Journal of Optimization Theory and Applications (2026) 208:104
https://doi.org/10.1007/s10957-025-02926-8

Optimal Strategies in Controlled Markov Chains Under a Stochastic Discount Factor with Applications to Investment Models

A. Y. Golubin[1,2]

Received: 14 December 2024 / Accepted: 23 December 2025

Abstract

This paper proposes a new approach to the investigation of a class of controlled Markov chains with discrete time on a finite interval, which possess some monotonicity and concavity properties related to the transition function and the utility function of the decision maker. By introducing a stochastic discount factor based on the myopic optimal strategy, it is relatively easy to find an optimal strategy in the process modified by the stochastic discount. In a particular case of the model – multi-stage optimal investment problem under "hard" constraints including a call option model – this approach provides a simple way for calculation of optimal investment strategy in the modified process. In other words, the decision maker adopting the stochastic discount factor into the model can essentially simplify the determination of an optimal control for the Markov chain by using the myopic optimal strategies only, i.e., by solving a sequence of one-stage optimization problems without analyzing the more complicated dynamic programming equations.

Keywords Controlled Markov chain · Multi-stage optimal investment · Stochastic discount · Utility function · Call option

Mathematics Subject Classification 60J20 · 91G20

1 Introduction

A core body of research on controlled Markov chains resulted from Howard's book [7]. Since the corresponding equations of dynamic programming are difficult to be solved,

Communicated by Kok Lay Teo.

✉ A. Y. Golubin
agolubin@hse.ru

1 Higher School of Economics, ul. Tallinskaya 34, Moscow, Russia

2 Design Information Technologies Center, Russian Academy of Sciences, Odintsovo, Moscow region, Russia

Published online: 18 February 2026

there appear works on approximate solutions to the problem: [9, section 2.5] studies an approximation of an original controlled Markov processes in continuous time on a finite interval by appropriately chosen controlled finite state Markov chains in discrete time; the author of [17] investigates a similar problem with a countable state space by considering a finite-state and finite-action model as an approximation. Applications of this approach to an inventory model, a cash-balance model, and a growth model are given there. Also, Runge-Kutta method for time discretization in optimal control models has recently been widely used (see, e.g., [18]). Investment processes with stage-by-stage probabilistic constraints are studied in [6, 10], the latter also admits a bankruptcy of the investor. Book [4] deals with financial investment markets, including the study of controlled Markov chains in the frame of options of different types. The authors of [15] studied portfolio insurance strategies to protect investors against market movements by providing an initially specified guarantee during the investment period. As was noted in that paper, this kind of a protection mechanism is especially important for the systems with pension plans modeled by a Markov chain. In [16, section 4.3], an eco-finance network is treated as a time-discrete control problem. It describes the economical interactions between several agents (countries, or companies, etc.) aimed at minimizing their emissions caused by technologies, and using financial means. Note that many papers have been devoted to the study of controlled Markov chains on an infinite time interval with discounted or averaged cost as the functional to be minimized (see, e.g., survey [1] and references therein).

In this paper, we investigate a problem of finding an optimal strategy in controlled Markov chains with discrete time and a finite horizon, that maximizes the expected utility of the final state. The novelty is to simplify the corresponding equations of dynamic programming, i.e., we modify this problem by introducing a stochastic discount factor related to a myopic strategy and then obtain a relatively simple solution to the modified problem. The applications to finance models within the proposed discount approach to investment problems, including the so-called call option, are presented. Practitioners may be interested in the suggested approach since that simplifies the solving of the dynamic programming optimality equations in Markov decision processes with discrete time, which describe the models of: inventory, investment, pension plans, and eco-finance network (see [4, 9, 15–17]). Converting the problems into a sequence of one-stage problems makes them much easier to be solved and, therefore, does not involve essential computation errors. As a result, it seems that the presented work suggests some novel method.

2 Initial Problem Formulation

Consider a homogeneous controlled Markov chain X_t, $t = 0, \dots, T$, with the state space X belonging to R.

Definition 2.1 In the following, the term "state space" is understood as an open convex set X in R containing all possible values of (X_t).

This definition does not seem restrictive, since all possible states of the controlled process (X_t) are included into X. At the same time, such a notion allows for determining the convexity of some subsets of X and differentiation of some functions defined on X, which will be used further.

The dynamics equation is

$$X_{t+1} = F(X_t, a, \xi_t),\ X_0 = x_0,$$

where: $F(\cdot, \cdot, \cdot)$ is a Borel-measurable function, stochastic vectors $\xi_t,\ t = 0, \ldots, T-1$, are independent and identically distributed (i.i.d.), decisions a belong to a convex compact set of admissible decisions $A \subset R^n$ defined as $A = \{a \in R^n : h_j(a) \geq 0,\ j = 1, \ldots, m\}$, and $h_j(a)$ are given continuous concave functions. Suppose

$$\text{the transition function } F(x, a, z) \text{ and } F'_x(x, a, z) > 0 \text{ are continuous in } (x, a) \in X \times A \text{ for any fixed } z \in \mathbf{S}, \tag{1}$$

where $\mathbf{S} = \operatorname{supp} \xi_t$ is the support of ξ_t, i.e., the smallest closed set in R^n such that $P(\xi_t \in \operatorname{supp} \xi_t) = 1$.

Remark 2.1 The assumed property that the function $F(x, a, z)$ is increasing in x is justified by the following: further $X_t = x$ will be regarded as a current capital of the decision maker, so a greater value x naturally leads to a greater value X_{t+1} (in a probabilistic sense).

We suppose that all the random variables are defined on a complete probability space $(\Omega, \mathcal{F}, P)$. On this space we define the natural filtration $(\mathcal{F}_t)_{t \geq 0}$, where $\mathcal{F}_t$ is the sigma-algebra $\sigma(\xi_0, \ldots, \xi_{t-1})$. An admissible strategy is a sequence $\pi = (a_0, \ldots, a_{T-1})$, where a_t is $\mathcal{F}_t$-predictable; for each $t = 0, \ldots, T-1$, the values of a_t belong to A. Denote the set of all admissible strategies by Π.

The problem we begin with is a problem of maximizing the expected utility of the final state

$$\text{maximize } E\, U(X_T) \text{ subject to } \pi \in \Pi. \tag{2}$$

Here and throughout the rest of the paper, the utility function $U(x)$ is continuous, concave, and $U'(x) > 0$ on the related domain.

Definition of a stochastic discount factor.

First of all, define a process of "utilities", $Y_t = U(X_t),\ t = 0, \ldots, T$. Its transition function is $Q(y, a, z) = U\left(F(U^{-1}(y), a, z)\right)$, and $Y_0 = U(x_0)$. In this case, condition (1) is equivalent to

$$\text{the transition function } Q(y, a, z) \text{ and } Q'_y(y, a, z) > 0 \text{ are continuous in}$$

$$(y, a) \in U(X) \times A \text{ for any fixed } z \in \mathbf{S},$$

where $U(X)$ denotes the set of values of the function $U(x),\ x \in X$. To prove the inequality above, note that

$$Q'_y(y, a, z) = \frac{\partial}{\partial y} U\left(F(U^{-1}(y), a, z)\right) =$$

$$U'\left(F(U^{-1}(y), a, z)\right) F'_x(U^{-1}(y), a, z)/U'(U^{-1}(y)) > 0.$$

Define also a myopic Markov strategy $\hat{\pi} = (\hat{a}_0(x), \dots, \hat{a}_{T-1}(x))$, where each decision $\hat{a}_t(x)$ is a solution to the problem

$$\text{maximize } E\,[U(X_{t+1})|X_t = x] \equiv E\,U(F(x, a, \xi_t)) \text{ subject to } a \in A \tag{3}$$

for each fixed $t = 0, \dots, T-1$, and $x \in X$. Evidently, $\hat{a}_t(x)$ are identical functions of x, since $\{X_t\}$ is a homogeneous controlled Markov chain.

