ as $t \to \infty$. Thus, $\Psi_{\Pi}(t) \to 0$ as $t \to +\infty$.

(d) $\Psi_{\Pi}(t) = \int_{0}^{\infty} (x-t)^+d\Pi(x) \geq \int_{0}^{\infty} (x-t)d\Pi(x) = \Psi_{\Pi}(0) - t$. \qed

**Theorem 2.** Let $g : \mathbb{R}^+ \to \mathbb{R}^+$ be a function with the following properties

(a) $g(x)$ is nonincreasing,

(b) $g(x)$ is convex;

(c) $\lim_{x \to +\infty} g(x) = 0$;

(d) $g(x) \geq g(0) - x$.

Then there exists $\Pi \in C$ such that $g(x) = \Psi_{\Pi}(x)$.

**Proof.** From the convexity of $g$ it follows that the right derivative of $g$, $g'(x+)$, is a right continuous nondecreasing function. Since $g(x) \geq g(0) - x$ and $g(x)$ is nonincreasing, then $-1 \leq g'(x+) \leq 0$. Also we have $g'(x+) \to 0$ as $x \to +\infty$. To justify the last statement, notice that $g'(x+)$ is a monotone bounded function, and consequently, it must converge to some finite value $a$ as $x \to \infty$. However, $a \neq 0$ would contradict the requirement in (c).

Consider the function $F(x) := 1 + g'(x+)$ defined on $\mathbb{R}^+$. From the above statements, it follows that

(i) $0 \leq F(x) \leq 1$;

(ii) $F(x)$ is a right-continuous and nondecreasing;
(iii) \( F(x) \to 1 \) as \( x \to \infty \).

Thus, \( F(x) \) satisfies all the properties of a distribution function. Denote by \( \Pi \) the law with the distribution function \( F(x) \). Then

\[
\Psi_\Pi(0) = E_\Pi[X] = \int_0^{+\infty} (1-F(x))dx = \int_0^{+\infty} (-g'(x+))dx \geq g(0) - \lim_{x \to +\infty} g(x) = g(0).
\]

The second to last equality is justified in this case since \( g(x) \) is a convex function and thus its right and left derivatives might differ only on a countable set of points. Indeed, since \( g'(x+) \) is a nondecreasing function, for it to have an uncountable number of jumps, it must be possible to split \( \mathbb{R} \) into an uncountable number of intervals of nonzero length. The last is not possible, however, since the set of rational numbers is countable and dense in \( \mathbb{R} \).

Since \( \Psi_\Pi'(x+) = -(1-F(x)) = g'(x+) \), \( \Psi_\Pi(x) \) and \( g(x) \) are both continuous and \( \Psi_\Pi(0) = g(0) \), it follows that \( \Psi_\Pi(x) = g(x) \) for any \( x \in \mathbb{R}^+ \). \( \square \)

Since there is a one to one correspondence between laws of random variables and their \( \Psi \) - transforms, it follows that conditions that a sequence of laws of random variables should satisfy can always be reformulated in terms of conditions on the sequence of \( \Psi \) - transforms of these random variables. Theorem 3 establishes sufficient and necessary conditions that the sequence of \( \Psi \) - transforms should satisfy so that the corresponding random variables could form a martingale. However prior to proving that theorem, we will need the following auxiliary results.

**Definition 2.** Let \( \Pi_1, \Pi_2 \in \mathcal{C} \). Define the relation \( \Pi_1 \mathcal{R} \Pi_2 \) if and only if there exist random variables \( X_1 \) and \( X_2 \), defined at the same probability space, such that \( \{X_1, X_2\} \) is a martingale and \( X_1 \) has a marginal law \( \Pi_1 \), while \( X_2 \) has a marginal law \( \Pi_2 \).

**Proposition 2.** (a) If \( \Pi_1 \mathcal{R} \Pi_2 \) and \( \Pi_2 \mathcal{R} \Pi_1 \) then \( \Pi_1 = \Pi_2 \);
(b) \( \Pi_1 \mathcal{R} \Pi_1 \);
(c) If \( \Pi_1 \mathcal{R} \Pi_2 \) and \( \Pi_2 \mathcal{R} \Pi_3 \), then \( \Pi_1 \mathcal{R} \Pi_3 \).

**Proof.** (a) If \( \{X_1, B_1\}, \{X_2, B_2\} \) is a martingale with marginal laws \( \Pi_1 \) and \( \Pi_2 \) correspondingly, and \( f \) is a convex function, defined on \( \mathbb{R} \), then \( f(X_1) = f(E[X_2|B_1]) \leq E[f(X_2)|B_1] \) by conditional Jensen's inequality, and thus \( E[f(X_1)] \leq E[f(X_2)] \). \( (X-t)^+ \) is a convex function of \( X \), so \( \Psi_{\Pi_1}(t) \leq \Psi_{\Pi_2}(t) \) for each \( t \in \mathbb{R}^+ \). From \( \Pi_2 \mathcal{R} \Pi_1 \), by the same argument we obtain \( \Psi_{\Pi_2}(t) \leq \Psi_{\Pi_1}(t) \) for each \( t \in \mathbb{R}^+ \). Thus
\( \Psi_{\Pi_1}(t) = \Psi_{\Pi_2}(t) \) and by Theorem 1, \( \Pi_1 = \Pi_2 \).

(b) \( \{X_1, X_1\} \) is always a martingale.

(c) Let \( \{X_1, X_2\} \) be a martingale with the law \( \mathcal{L}(X_1, X_2) \) defined on \( \mathbb{R} \times \mathbb{R} \) with support in \( \mathbb{R}^+ \times \mathbb{R}^+ \) and marginal laws \( \Pi_1 \) and \( \Pi_2 \) respectively. Let \( \{\bar{X}_2, X_3\} \) be a martingale with the law \( \mathcal{L}(\bar{X}_2, X_3) \) defined on \( \mathbb{R} \times \mathbb{R} \) with support in \( \mathbb{R}^+ \times \mathbb{R}^+ \) and marginal laws \( \Pi_2 \) and \( \Pi_3 \). Then on \( \mathbb{R} \) there exist conditional distributions \( \mathcal{L}(X_1|X_2) \) for \( \mathcal{L}(X_1, X_2) \) and \( \mathcal{L}(X_3|\bar{X}_2) \) for \( \mathcal{L}(\bar{X}_2, X_3) \). Since the marginal laws of \( X_2 \) and \( \bar{X}_2 \) are equal, then by Vorob'ev-Berkes-Philipp theorem [14, p.7], we can define a law \( \mathcal{L}(X_1, X_2, X_3) \) on \( \mathbb{R} \times \mathbb{R} \times \mathbb{R} \), such that \( X_1 \) and \( X_3 \) are conditionally independent given \( X_2 \), that is

\[
\mathcal{L}((X_1, X_3)|X_2) = \mathcal{L}(X_1|X_2) \times \mathcal{L}(X_3|\bar{X}_2).
\]

Denote by \( B_i, \ i = 1, 2, 3 \), the smallest \( \sigma \)-algebra for which all \( X_k, \ k \leq i \), are measurable. Then \( \{X_i, B_i\}_{i=1}^3 \) is a martingale with marginal laws \( \Pi_1, \Pi_2 \) and \( \Pi_3 \) correspondingly. Note that the martingale thus constructed is also a Markov process. \( \Box \)

**Theorem 3.** Let \( \Pi_1, \Pi_2, \ldots, \Pi_n \in \mathcal{C} \) and \( E_{\Pi_i}[X] = E_{\Pi_i}[X] \) for each \( i = 1, \ldots, n \). Then, there exists a martingale \( \{X_1, X_2, \ldots, X_n\} \), such that \( X_i \) has marginal law \( \Pi_i \) for each \( i \) if and only if the sequence of functions \( \Psi_{\Pi_1}, \Psi_{\Pi_2}, \ldots, \Psi_{\Pi_n} \) is nondecreasing.

**Remark.** After the paper was written we discovered that this result can be obtained from Kertz and Rosler [24], where the key step in the proof is the theorem proved in Strassen [43]. Since our approach is completely different, we include our original proof.

**Proof.** The only if part is a direct consequence of the fact that \( (X - t)^+ \) is a convex function of \( X \) and Jensen's inequality for convex functions of martingales [13, p.277]. The proof of this statement is also in the proof of part (a) of Proposition 2.

In the other direction, we have to show that it is possible to define a joint distribution of \( X_1, X_2, \ldots, X_n \) with the given marginal laws (since they are uniquely determined by \( \Psi_{\Pi_1}, \Psi_{\Pi_2}, \ldots, \Psi_{\Pi_n} \)), so that this sequence is a martingale. It is enough to show how to define a joint distribution of \( X_1 \) and \( X_2 \), since afterwards we can proceed recursively to form a martingale which is a Markov process as in the proof of Proposition 2.

First we are going to define a sequence \( X_1, Y_1, Y_2, \ldots, Y_n, \ldots \) so that this sequence is a martingale and \( \Psi_{Y_n} \to \Psi_{\Pi_2} \). Then we will show that there exists \( Y_\infty \), such that
\( \{Y_n\}_{1 \leq n \leq \infty} \) is a martingale and \( Y_\infty \) has a distribution \( \Pi_2 \). Then \( X_2 \) can be defined as \( X_2 = Y_\infty \).

Let \( Y_0 = X_1 \) and let us form a martingale \( Y_0, Y_1, Y_2, \ldots, Y_n \) such that it is also a Markov process. Then to specify joint distributions we only need to define the joint distribution of \( Y_i \) and \( Y_{i+1} \) for each \( i \). Now, the joint distribution of \( Y_i \) and \( Y_{i+1} \) is uniquely determined as soon as we know for each \( y \in \mathbb{R}^+ \) the conditional distribution of \( Y_{i+1} \) given \( Y_i = y \). Thus we are going to define the conditional distribution for \( Y_1 \), then for \( Y_2 \) and so on. We know that \( \Psi_{X_1} = \Psi_{\Pi_1} \leq \Psi_{\Pi_2} \). Now \( Y_1 \) will have

\[
\begin{align*}
\Psi \\
\Psi_{X_2} \\
\Psi_{X_1} \\
a & b & t
\end{align*}
\]

Figure 2-2: Construction of \( \Psi_{Y_1} \).

its \( \Psi \)-transform, \( \Psi_{Y_1} \), equal to \( \Psi_{X_1} \) everywhere except on \((a, b)\) where it will be a straight line (Figure 2-2). We choose \((a, b)\) so that the line which goes through the points \((a, \Psi_{X_1}(a))\), \((b, \Psi_{X_1}(b))\) is below the function \( \Psi_{\Pi_2}(t) \) (the specific of \((a, b)\) is discussed later.) What distribution will have its \( \Psi \)-transform equal to the transform of \( \Pi_1 \) everywhere except \((a, b)\) and being a straight line on \((a, b)\)? From Theorem 1 (a), we see that \( \Psi_{Y_1} \) should have no weight on \((a, b)\) but the weight of \([a, b]\) should be the same as for the distribution \( \Pi_1 \). Thus we are going to take all the weight of \((a, b)\) and redistribute it between \( \{a\} \) and \( \{b\} \). We have to do it in such a way that \( \{X_1, Y_1\} \) is a martingale.

The law \( \mathcal{L}(X_1) \) is uniquely determined by \( \Psi_{X_1} \) and the law \( \mathcal{L}(Y_1) \) is uniquely determined by \( \Psi_{Y_1} \). We define the law \( P := \mathcal{L}(X_1, Y_1) \) with support in \( \mathbb{R}^+ \times \mathbb{R}^+ \) in the following way. As the marginal law of \( X_1 \) take \( \Pi_{X_1} \). Then for each \( X_1 \in \mathbb{R}^+ \) define the conditional distribution of \( Y_1 \) given \( X_1 \). First, if \( X_1 \in \{\mathbb{R}^+ \setminus \{a, b\}\} \), take \( Y_1 = X_1 \). Now, if \( X_1 \in (a, b) \) we should have \( P(Y_1 \in \{a\} \cup \{b\} | X_1 \in (a, b)) = 1 \) for \( \Psi_{Y_1}(t) \) to be a straight line on \((a, b)\) and equal to \( \Psi_{X_1}(t) \) everywhere else. Let us denote \( \pi_{X_1}(X_1, Y_1) := X_1 \).

Thus for any \( c \in (a, b) \) we want the joint distribution of \( X_1, Y_1 \) to be such that
the following holds

\[ aP(Y_1 = a|X_1 = c) + bP(Y_1 = b|X_1 = c) = c, \]

\[ P(Y_1 = a|X_1 = c) + P(Y_1 = b|X_1 = c) = 1. \]

Consequently,

\[ P(Y_1 = a|X_1 = c) = \frac{b - c}{b - a}, \quad P(Y_1 = b|X_1 = c) = \frac{c - a}{b - a}. \]

Thus the conditional distribution of \( Y_1 \) given \( X_1 \) is defined and \( \{X_1, Y_1\} \) is a martingale. In the same manner we can define the conditional distribution of \( Y_{i+1} \) given \( Y_i \), for \( i = 1, 2, \ldots \).

The next step is to show that \( \Psi_{X_2}(t) \) can be approached by \( \Psi_{Y_n}(t) \) with any given precision, that is that we can choose \( Y_1, Y_2, \ldots, Y_n \) so that \( \Psi_{Y_n}(t) \to \Psi_{X_2}(t) \). For each \( i, \Psi_{Y_{i+1}}(t) \) is different from \( \Psi_{Y_i}(t) \) only on a certain interval \((a_i, b_i)\), where \( \Psi_{Y_{i+1}}(t) \) is a straight line. Thus the choice of the sequence \( \{Y_n\} \) is equivalent to the choice of the sequence \( \{(a_n, b_n)\} \).

![Figure 2-3: The choice of \((a_1, b_1)\).](image)

Let

\[ U := \{t \in \mathbb{R}^+ : \Psi_{X_2}(t) > \Psi_{X_1}(t)\}. \]

Let \( \{t_n\} \) be a countable dense set in \( U \). At each \( t_i, i = 1, 2, \ldots \), draw a tangent line \( y_i(t) \) to \( \Psi_{X_2}(t) \). (See Figure 2-3.) Let us first assume that \( \Psi_{X_2}(t_i) \neq 0 \). Then \( \Psi_{X_2}'(t_i^+) < 0 \) and the tangent line intersects \( \Psi_{Y_{i-1}}(t) \) at two different points, \( a_i \in \mathbb{R}^+ \) and \( b_i \in \mathbb{R}^+ \), since \( \Psi_{Y_i}(t) \to 0 \) as \( t \to 0 \) and \( \Psi_{Y_i}(0) = 1 \), while \( y_i(t) \) is below \( \Psi_{X_2}(t) \) and thus \( y_i(0) \leq 1 \). Thus the choice of \( \{t_n\} \) is equivalent to the choice of \( \{(a_n, b_n)\} \). We define \( \Psi_{Y_{i+1}} = \max \{y_i(t), \Psi_{Y_i}(t)\} \). In case \( \Psi_{X_2}(t_i) = 0, \Psi_{X_2}'(t_i) = 0 \) and we take
\[ \Psi_{Y_{t+1}} := \Psi_{Y_t}. \]

Thus the sequence \( \{\Psi_{Y_n}\} \) is an increasing sequence of continuous functions converging to a continuous function \( \Psi_{X_2} \) on a countable dense set. Consequently, \( \Psi_{Y_n} \to \Psi_{X_2} \) pointwise on \( \mathbb{R}^+ \).

Since \( \Psi_{X_1}(t) \to 0 \) and \( \Psi_{X_2} \to 0 \) as \( t \to +\infty \), then for any \( \epsilon > 0 \) there exist \( T \in \mathbb{R}^+ \) such that for any \( t > T \), \( |\Psi_{X_1}(t) - \Psi_{X_1}(t)| < \epsilon \). Then for any \( n \) and any \( t > T \), \( |\Psi_{X_2}(t) - \Psi_{Y_n}(t)| < \epsilon \). Taking into account that \( \{\Psi_{Y_n}\} \) converges uniformly to \( \Psi_{X_2} \) on \([0, T]\), we conclude that \( \Psi_{Y_n} \) converges uniformly to \( \Psi_{X_2} \) on \( \mathbb{R}^+ \).

Now, since \( \Psi_{Y_n} \to \Psi_{X_2} \) uniformly on \( \mathbb{R}^+ \) and the right derivatives of \( \Psi_{Y_n}(t) \), \( n = 1, 2, \ldots \), and \( \Psi_{X_2} \) are monotone functions, then also \( \Psi_{Y_n}'(t+) \to \Psi_{X_2}'(t+) \). Thus, the distribution functions of \( Y_n \), \( n = 1, 2, \ldots \), converge pointwise to the distribution function of \( X_2 \).

Now we are going to prove that \( \{Y_n\} \) is uniformly integrable. For that we need to show that for any \( \epsilon > 0 \) there exists \( t_0 > 0 \), such that for all \( t > t_0 \),

\[
\sup_n E \left[ Y_n 1_{\{Y_n > t\}} \right] < \epsilon.
\]

For each \( t \) and \( n \) we have

\[
\Psi_{Y_n}(t) \leq \Psi_{X_2}(t)
\]

Thus

\[
\int_0^{+\infty} (Y_n - t)^+ d\Pi_{Y_n} \leq \int_0^{+\infty} (X_2 - t)^+ d\Pi_{X_2},
\]

\[
\int_t^{+\infty} Y_n d\Pi_{Y_n} \leq \int_t^{+\infty} X_2 d\Pi_{X_2} + t |\Pi_{Y_n}([t, +\infty]) - \Pi_{X_2}([t, +\infty])|.
\]

Since \( \int_t^{+\infty} X_2 d\Pi_{X_2} \to 0 \) as \( t \to +\infty \) then we can take \( t_0 \) such that \( \int_{t_0}^{+\infty} X_2 d\Pi_{X_2} < \epsilon/3 \) for all \( t > t_0 \). Now since \( \Pi_{Y_n}([t_0, +\infty]) \to \Pi_{X_2}([t_0, +\infty]) \), there exists \( n_0 \) such that for all \( n > n_0 \), \( |\Pi_{Y_n}([t_0, +\infty]) - \Pi_{X_2}([t_0, +\infty])| < \frac{\epsilon}{3t_0} \). Thus for any \( n > n_0 \) and \( t > t_0 \) we have

\[
E \left[ Y_n 1_{\{Y_n > t\}} \right] \leq E \left[ Y_n 1_{\{Y_n > t_0\}} \right] < \frac{\epsilon}{3} + t_0 \frac{\epsilon}{3t_0} = \frac{2\epsilon}{3}.
\]

Thus

\[
\sup_{n > n_0} E \left[ Y_n 1_{\{Y_n > t\}} \right] < \frac{2\epsilon}{3}.
\]

For \( Y_1, Y_2, \ldots, Y_{n_0} \) we can always choose \( t_1 \) big enough so that for all \( t > t_1 \), \( E \left[ Y_i 1_{\{Y_i > t\}} \right] < \frac{\epsilon}{3} \), \( i = 1, 2, \ldots, n_0 \), since all \( Y_i \) have finite expectations. Consequently, if we denote
$$t^* := \max\{t_0, t_1\}, \text{ then}$$

$$\sup_n E \left[ Y_n 1_{\{Y_n > t\}} \right] < \epsilon,$$

for all $$t > t^*$$.

Thus, $$Y_1, Y_2, \ldots, Y_n$$ is uniformly integrable, and, consequently [13, p.283], right closable. The last means, that there exists a random variable $$Y_\infty$$ such that $$E[Y_\infty | \mathcal{B}_n] = Y_n$$ for all $$n$$, where $$\mathcal{B}_n$$ designates the smallest $$\sigma$$-algebra for which $$Y_1, \ldots, Y_n$$ are all measurable. Then by Doob's theorem [13, p.285] $$Y_n$$ converges to $$Y_\infty$$ a.s., and consequently the distribution functions of $$Y_1, Y_2, \ldots, Y_n$$ also converge to the distribution function of $$Y_\infty$$. But then $$Y_\infty$$ and $$X_2$$ have the same distribution function, and thus the same $$\Psi$$-transform. So we take $$X_2 = Y_\infty$$ and $$X_1, X_2$$ is a martingale.

The proof is complete.

Now we are ready to formulate the conditions that the set

$$\mathcal{S} := \{(k_{ij}, C_{ij}) \mid \{k_{ij}, C_{ij}\} \in \mathbb{R}^+, i \in \{1, 2, \ldots, n\}, j \in \{1, 2, \ldots, U(t)\}\}$$  \hspace{1cm} (2.16)

should satisfy, so that there exists a martingale $$\{X_1, X_2, \ldots, X_n\}$$, such that (2.3) holds.

**Proposition 3.** The set $$\mathcal{S}$$, as defined in (2.16), satisfies the condition of martingale existence if and only if there exists a nondecreasing sequence of convex functions, $$\{g_t(x)\}_{t \in \{1, \ldots, n\}} : \mathbb{R}^+ \to \mathbb{R}^+$$, such that for each $$t \in \{1, 2, \ldots, n\}$$

(a) $$g_t(0) = g_1(0);$$

(b) $$g_t(x) \geq g_t(0) - x;$$

(c) $$g_t(x) \to 0 \text{ as } x \to \infty;$$

(d) $$g_t(k_{tj}) = C_{tj} \text{ for each } j \in \{1, 2, \ldots, U(t)\}.$$

**Proof.** "$$\Rightarrow$$" By Theorem 2 for each $$g_t$$ there exists $$\Pi_t \in \mathcal{C}$$ such that $$\Psi_{\Pi_t} = g_t$$. Then, since the sequence of functions $$\Psi_{\Pi_1}, \Psi_{\Pi_2}, \ldots, \Psi_{\Pi_n}$$ is nondecreasing and $$\Psi_{\Pi_t}(0) = \Psi_{\Pi_1}(0)$$ for each $$t$$, from Theorem 3 it follows that the martingale satisfying all the points of the set $$\mathcal{S}$$ exists.

"$$\Leftarrow$$" Suppose that there exists a martingale $$\{X_1, X_2, \ldots, X_n\}$$, such that (2.3) holds. Then we can take $$g_t(x) := \Psi_{X_t}(x)$$ for $$t = 1, 2, \ldots, n$$. By Theorem 3 the sequence of functions $$\{g_t\}_{t = 1, \ldots, n}$$ is nondecreasing and by Proposition 1 conditions (a)-(c) hold. By definition, $$g_t(k_{tj}) = \Psi_{X_t}(k_{tj}) = C_{tj}$$ for all $$(k_{tj}, C_{tj}) \in \mathcal{S}$$.
Definition 3. (a) $I_t$ denote the subset of $S$ corresponding to options with maturity $t$ and higher, that is

$$I_t := \{(k_{ij}, C_{ij}) \mid (k_{ij}, C_{ij}) \in S, i \geq t\}. \quad (2.5)$$

(b) 

$$I_t^\infty := \{(0, \infty)\} \cup \{(\infty, 0)\} \cup I_t. \quad (2.6)$$

(c) Denote by $A_t \in \mathbb{R}^{2+}$ the convex hull of $I_t^\infty$.

(d) Denote by $\partial A_t$ the lower border of $A_t$.

Theorem 4. Let $k_{t1} = 0$ and $C_{t1} = C_{11}$ for each $(k_{t1}, C_{t1}) \in S$. Then there exists a martingale satisfying the set $S$ if and only if

(a) $C \geq C_{11} - k$ for any $(k, C) \in A_1$.

(b) for each $t$, none of the points $(k_{ij}, C_{ij}), j = 1, 2, \ldots, U(t)$, is in the interior of $A_t$.

(c) if $\{(k^1, C^1), (k^2, C^2)\} \in \partial A_t \cap S$ and $k^1 < k^2$, then $C^2 < C^1$.

**Proof.** "$\Rightarrow$" Let the function $\tilde{g}_t : \mathbb{R}^+ \to \mathbb{R}^+$, coincide with $\partial A_t$, for each $t = 1, \ldots, n$. Then the sequence of functions $\{\tilde{g}_t\}$ satisfies all the properties of Proposition 3 except of property (c). Each function $\tilde{g}_t$ is piecewise linear with slopes of all, but the last, line segments being negative. The last line segment has a slope zero. Let $x_t := \min\{x \mid x \in \mathbb{R}^+, \tilde{g}_t(x) = 0\}$. Denote by $s$ the set of all slopes of functions $\tilde{g}_t$, $t = 1, 2, \ldots, n$, and let $s^* := \max\{s \mid s \in s, s < 0\}$. Define the function

$$g_t(x) := \begin{cases} 
\tilde{g}_t(x), & x \leq x_t, \\
\tilde{g}_t(x_t) + s(x - x_t), & \{x \geq x_t \mid g_t(x_t) + s(x - x_t) \geq 0\}, \\
0, & \text{otherwise}.
\end{cases}$$

Then the sequence of function $g_t(x)$ satisfies all the properties of Proposition 3 and, thus, the martingale exists.

"$\Leftarrow$" If $X_1, X_2, \ldots, X_n$ is a martingale, satisfying all the points of $S$, then for any $(k, C) \in A_1, C \geq \Psi_{X_1}(k) \geq C_{11} - k$. Assume that there exists a point $(k_{t^*}, C_{t^*})$,
\( t^* \in \{1, \ldots, n\}, j^* \in 1, \ldots, U(t^*) \), which is in the interior of \( \mathcal{A}_{t^*} \) (Figure 2-4.) Take points \((k^l, C^l) \in I^\infty_{t^*}, (k^r, C^r) \in I^\infty_{t^*}\) such that they are on the border of \( \mathcal{A}_{t^*} \) and \( k^l \) is the closest to \( k_{t^*, j^*} \) from the left, while \( k^r \) is the closest to \( k_{t^*, j^*} \) from the right. Then \((k_{t^*, j^*}, C_{t^*, j^*})\) lies strictly above the interval connecting \((k^l, C^l)\) and \((k^r, C^r)\). Since \( \Psi_{X_{t^*}} \leq \Psi_{X_{t^*, t^*}} \leq \ldots \leq \Psi_{X_{n}}, \) then \( \Psi_{X_{t^*}}(k^l) \leq C^l \) and \( \Psi_{X_{t^*}}(k^r) \leq C^r \), thus \((k_{t^*, j^*}, C_{t^*, j^*})\) is also strictly above the interval connecting \((k^l, \Psi_{X_{t^*}}(k^l))\) and \((k^r, \Psi_{X_{t^*}}(k^r))\), which contradicts the condition that \( \Psi_{X_{t^*}} \) is a convex function.

\[ \square \]

Figure 2-4: For the proof of Theorem 4.

### 2.2 Upper and lower bounds on option prices

We return to the problem of determining the domain of possible values for \( C_{t^*} = E \left[ (X_{t^*} - k_{t^*})^+ \right] \) for a given \( k_{t^*} \), based on the set of points \( S \). We are going to rely on Theorem 4 to define the tight upper and lower bounds on \( C_{t^*} \).

#### 2.2.1 Derivation

In this section we will derive formulas for exact upper and lower bounds. For that we will rely on the theorems of the previous section.

Let \( \tilde{\mathcal{S}} = S \cup (k_{t^*}, C_{t^*}) \) and define \( \tilde{I}_t, \tilde{I}^\infty_t, \) and \( \tilde{\mathcal{A}}_t \) by using \( \tilde{\mathcal{S}} \) instead of \( S \) in Definition 3. If \((k_{t^*}, C_{t^*}) \in \text{int}(\mathcal{A}_{t^*})\), then \( \tilde{\mathcal{A}}_{t^*} = \mathcal{A}_{t^*} \) and \((k_{t^*}, C_{t^*}) \in \text{int}(\tilde{\mathcal{A}}_{t^*})\), which violates the requirements of Theorem 4. Consequently, we obtain:

**Corollary 1.** The tight upper bound on the price of an option with maturity \( t^* \) and a strike price \( k \) is given by \( g_n^u(k) \), where \( g_n^u : \mathbb{R}^+ \to \mathbb{R}^+ \), is a function which coincides with the lower border of \( \mathcal{A}_{t^*} \).
To obtain the tight lower bound on \( q^* \), we need the following notations and lemmas first.

**Notations:** Let \( P := (x, y) \) be the point in \( \mathbb{R}^+ \times \mathbb{R}^+ \). Define the function \( X : \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow \mathbb{R}^+ \), which maps each point \( P \) to its first coordinate, that is

\[
X(P) := x.
\]

In the same way define \( Y : \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow \mathbb{R}^+ \), so that

\[
Y(P) := y.
\]

Let \( P_1 = (x_1, y_1), P_2 = (x_2, y_2) \in \mathbb{R}^+ \times \mathbb{R}^+ \) and \( x_1 < x_2 \). Let \( \overline{l}_{12} : \mathbb{R}^+ \rightarrow \mathbb{R} \) be a linear function, such that \( \overline{l}_{12}(x_1) = y_1 \) and \( \overline{l}_{12}(x_2) = y_2 \).

Denote by \( \overline{P_1 P_2}^+, \overline{P_1 P_2}^-, \overline{P_1 P_2}, \overline{P_1}^-, \overline{P_1}^+ : \mathbb{R}^+ \rightarrow \mathbb{R}^+ \) the functions

\[
\begin{align*}
\overline{P_1 P_2}^+(x) & := \begin{cases} 
\max \{ \overline{l}_{12}(x), 0 \}, & x \geq x_2, \\
0, & x < x_2.
\end{cases} \\
\overline{P_1 P_2}^-(x) & := \begin{cases} 
\max \{ \overline{l}_{12}(x), 0 \}, & x \leq x_1, \\
0, & x > x_1.
\end{cases} \\
\overline{P_1 P_2}(x) & := \begin{cases} 
\overline{l}_{12}(x), & x_1 \leq x \leq x_2, \\
0, & x \in [0, x_1) \cup (x_2, +\infty).
\end{cases} \\
\overline{P_1}^-(x) & := \begin{cases} 
y_1, & x \leq x_1, \\
0, & x > x_1.
\end{cases} \\
\overline{P_1}^+(x) & := \begin{cases} 
y_1, & x \geq x_1, \\
0, & x < x_1.
\end{cases}
\end{align*}
\]

See also Figure 2-5.

Let \( \mathcal{P}_t^{\text{past}}, \mathcal{P}_t^{\text{future}}, \mathcal{P}_t^{\text{present}} \) denote the subsets of \( \mathcal{S} \), which correspond to maturities smaller, larger and equal to \( t \) respectively, that is

\[
\begin{align*}
\mathcal{P}_t^{\text{past}} & := \{(k_{ij}, C_{ij}) \mid (k_{ij}, C_{ij}) \in \mathcal{S}, \ l < t\}, \\
\mathcal{P}_t^{\text{future}} & := \{(k_{ij}, C_{ij}) \mid (k_{ij}, C_{ij}) \in \mathcal{S}, \ l > t\}, \\
\mathcal{P}_t^{\text{present}} & := \{(k_{ij}, C_{ij}) \mid (k_{ij}, C_{ij}) \in \mathcal{S}, \ l = t\}.
\end{align*}
\]
the lower boundary of the convex hull $A_t$, that is
\[
P^\theta_t := (P^\text{present}_t \cup P^\text{future}_t) \cap \partial A_t = (P^\text{future}_t \cap \partial A_t) \cup P^\text{present}_t.
\]
The last equality holds since $P^\text{present}_t \in \partial A_t$ by Theorem 4. Let us enumerate points of $P^\theta_t$:
\[
P^\theta_t := \{ P^\theta_{t_1}, P^\theta_{t_2}, \ldots, P^\theta_{\partial(t)} \}.
\] (2.10)
Here $\partial(t)$ denotes the number of elements of $P^\theta_t$.

