Chapter 3

Log-Optimal Portfolio Selection Strategies with Proportional Transaction Costs

László Györfi

Department of Computer Science and Information Theory, Budapest University of Technology and Economics. H-1117, Magyar tudósok körútja 2., Budapest, Hungary , gyorfi@shannon.szit.bme.hu

Harro Walk

Institute of Stochastics and Applications, Universität Stuttgart, Pfaffenwaldring 57, 70569 Stuttgart, Germany. walk@mathematik.uni-stuttgart.de

Discrete time growth optimal investment in stock markets with proportional transactions costs is considered. The market process is a sequence of daily relative prices (called returns), and it is modelled by a first order Markov process. Assuming that the distribution of the market process is known, we show sequential investment strategies such that, in the long run, the growth rate on trajectories achieves the maximum with probability 1. Investment with consumption and with fixed transaction cost where the cost depends on the number of the shares involved in the transaction is also analyzed.

3.1. Introduction

The purpose of this chapter is to investigate sequential investment strategies for financial markets such that the strategies are allowed to use information collected from the past of the market and determine, at the beginning of a trading period, a portfolio, that is, a way to distribute their current capital among the available assets. The goal of the investor is to maximize his wealth on the long run. If there is no transaction cost then the only assumption used in the mathematical analysis is that the daily price relatives form a stationary and ergodic process. Under this assumption the best strategy (called log-optimum strategy) can be constructed in full knowledge of the distribution of the entire process, see [Algoet and Cover (1988)]. Moreover, [Györfi and Schäfer (2003)], [Györfi *et al.* (2006)] and [Györfi *et al.* (2008)] constructed empirical (data driven) growth optimum strategies in case of unknown distributions. The empirical results show that the performance of these empirical investment strategies measured on past NYSE data is solid, and sometimes even spectacular.

The problem of optimal investment with proportional transaction cost has been essentially formulated and studied in continuous time only (cf. [Akien *et al.* (2001)], [Davis and Norman (1990)], [Eastham and Hastings (1988)], [Korn (1998)], [Morton and Pliska (1995)], [Palczewski and Stettner (2006)], [Pliska and Suzuki (2004)], [Shreve *et al.* (1991)], [Shreve and Soner (1994)], [Taksar *et al.* (1988)]).

Papers dealing with growth optimal investment with transaction costs in discrete time setting are seldom. [Iyengar and Cover (2000)] formulated the problem of horse race markets, where in every market period one of the assets has positive pay off and all the others pay nothing. Their model included proportional transaction costs and they used a long run expected average reward criterion. There are results for more general markets as well. [Sass and Schäl (2010)] investigated the numeraire portfolio in context of bond and stock as assets. [Iyengar (2002, 2005)] investigated growth optimal investment with several assets assuming independent and identically distributed (i.i.d.) sequence of asset returns. [Bobryk and Stettner (1999)] considered the case of portfolio selection with consumption, when there are two assets, a bond and a stock. Furthermore, long run expected discounted reward and i.i.d asset returns were assumed. In the case of discrete time and non i.i.d. market process, [Schäfer (2002)] considered the maximization of the long run expected average growth rate with several assets and proportional transaction costs, when the asset returns follow a stationary Markov process. [Györfi and Vajda (2008)] extended the expected growth optimality mentioned above to the almost sure (a.s.) growth optimality.

In this chapter we study the problem of discrete time growth optimal investment in stock markets with proportional, fixed transactions costs and consumption. In Section 3.2 the mathematical setup is introduced. Section 3.3 shows the empirical simulated results of two heuristic algorithms using NYSE data. If the market process is first order Markov process and the distribution of the market process is known, then we show simple sequential investment strategies such that, in the long run, the growth rate on trajectories achieves the maximum with probability 1 in Section 3.4 and Section

3.6 (Proofs). Finally Section 3.5 studies the portfolio selection strategies with consumption and fixed transaction cost.

3.2. Mathematical setup: investment with proportional transaction cost

Consider a market consisting of d assets. The evolution of the market in time is represented by a sequence of market vectors $\mathbf{s}_1, \mathbf{s}_2, \ldots \in \mathbb{R}_+^d$, where

$$\mathbf{s}_i = (s_i^{(1)}, \ldots, s_i^{(d)})$$

such that the j-th component $s_i^{(j)}$ of $\mathbf{s}_i$ denotes the price of the j-th asset at the end of the i-th trading period. ($s_0^{(j)} = 1$.)

In order to apply the usual prediction techniques for time series analysis one has to transform the sequence $\{\mathbf{s}_i\}$ into a sequence of return vectors $\{\mathbf{x}_i\}$ as follows:

$$\mathbf{x}_i = (x_i^{(1)}, \ldots, x_i^{(d)})$$

such that

$$x_i^{(j)} = \frac{s_i^{(j)}}{s_{i-1}^{(j)}}.$$

Thus, the j-th component $x_i^{(j)}$ of the return vector $\mathbf{x}_i$ denotes the amount obtained at the end of the i-th trading period after investing a unit capital in the j-th asset.

The investor is allowed to diversify his capital at the beginning of each trading period according to a portfolio vector $\mathbf{b} = (b^{(1)}, \ldots b^{(d)})^T$. The j-th component $b^{(j)}$ of $\mathbf{b}$ denotes the proportion of the investor's capital invested in asset j. Throughout the chapter we assume that the portfolio vector $\mathbf{b}$ has nonnegative components with $\sum_{j=1}^d b^{(j)} = 1$. The fact that $\sum_{j=1}^d b^{(j)} = 1$ means that the investment strategy is self financing and consumption of capital is excluded (besides Section 3.5). The non-negativity of the components of $\mathbf{b}$ means that short selling and buying stocks on margin are not permitted. To make the analysis feasible, some simplifying assumptions are used that need to be taken into account. We assume that assets are arbitrarily divisible and all assets are available in unbounded quantities at the current price at any given trading period. We also assume that the behavior of the market is not affected by the actions of the investor using the strategies under investigation.

For $j \leq i$ we abbreviate by $\mathbf{x}_j^i$ the array of return vectors $(\mathbf{x}_j, \ldots, \mathbf{x}_i)$. Denote by Δ_d the simplex of all vectors $\mathbf{b} \in \mathbb{R}_+^d$ with nonnegative components summing up to one. An *investment strategy* is a sequence $\mathbf{B}$ of functions

$$\mathbf{b}_i : \left(\mathbb{R}_+^d\right)^{i-1} \to \Delta_d \ , \qquad i = 1, 2, \ldots$$

so that $\mathbf{b}_i(\mathbf{x}_1^{i-1})$ denotes the portfolio vector chosen by the investor on the i-th trading period, upon observing the past behavior of the market. We write $\mathbf{b}(\mathbf{x}_1^{i-1}) = \mathbf{b}_i(\mathbf{x}_1^{i-1})$ to ease the notation.

In this section our presentation of the transaction cost problem utilized the formulation in [Kalai and Blum (1997)] and [Schäfer (2002)] and [Györfi and Vajda (2008)]. Let S_n denote the gross wealth at the end of trading period n, $n = 0, 1, 2, \cdots$, where without loss of generality let the investor's initial capital S_0 be 1 dollar, while N_n stands for the net wealth at the end of trading period n. Using the above notations, for the trading period n, the net wealth N_{n-1} can be invested according to the portfolio $\mathbf{b}_n$, therefore the gross wealth S_n at the end of trading period n is

$$S_n = N_{n-1} \sum_{j=1}^{d} b_n^{(j)} x_n^{(j)} = N_{n-1} \langle \mathbf{b}_n \, , \, \mathbf{x}_n \rangle \, ,$$

where $\langle \cdot \, , \, \cdot \rangle$ denotes inner product.

At the beginning of a new market day $n + 1$, the investor sets up his new portfolio, i.e. buys/sells stocks according to the actual portfolio vector $\mathbf{b}_{n+1}$. During this rearrangement, he has to pay transaction cost, therefore at the beginning of a new market day $n + 1$ the net wealth N_n in the portfolio $\mathbf{b}_{n+1}$ is less than S_n.

The rate of proportional transaction costs (commission factors) levied on one asset are denoted by $0 < c_s < 1$ and $0 < c_p < 1$, i.e., the sale of 1 dollar worth of asset i nets only $1 - c_s$ dollars, and similarly we take into account the purchase of an asset such that the purchase of 1 dollar's worth of asset i costs an extra c_p dollars. We consider the special case when the rate of costs are constant over the assets.

Let's calculate the transaction cost to be paid when select the portfolio $\mathbf{b}_{n+1}$. Before rearranging the capitals, at the j-th asset there are $b_n^{(j)} x_n^{(j)} N_{n-1}$ dollars, while after rearranging we need $b_{n+1}^{(j)} N_n$ dollars. If $b_n^{(j)} x_n^{(j)} N_{n-1} \geq b_{n+1}^{(j)} N_n$ then we have to sell and the transaction cost at the j-th asset is

$$c_s \left(b_n^{(j)} x_n^{(j)} N_{n-1} - b_{n+1}^{(j)} N_n \right) ,$$

otherwise we have to buy and the transaction cost at the j-th asset is

$$c_p\left(b_{n+1}^{(j)}N_n - b_n^{(j)}x_n^{(j)}N_{n-1}\right).$$

Let x^+ denote the positive part of x. Thus, the gross wealth S_n decomposes to the sum of the net wealth and cost in the following - self-financing - way

$$\begin{aligned} N_n = S_n &- \sum_{j=1}^{d} c_s\left(b_n^{(j)}x_n^{(j)}N_{n-1} - b_{n+1}^{(j)}N_n\right)^+ \\ &- \sum_{j=1}^{d} c_p\left(b_{n+1}^{(j)}N_n - b_n^{(j)}x_n^{(j)}N_{n-1}\right)^+, \end{aligned}$$

or equivalently

$$\begin{aligned} S_n = N_n &+ c_s\sum_{j=1}^{d} \left(b_n^{(j)}x_n^{(j)}N_{n-1} - b_{n+1}^{(j)}N_n\right)^+ \\ &+ c_p\sum_{j=1}^{d} \left(b_{n+1}^{(j)}N_n - b_n^{(j)}x_n^{(j)}N_{n-1}\right)^+. \end{aligned}$$

Dividing both sides by S_n and introducing ratio

$$w_n = \frac{N_n}{S_n},$$

$0 < w_n < 1$, we get

$$\begin{aligned} 1 = w_n &+ c_s\sum_{j=1}^{d} \left(\frac{b_n^{(j)}x_n^{(j)}}{\langle \mathbf{b}_n\,,\,\mathbf{x}_n\rangle} - b_{n+1}^{(j)}w_n\right)^+ \\ &+ c_p\sum_{j=1}^{d} \left(b_{n+1}^{(j)}w_n - \frac{b_n^{(j)}x_n^{(j)}}{\langle \mathbf{b}_n\,,\,\mathbf{x}_n\rangle}\right)^+. \end{aligned} \tag{3.1}$$

For given previous return vector $\mathbf{x}_n$ and portfolio vector $\mathbf{b}_n$, there is a portfolio vector $\tilde{\mathbf{b}}_{n+1} = \tilde{\mathbf{b}}_{n+1}(\mathbf{b}_n, \mathbf{x}_n)$ for which there is no trading:

$$\tilde{b}_{n+1}^{(j)} = \frac{b_n^{(j)}x_n^{(j)}}{\langle \mathbf{b}_n\,,\,\mathbf{x}_n\rangle} \tag{3.2}$$

such that there is no transaction cost, i.e., $w_n = 1$.