Remark 2.2 The introduced myopic strategy is explained by a desire of the decision maker to improve his/her situation only at the next stage, regardless of achieving the maximum of $E\,U(X_T)$ (see(2)). As we will see in section 3, this strategy turns out to be optimal in a discounted modification of problem (2).

In the terms introduced above, problem (3) is equivalent to

$$\text{maximize } E\,[Y_{t+1}|Y_t = y] \equiv E\,U\left(F(U^{-1}(y), a, \xi_t\right) \text{ subject to } a \in A,$$

$t = 0, \dots, T-1$, with $y = U(x)$, where $x = X_t$ is the current state of the process (X_t). This leads to the corresponding myopic strategy in the utility process, $\hat{\pi} = (\hat{a}_0(y), \dots, \hat{a}_{T-1}(y))$, where $y \in U(X)$ – the set of values of $U(x),\ x \in X$.

Let $\hat{Y}_t, t = 0, \dots, T$, be the utility process under the myopic strategy $\hat{\pi}$. By construction, $\hat{Y}_t = U(\hat{X}_t)$, where $(\hat{X}_t)$ is the process under the myopic strategy $\hat{\pi}$. Also, $\hat{Y}_{t+1} = Q(\hat{Y}_t, \hat{a}_t(\hat{Y}_t), \xi_t) \stackrel{def}{=} Q(\hat{Y}_t, \xi_t)$. Thus, $(\hat{Y}_t)$ is a Markov chain with the corresponding transition function $Q(y, z)$.

Define

$$\begin{aligned} g(y) \stackrel{def}{=} E\,[\hat{Y}_{t+1}|\hat{Y}_t = y] = \max_{a\in A} E\,Q(y, a, \xi_t) = \\ \max_{a\in A} E\,U\left(F(U^{-1}(y), a, \xi_t)\right) \end{aligned} \tag{4}$$

as an *optimal one-step utility*, and suppose that

$$g(y) \text{ is concave on } U(X). \tag{5}$$

The stochastic discount factor is the product of derivatives

$$N_T \stackrel{def}{=} \prod_{t=0}^{T-1} g'(\hat{Y}_t).[1] \tag{6}$$

Remark 2.3 (*On economic sense of the discount factor* N_T) Since $\hat{Y}_T = Q\left(\hat{Y}_{T-1}, \xi_{T-1}\right) = Q\left(Q(\hat{Y}_{T-2}, \xi_{T-2}), \xi_{T-1}\right)$ and so on for

[1] Here and further, for definiteness, the notation $g'(y)$ is understood as the right derivative.

$\hat{Y}_{T-3}, \ldots, \hat{Y}_0 = y_0$, formal differentiating the superposition of the functions with respect to the initial utility gives

$$\frac{\partial \hat{Y}_T(\omega, y_0)}{\partial y_0} = \prod_{t=0}^{T-1} Q'_y(\hat{Y}_t, \xi_t).$$

If each multiplier $Q'_y(\hat{Y}_t, \xi_t)$ is averaged with respect to ξ_t, we get

$$\prod_{t=0}^{T-1} g'(\hat{Y}_t) = N_T.$$

Thus, the discount factor N_T can be called *a step-wise marginal utility* (under the myopic $\hat{\pi}$).
In the case of a simple linear dynamics equation $\hat{X}_{t+1}=\hat{X}_t\xi_t$, where ξ_t are i.i.d. stochastic values, with a linear utility function $U(x) \equiv x$, we have
$\hat{Y}_T = y_0 \prod_{t=0}^{T-1} \xi_t$. Denoting $\rho = E\,\xi_t$, we get that the discount factor $N_T = \prod_{t=0}^{T-1} \rho = \rho^T$ – a constant exponentially increasing with T, i.e., a "usual" discount factor widely used in economic models.

Lemma 2.1 *Let (1) and (5) be met. Then, the function $g(y)$ (see (4)) is concave and $g'(y) > 0$.*

Proof The concavity of $g(y)$ is stated in (5), and the inequality $Q'_y(y, a, z) > 0$ is proved above. By, e.g., [3], the directional derivative in y of $\max_{a \in A} Q(y, a, z)$ with a direction $\nu = 1$, i.e., the right derivative, is

$$\frac{\partial \max_{a \in A} Q(y, a, z)}{\partial \nu} = \max_{a \in A(y,z)} \frac{\partial Q(y, a, z)}{\partial y} > 0,$$

where $A(y, z) = \{a' \in A : Q(y, a', z)\} = \max_{a \in A} Q(y, a, z)$. Hence, taking into account the definition of $g(y)$ in (4), we have that $g'(y) > 0$ on the related domain. □

The following lemma provides a sufficient condition for concavity of $g(y)$ in terms of the transition function $Q(y, a, z) = U\left(F(U^{-1}(y), a, z)\right)$ of the chain (Y_t).

Lemma 2.2 *Let $Q(y, a, z)$ be concave in (y, a). Then, the function $g(y)$ (see (4)) is concave.*

Proof Denote $y_\alpha = \alpha y_1 + (1-\alpha)y_2$ and $a_\alpha = \alpha a_1 + (1-\alpha)a_2$. By the assumption, a function $w(y, a) = E\,Q(y, a, \xi_t)$ is concave in (y, a), hence
$\alpha w(y_1, a_1) + (1-\alpha)w(y_2, a_2) \le w(y_\alpha, a_\alpha)$ for any $a_i \in A$ and $\alpha \in [0, 1]$. The right hand side of the inequality does not exceed $\max_{a \in A} w(y_\alpha, a) = g(y_\alpha)$.
Let $a^0(y_i) \in \operatorname{argmax}\{w(y_i, a) \,|\, a \in A\}$, $i = 1, 2$. Then,
$\alpha w(y_1, a^0(y_1)) + (1-\alpha)w(y_2, a^0(y_2)) = \alpha g(y_1) + (1-\alpha)g(y_2) \le g(y_\alpha)$. Thus, $g(y) = \max_{a \in A} w(y, a)$ is concave. □

Consider an example illustrating Lemmas 2.1 and 2.2. Let X_t be the investor capital at moment t, the dynamics equation be

$$X_{t+1} = X_t \langle a, \xi_t \rangle, \; X_0 = x_0 \; > 0,$$

where: $a \in A = \{a \in R^n : a_i \geq 0, \sum_1^n a_i = 1\}$, ξ_t are non-negative i.i.d. stochastic vectors equal to the rates of return, and $\langle \cdot, \cdot \rangle$ denotes the scalar product. The utility function is $U(y) = \ln y$. Then, the transition function of the utility process is $Q(y, a, z) = U\left(F(U^{-1}(y), a, z)\right) = y + \ln\langle a, z \rangle$ which is concave in (y, a) and increasing in y. Note that in the view of Lemma 2.1, the stochastic discount factor

$$N_T = \prod_{t=0}^{T-1} g'(\hat{Y}_t) \; > 0 \text{ almost sure (a.s.).}$$

3 A Modified Problem Formulation

Instead of problem (2), we now turn to the modified problem of maximizing the final utility under the stochastic discount factor,

$$\text{maximize } J^N[\pi] \equiv E\,[U(X_T)/N_T] \text{ subject to } \pi \in \Pi, \tag{7}$$

where N_T is defined in (6).

The very formulation resembles the model with a deterministic inflation discount given in Remark 2.4, but with the stochastic discount N_T introduced. Dividing the goal functional by N_T, we "rid" the model of the step-wise marginal utility under the myopic strategy (see Remark 2.4). As we will see further, problem (7) admits a relatively simple solution, namely, the myopic strategy $\hat{\pi}$.