Remark 5. In new notations, the function $g^u_t(k)$, defined in Corollary 1, can be defined as
\[
g^u_t(k) = \max \left\{ P^\theta_{t_j} P^\theta_{t_{j+1}}, P^\theta_{t_{\partial(t)}}^+, j = 1, 2, \ldots, \partial(t^*) - 1 \right\}(k).
\] (2.11)

Lemma 1. Let $0 = x_1 \leq x_2 \leq \ldots \leq x_n$ and $I = \cup_{1 \leq i \leq n}(x_i, y_i) \in \mathbb{R}^+ \times \mathbb{R}^+$ be the set of points such that there exists a convex nonincreasing function $f : \mathbb{R}^+ \to \mathbb{R}^+$, such that $f(x_i) = y_i$ for each $i$. Let
\[
g^d := \max \left\{ P_i P_{i+1}^-, P_i P_{i+1}^+, P_n^-, i = 1, 2, \ldots, n - 1 \right\}.
\]
Then for any $x^* > 0$ there exists $\tilde{f} : \mathbb{R}^+ \to \mathbb{R}^+$ such that it is convex and nonincreasing and goes through the set of points $I \cup (x^*, g^d(x^*))$. There does not exist a convex
nonincreasing function which goes through the set $I \cup (x^*, y^*)$ if $y^* < g^d(x^*)$.

![Figure 2-6: Function $g^d$ in the single maturity case](image)

![Figure 2-7: Function $g^d$ in the case of present and future prices. Here $P_1$, $P_3$, $P_4$ and $P_6$ are present constraint points, while $P_2$ and $P_5$ - future.](image)

**Remark 6.** Figure 2-6 illustrates the definition of the function $g^d$ for a given set of points.

**Proof.** Denote by $P^*$ the point $(x^*, y^*)$ and suppose first that $x_j \leq x^* \leq x_{j+1}$ for some $j \leq n - 2$. Consider function

$$h(x) := \max \left\{ P_1 P_2, \ldots, P_{j-2} P_{j-1} P_{j+1}^+, P_{j-1} P_j^+, P_{j+1} P_{j+2}^-, P_{j+2} P_{j+3}, \ldots, P_{n-1} P_n, P_{n-1} P_n^+ \right\}.$$  

This function is convex and nonincreasing and $h(x^*) = \max \left\{ P_{j-1} P_j^+, P_{j+1} P_{j+2}^- \right\} (x^*) \leq g^d(x^*)$.

If $x_{n-1} \leq x^* \leq x_n$, then $h(x) := \max \left\{ P_1 P_2, \ldots, P_{n-3} P_{n-2} P_n, P_{n-2} P_{n-1}^+, P_n^-, P_n^+ \right\}$ is convex and nonincreasing and $h(x^*) = \max \left\{ P_{n-2} P_{n-1}^+, P_n^- \right\} (x^*) \leq g^d(x^*)$.

At last, if $x^* > x_n$, we can take $h(x) := \max \left\{ P_1 P_2, \ldots, P_{n-2} P_{n-1}, P_{n-1} P_n^+ \right\}$ as a convex nonincreasing function such that $h(x^*) = P_{n-1} P_n^+ (x^*) \leq g^d(x^*)$.

The statement that if $y^* < g^d(x^*)$, then there does not exist a function with required properties, is obvious. \hfill $\square$

**Lemma 2.** Let $I = \cup_{1 \leq i \leq n} (x_i, y_i) \in \mathbb{R}^+ \times \mathbb{R}^+$ be the set of points as defined in Lemma 1 and $\mathcal{P}^1$ and $\mathcal{P}^2$ are any subsets of $I$, such that $\mathcal{P}^1 \cup \mathcal{P}^2 \equiv I$. Let

$$g^d = \max \left\{ \frac{P_{i+1}}{P_i} 1_{\{i+1 \in \mathcal{P}^1\}}, \frac{P_{i+1}}{P_i} 1_{\{i+1 \in \mathcal{P}^2\}}, \frac{P_n}{P_{i}} 1_{\{n \in \mathcal{P}^1\}}, \frac{P_n}{P_{i}} 1_{\{n \in \mathcal{P}^2\}} \right\}, i = 1, \ldots, n - 1,$$

where $1_{\{\cdot\}}$ is an indicator function. Then

(a) For any $x^* > 0$ there exists a convex nonincreasing function $\tilde{f} : \mathbb{R}^+ \rightarrow \mathbb{R}^+$, which goes through the set of points $\mathcal{P}^1 \cup (x^*, g^d(x^*))$, and below the set of points $\mathcal{P}^2$.

(b) There does not exist a convex nonincreasing function which goes through $\mathcal{P}^1 \cup (x^*, y^*)$ and below $\mathcal{P}^2$ if $y^* < g^d(x^*)$. 

36
Remark 7. Figure 2-7 demonstrates the construction of the function $g^d$ for a given set of points. Here $P^1 = \{P_1, P_3, P_4, P_6\}$ and $P^2 = \{P_2, P_3\}$.

Proof. (b) Assume that there exists a convex nonincreasing function $\tilde{f}$ which goes through $P^1 \cup (x^*, y^*)$ and below $P^2$ while $y^* < g^d(x^*)$. Let $y^* < \frac{P_i P_{i+1}^+}{P_i P_{i+1}^-} 1_{\{P_{i+1} \in P^1\}}$ for some $i \in \{1, \ldots, n-1\}$. Then we know that $x^* \geq x_{i+1}$ and $\tilde{f}(x_{i+1}) = y_{i+1}$. The last is true since $1_{\{P_{i+1} \in P^1\}} \neq 0$ and thus $P_{i+1} \in P^1$. Taking into account that $\tilde{f}(x_i) \leq y_i$, we have $x_i, \tilde{f}(x_i), \tilde{f}(x_{i+1}) \geq \frac{P_i P_{i+1}^+}{P_i P_{i+1}^-}$. Since $\tilde{f}$ is convex, it must be that $y^* \geq \frac{P_i P_{i+1}^+}{P_i P_{i+1}^-} (x^*)$ and consequently $y^* \geq \frac{P_i P_{i+1}^+}{P_i P_{i+1}^-} (x^*)$. That contradicts to our initial assumption. In the same way it can be shown that the assumption $y^* < \frac{P_i P_{i+1}^-}{P_i P_{i+1}^+} 1_{\{P_{i+1} \in P^1\}}$ for some $i = 1, \ldots, n-1$, leads to a contradiction. The requirement $y^* \geq \frac{P_i P_{i+1}^+}{P_i P_{i+1}^-} 1_{\{P_{i} \in P^1\}} (x^*)$ is clear, since otherwise the function is not nonincreasing.

(a) Let

$$x_{i_1} = \max_{i=1, \ldots, n} \{x_i \mid x_i \in P^1, x_i \leq x^*\}$$

and

$$x_{i_2} = \min_{i=1, \ldots, n} \{x_i \mid x_i \in P^1, x_i \geq x^*\}.$$ 

Note that $x_{i_1} \neq \emptyset$ since $x_1 = 0 \leq x^*$. If $x_{i_2} \neq \emptyset$ then there are three different cases possible. If $1 < i_1 < i_2 < n$, then as a convex nonincreasing function which goes through all the points of $P^1$ and below all the points of $P^2$ we can take

$$h = \max \left\{ \frac{P_1 P_2}{P_{i_1-1} P_{i_1}^+}, \frac{P_1 P_2}{P_{i_1-1} P_{i_1}^-}, \frac{P_{i_2} P_{i_2+1}}{P_{i_2} P_{i_2+1}^+}, \frac{P_{i_2} P_{i_2+1}}{P_{i_2} P_{i_2+1}^-}, \ldots, \frac{P_n}{P_n^+} \right\}.$$

Note that $h(x^*) = \max \left\{ \frac{P_1 P_2}{P_{i_1-1} P_{i_1}^+}, \frac{P_{i_2} P_{i_2+1}}{P_{i_2} P_{i_2+1}^+} \right\} (x^*) \leq g^d(x^*)$.

If $1 = i_1 < i_2 < n$, take

$$h = \max \left\{ \frac{P_2}{P_{i_2} P_{i_2+1}^+}, \frac{P_{i_2} P_{i_2+1}}{P_{i_2} P_{i_2+1}^-}, \ldots, \frac{P_n}{P_n^+} \right\}.$$ 

Then $h(x^*) = \frac{P_{i_2} P_{i_2+1}}{P_{i_2} P_{i_2+1}^-} (x^*) \leq g^d(x^*)$.

Lastly, if $1 < i_1 \leq i_2 = n$, take

$$h = \max \left\{ \frac{P_1 P_2}{P_{i_1-1} P_{i_1}^+}, \frac{P_{i_1-1} P_{i_1}^-}{P_{i_1-1} P_{i_1}^+}, \frac{P_n}{P_n^-}, \frac{P_n}{P_n^+} \right\}.$$ 

Here $h(x^*) = \max \left\{ \frac{P_{i_1-1} P_{i_1}^-}{P_{i_1-1} P_{i_1}^+}, \frac{P_n}{P_n^-} \right\} (x^*) \leq g^d(x^*)$. 

37
In case \( x_{i2} = \emptyset \), the function
\[
h = \max\{P_1 P_2, \ldots, P_{i-1} P_{i1}, P_{i1}^{-1} P_{i2}^+\}
\]
has all the required properties and \( h(x^*) = P_{i1}^{-1} P_{i2}^+(x^*) \leq g^d(x^*) \).
The proof is complete.

**Lemma 3.** Let \( I \) and \( \mathcal{P}^1 \) and \( \mathcal{P}^2 \) are as defined in lemma 4. Then there exists a convex nonincreasing function \( \tilde{f} : \mathbb{R}^+ \to \mathbb{R}^+ \) which goes through the set \( \mathcal{P}^1 \cup (x^*, y^*) \) and below the set \( \mathcal{P}^2 \) if and only if \( g^d(x^*) \leq y^* \leq g^u(x^*) \), where \( g^d(x^*) \) is as defined in lemma 4 and \( g^u = \max\{P_1 P_2, \ldots, P_{n-1} P_n, P_n^+\} \).

**Proof.** It has been shown that this condition is necessary. Let us show that it is sufficient. Let \( f^u(x) \) be a convex nonincreasing function which goes through the set \( \mathcal{P}^1 \cup (x^*, g^u(x^*)) \) and below \( \mathcal{P}^2 \) and let \( f^d(x) \) be a convex nonincreasing function which goes through the set \( \mathcal{P}^1 \cup (x^*, g^d(x^*)) \) and below \( \mathcal{P}^2 \). Then if we take a linear combination \( \alpha f^u + (1 - \alpha) f^d \), where \( 0 \leq \alpha \leq 1 \), we will obtain a convex nonincreasing function which goes through \( \mathcal{P}^1 \cup (x^*, \alpha g^u(x^*) + (1 - \alpha) g^d(x^*)) \) and below \( \mathcal{P}^2 \).

**Proposition 4.** Let the set \( S \), as defined in (2.16), satisfies the no-arbitrage condition and \( k_{t1} = 0 \) and \( C_{t1} = C_{11} \) for each \( (k_{t1}, C_{t1}) \in S \). Then the tight upper and lower bounds on the price \( C^* \) of an option with a maturity \( t^* \) and a strike price \( k^* \) are given by

\[
g^u_{t^*}(k^*) = \max\left\{ P_{t_{j1}} P_{t_{j+1}}, P_{t_{j1}}^{-1} P_{t_{j+1}}^+ : j = 1, 2, \ldots, \partial(t^*) - 1 \right\} (k^*), \tag{2.12}
\]

\[
g^d_{t^*}(k^*) := \max_{t = 1, 2, \ldots, t^*} g^d_{t(t)}(k^*), \tag{2.13}
\]

correspondingly, where

\[
g^d_{t(t)}(k^*) = \max\left\{ C_{11} - k, \frac{P_{t_{j1}} P_{t_{j+1}}}{P_{t_{j1}} P_{t_{j+1}}^{-1}} 1\{p_{t_{j+1}} \in p_{t}^{\text{present}}\}, \frac{P_{t_{j1}} P_{t_{j+1}}}{P_{t_{j1}} P_{t_{j+1}}^{-1}} 1\{p_{t_{j+1}} \in p_{t}^{\text{present}}\} : j = 1, 2, \ldots, \partial(t^*) - 1 \right\} (k^*). \tag{2.14}
\]

**Proof.** That \( g^d_{t^*}(k^*) \leq C^* \leq g^u_{t^*}(k^*) \) is a necessary condition follows from Corollary 1, Lemma 2 and Proposition 3.
Let \( g_t^u(k) \) designate the function which gives the tight upper bound on the price of the option with a maturity \( t \) and a strike price \( k \) (see definition in Corollary 1). Let us also redenote \((k^*, C^*)\) as \((k^*_{t_1}, C^*_{t_1})\), so that it is clear to which maturity this point belongs.

Assume that \( g_t^u(k^*_{t_1}) < C^*_{t_1} < g_t^u(k^*_{t_1}) \) but the no-arbitrage condition is not satisfied once \((k^*_{t_1}, C^*_{t_1})\) is adjoined to \( S \) to options of maturity \( t^* \). Let \( \tilde{P}_{t_1}^{\text{present}} \) and \( \tilde{P}_{t_1}^{\partial} \) be as defined in (2.9) but with \( S \cup (k^*_{t_1}, C^*_{t_1}) \) taken instead of \( S \). Since \( C_{t_1} - k^*_{t_1} \leq g_{t_1}^d(k^*_{t_1}) \leq C^*_{t_1} \), then by Theorem 4 there must exist \( t_1 \in \{1, \ldots, n\} \) such that \( \tilde{P}_{t_1}^{\text{present}} \notin \tilde{P}_{t_1}^{\partial} \). The last means that there is at least one point of \( \tilde{P}_{t_1}^{\text{present}} \) which is in the interior of the convex hull \( \tilde{A}_{t_1} \), and thus there does not exist a convex nonincreasing function which goes through \( \tilde{P}_{t_1}^{\text{present}} \) and below \( \tilde{P}_{t_1}^{\partial} \).

If \( t_1 > t^* \), then, by definition, \( \tilde{A}_{t_1} \) and \( \tilde{P}_{t_1}^{\text{present}} \) do not depend on \((k^*_{t_1}, C^*_{t_1})\). If \( t_1 = t^* \), then we have a contradiction with Lemma 3, since \( g_{t_1}^d(k^*_{t_1}) \leq C^*_{t_1} \leq g_t^u(k^*_{t_1}) \) and thus there must exist a convex nonincreasing function which goes through \( \tilde{P}_{t_1}^{\text{present}} \) and below \( \tilde{P}_{t_1}^{\partial} \). If \( t_1 < t^* \) and \( C^*_{t_1} > g_t^u(k^*_{t_1}) \), then adjoining \((k^*_{t_1}, C^*_{t_1})\) changes neither \( P_{t_1}^{\text{present}} \) nor \( P_{t_1}^{\partial} \). Then it must be that \( C^*_{t_1} \leq g_t^u(k^*_{t_1}) \). But then \( g_{t_1}^d(k^*_{t_1}) \leq C^*_{t_1} \leq g_t^u(k^*_{t_1}) \) and, again, by Lemma 3 there must exist a convex nonincreasing function which goes through \( \tilde{P}_{t_1}^{\text{present}} \) and below \( \tilde{P}_{t_1}^{\partial} \). Contradiction.

Remark 8. The tight upper and lower bounds on the price of the option with maturity \( t^* \) such that \( t < t^* < t + 1 \), and strike price \( k \), are given by \( g_t^u(k) \) and \( g_t^d(k) \) respectively.

Corollary 2. \( g_t^d(k) : \mathbb{R}^+ \to \mathbb{R}^+ \), is a continuous piecewise-linear nonincreasing function which turns to zero for \( k \) large enough.

Proof. Notice that

\[
\begin{align*}
g_t^d(k^*) &= \max \left\{ C_t^1 - k, \max \left\{ \frac{\tilde{P}_{t_j} - \tilde{P}_{t_j}^+}{\tilde{P}_{t_j}^+ - \tilde{P}_{t_j-1}^+} \right\} \{ p_{t_j}^+ \} \right\} \{ p_{t_j}^+ \in P_{t_j}^{\text{present}} \} \\
&\max \left\{ \frac{\tilde{P}_{t_0} - \tilde{P}_{t_0}^+}{\tilde{P}_{t_0}^+ - \tilde{P}_{t_0-1}^+} \right\} \{ p_{t_0}^+ \in P_{t_0}^{\text{present}} \}, \quad j = 1, 2, \ldots, \partial(t) - 1 \right\}.
\end{align*}
\]

Here \( h_{t_j}(k) := \max \left\{ \frac{\tilde{P}_{t_j} - \tilde{P}_{t_j}^+}{\tilde{P}_{t_j}^+ - \tilde{P}_{t_j-1}^+} \right\} \{ p_{t_j}^+ \} \) is a continuous piecewise-linear, nonincreasing function for each \( j = 1, 2, \ldots, \partial(t) - 1 \). The same is true for the function \( h_{t_0}(k) := \max \left\{ \frac{\tilde{P}_{t_0} - \tilde{P}_{t_0}^+}{\tilde{P}_{t_0}^+ - \tilde{P}_{t_0-1}^+} \right\} \{ p_{t_0}^+ \} \). Each function \( h_{t_j}(k), j = 1, 2, \ldots, \partial(t) \), turns to zero for \( k \) large enough. Indeed, for any two points \( \{P_1, P_2\} \in P_{t_j}^{\partial} \), such that
$X(P_1) < X(P_2)$, it must be $Y(P_1) > Y(P_2)$ for a convex function going through $(X(P_1), Y(P_1))$ and $(X(P_2), Y(P_2))$ and approaching zero as $k \to +\infty$ to exist.

Thus $g^d_t(k)$, as a maximum of a subset of these functions and $f(k) = C_{11} - k$, is also continuous, piecewise-linear, nonincreasing and turns to zero for $k$ large enough. $g^d_t(k)$, in its turn, is a maximum of $g^d_{(t)}(k)$, $t = 1, \ldots, t^*$, and, consequently, also has the specified properties. □

2.2.2 Algorithm

Given the set of options

$$S := \{(k_{ij}, C_{ij}) \mid (k_{ij}, C_{ij}) \in \mathbb{R}^{2+}, k_{t1} = 0, C_{t1} = C_{11}, t \in \{1, \ldots, n\}, j \in \{1, \ldots, v(t)\}\}, \quad (2.16)$$

which satisfies the no-arbitrage condition, find the tight upper and lower bounds on the price $C^*$ of an option with maturity $t^* \in \{1, 2, \ldots, n\}$ and a strike price $k^* > 0$.

Let

$$S_t := \{(k_{ij}, C_{ij}) \mid (k_{ij}, C_{ij}) \in S, l = t\}$$

represent the set of options of maturity $t$ and

$$S_{\geq t} := \bigcup_{l \geq t} S_l$$

represent the set of options of maturities $t$ and larger. Let $N$ be the total number of options.

Upper Bound

1. Find the subset of points of $S_{\geq t^*}$, which are on the border of the convex hull of $(S_{\geq t^*} \cup (\infty, 0) \cup (0, \infty))$. Denote this set by $B_{t^*}$. (This step requires $O(N^2)$ operations.)

2. Find the points

$$\left\{(k_{\text{left}}, C_{\text{left}}) \mid (k_{\text{left}}, C_{\text{left}}) \in B_{t^*}, k_{\text{left}} = \max_{ij} \{k_{ij} \mid k_{ij} \leq k^*, (k_{ij}, C_{ij}) \in B_{t^*}\} \right\}$$

and

$$\left\{(k_{\text{right}}, C_{\text{right}}) \mid (k_{\text{right}}, C_{\text{right}}) \in B_{t^*}, k_{\text{right}} = \min_{ij} \{k_{ij} \mid k_{ij} \geq k^*, (k_{ij}, C_{ij}) \in B_{t^*}\} \right\},$$

that is, the point of $B_{t^*}$, which is the closest to $(k^*, C^*)$ from the left and the
one which is the closest from the right. Notice, that \((k_{t}^{\text{left}}, C_{t}^{\text{left}})\) always exists, since \(k_{t} = 0\) for any \(t\). (This step requires \(O(N)\) operations.)

3. The upper bound on \(C^{*}\) is given by

\[
g^{u} := \begin{cases} 
C_{t}^{\text{left}} - C_{t}^{\text{right}} & k^{*} + \frac{k_{t}^{\text{left}} C_{t}^{\text{right}} - k_{t}^{\text{right}} C_{t}^{\text{left}}}{k_{t}^{\text{left}} - k_{t}^{\text{right}}}, \\
C_{t}^{\text{left}}, & (k^{\text{right}}, C^{\text{right}}) \neq \emptyset, \\
0, & (k^{\text{right}}, C^{\text{right}}) = \emptyset.
\end{cases}
\]

**Lower Bound**

1. For each \(t = 1, 2, \ldots, t^{*}\) find the subset of points of \(S_{\geq t}\), which are on the border of the convex hull of \((S_{\geq t} \cup (\infty, 0) \cup (0, \infty))\). Denote this set by \(B_{t}\). Keep in memory \(B_{t}^{\text{present}} := B_{t} \cap S_{t}\) and \(B_{t}^{\text{future}} := B_{t} \setminus B_{t}^{\text{present}}\). (This step requires \(O(N^{2})\) operations.)

2. Find the points

\[
\left\{ \left( k_{t}^{\text{left}}, C_{t}^{\text{left}} \right) \left| \left( k_{t}^{\text{left}}, C_{t}^{\text{left}} \right) \in B_{t}^{\text{present}}, k_{t}^{\text{left}} = \max_{j} \left\{ k_{tj} | k_{tj} \leq k^{*} \right\} \right. \right\} \quad \text{and}
\]

\[
\left\{ \left( k_{t}^{\text{right}}, C_{t}^{\text{right}} \right) \left| \left( k_{t}^{\text{right}}, C_{t}^{\text{right}} \right) \in B_{t}^{\text{present}}, k_{t}^{\text{right}} = \min_{j} \left\{ k_{tj} | k_{tj} \geq k^{*} \right\} \right. \right\},
\]

that is, points of \(B_{t}^{\text{present}}\), which are the closest to \((k^{*}, C^{*})\) from the left and from the right. Now find the points

\[
\left\{ \left( \tilde{k}_{t}^{\text{left}}, \tilde{C}_{t}^{\text{left}} \right) \left| \left( \tilde{k}_{t}^{\text{left}}, \tilde{C}_{t}^{\text{left}} \right) \in B_{t}, \tilde{k}_{t}^{\text{left}} = \max_{j} \left\{ k_{tj} | k_{tj} < k_{t}^{\text{left}} \right\} \right. \right\} \quad \text{and}
\]

\[
\left\{ \left( \tilde{k}_{t}^{\text{right}}, \tilde{C}_{t}^{\text{right}} \right) \left| \left( \tilde{k}_{t}^{\text{right}}, \tilde{C}_{t}^{\text{right}} \right) \in B_{t}, \tilde{k}_{t}^{\text{right}} = \min_{j} \left\{ k_{tj} | k_{tj} > k_{t}^{\text{right}} \right\} \right. \right\}.
\]

If any of the above points does not exist, assign it an empty set type. (This step requires \(O(N)\) operations.)

3. Let

\[
g_{t}^{\text{left}} := \begin{cases} 
C_{t}^{\text{left}} - \tilde{C}_{t}^{\text{left}} - k^{*} + \frac{k_{t}^{\text{left}} C_{t}^{\text{right}} - \tilde{k}_{t}^{\text{left}} C_{t}^{\text{left}}}{k_{t}^{\text{left}} - \tilde{k}_{t}^{\text{left}}}, \\
C_{t}^{\text{left}}, & (\tilde{k}_{t}^{\text{left}}, \tilde{C}_{t}^{\text{left}}) \neq \emptyset, (k_{t}^{\text{left}}, C_{t}^{\text{left}}) \neq \emptyset, \\
0, & (\tilde{k}_{t}^{\text{left}}, \tilde{C}_{t}^{\text{left}}) = \emptyset, (k_{t}^{\text{left}}, C_{t}^{\text{left}}) \neq \emptyset, \\
0, & (\tilde{k}_{t}^{\text{left}}, \tilde{C}_{t}^{\text{left}}) \neq \emptyset, (k_{t}^{\text{left}}, C_{t}^{\text{left}}) = \emptyset.
\end{cases}
\]

41
In analogous way (just changing "left" to "right") define \( g_t^{\text{right}} \). Let

\[
g_t := \max \left\{ g_t^{\text{left}}, g_t^{\text{right}}, 0 \right\}.
\]

![Graph](image)

**Figure 2-8**: Determining \( g_t \): black points are points of \( B_t^{\text{present}} \) and white points are points of \( B_t^{\text{future}} \).

4. Then the exact lower bound is given by

\[
g_t^d := \max \left\{ C_{11} - k^*, \max_{t=1,...,t^*} g_t \right\}.
\]

The total running time is \( O(N^2) \) for both upper and lower bounds.

**Remark 9.** Notice, that it is very easy to adjust the algorithm for the case, when the price of the stock is not available. That may be useful since, for example, prices of options and stocks reported at the end of the day do not correspond to exactly the same time. We are not going to discuss the details of this adjustment since they are trivial, however in the following computational example the price of the stock is not taken into account.

### 2.2.3 Computational example

The data on VeriZonCM call options is taken from The Wall Street Journal of October 22 2002 to derive exact bounds on the price of an option with maturity 80 days. The price of the stock is $37.75. Figure 2-9 represents available options and gives exact bounds on the price of an option with a strike price 34. Part (a) of the figure reports the results for the case when only options of the corresponding maturity are used. Part (b) demonstrates bounds for the case when options of other maturities are also taken into account.
Tables 2.1 and 2.2 compare the bounds for other strike prices. Note, that not always the information on options of other maturities improves the results. Let, as before, \( t^* \) denotes the maturity of an option of interest. The more options of maturity \( t^* \) are available, the more determined is the function \( f \), which goes through the set \( S_{t^*} \) of strike prices and prices of options of maturity \( t^* \), the less likely it is that options of other maturities will put additional constraints on \( f \) and, thus, influence the bounds.

Note, that if prices of options of maturities other than \( t^* \) are available only for the same set of strike prices as the one for maturity \( t^* \), and the discounting factor for these maturities is 1, then options of other maturities can not influence the bounds. Of course, they are valuable if we are interested in the price of an option with maturity not among given.

(a) Only options with maturity 80 days are used to find bounds. The range is [5.10, 5.80].

(b) All given options are used to find bounds. The range is [5.54, 5.80].

Figure 2-9: Bounds on the price of an option with maturity 80 days and a strike price 36.

2.3 Bounds on option if the variance of the underlying asset is known

Our next problem is to determine the tight upper and lower bounds on the price of an option, if besides the prices of other options, we are also given a condition on
<table>
<thead>
<tr>
<th>strike price</th>
<th>32</th>
<th>33</th>
<th>34</th>
<th>35</th>
<th>36</th>
<th>37</th>
<th>38</th>
<th>39</th>
<th>40</th>
<th>41</th>
<th>42</th>
</tr>
</thead>
<tbody>
<tr>
<td>lower bound</td>
<td>6.60</td>
<td>5.98</td>
<td>5.54</td>
<td>5.10</td>
<td>4.40</td>
<td>3.70</td>
<td>3.00</td>
<td>2.30</td>
<td>2.00</td>
<td>1.38</td>
<td>0.76</td>
</tr>
<tr>
<td>upper bound</td>
<td>7.20</td>
<td>6.50</td>
<td>5.80</td>
<td>5.10</td>
<td>4.66</td>
<td>4.22</td>
<td>3.78</td>
<td>3.34</td>
<td>2.90</td>
<td>2.46</td>
<td>2.02</td>
</tr>
</tbody>
</table>

Table 2.1: Bounds on the price of an option with maturity 80 days and a given strike price based on all available options.

<table>
<thead>
<tr>
<th>strike price</th>
<th>32</th>
<th>33</th>
<th>34</th>
<th>35</th>
<th>36</th>
<th>37</th>
<th>38</th>
<th>39</th>
<th>40</th>
<th>41</th>
<th>42</th>
</tr>
</thead>
<tbody>
<tr>
<td>lower bound</td>
<td>6.60</td>
<td>5.60</td>
<td>5.10</td>
<td>5.10</td>
<td>4.40</td>
<td>3.70</td>
<td>3.00</td>
<td>2.30</td>
<td>1.60</td>
<td>0.90</td>
<td>0.20</td>
</tr>
<tr>
<td>upper bound</td>
<td>7.20</td>
<td>6.50</td>
<td>5.80</td>
<td>5.10</td>
<td>5.10</td>
<td>5.10</td>
<td>5.10</td>
<td>5.10</td>
<td>5.10</td>
<td>5.10</td>
<td>5.10</td>
</tr>
</tbody>
</table>

Table 2.2: Bounds on the price of an option with maturity 80 days and a given strike price based only on the options of the corresponding maturity.

the variance of the underlying stock. More precisely, suppose that \( \text{Var}(X_t) \) for some \( t \in \{1, 2, \ldots, n\} \) is known. We are interested in finding tight upper and lower bounds on the price \( C^* \) of an option with maturity \( t^* \) and strike price \( k^* \).

Notice that since \( E[X_t] \) is known, the condition on the variance is equivalent to the condition on \( E[X_t^2] \). Let \( \mu_1 \) and \( \mu_2 \) be two measures which both make the process \( X_1, X_2, \ldots, X_n \) a martingale satisfying the set \( S \).

Define the measure \( \mu := \alpha \mu_1 + (1 - \alpha) \mu_2 \) for some \( \alpha \in (0, 1) \). Then under \( \mu \)
\( X_1, X_2, \ldots, X_n \) is also a martingale and it satisfies the set \( S \). Since also

\[
E_\mu[X_t^2] = \alpha E_{\mu_1}[X_t^2] + (1 - \alpha) E_{\mu_2}[X_t^2] \quad \text{and}
\]

\[
E_\mu[(X_{t^*} - k^*)^+] = \alpha E_{\mu_1}[(X_{t^*} - k^*)^+] + (1 - \alpha) E_{\mu_2}[(X_{t^*} - k^*)^+],
\]

then the set \( V := \{(\text{Var}(X_t), C^*)\} \in \mathbb{R}^{2+} \) of possible combinations of \( \text{Var}(X_t) \) and \( C^* \) must be convex. Suppose that the range of possible values for \( C^* \), based solely on the set \( S \), is \([a, b]\) and the range of possible values for \( \text{Var}(X_t) \), based on the same set, is \([c, d]\). Then \( V \) must lie within the rectangle \([c, d] \times [a, b] \in \mathbb{R}^{2+}\) and functions \( f_1, f_2 : [c, d] \to [a, b] \), which form the upper and lower bounds of the set, must be nondecreasing (Figure 2.10.) If we know \( f_1 \) and \( f_2 \), as well as the interval \([c, d]\), then
for a given value of $\text{Var}(X_t)$, we know the range of possible values of $C^*$, such that the martingale $X_1, X_2, \ldots, X_n$, that satisfies the set $S \cup (k^*, C^*)$ and has $\text{Var}(X_t)$ as given, exists. On the other hand, if for each $C^* \in [a, b]$ we know the bounds on $\text{Var}(X_t)$, then we know the interval $[c, d]$ and functions $f_1$ and $f_2$.

In the problem of determining bounds on $\text{Var}(X_t)$ if the set $S$ and $(k^*, C^*)$ are given, the special role of the option $(k^*, C^*)$ disappears. Thus the last problem is equivalent to the problem of determining exact bounds on the variance of the price of the underlying asset at a certain time, if the set $S$ of priced options is given.