For arbitrary portfolio vectors $\mathbf{b}_n$, $\mathbf{b}_{n+1}$, and return vector $\mathbf{x}_n$ there exist unique cost factors $w_n \in [0, 1)$, i.e., the portfolio is self financing. The value of cost factor w_n at day n is determined by portfolio vectors $\mathbf{b}_n$ and $\mathbf{b}_{n+1}$ as well as by return vector $\mathbf{x}_n$, i.e.

$$w_n = w(\mathbf{b}_n, \mathbf{b}_{n+1}, \mathbf{x}_n),$$

for some function w. If we want to rearrange our portfolio substantially, then our net wealth decreases more considerably, however, it remains positive. Note also, that the cost does not restrict the set of new portfolio vectors, i.e., the optimization algorithm searches for optimal vector $\mathbf{b}_{n+1}$ within the whole simplex Δ_d. The value of the cost factor ranges between

$$\frac{1-c_s}{1+c_p} \le w_n \le 1.$$

Without loss of generality we consider the special case of $c_s = c_p =: c$. Then

$$\begin{aligned} & c_s \left(b_n^{(j)} x_n^{(j)} N_{n-1} - b_{n+1}^{(j)} N_n \right)^+ + c_p \left(b_{n+1}^{(j)} N_n - b_n^{(j)} x_n^{(j)} N_{n-1} \right)^+ \\ &= c \left| b_n^{(j)} x_n^{(j)} N_{n-1} - b_{n+1}^{(j)} N_n \right| . \end{aligned}$$

Starting with an initial wealth $S_0 = 1$ and $w_0 = 1$, wealth S_n at the closing time of the n-th market day becomes

$$\begin{aligned} S_n &= N_{n-1} \langle \mathbf{b}_n \, , \, \mathbf{x}_n \rangle \\ &= w_{n-1} S_{n-1} \langle \mathbf{b}_n \, , \, \mathbf{x}_n \rangle \\ &= \prod_{i=1}^n [w(\mathbf{b}_{i-1}, \mathbf{b}_i, \mathbf{x}_{i-1}) \, \langle \mathbf{b}_i \, , \, \mathbf{x}_i \rangle]. \end{aligned}$$

Introduce the notation

$$g(\mathbf{b}_{i-1}, \mathbf{b}_i, \mathbf{x}_{i-1}, \mathbf{x}_i) = \log(w(\mathbf{b}_{i-1}, \mathbf{b}_i, \mathbf{x}_{i-1}) \, \langle \mathbf{b}_i \, , \, \mathbf{x}_i \rangle),$$

then the average growth rate becomes

$$\begin{aligned} \frac{1}{n} \log S_n &= \frac{1}{n} \sum_{i=1}^n \log(w(\mathbf{b}_{i-1}, \mathbf{b}_i, \mathbf{x}_{i-1}) \, \langle \mathbf{b}_i \, , \, \mathbf{x}_i \rangle) \\ &= \frac{1}{n} \sum_{i=1}^n g(\mathbf{b}_{i-1}, \mathbf{b}_i, \mathbf{x}_{i-1}, \mathbf{x}_i). \end{aligned} \tag{3.3}$$

Our aim is to maximize this average growth rate.

In the sequel $\mathbf{x}_i$ will be random variable and is denoted by $\mathbf{X}_i$, and we assume the following
Conditions:

(i) $\{\mathbf{X}_i\}$ is a homogeneous and first order Markov process,

(ii) the Markov kernel is continuous, which means that for $\mu(B|\mathbf{x})$ being the Markov kernel defined by

$$\mu(B|\mathbf{x}) := \mathbb{P}\{\mathbf{X}_2 \in B \mid \mathbf{X}_1 = \mathbf{x}\}$$

we assume that the Markov kernel is continuous in total variation, i.e.,

$$V(\mathbf{x}, \mathbf{x}') := \sup_{B \in \mathcal{B}} |\mu(B|\mathbf{x}) - \mu(B|\mathbf{x}')| \to 0$$

if $\mathbf{x}' \to \mathbf{x}$ such that $\mathcal{B}$ denotes the family of Borel σ-algebra, further

$$V(\mathbf{x}, \mathbf{x}') < 1 \text{ for all } \mathbf{x}, \mathbf{x}',$$

(iii) and there exist $0 < a_1 < 1 < a_2 < \infty$ such that $a_1 \leq X^{(j)} \leq a_2$ for all $j = 1, \ldots, d$.

We note that Conditions (ii) and (iii) imply uniform continuity of V and thus

$$\max_{\mathbf{x}, \mathbf{x}'} V(\mathbf{x}, \mathbf{x}') < 1. \tag{3.4}$$

For the usual stock market daily data, Condition (iii) is satisfied with $a_1 = 0.7$ and with $a_2 = 1.2$ (cf. [Fernholz (2000)]).

In the realistic case that the state space of the Markov process $(\mathbf{X}_n)$ is a finite set D of rational vectors (components being quotients of integer-valued \$-amounts) containing $\mathbf{e} = (1, \ldots, 1)$, the second part of (ii) is fulfilled under the plausible assumption $\mu(\{\mathbf{e}\}|\mathbf{x}) > 0$ for all $\mathbf{x} \in D$. Another example for finite state Markov process is when one rounds down the components of $\mathbf{x}$ to a grid applying, for example, a grid size 0.00001.

Let's use the decomposition

$$\frac{1}{n} \log S_n = I_n + J_n, \tag{3.5}$$

where I_n is

$$\frac{1}{n} \sum_{i=1}^{n} (g(\mathbf{b}_{i-1}, \mathbf{b}_i, \mathbf{X}_{i-1}, \mathbf{X}_i) - \mathbb{E}\{g(\mathbf{b}_{i-1}, \mathbf{b}_i, \mathbf{X}_{i-1}, \mathbf{X}_i)|\mathbf{X}_1^{i-1}\})$$

and

$$J_n = \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}\{g(\mathbf{b}_{i-1}, \mathbf{b}_i, \mathbf{X}_{i-1}, \mathbf{X}_i)|\mathbf{X}_1^{i-1}\}.$$

I_n is an average of martingale differences. Under the condition (iii), the random variable $g(\mathbf{b}_{i-1}, \mathbf{b}_i, \mathbf{X}_{i-1}, \mathbf{X}_i)$ is bounded, therefore I_n is an average of bounded martingale differences, which converges to 0 almost surely, since according to the Chow Theorem (cf. Theorem 3.3.1 in [Stout (1974)])

$$\sum_{i=1}^{\infty} \frac{\mathbb{E}\{g(\mathbf{b}_{i-1}, \mathbf{b}_i, \mathbf{X}_{i-1}, \mathbf{X}_i)^2\}}{i^2} < \infty$$

implies that

$$I_n \to 0$$

almost surely. Thus, the asymptotic maximization of the average growth rate $\frac{1}{n} \log S_n$ is equivalent to the maximization of J_n.

Under the condition (i), we have that

$$\begin{aligned} & \mathbb{E}\{g(\mathbf{b}_{i-1}, \mathbf{b}_i, \mathbf{X}_{i-1}, \mathbf{X}_i) | \mathbf{X}_1^{i-1}\} \\ = \; & \mathbb{E}\{\log(w(\mathbf{b}_{i-1}, \mathbf{b}_i, \mathbf{X}_{i-1}) \langle \mathbf{b}_i \, , \, \mathbf{X}_i \rangle) | \mathbf{X}_1^{i-1}\} \\ = \; & \log w(\mathbf{b}_{i-1}, \mathbf{b}_i, \mathbf{X}_{i-1}) + \mathbb{E}\{\log \langle \mathbf{b}_i \, , \, \mathbf{X}_i \rangle \, | \mathbf{X}_1^{i-1}\} \\ = \; & \log w(\mathbf{b}_{i-1}, \mathbf{b}_i, \mathbf{X}_{i-1}) + \mathbb{E}\{\log \langle \mathbf{b}_i \, , \, \mathbf{X}_i \rangle \, | \mathbf{b}_i, \mathbf{X}_{i-1}\} \\ \stackrel{\text{def}}{=} \; & v(\mathbf{b}_{i-1}, \mathbf{b}_i, \mathbf{X}_{i-1}), \end{aligned}$$

therefore the maximization of the average growth rate $\frac{1}{n} \log S_n$ is asymptotically equivalent to the maximization of

$$J_n = \frac{1}{n} \sum_{i=1}^{n} v(\mathbf{b}_{i-1}, \mathbf{b}_i, \mathbf{X}_{i-1}). \tag{3.6}$$

The terms in the average J_n have a memory, which transforms the problem into a dynamic programming setup (cf. [Merhav *et al.* (2002)]).

3.3. Experiments on heuristic algorithms

In this section we experimentally study two heuristic algorithms, which performed well without transaction cost (cf. Chapter 2 of this volume).

Algorithm 1. For transaction cost, one may apply the log-optimal portfolio

$$\mathbf{b}_n^*(\mathbf{X}_{n-1}) = \arg\max_{\mathbf{b}(\cdot)} \mathbb{E}\{\ln \langle \mathbf{b}(\mathbf{X}_{n-1}) \, , \, \mathbf{X}_n \rangle \mid \mathbf{X}_{n-1}\}$$

or its empirical approximation. For example, we may apply the kernel based log-optimal portfolio selection introduced by [Györfi *et al.* (2006)] as follows: Define an infinite array of experts $\mathbf{B}^{(\ell)} = \{\mathbf{b}^{(\ell)}(\cdot)\}$, where ℓ is a positive integer. For fixed positive integer ℓ, choose the radius $r_\ell > 0$ such that

$$\lim_{\ell\to\infty} r_\ell = 0.$$

Then, for $n > 1$, define the expert $\mathbf{b}^{(\ell)}$ as follows. Put

$$\mathbf{b}_n^{(\ell)} = \arg\max_{\mathbf{b}\in\Delta_d} \sum_{\{i<n:\|\mathbf{x}_{i-1}-\mathbf{x}_{n-1}\|\le r_\ell\}} \ln\langle \mathbf{b}\,,\, \mathbf{x}_i\rangle\ , \tag{3.7}$$

if the sum is non-void, and $\mathbf{b}_0 = (1/d, \ldots, 1/d)$ otherwise, where $\|\cdot\|$ denotes the Euclidean norm.

Similarly to Chapter 2 of this volume, these experts are aggregated (mixed) as follows: let $\{q_\ell\}$ be a probability distribution over the set of all positive integers ℓ such that for all ℓ, $q_\ell > 0$. Consider two types of aggregations:

- Here the initial capital $S_0 = 1$ is distributed among the expert according to the distribution $\{q_\ell\}$, and the expert makes the portfolio selection and pays for transaction cost individually. If $S_n(\mathbf{B}^{(\ell)})$ is the capital accumulated by the elementary strategy $\mathbf{B}^{(\ell)}$ after n periods when starting with an initial capital $S_0 = 1$, then, after period n, the investor's aggregated wealth after period n is

$$S_n = \sum_\ell q_\ell S_n(\mathbf{B}^{(\ell)}). \tag{3.8}$$

- Here $S_n(\mathbf{B}^{(\ell)})$ is again the capital accumulated by the elementary strategy $\mathbf{B}^{(\ell)}$ after n periods when starting with an initial capital $S_0 = 1$, but it is virtual figure, i.e., the experts make no trading, its wealth is just the base of aggregation. Then, after period n, the investor's aggregated portfolio becomes

$$\mathbf{b}_n = \frac{\sum_\ell q_\ell S_{n-1}(\mathbf{B}^{(\ell)})\mathbf{b}_n^{(\ell)}}{\sum_\ell q_\ell S_{n-1}(\mathbf{B}^{(\ell)})}. \tag{3.9}$$

  Moreover, the investor's capital is

$$S_n = S_{n-1}\langle \mathbf{b}_n\,,\, \mathbf{x}_n\rangle w(\mathbf{b}_{n-1}, \mathbf{b}_n, \mathbf{x}_{n-1}),$$

  so only the aggregated portfolio pays for the transaction cost.

In Chapter 2 of this volume we proved that without transaction cost the two aggregations are equivalent. However, in case of transaction cost the aggregation (3.9) is much better.

Algorithm 2. We may introduce a suboptimal algorithm, called *naive portfolio*, by a one-step optimization as follows: put $\mathbf{b}_1 = \{1/d, \dots, 1/d\}$ and for $n \geq 1$,

$$\mathbf{b}_n^{(\ell)} = \arg\max_{\mathbf{b} \in \Delta_d} \sum_{\{i<n: \|\mathbf{x}_{i-1} - \mathbf{x}_{n-1}\| \leq r_\ell\}} \left(\ln \langle \mathbf{b}\, ,\, \mathbf{x}_i \rangle + \ln w(\mathbf{b}_{n-1}, \mathbf{b}, \mathbf{x}_{n-1}) \right), \tag{3.10}$$

if the sum is non-void, and $\mathbf{b}_0 = (1/d, \dots, 1/d)$ otherwise. These elementary portfolios are mixed as in (3.8) or (3.9). Obviously, this portfolio has no global optimality property.