Theorem 3.1 *Let assumptions (1) and (5) be met. Then, an optimal strategy in problem (7) is the myopic strategy $\hat{\pi}$ defined in (3).*

Proof As before, let $\mathcal{F}_t$ be the sigma-algebra $\sigma(\xi_0, \ldots, \xi_{t-1})$. By construction of $g(y) = E\,[\hat{Y}_{t+1}|\hat{Y}_t = y] = \max_{a \in A} E\, Q(y, a, \xi_t)$, we have that the conditional expectation with respect to $\mathcal{F}_t$ (see, e.g., [4] for the definition) is $E\,[Y_{t+1}|\mathcal{F}_t] \leq g(Y_t)$ and $E\,[\hat{Y}_{t+1}|\mathcal{F}_t] = g(\hat{Y}_t)$ a.s. Hence,
$E\,[Y_{t+1} - \hat{Y}_{t+1}|\mathcal{F}_t] \leq g(Y_t) - g(\hat{Y}_t) \leq g'(\hat{Y}_t)(Y_t - \hat{Y}_t)$ a.s., the latter inequality holds by the stated concavity of $g(y)$. Following this way, we have

$$\begin{aligned} E\,[(Y_{t+1} - \hat{Y}_{t+1})/g'(\hat{Y}_t)|\mathcal{F}_t] &\leq Y_t - \hat{Y}_t, \\ E\,[(Y_t - \hat{Y}_t)/g'(\hat{Y}_{t-1})|\mathcal{F}_{t-1}] &\leq Y_{t-1} - \hat{Y}_{t-1}, \\ &\ldots \end{aligned}$$

Therefore, $Z_t \stackrel{def}{=} (Y_t - \hat{Y}_t)/\prod_{s=0}^{t-1} g'(\hat{Y}_s)$ is such that $E\,[Z_{t+1}|\mathcal{F}_t] \leq Z_t$ a.s., i.e., Z_t is a super-martingale (see, e.g., [4]) under any admissible strategy π. Indeed,

$$E\left[(Y_{t+1} - \hat{Y}_{t+1})/\prod_{s=0}^{t} g'(\hat{Y}_s)|\mathcal{F}_t\right] = E\left[(Y_{t+1} - \hat{Y}_{t+1})|\mathcal{F}_t\right]/\prod_{s=0}^{t} g'(\hat{Y}_s) \leq (Y_t - \hat{Y}_t)/\prod_{s=0}^{t-1} g'(\hat{Y}_s)$$

since $E\left[(Y_{t+1} - \hat{Y}_{t+1})|\mathcal{F}_t\right]/g'(\hat{Y}_t) \leq Y_t - \hat{Y}_t$ a.s. As easily seen, Z_t converts into a martingale under $\pi = \hat{\pi}$, which means in particular that $E\,Y_T \leq E\,\hat{Y}_T$, i.e., the myopic strategy $\hat{\pi}$ is optimal in (7). □

Remark 3.1 Actually, the proof of Theorem 3.2 establishes a result stronger then just the optimality of $\hat{\pi}$. It is proved that $Z_t = (Y_t - \hat{Y}_t)/\prod_{s=0}^{t-1} g'(\hat{Y}_s)$ is a super-martingale under any admissible strategy π, and converts into a martingale under $\pi = \hat{\pi}$. The optimality of $\hat{\pi}$ is only a trivial consequence of this fact.

Note also that an explicit form of the discount factor N_T is not needed for constructing the myopic optimal strategy (see (3)) in (7). Only concavity of the function $g(y)$ and the inequality $g'(y) > 0$ are of importance.

4 The Case of a Non-Homogeneous Controlled Markov Chain

In this case, the dynamics equation is

$$X_{t+1} = F(t, X_t, a, \xi_t),\ X_0 = x_0,$$

where stochastic vectors $\xi_t,\ t = 0, \ldots, T-1$, are independent. The set of admissible decisions a at each moment t with a current state x of the process is a compact convex set $A_{t,x} \subset R^n$ defined as
$A_{t,x} = \{a \in R^n : h_j(t, x, a) \geq 0,\ j = 1, \ldots, m\}$, where $h_j(t, x, a)$ are concave in a and continuous in (x, a).

Suppose that the derivatives

$$\frac{\partial h_j(t,x,a)}{\partial x} \geq 0,\ j = 1, \ldots, m,\ \text{and continuous in } (x, a) \in X \times R^n \text{ for any fixed } t. \tag{8}$$

Suppose also that

$$\text{the transition function } F(t, x, a, z) \text{ and } F'_x(t, x, a, z) > 0 \text{ are continuous in } (x, a). \tag{9}$$

Like in section 2, define the utility process $Y_t = U(X_t)$ with the corresponding transition function $Q(t, y, a, z) = U\left(F(t, U^{-1}(y), a, z)\right)$, where decisions

$a \in \hat{A}_{t,y} \stackrel{def}{=} A_{t,U^{-1}(y)} = \{a \in R^n : h_j(t, U^{-1}(y), a) \geq 0, j = 1, \ldots, m\}$ for $y \in U(X)$. In the terms of the process Y_t, conditions (8) and (9) imply that the same properties hold for $\hat{h}_j(t, y, a) \stackrel{def}{=} h(t, U^{-1}(y), a)$ and $Q(t, y, a, z)$.

Suppose, in addition, that a set

$$\hat{A}_t \stackrel{def}{=} \{(a, y) : a \in \hat{A}_{t,y},\ y \in U(X)\} \text{ is convex for all } t. \tag{10}$$

Note that convexity of $\hat{A}_t$ implies convexity in a of its projection $\hat{A}_{t,y}$ for any fixed (t, y).

As before, the problem we begin with is a problem of maximizing the expected utility of the final state

$$\text{maximize } E\,U(X_T) \text{ subject to } \pi \in \Pi. \tag{11}$$

Definition of a stochastic discount factor.

First of all, define a myopic Markov strategy $\hat{\pi} = (\hat{a}_0(x), \ldots, \hat{a}_{T-1}(x))$, where each decision $\hat{a}_t(x)$ is a solution to the problem

$$\begin{aligned} \text{maximize } E\,[U(X_{t+1})|X_t = x] \equiv E\,U(F(t, x, a, \xi_t)) \\ \text{subject to } a \in A_{t,x},\ t = 0, \ldots, T-1. \end{aligned} \tag{12}$$

Let $\hat{X}_t, t = 0, \ldots, T$, be the process under the myopic strategy $\hat{\pi}$. In the terms of the utility process $Y_t = U(X_t)$, problem (12) is equivalent to

$$\begin{aligned} \text{maximize } E\,[Y_{t+1}|Y_t = y] \equiv E\,U\left(F(t, U^{-1}(y), a, \xi_t)\right) \\ \text{subject to } a \in \hat{A}_{t,y},\ t = 0, \ldots, T-1, \end{aligned}$$

with the set $\hat{A}_{t,y}$ defined above. This leads to a myopic strategy in the utility process, $\tilde{\pi} = (\tilde{a}_0(y), \ldots, \tilde{a}_{T-1}(y))$, where $\tilde{a}_t(y) = \hat{a}_t(U^{-1}(y))$. Thus, $\hat{Y}_t = U(\hat{X}_t), t = 0, \ldots, T$, is the utility process under the myopic strategy $\tilde{\pi}$. Then $\hat{Y}_{t+1} = U\left(F(t, U^{-1}(\hat{Y}_t), \tilde{a}_t(\hat{Y}_t), \xi_t)\right) = Q(t, \hat{Y}_t, \xi_t)$ and $(\hat{Y}_t)$ is a Markov chain with the corresponding transition function $Q(t, y, z)$.