### 2.4 Bounds on variance

In this section we derive a formula for the exact upper bound on the variance of the underlying asset at a given time, and develop an algorithm for the lower bound.

**Definition 4.** Let $\Pi^S_t \in C$ designate the class of probability distributions $\{\Pi\} \in C$, such that there exists a martingale $X_1, X_2, \ldots, X_t^*, \ldots, X_n$, satisfying the set $S$, with $X_{t^*}$ distributed according to $\Pi$. Let

$$\Psi^S_{t^*} = \{ \Psi^*_\Pi(k) \mid \Pi \in \Pi^S_{t^*} \},$$  \hspace{1cm} (2.17)

$$\Psi^\text{max}_{t^*} = \left\{ \Psi^\text{max}_{t^*}(k) \mid \Psi^\text{max}_{t^*}(k) = \Psi^\text{max}_{\Pi^\text{max}}(k), \Pi^\text{max} = \arg\max_{\Pi \in \Pi^S_{t^*}} \int_0^\infty x^2 d\Pi(x) \right\},$$  \hspace{1cm} (2.18)

$$\Psi^\text{min}_{t^*} = \left\{ \Psi^\text{min}_{t^*}(k) \mid \Psi^\text{min}_{t^*}(k) = \Psi^\text{min}_{\Pi^\text{min}}(k), \Pi^\text{min} = \arg\min_{\Pi \in \Pi^S_{t^*}} \int_0^\infty x^2 d\Pi(x) \right\},$$  \hspace{1cm} (2.19)

Recall, that the distribution $\Pi$ is uniquely determined by its transform $\Psi^*_\Pi(k)$.
(Theorem 1 (a)), specifically, \( \Pi((k, +\infty]) = -\Psi_t^\prime(k+) \). Thus, the upper bound on 
\( E_{\Pi_t}^{\mathbb{E}}[X_t^2] \) is also uniquely determined by \( \Psi_t^{\max}(k) \). The same is true for the lower bound.

By definition \( \{\Psi_t^{\max}, \Psi_t^{\min}\} \subseteq \Psi_t^S \). Thus, in order to find a function which belongs to \( \Psi_t^{\max} \), or \( \Psi_t^{\min} \), it will be useful first to give a qualitative characteristic to functions of \( \Psi_t^S \).

**Theorem 5.** Let \( S \) be the set of options that satisfies the no-arbitrage condition (Definition 1) and \( k_{t1} = 0 \) and \( C_{t1} = C_{11} \) for each \( (k_{t1}, C_{t1}) \in S \). Let \( f(x) : \mathbb{R}^+ \to \mathbb{R}^+ \) be a convex nonincreasing function, such that \( \lim_{x\to\infty} f(x) = 0 \). Then \( f \in \Psi_t^S \) if and only if \( g_t^d \leq f \leq g_t^u \), where \( g_t^d \) and \( g_t^u \) are as defined in Proposition 4.

**Proof.** " \( \iff \) " It follows from Proposition 4.

" \( \Rightarrow \) " By Theorem 4 the martingale satisfying \( S \) exists if none of the points \( (k_{t1}, C_{t1}) \) is in the interior of \( A_t \), and for each of these points \( C_{t1} \geq C_{11} - k_{t1} \). Let, \( S^{\text{uf}} = S \cup \{(k, f(k))\}_{k \geq 0}, \) so that each \( (k, f(k)) \) is assigned maturity \( t^* \) within \( S^{\text{uf}} \). Define \( A_t^f \) as \( A_t \) is defined in Definition 3 but with \( S^{\text{uf}} \) taken instead of \( S \). We want to show that for the set \( S^{\text{uf}} \) the conditions of Theorem 4 are still satisfied.

First of all, \( f(k) \geq g_t^d(k) \geq C_{11} - k \). Adjoining points of \( f \) can not influence convex hulls \( A_{t+1}, A_{t+2}, \ldots, A_{t(t^*)} \), since the set \( \{(k, f(k))\} \) corresponds to a smaller maturity.

Since \( f(k) = C \) for each \( (k, C) \in \mathcal{P}_t^{\text{present}} (g_t^d(k) = g_t^u(k) = C \forall (k, C) \in \mathcal{P}_t^{\text{present}}) \), \( f(k) \leq C \) for each \( C \in \mathcal{P}_t^{\text{future}} (f \leq g_t^{2,2}) \) and \( f \) is a convex function, then it must form the border of the convex hull \( A_t^f \).

The only thing left to show is that none of the points \( (k_{t1}, C_{t1}) \in S \) is in the interior of \( A_t^f \) for \( t < t^* \). Let us assume that some point \( (k_{lm}, q_{lm}) \), \( l < t^* \), is in the interior of \( A_t^f \). Then, there exist points \( (k', q') \), and \( (k'', q'') \) to the left and to the right of \( (k_{lm}, q_{lm}) \) corresponding to options of maturity \( l \) or higher, such that the line passing through these points is strictly below \( (k_{lm}, q_{lm}) \). Notice that at least one of the two points, \( (k', q') \), and \( (k'', q'') \), must be a point of \( \{(k, f(k))\}_{k \geq 0} \), since prior to adjoining \( \{(k, f(k))\}_{k \geq 0} \), all the points of \( S \) satisfied conditions of Theorem 4. Suppose that \( (k', q') \in \{(k, f(k))\}_{k \geq 0} \), but \( (k'', q'') \notin \{(k, f(k))\}_{k \geq 0} \). Then the line connecting \( (k', q') \), and \( (k'', q'') \) can not go below \( (k_{lm}, q_{lm}) \). Indeed, adjoining the only point \( (k', q') \), as a point corresponding to maturity \( t^* \), with \( q' \) being within the bounds on the price of an option with maturity \( t^* \) and a strike price \( k' \), can not lead to the violation of conditions of martingale existence. Thus both, \( (k', q') \) and \( (k'', q'') \),
must be in $\{(k, f(k))\}_{k \geq 0}$. Note, however, that $f(k_{l,m}) \geq q_{l,m}$ since $l \leq t^*$. But then we have a contradiction with the requirement that $f$ is convex. \qed

**Proposition 5.** Let $S$ be the set of options as defined in Theorem 5 and $M > 0$ be the upper bound on the price of the underlying asset at time $t^*$. Then $\Psi_{t^*}^{\text{max}} : \mathbb{R}^+ \to \mathbb{R}^+$, coincides with the lower border of the convex hull of $\{I_{t^*} \cup (M,0)\}$ (2.5). If $M < \infty$, then the exact upper bound on $\text{Var}(X_{t^*})$ is

$$V_{t^*}^{\text{max}} = \sum_j k_j^2 \left( \Psi_{t^*}^{\text{max}}(k_j^+) - \Psi_{t^*}^{\text{max}}(k_j^-) \right) + M^2 \left( \Psi_{t^*}^{\text{max}}(M+) - \Psi_{t^*}^{\text{max}}(M-) \right) - C_{11}$$

where $\Psi_{t^*}^{\text{max}}(k_j^+)$ is the right derivative of $\Psi_{t^*}^{\text{max}}$ at $k_j$ and $k_j = X(P_{t^*}^{\partial_j}, j = 1, \ldots, \partial(t^*)$. If $M = +\infty$, then the upper bound is $+\infty$.

**Proof.** Notice, that $P(X_{t^*} > M) = 0 \iff E[(X_{t^*} - M)^+] = 0$. Thus the requirement $X_{t^*} < M$ will be automatically fulfilled if the option $(M,0)$ of maturity $t^*$ is added to the set $S$.

Let $g_{t^*}^u(k) : \mathbb{R}^+ \to \mathbb{R}^+$, be the function which coincides with the lower border of the convex hull of $\{I_{t^*} \cup (M,0)\}$. Since $\Psi_{t^*}^{\text{max}} \in \Psi_{t^*}^S$, then, by Theorem 5, for any $\Psi_{t^*}^{\text{max}} \in \Psi_{t^*}^{\text{max}}$ it must be that $\Psi_{t^*}^{\text{max}} \leq g_{t^*}^u$. At the same time, $g_{t^*}^u \in \Psi_{t^*}^S$.

Now if $X_1$ and $X_2$ are two random variables such that $\Psi_{C(X_1)} \leq \Psi_{C(X_2)}$, then by Theorem 3 we can define a joint distribution of $X_1$ and $X_2$ so that $X_1, X_2$ is a martingale. Consequently, if $\Psi_{C(X_1)} \leq \Psi_{C(X_2)}$, then $E[X_1^2] \leq E[X_2^2]$ by Jensen’s inequality for convex functions of martingales [13, p.277]. We conclude that $\Psi_{t^*}^{\text{max}} \equiv \{g_{t^*}^u\}$.

Let $\Pi \in \mathcal{C}$ be the law with its $\Psi$-transform equal to $\Psi_{t^*}^{\text{max}}$. Since $\Psi_{t^*}^{\text{max}}$ is a piecewise-linear function, then by Theorem 1, $\Pi$ is an atomic distribution. The weight of each atom $\{k_j\}$ is equal to the difference between the right and left derivatives of $\Psi_{t^*}^{\text{max}}$ at $k_j$. $\Psi_{t^*}^{\text{max}} = g_{t^*}^u$ can change its derivative only at points $k_j = X(P_{t^*}^{\partial_j}), j = 1, 2, \ldots, \partial(t^*)$, and $k_j = M$. Thus we obtain (2.20).

Let us consider the case when $M = +\infty$. Let $k' > 0$ be the strike price of an option of maturity $t^*$, such that the tight lower bound on its price is zero. Then adding a point $\{(k_n,0)\}, k_n > k'$, to the set of points corresponding to maturity $t^*$ will not violate the conditions of martingale existence. Consequently, the function $\Psi_{t^*}^{(n)}$ which is a lower border of the convex hull of points $I_{t^*} \cup \{(k_n,0)\}$ is in $\Psi_{t^*}^S$. Note however that we can make $k_n^2 \left( \Psi_{t^*}^{(n)}(k_n^+) - \Psi_{t^*}^{(n)}(k_n^-) \right) = -k_n^2 \Psi_{t^*}^{(n)}(k_n^-) \sim k_n$ be
arbitrarily large, by choosing $k_n > k'$ sufficiently large. Thus, there is no upper bound on $\text{Var}(X_t)$ in this case. 

Now we will turn to the problem of determining the lower bound.

**Definition 5.** Let

\[ \mathcal{P}^g := (\mathcal{P}^\text{present}_t \cup \mathcal{P}^\text{past}_t) \cap \{ (k, g^d_t(k)) \}_{k \geq 0} \]

represent the set of present and past constraint points $(k, C) \in \mathbb{R}^2$, such that $C = g^d_t(k)$,

\[ \mathcal{P}^\text{past} := \mathcal{P}^\text{past}_t \cap \mathcal{P}^g := \{ P^p_1, P^p_2, \ldots, P^p_{n_{\text{past}}} \}, \]

\[ \mathcal{P}^\text{future} := \mathcal{P}^\text{future}_t \cap \mathcal{P}^g := \{ P^f_1, P^f_2, \ldots, P^f_{n_{\text{future}}} \}, \]

\[ \mathcal{P}^\text{present} := \mathcal{P}^\text{present}_t := \{ P_1, P_2, \ldots, P_{n_{\text{present}}} \} \]

**Remark.** We introduce $\mathcal{P}^\text{future}$, since only future points from this set can influence the optimal solution. We introduce $\mathcal{P}^\text{past}$ since past prices from this set will have a special role in constructing $\Psi^\text{min}_t$, while others might influence the solution only through the function $g^d_t$.

**Proposition 6.** $\Psi^\text{min}_t(k) : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ is a piecewise linear function. It contains no more than $4n$ line segments, where $n$ is the number of elements of $\mathcal{P}^g$.

**Proof.** Assume that $(\Psi^\text{min}_t)^\prime(k) \neq 0$ on $[a, b]$. Then by drawing tangent lines to $\Psi^\text{min}_t(k)$ at points $k = a$ and $k = b$, and continuing $\Psi^\text{min}_t(k)$ by these tangent lines in the direction of $\frac{a+b}{2}$, the function can be reduced while remaining convex and nonincreasing, since it is bounded from below only by a piecewise linear function $g^d_t(k)$. Here and further the following lemma will be helpful.

**Lemma 4.** Each line segment of $\Psi^\text{min}_t(k)$ contains $\{ P \} \in \mathcal{P}^g$ or follows for a while a line segment of $g^d_t(k)$.

**Proof.** If none of the above conditions holds for a given segment, then this segment can be moved parallelly to itself till it hits past or present point or a linear constraint parallel to it. As a result, the function will be decreased, while remaining convex and nonincreasing. 

Returning to Proposition 6, notice, that if the number of present and past constraint points on $g^d_t$ is $n$, then $g^d_t$ contains no more than $2n$ line segments itself (see Proposition 4 and Corollary 2.) Consequently, no more than $2n$ segments of $\Psi^\text{min}_t$. 

48
can follow line segments of $g^d_t$. At the same time, $n$ points of $P^g$ can touch no more than $2n$ line segments of $\Psi_t^{\min}$. Hence, by Lemma 4, $\Psi_t^{\min}$ has no more than $4n$ line segments.

Notice, that if there are no past constraint points, then every line segment of $g^d_t$ with a nonzero slope goes through some $P \in P^{\text{present}}$. Thus we have the following corollary.

**Corollary 3.** In the absence of past points, each line segment of $\Psi_t^{\min}$ with nonzero slope must contain $P \in P^{\text{present}}$. $\Psi_t^{\min}$ consists of no more than $n + 1$ line segments in this case.

**Definition 6.** (a) We will call a line segment $(Q_1 Q_2) \subset \{(t, f(t))\}_{t \geq 0} \in \mathbb{R}^{2+}$ a maximum line segment of function $f$, if there does not exist a line segment $(\bar{Q}_1 \bar{Q}_2) \subset \{(t, f(t))\}_{t \geq 0}$, such that $(Q_1 Q_2) \subset (\bar{Q}_1 \bar{Q}_2)$, but $(Q_1 Q_2) \neq (\bar{Q}_1 \bar{Q}_2)$.

(b) Let $l(Q_1 Q_2)$ denote the length of $(Q_1 Q_2)$.

**Lemma 5.** Let $\Psi \in \Psi_t^{\min}$ and let $(Q_1 Q_2)$ be a maximum line segment of $\Psi$ such that

(a) $\{(Q_1 Q_2) \setminus (\{Q_1\} \cup \{Q_2\})\} \cap \{k, g^d_t(k)\}_{k \geq 0} = \{P\} \in P^g$.

(b) $\{(Q_1 Q_2) \setminus P\} \cap \{P^{\text{future}} \cup P^{\text{present}}\} = \emptyset$.

Then $l(Q_1 P) = l(PQ_2)$.

**Figure 2-11:** To the proof of Lemma 5.

**Proof.** For each $\Psi \in \Psi_t^S$, let $M_\Psi := \int_0^\infty x^2 d\Pi(x)$, where $\Pi$ has a $\Psi$-transform $\Psi$. From (a) and (b) it follows that the segment $(Q_1 Q_2)$ can be perturbed a little by rotating it around $P$ (Figure 2-11, left.) Consequently, the partial derivative of $M_\Psi$ with respect to the angle $\alpha$ must be zero for the observed value of $\alpha$, if $\Psi \in \Psi_t^{\min}$.

Let $d$ designate the segment $(Q_1 Q_2)$, $d_1$ the segment of $\Psi$ coming to $Q_1$ from the left, and $d_2$ the segment going out of $Q_2$ to the right. Let also $s$, $s_1$ and $s_2$ designate
the corresponding slopes. Then we must have that \( \frac{\partial M_2}{\partial s}(s_1, s_2, s) = 0 \). Now

\[
\frac{\partial M_2}{\partial s} = \frac{\partial}{\partial s} \left[ (s - s_1)x_1^2 + (s_2 - s)x_2^2 \right] = x_1^2 + (s - s_1)2x_1 \frac{\partial x_1}{\partial s} + (s_2 - s)2x_2 \frac{\partial x_2}{\partial s} - x_2^2.
\]

Let us find \( \frac{\partial x_1}{\partial s} \) and \( \frac{\partial x_2}{\partial s} \). From Figure 2-11 (right) we see that \((x - x_1)s = (x - x_1 - a)s_1\), \(a\) remains constant, when \((Q_1Q_2)\) is perturbed. Thus

\[
-\frac{\partial x_1}{\partial s} + (x - x_1) = -\frac{\partial x_1}{\partial s} s_1 \implies \frac{\partial x_1}{\partial s} = \frac{x_1 - x}{s_1 - s}.
\]

In the same way we obtain that \( \frac{\partial x_2}{\partial s} = \frac{x_2 - x}{s_2 - s} \). Hence

\[
\frac{\partial M_2}{\partial s} = x_1^2 + (s - s_1)2x_1 \frac{x_1 - x}{s_1 - s} + (s_2 - s)2x_2 \frac{x_2 - x}{s_2 - s} - x_2^2
\]

\[
= x_1^2 - 2x_1(x_1 - x) + 2x_2(x_2 - x) - x_2^2
\]

\[
= x_2^2 - 2xx_2 + (2xx_1 - x_2^2) = 0
\]

Solving for \(x_2\), we find \(x_2 = x \pm \sqrt{x^2 - 2xx_1 + x_1^2} = x \pm |x - x_1| = x_1; 2x - x_1\). Since \(x_2 \neq x_1\), then \(x_2 = 2x - x_1\) or \(x - x_1 = x_2 - x\). \(\square\)

**Definition 7.** (a) Let \(U = \{(x, y) \in \mathbb{R}^2^+ \mid g^d(x) \leq y(x) \leq g^u(x)\}\),

(b) For a continuous piecewise linear function \(u : \mathbb{R}^+ \rightarrow \mathbb{R}^+\), which changes its derivative in a finite number of points \(k_1, k_2, \ldots, k_n\), let

\[
\Phi_a(u) = \sum_{k_i \geq a} (u'(k_i^+) - u'(k_i^-))k_i^2.
\]

We are going to propose an algorithm for the single maturity case first, since in that case the algorithm is much simpler and an exact lower bound can be found. For the case, when there are also options of higher maturities, an algorithm is slightly harder, but an exact lower bound can be found as well. For the case of past points we propose an approximation algorithm, which converges to the optimal bound as the precision increases.

I. Present Prices.

**Definition 8.** For \(\{P_0, P_1, \ldots, P_n\} \in \mathcal{P}^{\text{present}}\) let for \(i = 1, \ldots, n - 1\)

- \(d_i = \{P_i + t(P_i - P_{i-1}) \mid t \in \mathbb{R}^+\} \cap U\),
- \(d_0 = \{P_0 + t(1, -1) \mid t \in \mathbb{R}^+\} \cap U\),

50
\[ B(d_i) = P_i; \quad E(d_i) = \{ \{ P^* \} \in d_i | X(P^*) \geq X(P) \forall P \in d_i \} \]

\[ \text{Int}(d_i) = d_i \setminus (\{ B(d_i) \} \cup \{ E(d_i) \}) \]

- \( s_i \) designate the slope of \( d_i \).
- For \( i = 0, \ldots, n - 1 \), let

\[ \xi_i := \min_{\Psi} \left( \Phi_{X(P_i)}(\Psi) \right| \Psi \in \Psi^S, \frac{d\Psi}{dt}(X(P_i)-) = s_i \), \quad (2.22) \]

\[ \Psi_{\xi_i} = \arg\min_{\Psi} \left( \Phi_{X(P_i)}(\Psi) \right| \Psi \in \Psi^S, \frac{d\Psi}{dt}(X(P_i)-) = s_i \right. \).

Let \( f_{\xi_i} : [X(P_i), +\infty) \to [0, Y(P_i)] \), be the restriction of \( \Psi_{\xi_i} \) to the interval \([X(P_i), +\infty), \) that is \( f_{\xi_i} = \Psi_{\xi_i} \) on \([X(P_i), +\infty) \). Then

\[ \xi_i = (f'_{\xi_i}(X(P_i)+) - s_i) X(P_i)^2 + \sum_{i=1}^{n} \left( f'_{\xi_i}(k_i+) - f'_{\xi_i}(k_i-) \right) k_i^2 \), \quad (2.23) \]

where \( X(P_i) < k_1 < k_2 \cdots < k_n \) are points, where \( f_{\xi} \) changes its derivative.

Notice, that \( \xi_0 = \min_{\Psi \in \Psi^S} M_{\Psi} \).

**Algorithm for the single maturity case.**

We will first show how to determine \( \xi_{n-1} \), and then \( \xi_i \), if \( \xi_{i+1}, \xi_{i+2}, \ldots, \xi_{n-1} \) are known. Notice, that \( \xi_i \) is fully determined by \( f_{\xi_i} \).

![Figure 2-12: Construction of \( f_{\xi_{n-1}} \).](image)

Construction of \( f_{\xi_{n-1}} \) (see Figure 2-12): draw a line \( l_{n-1} \) parallel to the line segment \( d_{n-1} \) to the right of \( P \), so that this line is at the same distance from \( P_n \) as the line containing \( d_{n-1} \). Two cases are possible:
\* \(l_{n-1}\) does not intersect \(d_n\) (Figure 2-12, left.) Denote by \(A_n\) the point of the intersection of \(l\) and the axis \(k\). Draw a line through \(A_n\) and \(P_n\). This line must intersect \(d_{n-1}\) at some point \(A_{n-1}\). Then \(f_{\xi_{n-1}}\) consists of line segments \((P_{n-1}, A_{n-1}), (A_{n-1}, A_n)\) and \((A_n, (+\infty, 0))\).

\* \(l_{n-1}\) intersects \(d_n\) (Figure 2-12, right.) Denote by \(B\) the point of the intersection of \(d_n\) with axis \(k\). Then \(f_{\xi_{n-1}}\) is given by line segments \((P_{n-1}, B)\) and \((B, (+\infty, 0))\).

Once, \(f_{\xi_{n-1}}\) is constructed, it is easy to calculate \(\xi_{n-1}\).

Now we want to determine \(\xi_i\), if \(\xi_{i+1}, \ldots, \xi_{n-1}\) are known. There will be three candidates to the optimal solution \(f^1_{\xi_i}, f^2_{\xi_i}\) and \(f^3_{\xi_i}\), among which we will choose the one with the smallest \(\xi_i\). (See Figure 2-14, for example.)

1. \(f^1_{\xi_i}(k)\) is given by

\[
f^1_{\xi_i}(k) = \begin{cases} P_i P_{i+1}(k), & k \in [X(P_i), X(P_{i+1})] \\ f_{\xi_{i+1}}(k), & k \in [X(P_{i+1}), (+\infty, 0)]. \end{cases}
\] (2.24)

2. \(f^2_{\xi_i}(k)\) is given by

\[
f^2_{\xi_i}(k) = \begin{cases} B(d_i) E(d_i)(k), & k \in [X(B(d_i)), X(E(d_i))], \\ E(d_i) P_{i+2}(k), & k \in (X(E(d_i)), X(P_{i+2})], \\ f_{\xi_{i+2}}(k), & k \in [X(P_{i+2}), (+\infty, 0)]. \end{cases}
\] (2.25)

3. \(f^3_{\xi_i}\) will be constructed as follows. Let \(l_i\) be the line containing \(d_i\). Draw the line \(l_{i+1}\) parallel to \(l_i\) to the right of \(P_{i+1}\) at the same distance from \(P_{i+1}\) as \(d_i\). Two cases are possible:

(a) \(P_{i+2}\) is to the left of \(l_{i+1}\) (see Figure 2-13.) Then \(f^3_{\xi_i} = f^2_{\xi_i}\).

(b) \(P_{i+2}\) is to the right of \(l_{i+1}\). Then repeat the procedure in 3, taking \(i + 1\) instead of \(i\) (see Figure 2-14.) We stop when either

\* there is a point \(P_{k+2}\) to the left of \(l_{k+1}\) for some \(k > i\) (Figure 2-15.) In that case \(f^3_{\xi_i} = f_{\xi_{k+2}}\) on \([X(P_{k+2}), +\infty]\). Thus we have to define \(f^3_{\xi_i}\) only on \([X(P_i), X(P_{k+2})]\). Let the point \(A_k\) be the intersection of \(l_k\) with the function \(g^d(k)\) (Figure 2-15.) Define \(f^3_{\xi_i}(k) = g^d(k)\) on
[\{X(A_{k}), X(P_{k+2})\}]. Draw a line \(L_k\) through \(A_k\) and \(P_k\). This line must intersect \(l_{k-1}\) at some point \(A_{k-1}\). Then the line segment \((A_{k-1}, A_k)\) defines the function \(f^3_{\xi_i}\) on the interval \([X(A_{k-1}), X(A_k)]\). Continue like that till you reach \(A_i \in l_i\).

- there are no more points to the right of \(l_{k+1}\) (Figure 2-14.) Then it means that \(k+1 = n\) and we have to define \(f^3_{\xi_i}\) on \([X(P_i), +\infty]\). Let \(A_n\) be the intersection of \(l_n\) and the function \(g^d(k)\). Define \(f^3_{\xi_i}(k) = g^d(k)\).
on \([X(A_n), +\infty]\). Draw a line \(l_n\) through \(A_n\) and \(P_n\). It must intersect \(l_{n-1}\) at some point \(A_{n-1}\). Then the line segment \((A_{n-1}, A_n)\) defines the function \(f_{\xi_i}^3\) on the interval \([X(A_{n-1}), X(A_n)]\). Continue like that till you reach \(A_i \in l_i\).

Now, once \(f_{\xi_i}\) is constructed, calculate \(\xi_i\).

**Proof.** We want to show, that this algorithm leads to an optimal solution. By Corollary 3, \(\Psi^{\text{min}}(k)\) consists of no more than \(n + 1\) line segments. Moreover, each line segment with a nonzero slope must go through one of the constraint points. If the first (from the left) line segment of \(\Psi^{\text{min}}\) goes only through \(P_0\), but not \(P_1\), then this segment must lie on \(d_0\) (see Figure 2-16.) Indeed, otherwise, the function could be strictly decreased. Thus the next segment must depart from \(d_0\) at some point \(A_0\).

If \(A_0 \in \text{Int}(l_0)\), then by Lemma 5, this segment must be split by \(P_1\) into two equal parts. It means that the end \(A_1\) of that segment must lie on the line \(l_1\) parallel to \(d_0\) at the same distance from \(P_1\) as \(d_0\). However if \(P_2\) is to the left of \(l_1\), no such line segment would exist. The part of the segment to the right of \(P_1\) would always be shorter, than its part to the left, and in order to decrease \(M_\psi\), we would have to turn this segment in the clockwise direction till we hit \(P_2\).

If \(P_2\) is to the right of \(l_1\), then for some starting points \(A_0 \in d_0\), there exist line segments with required properties and ends \(A_1 \in U \cap l_1\). To continue the proof, we would have to apply the same arguments to segments, departing from \(U \cap l_1\), as were applied to segments departing from \(d_0\). Lemma 5 does not apply if \(A_0 = B(d_0)\) or \(A_0 = E(d_0)\). If \(A_0 = B(d_0)\), then we know that the first line segment of the function goes through \(P_0\) and \(P_1\), and the next departs from \(d_1\). The same arguments apply, as when we started with \(d_0\). If \(A_0 = E(d_0)\), then we know that the first line segment of the function is \((P_0, E(d_0))\), while the next goes through \(P_1\) and \(P_2\). The third segment would have to depart from \(d_2\), and the same arguments would apply, as when we had to depart from \(d_0\).

The running time of the algorithm is \(O(n^2)\), if \(n\) is the number of present prices.

**II. Future prices.**

From now on we will assume that \(\mathcal{P}^{\text{future}}\) contains only those future points, which are strictly below the border of the convex hull of \(\mathcal{P}^{\text{present}} \cup \{(0, +\infty)\} \cup \{(+\infty, 0)\}\).
Figure 2-16: To the proof of the algorithm for the single maturity case.

Future points which are on the border of of that convex hull do not influence the optimal solution, but might introduce ambiguity.

By Corollary 3, in the absence of past constraint points, each line segment of $\Psi_{t_{min}}$ with a nonzero slope must contain $P \in \mathcal{P}_{\text{present}}$. Thus the optimal solution (the lower bound on the variance) will not change if we reduce $\mathcal{P}_{\text{future}}$ to $\mathcal{P}_{\text{future}}$.

$$
\mathcal{P}_{\text{future}} := \mathcal{P}_{\text{future}} \cap \mathcal{P}_{\text{u}}, \text{ where }
\mathcal{P}_{\text{u}} := \left\{ p_{t+j} \in \mathcal{P}_{\text{u}} \, \mid \, \left( p_{t+k-1} \vee p_{t+k} \vee p_{t+k+1} \right) \in \mathcal{P}_{\text{present}} \right\}. \tag{2.26}
$$

that is, $\mathcal{P}_{\text{u}}$ consists of points of $\mathcal{P}_{\text{u}}$, such that either the point itself or one of its closest (from the left or from the right) neighbors is a present point. Also, let us define for present points $\{P_1, P_2, \ldots, P_{n_{\text{present}}}\} = \mathcal{P}_{\text{present}}$

$$
U_{P_i}^\text{left} = \left\{ P \in U \mid X(P_{i-1}) \leq X(P) \leq X(P_i), (P, P_i) \in U \right\},
U_{P_i}^\text{right} = \left\{ P \in U \mid X(P_i) \leq X(P) \leq X(P_{i+1}), (P, P_i) \in U \right\}. \tag{2.27}
$$

See Figure 2-17 below.

Notice, that we can reduce the domain of $\{(k, \Psi_{\text{min}}(k))\}_{k \geq 0} \subset U \subset \mathbb{R}^2$ to

$$
\hat{U} = \bigcup_{P_i \in \mathcal{P}_{\text{present}}} \left( U_{P_i}^\text{left} \cup U_{P_i}^\text{right} \right).
$$

without influencing the optimal solution (the lower bound.)

**Definition 9.** • Let $\mathcal{H} : \{0, 1, 2, \ldots, n\} \rightarrow \{1, 2, \ldots, n\}$ be the function that maps $i \in \{0, 1, 2, \ldots, n\}$ into $i^* \in \{1, \ldots, n\}$, such that $P_{i^*} \in \mathcal{P}_{\text{u}}$ is the present point which is the closest to $P_{i^*}$ from the right. If there is no present point to the right of $P_{i^*}$, then we can put $\mathcal{H}(i) = i$. Let $\mathcal{H}^2(i) := \mathcal{H}(\mathcal{H}(i))$ and by
Figure 2-17: To the definition of $U^\text{left}_P$ and $U^\text{right}_P$. Here $\{P_1, P_2, P_5, P_6\} \in \mathcal{P}^{\text{present}}$, while $\{P_3, P_4\} \in \mathcal{P}^u \cap \mathcal{P}^{\text{future}}$.

induction define $\mathcal{H}^k(i)$.

- Let $d_0 = \{P_0 + t(1, -1) \mid t \in \mathbb{R}^+\} \cap U^\text{left}_{\mathcal{H}(1)}$. For $\{P_1^u, P_2^u, \ldots, P_n^u\}$ define for each pair of consecutive points $(P^u_{i-1}, P^u_i)$, such that $P^u_i \lor P^u_{i-1} \in \mathcal{P}^{\text{present}}$,
  $$d_i = \{P^u_i + t(P^u_i - P^u_{i-1}) \mid t \in \mathbb{R}^+\} \cap U^\text{left}_{\mathcal{H}(i)}.$$
  $$B(d_i) = (\{P^*\} \in d_i | X(P^*) \leq X(P) \forall P \in d_i),$$
  $$E(d_i) = (\{P^*\} \in d_i | X(P^*) \geq X(P) \forall P \in d_i),$$
  $$Int(d_i) = d_i \setminus (\{B(d_i)\} \cup \{E(d_i)\}).$$

- Define $s_i, \xi_i, f_{\xi_i}$ for $i = 0, 1, 2, \ldots, n_u - 1$ as they were defined in Definition 8, but taking $B(d_i)$ instead of $P_i$.