Next we present some numerical results for transaction cost obtained by applying the kernel based semi-log-optimal algorithm to the 19 assets of the second NYSE data set as in Chapter 2 of this volume. We take a finite set of of experts of size L. In the experiment we selected $L = 10$. Choose the uniform distribution $q_\ell = 1/L$ over the experts in use, and the radius

$$r_\ell^2 = 0.0002 \cdot d(1 + \ell/10), \qquad \text{for } \ell = 1, \dots, L\ .$$

Table 3.1 summarizes the average annual yield achieved by each expert at the last period when investing one unit for the kernel-based log-optimal portfolio. Experts are indexed by $\ell = 1 \dots 10$ in rows. The second column contains the average annual yields of experts for kernel based log-optimal portfolio if there is no transaction cost, and in this case the results of the two aggregations are the same: 35%. Mention that, out of the 19 assets, MORRIS had the best average annual yield, 20%, so, for no transaction cost, with

Table 3.1. The average annual yields of the individual experts for kernel strategy and of the aggregations with $c = 0.0015$.

ℓ	$c = 0$	Algorithm 1	Algorithm 2
1	31%	-22%	18%
2	34%	-22%	10%
3	35%	-24%	9 %
4	35%	-23%	14%
5	34%	-21%	13%
6	35%	-19%	13%
7	33%	-20%	12%
8	34%	-18%	8 %
9	37%	-17%	6 %
10	34%	-18%	11%
Wealth Agg. (3.8)	35%	**-19%**	**13%**
Portfolio Agg. (3.9)	35%	**-15%**	**17%**

kernel based log-optimal portfolio we have a spectacular improvement. The third and fourth columns contain the average annual yields of experts for kernel based log-optimal portfolio if the commission factor is $c = 0.0015$. Notice that the growth rate of the Algorithm 1 is negative, and the growth rate of the Algorithm 2 is poor, too, it is less than the growth rate of the best asset, and the results of aggregations are different.

In Table 3.2 we have got similar results for nearest neighbor strategy, where ℓ is the number of nearest neighbors. As we mentioned in Chapter 2 of this volume, the time varying portfolio is very undiversified such that the subset of assets with non-zero weight is changing from time to time, which makes the problem of transaction cost challenging. Moreover, the better the nearest neighbor strategy is without transaction cost, the worse it is with transaction cost, and the main reasoning of this fact is that for the good time varying portfolio, the portfolio vector component is very fluctuating, and so the proper handling of the transaction cost is still an open question and an important direction of the further research.

3.4. Growth optimal portfolio selection algorithms

An essential tool in the definition and investigation of portfolio selection algorithms under transaction costs are optimality equations of Bellman type. First we present an informal and heuristic way to them in our context of portfolio selection. Later on a rigorous treatment will be given.

Table 3.2. The average annual yields of the individual experts for nearest neighbor strategy and of the aggregations with $c = 0.0015$.

ℓ	$c = 0$	Algorithm 1	Algorithm 2
50	31%	-35%	-14%
100	33%	-33%	3%
150	38%	-29%	3%
200	38%	-28%	9%
250	37%	-28%	9%
300	41%	-26%	7%
350	39%	-26%	9%
400	39%	-26%	10%
450	39%	-25%	14%
500	42%	-23%	14%
Wealth Agg. (3.8)	39%	**-25%**	**11%**
Portfolio Agg. (3.9)	39%	**-23%**	**11%**

Let us start with a finite-horizon problem concerning J_N defined by (3.6): For fixed integer $N > 0$, maximize

$$\mathbb{E}\{N \cdot J_N \mid \mathbf{b}_0 = \mathbf{b}, \mathbf{X}_0 = \mathbf{x}\} = \mathbb{E}\left\{\sum_{i=1}^{N} v(\mathbf{b}_{i-1}, \mathbf{b}_i, \mathbf{X}_{i-1}) \mid \mathbf{b}_0 = \mathbf{b}, \mathbf{X}_0 = \mathbf{x}\right\}$$

by suitable choice of $\mathbf{b}_1, \ldots, \mathbf{b}_N$. For general problems of dynamic programming (dynamic optimization), on page 89 [Bellman (1957)] formulates his famous principle of optimality as follows: "An optimality policy has the property that whatever the initial state and initial decisions are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision."

By this principle, which for stochastic models is not so obvious as it seems (cf. pp. 14, 15 in [Hinderer (1970)]), one can show the following. Let the functions $G_0, G_1, \ldots, G_N$ on $\Delta_d \times [a_1, a_2]^d$ be defined by the so-called dynamic programming equations (optimality equations, Bellman equations)

$$G_N(\mathbf{b}, \mathbf{x}) := 0,$$
$$G_n(\mathbf{b}, \mathbf{x}) := \max_{\mathbf{b}'} \left[v(\mathbf{b}, \mathbf{b}', \mathbf{x}) + \mathbb{E}\{G_{n+1}(\mathbf{b}', \mathbf{X}_2) \mid \mathbf{X}_1 = \mathbf{x}\}\right]$$

$(n = N-1, N-2, \ldots, 0)$ with maximizer $\mathbf{b}'_n = g_n(\mathbf{b}, \mathbf{x})$. Setting

$$F^n := G_{N-n}$$

$(n = 0, 1, \ldots, N)$, one can write these backward equations in the forward form

$$F^0(\mathbf{b}, \mathbf{x}) := 0,$$
$$F^n(\mathbf{b}, \mathbf{x}) := \max_{\mathbf{b}'} \left[v(\mathbf{b}, \mathbf{b}', \mathbf{x}) + \mathbb{E}\{F^{n-1}(\mathbf{b}', \mathbf{X}_2) \mid \mathbf{X}_1 = \mathbf{x}\}\right] \quad (3.11)$$

$(n = 1, 2, \ldots, N)$ with maximizer $f_n(\mathbf{b}, \mathbf{x}) = g_{N-n}(\mathbf{b}, \mathbf{x})$. Then the choices $\mathbf{b}_n = f_n(\mathbf{b}_{n-1}, \mathbf{X}_{n-1})$ are optimal.

For the situations, which are favorite for the investor, one has $F^n(\mathbf{b}, \mathbf{x}) \to \infty$ as $n \to \infty$, which does not allow distinguishing between the qualities of competing choice sequences in the infinite-horizon case. If one considers (3.11) as a Value Iteration formula, then the underlying Bellman type equation

$$F^\infty(\mathbf{b}, \mathbf{x}) = \max_{\mathbf{b}'} \{v(\mathbf{b}, \mathbf{b}', \mathbf{x}) + \mathbb{E}\{F^\infty(\mathbf{b}', \mathbf{X}_2) \mid \mathbf{X}_1 = \mathbf{x}\}\}$$

has, roughly speaking, the degenerate solution $F^\infty = \infty$. Therefore one uses a discount factor $0 < \delta < 1$ and arrives at the discounted Bellman equation

$$F_\delta(\mathbf{b},\mathbf{x}) = \max_{\mathbf{b}'} \{v(\mathbf{b},\mathbf{b}',\mathbf{x}) + (1-\delta)\mathbb{E}\{F_\delta(\mathbf{b}',\mathbf{X}_2) \mid \mathbf{X}_1 = \mathbf{x}\}\}. \tag{3.12}$$

Its solution allows to solve the discounted problem maximizing

$$\mathbb{E}\left\{\sum_{i=1}^{\infty}(1-\delta)^i v(\mathbf{b}_{i-1},\mathbf{b}_i,\mathbf{X}_{i-1}) \mid \mathbf{b}_0 = \mathbf{b}, \mathbf{X}_0 = \mathbf{x}\right\}$$
$$= \sum_{i=1}^{\infty}(1-\delta)^i \mathbb{E}\left\{v(\mathbf{b}_{i-1},\mathbf{b}_i,\mathbf{X}_{i-1}) \mid \mathbf{b}_0 = \mathbf{b}, \mathbf{X}_0 = \mathbf{x}\right\}.$$

The classic Hardy-Littlewood theorem (see, e.g., Theorem 95, together with Theorem 55 in [Hardy (1949)]) states that for a real valued bounded sequence a_n, $n = 1, 2, \dots$,

$$\lim_{\delta\downarrow 0} \delta \sum_{i=0}^{\infty}(1-\delta)^i a_i$$

exists if and only if

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} a_i$$

exists and that then the limits are equal. Therefore, for maximizing

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\left\{v(\mathbf{b}_{i-1},\mathbf{b}_i,\mathbf{X}_{i-1}) \mid \mathbf{b}_0 = \mathbf{b}, \mathbf{X}_0 = \mathbf{x}\right\},$$

(if it exists), it is important to solve the equation (3.12) for small δ. This principle results in Rule 1 below. Letting $\delta \downarrow 0$, (3.12) with solution F_δ^* leads to the non-discounted Bellman equation

$$\lambda + F(\mathbf{b},\mathbf{x}) = \max_{\mathbf{b}'} \{v(\mathbf{b},\mathbf{b}',\mathbf{x}) + \mathbb{E}\{F(\mathbf{b}',\mathbf{X}_2) \mid \mathbf{X}_1 = \mathbf{x}\}\}. \tag{3.13}$$

The interpretation of (3.11) as Value Iteration motivates solving (3.12) and (3.13) also by Value Iterations $F_{\delta,n}$ (see below) and F'_n with discount factors $\delta_n \downarrow 0$ (see Rule 4). As to the corresponding problems in Markov control theory we refer to [Hernández-Lerma and Lassere (1996)].

[Györfi and Vajda (2008)] studied the following two optimal portfolio selection rules. Let $0 < \delta < 1$ denote a discount factor. Let the discounted Bellman equation (3.12). One can show that this discounted Bellman equation (3.12) and also the more general Bellman equation (3.19) below, have a unique solution (cf. [Schäfer (2002)] and the proof of Proposition 3.1

below). Concerning the discounted Bellman equation (3.12), the so-called Value Iteration may result in the solution: for fixed $0 < \delta < 1$, put

$$F_{\delta,0} = 0$$

and

$$\begin{aligned}&F_{\delta,k+1}(\mathbf{b},\mathbf{x})\\ &= \max_{\mathbf{b}'} \left\{ v(\mathbf{b},\mathbf{b}',\mathbf{x}) + (1-\delta)\mathbb{E}\{F_{\delta,k}(\mathbf{b}',\mathbf{X}_2) \mid \mathbf{X}_1 = \mathbf{x}\} \right\},\end{aligned}$$

$k = 0, 1, \ldots$. Then Banach's fixed point theorem implies that the value iteration converges uniformly to the unique solution.

Rule 1. [Schäfer (2002)] introduced the following non-stationary rule. Put

$$\bar{\mathbf{b}}_1 = \{1/d, \ldots, 1/d\}$$

and

$$\bar{\mathbf{b}}_{i+1} = \arg\max_{\mathbf{b}'} \left\{ v(\bar{\mathbf{b}}_i,\mathbf{b}',\mathbf{X}_i) + (1-\delta_i)\mathbb{E}\{F_{\delta_i}(\mathbf{b}^{'},\mathbf{X}_{i+1})|\mathbf{X}_i\}\right\},$$

for $1 \le i$, where $0 < \delta_i < 1$ is a discount factor such that $\delta_i \downarrow 0$. [Schäfer (2002)] proved that for the conditions (i), (ii) (in a weakened form) and (iii) and under some mild conditions on δ_i's for Rule 1, the portfolio $\{\bar{\mathbf{b}}_i\}$ with capital $\bar{S}_n$ is optimal in the sense that for any portfolio strategy $\{\mathbf{b}_i\}$ with capital S_n,

$$\liminf_{n\to\infty} \left(\frac{1}{n}\mathbb{E}\{\log \bar{S}_n\} - \frac{1}{n}\mathbb{E}\{\log S_n\} \right) \ge 0.$$

[Györfi and Vajda (2008)] extended this optimality in expectation to path-wise optimality such that under the same conditions

$$\liminf_{n\to\infty} \left(\frac{1}{n}\log \bar{S}_n - \frac{1}{n}\log S_n \right) \ge 0$$

a.s.