Define

$$g_t(y) \stackrel{def}{=} E\,[\hat{Y}_{t+1}|\hat{Y}_t = y] = \max_{a \in \tilde{A}_{t,y}} E\,Q(t, y, a, \xi_t) \tag{13}$$

as an *optimal one-step utility*, where $y = U(x)$ is the utility of the current capital $X_t = x$.

$$\text{Suppose that } g_t(y) \text{ is concave in } y \text{ on the set } U(X). \tag{14}$$

Lemma 4.1 *Let conditions (8)-(9) and (14) be met, and the Slater condition be satisfied, i.e., for any t and $x \in X$ there exists $a = a_{t,x}$ such that $h_j(t, x, a) > 0$, $j = 1, \ldots, m$.*

Then, the functions $g_t(y)$ defined by (13) are concave in y and

$$g_t'(y) > 0.$$

Proof Concavity of $g_t(y)$ is supposed in (14). In terms of the utility process, the Slater condition above is equivalent to the following: for any $y \in U(X)$ there exists $\tilde{a}_{t,y}$ such that $\tilde{h}_j(t, y, \tilde{a}_{t,y}) = h_j(t, U^{-1}(y), \tilde{a}_{t,y}) > 0$, $j = 1, \dots, m$, where $\tilde{a}_{t,y} = a_{t,U^{-1}(y)}$. According to [14, Theorem 7.3, p. 101], the sign of $g_t'(y)$ coincides with the sign of $L_y'(t, y, \bar{a}, \bar{\mu})$ at some $\bar{a} \in \tilde{A}_{t,y}$ and $\bar{\mu} = (\bar{\mu}_1, \dots, \bar{\mu}_n) \geq 0$. Here, the Lagrangian $L(t, y, a, \mu) = E\, Q(t, y, a, \xi_t) + \sum_{j=1}^m \mu_j \tilde{h}_j(t, y, a)$. Taking into consideration that

$$\frac{\partial E\, Q(t, y, a, \xi_t)}{\partial y} > 0 \text{ (see (9)) and } \frac{\partial \tilde{h}_j(t, y, a)}{\partial y} \geq 0 \text{ (see (8))},$$

we complete the proof. □

The following lemma provides a sufficient condition for concavity of $g_t(y)$ in terms of the transition function $Q(t, y, a, z)$ and sets $\tilde{A}_{t,y}$ of admissible decisions of the chain (Y_t).

Lemma 4.2 *Let $Q(t, y, a, z) = U\left(F(t, U^{-1}(y), a, z)\right)$ be concave in (y, a) and (10) be satisfied. Then, the functions $g_t(y)$ are concave in y at any fixed t.*

Proof By the assumption, a function $w_t(y, a) = E\, Q(t, y, a, \xi_t)$ is concave in (y, a). Fix any $y_1, y_2 \in U(X)$ and $a_1 \in \tilde{A}_{t,y_1}$, $a_2 \in \tilde{A}_{t,y_2}$. Hence,
$\alpha w_t(y_1, a_1) + (1-\alpha) w_t(y_2, a_2) \leq w_t(y_\alpha, a_\alpha)$ for any $\alpha \in [0, 1]$, where we denote $y_\alpha = \alpha y_1 + (1-\alpha) y_2$ and $a_\alpha = \alpha a_1 + (1-\alpha) a_2$. By concavity of $\tilde{A}_t$, the decision $a_\alpha \in \tilde{A}_{t,y_\alpha}$. Therefore,

$$\alpha w_t(y_1, a_1) + (1-\alpha) w_t(y_2, a_2) \leq \max_{a \in \tilde{A}_{t,y_\alpha}} w_t(y_\alpha, a).$$

Since a_1 and a_2 can be chosen arbitrary, we have $g_t(y) = \max_{a \in \tilde{A}_{t,y}} w_t(y, a)$ is concave in y. □

Similar to (6), the stochastic discount factor is the product of derivatives

$$N_T \stackrel{def}{=} \prod_{t=0}^{T-1} g_t'(\hat{Y}_t) > 0 \text{ a.s.} \tag{15}$$

Instead of problem (11), we now study a modified problem of maximizing the final utility under the stochastic discount factor

$$\text{maximize } J^N[\pi] \equiv E\,[U(X_T)/N_T] \text{ subject to } \pi \in \Pi, \tag{16}$$

where N_T is defined in (15).

Theorem 4.1 *Let assumptions (8)-(10) and (14) be met. Then, an optimal strategy in problem (16) is the myopic strategy $\hat{\pi}$ defined in (12).*

Proof The proof repeats the reasoning in the proof of Theorem 3.1. □

5 A Multi-Stage Portfolio Investment Problem with a "Hard" Constraint on the Final Capital

Here the dynamics equation is

$$X_{t+1} = X_t \langle a_t, \xi_t \rangle, \ t = 0, \dots, T-1,$$

$X_0 = x_0 > 0$ is the initial investor's capital, $\langle x, y \rangle$ denotes the scalar product $\sum_{i=1}^{n} x_i y_i$, stochastic vectors $\xi_t, \ t = 0, \dots, T-1$, are positive i.i.d., and ξ_t^i is the rate of return of the i-th asset at moment t. A decision a_t is an n-dimensional vector of proportions of the current capital chosen from the set $A = \{a \in R^n : \ a_i \geq 0, \sum_1^n a_i = 1\}$, that is, the "short sales" are not allowed (see, e.g., [10]). In addition, we impose a constraint on the final capital, $X_T \geq \beta \ (> 0)$ a.s. Introduce a problem

$$\text{maximize } J[\pi] \equiv E\, U(X_T) \text{ subject to } X_T \geq \beta \ \text{ a.s.} \tag{17}$$

In this problem, the vector a_t belongs to a set of admissible decisions $A_{t,x}$, where $t = 0, \dots, T-1$, is the number of stage and x is the current state of the process. The rest of the section is devoted to finding the form of the sets $A_{t,x}$.

Lemma 5.1 *Let* $v \stackrel{def}{=} \max_{a \in A} \min_{z \in \mathbf{S}} \langle a, z \rangle < \infty$, *where* $\mathbf{S} = \operatorname{supp} \xi_t$ *is the support of* ξ_t *(see (1)), i.e., roughly saying, the set of possible values of this stochastic vector. Condition*

$$x_0 > \beta / v^T \tag{18}$$

is necessary and sufficient for non-emptiness of the interiors $Int\, A_{t,x}$ *of the sets of admissible decisions*

$$A_{t,x} = \{a \in A : x \min_{z \in \mathbf{S}} \langle a, z \rangle \geq \beta / v^{T-t-1}\}, t = 0, \dots, T-1. \tag{19}$$

Proof Begin with $A_{T-1,x} = \{a \in A : x \min_{z \in \mathbf{S}} \langle a, z \rangle \geq \beta\}$. It is easily seen that $Int\, A_{T-1,x} \neq \emptyset$ if and only if $x > \beta/v$. Then, the interior of $A_{T-2,x} = \{a \in A : x \min_{z \in \mathbf{S}} \langle a, z \rangle \geq \beta/v\}$ is non empty if and only if $x > \beta/v^2$. Following this reasoning, we have that the interior of $A_{0,x_0} = \{a \in A : x_0 \min_{z \in \mathbf{S}} \langle a, z \rangle \geq \beta/v^{T-1}\}$ is non-empty if and only if $x_0 > \beta/v^T$. □

Remark 5.1 If $\mathbf{S}$ is convex then by the known theorem of the game theory (see, e.g., [5]) we have: since $\langle a, z \rangle$ is concave in a and convex in z, then the constant v in (18) can be easily determined

$$v = \max_{a \in A} \min_{z \in \mathbf{S}} \langle a, z \rangle = \min_{z \in \mathbf{S}} \max_{a \in A} \langle a, z \rangle = \min_{z \in \mathbf{S}} \max\{z_1, \dots, z_n\}.$$

Consider also the case of **the binomial market**, where: $n = 2$, $\xi_t = (m_1, \xi_2^t)$, $P\{\xi_2^t = M\} = \rho$, $P\{\xi_2^t = m\} = 1 - \rho$, and $0 < m < m_1 < M$. In this case, $\mathbf{S} = \operatorname{supp} \xi_t$ consists of two points, (m_1, m) and (m_1, M), and, hence, is not convex. However, an antagonistic game with the gain matrix (see e.g., [5])

$$\mathbf{M} = \begin{pmatrix} m_1 \, m_1 \\ m \, M \end{pmatrix}$$

has a solution in the set of pure strategies because

$$\max_{i=1,2} \min_{j=1,2} m_{ij} = \min_{j=1,2} \max_{i=1,2} m_{ij} = m_1.$$

Hence, we simply have $v = \max_{a \in A} \min_{z \in \mathbf{S}} \langle a, z \rangle = \min_{z \in \mathbf{S}} \max_{a \in A} \langle a, z \rangle = m_1$.