Then, as in the case of a single maturity, we will have to find the sequence of $\{\xi_i\}$ by a backward recursion. However, we don’t need to know $\xi_i$, if $d_i$ corresponds to two consecutive future points, since we know that each line segment of $\Psi^{\text{min}}$ must contain at least one present point.

**Algorithm for the case of present and future prices.**

First we have to determine $\xi_{n-1}$. If the last point in $\mathcal{P}^u$ is a present point, then $\xi_{n-1}$ must be found in exactly the same way as in the case of a single maturity. If the last point of $\mathcal{P}^u$ is a future point, then $\xi_{n-1}$ does not have to be determined, since $\Psi^{\text{min}}$ can not have a line segment, which does not go through at least one of the present points. In that case we have to determine $\xi_{n-2}$ instead of $\xi_{n-1}$. Draw the line $l_{n-1}$ parallel to $d_{n-2}$ to the right of $P^u_{n-1}$ at the same distance from that point as the line
containing $d_{n-2}$ (Figure 2-18). Let

\[
\{(a, 0)\} := \left\{ P_{n-2}^u t + P_{n-1}^u (1 - t) \right\}_{t \in \mathbb{R}} \cap \{(k, 0)\}_{k \geq 0}, \quad \{(b, 0)\} := \left\{ P_{n-1}^u t + P_n^u (1 - t) \right\}_{t \in \mathbb{R}} \cap \{(k, 0)\}_{k \geq 0},
\]

\[
\bar{c} := \begin{cases} 
  c, & c \in [a, b], \\
  a, & c < a, \\
  b, & c > b.
\end{cases}
\]

Draw a line through $P_{n-1}^u$ and $\{(\bar{c}, 0)\}$. This line must intersect $d_{n-1}$ at some point $A$. Then $f_{\xi_{n-2}}$ consists of line segments $(P_{n-2}^u, A, (\bar{c}, 0))$ and $((\bar{c}, 0), (+\infty, 0))$.

\[
\begin{tikzpicture}
  \node[anchor=north west] at (0,0) {Figure 2-18: Construction of $f_{\xi_{n-2}}$. Here $P_{n-1}^u$ is a present point, $P_n^u$ is a future point, and $P_{n-2}^u \in \mathcal{P}^u$.

Now we have to determine $\xi_i$ if $\xi_{i+1}, \xi_{i+2}, \ldots \xi_{n-1}$ are given. We are going to find a finite number of candidates to $f_{\xi_i}$, among which we will choose the one with the smallest $\xi_i$. The procedure is recursive. At each step we add at most two candidates to the optimal $f_{\xi_i}$, and the number of steps is at most the number of present points to the right of $P_i^u$.

step 1. Let $l_i = d_i$. Notice that $P_{\mathcal{H}(i)}^u$ is the present point closest to $P_i^u$ from the right and $B(d_i)$ always lies on the same line as $d_{\mathcal{H}(i)}$, while $E(d_i)$ on the same line as $d_{\mathcal{H}(i)+1}$. See Figure 2-19.

Because of this property of $d_i$, there are always two candidates to $f_{\xi_i}$ to be added at the first step. The first candidate is

\[
f_{\xi_i}^{11}(k) = \begin{cases} 
  B(d_i)B(d_{\mathcal{H}(i)}) (k), & k \in [X(B(d_i)), X(B(d_{\mathcal{H}(i)}))] \\
  f_{\mathcal{H}(i)+1}^{\xi_i}(k), & k \in [X(B(d_{\mathcal{H}(i)})), (+\infty, 0)].
\end{cases}
\] (2.28)
Figure 2-19: Determining $\xi_i$ in the case when there are future prices. Here $P^u_{i+1}$ and $P^u_{i+2}$ are future prices, $P^u_{\mathcal{H}(i)}$ and $P^u_{\mathcal{H}^2(i)}$ are present prices, while $P^u_i$ might be a future or present point.

The second candidate is

$$ f_{\xi_i}^{12}(k) = \begin{cases} 
\frac{B(d_i)E(d_i)(k)}{E(d_i)B(d_{\mathcal{H}(i)+1})(k)}, & k \in [X(B(d_i)), X(E(d_i))], \\
\frac{E(d_i)B(d_{\mathcal{H}(i)+1})(k)}{E(d_i)B(d_{\mathcal{H}(i)+1})(k)}, & k \in (X(E(d_i)), X(B(d_{\mathcal{H}(i)+1}))), \\
f_{\xi_{\mathcal{H}(i)+1}}(k), & k \in [B(X(d_{\mathcal{H}(i)+1})), (+\infty, 0)]. 
\end{cases} \quad (2.29) $$

**step 2.** Let $\tilde{l}_2$ be the symmetric reflection of $l_1$ with respect to $P^u_{\mathcal{H}(i)}$, that is

$$ \tilde{l}_2 := \{ P + 2 \cdot ((P^u_{\mathcal{H}(i)}) - P) \mid P \in l_1 \}.$$

Let

$$ l_2 = \tilde{l}_2 \cap U^\text{right}_{P^u_{\mathcal{H}(i)}} \cap U^\text{left}_{P^u_{\mathcal{H}^2(i)}}. $$

If $l_2 = \emptyset$, then there are no more candidates to the optimal solution.

If $l_2 \neq \emptyset$, then check if $B(l_2)$ is on the line containing $d_{\mathcal{H}^2(i)}$. If this is the case, then the first candidate to the optimal solution to be added at this step is $f_{\xi_i}^{21}$, such that on $[X(B(l_2)), (+\infty, 0)]$

$$ f_{\xi_i}^{21}(k) = \begin{cases} 
\frac{B(l_2)B(d_{\mathcal{H}^2(i)})(k)}{B(l_2)B(d_{\mathcal{H}^2(i)})(k)}, & k \in [X(B(l_2)), X(B(d_{\mathcal{H}^2(i)}))], \\
f_{\xi_{\mathcal{H}^2(i)}}(k), & k \in [B(X(d_{\mathcal{H}^2(i)})), (+\infty, 0)]. 
\end{cases} \quad (2.30) $$

To define $f_{\xi_i}^{21}$ on $[X(B(l_1)), X(B(l_2))],$ draw a halfline, starting at $B(l_2)$ and going through $\mathcal{H}^1(P_i)$. Find the intersection of this line with $l_1$. 

58
The second candidate will exist if $E(l_2)$ is on the line containing $d_{\mathcal{H}^2(i)+1}$. It can be defined in the same manner as $f^{21}_{\xi_1}$.

last step: Continue like that till we reach $l_k = \emptyset$ for some $k$ or there are no more present points to the right of $l_m$. If there are no more points to the right of $l_m$ and $l_m$ does not intersect the axis $k$, then there are no more candidates. If $(0, a)$ is the intersection of $l_m$ with the axis, then define $f^{m1}_{\xi_1}(k)$ to be 0 at $[a, +\infty)$ and to build $f^{m1}_{\xi_1}(k)$ on $[B(l_1), a]$, start moving from $(a, 0)$ by a straight line in the direction of $P^u_{\mathcal{H}^m(i)}$ (the closest present point from the left) till you reach $l_{m-1}$. Take the intersection point and move by a straight line in the direction of $P^u_{\mathcal{H}^{m-1}(i)}$ till you reach $l_{m-2}$ and so on, till you reach $l_1$.

The running time of the algorithm is $O(n^2_{\text{present}})$.

Proof. The proof is very similar to the proof of the algorithm for the single maturity case. By Corollary 3 each line segment of $\Psi^{\text{min}}$ with a nonzero slope must contain a present point. By Lemma 5 that segment must be split by that present point into two equal parts, unless it contains some other present or future points.

![Figure 2-20: To the proof of the algorithm for future prices. Here $P^u_0$, $P^u_2$, $P^u_4$ are present prices, while $P^u_1$ and $P^u_3$ are future prices.](image-url)

If the first (from the left) line segment of $\Psi^{\text{min}}$ goes only through $P^u_0$, but not $P^u_1$, then this segment must contain $d_0$ (see Figure 2-20.) Indeed, otherwise, the function could be strictly decreased. In this case the next segment must depart from $d_0$ at some point $A_0$. Even if $B(d_0) \neq P_0$, this next segment can not depart from $(P_0, B(d_0))$, since then it would not be able to reach $P^u_{\mathcal{H}(0)}$ without violating constraints.

If $A_0 \in \text{Int}(d_0)$, then this segment must be split by $P^u_{\mathcal{H}(i)}$ into two equal parts. Thus, the end of that segment must lie on the symmetric reflection of $d_0$ with respect to $P^u_{\mathcal{H}(0)}$. We call it $\tilde{l}_1$. However if $\tilde{l}_1$ has an empty intersection with $U^\text{right}_{P^u_{\mathcal{H}(0)}}$, then this
segment violates some of the constraints and thus $A_0 \notin \text{Int}(d_0)$. If $\tilde{t}_1 \cap U^\text{right}_{P^u_{\mathcal{H}(0)}} \neq \emptyset$, but $\tilde{t}_1 \cap U^\text{left}_{P^u_{\mathcal{H}(0)}} = \emptyset$, then there is no line segment that starts at $\tilde{t}_1$ and goes through $P^u_{\mathcal{H}(0)}$ (the next present point). Thus, again, it is not possible that $A_0 \in \text{Int}(d_0)$.

If $\tilde{t}_1 \cap U^\text{right}_{P^u_{\mathcal{H}(0)}} \cap U^\text{left}_{P^u_{\mathcal{H}(0)}} = l_1 \neq \emptyset$ then we should apply the same arguments to $l_1$, as were applied to $d_0$.

Now, if $A_0 = B(d_0)$, then $A_0$ is on the line containing $d_{\mathcal{H}(0)}$ as well as $P^u_{\mathcal{H}(0)}$. Thus the function will depart from the segment that goes through $A_0$ and $P^u_{\mathcal{H}(0)}$ at some point of $d_{\mathcal{H}(0)}$. If $A_0 = E(d_0)$, then $A_0$ is on the line containing $d_{\mathcal{H}(0)+1}$ as well as $P^u_{\mathcal{H}(0)}$. Thus in this case the function will depart from the segment that goes through $A_0$ and $P^u_{\mathcal{H}(0)}$ at some point of $d_{\mathcal{H}(0)+1}$. Continue the proof by induction. $\square$

We will formulate one additional Lemma, which can reduce the number of candidates to the optimal solution even more in the case future prices are present. It holds if past prices are also in consideration.

**Lemma 6.** Let $\Psi \in \Psi^{\min}$ and $(Q_1Q_2)$ be a maximum line segment of $\Psi$ such that

(a) $(Q_1Q_2) \setminus \{Q_1 \cup \{Q_2\}\} \cap \{(k, g^d(k))\}_{k \geq 0} = \{P_1\} \in \mathcal{P}^g$.

(b) $(Q_1Q_2) \setminus P_1 \cap \mathcal{P}^{\text{present}} = \emptyset$;

(c) $(Q_1Q_2) \cap \mathcal{P}^{\text{future}} = \{P^f\}$.

Then $P^f$ lies in the one of the two segments $(Q_1P_1)$ and $(P_1Q_2)$ which is shorter than the other.

![Figure 2-21: To the proof of Lemma 6.](image)

**Proof.** Notice, that when we turn $(Q_1Q_2)$ around $P_1$ in the clockwise direction (Figure 2-21), the length of $(P_1Q_1)$ increases while the length of $(P_1Q_2)$ decreases, since the function is nonincreasing. Assume that $l(Q_1P_1) > l(P_1Q_2)$ and $\{P^f\} \in (Q_1P_1)$. Then turning $(Q_1Q_2)$ around $P$ in the counterclockwise direction will reduce $l(Q_1P_1)$ and increase $l(P_1Q_2)$, thus reducing $\mathcal{M}_\Psi$ according to Lemma 5. But then $\{P^f\}$ will not belong to $(Q_1Q_2)$. In the same way we can show that if we assume that $\{P^f\} \in (P_1Q_2)$ and $(P_1Q_2)$ is longer than $(Q_1P_1)$, we come to a contradiction. $\square$
III. Past prices. If \( \{(k, g^d(k))\}_{k \geq 0} \) contains past price constraint points, then the problem becomes much more complicated because of the following reasons.

(a) A line segment of \( g^d(k) \) does not necessarily lie on the line that goes through some present price (it is always the case when there are not past prices.)

(b) As stated by Lemma 4, there might be segments \( \Psi_i^{\min} \) that go only through past points, but not present. Thus in the case that we have past prices on \( g^d(k) \) as picks, much more candidates to the optimal solution arise. Indeed, here we have to choose through which of the past constraint points the function will follow and which it will miss. Thus, all such combinations must be considered. The time for the solution would grow exponentially with the number of past constraint points.

The problem can be resolved in the following way.

**Definition 10.** Let \( \mathcal{P}^{\text{present}} = \{P_1, P_2, \ldots, P_n\} \) and \( k_j := X(P_j), j = 1, \ldots, n \). Designate

\[
\alpha_j^- = \min_{\psi} \left( \Psi^j(k_j^-) | \psi \in \Psi^S_i \right) \quad \text{and} \quad \alpha_j^+ = \max_{\psi} \left( \Psi^j(k_j^-) | \psi \in \Psi^S_i \right).
\]

For each \( j = 2, \ldots, n \) define the function \( \xi_j(\alpha) : [\alpha_j^-, \alpha_j^+] \rightarrow \mathbb{R}^+ \):

\[
\xi_j(\alpha) := \min_{\psi} \left( \Phi_{k_j}(\psi) \bigg| \psi \in \Psi^S_i, \frac{d\Psi}{dk}(k_j) = \alpha \right).
\]  

(2.31)

**Algorithm for the multiple maturities case.** Let us show how to determine \( \xi_j(\alpha) \), if \( \xi_{j+1}(\alpha), \ldots, \xi_n(\alpha) \) are known.

![Figure 2-22:](image)

- Let \( R_j := \{P \mid P \in \mathcal{P}^{\text{past}} \cap \mathcal{P}^S, X(P_j) < X(P) < X(P_{j+1})\} \). Select a subset \( R_{j1} \) of \( R_j \).
• Temporarily modify $g^u(k)$ and $g^d(k)$ within the interval $[X(P_j), X(P_{j+1})]$, as if $R_{j,1} \in P_{\text{present}}$.

• Choose a point $Q_0$ on the line, which goes through $P_j$ under an angle $\alpha_0 \in [\alpha^{-}_j, \alpha^{+}_j]$, to the left of $P_j$ (Figure 2-22.) Assume that the function $\Psi$ changes its derivative last time before $P_j$ at this point.

• Construct a piecewise linear function $f^\alpha_{j0}(k)$ on $[X(Q_0), X(P_{j+1})]$, so that it starts at $Q_0$, goes through all the points of $R_{j1}$ and enters $P_{j+1}$ under an angle $\alpha \in [\alpha^{-}_{j+1}, \alpha^{+}_{j+1}]$. Each line segment of $f^\alpha_{j0}$ must either have a nonempty line segment overlap with $g^d(k)$, or contain a point of $\{P_j \cup R_{j1}\}$. If the last is true, this segment must be split by that point into two equal parts, unless it has a nonempty line segment overlap with $g^u(k)$.

Suppose that $Q_i$ is either $Q_0$ or the end of the last constructed segment and we want to construct the next segment, $(Q_i, Q_{i+1})$. Start moving from $Q_i$ by the straight line in the direction of the closest from the right point of $\{P_j \cup R_{j1}\}$, call it $P^*$.

The following cases might occur:

(a) We hit $g^d(k)$ before we reach $Q_{i+1}$, such that $(Q_i, Q_{i+1})$ is split by $P^*$ into two equal parts, while $(Q_i, P^*)$ is not on the line containing a nonempty line segment overlap with $g^u(k)$. Then this solution must be disregarded and another initial point $Q_0$ on the line with the angle $\alpha_0$ be chosen.

(b) $(Q_i, P^*)$ is on the line containing a nonempty line segment overlap with $g^u(k)$. Then after passing $P^*$ we have to introduce a new variable, which is the distance that we follow that segment till we leave it. Here Lemma 6 should be taken into account.

(c) If $(Q_i, P^*)$ does not lie below the function $g^u(k)$, then that solution must be disregarded and a new point $Q_0$ chosen.

(d) If we reach $Q_{i+1}$, such that $l(Q_i, P^*) = l(P^*, Q_{i+1})$ exactly when we hit $g^d(k)$, then we have to follow the constraint of $g^d(k)$. This adds a new variable over which to optimize. That variable is the distance that we follow the constraint till we leave it.

(e) None of the above cases occurred and we reached $Q_{i+1}$, such that $l(Q_i, P^*) = l(P^*, Q_{i+1})$. 

62
• After we have passed all the points of $R_{j1}$ and stopped at the point such that the last segment is split into two equal parts by $P_{j+1}$, we can calculate $\xi_j(R_{j1}, \alpha_0, Q_0)$.

• Now $\xi_j(R_{ji}, \alpha_0) = \min_{Q_0} \xi_j(R_{ji}, \alpha_0, Q_0)$ and $\xi_j(\alpha_0) = \min_{R_{ji} \in R_j} \xi_j(R_{ji}, \alpha_0)$. If for some $\alpha_0 \in [\alpha_j^-, \alpha_j^+]$ no $f_j^{\alpha_0, Q_0}(k)$ with required properties exists, then assign $\xi_j(\alpha_0) = +\infty$.

For construction of $f_n^{\alpha_0, Q_0}(k)$ the same principles apply, except that if after passing all the points of $\{P_n \cup R_{n1}\}$, we found ourselves in case (e), then that solution must also be disregarded.

*Remark.* The set of all possible starting points $Q_0$ is determined by $\alpha_0$ as well as constraints $g^*(k)$ and $g^d(k)$. Then this set should be divided into nonintersecting intervals, the number of which depends on the desired precision of the solution.

If $n$ is the number of present points and $k_i$ is the number of past points which lie on $g^d(k)$ and are inbetween the two present points $P_i$ and $P_{i+1}$, or more precisely, the number of past points which belong to $\{(k, g_d(k)) \mid X(P_i) \leq k \leq X(P_{i+1})\}$, then the running time of the algorithm is $O(\sum_{i=1}^n 2^{k_i})$.

Let us give an example demonstrating how the optimal $\Psi_{\min}$ (Figure 2-23) is constructed.

![Example Diagram](image)

*Figure 2-23: Example.*

Let $P_0$, $P_1$ and $P_2$ be the present points, $P^f$ the future point and $P^p$ the past point. We will consider the case, when the function goes through $P^p$ and the case when it does not.

(a) The function goes through $P^p$. We choose $Q_0$ on $d_0$, the initial segment with the slope -1, and draw a segment $(Q_0Q_1)$ through $P_1$, so that $P_1$ splits this segment into
two equal parts. From \( Q_1 \) we draw a new segment through \( P^p \). Suppose that segment goes below \( P^f \). Then that solution can not be optimal since the segment hits \( g^d(k) \) before it reaches the point such that \( P^p \) splits it into two equal parts. Thus that solution must be disregarded. In case the segment also goes through \( P^f \), by Lemma 2-21, the solution again can not be optimal, since the length of \((Q_1P^p)\), containing \( P^f \), is greater than the part of the segment to the right of \( P^p \).

A boundary case \( Q_0 = P_0 \) must be considered separately. In this case we have to choose a starting point on the segment \( (P_1, S) \). However if we choose a point above the line segment going through \( S \) and \( P^p \), the rule of equal length will not be satisfied, thus the solution will not be optimal. Consequently, we have to consider only the boundary case, that is, when we choose a point on the line \( SP^p \). But that line also goes through \( P_2 \).

The other boundary case, when \( Q_0 = E(d_0) \) can not be optimal by Lemma 2-21, since \((E(d_0), P_1)\) is shorter than \((P_1, P^f)\).

Consequently if we require that the function goes through the past price, we have only one candidate to the optimal solution. That is the function which consists of two line segments: \((P_0S)\) and the segment which goes through \( S \) and \( P_2 \).

(b) Let us consider what candidates to the optimal solution we will have in the case that the function does not go through \( P^p \). First, notice that again we have to consider only the boundary case \( Q_0 = P_0 \). Thus we have to choose point \( \tilde{Q}_1 \) on \((P_1S)\) which will lead to an optimal solution. Notice that that will actually include the candidate of (a). Here we can see that \((\tilde{Q}_1P_2)\) will always be shorter than the part of the same segment to the right of \( P_2 \). Thus in that case we can immediately say that the \( \Psi_{\min} \) is the function which consists of two segments: one of them contains \( P_0 \) and \( P_1 \) and the other contains \( P^f \) and \( P_2 \).

Remark 10. A question arises whether there is a more efficient algorithm for determining the lower bound if options of past maturities are in consideration.
Chapter 3

Option Pricing Without Price Dynamics, the Two Dimensional Case.

This chapter considers the problem of determining bounds on options on two different assets if prices of options on individual assets are known.

Let \( \{ S^1(t) \}_{t \geq 0} \) be a stochastic process describing the price process of the first asset, and \( \{ S^2(t) \}_{t \geq 0} \) that of the second asset. If \((k_{ij}^h, C_{ij}^h, t_i), i = 1, \ldots, n, j = 1, 2, \ldots, U(i, h),\) represent the set of options on asset \( S^h, \) with \( t_i \) being the maturity of an option, \( k_{ij}^h \) its strike price, and \( C_{ij}^h \) its price, then we must have

\[
C_{ij}^h = e^{-\int_0^{t_i} r(s) ds} E_Q \left[ (S^h(t_i) - k_{ij}^h)^+ \right] = E_Q \left[ \left( X_i^h - k_{ij}^h e^{-\int_0^{t_i} r(s) ds} \right)^+ \right], \tag{3.1}
\]

where \( \{ X_i^h \}_{t \geq 0} \) is a martingale under \( Q. \)

\textit{Remark 11.} As in Section 2.1, without loss of generality, the risk-free rate \( r(s) \) can be assumed 0. Thus we can change times \( t_i \) to their indexes.

Let \( \mathcal{T} := \{(h, t, j)| h \in \{1, 2\}, t \in \{1, 2, \ldots, n\}, j = 1, 2, \ldots, U(t, h)\}. \) Define a family of measurable functions \( f_{ij}^h : \mathbb{R}^+ \to \mathbb{R}^+ \) as

\[
f_{ij}^h(X_t) = (X_i^h - k_{ij}^h)^+, \quad (h, t, j) \in \mathcal{T}. \tag{3.2}
\]

Then the problem can be formulated as follows:
Given the set \((k^h_{tj}, C^h_{tj}), (h, t, j) \in \mathcal{T}\), and a continuous piecewise linear function \(g : \mathbb{R}^2 \to \mathbb{R}\), find the maximal and minimal possible values of \(E[g(X_{t^*})]\), for a given \(t^* \in \{1, 2, \ldots, n\}\), under the condition that \(X_1, X_2, \ldots, X_n\) is a two-dimensional nonnegative martingale, such that

\[
E \left[ f^h_{ij}(X_t) \right] = C^h_{tj}, \quad (h, t, j) \in \mathcal{T}
\]  

(3.3)

**Definition 11.** We call function \(g : \mathbb{R}^d \to \mathbb{R}\) piecewise-linear, if there exists a partition of \(\mathbb{R}^d\) by a finite number of hyperplanes (lines if \(d = 2\)) into nonoverlapping subsets (regions if \(d = 2\)), such that in the interior of each subset \(g\) is an affine function.

The idea of our approach is to reduce the class of martingale distributions over which we maximize (or minimize) as much as possible, while making sure that the optimal solution does not change. Thus in Subsection 3.1.1 we show that it is enough to consider martingales with the Markov property only. In Subsection 3.1.2 we prove that conditions (3.3) for \(t > t^*\) can be reduced to separate conditions on the first and second coordinates of \(X_{t^*}\). Finally in Subsection 3.1.3 we show how to reduce the general uncountable distribution to the distribution on a finite set, while satisfying all the constraints (3.3) and achieving the same maximum (minimum) of \(E[g(X_{t^*})]\).

In Section 3.2 we propose an algorithm which leads to an exact solution. However, the number of variables in the linear optimization problem, which must be solved at the end, is proportional to the number of different \(k_{tj}\) with \(t \leq t^*\) to the power of \(t^*\). Thus, this method works only if \(t^*\) is not too large. Section 3.3 provides a computational example.

In Section 3.4 we discuss a second method that gives an approximate solution, which converges to the exact solution as the precision increases. \(g : \mathbb{R}^2 \to \mathbb{R}\) in this case can be any continuous function. The number of unknown variables, encountered in the linear optimization problem, grows linearly with \(t^*\) in this case. The algorithm is given in Subsection 3.4.1. In Subsection 3.4.2 we propose an improved approximation algorithm for the case when \(g(X_{t^*})\) is a piecewise linear function.

In Section 3.5 we notice, that the exact solution proposed for the two-dimensional case, leads to a qualitatively new solution in the one-dimensional case. While the
efficiency of the algorithm for the one-dimensional case proposed in Chapter 2 is higher, the new approach enables to solve efficiently a more general problem, where the payoff function \( g(X_{t^*}) \) is an arbitrary continuous piecewise linear function.

In Section 3.6 we notice that all approaches developed for the two-dimensional case are applicable to the multiple dimensional problem. However, the number of variables in the linear programming problem that has to be solved grows exponentially not only with \( t^* \), but with the dimensionality of the problem.

Finally in Section 3.7 we consider the problem of computing upper and lower bounds on the price of a European basket call option, given prices of other similar baskets with the same maturities. In mathematical terms:

given \( C \in \mathbb{R}^{m^+}, k \in \mathbb{R}^{n^+}, w_i \in \mathbb{R}^n, i = 1, \ldots, m \) and \( k_0 > 0, w_0 \in \mathbb{R}^{m^+} \), find the upper and lower bounds on

\[
E \left( w_0^T X - k_0 \right)^+, \tag{3.4}
\]

with respect to a distribution of an \( n \)-dimensional random variable \( X \) with finite expectation and support in \( \mathbb{R}^{n^+} \) under the condition

\[
E \left( w_i^T X - k_i \right)^+ = C_i, \quad i = 1, \ldots, m. \tag{3.5}
\]

The solution to this problem also immediately follows from the approach developed for the exact solution of the two-dimensional multiple maturities problem. Again, a linear programming problem has to be solved to obtain the bounds. The number of variables is bounded by \( \binom{2n+m}{n} \).

### 3.1 Structural theorems

In this section, we formulate and prove the theorems on which the algorithm will be based.

#### 3.1.1 Restriction to Markov martingales

**Proposition 7.** Let \( X_1, X_2, \ldots, X_n \) be a martingale with values in \( \mathbb{R}^d \) and let \( \mu_i \) be the marginal distribution of \( X_i \). Then, there exists a martingale with the Markov property \( \tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_n \), such that the marginal distribution of \( \tilde{X}_i \) is \( \mu_i \) for all \( i = 1, 2, \ldots, n \).
In the proof we will use an existence theorem for Markov processes due to Kolmogorov [23, p. 120]:

Fix a time scale $T$ starting at 0, a Borel space $(S, S)$, a probability measure $\nu$ on $S$ and a family of probability kernels $\mu_{s,t}$ on $S$, $s \leq t$ in $T$, satisfying

$$\mu_{s,u} = \mu_{s,t}\mu_{t,u}, \quad s \leq t \leq u.$$ 

Then there exists an $S$-valued Markov process $X$ on $T$ with initial distribution $\nu$ and transition kernels $\mu_{s,t}$.

**Proof.** For specifying the distribution of a discrete time Markov process, it is enough to specify the transition probabilities as well as the initial distribution.

Let $\nu = \mathcal{L}(X_1)$, $\mu_{1,2} = \mathcal{L}(X_2|X_1)$, $\ldots$, $\mu_{n-1,n} = \mathcal{L}(X_n|X_{n-1})$. Then by the existence theorem for Markov processes due to Kolmogorov there exists a Markov process $\tilde{X}_1, \ldots, \tilde{X}_n$ such that $\mathcal{L}(X_1) = \nu$, $\mathcal{L}(\tilde{X}_2|\tilde{X}_1) = \mu_{1,2}$, $\ldots$, $\mathcal{L}(\tilde{X}_{n-1}|\tilde{X}_{n}) = \mu_{n-1,n}$.

It remains to show that the so-defined process $\tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_n$ is also a Martingale. Let $\sigma_k$ be the smallest $\sigma$-algebra for which $\tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_k$ are all measurable. We need to prove that $E[\tilde{X}_{k+1}|\sigma_k] = \tilde{X}_k$. From the Markov property of $\tilde{X}_1, \ldots, \tilde{X}_k$, it follows that $E[\tilde{X}_{k+1}|\sigma_k] = E[\tilde{X}_{k+1}|\tilde{X}_k]$. By definition $E[\tilde{X}_{k+1}|\tilde{X}_k] = E[X_{k+1}|X_k]$.

On the other hand, since $X_1, X_2, \ldots, X_n$ is a martingale, we have

$$E[E[X_{k+1}|X_k, X_{k-1}, \ldots, X_1]|X_k] = E[X_k|X_k] = X_k.$$  

Also

$$E[E[X_{k+1}|X_k, X_{k-1}, \ldots, X_1]|X_k] = E[X_{k+1}|X_k].$$  

Thus $E[X_{k+1}|X_k] = X_k$. Consequently, $E[\tilde{X}_{k+1}|\tilde{X}_k] = \tilde{X}_k$ and $E[\tilde{X}_{k+1}|\sigma_k] = \tilde{X}_k$. \qed

**Definition 12.** A stochastic process which is a martingale and a Markov process will be called a Markov martingale.

**Corollary 4.** If we restrict problem (3.3) to Markov martingales only, the optimal solution to the problem does not change.

### 3.1.2 Treatment of future conditions

By future conditions we mean constraints introduced by options of maturities higher than $t^*$.

We have studied the one-dimensional case in Chapter 2. There we show that for a martingale to exist, the sequence of $\Psi$-transforms of its marginal distributions must be nondecreasing in time. For the treatment of future conditions in the
two-dimensional case, something similar applies: it is enough to ensure that the \(\Psi\)-transforms of \(\{X^1_t\}\) and \(\{X^2_t\}\) form nondecreasing sequences.

Let

\[
S^h_{\geq t^*} = \{(k^h_{t_j}, C^h_{t_j}) \mid t^* \leq t \leq n, t \in \{1, 2, \ldots, U(t, h)\}\}, \quad h = 1, 2 \quad \text{and}
\]

\[
\mathcal{F}^h = \{(k^h_{t_j}, C^h_{t_j}) \mid t^* < t \leq n, (k^h_{t_j}, C^h_{t_j}) \text{ is a vertex of the convex hull of} \quad S^h_{\geq t^*} \cup \{(0, +\infty), (+\infty, 0)\}\}
\]

represent future points which are on the border of the convex hull of \(S^h_{\geq t^*}\).