Rule 2. [Györfi and Vajda (2008)] introduced a portfolio with *stationary* (time invariant) recursion. For any integer $1 \le k$, put

$$\mathbf{b}_1^{(k)} = \{1/d, \ldots, 1/d\}$$

and

$$\mathbf{b}_{i+1}^{(k)} = \arg\max_{\mathbf{b}'} \left\{ v(\mathbf{b}_i^{(k)},\mathbf{b}',\mathbf{X}_i) + (1-\delta_k)\mathbb{E}\{F_{\delta_k}(\mathbf{b}^{'},\mathbf{X}_{i+1})|\mathbf{X}_i\}\right\},$$

for $1 \le i$, where $0 < \delta_k < 1$. The portfolio $\mathbf{B}^{(k)} = \{\mathbf{b}_i^{(k)}\}$ is called the portfolio of expert k with capital $S_n(\mathbf{B}^{(k)})$. Choose an arbitrary probability distribution $q_k > 0$, and introduce the combined portfolio with its capital

$$\tilde{S}_n = \sum_{k=1}^{\infty} q_k S_n(\mathbf{B}^{(k)}).$$

[Györfi and Vajda (2008)] proved that under the above mentioned conditions, for Rule 2,

$$\lim_{n\to\infty} \left(\frac{1}{n} \log \bar{S}_n - \frac{1}{n} \log \tilde{S}_n \right) = 0$$

a.s. Notice that maybe non of the averaged growth rates $\frac{1}{n}\log \bar{S}_n$ and $\frac{1}{n}\log \tilde{S}_n$ are convergent to a constant, since we didn't assume the ergodicity of $\{\mathbf{X}_i\}$.

Next we introduce further portfolio selection rules. According to Proposition 3.1 below a solution $(\lambda = W_c^*, F)$ of the (non-discounted) Bellman equation (3.13) exists, where $W_c^* \in \mathbb{R}$ is unique according to Proposition 3.2 below. W_c^* is the maximum growth rate (see Theorem 3.1 below).

Rule 3. Introduce a stationary rule such that put

$$\mathbf{b}_1^* = \{1/d, \dots, 1/d\}$$

and

$$\mathbf{b}_{i+1}^* = \arg\max_{\mathbf{b}'} \left\{ v(\mathbf{b}_i^*, \mathbf{b}', \mathbf{X}_i) + \mathbb{E}\{F(\mathbf{b}', \mathbf{X}_{i+1}) | \mathbf{X}_i\} \right\}. \tag{3.14}$$

Theorem 3.1. *Under the Conditions (i), (ii) and (iii), if S_n^* denotes the wealth at period n using the portfolio $\{\mathbf{b}_n^*\}$ then*

$$\lim_{n\to\infty} \frac{1}{n} \log S_n^* = W_c^*$$

a.s., while if S_n denotes the wealth at period n using any other portfolio $\{\mathbf{b}_n\}$ then

$$\limsup_{n\to\infty} \frac{1}{n} \log S_n \le W_c^*$$

a.s.

Remark 3.1. There is an obvious question, how to ensure that $W_c^* > 0$? Next we show a simple sufficient condition for $W_c^* > 0$. We prove that if the best asset has positive growth rate then $W_c^* > 0$, for any c. Consider the uniform static portfolio (uniform index), i.e., at time $n = 0$ we apply the uniform portfolio and later on there is no trading. It means that the wealth at time n is defined by

$$S_n = S_0 \frac{1}{d} \sum_{j=1}^{d} s_n^{(j)}.$$

Apply the following simple bounds

$$S_0 \frac{1}{d} \max_j s_n^{(j)} \leq S_n \leq S_0 \max_j s_n^{(j)}.$$

These bounds imply that

$$\begin{aligned}
\limsup_{n\to\infty} \frac{1}{n} \ln S_n &= \limsup_{n\to\infty} \max_j \frac{1}{n} \ln s_n^{(j)} \\
&\geq \max_j \limsup_{n\to\infty} \frac{1}{n} \ln s_n^{(j)} \\
&=: \max_j W^{(j)} > 0.
\end{aligned}$$

Thus,

$$W_c^* \geq \max_j W^{(j)} > 0.$$

Remark 3.2. For i.i.d. (independent identically distributed) market process, [Iyengar (2002, 2005)] observed that even in discrete time setup there is no trading with positive probability, i.e.,

$$\mathbb{P}\{\tilde{\mathbf{b}}_{n+1}(\mathbf{b}_n^*, \mathbf{X}_n) = \mathbf{b}_{n+1}^*\} > 0,$$

where the no-trading portfolio $\tilde{\mathbf{b}}_{n+1}$ has been defined by (3.2). Moreover, one may get an approximately optimal selection rule, if $\mathbf{b}_{n+1}^*$ is restricted on an appropriate neighborhood of $\tilde{\mathbf{b}}_{n+1}(\mathbf{b}_n^*, \mathbf{X}_n)$.

Remark 3.3. The problem is more simple if the market process is i.i.d. Then, on the one hand v has the form

$$\begin{aligned}
v(\mathbf{b}, \mathbf{b}', \mathbf{x}) &= \log w(\mathbf{b}, \mathbf{b}', \mathbf{x}) + \mathbb{E}\{\log \langle \mathbf{b}', \mathbf{X}_2 \rangle \,|\mathbf{X}_1 = \mathbf{x}\} \\
&= \log w(\mathbf{b}, \mathbf{b}', \mathbf{x}) + \mathbb{E}\{\log \langle \mathbf{b}', \mathbf{X}_2 \rangle\},
\end{aligned}$$

while the Bellman equation (3.13) looks like as follows:

$$\begin{aligned} W_c^* + F(\mathbf{b}, \mathbf{x}) &= \max_{\mathbf{b}'} \{v(\mathbf{b}, \mathbf{b}', \mathbf{x}) + \mathbb{E}\{F(\mathbf{b}', \mathbf{X}_2) \mid \mathbf{X}_1 = \mathbf{x}\}\} \\ &= \max_{\mathbf{b}'} \{v(\mathbf{b}, \mathbf{b}', \mathbf{x}) + \mathbb{E}\{F(\mathbf{b}', \mathbf{X}_2)\}\}. \end{aligned}$$

This problem was studied by [Iyengar (2002, 2005)]. As to Theorem 3.1, also conditional expectation in context of F in (3.14) simplifies to expectation, and its proof shows that the last assumption in Condition (ii) can be omitted. For Theorem 3.2 the analogue holds.

Remark 3.4. Use of portfolio $\mathbf{b}_n^*$ in Theorem 3.1 needs a solution of the non-discounted Bellman equation (3.13). For this, an iteration procedure is given in Lemma 3.2 below.

Remark 3.5. In practice, the conditional expectations are unknown and they can be replaced by estimates. It's an open problem what is the loss in growth rate if we apply estimates in the Bellman equation

$$\begin{aligned} W_c^* + F(\mathbf{b}, \mathbf{x}) = \max_{\mathbf{b}'} \{ &\log w(\mathbf{b}, \mathbf{b}', \mathbf{x}) + \mathbb{E}\{\log \langle \mathbf{b}', \mathbf{X}_2 \rangle \,|\mathbf{X}_1 = \mathbf{x}\} \\ &+ \mathbb{E}\{F(\mathbf{b}', \mathbf{X}_2) \mid \mathbf{X}_1 = \mathbf{x}\}\}. \end{aligned}$$

Rule 4. Choose a sequence $0 < \delta_n < 1, n = 1, 2, \dots$ such that

$$\delta_n \downarrow 0, \quad \sum_n \delta_n = \infty, \quad \frac{\delta_{n+1}}{\delta_n} \to 1 \quad (n \to \infty),$$

e.g., $\delta_n = \frac{1}{n+1}$. Set

$$F_1' := 0,$$

and iterate

$$F_{n+1}' := M_{\delta_n} F_n' - \max_{\mathbf{b}, \mathbf{x}} (M_{\delta_n} F_n')(\mathbf{b}, \mathbf{x}) \quad (n = 1, 2, \dots)$$

with

$$(M_{\delta_n} F)(\mathbf{b}, \mathbf{x}) := \max_{\tilde{\mathbf{b}}} \left\{ v(\mathbf{b}, \tilde{\mathbf{b}}, \mathbf{x}) + (1 - \delta_n) \mathbb{E}\{F(\tilde{\mathbf{b}}, \mathbf{X}_2) \mid \mathbf{X}_1 = \mathbf{x}\} \right\}, \ F \in C.$$

Put

$$\mathbf{b}_1' = \{1/d, \dots, 1/d\}$$

and

$$\mathbf{b}_{i+1}' = \arg\max_{\tilde{\mathbf{b}}} \{ v(\mathbf{b}_i', \tilde{\mathbf{b}}, \mathbf{X}_i) + (1 - \delta_i) \mathbb{E}\{F_i'(\tilde{\mathbf{b}}, \mathbf{X}_{i+1}) | \mathbf{X}_i\}\},$$

for $1 \leq i$. This non-stationary rule can be interpreted as a combination of the value iteration and Rule 1.

Theorem 3.2. *Under the Conditions (i), (ii) and (iii), if S'_n denotes the wealth at period n using the portfolio $\{\mathbf{b}'_n\}$ then*

$$\lim_{n\to\infty} \frac{1}{n} \log S'_n = W^*_c$$

a.s.

Note that according to Theorem 3.1, if S_n denotes the wealth at period n using any portfolio $\{\mathbf{b}_n\}$ then

$$\limsup_{n\to\infty} \frac{1}{n} \log S_n \leq W^*_c$$

a.s.

3.5. Portfolio selection with consumption

For a real number x, let x^+ be the positive part of x. Assume that at the end of trading period n there is a consumption $c_n \geq 0$. For the trading period n the initial capital is S_{n-1}, therefore

$$S_n = \left(S_{n-1} \langle \mathbf{b}_n , \mathbf{x}_n \rangle - c_n\right)^+ .$$

If $S_j > 0$ for all $j = 1, \ldots, n$ then we show by induction that

$$S_n = S_0 \prod_{i=1}^{n} \langle \mathbf{b}_i , \mathbf{x}_i \rangle - \sum_{k=1}^{n} c_k \prod_{i=k+1}^{n} \langle \mathbf{b}_i , \mathbf{x}_i \rangle , \tag{3.15}$$

where the empty product is 1, by definition. For $n = 1$, (3.15) holds. Assume (3.15) for $n - 1$:

$$S_{n-1} = S_0 \prod_{i=1}^{n-1} \langle \mathbf{b}_i , \mathbf{x}_i \rangle - \sum_{k=1}^{n-1} c_k \prod_{i=k+1}^{n-1} \langle \mathbf{b}_i , \mathbf{x}_i \rangle .$$

Then

$$\begin{aligned} S_n &= S_{n-1} \langle \mathbf{b}_n , \mathbf{x}_n \rangle - c_n \\ &= \left(S_0 \prod_{i=1}^{n-1} \langle \mathbf{b}_i , \mathbf{x}_i \rangle - \sum_{k=1}^{n-1} c_k \prod_{i=k+1}^{n-1} \langle \mathbf{b}_i , \mathbf{x}_i \rangle \right) \langle \mathbf{b}_n , \mathbf{x}_n \rangle - c_n \\ &= S_0 \prod_{i=1}^{n} \langle \mathbf{b}_i , \mathbf{x}_i \rangle - \sum_{k=1}^{n} c_k \prod_{i=k+1}^{n} \langle \mathbf{b}_i , \mathbf{x}_i \rangle . \end{aligned}$$

One has to emphasize that (3.15) holds for all n iff $S_n > 0$ for all n, otherwise there is a ruin. In the sequel, we study the average growth rate under no ruin and the probability of ruin.