6 A Modified Problem of the Multi-Stage Portfolio Investment with a "Hard" Constraint

Instead of problem (17) which, to the best of the author's knowledge, is not solved fully, i.e., for a general distribution of ξ_t, we now turn to the corresponding modified problem.

$$\text{maximize } J^N[\pi] \equiv E\,[U(X_T)/N_T] \text{ subject to } X_T \geq \beta \text{ a.s.,} \tag{20}$$

where: $U(x)$ is the utility function, N_T defined in (15) is

$$N_T = \prod_{t=0}^{T-1} g_t'(\hat{Y}_t), \;\; \text{where } g_t(y) = E\,U\left(U^{-1}(y)\langle \hat{a}_t(U^{-1}(y)), \xi_t\rangle\right),$$

and $\hat{\pi} = (\hat{a}_0, \ldots, \hat{a}_{T-1})$ is the myopic strategy,

$$\hat{a}_t(x) \in \operatorname{argmax}\{E\,U(x\langle a, \xi_t\rangle) \,|\, a \in A_{t,x}\}, \tag{21}$$

with $A_{t,x}$ defined in (19). As before, define a utility process $Y_t = U(X_t)$, $t = 0, \ldots, T$, with the transition function

$Q(t, y, a, z) = U\left(F(t, U^{-1}(y), a, z)\right) = U\left(U^{-1}(y)\langle a, z\rangle\right)$ and the sets of admissible decisions

$$\tilde{A}_{t,y} = A_{t,U^{-1}(y)} = \{a \in A : U^{-1}(y) \min_{z \in \mathbf{S}} \langle a, z \rangle \geq \beta / v^{T-t-1}\}.$$

In terms of the utility process $Y_t = U(X_t)$, problem (12) is equivalent to

$$\text{maximize } E\,[Y_{t+1} | Y_t = y] \equiv E\,U\left(F(t, U^{-1}(y), a, \xi_t)\right) \text{ subject to } a \in \tilde{A}_{t,y},$$

$t = 0, \ldots, T-1$. This leads to a myopic strategy in the utility process, $\tilde{\pi} = (\tilde{a}_0(y), \ldots, \tilde{a}_{T-1}(y))$, where $\tilde{a}_t(y) = \hat{a}_t(U^{-1}(y))$ (see (21)). Thus, $\hat{Y}_t = U(\hat{X}_t)$, $t = 0, \ldots, T$, is the utility process under the myopic strategy $\tilde{\pi}$.

Now we study the convexity property of the set $\tilde{A}_t$ (condition (10)), which in this case has the form $\tilde{A}_t = \{(a, y) : U^{-1}(y) \min_{z \in \mathbf{S}} \langle a, z \rangle \geq \beta / v^{T-t-1}\}$. Let us recall the definition of risk aversion function [11]: let $U(x)$ be twice differentiable, the risk aversion function is

$$R(x) \stackrel{def}{=} -\frac{U''(x)}{U'(x)}.$$

Lemma 6.1 *If $U(x)$ is such that*

$$R(x) \leq 2/x \text{ on } X,$$

then $\tilde{A}_t$ is convex in (y, a).

Proof Rewrite the expression for $\tilde{A}_t$ as

$$\{(a, y) : \min_{z \in \mathbf{S}} \langle a, z \rangle - \frac{1}{U^{-1}(y)} \beta / v^{T-t-1} \geq 0\} =$$
$$\{(a, y) : \min_{z \in \mathbf{S}} \langle a, z \rangle + u(U^{-1}(y)) \beta / v^{T-t-1} \geq 0\},$$

where $u(x) = -1/x$ with its risk aversion function $R_u(x) = 2/x$. According to [11], $u(U^{-1}(y))$ is concave as $R(x) \leq R_u(x)$. Noting that $\min_{z \in \mathbf{S}} \langle a, z \rangle$ is a concave function in a, we complete the proof. □

Let us verify the fulfillment of conditions (8)-(10) and (14) in Theorem 4.1 for the considered case. Condition (8) evidently holds (see (19)). Condition (9) of the increasing in x of $F(t, x, a, z)$ is satisfied (see the dynamics equation in section 5). Condition (10) is met by virtue of Lemma 6.1. Condition (14) is guaranteed by the following assumption (see Lemma 4.2):
$U(U^{-1}(y)\langle a, z \rangle)$ is concave in (y, a), which is evidently satisfied if the utility function $U(x)$ is such that

$$U(U^{-1}(y)w) \text{ is concave in } (y, w) \in U(X) \times (0, \infty). \tag{22}$$

Theorem 6.1 *Let assumptions (18), the inequality $R(x) \leq 2/x$ on X, and (22) be met. Then an optimal strategy in problem (20) is the myopic strategy $\hat{\pi}$ defined in (21).*

Proof According to the theorem's assumptions, the sets $\tilde{A}_{t,y}$ have non-empty interiors, $\tilde{A}_t$ are convex, and the function $Q(t, y, a, z)$ is concave in (y, a), with $Q'_y(t, y, a, z) > 0$. The latter implies that $N_T > 0$ a.s. So, the proof repeats the reasoning in the proof of Theorem 4.1. □

6.1 The Case of a Logarithmic Utility Function

Let $U(x) = \ln x$. Then $U(U^{-1}(y)w) = y + \ln w$, so (22) is satisfied and $R(x) = 1/x < 2/x$. By Theorem 6.1, if (18) holds then the myopic strategy $\hat{\pi}$ solves problem (20), where now

$$N_T = \prod_{t=0}^{T-1} g_t'(\hat{Y}_t),\ \ g_t(y) = y + E\ \ln\langle \hat{a}_t(e^y), \xi_t\rangle,$$

$\hat{\pi} = (\hat{a}_0, \dots, \hat{a}_{T-1})$ is the myopic strategy defined by

$$\hat{a}_t(x) \in \operatorname{argmax}\{E\ \ln\langle a, \xi_t\rangle \,|\, a \in A_{t,x}\}, \tag{23}$$

the sets of admissible decisions $A_{t,x}$ are defined in (19).

The case of a two-asset market. Consider a particular kind of problem (20), where the number of assets $n = 2$, the first is risk-less with a return $\xi_1^t = m_1$, the second is risky with a stochastic return ξ_2^t such that $0 < m \overset{def}{=} \inf\{\operatorname{supp}\xi_2^t\} < m_1 < M \overset{def}{=} \sup\{\operatorname{supp}\xi_2^t\} \le \infty$. The feasibility condition in (18), $x_0 \ge \beta/v^T$, is now $x_0 \ge \beta/m_1^T$. Notice that according to Definition 2.1, the "state space" X of this chain is supposedly open and convex, so the operation of differentiation is admissible for appropriate functions on X.

Denote $\min\{x, y\} = x \wedge y$ and $\max\{x, 0\} = (x)_+$. The optimization problem for finding a myopic strategy takes the form

$$\begin{gathered}\text{maximize } E\ln[(1-a)m_1 + a\xi_t^2] \\ \text{subject to } 0 \le a \le 1 \wedge \left\{\left(m_1 - \frac{\beta}{x m_1^{T-t-1}}\right)/(m_1 - m)\right\},\end{gathered} \tag{24}$$

where the constraint (24) on a is equivalent in this case to (19).