**Remark 12.** In formulating all further results we will assume that \((k^1_{t_j}, C^1_{t_j})\), 
\((1, t, j) \in \mathcal{T}\) and \((k^2_{t_j}, C^2_{t_j})\), \((2, t, j) \in \mathcal{T}\) separately do satisfy the no-arbitrage condition, that is that there exist one-dimensional martingales \(X^1_t\) and \(X^2_t\) such that

\[
E\left[(X^h_t - k^h_{t_j})^+\right] = C^h_{t_j}, \quad h = 1, 2.
\]

**Theorem 6.** Let \(X_1, X_2, \ldots, X_t\) be a two-dimensional martingale satisfying

\[
E\left[f^h_{t_j}(X_t)\right] = C^h_{t_j}, \quad \text{for all } \{(h, t, j) \in \mathcal{T} \mid t \leq t^*\}. \quad \text{If also for each } h \in \{1, 2\}
\]

\[
E\left[(X^h_t - k^h_{fut,j})^+\right] \leq C^h_{fut,j}
\]

(3.6)

is satisfied for all \((k^h_{fut,j}, C^h_{fut,j}) \in \mathcal{F}^h\), then there exist two-dimensional random variables \(X^1_{t^*+1}, \ldots, X_n\) such that

(a) \(X_1, X_2, \ldots, X_{t^*}, X_{t^*+1}, \ldots, X_n\) is a martingale;

(b) all the conditions \(E\left[f^h_{t_j}(X_t)\right] = C^h_{t_j}\) are satisfied for all \((h, t, j) \in \mathcal{T}\).

**Proof.** From the conditions of the theorem it follows that for each \(h = 1, 2\), \(\Psi_X^h\) is below the border of the convex hull \(A^h\). Since also the set \((k^h_{t_j}, C^h_{t_j})\), \((h, t, j) \in \mathcal{T}\) satisfies the no-arbitrage condition for each \(h \in \{1, 2\}\), then by Theorem 5 there exists a one-dimensional martingale \(\tilde{X}^h_t, \tilde{X}^h_{t^*+1}, \ldots, \tilde{X}^h_n\), such that it satisfies the set \((k^h_{t_j}, C^h_{t_j})\), \(t \geq t^*, \quad j = 1, \ldots, U(t, h)\) and \(\Psi_X^h = \Psi_X^h\).

We will extend the martingale \(X_1, X_2, \ldots, X_{t^*}\) to the martingale \(X_1, X_2, \ldots, X_{t^*}, X_{t^*+1}, \ldots, X_n\), such that \(X_t, X_{t^*+1}, \ldots, X_n\) is a Markov process. To define a joint distribution of \(X_t\) and \(X_{t+1}\) for \(t \geq t^*\), allow each coordinate to evolve independently of each other. That is, for each \(h = 1, 2\) define \(\mathcal{L}\left(X^h_t, X^h_{t+1}\right)\), as \(\mathcal{L}\left(\tilde{X}^h_t, \tilde{X}^h_{t+1}\right)\) for
\[ t = t^*, t^* + 1, \ldots, n. \]

Theorem 6 allows us to reduce "future conditions" to conditions at time \( t^* \). This is a significant simplification of the problem, since difficulties come from the requirement that the martingale satisfies \( E \left[ f^h_{t^*}(X_t) \right] = C^h_{t^*} \), \((h, t, j) \in T\).

Corollary 5. Our main problem is equivalent to the following: find the maximum (minimum) of \( E \left[ g(X_t^+) \right] \) under the condition that \( X_1, X_2, \ldots, X_{t^*} \) is a Markov Martingale such that

(a) \( E \left[ (X^h_t - k^h_{t,j})^+ \right] = C^h_{t,j} \) for all \( t \leq t^* \) and \( j \leq U(t, h) \) and

(b) \( E \left[ (X^h_t - k^h_{fut,j})^+ \right] \leq C^h_{fut,j} \) for all \( (k^h_{fut,j}, C^h_{fut,j}) \in \mathcal{F}^h \).

Definition 13. Let \( f^h_{fut,j} : \mathbb{R}^2 \rightarrow \mathbb{R}^+ \), \( f^h_{fut,j}(X) = (X^h - k^h_{fut,j})^+ \), be a family of functions, where each function corresponds to some \( k^h_{fut,j} \), such that \( (k^h_{fut,j}, C^h_{fut,j}) \in \mathcal{F}^h \), \( h = 1, 2 \).

The constraints from the past introduce much more difficulty to the problem than the future constraints. It is not possible to reduce past martingale conditions to the two separate conditions on \( X^1_t \) and \( X^2_t \). In other words, for a given distribution of \( X_{t^*} \), and two sequences of marginal laws \( \{\Pi_t\}_{t=1,\ldots,t^*} \) and \( \{Q_t\}_{t=1,\ldots,t^*} \) of one-dimensional processes, such that \( \Pi_t = \mathcal{L}(X^1_t) \) and \( Q_{t^*} = \mathcal{L}(X^2_{t^*}) \), the existence of one-dimensional martingales with the above sequences of marginal laws, is not sufficient for the existence of a two-dimensional martingale \( X_1, X_2, \ldots, X_{t^*} \) such that \( \mathcal{L}(X^1_t) = \Pi_t \) and \( \mathcal{L}(X^2_t) = Q_t \) for \( t = 1, \ldots, t^* \). To illustrate this point we offer the following counterexample.

Counterexample. Suppose that \( X_2 \) has an atomic distribution with two atoms \( a \) and \( b \), as shown on Figure 3-1, where each atom has weight of \( 1/2 \). Suppose that the marginal law of \( X^1_t \) is \( P(X^1_t = x_c) = 1 \), where \( x_c = \frac{1}{2}(x_a + x_b) \) and the marginal law of \( X^2_t \) is \( P(X^2_t = y_a) = P(X^2_t = y_b) = \frac{1}{2} \). Obviously, we can define a joint distribution of \( X^1_t, X^2_t \), so that it is a one-dimensional martingale. The same is true for \( X^1_t, X^2_t \). However, it is not possible to define a joint distribution of \( X_1, X_2 \), so that it is a two-dimensional martingale.

Thus, the condition that \( X_1, X_2, \ldots, X_{t^*} \) is a martingale, is truly a "joint" condition on \( \{X^1_t\} \) and \( \{X^2_t\} \).
3.1.3 Restriction to discrete distributions

In what follows, the function $g$ is defined as $g : \mathbb{R}^{+2} \to \mathbb{R}^{+}$, $(x, y) \to |\alpha x + \beta y - k_g|^{+}$. However, an analogous solution would apply if $g : \mathbb{R}^{2} \to \mathbb{R}$ is any continuous piecewise-linear function.

We will add an additional constraint to our problem that all measures have support in $[0, L]^2$ for some $L > 0$. In other words, that for $h = 1, 2$

$$P\left(X_1^h, X_2^h, \ldots, X_n^h \leq L \right) = 1.$$

**Definition 14.** Consider the graph $G$ (Figure 3-2), which is contained in $Q = \{(x, y) \in \mathbb{R}^{2+} | 0 \leq x, y \leq L\}$, and has vertices on intersections of the lines:

(a) $x = k_{1j}^1$, for $t \leq t^*$, $j \leq U(t, 1)$,  
(a') $x = k^1$ for all $(k^1, (k^1, C^1) \in \mathcal{F}^h)$,

(b) $y = k_{2j}^2$, for $t \leq t^*$, $j \leq U(t, 2)$,  
(b') $y = k^2$ for all $(k^2, (k^2, C^2) \in \mathcal{F}^h)$,

(c) $\alpha x + \beta y = k_g$,

(d) segments of $Q$: from $(0, 0)$ to $(0, L)$, $(0, L)$ to $(L, L)$, $(L, L)$ to $(L, 0)$, $(0, 0)$ to $(L, 0)$.

Each edge of the graph belongs to one of the above lines (or segments). The intersection between two edges is either empty or consists of one vertex point.

Let $G_V$ designate all the vertices of the graph and $G_E$ the set of points which lie on edges of $G$. Let us assume that the edges are enumerated. We will thus write $e$ for an edge numbered $e$. Let $\mathcal{E}$ be the set of all integers such that each of them represents some edge. Also, let us define the beginning and the end of an edge in the following way: if an edge $e$ connects vertices $(x_1, y_1)$ and $(x_2, y_2)$, then let the beginning of it
be

\[
B(\epsilon) = \begin{cases} 
(x_1, y_1), & x_1 < x_2 \\
(x_2, y_2), & x_1 = x_2, y_1 < y_2.
\end{cases}
\]

Let, correspondingly, \(E(\epsilon)\) denote the end of edge \(\epsilon\).

Let \(Edge : G_E \rightarrow \mathcal{E}\) be the function which attributes to each point of an edge the index of that edge. Since vertices are contained on several edges, we have to specify which edge corresponds to each vertex.

**Theorem 7.** Let \(X_1, X_2, \ldots, X_t^*\) be a nonnegative Markov martingale bounded by \(L\). Then there exists a nonnegative Markov martingale \(\{\tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_t^*\}\), with the state space in \(G_E\), such that it satisfies

\[
E\left[f_{ij}^h(X_t)\right] = E\left[f_{ij}^h(\tilde{X}_t)\right], \ t \leq t^*, \ (h, t, j) \in T,
\]

\[
E\left[f_{fut,j}^h(X_t)\right] = E\left[f_{fut,j}^h(\tilde{X}_t)\right], \ j = 1, 2, \ldots, U(fut, t^*),
\]

\[
E[g(X_{t^*})] = E[g(\tilde{X}_{t^*})].
\]

We will use the following two lemmas in the proof of the theorem.

**Lemma 7.** Let \(X_1, X_2\) be a martingale with values in \(\mathbb{R}^d\). Then there exists a continuous time martingale \(\{\tilde{X}_t\}_{t \in [1,2]}\) with continuous paths, such that \(\tilde{X}_1 = X_1\) and \(\tilde{X}_2 = X_2\).

In the proof of the lemma we will use the Martingale version of Skorohod embedding theorem [23, p.229]: Let \(M_n\) be a martingale with \(M_0 = 0\) and induced filtration \((\mathcal{G}_n)\). Then there exist a Brownian motion \(B\) and associated optional times \(0 = \tau_0 \leq \tau_1 \leq \cdots\) such that \(M_n = B_{\tau_n}\) a.s. for all \(n\).
Proof. By the martingale version of Skorohod embedding theorem [23, p.229], there exists a Brownian motion $B$ and optional time $\tau \geq 0$, such that $X_2 - X_1 = B_\tau$ a.s. Let $M_t = X_t + B_{t \wedge \tau}$, where $t \wedge \tau := \min(t, \tau)$. Then by optional stopping theorem $M_t$ is a martingale. This martingale is bounded and, thus, uniformly integrable. Consequently, there exists $M_\infty$ such that $\{M_t\}_{t \in [0, +\infty)}$ is a martingale and $M_t$ converges to $M_\infty$ almost surely by Doob's theorem [13, p.285].

Let $h : [0, +\infty) \to [1, 2]$ be a nondecreasing continuous function and let $\tilde{M}_t \equiv M_{h(t)}$, $t \in [1, 2]$. Then $\{\tilde{M}_1, \tilde{M}_2\}$ has the same joint distribution as $\{X_1, X_2\}$ and $\{\tilde{M}_t\}_{t \in [1, 2]}$ is a continuous time martingale with continuous paths. \hfill \Box

Lemma 8. Let $\{X_t\}_{t \in [1, 2]}$ be a continuous time martingale with values in $\mathbb{R}^d$. Let $h : \mathbb{R}^d \to \mathbb{R}$ be a continuous piecewise linear function. Let $\mathbb{R}^d = \bigcup_i D_i$, and $h$ is affine on each $D_i$, $i = 1, 2, \ldots, n$. Then for each $i$

$$E \left[ h(X_1) \mid \{X_t\}_{t \in [1, 2]} \in D_i \right] = E \left[ h(X_2) \mid \{X_t\}_{t \in [1, 2]} \in D_i \right].$$

Proof. The result immediately follows from conditional Jensen's inequality applied to $h$ and $-h$ restricted to $D_i$ [13, p.277]. \hfill \Box

Proof of Theorem 7. Start at time $t^*$ and proceed reverse recursively in the following way. Let $B_t$, $t \in [0, +\infty)$, be a two-dimensional Brownian motion starting at 0. Let $\tilde{X}_t = X_{t^*} + B_{t^*-t}$ for all $t \geq t^*$. Then $X_1, X_2, \ldots, X_{t^*}, \tilde{X}_t$ is a martingale. If $\tau_1$ is the first hitting time of $\tilde{X}_t$ on $G_E$, then by optional stopping theorem $X_1, X_2, \ldots, X_{t^*}, \tilde{X}_{(t^*+1)\wedge \tau_1}, \tilde{X}_{(t^*+2)\wedge \tau_1}, \ldots$ is also a martingale. Since this martingale is bounded, it is uniformly integrable, and thus [13, p.283, p.285] there exists $\tilde{X}_\infty$, such that $E \left[ \tilde{X}_\infty \mid B_n \right] = \tilde{X}_n$ for all $n = t^*, t^* + 1, \ldots$, and $\tilde{X}_n$ converges to $\tilde{X}_\infty$ a.s. Let $\tilde{X}_{t^*} = \tilde{X}_\infty$. Then $X_1, X_2, \ldots, X_{t^*-1}, \tilde{X}_{t^*}$ is a martingale, such that $P \left( \tilde{X}_{t^*} \in G_E \right) = 1$ and it takes the same values on functionals as $X_1, X_2, \ldots, X_{t^*}$.

Now let us make the discrete time martingale $X_1, X_2, \ldots, X_{t^*-1}, \tilde{X}_{t^*}$ be continuous on $[t^*-1, t^*]$ and have continuous paths. That is, by Lemma 7 we can define the martingale $\{\tilde{X}_t\}_{t \in [t^*-1, t^*]}$, such that $\tilde{X}_{t^*-1} = X_{t^*-1}, \tilde{X}_{t^*} = \tilde{X}_{t^*}$ and it is continuous. Let $\tau_2$ be the stopping time which stops the martingale when it first hits $G_E$ after $t^*-1$. Then by Lemma 8, $X_{t^*-1} \left( = \tilde{X}_{t^*-1} \right)$ and $\tilde{X}_{t^*}$ take the same values on functionals but $P \left( \tilde{X}_{t^*-1} \in G_E \right) = 1$. Since $X_1, X_2, \ldots, X_{t^*-1}, \tilde{X}_{t^*}, \tilde{X}_{t^*}$ is a martingale, so is $X_1, X_2, \ldots, X_{t^*-2}, \tilde{X}_{t^*}, \tilde{X}_{t^*}$. Define $\tilde{X}_{t^*-1} = \tilde{X}_{t^*}$. We continue in this way until we define $\tilde{X}_1$. \hfill \Box
**Definition 13.** We will say that a random process \( X_1, X_2, \ldots, X_{t^*} \) has property \( P_G \) if and only if all of the following hold

(a) \( P(X_{t^*} \in G_V) = 1 \);
(b) \( P(X_t \in G_E) = 1 \) for all \( t = 1, 2, \ldots, t^* \);
(c) \( X_1 \) has an atomic distribution with at most one atom per edge;
(d) If \( \epsilon_1, \epsilon_2, \ldots, \epsilon_t \in \mathcal{E} \) with \( t < t^* \), then

\[
\mathcal{L} (X_{t+1} | [\text{Edge}(X_1), \text{Edge}(X_2), \ldots, \text{Edge}(X_t)]) = [\epsilon_1, \epsilon_2, \ldots, \epsilon_t]
\]

has an atomic measure with at most one atom per edge.

**Theorem 8.** Let \( X_1, X_2, \ldots, X_{t^*} \) be a nonnegative Markov Martingale bounded by \( L \). Then there exists a nonnegative Markov Martingale \( \{\tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_{t^*}\} \), bounded by \( L \), such that it has property \( P_G \) and satisfies

\[
E \left[ f^h_{ij}(X_t) \right] = E \left[ f^h_{ij}(\tilde{X}_t) \right], \quad t \leq t^*, \quad (h, t, j) \in T,
\]

\[
E \left[ f^h_{fut,j}(X_{t^*}) \right] = E \left[ f^h_{fut,j}(\tilde{X}_{t^*}) \right], \quad j = 1, 2, \ldots, U(fut, h),
\]

\[
E [g(X_{t^*})] = E [g(\tilde{X}_{t^*})].
\]

**Proof.** From Theorem 7 we know that there exists a martingale \( \tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_{t^*} \) with a state space in \( G_E \), taking the same values on functionals \( E[f^h_{ij}(\cdot)] \), \( E[f^h_{fut,j}(\cdot)] \) and \( E[g(\cdot)] \) as \( X_1, X_2, \ldots, X_{t^*} \).

We will use the process \( \tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_{t^*} \) with the state space in \( G_E \) to form a process \( \tilde{X}_1, \ldots, \tilde{X}_{t-1} \) such that

\[
\mathcal{L} (\tilde{X}_t | [\text{Edge}(X_1), \text{Edge}(X_2), \ldots, \text{Edge}(X_{t-1})]) = [\epsilon_1, \epsilon_2, \ldots, \epsilon_{t-1}]
\]

has an atomic distribution with at most one atom per edge for all \( t = 1, 2, \ldots, t^* \).

The proof uses an induction argument. Start with \( t = 1 \). On the probability space, where the process \( \tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_{t^*} \) is formed, define a random variable \( X_0 \) with an atomic distribution, such that \( X_0 = Z_{1t} = E \left[ \tilde{X}_1 \big| \text{Edge} (\tilde{X}_1) = \epsilon \right] \) on \( \tilde{X}_1^{-1}[\text{Edge}^{-1}(\epsilon)] \). Thus,

\[
P (X_0 = Z_{1t}) = P \left( \text{Edge} (\tilde{X}_1) = \epsilon \right).
\]

But then \( X_0, \tilde{X}_1, \ldots, \tilde{X}_{t^*} \) is a martingale, \( X_0 \) has an atomic measure with one atom
per edge of $G_E$ and $X_0$ takes the same values on the functionals as $\tilde{X}_1$ by Lemma 8. Let $\tilde{X}_1 = X_0$. Then $\tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_{t^*}$ is a martingale.

Now in a similar way we can change $\tilde{X}_2$ to $\tilde{X}_2$ such that $\mathcal{L} \left( \tilde{X}_2 | Edge (\tilde{X}_1) = \epsilon \right)$ has an atomic measure with at most one atom per edge. For that it is enough to consider $\mathcal{L} \left( \tilde{X}_2 | Edge (\tilde{X}_1) = \epsilon \right) = \mathcal{L} \left( \tilde{X}_2 | \tilde{X}_1 = Z_{1\epsilon} \right)$ and for each $Z_{1\epsilon}, \epsilon \in \mathcal{E}$, to apply the same procedure, as was done for time 1, to obtain $\mathcal{L} \left( \tilde{X}_2 | \tilde{X}_1 = Z_{1\epsilon} \right)$ with the required properties.

Thus, we can form a Markov Martingale $\tilde{X}_1, \ldots, \tilde{X}_{t^*-1}, \tilde{X}_{t^*}$ such that $\mathcal{L} \left( \tilde{X}_t | [Edge(X_1), Edge(X_2), \ldots, Edge(X_{t-1})] = [\epsilon_1, \epsilon_2, \ldots, \epsilon_{t-1}] \right)$ has an atomic measure with at most one atom per edge for all $t = 1, 2, \ldots, t^*$. Define a random variable $\tilde{X}_{t^*}$ by specifying its conditional probabilities conditioned on $\tilde{X}_{t^*}$:

$$P \left( \tilde{X}_{t^*} = B(\epsilon) | \tilde{X}_{t^*} = Z_{t^*\epsilon} \right) + P \left( \tilde{X}_{t^*} = E(\epsilon) | \tilde{X}_{t^*} = Z_{t^*\epsilon} \right) = 1.$$

$$B(\epsilon)P \left( \tilde{X}_{t^*} = B(\epsilon) | \tilde{X}_{t^*} = Z_{t^*\epsilon} \right) + E(\epsilon)P \left( \tilde{X}_{t^*} = E(\epsilon) | \tilde{X}_{t^*} = Z_{t^*\epsilon} \right) = Z_{t^*\epsilon},$$

that is we redistribute the weight between the vertices of the edge so that $\tilde{X}_{t^*}, \tilde{X}_{t^*}$ is a martingale.

Then $\tilde{X}_1, \ldots, \tilde{X}_{t^*}, \tilde{X}_{t^*}$ is a martingale with $P \left( \tilde{X}_{t^*} \in G_V \right) = 1$. By Lemma 8 $\tilde{X}_{t^*}$ takes the same values on functionals as $\tilde{X}_{t^*}$. Consequently, Markov Martingale $\tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_{t^*-1}, \tilde{X}_{t^*}$ satisfies all the required properties of Theorem 8. The proof is complete. 

Thus, to solve Problem (3.3), it is enough to consider only Markov martingales $X_1, X_2, \ldots, X_{t^*}$ satisfying the property $P_G$. Given, that distributions with the $P_G$ property can be described with finitely many parameters, it is a finite dimensional optimization problem.

Remark 13. General measures belong to an infinite-dimensional space which is even uncountable. So the above result is important. However we will see that the number of parameters grows exponentially with $t^*$. So solution methods described below will work only if $t^*$ is not too large.

The previous results reduce Problem (3.3) to a finite dimensional linear programming problem that we explore in the next section.

75
### 3.2 Algorithm

Let us assume that $X_1, X_2, \ldots, X_{t^*}$ satisfy condition $P_G$. Let for $t \leq t^*$ and for $(e_1, e_2, \ldots, e_t) \in \mathcal{E}$

$$P_{t; t_1, t_2, \ldots, t_t} = P (\{ \text{Edge}(X_1), \text{Edge}(X_2), \ldots, \text{Edge}(X_t) = [e_1, e_2, \ldots, e_t] \}). \quad (3.7)$$

Let $Z_{t; t_1, t_2, \ldots, t_t} \in \mathbb{R}^{2^+}$ denote the place of the unique atom of the conditional distribution

$$\mathcal{L} (X_t \mid [\text{Edge}(X_1), \ldots, \text{Edge}(X_t)] = [e_1, e_2, \ldots, e_t]). \quad (3.8)$$

Notice, that we can write martingale conditions in terms of (3.7) and (3.8). Indeed, for $t = 2, 3, \ldots, t^*$, the martingale condition is

$$Z_{t-1; t_1, t_2, \ldots, t_{t-1}} = E [X_{t-1} \mid \{ \text{Edge}(X_1), \ldots, \text{Edge}(X_{t-1}) = [e_1, e_2, \ldots, e_{t-1}] \} = E [X_t \mid \{ \text{Edge}(X_1), \ldots, \text{Edge}(X_{t-1}) = [e_1, e_2, \ldots, e_{t-1}] \} =$$

$$\sum_{e \in \mathcal{E}} Z_{t; t_1, t_2, \ldots, t_t} P (\{ \text{Edge}(X_1), \ldots, \text{Edge}(X_t) = [e_1, e_2, \ldots, e_t] \} [\text{Edge}(X_1), \ldots, \text{Edge}(X_{t-1}) = [e_1, \ldots, e_{t-1}]) =$$

$$= \sum_{e \in \mathcal{E}} Z_{t; t_1, t_2, \ldots, t_t} P_{t; t_1, t_2, \ldots, t_{t-1}, t_t} / P_{t-1; t_1, t_2, \ldots, t_{t-1}} \quad (3.9)$$

or, equivalently,

$$Z_{t-1; t_1, t_2, \ldots, t_{t-1}} P_{t-1; t_1, t_2, \ldots, t_{t-1}} = \sum_{e \in \mathcal{E}} Z_{t; t_1, t_2, \ldots, t_{t-1}, e} P_{t; t_1, t_2, \ldots, t_{t-1}, e} \quad (3.9)$$

Let $z_{t; t_1, t_2, \ldots, t_t} \in [0, 1]$ denote the one-dimensional linear parameterization. That is

$$Z_{t; t_1, t_2, \ldots, t_t} = B(t_t) + (E(t_t) - B(t_t)) z_{t; t_1, t_2, \ldots, t_t}.$$ 

Notice that $f^1_{ij}$, $f^2_{ij}$ and $g$ are affine when restricted to one of the edges, that is

$$f^h_{ij} (Z_{t; t_1, \ldots, t_{t-1}, e}) = a^h_{ij} (e) z_{t; t_1, t_2, \ldots, t_{t-1}, e} + b^h_{ij} (e) \quad (3.10)$$

$$g (Z_{t; t_1, \ldots, t_{t-1}, e}) = a_g (e) z_{t; t_1, t_2, \ldots, t_{t-1}, e} + b_g (e) \quad (3.11)$$

for some coefficients $a^h_{ij} (e)$, $b^h_{ij} (e)$, $a_g (j)$, $b_g (e)$, which are easy to determine. Let also $a^h_{fut, j} (e)$ and $b^h_{fut, j} (e)$ represent corresponding coefficients for $f^h_{fut, j}$ (Definition 13.)
Then our optimization problem can be rewritten as:

find a collection of \( P_{t; t_1, t_2, \ldots, t_t} \geq 0 \) and \( z_{t; t_1, t_2, \ldots, t_t} \in [0, 1] \), \( t = 0, 1, \ldots, t^* \), \( \{ e_1, e_2, \ldots, e_t \} \in \mathcal{E} \), such that the following conditions are satisfied

(a) \( \sum_{t \in \mathcal{E}} P_{1; t} = 1 \);
(b) transition probabilities condition:

\[
P_{t-1; t_1, t_2, \ldots, t_{t-1}, t_t} = \sum_{t \in \mathcal{E}} P_{t; t_1, t_2, \ldots, t_{t-1}, t_t}, \quad \text{for all } t \leq t^* \text{ and } e_1, e_2, \ldots, e_{t-1} \in \mathcal{E}
\]

(c) martingale condition (3.9):

\[
Z_{t-1; t_1, t_2, \ldots, t_{t-1}, t_t} P_{t-1; t_1, t_2, \ldots, t_{t-1}, t_t} = \sum_{t \in \mathcal{E}} Z_{t; t_1, t_2, \ldots, t_{t-1}, t_t} P_{t; t_1, t_2, \ldots, t_{t-1}, t_t}
\]

for all \( t \leq t^* \) and \( e_1, e_2, \ldots, e_{t-1} \in \mathcal{E} \)

(d) the past and present prices constraints:

\[
\sum_{t_1, t_2, \ldots, t_{t-1}, t_t \in \mathcal{E}} \left( a_{ij}^h(e) z_{t; t_1, t_2, \ldots, t_{t-1}, t_t} + b_{ij}^h(e) \right) P_{t; t_1, t_2, \ldots, t_{t-1}, t_t} = C_{ij}^h
\]

for all \( ((h, t, j) \in \mathcal{T} \mid t \leq t^*) \)

(e) future prices conditions:

\[
\sum_{t_1, t_2, \ldots, t_{t^*}, t_{t+1} \in \mathcal{E}} \left( a_{ij}^{h(j)}(e) z_{t^*; t_1, t_2, \ldots, t_{t^*}, t_{t+1}, t_t} + b_{ij}^{h(j)}(e) \right) P_{t^*; t_1, t_2, \ldots, t_{t^*}, t_{t+1}, t_t} \leq C_{ij, j^*}^h
\]

for all \( ((h, t^*, j) \in \mathcal{T} \mid (k_{ij}^{h(j)}, C_{ij, j^*}^h) \in \mathcal{F}^h) \)

and \( g = \sum_{t_1, \ldots, t_{t^*}, t_{t+1} \in \mathcal{E}} \left( a_g(e) z_{t^*; t_1, \ldots, t_{t^*}, t_{t+1}, t_t} + b_g(e) \right) P_{t^*; t_1, \ldots, t_{t^*}, t_{t+1}, t_t} \) is maximized (minimized).

The above system is not linear. By making the change of variables

\[
\hat{z}_{t; t_1, t_2, \ldots, t_{t-1}, t_t} = z_{t; t_1, t_2, \ldots, t_{t-1}, t_t} P_{t; t_1, t_2, \ldots, t_t}
\]

we make it linear. The domain for \( (\hat{z}_{t; t_1, t_2, \ldots, t_{t-1}, t_t}, P_{t; t_1, t_2, \ldots, t_t}) \) is represented on Figure 3.2. Thus all the constraints as well as the objective function are linear.
Figure 3-3: Linear inequalities on $\hat{z}_{t_1,t_2,\ldots,t_{i-1},t_i}$ and $P_{t_1,t_2,\ldots,t_i}$.

3.3 Computational example

The data on VerizonCM and Cisco call options (Table 3.1) is taken from The Wall Street Journal of October 22, 2002 to derive exact bounds on the price of an option with a payoff function $(V + C - k)^+$, where $V$ is the price of the VerizonCM stock at the maturity, $C$ is the price of Cisco stock and $k$ is the strike price. The prices of VerizonCM and Cisco stocks on that day were $37.75$ and $11.22$ correspondingly.

Table 3.2 represents exact bounds on the price of the option with maturity 52, if all given options are taken into account. Table 3.3 gives bounds on the price of the same option, if only options of maturity 52 are taken into account. As expected, these bounds are looser.

As can be noticed, bounds for strike prices smaller than 20 are extremely tight. That makes sense, since the probability that the sum of prices of the two given stocks will go below 20, is almost zero. Thus, the option will be exercised almost surely and buying the option is almost equivalent to buying the two stocks.

3.4 An approximation approach

In this section, we develop an approximation approach which allows to get rid of the exponential growth of the number of variables with $t^*$.

Recall that we consider only martingales $X_t$ such that $P(X_t \leq L) = 1$ for some $L > 0$. For the approximation approach we will allow $g : \mathbb{R}^2 \to \mathbb{R}$, to be any function continuous on $[0,L] \times [0,L]$. Choose $\epsilon > 0$ and designate by $L_\epsilon$ the set of all vertices of the $\epsilon$-square lattice in $[0,L] \times [0,L] \subset \mathbb{R}^2$.

**Theorem 9.** Let $X_1, X_2, \ldots, X_t$ be a two-dimensional Markov Martingale with values in $[0,L] \times [0,L]$. Then there exists a two-dimensional Markov Martingale
<table>
<thead>
<tr>
<th>maturity</th>
<th>strike price</th>
<th>price</th>
<th>maturity</th>
<th>strike price</th>
<th>price</th>
</tr>
</thead>
<tbody>
<tr>
<td>17</td>
<td>30.0</td>
<td>7.80</td>
<td>17</td>
<td>10.0</td>
<td>1.65</td>
</tr>
<tr>
<td>17</td>
<td>37.5</td>
<td>3.70</td>
<td>17</td>
<td>12.5</td>
<td>0.35</td>
</tr>
<tr>
<td>17</td>
<td>40.0</td>
<td>0.95</td>
<td>52</td>
<td>12.5</td>
<td>0.65</td>
</tr>
<tr>
<td>17</td>
<td>42.5</td>
<td>0.35</td>
<td>80</td>
<td>7.50</td>
<td>3.75</td>
</tr>
<tr>
<td>52</td>
<td>40.0</td>
<td>1.80</td>
<td>80</td>
<td>10.0</td>
<td>2.15</td>
</tr>
<tr>
<td>80</td>
<td>30.0</td>
<td>8.60</td>
<td>80</td>
<td>12.5</td>
<td>0.80</td>
</tr>
<tr>
<td>80</td>
<td>35.0</td>
<td>5.10</td>
<td>171</td>
<td>15.0</td>
<td>0.65</td>
</tr>
<tr>
<td>80</td>
<td>40.0</td>
<td>2.20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>171</td>
<td>42.5</td>
<td>2.00</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3.1: Data.