By definition,

$$\begin{aligned}\mathbb{P}\{\text{ ruin }\} &= \mathbb{P}\left\{\bigcup_{n=1}^{\infty}\{S_n = 0\}\right\}\\ &= \mathbb{P}\left\{\bigcup_{n=1}^{\infty}\left\{S_0\prod_{i=1}^{n}\langle \mathbf{b}_i\,,\,\mathbf{x}_i\rangle - \sum_{k=1}^{n}c_k\prod_{i=k+1}^{n}\langle \mathbf{b}_i\,,\,\mathbf{x}_i\rangle \le 0\right\}\right\},\end{aligned}$$

therefore

$$\begin{aligned}\mathbb{P}\{\text{ ruin }\} &= \mathbb{P}\left\{\bigcup_{n=1}^{\infty}\left\{\prod_{i=1}^{n}\langle \mathbf{b}_i\,,\,\mathbf{x}_i\rangle\left(S_0 - \sum_{k=1}^{n}\frac{c_k}{\prod_{i=1}^{k}\langle \mathbf{b}_i\,,\,\mathbf{x}_i\rangle}\right)\le 0\right\}\right\}\\ &\le \mathbb{P}\left\{\bigcup_{n=1}^{\infty}\left\{\prod_{i=1}^{n}\langle \mathbf{b}_i\,,\,\mathbf{x}_i\rangle\left(S_0 - \sum_{k=1}^{\infty}\frac{c_k}{\prod_{i=1}^{k}\langle \mathbf{b}_i\,,\,\mathbf{x}_i\rangle}\right)\le 0\right\}\right\}\\ &\le \mathbb{P}\left\{S_0 \le \sum_{k=1}^{\infty}\frac{c_k}{\prod_{i=1}^{k}\langle \mathbf{b}_i\,,\,\mathbf{x}_i\rangle}\right\} \qquad (3.16)\end{aligned}$$

and

$$\begin{aligned}\mathbb{P}\{\text{ ruin }\} &= \mathbb{P}\left\{\bigcup_{n=1}^{\infty}\left\{\prod_{i=1}^{n}\langle \mathbf{b}_i\,,\,\mathbf{x}_i\rangle\left(S_0 - \sum_{k=1}^{n}\frac{c_k}{\prod_{i=1}^{k}\langle \mathbf{b}_i\,,\,\mathbf{x}_i\rangle}\right)\le 0\right\}\right\}\\ &\ge \max_n \mathbb{P}\left\{\prod_{i=1}^{n}\langle \mathbf{b}_i\,,\,\mathbf{x}_i\rangle\left(S_0 - \sum_{k=1}^{n}\frac{c_k}{\prod_{i=1}^{k}\langle \mathbf{b}_i\,,\,\mathbf{x}_i\rangle}\right)\le 0\right\}\\ &= \mathbb{P}\left\{S_0 \le \sum_{k=1}^{\infty}\frac{c_k}{\prod_{i=1}^{k}\langle \mathbf{b}_i\,,\,\mathbf{x}_i\rangle}\right\}. \qquad (3.17)\end{aligned}$$

(3.16) and (3.17) imply that

$$\mathbb{P}\{\text{ ruin }\} = \mathbb{P}\left\{S_0 \le \sum_{k=1}^{\infty}\frac{c_k}{\prod_{i=1}^{k}\langle \mathbf{b}_i\,,\,\mathbf{x}_i\rangle}\right\}.$$

Under no ruin, on the one hand we get the upper bound on the average growth rate

$$
\begin{aligned}
W_n &= \frac{1}{n}\ln S_n \\
&= \frac{1}{n}\ln\left(S_0\prod_{i=1}^{n}\langle \mathbf{b}_i\,,\,\mathbf{x}_i\rangle - \sum_{k=1}^{n} c_k \prod_{i=k+1}^{n}\langle \mathbf{b}_i\,,\,\mathbf{x}_i\rangle\right) \\
&\le \frac{1}{n}\ln S_0\prod_{i=1}^{n}\langle \mathbf{b}_i\,,\,\mathbf{x}_i\rangle \\
&= \frac{1}{n}\sum_{i=1}^{n}\ln\langle \mathbf{b}_i\,,\,\mathbf{x}_i\rangle + \frac{1}{n}\ln S_0.
\end{aligned}
$$

On the other hand we have the lower bound

$$
\begin{aligned}
W_n &= \frac{1}{n}\ln S_n \\
&= \frac{1}{n}\ln\left(S_0\prod_{i=1}^{n}\langle \mathbf{b}_i\,,\,\mathbf{x}_i\rangle - \sum_{k=1}^{n} c_k \prod_{i=k+1}^{n}\langle \mathbf{b}_i\,,\,\mathbf{x}_i\rangle\right) \\
&= \frac{1}{n}\ln\prod_{i=1}^{n}\langle \mathbf{b}_i\,,\,\mathbf{x}_i\rangle\left(S_0 - \sum_{k=1}^{n}\frac{c_k}{\prod_{i=1}^{k}\langle \mathbf{b}_i\,,\,\mathbf{x}_i\rangle}\right) \\
&\ge \frac{1}{n}\ln\prod_{i=1}^{n}\langle \mathbf{b}_i\,,\,\mathbf{x}_i\rangle\left(S_0 - \sum_{k=1}^{\infty}\frac{c_k}{\prod_{i=1}^{k}\langle \mathbf{b}_i\,,\,\mathbf{x}_i\rangle}\right) \\
&= \frac{1}{n}\sum_{i=1}^{n}\ln\langle \mathbf{b}_i\,,\,\mathbf{x}_i\rangle + \frac{1}{n}\ln\left(S_0 - \sum_{k=1}^{\infty}\frac{c_k}{\prod_{i=1}^{k}\langle \mathbf{b}_i\,,\,\mathbf{x}_i\rangle}\right),
\end{aligned}
$$

therefore under no ruin the asymptotic average growth rate with consumption is the same as without consumption:

$$
W_n = \frac{1}{n}\ln S_n \approx \frac{1}{n}\sum_{i=1}^{n}\ln\langle \mathbf{b}_i\,,\,\mathbf{x}_i\rangle\,.
$$

Consider the case of constant consumption, i.e., $c_n = c > 0$. Then there is no ruin if

$$
S_0 > c\sum_{k=1}^{\infty}\frac{1}{\prod_{i=1}^{k}\langle \mathbf{b}_i\,,\,\mathbf{x}_i\rangle}.
$$

Because of the definition of the average growth rate we have that

$$
W_k \approx \frac{1}{k}\ln\prod_{i=1}^{k}\langle \mathbf{b}_i\,,\,\mathbf{x}_i\rangle\,,
$$

which implies that

$$\sum_{k=1}^{\infty} \frac{1}{\prod_{i=1}^{k} \langle \mathbf{b}_i \, , \, \mathbf{x}_i \rangle} \approx \sum_{k=1}^{\infty} e^{-kW_k}.$$

Assume that our portfolio selection is asymptotically optimal, which means that

$$\lim_{n\to\infty} W_n = W^*.$$

Then

$$\sum_{k=1}^{\infty} \frac{1}{\prod_{i=1}^{k} \langle \mathbf{b}_i \, , \, \mathbf{x}_i \rangle} \approx \sum_{k=1}^{\infty} e^{-kW^*} = \frac{e^{-W^*}}{1 - e^{-W^*}}.$$

This approximation implies that the ruin probability can be small only if

$$S_0 > c\frac{e^{-W^*}}{1 - e^{-W^*}}.$$

A special case of this model is when there is only one risk-free asset:

$$S_n = (S_{n-1}(1 + r) - c)^+$$

with some $r > 0$. Obviously, there is no ruin if $S_0 r > c$. It is easy to verify that this assumption can be derived from the general condition if

$$e^{W^*} = 1 + r.$$

The ruin probability can be decreased if the consumptions happen in blocks of size N trading periods. Let S_n denote the wealth at the end of n-th block. Then

$$S_n = \left(S_{n-1} \prod_{j=(n-1)N+1}^{nN} \langle \mathbf{b}_j \, , \, \mathbf{x}_j \rangle - Nc \right)^+ .$$

Similarly to the previous calculations, we can check that under no ruin the average growth rates with and without consumption are the same. Moreover

$$\mathbb{P}\{ \text{ ruin } \} = \mathbb{P} \left\{ S_0 \le cN \sum_{k=1}^{\infty} \frac{1}{\prod_{i=1}^{kN} \langle \mathbf{b}_i \, , \, \mathbf{x}_i \rangle} \right\}.$$

This ruin probability is a monotonically decreasing function of N, and for large N the exact condition of no ruin is the same as the approximation mentioned above.

This model can be applied for the analysis of portfolio selection strategies with fixed transaction cost such that c_n is the transaction cost to be paid when change the portfolio $\mathbf{b}_n$ to $\mathbf{b}_{n+1}$. In this case the transaction cost c_n depends on the number of shares involved in the transaction.

Let's calculate c_n. At the end of the n-th trading period and before paying for transaction cost the wealth at asset j is $S_{n-1}b_n^{(j)}x_n^{(j)}$, which means that the number of shares j is

$$m_n^{(j)} = \frac{S_{n-1}b_n^{(j)}x_n^{(j)}}{S_n^{(j)}}.$$

In the model of fixed transaction cost, we assume that $m_n^{(j)}$ is integer. If one changes the portfolio $\mathbf{b}_n$ to $\mathbf{b}_{n+1}$ then the wealth at asset j should be $S_{n-1}\langle \mathbf{b}_n\,,\,\mathbf{x}_n\rangle b_{n+1}^{(j)}$, so the number of shares j should be

$$m_{n+1}^{(j)} = \frac{S_{n-1}\langle \mathbf{b}_n\,,\,\mathbf{x}_n\rangle b_{n+1}^{(j)}}{S_n^{(j)}}.$$

If $m_{n+1}^{(j)} < m_n^{(j)}$ then we have to sell, and the wealth what we got is

$$\sum_{j=1}^d \left(m_n^{(j)} - m_{n+1}^{(j)}\right)^+ S_n^{(j)} = \sum_{j=1}^d \left(S_{n-1}b_n^{(j)}x_n^{(j)} - S_{n-1}\langle \mathbf{b}_n\,,\,\mathbf{x}_n\rangle b_{n+1}^{(j)}\right)^+.$$

If $m_{n+1}^{(j)} > m_n^{(j)}$ then we have to buy, and the wealth what we pay is

$$\sum_{j=1}^d \left(m_{n+1}^{(j)} - m_n^{(j)}\right)^+ S_n^{(j)} = \sum_{j=1}^d \left(S_{n-1}\langle \mathbf{b}_n\,,\,\mathbf{x}_n\rangle b_{n+1}^{(j)} - S_{n-1}b_n^{(j)}x_n^{(j)}\right)^+.$$

Let $C > 0$ be the fixed transaction cost, then the transaction fee is

$$c_n = c_n(\mathbf{b}_{n+1}) = C\sum_{j=1}^d \left|m_n^{(j)} - m_{n+1}^{(j)}\right|.$$

The portfolio selection $\mathbf{b}_{n+1}$ is self-financing if

$$\begin{aligned}&\sum_{j=1}^d \left(S_{n-1}b_n^{(j)}x_n^{(j)} - S_{n-1}\langle \mathbf{b}_n\,,\,\mathbf{x}_n\rangle b_{n+1}^{(j)}\right)^+\\ \geq &\sum_{j=1}^d \left(S_{n-1}\langle \mathbf{b}_n\,,\,\mathbf{x}_n\rangle b_{n+1}^{(j)} - S_{n-1}b_n^{(j)}x_n^{(j)}\right)^+ + c_n.\end{aligned}$$

$\mathbf{b}_{n+1}$ is an admissible portfolio if $m_{n+1}^{(j)}$ is integer for all j and it satisfies the self-financing condition. The set of admissible portfolios is denoted by $\Delta_{n,d}$.

Taking into account the fixed transaction cost, a kernel based portfolio selection can be defined as follows: choose the radius $r_{k,\ell} > 0$ such that for any fixed k,

$$\lim_{\ell \to \infty} r_{k,\ell} = 0.$$

For $n > k+1$, introduce the expert $\mathbf{b}^{(k,\ell)}$ by

$$\mathbf{b}_{n+1}^{(k,\ell)} = \underset{\mathbf{b} \in \Delta_{n,d}}{\arg\max} \sum_{i \in J} \ln \left\{ (S_{n-1}^{(k,\ell)} \left\langle \mathbf{b}_n^{(k,\ell)} , \mathbf{x}_n \right\rangle - c_n(\mathbf{b})) \langle \mathbf{b} , \mathbf{x}_i \rangle \right\} ,$$

if the sum is non-void, and $\mathbf{b}_0 = (1/d, \ldots, 1/d)$ otherwise, where

$$J = \left\{ k < i \le n : \|\mathbf{x}_{i-k}^{i-1} - \mathbf{x}_{n-k+1}^{n}\| \le r_{k,\ell} \right\}.$$

Combine the elementary portfolio strategies $\mathbf{B}^{(k,\ell)} = \{\mathbf{b}_n^{(k,\ell)}\}$ as in (3.9).