This problem has an explicit solution in a particular case of a **binomial market**, where the risky return is such that $P\{\xi_2^t = M\} = \rho$, $P\{\xi_2^t = m\} = 1 - \rho$, and $0 < m < m_1 < M$. Since the objective function at any fixed t is concave in a on the non-empty interval, we get

$$\hat{a}_t(x) = (1 \wedge (a^*)_+) \wedge \left\{\left(m_1 - \frac{\beta}{x m_1^{T-t-1}}\right)/(m_1 - m)\right\},\ t = 0, \dots, T-1,$$

$$\text{where } a^* = m_1\left(\frac{\rho}{m_1 - m} - \frac{1-\rho}{M - m_1}\right)$$

is the point at which the derivative of the objective function becomes zero. According to Theorem 6.1, if $x_0 \ge \beta/m_1^T$ then the optimal in (20) myopic portfolios are $(1 - \hat{a}_t(x), \hat{a}_t(x))$, $t = 0, \dots, T-1$, with $\hat{a}_t(x)$ determined above.

6.2 The Case of a Linear Utility Function

Let $U(x) = x$. The discount factor in this case is

$$N_T = \prod_{t=0}^{T-1} g_t'(\hat{Y}_t),\ \ g_t(y) = y\langle \hat{a}_t(y), E\,\xi_t\rangle, \text{ where } y = x,$$

$\hat{\pi} = (\hat{a}_0, \ldots, \hat{a}_{T-1})$ is the myopic strategy,

$$\hat{a}_t(x) \in \operatorname{argmax}\{\langle a, E\,\xi_t\rangle \mid a \in A_{t,x}\},$$

with $A_{t,x}$ defined in (19).

Since $U(U^{-1}(y)w) = yw$, (22) is not satisfied. However, below we directly prove concavity of the function $g_t(y)$ that can be rewritten as

$$g_t(y) = \max_{a \in N(y)} L(a),$$

where the sets $N(y) = \{a \in R^n : a_i \geq 0, \sum_1^n a_i = y, \min_{z \in S}\langle a, z\rangle \geq \beta/v^{T-t-1}\}$ and the function $L(a) = \langle a, E\,\xi_t\rangle$.

Lemma 6.2 *The function $g_t(y)$ is concave in $y \in X$.*

Proof Let $y_1, y_2 \in X$ and $a_1 \in N(y_1)$, $a_2 \in N(y_2)$. Since $L(a)$ is concave,

$$\alpha L(a_1) + (1-\alpha)L(a_2) \leq L(\alpha a_1 + (1-\alpha)a_2) \text{ for any } \alpha \in [0, 1]. \tag{25}$$

As $\alpha a_1 + (1-\alpha)a_2 \in N(\alpha y_1 + (1-\alpha)y_2)$, from (25) we have
$\alpha L(a_1) + (1-\alpha)L(a_2) \leq g_t(\alpha y_1 + (1-\alpha)y_2)$. Hence
$\alpha g_t(y_1) + (1-\alpha)g_t(y_2) \leq g_t(\alpha y_1 + (1-\alpha)y_2)$. □

Thus, if (18) holds, then the myopic strategy is optimal in (20) in the case of the linear utility function.

7 The Case of the Call Option in ($B - S$) market

The call option is a contract between the buyer and the seller of the call option to exchange a security at an agreed price [8]. The buyer of the call option has the right, but not the obligation, to buy an agreed quantity of a particular commodity or financial instrument from the seller of the option at or before time T (the expiration date) for a certain price. The buyer has no profit if he/she requests a payment at a moment before a constant K (a strike price). In the interval $[K, T]$, the buyer receives the profit $S_T - K$, where S_T is the price of the risky asset. Thus, the function of the buyer's profit (the payment function) is $(s-K)_+ \stackrel{def}{=} \max\{0, s-K\}$.

Similar to the case of the two-asset market, assume that the market has only two assets with rates of return $\xi_t = (\xi_1^t, \xi_2^t)$. The first are risk-less, $\xi_1^t = m_1$ a.s., ξ_2^t are i.i.d. and such that
$0 < m \overset{def}{=} \inf\{\operatorname{supp} \xi_2^t\} < m_1 < M \overset{def}{=} \sup\{\operatorname{supp} \xi_2^t\} < \infty$ – in distinction to the discrete stochastic values ξ_2^t introduced in the binomial market.

Define the dynamics equations for this model:
the investor's capital is $X_{t+1} = X_t\langle a_t, \xi_t\rangle,\ X_0 = x_0 > 0$,
the price for the risk-less asset is $B_{t+1} = B_t m_1,\ B_0 = m_1 > 0$,
the price for the risky asset is $S_{t+1} = S_t\xi_2^t,\ S_0 = s_0 > 0$.
Let the payment function be $f(s) = (s - K)_+$, and the investor's utility function be $U(x) = \ln x$.

The investigated problem is the maximization of the expected utility of the final buyer's capital under the stochastic discount

$$\text{maximize } E\left[\ln(X_T)/N_T\right] \text{ subject to } X_T \geq f(S_T) \text{ a.s.}, \tag{26}$$

$$N_T = \prod_{t=0}^{T-1} \frac{\partial g_t(\hat{Y}_t, S_t)}{\partial y},\ \ g_t(y, s) = y + E\ \ln\langle \hat{a}_t(e^y, s), \xi_t\rangle \text{ (see section 6.1).}$$

Here $Y_t = \ln(X_t)$, $\hat{Y}_t$ is the utility process under a myopic strategy
$\tilde{\pi} = (\hat{a}_0(e^y, s), \ldots, \hat{a}_{T-1}(e^y, s))$ with $\hat{a}_t(x, s) = (1 - \hat{a}_t^1(x, s), \hat{a}_t^1(x, s))$ such that

$$\hat{a}_t^1(x, s) \in \operatorname{argmax}\{E \ln[(1 - a)m_1 + a\xi_t^2] \mid a \in A_{t,x,s}\}. \tag{27}$$

The sets $A_{t,x,s}$ of admissible decisions under current t, $X_t = x$, and $S_t = s$ are found in the following lemma, where we denote $x \vee y = \max\{x, y\}$ and $x \wedge y = \min\{x, y\}$.

Lemma 7.1

$$A_{t,x,s} = \left[0 \vee \frac{m_1^{t+1} d(t, sM) - xm_1}{x(M - m_1)},\ \frac{xm_1 - m_1^{t+1} d(t, sm)}{x(m_1 - m)} \wedge 1\right], \tag{28}$$

all the sets $A_{t,x,s}$ have non-empty interiors $Int\ A_{t,x,s}$ if and only if $x_0 > m_1 d(0, s_0)$, where:

$$d(t, s) \overset{def}{=} \frac{1}{m_1^{T+1}} \sum_{i=0}^{T-t} C_{T-t}^i q^i (1 - q)^{T-t-i} f(sm^i M^{T-t-i}), t = 0 \ldots, T - 1,$$

$C_j^i = j!/[(j - i)!i!]$*, and* $q = (M - m_1)/(M - m)$.

Proof [2] According to [4, Corollary 7.9], and taking into account the convexity of $f(s) = (s - K)_+$ and Jensen's inequality, we have that for any $0 \leq t \leq T - 1$ the

[2] The formulation and proof of Lemma 7.1 are given by E.A. Shelemech

minimal investor's capital X_t^* at moment t with $S_t = s$ guaranteeing the inequality $X_T \geq f(S_T)$ a.s. is such that

$$X_t^*/m_1^{t+1} = d(t, s)$$

(see the definition of $d(t, s)$ above). Due to the reasoning at the beginning of the proof, a decision $a \in [0, 1]$ is admissible if and only if

$$0 \leq \min_{z \in [m, M]} \frac{x[(1-a)m_1 + az]}{m_1^{t+1}} - d(t, sz) =$$

$$\min \left\{ \frac{x[(1-a)m_1 + am]}{m_1^{t+1}} - d(t, sm), \ \frac{x[(1-a)m_1 + aM]}{m_1^{t+1}} - d(t, sM) \right\}, \quad (29)$$

$t = 0, \ldots, T-1$. The latter equality follows from that the minimized function is concave in z. Trivial rearrangements of (29) give the lemma proof. □

Remark 7.1 The parametrization of the set $A_{t,x}$ introduced in sections 5-6 and that of the set $A_{t,x,s}$ are different: the former uses $t \in \{0, \ldots, T-1\}$ and $x \in X$, while the latter uses the three-dimensional parameter (t, x, s) with $s \in \operatorname{supp} S_t$. However, when studying properties of these sets under the fixed parameters, we can exploit a similarity of problems (20) and (26), and use Theorem 6.1, after verifying its assumptions with respect to the considered case.