<table>
<thead>
<tr>
<th>strike price</th>
<th>0</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>45</th>
<th>50</th>
<th>55</th>
<th>60</th>
<th>70</th>
</tr>
</thead>
<tbody>
<tr>
<td>lower bound</td>
<td>48.97</td>
<td>38.98</td>
<td>28.99</td>
<td>19.99</td>
<td>9.02</td>
<td>4.03</td>
<td>0.89</td>
<td>0.57</td>
<td>0.04</td>
<td>0.01</td>
</tr>
<tr>
<td>upper bound</td>
<td>48.97</td>
<td>39.02</td>
<td>29.31</td>
<td>19.47</td>
<td>10.56</td>
<td>7.17</td>
<td>3.91</td>
<td>2.08</td>
<td>1.93</td>
<td>1.62</td>
</tr>
</tbody>
</table>

Table 3.2: Exact lower and upper bounds on the price of an option on the linear combination of two stocks. Maturity is 52. Options of all maturities are taken into account.

<table>
<thead>
<tr>
<th>strike price</th>
<th>0</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>45</th>
<th>50</th>
<th>55</th>
<th>60</th>
<th>70</th>
</tr>
</thead>
<tbody>
<tr>
<td>lower bound</td>
<td>48.97</td>
<td>38.98</td>
<td>28.97</td>
<td>18.97</td>
<td>8.97</td>
<td>3.97</td>
<td>0.25</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Table 3.3: Exact lower and upper bounds on the price of an option on the linear combination of two stocks. Maturity is 52. Only options of this maturity are taken into account.
\( \tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_{t^*} \) such that for all \( t = 1, \ldots, t^* \), \( P \left( \tilde{X}_t \in L_t^t \right) = 1 \) and

\[
P \left( \left| \tilde{X}_t^h - X_t^h \right| \leq t \epsilon \right) = 1, \quad h = 1, 2.
\]

Here \( L_t^t \) is an \( \epsilon \)-square lattice in \([0, L + (t - 1)\epsilon] \times [0, L + (t - 1)\epsilon] \).

**Proof.** Let \( B_1 \subset B_2 \subset \cdots \subset B_{t^*} \) be the family of \( \sigma \)-algebras corresponding to the martingale \( X_1, X_2, \ldots, X_{t^*} \). Let \( \Delta_t := X_t - X_{t-1} \) for \( t = 2, 3, \ldots, t^* \). Then

\[
X_t = X_1 + \Delta_2 + \Delta_3 + \cdots + \Delta_t
\]

and \( E[\Delta_t | B_{t-1}] = 0 \).

![Figure 3-4: Lattice \( \tilde{L}_\epsilon \).](image)

Let \( \tilde{L}_\epsilon \) designate the set of all vertices of the \( \epsilon \)-square lattice in \([-L, L] \times [-L, L] \in \mathbb{R}^2 \) (Figure 3-4). And let \( S^j : [-L, L] \times [-L, L] \to \tilde{L}_\epsilon, j = 1, 2, 3, 4 \), be the family of functions, such that each function puts in correspondence to \( x \in [-L, L] \times [-L, L] \) one of the closest to it four vertices of \( \tilde{L}_\epsilon \).

Notice that \( \Delta_t \in [-L, L] \times [-L, L] \) for all \( t = 2, \ldots, t^* \). Let us define a random variable \( \tilde{\Delta}_t \) such that conditioned on \( \Delta_t \), it can take only four values: \( \{ S^j(\Delta_t) - \Delta_t, j = 1, 2, 3, 4 \} \) and such that \( E\left[ \tilde{\Delta}_t \mid B_t \right] = 0 \). Moreover, we add a requirement, that conditioned on \( B_t \), \( \tilde{\Delta}_t \) is independent of any other random variables in consideration, including \( \tilde{\Delta}_1, \tilde{\Delta}_2, \ldots, \tilde{\Delta}_{t-1} \). While \( \tilde{\Delta}_2, \tilde{\Delta}_3, \ldots, \tilde{\Delta}_{t^*} \) can be defined based on \( \Delta_2, \Delta_3, \ldots, \Delta_{t^*} \) as described above, define \( \tilde{\Delta}_1 \) in a similar way, but take \( \Delta_1 = X_1 \).

Then we have \( \Delta_t + \tilde{\Delta}_t \in \tilde{L}_\epsilon \) and \( P \left( \left| \tilde{\Delta}_t^h \right| \leq \epsilon \right) = 1, h = 1, 2, t = 1, 2, \ldots, t^* \). Let \( \sigma \left( \tilde{\Delta}_1, \tilde{\Delta}_2, \ldots, \tilde{\Delta}_t \right) \) denote the smallest \( \sigma \)-algebra for which \( \tilde{\Delta}_1, \tilde{\Delta}_2, \ldots, \tilde{\Delta}_t \) are all
measurable and
\[ \tilde{B}_t := \sigma \left( B_t \cup \sigma \left( \tilde{\Delta}_1, \tilde{\Delta}_2, \ldots, \tilde{\Delta}_t \right) \right), \quad t = 1, 2, \ldots, t^*, \]
\[ \tilde{X}_1 = X_1 + \tilde{\Delta}_1, \]
\[ \tilde{X}_2 = X_2 + \tilde{\Delta}_1 + \tilde{\Delta}_2, \]
\[ \vdots \]
\[ \tilde{X}_{t^*} = X_{t^*} + \tilde{\Delta}_1 + \tilde{\Delta}_2 + \cdots + \tilde{\Delta}_{t^*}. \]

\( \tilde{X}_1 \in L_\epsilon \) is obvious. Now
\[ \tilde{X}_2 = X_2 + \tilde{\Delta}_1 + \tilde{\Delta}_2 = X_1 + (X_2 - X_1) + \tilde{\Delta}_1 + \tilde{\Delta}_2 = (X_1 + \tilde{\Delta}_1) + (\Delta_2 + \tilde{\Delta}_2) \in \tilde{L}_\epsilon. \]

But since also \( P \left( \left| \tilde{X}_2^h - X_2^h \right| \leq 2\epsilon \right) = 1, \ h = 1, 2, \) and \( X_2 \in L_\epsilon \subset \tilde{L}_\epsilon, \) it must be that \( \tilde{X}_2 \in L_\epsilon^2. \)

Continuing recursively we prove the theorem. \( \square \)

If \( X_1, X_2, \ldots, X_n \) is a two-dimensional martingale, that satisfies \( E \left[ f_{ij}^h (X_t) \right] = C_{ij}^h \)
for \( (t, j, h) \in T, \) and \( \tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_n \) is a two-dimensional martingale, such that \( P \left( \left| \tilde{X}_t^h - X_t^h \right| \leq t\epsilon \right), \ h = 1, 2, t = 1, \ldots, n, \) then we must have
\[
\left| E \left[ f_{ij}^h \left( \tilde{X}_t \right) \right] - C_{ij}^h \right| \leq \epsilon t, \quad h = 1, 2, \quad t = 1, \ldots, n \tag{3.12}
\]

and
\[
\left| E \left[ g \left( \tilde{X}_{t^*} \right) \right] - E \left[ g \left( X_{t^*} \right) \right] \right| \leq M \sqrt{2} \epsilon t^* \tag{3.13}
\]

where \( M \) is the smallest number, such that
\[
|g(x_2) - g(x_1)| \leq M \| x_1 - x_2 \|_2
\]
for all \( x_1, x_2 \geq 0, \)

Notice that if \( g(x, y) = |ax + \beta y - k| \), then \( \sqrt{2} M = a + \beta. \)

**Theorem 10.** The following problem has its optimal solution converging to the optimal solution of Problem (3.3) as \( \epsilon \to 0: \)
Minimize (maximize) $E \left[ g \left( \tilde{X}_{t^*} \right) \right]$ over all Markov martingales $\tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_{t^*}$, such that $P \left( \tilde{X}_t \in L_t^\epsilon \right) = 1$ and

$$
|E \left[ f_{tj}^h \left( \tilde{X}_t \right) \right] - C_{tj}^h | \leq \epsilon t, \quad \{(h, t, j) \in T \mid t \leq t^* \}, \quad (3.14)
$$

$$
E \left[ f_{jut,j}^h \left( \tilde{X}_{t^*} \right) \right] - C_{jut,j}^h \leq \epsilon t^*, \quad \{(h, j_{ut}, j) \in T \mid (k_{jut,j}^h, C_{jut,j}^h) \in F^h \}. \quad (3.15)
$$

Proof. Let us prove the theorem for the case when we want to maximize $E[g(\mathbf{X}_{t^*})]$. For the minimization problem the proof is equivalent.

Let $g^*$ be the optimal solution to our original problem and $g^\epsilon$ the supremum as $\epsilon \to 0$ of optimal solutions to problems with martingales' state spaces restricted to $L_t^\epsilon$ and constraints loosened to (3.14) and (3.15). From Theorem 9 and (3.13) it immediately follows that $g^\epsilon \geq g^*$. Suppose that there exists $\delta > 0$, such that $g^\epsilon \geq g^* + \delta$. Then there exists a sequence of $\epsilon$-s, $\{\{\epsilon_n\} \mid \epsilon_1 > \epsilon_2 > \cdots \epsilon_n > \cdots \mid \lim_{n \to \infty} \epsilon_n \to 0\}$, such that the limit of optimal solutions to problems corresponding to $\epsilon_1, \epsilon_2, \ldots, \epsilon_n, \ldots$, is equal to $g^\epsilon$. Let $\{X_{t^n}^\epsilon\}_{n \geq 1}$ be the sequence of martingales corresponding to the sequence $\{\epsilon_n\}$, such that $\{X_{t^n}^\epsilon\}_{n \geq 1}$ defines an optimal solution to the $\epsilon_n$-problem.

By definition of $\{X_{t^n}^\epsilon\}$, it takes values in $L_{\epsilon_n}^\epsilon$ and

$$
|E \left[ f_{tj} \left( X_{t^n}^\epsilon \right) \right] - C_{tj} | \leq \epsilon_t n, \quad t = 1, 2, \ldots, t^*.
$$

Let $\mu^{\epsilon_n}$ designate the law of $\{X_{t^n}^\epsilon\}_{t=1,\ldots,t^*}$. Then $\mu_{\epsilon_1}, \mu_{\epsilon_2}, \ldots, \mu_{\epsilon_n}, \ldots$ have support in $[0, L + t^* \epsilon_1]^2$, which is a compact. It follows that the sequence of laws $\{\mu_{\epsilon_n}\}_{n \geq 1}$ is uniformly tight. Consequently, there exists a subsequence $\mu_{\epsilon_n(\epsilon)} \to \mu$ for some law $\mu$ [13, p.230]. But then we have an admissible solution to our original problem with an optimal value greater than $g^* + \delta$. Contradiction.

\[ \square \]

### 3.4.1 Algorithm

Let us enumerate all the nodes of $L_\epsilon$ and denote by $N$ the set of all the indexes. Designate by $Z_n \in L_\epsilon$, the n-th node of $L_\epsilon$. Let

$$
P_{t,n} = P \left( \tilde{X}_t = Z_n \right) \quad \text{and}
$$

82
\[ P_{t; n_1, n_2} = P \left( \tilde{X}_t = Z_{n_1}; \tilde{X}_{t+1} = Z_{n_2} \right). \]

Notice that the distribution of \( \tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_{t^*} \) is uniquely determined by \( P_{t; n} \) and \( P_{t; n_1, n_2} \), \( t \in \{1, 2, \ldots, t^*\}, n, n_1, n_2 \in \mathbb{N} \).

Then the objective is to find a collection of \( P_{t; n} \in [0, 1] \) and \( P_{t; n_1, n_2} \in [0, 1], t \in \{1, 2, \ldots, t^*\}, n, n_1, n_2 \in \mathbb{N} \), such that

(a) \( \sum_{n \in \mathbb{N}} P_{t; n} = 1 \),
(b) \( \sum_{n \in \mathbb{N}} P_{t; n_1, n} = P_{t; n_1} \quad t < t^*, n_1 \in \mathbb{N} \),
(c) \( \sum_{n \in \mathbb{N}} P_{t; n, n_1} = P_{t+1; n_1} \quad t < t^*, n_1 \in \mathbb{N} \),
(d) \( \sum_{n \in \mathbb{N}} P_{t; n_1, n} Z_n = P_{t; n_1} Z_{n_1} \) for all \( t < t^*, n_1 \in \mathbb{N} \),
(e) \( \sum_{n \in \mathbb{N}} P_{t; n} f_{ij}^h(Z_n) - C_{ij}^h \leq t \epsilon \), \( \sum_{n \in \mathbb{N}} P_{t; n} f_{ij}^h(Z_n) - C_{ij}^h \geq -t \epsilon \) for \( \{(h, t, j) \in \mathcal{T} \mid t \leq t^*\} \),
(f) \( \sum_{n \in \mathbb{N}} P_{t; n} f_{\text{fut}, j}^h(Z_n) - C_{\text{fut}, j}^h \leq t^* \epsilon \) for \( \{(h, t, j) \in \mathcal{T} \mid (k_{\text{fut}, j}^h, C_{\text{fut}, j}^h) \in \mathcal{I}^h\} \),

and \( \sum_{n \in \mathbb{N}} P_{t; n} g(Z_n) \) is maximized (minimized).

Notice, that in this problem the only variables are the probabilities \( P_{t; n} \) and \( P_{t; n_1, n_2} \), \( t \in \{1, 2, \ldots, t^*\}, n, n_1, n_2 \in \mathbb{N} \). The \( Z_n, n \in \mathbb{N} \), have already been assigned.

### 3.4.2 Approximation approach for the special case of a payoff function

We can combine ideas of the exact and approximation approaches to obtain a more efficient algorithm for the case when \( g \) is a piecewise linear continuous function. In particular, from theorems developed for finding the exact solution we know that it is enough to consider martingales with the state space in \( G_E \) (edges of Graph \( G \)) only. From the general approximation approach (Theorem 9), we know that for each martingale \( \{X_t\}_{t=1, \ldots, t^*} \) with the state space in \([0, L] \times [0, L]\), there exists a martingale \( \tilde{X}_t \) with the state space in \( L^t \) (nodes of the \( \epsilon \)-square lattice in \([0, L+(t-1)\epsilon] \times [0, L+(t-1)\epsilon]\)), such that \( P \left| X_t^h - \tilde{X}_t^h \right| \leq 1 \) for \( t = 1, 2, \ldots, t^* \). Thus, if in the approximation formulation of the problem we allow \( \tilde{X}_t \) to take values in

\[ \Lambda^t \epsilon := \{(x, y) \in L^t \mid \exists (x_0, y_0) \in G_E : |x - x_0| \leq t \epsilon, |y - y_0| \leq t \epsilon\}, \]

instead of \( L^t \), the optimal solution to the approximation problem should not change.

Thus we have the following corollary:
Corollary 6. The following problem has its optimal solution converging to the optimal solution of problem (3.3) as $\epsilon \to 0$:
Minimize (maximize) $E \left[ g \left( \tilde{X}_{t^*} \right) \right]$ over all Markov martingales $\tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_{t^*}$, such that $P \left( \tilde{X}_t \in \Lambda_t \right) = 1$ and
\[
\left| E \left[ f_{t,j}^h \left( \tilde{X}_t \right) \right] - C_{t,j}^h \right| \leq \epsilon t, \quad \{(h,t,j) \in \mathcal{T} \mid t \leq t^*\},
\]
\[
E \left[ f_{fut,j}^h \left( \tilde{X}_{t^*} \right) \right] - C_{fut,j}^h \leq \epsilon t^*, \quad \{(h,fut,j) \in \mathcal{T} \mid (k_{fut,j}^h, C_{fut,j}^h) \in \mathcal{F}^h\}.
\]

For the discussions of running times see Section 3.6.

3.5 Application of the two-dimensional approach to the one-dimensional case

We want now to return to the problem formulated in Section 2.1, that is, determining exact bounds on the price of a European call option, given the set of priced options of different maturities.

The incentive is to give a different solution, which is based on the approach found for the two-dimensional case. This approach not only allows to find a completely different solution, but also extends the result to a more general case. Specifically, exact bounds on the option with an arbitrary piecewise linear continuous payoff function can be found.

We will assume again that the martingale has a support in $[0,L]$.
Let $\mathcal{T} = \{(t,j) \mid t \in \{1,2,\ldots,n\}, j \in \{1,2,\ldots,U(t)\}\}$ and $(k_{t,j}, C_{t,j}), (t,j) \in \mathcal{T}$ be the set of options, satisfying the no-arbitrage condition. Denote
\[
\mathcal{S}_{2^t} = \{(k_{t,j}, C_{t,j}) \mid (t,j) \in \mathcal{T}, t \geq t^*\} \cup \{(L,0)\},
\]
\[
\mathcal{F} = \{(k_{t,j}, C_{t,j}) \mid t > t^*, (k_{t,j}, C_{t,j}) \text{ is a vertex of the convex hull of } \mathcal{S}_{2^t} \cup \{(0,0)\}\},
\]
\[
G = \{0\} \cup \mathcal{L} \cup \{(k_{t,j} \mid (t,j) \in \mathcal{T}, t < t^*\} \cup \{(k_{t,j}, C_{t,j}) \in \mathcal{F}\}.
\]
Let $g : [0,L] \to \mathbb{R}$ be a continuous piecewise-linear function and
\[
K_g = \{k \in [0,L] \mid g'(k-) \neq g'(k+)\} \text{ the set of points, where } g \text{ changes its derivative.}
\]

84
Define a family of measurable functions: \( f_{tj} : \mathbb{R}^+ \rightarrow \mathbb{R}^+ \), \((t, j) \in T\) as
\[
f_{tj}(X_t) = (X_t - k_{tj})^+.
\]

**Theorem 11.** Let \( X_1, X_2, \ldots, X_{t^*}, \ldots, X_n \) be a nonnegative martingale bounded by \( L \). Then there exists a nonnegative Markov martingale \( \bar{X}_1, \bar{X}_2, \ldots, \bar{X}_{t^*}, \ldots, \bar{X}_n \), such that
\[
P \left( \{ \bar{X}_t \}_{t=1,2,\ldots,t^*} \in \{ G \cup K_g \} \right) = 1 \text{ and}
\]
\[
E[f_{tj}(X_t)] = E[f_{tj}(\bar{X}_t)] \quad \text{for all} \quad (t, j) \in T,
\]
\[
E[g(X_{t^*})] = E[g(\bar{X}_{t^*})].
\]

**Proof.** As in the two dimensional case, we can restrict the consideration to Markov martingales only, based on Proposition 7. To obtain the martingale \( \bar{X}_1, \ldots, \bar{X}_{t^*}, \ldots, \bar{X}_n \) with the state space in \( \{ G \cup K_g \} \), but taking the same values on the functionals as \( X_1, X_2, \ldots, X_{t^*}, \ldots, X_n \), we can proceed exactly as in the proof of Theorem 7, except that here we have to use one-dimensional Brownian motions, instead of two-dimensional Brownian motions. \( \square \)

Let \( \mathcal{K} = \{ G \cup K_g \} = \{ k_1, k_2, \ldots, k_n, \ldots, k_N \} \) and \( \mathcal{N} \) denote the set of all the indexes \( 1, 2, \ldots, N \). Let
\[
P_{t,n} = P \left( \bar{X}_t = k_n \right) \quad \text{and}
\]
\[
P_{t,n_1,n_2} = P \left( \bar{X}_t = k_{n_1}, \bar{X}_{t+1} = k_{n_2} \right).
\]

Then the problem of determining exact upper (lower) bounds on the price of an option with maturity \( t^* \) and payoff function \( g \), consists of finding the collection of \( P_{t,n} \in [0, 1] \) and \( P_{t,n_1,n_2} \in [0, 1] \), such that

(a) \( \sum_{j \in J} P_{1,n} = 1, \)

(b) \( \sum_{n \in \mathcal{N}} P_{t,n_1,n} = P_{t,n_1} \quad \text{for all} \quad t < t^*, \quad n_1 \in \mathcal{N}, \)

(c) \( \sum_{n \in \mathcal{N}} P_{t,n_1,n_1} = P_{t+1,n} \quad \text{for all} \quad t < t^*, \quad n_1 \in \mathcal{N}, \)

(d) \( \sum_{n \in \mathcal{N}} P_{t,n_1,n} k_n = P_{t,n_1} k_{n_1} \quad \text{for all} \quad t < t^*, \quad n_1 \in \mathcal{N}, \)

(e) \( \sum_{n \in \mathcal{N}} P_{t,n} f_{tj}(k_n) = C_{tj} \quad \text{for all} \quad t \leq t^*, (t, j) \in T, \)

(f) \( \sum_{n \in \mathcal{N}} P_{t,n} f_{t\ast,j}(k_n) \leq C_{fut,j} \)

and \( \sum_{n \in \mathcal{N}} P_{t^*,n} g(k_n) \) is maximized (minimized).
Notice, that the number of variables in this linear programming problem is $O(K^2t^*)$, so this approach is less efficient, than the approach developed in Chapter 2, for the problem of determining tight bounds on the price of European call option. However, this approach works for more general payoff functions.

### 3.6 Generalization to a multiple dimensional case

The exact solution, the approximation approach for an arbitrary payoff function $g$ and the approximation approach for a continuous piecewise linear payoff function $g$, can all be extended from a two dimensional case to a multiple dimensional case. All theorems and definitions of this chapter can be reformulated for an $n$-dimensional case, and proofs will be identical up to a dimensionality.

Let us estimate the complexity of the problem for each approach. Let $n$ be the number of different assets. Let $K^h$ be the number of different strike prices of options on asset $h$. Remember, that we don’t need to take into account options of maturities greater than $t^*$, unless they are on the lower border of the convex hull of options of maturity $t^*$ or higher (more precisely, on the border of of the convex hull of $S^h_{t^*} \cup \{(0, +\infty), (+\infty, 0)\}$.) For the purposes of estimating the complexity of the problem, we can assume that $K^h$ is the same across different assets, that is $K^h = K$ for all $h = 1, 2, \ldots, n$. Let $N = L/c$ in approximation methods. Then Table 3-5 represents the complexity of the final linear programming program if different approaches are used. It is not relevant for the one-dimensional case, though, except for the general approximation algorithm.

<table>
<thead>
<tr>
<th>number of variables</th>
<th>exact solution</th>
<th>general case approximation algorithm</th>
<th>special case approximation algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>$O(K^nt^*)$</td>
<td>$O(N^{2n}t^*)$</td>
<td>$O(N^{n-1}K^*)^2$</td>
<td></td>
</tr>
</tbody>
</table>

Figure 3-5: The complexity of the final linear-programming problem for different approaches.

Notice, that the problem for the one-dimensional case and a continuous piecewise linear payoff function, is qualitatively different from the multiple dimensional case. The complexity of the exact solution is $O(K^2t^*)$ in that case, not $O(K^*)$, as the table shows. This is so, since in a one dimensional case we immediately know the
locations of atoms of an optimal martingale distribution and, thus, the number of
variables (weights of atoms and transition probabilities) grows linearly with $t^*$. In
a multiple dimensional case, to obtain an atomic distribution at time $t^*$, we have to
condition on which edges (or faces) of the graph have been visited up to that time.
Thus, the number of variables grows exponentially with $t^*$.

Remark. We conjecture that the minimum number of variables necessary to obtain
the exact solution to Problem 3.3 grows exponentially with $n$ and $t^*$. However, $K$
can be reduced, which we will do in our future work.

### 3.7 Static arbitrage bounds on basket option prices

We notice, that the idea of the approach developed in this chapter also works for the
problem of determining upper and lower bounds on the price of a European basket
call option, given prices of other similar baskets of the same maturity:

Given $C \in \mathbb{R}^{n+}$, $k \in \mathbb{R}^{m+}$, $w_i \in \mathbb{R}^n$, $i = 1, \ldots, m$ and $k_0 > 0$, $w_0 \in \mathbb{R}^{n+}$, find the
upper and lower bounds on

$$E\left(w_0^T X - k_0\right)^+, \quad (3.16)$$

with respect to distributions of an $n$-dimensional random variable $X$ with finite
expectation and support in $\mathbb{R}^{n+}$ under the condition

$$E\left(w_i^T X - k_i\right)^+ = C_i, \quad i = 1, \ldots, m. \quad (3.17)$$

To solve the problem we make an additional assumption that the support of $X$ is
bounded, that is, there exists $L > 0$, such that $P\left(X^h \leq L\right) = 1$ for any $h = 1, 2, \ldots, n$.

Let $f_i : \mathbb{R}^{n+} \to \mathbb{R}^+$ be the family of measurable functions defined as

$$f_i(X) = \left(w_i^T X - k_i\right)^+, \quad i = 0, 1, \ldots, m.$$ 

Let $\bigcup_j D_j = [0, L]^n$ be the partition of $[0, L]^n$, into subsets $D_j$ of $\mathbb{R}^{n+}$ such that in the
interior of each subset $D_j$ all functions $f_i$, $i = 0, 1, \ldots, m$ are affine.

Assume that $X$ is defined on some probability space $(\Omega, \mathcal{F}, P)$ and the distribution
of it is known. Let $Z_j = E\left[X|X \in D_j\right]$. Define a random variable $\tilde{X}$ on the same
probability space so that

\[ P \left( \tilde{X} = Z_j \mid X \in D_j \right) = 1. \]

Then \( \tilde{X} \) is uniquely defined and \( \{ \tilde{X}, X \} \) is a martingale. By Lemma 8, \( \tilde{X} \) and \( X \) take the same values on functionals (3.16) and (3.17). Thus to find exact upper and lower bounds on (3.16) it is enough to consider atomic distributions with one atom per each subset \( D_j \). Notice, that the number of atoms can not be smaller than the number of subsets \( D_j \) to repeat values achieved by an arbitrary distribution on functionals (3.16) and (3.17). Applying the same idea as in deriving the algorithm in Section 4.3. we can write a linear programming problem to find optimal bounds. However if we want to restrict ourselves to discrete distributions with only one atom per \( D_j \), then we have to maximize (minimize) over all locations of the atom within \( D_j \). It seems more plausible to increase the number of atoms, while fixing their locations and thus reducing the number of constraints in the linear programming problem.

Let \( G \subset [0, L]^n \) be the graph in \( \mathbb{R}^n \) formed by intersecting hyperplanes

\[
\begin{align*}
\omega^T_i X &= k_i, \quad i = 1, 2, \ldots, m, \\
X^h &= 0, \quad h = 1, 2, \ldots, n, \\
X^h &= L, \quad h = 1, 2, \ldots, n.
\end{align*}
\] (3.18)

Let us enumerate these hyperplanes and call them \( h_1, h_2, \ldots, h_{2n+m} \) correspondingly. We call a vector \( r \in [0, L]^n \) a vertex of \( G \) if there are \( n \) independent hyperplanes of (3.18) which intersect at \( r \). Let \( G_V \) designate the set of all vertices. Notice, that this set contains no more than \( \binom{2n+m}{n} \) elements.

**Corollary 7.** For each random variable \( X \) with support in \([0, L]^n\) there exists a random variable \( \tilde{X} \) such that \( P (\tilde{X} \in G_V) = 1 \) and \( E[f_i(X)] = E[f_i(\tilde{X})], \ i = 0, 1, \ldots, m. \)

**Proof.** Let \( X_0 = X \) and \( B_t, t \in [0, \infty) \) be an \( n \)-dimensional Brownian motion starting at 0. Define \( \hat{X}_t = X_0 + B_t \). Let \( \tau_1 \) be the first hitting time by \( \hat{X}_t \) on one of the hyperplanes (3.18). Then by optional stopping theorem [13, p. 281] \( \hat{X}_{t \wedge \tau_1} \) is a martingale. This martingale is bounded, thus uniformly integrable and there exists \( \hat{X}_\infty \) such that \( \{ \hat{X}_t \}_{t \in [0, +\infty]} \) is a martingale. By Doob's theorem [13, p.285] \( \hat{X}_t \) converges to \( \hat{X}_\infty \) a.s. Now \( P (X_1 \in \bigcup_{j=1}^{2n+m} h_j \cap [0, L]^n) = 1 \) and \( E[f_i(X_1)] = E[f_i(X_0)], \ i = 0, 1, \ldots, m, \) by Lemma 8.
Let us define $\tilde{X}_t$ by specifying its conditional distribution with respect to $X_1$. In particular, let $\mathcal{L}\left(\tilde{X}_t | X_1 \in h_j\right) = \mathcal{L}\left(X_1 + B_t^{h_j} | X_1 \in h_j\right)$, where $B_t^{h_j}$ is an $(n - 1)$ dimensional Brownian motion in the hyperplane $h_j$ starting at 0. Again we can choose a stopping time $\tau_2$ which is the first hitting time of $\tilde{X}_t$ on one of the hyperplanes $h_1, \ldots, h_{2n+m}$ which is different from $h_j$. If we continue recursively, a martingale $X_0, X_1, \ldots, X_n$, will be formed, such that $P(X_n \in G_V) = 1$ and $X_n$ takes the same values on the functionals as $X_0$.

Thus to find optimal bounds we have to solve a linear programming problem with unknown variables being weights of the atomic distribution. The atoms are located in the vertices of the defined graph.

**Remark.** The reduction of all problems considered in this section to finite dimensional linear programming problems is possible due to the piecewise linearity of constraints and the objective function.
Chapter 4

Portfolio Optimization With General Price Dynamics

4.1 Introduction

In this chapter we address a few general issues in dynamic portfolio allocation problem and propose an algorithm for the problem with the mean-variance objective function. While the price dynamics is assumed to be known, we do not make any assumptions about the distribution of the process.

We show that while the problem of maximizing $E[f(W_T)]$, where $f : \mathbb{R} \to \mathbb{R}$ is a given function, can always be solved by stochastic dynamic programming, this is not true for the mean-variance optimization. Specifically, the principle of optimality is violated, in the sense that the optimal policy on $[t, T]$ might not coincide with the optimal policy on $[0, T]$ when restricted to $[t, T]$.

We prove that as $T \to \infty$, there exists a policy which leads to a distribution of the terminal wealth with the property that the probability of underperforming any other policy tends to zero as $T \to \infty$.

At last we address the problem of maximizing an analogue to the mean-variance function of the terminal wealth in the multiperiod case. In that problem the optimal policy is coherent, which means that the problem can be in principle solved by dynamic programming. However the time required for computations would grow exponentially with the dimension of an information vector, that is the vector, which totally determines the distribution of future prices. We suggest a monte-carlo based method which is polynomial in a number of securities. Importantly, the number of
simulations that we need does not depend on the dimension of the information vector.

4.2 On the consistency of optimal policies

Let $\mathcal{T} \subseteq [0, T]$ be a linearly ordered set that represents trading dates, and $\{\mathcal{F}_t\}_{t \in \mathcal{T}}$ a family of $\sigma$-algebras with $\mathcal{F}_t \subset \mathcal{F}_s$ for $t \leq s$, such that $P_t$ is measurable with respect to $\mathcal{F}_t$ for each $t$. Let $\delta_{(t_1 \rightarrow t_2)} = \{\delta(t) | t \in \mathcal{T}, t_1 \leq t < t_2\}$ denote a self-financing strategy for the period $[t_1, t_2)$. Here $\delta(t)$, the policy at time $t$, is $\mathcal{F}_t$-measurable for each $t \in \mathcal{T}$. The wealth at time $T$, $W_T$, is determined, once $W_t$, $\delta_{[t \rightarrow T)}$ and $\{P_s\}_{s \in [t, T)}$ are given for some time $t \in \mathcal{T}$. Thus $W_T = W_T(W_t, \delta_{[t \rightarrow T)}, \{P_s\}_{s \in [t, T)})$.

**Definition.** We call an optimal policy $\delta_{(0 \rightarrow T)}$ on $[0, T]$ coherent if when restricted to $[t, T]$ it is also optimal for $[t, T]$ for any $t$.

Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a measurable function. Then the problem of maximizing $E[f(W_T)]$ has the following important property:

**Theorem 1.**

(a) The problem of maximizing $E[f(W_T)|\mathcal{F}_t]$ leads to a coherent policy. In other words, an optimal policy on $[t_2, T]$, $\delta_{(t_2 \rightarrow T)}$, which maximizes

$$E[f(W_T(W_{t_2}, \delta_{(t_2 \rightarrow T)}, \{P_s\}_{s \in [t_2, T]})) | \mathcal{F}_{t_2}],$$

defines the policy that at any time $t_1 \leq t_2$ will be considered optimal for the time interval $[t_2, T]$, given that $\mathcal{F}_{t_2}$ is realized up to $t_2$.

(b) The problem of maximizing $E[W_T|\mathcal{F}_t] - \gamma Var(W_T|\mathcal{F}_t)$ does not imply a coherent policy.

**Proof.**

(a)

$$E[f(W_T(W_{t_1}, \delta_{(t_1 \rightarrow T)}, \{P_s\}_{s \in [t_1, T]})) | \mathcal{F}_{t_1}] =$$

$$E[E[f(W_T(W_{t_2}, \delta_{(t_2 \rightarrow T)}, \{P_s\}_{s \in [t_2, T]})) | \mathcal{F}_{t_2}] | \mathcal{F}_{t_1}] \leq$$

$$E\left[\max_{\delta_{(t_2 \rightarrow T)}} E[f(W_T(W_{t_2}, \delta_{(t_2 \rightarrow T)}, \{P_s\}_{s \in [t_2, T]})) | \mathcal{F}_{t_2}] | \mathcal{F}_{t_1}\right] =$$

$$E[E[f(W_T(W_{t_2}, \delta^*_{(t_2 \rightarrow T)}, \{P_s\}_{s \in [t_2, T]})) | \mathcal{F}_{t_2}] | \mathcal{F}_{t_1}] =$$

$$E[f(W_T(W_{t_2}, \delta^*_{(t_2 \rightarrow T)}, \{P_s\}_{s \in [t_2, T]})) | \mathcal{F}_{t_1}] =$$
\[ E \left[ f \left( W_T \left( W_{t_1}, \delta_{[t_1 \rightarrow t_2]} \cup \delta^*_{[t_2 \rightarrow T]} \right), \{P_s\}_{s \in \{t_1, T\}} \right) | \mathcal{F}_{t_1} \right], \quad (4.1) \]

where
\[ \delta^*_{[t_2 \rightarrow T]} = \arg \max_{\delta_{[t_2 \rightarrow T]}} E \left[ f \left( W_T \left( W_{t_2}, \delta_{[t_2 \rightarrow T]} \right), \{P_s\}_{s \in \{t_2, T\}} \right) | \mathcal{F}_{t_2} \right] \]
and \( \delta_{[t_1 \rightarrow t_2]} \cup \delta^*_{[t_2 \rightarrow T]} \) denote the policy over the period \([t_1, T]\), which consists of the policy \( \delta_{[t_1 \rightarrow t_2]} \) on the interval \([t_1, t_2]\) and the policy \( \delta^*_{[t_2 \rightarrow T]} \) on the interval \([t_2, T]\).

From (4.1) it follows that an optimal policy on \([t_1, T]\) will contain an optimal policy on \([t_2, T]\).

\[ \square \]

**Remark.** Part (a) of the theorem implies that when maximizing the expected value of a nonrandom function of a terminal wealth, an optimal strategy at time \( t \) depends only on the wealth at time \( t \), \( W_t \), and the conditional distribution \( \mathcal{L} \left( \{P_s\}_{s \in \{t, T\}} | \mathcal{F}_t \right) \) of future prices, conditioned on the information obtained up to time \( t \), but does not depend on the policy that we apply up to time \( t \).

(b) We give a counterexample. Suppose that we can make decisions only at time \( t_1 \) and \( t_2 \) and we want to maximize \( E[W_T | \mathcal{F}_{t_1}] - Var[W_T | \mathcal{F}_{t_1}] \) (Figure 4.1.) Moreover, let us assume that at time \( t_1 \) we are unable to influence the change of the value of the portfolio from time \( t_1 \) to \( t_2 \), while at time \( t_2 \) we can deterministically decide on which of the two directions to go. Suppose \( p_1 = p_2 = 0.5 \) are the probabilities of moving up and down at time \( t_1 \). Let \( \delta_1 \) be the policy that chooses to go up in states I and II, and \( \delta_2 \) be the policy that chooses to go up in state II and down in state I. Then, \( \delta_1 \) maximizes \( E[W_T | \mathcal{F}_{t_2}] - Var[W_T | \mathcal{F}_{t_2}] \). However \( \delta_1 \) does not maximize \( E[W_T | \mathcal{F}_{t_1}] - Var[W_T | \mathcal{F}_{t_1}] \):

\[
E \left[ W_T(\delta_1) | \mathcal{F}_{t_1} \right] = 3 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = 2, \quad E \left[ W_T^2(\delta_1) | \mathcal{F}_{t_1} \right] = \frac{1}{2} \cdot 9 + \frac{1}{2} \cdot 1 = 5, \]
\[
E \left[ W_T(\delta_1) | \mathcal{F}_{t_1} \right] - Var[W_T(\delta_1) | \mathcal{F}_{t_1}] = 1; \]
\[
E \left[ W_T(\delta_2) | \mathcal{F}_{t_1} \right] = 2 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = \frac{3}{2}, \quad E \left[ W_T^2(\delta_2) | \mathcal{F}_{t_1} \right] = \frac{1}{2} \cdot 4 + \frac{1}{2} \cdot 1 = \frac{5}{2}, \]
\[
E \left[ W_T(\delta_2) | \mathcal{F}_{t_1} \right] - Var[W_T(\delta_2) | \mathcal{F}_{t_1}] = \frac{5}{4} > 1. \]

\[ \square \]
4.3 On the asymptotic optimality of the myopic policy

In this section, we propose a myopic policy that is asymptotically optimal as \( T \to \infty \), which suggests that for large time horizons and arbitrary price dynamics, single period optimization as opposed to multiperiod is sufficient.

From now on we will assume that portfolio rebalancing is done only at some discrete times \( t_0, t_1, \ldots, t_n \).

Let \( P_t^i \) designate the price of the \( i \)-th security at time \( t \) and \( P_t = (P_t^1, P_t^2, \ldots, P_t^N) \) prices of \( N \) securities at time \( t \). Without loss of generality we will consider integer times \( t \in N \) only and take \( P_0 = (1, 1, \ldots, 1) \), \( W_0 = 1 \) for convenience. Let \( R_t^i = P_t^i / P_{t-1}^i \) designate the return on asset \( i \) for the time interval \([t - 1, t]\). Then \( P_t^i = R_t^i R_{t-1}^i \ldots R_1^i \).

If at time \( t \) we choose the policy \( x_t = (x_t^1, x_t^2, \ldots, x_t^N) \geq 0 \), where \( \sum_{i=1}^N x_t^i = 1 \), and the value of the portfolio at that moment is \( W_t \), then

\[
W_{t+1} = W_t \left( \sum_{i=1}^N x_t^i R_{t+1}^i \right) = W_t (x_t' R_{t+1}).
\]

For different policies \( \delta \), \( x_t \) are different random variables measurable with respect to
\[ W_t(\delta) = \prod_{k=0}^{t-1} (x_k(\delta)' R_{k+1}). \]

Let us assume, that there exists \( \sigma_{up} > 0 \) such that for all \( t = 0, 1, 2, \ldots T, \) and \( x \geq 0 \) with \( \sum_{i=1}^{N} x_i = 1 \) we have

\[ \text{Var} (\log (x_t'R_{t+1}) \mid F_t) \leq \sigma_{up}^2, \]

that is, the volatility of returns on each time interval is bounded.

**Theorem 2.** Let \( \delta^{opt} \) be the policy which at each time \( t = 1, 2, \ldots \) maximizes

\[ E [\log (x_t(\delta)' R_{t+1}) \mid F_t] \]

and assume that condition (4.3) holds. Let \( 0 \leq \epsilon, p \leq 1. \) Then for any \( t \) and for any policy \( \delta \) which up to time \( t \) invests at least \( p \) percentage of times so that

\[ E [\log (x_t(\delta^{opt})' R_{t+1}) \mid F_t] \geq \epsilon + E [\log (x_t(\delta)' R_{t+1}) \mid F_t], \]

we have

\[ P \left( \frac{W_t(\delta^{opt})}{W_t(\delta)} \geq \exp \left\{ \frac{\epsilon p}{2} t \right\} \right) \geq 1 - \frac{16\sigma_{up}^2}{\epsilon^2 p^2} \left( \frac{1}{t} \right). \]

**Proof.** Define

\[ Y_t(\delta) := \log W_t = \begin{cases} 0, & t = 0 \\ \sum_{k=1}^{t} \log (x_{k-1}(\delta)' R_k), & t \geq 1 \end{cases} \]

where \( \delta \) is an arbitrary policy. Then

\[ \tilde{Y}_t(\delta) := Y_t(\delta^{opt}) - Y_t(\delta) \]

is a submartingale. Indeed, it follows from the definition of \( \delta^{opt} :\)

\[ E [Y_t(\delta^{opt}) - Y_{t-1}(\delta^{opt}) \mid F_{t-1}] \geq E [Y_t(\delta) - Y_{t-1}(\delta) \mid F_{t-1}] \quad \forall \delta. \]
Thus
\[ E \left[ Y_t(\delta^{opt}) - Y_t(\delta) | \mathcal{F}_{t-1} \right] \geq Y_{t-1}(\delta^{opt}) - Y_{t-1}(\delta) \quad \forall \delta. \]

Consequently, by Doob’s decomposition theorem [13, p.277] there exists a martingale \( \tilde{z}_t(\delta) \) and an increasing process \( \tilde{M}_t(\delta) \) such that
\[ \tilde{Y}_t(\delta) = \tilde{z}_t(\delta) + \tilde{M}_t(\delta). \quad (4.7) \]

Let \( \tilde{M}_0(\delta) = 0 \) and for \( t \geq 1 \)
\[ \tilde{M}_t(\delta) = \sum_{k=1}^{t} E \left[ \tilde{Y}_k(\delta) - \tilde{Y}_{k-1}(\delta) | \mathcal{F}_{k-1} \right], \quad (4.8) \]

then \( \tilde{M}_t(\delta) \) satisfies the properties of an increasing process, that is, it is \( \mathcal{F}_{t-1} \) measurable and \( \tilde{M}_{t+1}(\delta) \geq \tilde{M}_t(\delta) \). It is easy to check that \( \tilde{z}_t(\delta) = \tilde{Y}_t(\delta) - \tilde{M}_t(\delta) \) is a martingale:
\[ \tilde{z}_t(\delta) = \tilde{Y}_t(\delta) - \tilde{M}_t(\delta) = \tilde{Y}_t(\delta) - \sum_{k=1}^{t} E \left[ \tilde{Y}_k(\delta) - \tilde{Y}_{k-1}(\delta) | \mathcal{F}_{k-1} \right] = \quad (4.9) \]
\[ \sum_{k=1}^{t} \left[ \tilde{Y}_k(\delta) - E \left[ \tilde{Y}_k(\delta) | \mathcal{F}_{k-1} \right] \right] + \tilde{Y}_0(\delta). \]

Notice, that \( \tilde{Y}_0(\delta) = 0 \) by definition.

Take \( \delta^{opt} \) and \( \delta \) so that they satisfy (4.4). Then
\[ \tilde{M}_t(\delta) = \sum_{k=1}^{t} E \left[ \log \left( x_{k-1} (\delta^{opt})' R_k \right) | \mathcal{F}_{k-1} \right] - \sum_{k=1}^{t} E \left[ \log \left( x_{k-1}(\delta)' R_k \right) | \mathcal{F}_{k-1} \right] \geq \epsilon tp. (4.10) \]

By Chebyshev inequality:
\[ P \left( |\tilde{z}_t(\delta)| \geq \frac{\epsilon tp}{2} \right) \leq \frac{4E \left[ (\tilde{z}_t(\delta))^2 \right]}{\epsilon^2 t^2 p^2}. \]

Since \( \tilde{z}_0(\delta) = 0 \) and due to Martingale property we have
\[ E \left[ (\tilde{z}_t(\delta))^2 | \mathcal{F}_0 \right] = E \left[ \left( \sum_{k=1}^{t} (\tilde{z}_k(\delta) - \tilde{z}_{k-1}(\delta)) \right)^2 | \mathcal{F}_0 \right] = \]

96
\[
\sum_{k=1}^{t} E \left[ (\tilde{z}_k(\delta) - \tilde{z}_{k-1}(\delta))^2 \mid \mathcal{F}_0 \right] + 2 \sum_{k<t} E \left[ \left( (\tilde{z}_k(\delta) - \tilde{z}_{k-1}(\delta)) \left( \tilde{z}_t(\delta) - \tilde{z}_{t-1}(\delta) \right) \right) \mid \mathcal{F}_k \right] \mid \mathcal{F}_0 \\
\leq t \max_{k=1, \ldots, t} \left\{ E \left[ (\tilde{z}_k(\delta) - \tilde{z}_{k-1}(\delta))^2 \mid \mathcal{F}_0 \right] \right\}.
\]

Now by definitions (4.9), (4.8) and (4.6) and condition (4.3)

\[
E \left[ (\tilde{z}_k(\delta) - \tilde{z}_{k-1}(\delta))^2 \right] = E \left[ \left( \tilde{Y}_k(\delta) - \tilde{Y}_{k-1}(\delta) - E \left[ \tilde{Y}_k(\delta) - \tilde{Y}_{k-1}(\delta) \mid \mathcal{F}_{k-1} \right] \right)^2 \right] =
\]

\[
Var \left[ \log \left( x_{k-1}(\delta^{opt})' \mathbf{R}_k \right) - \log \left( x_{k-1}(\delta)' \mathbf{R}_k \right) \right] \leq 4\sigma_{up}^2.
\]

Consequently,

\[
P \left( \left| \tilde{z}_t(\delta) \right| \geq \frac{ctp}{2} \right) \leq \frac{16 \sigma_{up}^2 t}{e^{2} p^2} = \frac{16 \sigma_{up}^2}{e^{2} p^2} \left( \frac{1}{t} \right).
\]

(4.11)

Since \( \bar{M}_t(\delta) \geq ctp \) (4.10) and \( P \left( \tilde{z}_t(\delta) \leq -\frac{ctp}{2} \right) \leq \frac{16 \sigma_{up}^2}{e^{2} p^2} \left( \frac{1}{t} \right) \) (4.11), we have

\[
P \left( \tilde{Y}_t(\delta) \geq ctp - \frac{ctp}{2} \right) \geq 1 - \frac{16 \sigma_{up}^2}{e^{2} p^2} \left( \frac{1}{t} \right).
\]

Thus

\[
P \left( \log W_t(\delta^{opt}) - \log W_t(\delta) \geq \frac{ep}{2} \right) \geq 1 - \frac{16 \sigma_{up}^2}{e^{2} p^2} \left( \frac{1}{t} \right)
\]

and

\[
P \left( \frac{W_t(\delta^{opt})}{W_t(\delta)} \geq \exp \left\{ \frac{ep}{2} t \right\} \right) \geq 1 - \frac{16 \sigma_{up}^2}{e^{2} p^2} \left( \frac{1}{t} \right).
\]

Let us give an example. Suppose that policy 1 is the myopic policy which maximizes \( E[\log(W_{t+1} \mid \mathcal{F}_t)] \) at each time \( t = 0, 1, \ldots, T-1 \). Policy 2 maximizes the \( q \)-th moment of the terminal wealth, that is \( E \left[ W_T^q \mid \mathcal{F}_t \right] \) at \( t = 0, 1, 2 \ldots, T-1 \). Take \( q = 0.8 \) and the horizon \( T = 80 \).

Assume that at each time we have a choice between a stock and a bond. The stock follows a Geometric Brownian motion process with a drift \( \mu_1 = 0.12 \) and variance \( \sigma_1^2 = 0.16 \), while the bond has \( \mu_2 = 0.085 \) and \( \sigma_2^2 = 0 \).

The price of the stock at each time is given by

\[
S(t) = S(0) \exp \left( \mu_1 t - \frac{\sigma_1^2}{2} t + \sigma_1 B(t) \right).
\]

(4.12)
Policy 2 at each time \( t \) maximizes

\[
E[S^q(T)\mid S(t)] = S(t) \exp \left\{ \left( \mu_1 - \frac{\sigma_1^2}{2} \right)(T - t)q \right\} \mathbb{E} \left[ e^{qT - tN(0,1)} \right]
\]

\[
= S(t) \exp \left( (T - t)q \left( \mu_1 - \frac{\sigma_1^2}{2}(1 - q) \right) \right)
\]  \hspace{1cm} (4.13)

Since

\[
\mu_1 - \frac{\sigma_1^2}{2}(1 - q) = 0.12 - 0.08 \cdot 0.2 = 0.104,
\]

\[
\mu_2 - \frac{\sigma_2^2}{2}(1 - q) = \mu_2 = 0.085.
\]

then policy 2 will choose to invest in the stock at each time \( t \). According to policy 1, however, we should always invest in the bond, since \( \mu_1 - \frac{\sigma_1^2}{2} = 0.12 - 0.08 = 0.04 < 0.085 = \mu_2 \).

Let us compare the distributions of the terminal wealth in cases that policy 1 or policy 2 are applied. Take \( W(0) = \$1 \). Since

\[
\left( \mu_1 - \frac{\sigma_1^2}{2} \right)T = 0.04 \times 80 = 3.2,
\]

\[
\sigma_1 \times \sqrt{T} \approx 0.4 \times 8.9 \approx 3.56,
\]

and

\[
P \left( B(t) \leq \sqrt{t} \right) \approx 0.84,
\]

then

\[
P \left( (\mu_1 - \sigma_1^2/2) \cdot T + \sigma_1 B(t) \leq 3.2 + 3.56 = 6.76 \right) \approx 0.84,
\]

so with probability 84% we end up with less than \$824.2 = \exp(6.76)\) if we invest in the stock (with 50% chance we get less than \$24.0 = \exp(3.2))\). If we invest in the bond we end up with \$857.6 = \exp(\mu_2 T) = \exp(6.8)\) with probability one. Thus, an investment in the bond seems to be much more beneficial, than the investment in the stock, while according to maximization of the \( q \)-th moment, our choice is the opposite.
4.4 Multiperiod mean-variance optimization

Let us introduce the idea of a coherent policy in the case of the mean-variance objective function. The coherent mean-variance policy is the policy, that at time $T - 1$ maximizes

$$E[W_T | \mathcal{F}_{T-1}] - \gamma \text{Var}[W_T | \mathcal{F}_{T-1}],$$

(4.14)

at time $T - 2$ maximizes

$$E[W_T | \mathcal{F}_{T-2}] - \gamma \text{Var}[W_T | \mathcal{F}_{T-2}],$$

(4.15)

given that at time $T - 1$ the policy maximizing (4.14) will be applied and so on by recursion. We will be interested in finding this optimal policy.

We consider a market with $N$ risky securities and one risk-free security, at which funds can be borrowed or lent. Reinvestment can be done at times $\{0, 1, 2, \ldots, T\}$ and withdrawals or capital additions are ruled out. Short sales are allowed.

We will use the following notations:

(a) Let $r_f$ denote $1 + r$, where $r$ is a risk-free return for $[t-1, t]$ for each $t$. Let $r_t = R_t - r_f$ denote the vector of excess stocks returns to the risk-free return for the period $[t-1, t]$.

(b) Designate by $z_t$ an $\mathcal{F}_t$-measurable vector which totally determines the conditional distribution of $\{r_{kt}\}_{k \geq t}$, conditioned on $\mathcal{F}_t$. We assume that $z_t$ exists for each time $t$.

(c) Let $u_t$ be the vector with components representing the amounts of wealth being invested into different risky assets at time $t$. Since short sales are allowed, $u^i_t$, $i = 1, \ldots, N$, can also be negative.

(d) Designate by $\tilde{r}_t := r_t - r_f^{T-t}$ the vector of excess returns discounted to the end of the period $T$.

(e) Let $\Sigma_{t+1}$ be the covariance matrix of $\tilde{r}_{t+1}$ conditioned on $z_t$.

Then the recursive formula for wealth $W_t$ is

$$W_{t+1} = \left( W_t - \sum_{i=1}^{N} u^i_t \right) r_f + (u^i_t R_{t+1}).$$
4.4.1 Algorithm

The following algorithm provides a consistent estimator of the optimal policy $u_t$ at time $t$.

(a) Since $z_t$ is known at time $t$, let us assume that $z_t = z_t^0$. Simulate $k_1$ i.i.d. random variables $z_{t+1}^{(m)}$, $m = 1, 2, \ldots, k_1$, each with a distribution $\mathcal{L}(z_{t+1}^m | z_t = z_t^0)$.

(b) For each $z_{t+1}^{(m)}$ simulate $k_2$ i.i.d. paths

$$Z^{(ml)} := (z_{t+2}^{(ml)}, z_{t+3}^{(ml)}, \ldots, z_T^{(ml)}), \quad l = 1, 2, \ldots, k_2$$

with a distribution $\mathcal{L}(z_{t+2}, z_{t+3}, \ldots, z_T | z_{t+1} = z_{t+1}^{(m)})$.

(c) For each sequence $Z^{(ml)}$ find the corresponding sequence $\{a_s^{(ml)}, b_s^{(ml)}\}_{s=t+2, \ldots, T}$ where

$$a_{s+1}^{(ml)} = \frac{1}{2\gamma} \left( \sum_{s=s+1}^{T-1} E \left[ \hat{r}_{s+1} | z_s = z_s^{(ml)} \right] \right)' E \left[ \hat{r}_{s+1} | z_s = z_s^{(ml)} \right]$$

and

$$b_{s+1}^{(ml)} = 1 - \left( \sum_{s=s+1}^{T-1} E \left[ \hat{r}_{s+1} | z_s = z_s^{(m0)} \right] \right)' E \left[ \hat{r}_{s+1} | z_s = z_s^{(ml)} \right]$$

(d) For each sequence $\{a_s^{(ml)}, b_s^{(ml)}\}_{s=t+2}$ calculate

$$\tilde{\Phi}_{t+1}^{(ml)} := a_{t+2}^{(ml)} + b_{t+2}^{(ml)} \left( a_{t+3}^{(ml)} + b_{t+3}^{(ml)} \left( \cdots \left( a_{T-1}^{(ml)} + b_{T-1}^{(ml)} \left( a_T^{(ml)} \right) \right) \right) \right)$$

and evaluate

$$\hat{\Phi}_{t+1}^{(m)} := \frac{\tilde{\Phi}_{t+1}^{(ml)} + \tilde{\Phi}_{t+1}^{(m2)} + \cdots + \tilde{\Phi}_{t+1}^{(mk_2)}}{k_2}. \quad (4.16)$$

(e) The estimator of the optimal policy is then given by

$$\hat{u}_t (z_t^0) = \frac{\sum_{s=t+1}^{T-1} E \left[ \hat{r}_{t+1} | z_t^0 \right]}{2\gamma} \left\{ \Phi_{t+1}^{(ml)} \left( \hat{r}_{t+1} \left( z_{t+1}^{(ml)} \right) - E \left[ \hat{r}_{t+1} | z_t^0 \right] \right) | z_t^0 \right\}$$

**Remark.** The choice of $k_1$ and $k_2$ depends on the required precision of the solution.
4.4.2 Example.

Let \( \{r_t\} \) be a second-order autoregressive process

\[
    r_t = \phi_1 r_{t-1} + \phi_2 r_{t-2} + \varepsilon_t,
\]

with \( \{\varepsilon_t : t = 0, \pm 1, \pm 2, \ldots\} \) being a sequence of independent normal random variables with \( E[\varepsilon_t] = 0 \) and \( Var[\varepsilon_t] = \sigma^2 \). Suppose \( T = \varepsilon \) and rebalancing of the portfolio happens at times 0, 1, 2, 3. We assume also that \( r_0, r_{-1}, r_{-2}, \ldots \) are known. Then the algorithm goes as follows.

(a) Simulate \( k_1 \) random variables \( \varepsilon_1^{(m)} \) from a distribution \( N(0, \sigma) \). Find

\[
    r_1^{(m)} := \phi_1 r_0 + \phi_2 r_{-1} + \varepsilon_1^{(m)}.
\]

(b) For each \( \varepsilon_1^{(m)} \) simulate \( k_2 \) independent 2-dimensional random vectors \( (\varepsilon_2^{(ml)}, \varepsilon_3^{(ml)}) \), with components being independent and distributed as \( N(0, \sigma) \).

(c) For each vector find

\[
    r_2^{(ml)} := \phi_1 r_1^{(ml)} + \phi_2 r_0 + \varepsilon_2^{(ml)}, \quad r_3^{(ml)} := \phi_1 r_2^{(ml)} + \phi_2 r_1^{(ml)} + \varepsilon_3^{(ml)},
\]

\[
    a_2^{(ml)} := \frac{1}{2\gamma \sigma^2} \left( \phi_1 r_1^{(ml)} + \phi_2 r_0 \right), \quad b_2^{(ml)} := 1 - \frac{\varepsilon_2^{(ml)}}{\sigma} \left( \phi_1 r_1^{(ml)} + \phi_2 r_0 \right),
\]

\[
    a_3^{(ml)} := \frac{1}{2\gamma \sigma^2} \left( \phi_1 r_2^{(ml)} + \phi_2 r_1^{(ml)} \right), \quad b_3^{(ml)} := 1 - \frac{\varepsilon_3^{(ml)}}{\sigma} \left( \phi_1 r_2^{(ml)} + \phi_2 r_1^{(ml)} \right),
\]

\[
    a_4^{(ml)} := \frac{1}{2\gamma \sigma^2} \left( \phi_1 r_3^{(ml)} + \phi_2 r_2^{(ml)} \right).
\]

(d) Evaluate \( \tilde{\Phi}_1^{(ml)} := a_2^{(ml)} + b_2^{(ml)} \left( a_3^{(ml)} + b_3^{(ml)} a_4^{(ml)} \right) \) and \( \tilde{\Phi}_1^{(m)} := \frac{1}{k_2} \sum_{l=1}^{k_2} \tilde{\Phi}_1^{(ml)} \).

(e) Calculate \( \tilde{u}_0 := \frac{1}{2\gamma \sigma^2} \left( \phi_1 r_0 + \phi_2 r_{-1} - 2\gamma \sum_{m=1}^{k_1} \tilde{\Phi}_1^{(m)} \varepsilon_1^{(m)} \right) \).

4.4.3 Analysis

Let us prove that the above algorithm leads to a consistent estimator of the policy.

We will use the following lemma later

**Lemma 1.** The optimal policy is separable in wealth.

**Proof.** Let \( e \) be an \( N \)-dimensional (\( N \) is the number of stocks) unit vector. Then
we can write a recursive formula for the amount of wealth at time $t$ as

$$W_{t+1} = (W_t - (u_t' e)) r_f + (u_t' R_{t+1}) = (W_t - (u_t' e)) r_f + u_t' (r_{t+1} + \epsilon r_f) = W_t r_f + (u_t' r_{t+1}).$$

Thus by recursion

$$W_{t_n} = W_0 r_f^{n-1} + (u_0' r_1) r_f^{n-2} + \cdots + (u_{n-1}' r_n).$$

We are going to prove that optimal investments into risky assets at time $t_0$ do not depend on the initial wealth $W_0$. This then will imply the same result for each time $t_i$. Let $u_1, u_2, \ldots, u_{n-1}$ be optimal controls at times $t_1, t_2, \ldots, t_{n-1}$ respectively. An optimal control at time $t_0$ is $x$ which maximizes

$$E \left[ W_0 r_f^{n-1} + (x' r_1) r_f^{n-2} + (u_1' r_2) r_f^{n-3} + \cdots + (u_{n-1}' r_n) \right] - \gamma Var \left[ W_0 r_f^{n-1} + (x' r_1) r_f^{n-2} + \cdots + (u_{n-1}' r_n) \right].$$

Thus the optimal control at time $t_0$ does not depend on $W_0$.