3.6. Proofs

We split the statement of Theorem 3.1 into two propositions.

Proposition 3.1. *Under the Conditions (i), (ii) and (iii) the Bellman equation (3.13) has a solution* (W_c^*, F) *such that the function* F *is bounded and continuous, where*

$$\max_{\mathbf{b},\mathbf{x}} F(\mathbf{b}, \mathbf{x}) = 0.$$

Proof. Let C be the Banach space of continuous functions F defined on the compact set $\Delta_d \times [a_1, a_2]^d$ with the sup norm $\|\cdot\|_\infty$. For $0 \le \delta < 1$ and for $f \in C$, define the operator

$$(M_\delta f)(\mathbf{b}, \mathbf{x}) := \max_{\mathbf{b}'} \{ v(\mathbf{b}, \mathbf{b}', \mathbf{x}) + (1-\delta)\mathbb{E}\{ f(\mathbf{b}', \mathbf{X}_2) \mid \mathbf{X}_1 = \mathbf{x} \} \}. \tag{3.18}$$

By continuity assumption (ii) this leads to an operator

$$M_\delta : C \to C.$$

(See [Schäfer (2002)] p.114.)

The operator M_δ is continuous, even Lipschitz continuous with Lipschitz constant $1-\delta$. Indeed, for $f, f' \in C$ from the representation

$$(M_\delta f)(\mathbf{b}, \mathbf{x}) = v(\mathbf{b}, \mathbf{b}_f^*(\mathbf{b}, \mathbf{x}), \mathbf{x}) + (1-\delta)\mathbb{E}\{ f(\mathbf{b}_f^*(\mathbf{b}, \mathbf{x}), \mathbf{X}_2) \mid \mathbf{X}_1 = \mathbf{x} \}$$

and from the corresponding representation of $(M_\delta f')(\mathbf{b},\mathbf{x})$ one obtains

$$\begin{aligned}(M_\delta f')(\mathbf{b},\mathbf{x}) &\geq v(\mathbf{b},\mathbf{b}^*_f(\mathbf{b},\mathbf{x}),\mathbf{x}) + (1-\delta)\mathbb{E}\{f'(\mathbf{b}^*_f(\mathbf{b},\mathbf{x}),\mathbf{X}_2) \mid \mathbf{X}_1=\mathbf{x}\}\\ &\geq v(\mathbf{b},\mathbf{b}^*_f(\mathbf{b},\mathbf{x}),\mathbf{x}) + (1-\delta)\mathbb{E}\{f(\mathbf{b}^*_f(\mathbf{b},\mathbf{x}),\mathbf{X}_2) \mid \mathbf{X}_1=\mathbf{x}\}\\ &\quad -(1-\delta)\|f-f'\|_\infty\\ &= (M_\delta f)(\mathbf{b},\mathbf{x}) - (1-\delta)\|f-f'\|_\infty\end{aligned}$$

for all $(\mathbf{b},\mathbf{x}) \in \Delta_d \times [a_1,a_2]^d$, therefore

$$\|M_\delta f - M_\delta f'\|_\infty \leq (1-\delta)\|f-f'\|_\infty.$$

Thus, by Banach's fixed point theorem, the Bellman equation

$$\lambda + F(\mathbf{b},\mathbf{x}) = \max_{\mathbf{b}'}\left\{v(\mathbf{b},\mathbf{b}',\mathbf{x}) + (1-\delta)\mathbb{E}\{F(\mathbf{b}',\mathbf{X}_2) \mid \mathbf{X}_1=\mathbf{x}\}\right\}, \quad (3.19)$$

i.e.,

$$\lambda + F = M_\delta F$$

with $\lambda \in \mathbb{R}$, has a unique solution if $0<\delta<1$. (3.19) corresponds to (3.12) for $\lambda=0$, $0<\delta<1$ with the unique solution denoted by F_δ, and to (3.13) for $\lambda = W^*_c$ and $\delta=0$.

We notice

$$\sup_{0<\delta<1} \delta\|F_\delta\|_\infty \leq \max_{\mathbf{b},\mathbf{b}',\mathbf{x}} |v(\mathbf{b},\mathbf{b}',\mathbf{x})| < \infty,$$

(cf. [Schäfer (2002)], Lemma 4.2.3). Similarly to [Iyengar (2002)], put

$$m_\delta := \max_{(\mathbf{b},\mathbf{x})} F_\delta(\mathbf{b},\mathbf{x}), \quad (3.20)$$

where we get that

$$\sup_{0<\delta<1} \delta m_\delta < \infty.$$

Put

$$W^*_c := \limsup_{\delta\downarrow 0} \delta m_\delta$$

and

$$\tilde{F}_\delta(\mathbf{b},\mathbf{x}) := F_\delta(\mathbf{b},\mathbf{x}) - m_\delta. \quad (3.21)$$

Thus,

$$\max_{(\mathbf{b},\mathbf{x})} \tilde{F}_\delta(\mathbf{b},\mathbf{x}) = 0. \quad (3.22)$$

$\tilde{F}_\delta$ satisfies the Bellman equation (3.19) with $\lambda = \delta m_\delta$, therefore

$$\delta m_\delta + \tilde{F}_\delta = M_\delta \tilde{F}_\delta = M_0 \tilde{F}_\delta + (M_\delta \tilde{F}_\delta - M_0 \tilde{F}_\delta) \tag{3.23}$$

It is easy to check that

$$\|M_\delta \tilde{F}_\delta - M_0 \tilde{F}_\delta\|_\infty \leq \delta \|\tilde{F}_\delta\|_\infty. \tag{3.24}$$

By Lemma 3.1 below

$$\sup_{0<\delta<1} \|\tilde{F}_\delta\|_\infty < \infty. \tag{3.25}$$

Now we choose a sequence δ_n with $\delta_n \downarrow 0$ such that

$$\delta_n m_{\delta_n} \to W_c^*. \tag{3.26}$$

Lemma 3.1 further states that

$$\sup_{0<\delta<1} |\tilde{F}_\delta(\bar{\mathbf{b}}, \bar{\mathbf{x}}) - \tilde{F}_\delta(\mathbf{b}, \mathbf{x})| \to 0$$

(even uniformly with respect to $(\mathbf{b}, \mathbf{x})$, because of compactness of $\Delta_d \times [a_1, a_2]^d$) when $(\bar{\mathbf{b}}, \bar{\mathbf{x}}) \to (\mathbf{b}, \mathbf{x})$, i.e., there is equicontinuity for $\{\tilde{F}_\delta\}$, which together with (3.25) implies that there exist a subsequence δ_{n_l} and a function $\tilde{F} \in C$ such that $\tilde{F}_{\delta_{n_l}}$ converges in C to $\tilde{F}$ (cf. Ascoli-Arzelá theorem, [Yosida (1968)]). Thus, by continuity of M_0, we get the convergence of $M_0 \tilde{F}_{\delta_{n_l}}$ in C to $M_0 \tilde{F}$. Therefore

$$W_c^* + \tilde{F} = M_0 \tilde{F},$$

i.e., $\tilde{F} \in C$ solves the Bellman equation (3.13). $\tilde{F}$ is continuous on a compact set, therefore it is bounded, where

$$\max_{\mathbf{b}, \mathbf{x}} \tilde{F}(\mathbf{b}, \mathbf{x}) = 0.$$

□

Lemma 3.1. *If F_δ denotes the solution of the discounted Bellman equation (3.12) then (3.25) holds and it implies that*

$$\sup_{0<\delta<1} |F_\delta(\bar{\mathbf{b}}, \bar{\mathbf{x}}) - F_\delta(\mathbf{b}, \mathbf{x})| \to 0 \tag{3.27}$$

when $(\bar{\mathbf{b}}, \bar{\mathbf{x}}) \to (\mathbf{b}, \mathbf{x})$.

Proof. We use the decomposition

$$F_\delta(\bar{\mathbf{b}},\bar{\mathbf{x}}) - F_\delta(\mathbf{b},\mathbf{x}) = F_\delta(\bar{\mathbf{b}},\bar{\mathbf{x}}) - F_\delta(\mathbf{b},\bar{\mathbf{x}}) + F_\delta(\mathbf{b},\bar{\mathbf{x}}) - F_\delta(\mathbf{b},\mathbf{x}).$$

Concerning the first term in this decomposition we assumed that F_δ the solution of the discounted Bellman equation (3.12), therefore

$$\begin{aligned}
&F_\delta(\bar{\mathbf{b}},\bar{\mathbf{x}}) - F_\delta(\mathbf{b},\bar{\mathbf{x}}) \\
&= \max_{\mathbf{b}'} \left\{v(\bar{\mathbf{b}},\mathbf{b}',\bar{\mathbf{x}}) + (1-\delta)\mathbb{E}\{F_\delta(\mathbf{b}',\mathbf{X}_2) \mid \mathbf{X}_1 = \bar{\mathbf{x}}\}\right\} \\
&\quad - \max_{\mathbf{b}''} \{v(\mathbf{b},\mathbf{b}'',\bar{\mathbf{x}}) + (1-\delta)\mathbb{E}\{F_\delta(\mathbf{b}'',\mathbf{X}_2) \mid \mathbf{X}_1 = \bar{\mathbf{x}}\}\} \\
&\le \max_{\mathbf{b}'} \{v(\bar{\mathbf{b}},\mathbf{b}',\bar{\mathbf{x}}) + (1-\delta)\mathbb{E}\{F_\delta(\mathbf{b}',\mathbf{X}_2) \mid \mathbf{X}_1 = \bar{\mathbf{x}}\} \\
&\quad -(v(\mathbf{b},\mathbf{b}',\bar{\mathbf{x}}) + (1-\delta)\mathbb{E}\{F_\delta(\mathbf{b}',\mathbf{X}_2) \mid \mathbf{X}_1 = \bar{\mathbf{x}}\})\} \\
&= \max_{\mathbf{b}'} \{v(\bar{\mathbf{b}},\mathbf{b}',\bar{\mathbf{x}}) - v(\mathbf{b},\mathbf{b}',\bar{\mathbf{x}})\},
\end{aligned}$$

therefore

$$\sup_{0<\delta<1} |F_\delta(\bar{\mathbf{b}},\bar{\mathbf{x}}) - F_\delta(\mathbf{b},\bar{\mathbf{x}})| \le \max_{\mathbf{b}'} |v(\bar{\mathbf{b}},\mathbf{b}',\bar{\mathbf{x}}) - v(\mathbf{b},\mathbf{b}',\bar{\mathbf{x}})| \to 0 \qquad (3.28)$$

when $(\bar{\mathbf{b}},\bar{\mathbf{x}}) \to (\mathbf{b},\mathbf{x})$. Concerning the second term in this decomposition, we analogously get that

$$\begin{aligned}
F_\delta(\mathbf{b},\bar{\mathbf{x}}) - F_\delta(\mathbf{b},\mathbf{x}) &\le \max_{\mathbf{b}'} \{v(\mathbf{b},\mathbf{b}',\bar{\mathbf{x}}) + (1-\delta)\mathbb{E}\{F_\delta(\mathbf{b}',\mathbf{X}_2) \mid \mathbf{X}_1 = \bar{\mathbf{x}}\} \\
&\quad -(v(\mathbf{b},\mathbf{b}',\mathbf{x}) + (1-\delta)\mathbb{E}\{F_\delta(\mathbf{b}',\mathbf{X}_2) \mid \mathbf{X}_1 = \mathbf{x}\})\} \\
&\le \max_{\mathbf{b}'} \{v(\mathbf{b},\mathbf{b}',\bar{\mathbf{x}}) - v(\mathbf{b},\mathbf{b}',\mathbf{x})\} \\
&\quad +(1-\delta) \max_{\mathbf{b}'} \{\mathbb{E}\{F_\delta(\mathbf{b}',\mathbf{X}_2) \mid \mathbf{X}_1 = \bar{\mathbf{x}}\} \\
&\quad -\mathbb{E}\{F_\delta(\mathbf{b}',\mathbf{X}_2) \mid \mathbf{X}_1 = \mathbf{x}\}\}.
\end{aligned}$$