The set of admissible decisions $A_{t,x,s}$ in (28) can be represented by inequalities $h_j(t, x, s, a) \geq 0$ for the corresponding functions h_j, $j = 1, \ldots, 4$:

$$\begin{aligned} a &\geq 0, \\ \frac{m_1^{t+1} d(t, sM) - x m_1}{x(M - m_1)} &\geq 0, \\ -\frac{x m_1 - m_1^{t+1} d(t, sm)}{x(m_1 - m)} &\geq 0, \\ 1 - a &\geq 0. \end{aligned}$$

As easily seen, an analogue of (8) holds, the derivatives

$$\frac{\partial h_j(t, x, s, a)}{\partial x} \geq 0, \ j = 1, \ldots, 4,$$

and continuous in $(x, a) \in X \times R$ for any fixed t and s. Note also that the Slater's condition is met under fixed t, x, s.

From the expression of dynamics equation for the investor's capital, we have that (9) is satisfied, i.e., the transition function

$F(t, x, a, z)$ and $F'_x(t, x, a, z) > 0$ are continuous in (x, a). In terms of the utility process Y_t, the statement of Lemma 7.1 is rewritten as

$$\tilde{A}_{t,y,s} = A_{t,e^y,s} = \left[0 \vee \frac{m_1^{t+1} d(t, sM) - e^y m_1}{e^y (M - m_1)}, \ \frac{e^y m_1 - m_1^{t+1} d(t, sm)}{e^y (m_1 - m)} \wedge 1 \right].$$

Evidently, the interior of $\tilde{A}_{t,y,s}$ is not empty if and only if $y_0 > \ln(m_1 d(0, s_0))$. The next lemma establishes that an analogue of condition (10) takes place.

Lemma 7.2 *Let $\tilde{A}_{t,s} \overset{def}{=} \{(y, a) : y \in U(X), a \in \tilde{A}_{t,y,s}\}$, where $\tilde{A}_{t,y,s}$ is defined above. Then, the set $\tilde{A}_{t,s}$ is convex.*

Proof As is easily seen, the set $\tilde{A}_{t,y,s}$ can be represented as

$$0 \vee (b/e^y - c) \le a \le (-g/e^y + w) \wedge 1,$$

where b, c, g, and w are corresponding positive constants that do not depend on y and a. Let $(y_i, a_i) \in \tilde{A}_{t,s}$, $i = 1, 2$. Denote by
$(y_\alpha, a_\alpha) = (\alpha y_1 + (1 - \alpha) y_2, \alpha a_1 + (1 - \alpha) a_2)$, $\alpha \in [0, 1]$, and prove that

$$0 \vee (b/e^{y_\alpha} - c) \le a_\alpha \le (-g/e^{y_\alpha} + w) \wedge 1 \tag{30}$$

for any $\alpha \in [0, 1]$. The left inequality holds as the function $b/\exp(y) - c$ is convex in y. The right inequality holds since $-g/\exp(y) + w$ is concave in y. □

Remark 7.2 The proved convexity property of $\tilde{A}_{t,s}$ under the logarithmic utility function can be proved for a more general case of the utility function $U(x)$ such that its risk aversion function $R(x) \le 2/x$ (recall that for $U(x) = \ln x$ its $R(x) = 1/x < 2/x$). Indeed, following the reasoning in Lemma 6.1, we have (see (30)) that the function $b/U^{-1}(y) - c$ is convex in y and the function $-g/U^{-1}(y) + w$ is concave in y.

Theorem 7.1 *Let*

$$x_0 > m_1 d(0, s_0), \tag{31}$$

where the function $d(t, s)$ is given in Lemma 7.1. Then, an optimal strategy in problem (26) is the myopic strategy $\hat{\pi}$ defined in (27).

Proof Condition (22) is met (see the beginning of section 6.1) as the investor's utility function $U(x) = \ln x$. Condition (18) converts into (31) as follows from Lemma 7.1. As was shown above, the analogues of conditions (8), (9), (10) are satisfied in our case. Repeating the reasoning in the proof of Theorem 6.1, we complete the proof. □

In the particular case of **the two-asset market** described in section 6.1, where $P(\xi_2^t = M) = \rho$ and $P(\xi_2^t = m) = 1 - \rho$, the optimal myopic portfolios (see (27)) have an explicit form $\hat{\pi}(x, s) = (1 - \hat{a}_t^1(x, s), \hat{a}_t^1(x, s))$ given below. Denote the point at which the derivative of the objective function

$$E \ \ln[[(1 - a) m_1 + a \xi_2^t] = \ln[(1 - a) m_1 + am](1 - \rho) + \ln[(1 - a) m_1 + aM]\rho$$

vanishes by $a^* \stackrel{def}{=} m_1 \left[\frac{\rho}{m_1-m} - \frac{1-\rho}{M-m_1}\right]$. Since the objective function is concave in a, the optimal decisions are

$$\hat{a}_t^1(x,s) = \{a_1^*(t,x,s) \vee a^*\} \wedge a_2^*(t,x,s),$$

where (see (28))

$$a_1^*(t,x,s) \stackrel{def}{=} 0 \vee \frac{m_1^{t+1} d(t,sM) - x m_1}{x(M-m_1)},$$

$$a_2^*(t,x,s) \stackrel{def}{=} \frac{x m_1 - m_1^{t+1} d(t,sm)}{x(m_1-m)} \wedge 1.$$

8 The Second Type of the Discount Factor

In the frame of the optimal control problem of non-homogeneous Markov chain defined in section 4, we introduce another myopic strategy,
$\tilde{\pi} = (\tilde{a}_0(x), \ldots, \tilde{a}_{T-1}(x))$, where each decision $\tilde{a}_t(x)$, $t = 0, \ldots, T-2$, is a solution to the problem

$$\text{maximize } E\,[X_{t+1}|X_t = x] \equiv E\,F(t,x,a,\xi_t) \text{ subject to } a \in A_{t,x}, \tag{32}$$

and the final decision $\tilde{a}_{T-1}(x)$ is a solution to the problem

$$\text{maximize } E\,[U(X_T)|X_{T-1} = x] \equiv E\,U(F(T-1,x,a,\xi_{T-1})) \\ \text{subject to } a \in A_{T-1,x}. \tag{33}$$

As compared with the myopic strategies introduced in (3) and (12), this approach is "more straightforward": The decision maker seeks just the largest mean capital at the next stage up to the final stage, while in the previous models, he/she seeks the largest expected utility of the capital at each next stage.
As before, we assume that the utility function $U(x)$ is concave and $U'(x) > 0$ on X, and conditions (8),(9), along with the Slater condition (see Lemma 4.1) are satisfied for any $t = 0, \ldots, T-1$. Also, an analogue of condition (10) is supposed to be met, i.e., the set

$$A_t = \{(a,x) : a \in R^n, h_j(t,x,a) \geq 0, j = 1, \ldots m, x \in X\} \\ \text{is convex for all } 0 \leq t \leq T-1. \tag{34}$$

The condition in Lemma 4.2,

$$U\left(F(t, U^{-1}(y), a, z))\right) \text{ is concave in } (y,a) \text{ at each fixed } t, z,$$

converts into

$$F(t,x,a,z) \text{ is concave in } (x,a) \text{ for } t = 1, \ldots, T-2, \tag{35}$$

$$\text{and } U(F(T-1, x, a, z)) \text{ is concave in } (x, a). \tag{36}$$

Introduce the functions $r_t(x) \overset{def}{=} E[\tilde{X}_{t+1}|\tilde{X}_t = x], t = 0, \dots, T-2$, and $r_{T-1}(x) \overset{def}{=} E[U(\tilde{X}_T)|\tilde{X}_{T-1} = x]$, where $\tilde{X}_t$ is a process under the strategy $\tilde{\pi}$ defined above. Following the reasoning in the proofs of Lemmas 4.1-4.2, we have that $r_t(x)$ are concave in x, and $r_t'(x) > 0,\ t = 0, \dots, T-1$.