Let us first determine an optimal control at time $t$, if we assume, that future optimal controls $u_{t+1}(z_{t+1}), u_{t+2}(z_{t+2}), \ldots, u_{T-1}(z_{T-1})$ are known functions of information vectors $z_{t+1}, z_{t+2}, \ldots, z_{T-1}$. During the period $(t, T]$ we gain

$$(x' r_{t+1}) r_f^{T-t-1} + \sum_{s=t+1}^{T-1} (u_s' r_{s+1}) r_f^{T-s-1},$$

by making new investments to risky assets. $x$ here denotes investments made at time $t$. Define

$$W_T^{t+1} := \sum_{s=t+1}^{T-1} (u_s' r_{s+1}) r_f^{T-s-1}. \quad (4.17)$$

Then the optimization problem is to find $x$ which maximizes

$$E \left[ (x' r_{t+1}) r_f^{T-t-1} + W_T^{t+1} | z_t \right] - \gamma Var \left[ (x' r_{t+1}) r_f^{T-t-1} + W_T^{t+1} | z_t \right].$$

102
or with the notation \( \hat{\mathbf{r}}_t := \mathbf{r}_t \mathbf{r}_f^T \)

\[
E \left[ (\mathbf{x}' \hat{\mathbf{r}}_{t+1}) + W_{T+1}^{t+1} | \mathbf{z}_t \right] - \gamma Var \left[ (\mathbf{x}' \hat{\mathbf{r}}_{t+1}) + W_{T+1}^{t+1} | \mathbf{z}_t \right].
\]

The last is equivalent to maximizing

\[
E \left[ (\mathbf{x}' \hat{\mathbf{r}}_{t+1}) | \mathbf{z}_t \right] - \gamma Var \left[ (\mathbf{x}' \hat{\mathbf{r}}_{t+1}) | \mathbf{z}_t \right] - 2\gamma Cov \left[ (\mathbf{x}' \hat{\mathbf{r}}_{t+1}), W_{T+1}^{t+1} | \mathbf{z}_t \right]. \tag{4.18}
\]

By taking partial derivatives of (4.18) with respect to coordinates of \( \mathbf{x} \) to zero, we obtain a system of linear equations:

\[
E \left[ \hat{r}_t^i | \mathbf{z}_t \right] - 2\gamma \sum_{j=1}^N x_j Cov \left[ \hat{r}_t^i, \hat{r}_t^j | \mathbf{z}_t \right] - 2\gamma Cov \left[ \hat{r}_t^i, W_{T+1}^{t+1} | \mathbf{z}_t \right] = 0, \tag{4.19}
\]

\[
i = 1, \ldots, N,
\]

where \( \hat{r}_t^i \) is the \( i \)-th coordinate of \( \hat{r}_t \). We can rewrite (4.19) as

\[
E \left[ \hat{r}_{t+1} | \mathbf{z}_t \right] - 2\gamma \begin{pmatrix}
Cov \left[ \hat{r}_t^1, W_{T+1}^{t+1} | \mathbf{z}_t \right] \\
Cov \left[ \hat{r}_t^2, W_{T+1}^{t+1} | \mathbf{z}_t \right] \\
\vdots \\
Cov \left[ \hat{r}_t^N, W_{T+1}^{t+1} | \mathbf{z}_t \right]
\end{pmatrix} = 2\gamma \Sigma \hat{r}_{t+1} \mathbf{x}, \tag{4.20}
\]

Thus an optimal control at time \( t \) is given by

\[
\mathbf{u}_t = \frac{\Sigma_{t+1}^{-1}}{2\gamma} \begin{pmatrix}
\mathbf{E} \left[ \hat{r}_{t+1} | \mathbf{z}_t \right] - 2\gamma \begin{pmatrix}
Cov \left[ \hat{r}_t^1, W_{T+1}^{t+1} | \mathbf{z}_t \right] \\
Cov \left[ \hat{r}_t^2, W_{T+1}^{t+1} | \mathbf{z}_t \right] \\
\vdots \\
Cov \left[ \hat{r}_t^N, W_{T+1}^{t+1} | \mathbf{z}_t \right]
\end{pmatrix}
\end{pmatrix}.
\]

We can rewrite the last expression as

\[
\mathbf{u}_t = \frac{\Sigma_{t+1}^{-1}}{2\gamma} \left\{ E \left[ \hat{r}_{t+1} | \mathbf{z}_t \right] - 2\gamma E \left[ W_{T+1}^{t+1} (\hat{r}_{t+1} - \mathbf{E} \left[ \hat{r}_{t+1} | \mathbf{z}_t \right]) | \mathbf{z}_t \right] \right\}
\]

\[
= \frac{\Sigma_{t+1}^{-1}}{2\gamma} \left\{ E \left[ \hat{r}_{t+1} | \mathbf{z}_t \right] - 2\gamma E \left[ E \left[ W_{T+1}^{t+1} | \mathbf{z}_{t+1} \right] (\hat{r}_{t+1} - \mathbf{E} \left[ \hat{r}_{t+1} | \mathbf{z}_t \right]) | \mathbf{z}_t \right] \right\}. \tag{4.21}
\]
Here \( E \left[ W_{T+1}^t | z_{t+1} \right] \) only depends on \( z_{t+1} \). Let us define

\[
\Phi_{t+1}(z_{t+1}) := E \left[ W_{T+1}^t | z_{t+1} \right].
\]  

(4.22)

Thus \( \Phi(.) \) is a nonrandom function of \( z_t \). In new terms, (4.21) becomes

\[
u_t = \frac{\Sigma_{t+1}^{-1}}{2\gamma} \left\{ E \left[ \hat{r}_{t+1} | z_t \right] - 2\gamma E \left[ \Phi_{t+1}(z_{t+1}) \left( \hat{r}_{t+1} - E \left[ \hat{r}_{t+1} | z_t \right] \right) | z_t \right] \right\}
\]

(4.23)

If we knew \( \Phi_{t+1}(z_{t+1}) \) it would be easy to estimate \( u_t^i(z_t) \). Namely, we would simulate a necessary, for a given precision, number \( k_1 \) of i.i.d. random variables \( z_{t+1}^{(1)}, z_{t+1}^{(2)}, \ldots, z_{t+1}^{(m)} \), each with a distribution \( \mathcal{L}(z_{t+1}|z_t) \) and take as a consistent estimator

\[
\hat{u}_t(z_t^0) = \frac{\Sigma_{t+1}^{-1}}{2\gamma} \left\{ E \left[ \hat{r}_{t+1} | z_t^0 \right] - 2\gamma \sum_{m=1}^{k_1} \left[ \Phi_{t+1} \left( z_{t+1}^{(m)} \right) \left( \hat{r}_{t+1} \left( z_{t+1}^{(m)} \right) - E \left[ \hat{r}_{t+1} | z_t^0 \right] \right) \right] \right\}
\]

Notice that in case of known \( \Phi(.) \), even if \( z_t \) is a high dimensional vector, we can still easily calculate an optimal control, since the efficiency of Monte-Carlo simulation does not depend on the dimension of the space.

We are going to find an unbiased consistent estimator of function \( \Phi_t(.) \). First, let us find a recursive formula for \( \Phi_t(.) \). We have:

\[
\Phi_t(z_t) = E \left[ (u_t'\hat{r}_{t+1}) + W_{T+1}^t | z_t \right] = E \left[ (u_t'\hat{r}_{t+1}) + E \left[ W_{T+1}^t | z_{t+1} \right] | z_t \right] = 
\]

\[
E \left[ (u_t'\hat{r}_{t+1}) + \Phi_{t+1}(z_{t+1}) | z_t \right] = \frac{1}{2\gamma} \left( \left\{ \Sigma_{t+1}^{-1} E \left[ \hat{r}_{t+1} | z_t \right] \right\} ' E \left[ \hat{r}_{t+1} | z_t \right] \right) + 
\]

\[
E \left[ \Phi_{t+1}(z_{t+1}) \left( 1 - \left( \left\{ \Sigma_{t+1}^{-1} (\hat{r}_{t+1} - E \left[ \hat{r}_{t+1} | z_t \right]) \right\} ' E \left[ \hat{r}_{t+1} | z_t \right] \right) \right) \right] .
\]

Let

\[
a_{t+1} = \frac{1}{2\gamma} \left( \left\{ \Sigma_{t+1}^{-1} E \left[ \hat{r}_{t+1} | z_t \right] \right\} ' E \left[ \hat{r}_{t+1} | z_t \right] \right),
\]  

(4.24)

\[
b_{t+1} = 1 - \left( \left\{ \Sigma_{t+1}^{-1} (\hat{r}_{t+1} - E \left[ \hat{r}_{t+1} | z_t \right]) \right\} ' E \left[ \hat{r}_{t+1} | z_t \right] \right).
\]  

(4.25)
With these definitions we obtain

\[ \Phi_t = a_{t+1} + E \left[ b_{t+1} \Phi_{T-1} | z_t \right]. \]

Notice, that \( a_{t+1} \) and \( b_{t+1} \) are \( \mathcal{F}_{t+1} \) measurable, even more, they are known once \( z_t \) and \( z_{t+1} \) are known:

\[ a_{t+1} = a_{t+1}(z_t), \]
\[ b_{t+1} = b_{t+1}(z_t, z_{t+1}). \]

We also have

\[ u_{T-1}(z_{T-1}) = \frac{\Sigma_{t+1}^{-1}}{2\gamma} E \left[ \hat{x}_T | z_{T-1} \right], \]

\[ \Phi_{T-1}(z_{T-1}) = E \left[ W_T^{-1} | z_{T-1} \right] = E \left[ (u_{T-1}' \hat{x}_T) | z_{T-1} \right] = \frac{1}{2\gamma} \left( \Sigma_{T-1}^{-1} E \left[ \hat{x}_T | z_{T-1} \right] \right)' E \left[ \hat{x}_T | z_{T-1} \right]. \]

Thus

\[ \Phi_{T-1} = a_T, \quad (4.26) \]

\[ \Phi_s = a_{s+1} + E \left[ b_{s+1} \Phi_{s+1} | z_s \right], \quad s = T - 2, T - 3, \ldots, t \quad (4.27) \]

is a recursive formula for \( \Phi \). Define

\[ \tilde{\Phi}_t = a_{t+1} + b_{t+1} (\cdots (a_{T-3} + b_{T-3}(a_{T-2} + b_{T-2}(a_{T-1} + b_{T-1}(a_T)))). \quad (4.28) \]

**Theorem 3.**

\[ E \left[ \tilde{\Phi}_t | z_t = z^0_t \right] = \Phi_t (z^0_t). \quad (4.29) \]

**Proof.** By recursion

\[ E \left[ \tilde{\Phi}_t | \mathcal{F}_t \right] = E \left[ \cdots E \left[ E \left[ \tilde{\Phi}_t | \mathcal{F}_{T-1} \right] | \mathcal{F}_{T-2} \right] \cdots | \mathcal{F}_t \right] \]

105
Now $a_T, a_{T-1}, a_{T-2}, \ldots, a_t$ and $b_{T-1}, b_{T-2}, \ldots, b_t$ are $\mathcal{F}_{T-1}$ measurable, thus the above is equivalent to

$$E[\cdots E[a_{t+1} + b_{t+1} (\cdots (a_{T-2} + b_{T-2}(a_{T-1} + b_{T-1}a_T)))] | \mathcal{F}_{T-2}] | \cdots | \mathcal{F}_t]$$

Taking into account that all the coefficients above are $\mathcal{F}_{T-2}$ measurable except $a_{T-1} + b_{T-1}a_T$ and that $E[a_{T-1} + b_{T-1}a_T | \mathcal{F}_{T-2}] = \Phi_{T-2}$, we obtain

$$E\left[\Phi_t | \mathcal{F}_t\right] = E[\cdots E[a_{t+1} + b_{t+1} (\cdots (a_{T-2} + b_{T-2}\Phi_{T-2}))] | \mathcal{F}_{T-3}] | \cdots | \mathcal{F}_t].$$

Again, taking into account that in the above expression all the coefficients are $\mathcal{F}_{T-3}$ measurable except $a_{T-2} + b_{T-2}\Phi_{T-2}$ and that $E[a_{T-2} + b_{T-2}\Phi_{T-2} | \mathcal{F}_{T-3}] = \Phi_{T-3}$, we have

$$E\left[\tilde{\Phi}_t | \mathcal{F}_t\right] = E[\cdots E[a_{t+1} + b_{t+1} (\cdots (a_{T-3} + b_{T-3}\Phi_{T-3}))] | \mathcal{F}_{T-4}] | \cdots | \mathcal{F}_t].$$

Following the same procedure inductively we conclude that

$$E\left[\tilde{\Phi}_t | \mathcal{F}_t\right] = E\left[\tilde{\Phi}_t | z_t = z_0^t\right] = \Phi_t(z_0^t).$$

If components of $\tilde{r}_t$ are uncorrelated conditioned on $z_{t-1}$, all the formulas become much simpler. Notice that if the covariance matrix of $[\tilde{r}_t|z_{t-1}]$ does not depend on $z_{t-1}$, then by the Graham-Schmidt procedure it is easy to make the coordinates uncorrelated. Let us give formulas in this case

$$u_i^t = \frac{E[\tilde{r}_{i+1}^t | z_t] - 2\gamma E[\Phi_{t+1}(z_{t+1}) (\tilde{r}_{i+1}^t - E[\tilde{r}_{i+1}^t | z_t])] | z_t]}{2\gamma Var[\tilde{r}_{i+1}^t | z_t]}.$$  \hspace{2cm} (4.30)

$$a_{t+1} := \sum_{i=1}^{N} \left( \frac{E^2[\tilde{r}_{i+1}^t | z_t]}{2\gamma Var[\tilde{r}_{i+1}^t | z_t]} \right)$$  \hspace{2cm} (4.31)

$$b_{t+1} := 1 - \sum_{i=1}^{N} \left( \frac{\tilde{r}_{i+1}^t - E[\tilde{r}_{i+1}^t | z_t]}{Var[\tilde{r}_{i+1}^t(z_t)]} \right) E[\tilde{r}_{i+1}^t | z_t].$$  \hspace{2cm} (4.32)
\[ \Phi_t = a_{t+1} + E \left[ b_{t+1} \Phi_{T+1} | z_t \right]. \]

\[ \Phi_{T-1} = a_T. \]

Thus \( \bar{\Phi}_t(z_t^0) \) is an unbiased estimator of \( \Phi_t(z_t^0) \). However the variance of this estimator might be large. So in order to reduce it, we would have to simulate a sufficient number \( N \) of paths and to take an average of the \( N \) unbiased estimators, each of which is determined by a corresponding path. The average will also be an unbiased estimator of \( \Phi_t(z_t^0) \) and its variance decreases with \( N \) as \( \frac{1}{\sqrt{N}} \), since paths are independent conditioned on \( z_t^0 \).

If we compare this method with a dynamic programming (backward recursion) method, we will see that the recursion method actually also takes an average of \( \bar{\Phi} \)'s. However \( \bar{\Phi} \)'s in the recursion method are not independent (while in our method they are), and thus the variance of the average does not get reduced enough and the method is not efficient.

Notice that we assume that \( a_t \)'s and \( b_t \)'s are known once \( z_t \) and \( z_{t-1} \) are known. If we also have to estimate \( a_t \)'s and \( b_t \)'s, then it might introduce a small bias. However, since for a given \( z_t \) and \( z_{t-1} \), it is not a problem to estimate \( a_t \) and \( b_t \) as precisely as we want, we can assume that we use the exact \( a_t \)'s and \( b_t \)'s in the calculation of \( \bar{\Phi}_t \).

Let us evaluate how many estimators \( \bar{\Phi} \) one should take in order to obtain a certain degree of precision. If \( X_1, X_2, \ldots, X_n \) are i.i.d. random variables with \( Var[X_1] < \infty \) and \( E[X_1] = \mu \), then by Chebyshev inequality

\[
P \left( \left| \frac{X_1 + X_2 + \ldots + X_n}{n} - \mu \right| \geq \epsilon \right) \leq \frac{E \left[ \left( \frac{X_1 + X_2 + \ldots + X_n}{n} - \mu \right)^2 \right]}{\epsilon^2} = \frac{Var[X_1]}{\epsilon^2 n}. \tag{4.33} \]

Thus the number of estimators \( \bar{\Phi} \), that one should take for a precision \( \epsilon \) is of order of \( \frac{Var[\bar{\Phi}]}{\epsilon^2} \). So in order to know how many simulations of paths we will need, we must have an idea how big \( Var[\bar{\Phi}] \) is. One approach might be to simulate several paths and estimate \( Var[\bar{\Phi}] \) from there. However, this approach might fail to work which is why it is necessary to find also a theoretical bound on \( Var[\bar{\Phi}] \).

Indeed, consider a random variable \( X_1 \) which is 0 with probability 0.9999 and 10,000 with probability 0.0001. Then \( E[X_1] = 1 \) and \( Var[X_1] = 9,999 \). However if we simulate about 100 i.i.d. variables \( X_1, X_2, \ldots, X_{100} \), then we will most likely only
get zeros. Thus, we will conclude that $E[X_1] = 0$ and $Var[X_1] = 0$. This is why we need to have a theoretical bound on the variance of $X_1$ to be sure that $\frac{X_1 + \ldots + X_n}{n}$ is a reliable estimate of $E[X_1]$.

Assume, that for any $z_t$ in the feasible domain, $t = T - m, T - m + 1, \ldots, T$, the following condition is satisfied for some $C \geq 0$:

$$\left( \left\{ \sum_{t+1}^{t+1} E[\hat{r}_{t+1} | z_t] \right\}^T E[\hat{r}_{t+1} | z_t] \right) \leq C$$  \hspace{1cm} (4.34)

Notice that when coordinates of $r_t$ are uncorrelated, as $C$ we can take $C_1 N$, where $C_1$ is such that for any $z_t$ in the feasible domain, $t = T - m, T - m + 1, \ldots, T$, and any $i = 1, \ldots, N$

$$\frac{E^2[\hat{r}_{t+1}^i | z_t]}{Var[\hat{r}_{t+1}^i | z_t]} \leq C_1.$$  \hspace{1cm} (4.35)

**Theorem 4.** If condition (4.34) holds, then

$$Var[\Phi_{T-m} | z_{T-m-1}] \leq m^2 \frac{C^2}{4 \gamma^2} (1 + C)^{m-1}.$$  \hspace{1cm} (4.36)

**Proof.** From (4.28) follows

$$\Phi_{T-m} = b_{T-m+1} \ldots b_{T-4} b_{T-3} b_{T-2} b_{T-1} a_T + b_{T-m+1} \ldots b_{T-4} b_{T-3} b_{T-2} a_{T-1} + b_{T-m+1} \ldots b_{T-4} b_{T-3} a_{T-2} + \ldots \ldots + a_{T-m+1}.$$
Since \( a_T \leq \frac{C}{2\gamma} \), we have

\[
\begin{align*}
\text{Var} \left[ b_{T-m+1} \ldots b_{T-3} b_{T-2} b_{T-1} a_T \mid z_{T-m} \right] \leq \\
E \left[ b_{T-m+1}^2 \ldots b_{T-3}^2 b_{T-2}^2 b_{T-1}^2 a_T^2 \mid z_{T-m} \right] \leq \\
\frac{C^2}{4\gamma^2} E \left[ b_{T-m+1}^2 \ldots b_{T-3}^2 b_{T-2}^2 b_{T-1}^2 \mid z_{T-m} \right] = \\
\frac{C^2}{4\gamma^2} E \left[ E \left[ b_{T-m+1}^2 \ldots b_{T-3}^2 b_{T-2}^2 b_{T-1}^2 \mid \mathcal{F}_{T-2} \right] \mid \mathcal{F}_{T-m} \right] = \\
\frac{C^2}{4\gamma^2} E \left[ b_{T-m+1}^2 \ldots b_{T-3}^2 b_{T-2}^2 \ E \left[ b_T^2 \mid \mathcal{F}_{T-2} \right] \mid \mathcal{F}_{T-m} \right]. \quad (4.38)
\end{align*}
\]

The last equality holds since \( b_{T-m-1}, \ldots, b_{T-2} \) are all \( \mathcal{F}_{T-2} \) measurable. Now from the definition of \( b_t \) (4.25)

\[
E \left[ b_{t+1}^2 \mid z_t \right] = E \left[ \left( 1 - \left( \Sigma_{\hat{r}_{t+1}^{-1}} (\hat{r}_{t+1} - E[\hat{r}_{t+1} \mid z_t]) \right)' E[\hat{r}_{t+1} \mid z_t] \right)^2 \mid z_t \right] = \\
E \left[ 1 + \left( \Sigma_{\hat{r}_{t+1}^{-1}} (\hat{r}_{t+1} - E[\hat{r}_{t+1} \mid z_t]) \right)' E[\hat{r}_{t+1} \mid z_t] \right]^2 \mid z_t, \\
\text{since} \\
E \left[ \left( \Sigma_{\hat{r}_{t+1}^{-1}} (\hat{r}_{t+1} - E[\hat{r}_{t+1} \mid z_t]) \right)' E[\hat{r}_{t+1} \mid z_t] \right] \mid z_t = 0.
\]

Now

\[
E \left[ \left( \Sigma_{\hat{r}_{t+1}^{-1}} (\hat{r}_{t+1} - E[\hat{r}_{t+1} \mid z_t]) \right)' E[\hat{r}_{t+1} \mid z_t] \right]^2 \mid z_t = \\
E \left[ \left( E[\hat{r}_{t+1} \mid z_t]' \Sigma_{\hat{r}_{t+1}^{-1}} (\hat{r}_{t+1} - E[\hat{r}_{t+1} \mid z_t]) \right)^2 \mid z_t \right] = \\
E \left[ E[\hat{r}_{t+1} \mid z_t]' \Sigma_{\hat{r}_{t+1}^{-1}} (\hat{r}_{t+1} - E[\hat{r}_{t+1} \mid z_t]) \right] \left( \Sigma_{\hat{r}_{t+1}}^{-1} \right)^T E[\hat{r}_{t+1} \mid z_t] \left( \Sigma_{\hat{r}_{t+1}}^{-1} \right)^T E[\hat{r}_{t+1} \mid z_t] = \\
E[\hat{r}_{t+1} \mid z_t]' \Sigma_{\hat{r}_{t+1}}^{-1} \ E \left[ \left( \hat{r}_{t+1} - E[\hat{r}_{t+1} \mid z_t] \right)' \hat{r}_{t+1} - E[\hat{r}_{t+1} \mid z_t] \right)' \left( \Sigma_{\hat{r}_{t+1}}^{-1} \right)^T E[\hat{r}_{t+1} \mid z_t] = 
\]

109
\[
E [\hat{r}_{t+1} | z_t] \Sigma_{\hat{r}_{t+1}}^{-1} \Sigma_{\hat{r}_{t+1}} \left( \Sigma_{\hat{r}_{t+1}}^{-1} \right)^T E [\hat{r}_{t+1} | z_t] = \left( \left\{ \Sigma_{\hat{r}_{t+1}}^{-1} E [\hat{r}_{t+1} | z_t] \right\} \right)^T E [\hat{r}_{t+1} | z_t] \leq C
\]

Thus

\[
E \left[ b_{t+1}^2 | z_t \right] \leq 1 + C \tag{4.39}
\]

and going back to (4.38) we see that

\[
E \left[ b_{T-m+1}^2 \cdots b_{T-3}^2 b_{T-2}^2 b_{T-1}^2 a_T^2 | z_{T-m} \right] \leq \frac{C^2}{4\gamma^2} (1 + C) E \left[ b_{T-m+1}^2 \cdots b_{T-3}^2 b_{T-2}^2 | z_{T-m} \right].
\]

By applying the same technique, that is, successively conditioning on \( \mathcal{F}_{T-2}, \mathcal{F}_{T-3}, \ldots \) up to \( \mathcal{F}_{T-m} \) we eventually obtain

\[
E \left[ b_{T-m+1}^2 \cdots b_{T-3}^2 b_{T-2}^2 b_{T-1}^2 a_T^2 | z_{T-m} \right] \leq \frac{C^2}{4\gamma^2} (1 + C)^{m-1} \tag{4.40}
\]

Let us now apply this to the estimation of the variance of \( \tilde{\Phi}_{T-m} \)

\[
\text{Var} \left[ \tilde{\Phi}_{T-m} | z_{T-m} \right] \leq E \left[ \tilde{\Phi}_{T-m}^2 | z_{T-m} \right] =
\]

\[
E \left[ \begin{bmatrix}
  b_{T-m+1} \cdots b_{T-4} b_{T-3} b_{T-2} b_{T-1} a_T + \\
  b_{T-m+1} \cdots b_{T-4} b_{T-3} b_{T-2} a_{T-1} + \\
  b_{T-m+1} \cdots b_{T-4} b_{T-3} a_{T-2} + \\
  \cdots \\
  a_{T-m+1}
\end{bmatrix}^2 | z_{T-m+1} \right].
\]

By expanding the expression in parenthesis, we get \( m^2 \) terms, and for each of these terms, call them \( c_k, k = 1, \ldots, m^2 \), we have

\[
E [c_k | z_{T-m}] \leq E \left[ b_{T-m+1}^2 \cdots b_{T-3}^2 b_{T-2}^2 b_{T-1}^2 a_T^2 | z_{T-m} \right] \leq \frac{C^2}{4\gamma^2} (1 + C)^{m-1}
\]

**Remark.** It is important to note that
the variance's bound is much below the exponential in the number of securities. As we can see, it actually is polynomial in \( N \) (however if \( m \) is big, then the power is big too;)

- the bound is totally independent of the dimension of the information vector \( \mathbf{z}_t \).

Thus, even if we use much more predictive variables for this algorithm, this does not create a problem.

Let us consider a particular example. Assume that we have 50 stocks and we can rebalance once a month over a year. Then if the standard deviation of the return is 30\% per year, then the variance is 0.09, and the variance per month is

\[
\frac{0.09}{12} = 0.0075.
\]

Also if we take the expected excess return of 10\% per year then

\[
C_1 = \frac{(0.1/12)^2}{0.0075} \approx 0.0092, \quad NC = 0.46.
\]

Thus, if the risk aversion is, for example, 0.3, we get

\[
Var \left[ \tilde{\Phi}_{t_1} | \mathbf{z}_{t_1} \right] \leq m^2 \frac{N^2 C^2}{4 \gamma^2} (1 + NC)^{m-1} = 144 \frac{(0.46)^2}{4 \times 0.09} (1 + 0.46)^{11}
\approx 400 \times 0.0324 \times 94 \approx 7956. \quad (4.41)
\]

Remark. A question remains, if a better upper bound on the variance of \( \Phi \) can be found.

4.4.4 Numerical experiments

Let us consider the case, when returns to risky assets follow the discretized version of Ornstein-Uhlenbeck process

\[
\mathbf{r}_{n+1} = \mathbf{r}_n + \beta (\bar{\mathbf{r}} - \mathbf{r}_n) + \epsilon_n.
\]

Here average returns

\[
\bar{\mathbf{r}} = (0.12, 0.06, 0.08, 0.07, 0.16, 0.13, 0.06, 0.20, 0.18, 0.15)
\]
are given, as well as the covariance matrix of $\epsilon_n$. The standard deviation is

$$
\sigma = (0.09, 0.05, 0.06, 0.05, 0.12, 0.10, 0.06, 0.15, 0.14, 0.12)
$$

and the correlation matrix is represented in Table 4.1.

Without loss of generality, the risk-free interest rate is taken to be zero. The risk-aversion coefficient $\gamma$ is taken 0.5 and parameter $\beta$ in Ornstein-Uhlenbeck equation is 0.6.

<table>
<thead>
<tr>
<th></th>
<th>1.00</th>
<th>0.15</th>
<th>0.19</th>
<th>-0.13</th>
<th>-0.19</th>
<th>0.07</th>
<th>-0.42</th>
<th>-0.27</th>
<th>-0.18</th>
<th>0.29</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.15</td>
<td>1.00</td>
<td>0.31</td>
<td>-0.24</td>
<td>0.04</td>
<td>-0.21</td>
<td>0.42</td>
<td>-0.45</td>
<td>-0.19</td>
<td>-0.42</td>
<td></td>
</tr>
<tr>
<td>0.19</td>
<td>0.31</td>
<td>1.00</td>
<td>-0.08</td>
<td>0.27</td>
<td>0.05</td>
<td>0.07</td>
<td>-0.16</td>
<td>0.25</td>
<td>-0.29</td>
<td></td>
</tr>
<tr>
<td>-0.13</td>
<td>-0.24</td>
<td>-0.08</td>
<td>1.00</td>
<td>0.03</td>
<td>-0.23</td>
<td>0.26</td>
<td>0.46</td>
<td>0.15</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>-0.19</td>
<td>0.04</td>
<td>0.27</td>
<td>0.03</td>
<td>1.00</td>
<td>-0.34</td>
<td>0.06</td>
<td>0.08</td>
<td>-0.46</td>
<td>0.29</td>
<td></td>
</tr>
<tr>
<td>0.07</td>
<td>-0.21</td>
<td>0.05</td>
<td>-0.23</td>
<td>-0.34</td>
<td>1.00</td>
<td>-0.41</td>
<td>0.31</td>
<td>0.26</td>
<td>-0.24</td>
<td></td>
</tr>
<tr>
<td>-0.42</td>
<td>0.42</td>
<td>0.07</td>
<td>0.26</td>
<td>0.06</td>
<td>-0.41</td>
<td>1.00</td>
<td>-0.28</td>
<td>-0.06</td>
<td>-0.18</td>
<td></td>
</tr>
<tr>
<td>-0.27</td>
<td>-0.45</td>
<td>-0.17</td>
<td>0.46</td>
<td>0.08</td>
<td>0.31</td>
<td>-0.28</td>
<td>1.00</td>
<td>0.25</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>-0.18</td>
<td>-0.19</td>
<td>0.25</td>
<td>0.15</td>
<td>-0.46</td>
<td>0.26</td>
<td>-0.06</td>
<td>0.25</td>
<td>1.00</td>
<td>-0.39</td>
<td></td>
</tr>
<tr>
<td>0.29</td>
<td>-0.42</td>
<td>-0.29</td>
<td>0.08</td>
<td>0.29</td>
<td>-0.24</td>
<td>-0.19</td>
<td>0.09</td>
<td>-0.39</td>
<td>1.00</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.1: Correlation Matrix.

Since the key issue in calculating an optimal policy at time $t$ is the estimation of $\Phi_{t+1}$, we give an example of how many paths we might need to generate in order to estimate $\Phi$ with a certain precision. Precision is measured as the standard error $s_\Phi$ divided by the estimator of $\Phi$, $\hat{\Phi}$. Table 4.2 gives an estimation of $\Phi_1$ for different numbers of stocks and the number of periods being 5. Table 4.3 gives an estimation of $\Phi_1$ for different numbers of periods for 5 stocks.

At last, Table 4.4 compares relative errors for standard dynamic programming method and our method. As, we can see, for $8 \cdot 10^3$ paths, the relative error is about 20 times larger for the dynamic programming method. Moreover, when the number of paths is increased to $125 \cdot 10^3$, the relative error for the dynamic programming method becomes about 45 times larger. Thus, as expected, the error decreases with the increase of number of paths much faster, if independent paths are generated.
<table>
<thead>
<tr>
<th>N stocks</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\Phi}_1$</td>
<td>42.6</td>
<td>52.3</td>
<td>98.6</td>
<td>122.4</td>
<td>240.5</td>
<td>301.9</td>
<td>345.4</td>
<td>612.7</td>
<td>660.2</td>
</tr>
<tr>
<td>$\frac{s^{\hat{\Phi}_1}/\hat{\Phi}_1}{10^{-2}}$</td>
<td>0.7</td>
<td>0.8</td>
<td>0.9</td>
<td>0.9</td>
<td>1.1</td>
<td>1.2</td>
<td>1.2</td>
<td>1.3</td>
<td>1.3</td>
</tr>
<tr>
<td>N paths</td>
<td>$1 \cdot 10^4$</td>
<td>$1 \cdot 10^4$</td>
<td>$2 \cdot 10^4$</td>
<td>$3 \cdot 10^4$</td>
<td>$5 \cdot 10^4$</td>
<td>$7 \cdot 10^4$</td>
<td>$9 \cdot 10^4$</td>
<td>$10 \cdot 10^4$</td>
<td>$10 \cdot 10^4$</td>
</tr>
</tbody>
</table>

Table 4.2: Estimation of $\Phi_1$ for different number of risky assets and number of periods being 5.

<table>
<thead>
<tr>
<th>N periods</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\Phi}_1$</td>
<td>32.8</td>
<td>63.7</td>
<td>92.9</td>
<td>121.3</td>
<td>148.1</td>
</tr>
<tr>
<td>$\frac{s^{\hat{\Phi}_1}/\hat{\Phi}_1}{10^{-2}}$</td>
<td>0</td>
<td>0.8</td>
<td>0.9</td>
<td>0.9</td>
<td>1.0</td>
</tr>
<tr>
<td>N paths</td>
<td>$1 \cdot 5 \cdot 10^3$</td>
<td>$15 \cdot 10^3$</td>
<td>$30 \cdot 10^3$</td>
<td>$50 \cdot 10^3$</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.3: Estimation of $\Phi_1$ for different number of periods, and the number of stocks being 5.

<table>
<thead>
<tr>
<th>number of stocks</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>number of paths</td>
<td>$8 \cdot 10^3$</td>
<td>$8 \cdot 10^3$</td>
<td>$8 \cdot 10^3$</td>
<td>$8 \cdot 10^5$</td>
</tr>
<tr>
<td>$\frac{s^{\hat{\Phi}_1}/\hat{\Phi}_1}{10^{-2}}$ (our method)</td>
<td>0.6</td>
<td>0.7</td>
<td>1.0</td>
<td>1.2</td>
</tr>
<tr>
<td>$\frac{s^{\hat{\Phi}_1}/\hat{\Phi}_1}{10^{-2}}$ (dynamic programming)</td>
<td>12.2</td>
<td>14.3</td>
<td>18.5</td>
<td>20.1</td>
</tr>
<tr>
<td>number of paths</td>
<td>$125 \cdot 10^3$</td>
<td>$125 \cdot 10^3$</td>
<td>$125 \cdot 10^3$</td>
<td>$125 \cdot 10^3$</td>
</tr>
<tr>
<td>$\frac{s^{\hat{\Phi}_1}/\hat{\Phi}_1}{10^{-2}}$ (our method)</td>
<td>0.15</td>
<td>0.18</td>
<td>0.26</td>
<td>0.30</td>
</tr>
<tr>
<td>$\frac{s^{\hat{\Phi}_1}/\hat{\Phi}_1}{10^{-2}}$ (dynamic programming)</td>
<td>6.8</td>
<td>8.7</td>
<td>11.4</td>
<td>13.2</td>
</tr>
</tbody>
</table>

Table 4.4: Comparison of relative errors in $\Phi$-estimators, in cases of applying standard dynamic programming method and our method.
Bibliography