Moreover

$$\begin{aligned}
&\mathbb{E}\{F_\delta(\mathbf{b}',\mathbf{X}_2) \mid \mathbf{X}_1 = \bar{\mathbf{x}}\} - \mathbb{E}\{F_\delta(\mathbf{b}',\mathbf{X}_2) \mid \mathbf{X}_1 = \mathbf{x}\}\} \\
&= \mathbb{E}\{\tilde{F}_\delta(\mathbf{b}',\mathbf{X}_2) \mid \mathbf{X}_1 = \bar{\mathbf{x}}\} - \mathbb{E}\{\tilde{F}_\delta(\mathbf{b}',\mathbf{X}_2) \mid \mathbf{X}_1 = \mathbf{x}\}\} \\
&\le \|\tilde{F}_\delta\|_\infty V(\mathbf{x},\bar{\mathbf{x}}),
\end{aligned}$$

where the function V has been defined for Condition (ii). Thus,

$$\begin{aligned}
&\sup_{0<\delta<1} |F_\delta(\mathbf{b},\bar{\mathbf{x}}) - F_\delta(\mathbf{b},\mathbf{x})| \\
&\le \max_{\mathbf{b}'} |v(\mathbf{b},\mathbf{b}',\bar{\mathbf{x}}) - v(\mathbf{b},\mathbf{b}',\mathbf{x})| + \sup_{0<\delta<1} \|\tilde{F}_\delta\|_\infty V(\mathbf{x},\bar{\mathbf{x}}).
\end{aligned} \qquad (3.29)$$

The inequalities in (3.28) and (3.29) and boundedness of g and also of v (by Condition (iii)) yield

$$\sup_{0<\delta<1} |F_\delta(\bar{\mathbf{b}}, \bar{\mathbf{x}}) - F_\delta(\mathbf{b}, \mathbf{x})| \leq const + \sup_{0<\delta<1} \|\tilde{F}_\delta\|_\infty V(\mathbf{x}, \bar{\mathbf{x}})$$

for some $const < \infty$. Noticing

$$\sup_{(\mathbf{b},\mathbf{x}),(\bar{\mathbf{b}},\bar{\mathbf{x}})} |F_\delta(\bar{\mathbf{b}}, \bar{\mathbf{x}}) - F_\delta(\mathbf{b}, \mathbf{x})| = \sup_{(\mathbf{b},\mathbf{x}),(\bar{\mathbf{b}},\bar{\mathbf{x}})} |\tilde{F}_\delta(\bar{\mathbf{b}}, \bar{\mathbf{x}}) - \tilde{F}_\delta(\mathbf{b}, \mathbf{x})| = \|\tilde{F}_\delta\|_\infty$$

(by (3.21) and (3.22)), we then obtain

$$\sup_{0<\delta<1} \|\tilde{F}_\delta\|_\infty \leq const + \sup_{0<\delta<1} \|\tilde{F}_\delta\|_\infty \max_{\mathbf{x},\bar{\mathbf{x}}} V(\mathbf{x}, \bar{\mathbf{x}})$$

and thus (3.25) by (3.4). Condition (ii) and (3.25) yield that the right hand side of (3.29) converges to 0 when $(\bar{\mathbf{b}}, \bar{\mathbf{x}}) \to (\mathbf{b}, \mathbf{x})$. Then (3.28) and (3.29) imply (3.27). $\square$

Proposition 3.2. *Assume that the Bellman equation (3.13) has a solution (W_c^*, F) such that the function F is bounded. If S_n^* denotes the wealth at period n using the portfolio $\{\mathbf{b}_n^*\}$ then*

$$\lim_{n\to\infty} \frac{1}{n} \log S_n^* = W_c^*$$

a.s., while if S_n denotes the wealth at period n using any other portfolio $\{\mathbf{b}_n\}$ then

$$\limsup_{n\to\infty} \frac{1}{n} \log S_n \leq W_c^*$$

a.s. These statements imply that W_c^* in the Bellman equation (3.13) is unique.

Proof. We have to show that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} g(\mathbf{b}_i^*, \mathbf{b}_{i+1}^*, \mathbf{X}_i, \mathbf{X}_{i+1}) = W_c^*$$

a.s. and

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} g(\mathbf{b}_i, \mathbf{b}_{i+1}, \mathbf{X}_i, \mathbf{X}_{i+1}) \leq W_c^*$$

a.s. Because of the martingale difference argument in Section 3.2, these two limit relations are equivalent to

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} v(\mathbf{b}_i^*, \mathbf{b}_{i+1}^*, \mathbf{X}_i) = W_c^*$$

a.s. and

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} v(\mathbf{b}_i, \mathbf{b}_{i+1}, \mathbf{X}_i) \le W_c^*$$

a.s. (3.13) and (3.14) imply that

$$W_c^* + F(\mathbf{b}_i^*, \mathbf{X}_i) = v(\mathbf{b}_i^*, \mathbf{b}_{i+1}^*, \mathbf{X}_i) + \mathbb{E}\{F(\mathbf{b}_{i+1}^*, \mathbf{X}_{i+1}) \mid \mathbf{b}_{i+1}^*, \mathbf{X}_i\}, \tag{3.30}$$

while for any portfolio $\{\mathbf{b}_i\}$,

$$W_c^* + F(\mathbf{b}_i, \mathbf{X}_i) \ge v(\mathbf{b}_i, \mathbf{b}_{i+1}, \mathbf{X}_i) + \mathbb{E}\{F(\mathbf{b}_{i+1}, \mathbf{X}_{i+1}) \mid \mathbf{b}_{i+1}, \mathbf{X}_i\}. \tag{3.31}$$

Because of (3.30), we get that

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n} v(\mathbf{b}_i^*, \mathbf{b}_{i+1}^*, \mathbf{X}_i) \\
&= W_c^* + \frac{1}{n}\sum_{i=1}^{n} \left(F(\mathbf{b}_i^*, \mathbf{X}_i) - \mathbb{E}\{F(\mathbf{b}_{i+1}^*, \mathbf{X}_{i+1}) \mid \mathbf{b}_{i+1}^*, \mathbf{X}_i\}\right) \\
&= W_c^* + \frac{1}{n}\sum_{i=1}^{n} F(\mathbf{b}_i^*, \mathbf{X}_i) - \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\{F(\mathbf{b}_{i+1}^*, \mathbf{X}_{i+1}) \mid \mathbf{X}_1^i\} \\
&= W_c^* + \frac{1}{n}\sum_{i=2}^{n} \left(F(\mathbf{b}_i^*, \mathbf{X}_i) - \mathbb{E}\{F(\mathbf{b}_i^*, \mathbf{X}_i) \mid \mathbf{X}_1^{i-1}\}\right) \\
&\quad + \frac{1}{n}\left(F(\mathbf{b}_1^*, \mathbf{X}_1) - \mathbb{E}\{F(\mathbf{b}_{n+1}^*, \mathbf{X}_{n+1}) \mid \mathbf{X}_1^n\}\right).
\end{aligned}$$

By the condition of Theorem 3.1, the function F is bounded, therefore the Chow theorem can be applied for martingale differences, and so

$$\frac{1}{n}\sum_{i=1}^{n} v(\mathbf{b}_i^*, \mathbf{b}_{i+1}^*, \mathbf{X}_i) \to W_c^*$$

a.s. Similarly, because of (3.31), we get that

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n} v(\mathbf{b}_i, \mathbf{b}_{i+1}, \mathbf{X}_i) \\
&\le W_c^* + \frac{1}{n}\sum_{i=1}^{n} \left(F(\mathbf{b}_i, \mathbf{X}_i) - \mathbb{E}\{F(\mathbf{b}_{i+1}, \mathbf{X}_{i+1}) | \mathbf{b}_{i+1}, \mathbf{X}_i\}\right) \\
&= W_c^* + \frac{1}{n}\sum_{i=1}^{n} F(\mathbf{b}_i, \mathbf{X}_i) - \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\{F(\mathbf{b}_{i+1}, \mathbf{X}_{i+1}) | \mathbf{X}_1^i\} \\
&\to W_c^*
\end{aligned}$$

a.s. □

Corollary 3.1. *Assume the conditions of Proposition 3.1 and let m_δ defined by (3.20). Then*

$$\delta m_\delta \to W_c^* \text{ as } \delta \downarrow 0.$$

For each sequence $0 < \delta_n < 1$ with $\delta_n \downarrow 0$, the sequence $\tilde{F}_{\delta_n} \in C$ defined by (3.21) converges to a set of solutions F of the Bellman equation (3.13).

Proof. Since in the proof of Proposition 3.1 $\limsup_{\delta\downarrow 0} \delta m_\delta$ can be replaced by $\liminf_{\delta\downarrow 0} \delta m_\delta$, uniqueness of W_c^* yields existence of $\lim_{\delta\downarrow 0} \delta m_\delta = W_c^*$. For each sequence $\delta_n \downarrow 0$ a subsequence δ_{n_ℓ} exists such that $\tilde{F}_{\delta_{n_\ell}}$ converges in C to some solution F of (3.13). This proves the second assertion. □

For the proof of Theorem 3.2 we need the following lemma:

Lemma 3.2. *Assume Conditions (i), (ii) and (iii). Let δ_n and F_n' be as in Rule 4. Then F_n' converges in C to a set of solutions F of the Bellman equation (3.13), further*

$$w_n := \max_{\mathbf{b},\mathbf{x}} (M_{\delta_n} F_n')(\mathbf{b},\mathbf{x}) \to W_c^* \text{ as } n \to \infty.$$

Proof. We can write

$$F_{n+1}' = M_{\delta_n} F_n' - w_n \tag{3.32}$$

with the continuous operator $M_{\delta_n} : C \to C$ according to (3.18). It holds

$$\begin{aligned}
|F_{n+1}'(\bar{\mathbf{b}},\bar{\mathbf{x}}) - F_{n+1}'(\mathbf{b},\mathbf{x})| &= |(M_{\delta_n} F_n')(\bar{\mathbf{b}},\bar{\mathbf{x}}) - (M_{\delta_n} F_n')(\mathbf{b},\mathbf{x})| \\
&\le |(M_{\delta_n} F_n')(\bar{\mathbf{b}},\bar{\mathbf{x}}) - (M_{\delta_n} F_n')(\mathbf{b},\bar{\mathbf{x}})| \\
&\quad + |(M_{\delta_n} F_n')(\mathbf{b},\bar{\mathbf{x}}) - (M_{\delta_n} F_n')(\mathbf{b},\mathbf{x})| \\
&\le \max_{\mathbf{b}'} |v(\bar{\mathbf{b}},\mathbf{b}',\bar{\mathbf{x}}) - v(\mathbf{b},\mathbf{b}',\bar{\mathbf{x}})| \\
&\quad + \max_{\mathbf{b}'} |v(\mathbf{b},\mathbf{b}',\bar{\mathbf{x}}) - v(\mathbf{b},\mathbf{b}',\mathbf{x})| \\
&\quad + \max_{\mathbf{x},\bar{\mathbf{x}}} V(\mathbf{x},\bar{\mathbf{x}}) \|F_n'\|_\infty,
\end{aligned} \tag{3.33}$$

where the inequalities are obtained as in the proof of Lemma 3.1. Noticing

$$\max_{\mathbf{b},\mathbf{x}} F_n'(\mathbf{b},\mathbf{x}) = 0$$

and thus

$$\max_{(\mathbf{b},\mathbf{x}),(\bar{\mathbf{b}},\bar{\mathbf{x}})} |F_n'(\mathbf{b},\mathbf{x}) - F_n'(\bar{\mathbf{b}},\bar{\mathbf{x}})| = \|F_n'\|_\infty,$$

moreover, the boundedness of v implies that

$$\|F'_{n+1}\|_\infty \le const + \max_{\mathbf{x},\bar{\mathbf{x}}} V(\mathbf{x},\bar{\mathbf{x}})\|F'_n\|_\infty$$

with $const < \infty$. Then, by induction,

$$\|F'_n\|_\infty \le \frac{const}{1-\max_{\mathbf{x},\bar{\mathbf{x}}} V(\mathbf{x},\bar{\mathbf{x}})} =: K < \infty. \tag{3.34}$$

It can be easily checked that

$$\|M_{\delta_{n+1}}F'_{n+1} - M_{\delta_n}F'_{n+1}\|_\infty \le (\delta_n - \delta_{n+1})\|F'_{n+1}\|_\infty. \tag{3.35}$$

According to the proof of Proposition 3.1, the operator M_{δ_n} is Lipschitz continuous with Lipschitz constant $1-\delta_n$. Then

$$\begin{aligned}
&\|F'_{n+2} - F'_{n+1}\|_\infty \\
&= \|M_{\delta_{n+1}}F'_{n+1} - M_{\delta_n}F'_n\|_\infty \\
&\le \|M_{\delta_n}F'_{n+1} - M_{\delta_n}F'_n\|_\infty + \|M_{\delta_{n+1}}F'_{n+1} - M_{\delta_n}F'_{n+1}\|_\infty \\
&\le (1-\delta_n)\|F'_{n+1} - F'_n\|_\infty + \left(1 - \frac{\delta_{n+1}}{\delta_n}\right)\delta_n K.
\end{aligned}$$

By the condition on δ_n, we then obtain

$$\|F'_{n+1} - F'_n\|_\infty \to 0 \text{ as } n \to \infty, \tag{3.36}$$

(cf. Lemma 1(c) in [Walk and Zsidó (1989)]). Now let (δ_{n_k}) be an arbitrary subsequence of (δ_n). From (3.33) and (3.34) and Condition (ii) we obtain

$$\sup_i |F'_i(\bar{\mathbf{b}},\bar{\mathbf{x}}) - F'_i(\mathbf{b},\mathbf{x})| \to 0$$

when $(\bar{\mathbf{b}},\bar{\mathbf{x}}) \to (\mathbf{b},\mathbf{x})$, even uniformly with respect to $(\mathbf{b},\mathbf{x})$. This together with (3.34) yields existence of a subsequence $(\delta_{n_{k_\ell}})$ and of a function $\bar{F} \in C$ (bounded, where $\max_{\mathbf{b},\mathbf{x}} \bar{F}(\mathbf{b},\mathbf{x}) = 0$) such that

$$\|F'_{n_{k_\ell}} - \bar{F}\|_\infty \to 0 \text{ as } \ell \to \infty. \tag{3.37}$$

Thus, by continuity of M_0,

$$\|M_0F'_{n_{k_\ell}} - M_0\bar{F}\|_\infty \to 0 \text{ as } \ell \to \infty. \tag{3.38}$$

By (3.32),

$$F'_{n_{k_\ell}} + (F'_{n_{k_\ell}+1} - F'_{n_{k_\ell}}) = M_0F'_{n_{k_\ell}} + (M_{\delta_{n_{k_\ell}}}F'_{n_{k_\ell}} - M_0F'_{n_{k_\ell}}) - w_{n_{k_\ell}}.$$

(3.36) implies that

$$\|F'_{n_{k_\ell}+1} - F'_{n_{k_\ell}}\|_\infty \to 0.$$

By (3.24) and (3.34),

$$\|M_{\delta_{n_{k_\ell}}} F'_{n_{k_\ell}} - M_0 F'_{n_{k_\ell}}\|_\infty \leq \delta_{n_{k_\ell}} K \to 0.$$

This together with (3.37) and (3.38) yields convergence of $(w_{n_{k_\ell}})$ and

$$\lim_\ell w_{n_{k_\ell}} + \bar{F} = M_0 \bar{F}.$$

This means that $\bar{F}$ solves the Bellman equation (3.13) such that $\lim_\ell w_{n_{k_\ell}} = W_c^*$ (unique by Proposition 3.2). These convergence results yield the assertion. □

Proof of Theorem 3.2. According to Proposition 3.2 and its proof it is enough to show

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} v(\mathbf{b}'_i, \mathbf{b}'_{i+1}, \mathbf{X}_i) = W_c^* \tag{3.39}$$

a.s. Rule 4 yields

$$\begin{aligned} &w_n + F'_{n+1}(\mathbf{b}'_n, \mathbf{X}_n) \\ &= v(\mathbf{b}'_n, \mathbf{b}'_{n+1}, \mathbf{X}_n) + (1-\delta_n)\mathbb{E}\{F'_n(\mathbf{b}'_{n+1}, \mathbf{X}_{n+1}) \mid \mathbf{b}'_{n+1}, \mathbf{X}_n\}, \end{aligned}$$

where

$$w_n = \max_{\mathbf{b},\mathbf{x}} (M_{\delta_n} F'_n)(\mathbf{b}, \mathbf{x}).$$

Then

$$\begin{aligned} \frac{1}{n}\sum_{i=1}^{n} v(\mathbf{b}'_i, \mathbf{b}'_{i+1}, \mathbf{X}_i) &= \frac{1}{n}\sum_{i=1}^{n} w_i + \frac{1}{n}\sum_{i=1}^{n} \big(F'_{i+1}(\mathbf{b}'_i, \mathbf{X}_i) \\ &\quad -(1-\delta_i)\mathbb{E}\{F'_i(\mathbf{b}'_{i+1}, \mathbf{X}_{i+1}) \mid \mathbf{b}'_{i+1}, \mathbf{X}_i\}\big) \\ &= \frac{1}{n}\sum_{i=1}^{n} w_i \\ &\quad + \frac{1}{n}\sum_{i=1}^{n} \big(F'_i(\mathbf{b}'_{i+1}, \mathbf{X}_{i+1}) - \mathbb{E}\{F'_i(\mathbf{b}'_{i+1}, \mathbf{X}_{i+1}) \mid \mathbf{X}_1^i\}\big) \\ &\quad + \Big[\frac{1}{n}\sum_{i=1}^{n} \big(F'_{i+1}(\mathbf{b}'_i, \mathbf{X}_i) - F'_i(\mathbf{b}'_{i+1}, \mathbf{X}_{i+1})\big) \\ &\quad + \frac{1}{n}\sum_{i=1}^{n} \delta_i \mathbb{E}\{F'_i(\mathbf{b}'_{i+1}, \mathbf{X}_{i+1}) \mid \mathbf{X}_1^i\}\Big] \\ &=: A_n + B_n + C_n. \end{aligned}$$

By Lemma 3.2, $A_n \to W_c^*$. By (3.34) and Chow's theorem $B_n \to 0$ a.s. Further

$$\begin{aligned}|C_n| &\leq \frac{1}{n}\left|\sum_{i=1}^{n-1}\left(F'_{i+2}(\mathbf{b}'_{i+1},\mathbf{X}_{i+1}) - F'_i(\mathbf{b}'_{i+1},\mathbf{X}_{i+1})\right)\right| \\ &\quad + \frac{1}{n}|F'_2(\mathbf{b}'_1,\mathbf{X}_1)| + \frac{1}{n}|F'_n(\mathbf{b}'_{n+1},\mathbf{X}_{n+1})| + \frac{1}{n}\sum_{i=1}^{n}\delta_i K \\ &\to 0\end{aligned}$$

by (3.34) and (3.36) and $\delta_n \to 0$. Thus (3.39) is obtained. □

References

Akien, M., Sulem, A. and Taksar, M. I. (2001). Dynamic optimization of long-term growth rate for a portfolio with transaction costs and logaritmic utility, *Mathematical Finance* **11**, pp. 153–188.

Algoet, P. and Cover, T. (1988). Asymptotic optimality asymptotic equipartition properties of log-optimum investments, *Annals of Probability* **16**, pp. 876–898.

Bellman, R. (1957). *Dynamic Programming* (Princeton University Press, Princeton).

Bobryk, R. V. and Stettner, L. (1999). Discrete time portfolio selection with proportional transaction costs. *Probability and Mathematical Statistics* **19**, pp. 235–248.

Davis, M. H. A. and Norman, A. R. (1990). Portfolio selection with transaction costs, *Mathematics of Operations Research* **15**, pp. 676–713.

Eastham, J. and Hastings, K. (1988). Optimal impulse control of portfolios, *Mathematics of Operations Research* **13**, pp. 588–605.

Fernholz, E. R. (2000). *Stochastic Portfolio Theory* (Springer, New York).

Györfi, L., Lugosi, G. and Udina, F. (2006). Nonparametric kernel based sequential investment strategies, *Mathematical Finance* **16**, pp. 337–357.

Györfi, L. and Schäfer, D. (2003). Nonparametric prediction, in J. A. K. Suykens, G. Horváth, S. Basu, C. Micchelli and J. Vandevalle (eds.), *Advances in Learning Theory: Methods, Models and Applications* (IOS Press, NATO Science Series), pp. 339–354.

Györfi, L., Udina, F. and Walk, H. (2008). Nonparametric nearest-neighbor-based empirical portfolio selection strategies, *Statistics and Decisions* **22**, pp. 145–157.

Györfi, L. and Vajda, I. (2008). On the growth of sequential portfolio selection with transaction cost, in Y.Freund, L.Györfi, G.Turn and Th.Zeugmann (eds.), *Algorithmic Learning Theory* (Springer), pp. 108–122.

Hardy, G. H. (1949). *Divergent Series* (Oxford University Press, London).

Hernández-Lerma, O. and Lassere, J. B. (1996). *Discrete-Time Markov Control Processes: Basic Optimality Criteria* (Springer, New York).

Hinderer, K. (1970). *Foundations of Non-Stationary Dynamic Programming with Discrete Time Parameter* (Springer-Verlag, Berlin).

Iyengar, G. (2002). Discrete time growth optimal investment with costs, URL `http://www.columbia.edu/~gi10/Papers/stochastic.pdf`.

Iyengar, G. (2005). Universal investment in markets with transaction costs, *Mathematical Finance* **15**, pp. 359–371.

Iyengar, G. and Cover, T. (2000). Growths optimal investment in horse race markets with costs. *IEEE Transactions on Information Theory* **46**, pp. 2675–2683.

Kalai, A. and Blum, A. (1997). Universal portfolios with and without transaction costs, in *Proceedings of the 10th Annual Conference on Learning Theory*, pp. 309–313.

Korn, R. (1998). Portfolio optimization with strictly positive transaction cost and impulse control, *Finance and Stochastics* **2**, pp. 85–114.

Merhav, N., Ordentlich, E., Seroussi, G. and Weinberger, M. J. (2002). On sequential strategies for loss functions with memory. *IEEE Transactions on Information Theory* **48**, pp. 1947–1958.

Morton, A. J. and Pliska, S. R. (1995). Optimal portfolio manangement with transaction costs, *Mathematical Finance* **5**, pp. 337–356.

Palczewski, J. and Stettner, L. (2006). Maximization of portfolio growth rate under fixed and proportional transaction cost, *Communications in Information and Systems* **7**, pp. 31–58.

Pliska, S. R. and Suzuki, K. (2004). Optimal tracking for asset allocation with fixed and proportional transaction costs, *Quantitative Finance* **4**, pp. 223–243.

Sass, J. and Schäl, M. (2010). The numéraire portfolio under proportional transaction cost, Working paper.

Schäfer, D. (2002). *Nonparametric Estimation for Financial Investment under Log-Utility*, Ph.D. thesis, Universität Stuttgart, Shaker Verlag.

Shreve, S. E., Soner, H. and Xu, G. (1991). Optimal investment and consumption with two bonds and transaction costs, *Mathematical Finance* **1**, pp. 53–84.

Shreve, S. E. and Soner, H. M. (1994). Optimal investment and consumption with transaction costs, *Annals of Applied Probability* **4**, pp. 609–692.

Stout, W. F. (1974). *Almost sure convergence* (Academic Press, New York).

Taksar, M., Klass, M. and Assaf, D. (1988). A diffusion model for optimal portfolio selection in the presence of brokerage fees, *Mathematics of Operations Research* **13**, pp. 277–294.

Walk, H. and Zsidó, L. (1989). Convergence of the Robbins-Monro method for linear problems in a Banach space, *Journal of Mathematical Analysis and Applications* **139**, pp. 152–177.

Yosida, K. (1968). *Functional Analysis*, 2nd edn. (Springer-Verlag, Berlin).