Let us study a problem of maximizing the expected discounted utility of the investor's final capital (see section 4)

$$\max_{\pi} J^M[\pi] \equiv E[U(X_T)/M_T], \tag{37}$$

where the discount factor $M_T = \prod_{t=0}^{T-1} r_t'(\tilde{X}_t) > 0$ a.s.

Theorem 8.1 *Let conditions (8)-(9) and (34)-(36) be met. Then, an optimal strategy in problem (37) is the myopic strategy $\tilde{\pi}$ defined in (32)-(33).*

Proof The proof repeats the arguments in the proof of Theorem 4.1. □

Remark 8.1 Consider applications of the suggested approach to the multi-stage investment models with "hard" constraints, which are studied in sections 6-7. Now, the utility functions over stages 0, ..., $T-1$, are linear and at the final stage T the function $U(x)$ is supposed to be such that its risk aversion function $R(x) \le 2/x$ (see Lemma 6.1). Note that this inequality holds for the linear utility function, where $R(x) \equiv 0 \le 2/x$. To sum up, under conditions (8)-(9) and (34)-(36), the optimal decisions in the discounted problem (37) with the constraint $X_T \ge \beta$ a.s. are

$$\tilde{a}_t(x) \in \operatorname{argmax}\{\langle a, E\,\xi_t\rangle \,|\, a \in A_{t,x}\},$$

$t = 0, ..., T-2$, with $A_{t,x}$ defined in (19). The final decision $\tilde{a}_{T-1}(x) \in \operatorname{argmax}\{E\,U(x\langle a, \xi_T\rangle) \,|\, a \in A_{T-1,x}\}$ (see section 6).

Turn to the call option in the (B-S) market described in section 7. Below, we follow the notation introduced in that section. As was proved, the sets $A_{t,x,s}$ of admissible decisions at a current moment t, with a state $X_t = x$, and a current price of the risky asset $S_t = s$ are defined in (28). Let $x_0 > m_1 d(0, s_0)$ (see (31)), which means the non-emptiness of the interior of $A_{t,x,s}$. Under the linear utility function over stages 0, ..., $T-1$, the optimal myopic portfolios $(1-\tilde{a}_t^1, \tilde{a}_t^1)$ are such that the shares invested to the risky asset equal the right boundaries of the intervals $A_{t,x,s}$, i.e.,

$$\tilde{a}_t^1 = \frac{x_{m_1} - m_1^{t+1} d(t, sm)}{x(m_1 - m)} \wedge 1, t = 0, ..., T-2.$$

The final decision is determined as

$$\tilde{a}_{T-1}^1 \in \operatorname{argmax}\{E \ln[(1-a)m_1 + a\xi_{T-1}^2] \,|\, a \in A_{T-1,x,s}\}.$$

Let us examine the multi-stage investment model with the "hard" constraint $X_T \geq \beta$ a.s. and the logarithmic utility function $U(x) = \ln x$ in the two-asset market described in section 6.1. Recall that the sets of admissible decisions are

$$A_{t,x} = \{0 \leq a \leq 1 \wedge \{\left(m_1 - \frac{\beta}{x m_1^{T-t-1}}\right)/(m_1 - m)\},$$

$t = 0, \ldots, T-1$. Problem (32) converts into

$$\text{maximize } x(m_1 + a(E\,\xi_2^t - m_1)) \text{ subject to } a \in A_{t,x}.$$

It is naturally supposed that the mean of the risky asset $E\,\xi_2^t > m_1$, where m_1 is the return of the risk-less asset. Thus, the optimal myopic decisions in the strategy $\tilde{a}_t(x) = (1 - \tilde{a}_t^1(x), \tilde{a}_t^1(x))$ are

$$\tilde{a}_t^1(x) = 1 \wedge \{\left(m_1 - \frac{\beta}{x m_1^{T-t-1}}\right)/(m_1 - m),\ t = 0, \ldots, T-2, \tag{38}$$

with $\tilde{a}_t^1(x)$ being the right boundary of the interval $A_{t,x}$. The final decision $\tilde{a}_{T-1}^1(x)$ (see (33)) is a solution to the problem

maximize $E \ln x[(1-a)m_1 + a\xi_2^{T-1}]$ subject to $a \in A_{T-1,x}$, which can be rewritten as

$$\text{maximize} \int_m^M \ln(m_1(1-a) + ya)dF(y) \text{ subject to } a \in A_{T-1,x}, \tag{39}$$

where $F(y)$ is the distribution function of ξ_2^{T-1}.

9 Examples

In this section, we present an illustrative example of the study of the multi-stage investment model for the binomial market with the"hard" constraint, $X_T \geq \beta$ a.s., described in section 6.1. The goals are: to calculate the myopic optimal strategy in the problem (see (20))

$$\text{maximize } J^N[\pi] \equiv E\,[U(X_T)/N_T] \text{ subject to } X_T \geq \beta \text{ a.s.}, \tag{40}$$

and to calculate the myopic optimal strategy of the second type defined in section 8 (see (39)-(38)) in the problem with the maximization of the mean current capitals up to $T-2$ stage, and maximization of the final utility,

$$\text{maximize } J^M[\pi] \equiv E\,[U(X_T)/M_T] \text{ subject to } X_T \geq \beta \text{ a.s.} \tag{41}$$

Let the initial capital $x_0 = 1$, the risk-less asset has a return $\xi_1^t = m_1 = 1.15$ a.s., the risky asset ξ_2^t be such that $P\{\xi_2^t = M\} = \rho$, $P\{\xi_2^t = m\} = 1 - \rho$, where

Table 1

$x\backslash t$	1	2	3	4
1.5	0.4205	0.2182	0	0
2	0.4587	0.4587	0.4314	0.2308
2.5	0.4587	0.4587	0.4587	0.4587
3	0.4587	0.4587	0.4587	0.4587

Table 2

$x\backslash t$	1	2	3	4
1.5	0.4205	0.2182	0	0
2	0.7577	0.6059	0.4314	0.2308
2.5	0.9600	0.8386	0.6990	0.4587
3	1	0.9937	0.8774	0.4587

$m = 0.5, M = 2.5$, and $\rho = 0.5$. The investor's utility function is logarithmic, $U(x) = \ln x$, and the length of the horizon is $T = 5$.

Table 1 presents the optimal decisions $\hat{a}_t^1(x)$ in the myopic strategy $\hat{\pi} = (\hat{a}_0(x), \ldots, \hat{a}_4(x))$ for problem (40), where $\hat{a}_t(x) = (1 - \hat{a}_t^1(x), \hat{a}_t^1(x))$. In the table, the moment t of decision making runs over $\{1, \ldots, 4\}$. The values of the current states $X_t = x \in \{1.5, 2, 2.5, 3\}$ are chosen just to illustrate the dynamics of decision making, the optimal decisions are the solutions to problems (24). At the moment $t = 0$ under the fixed initial capital $x_0 = 1$, the optimal decision $\hat{a}_0^1(1) = 0.01$ is the right boundary of the interval $A_{0,1}$ (see (24)).
Note that the value

$$\hat{a}^1 = 0.4587 = m_1 \left(\frac{\rho}{m_1 - m} - \frac{1 - \rho}{M - m_1} \right)$$

is the point at which the derivative of the objective concave function in (24) vanishes. The other entries of Table 1 are either the left boundary (=0) or the right boundaries of the intervals $A_{t,x}$ defined in (24). Quite expected, the interval $A_{t,x}$ shrinks as t increases at a fixed x. On the other hand, $A_{t,x}$ becomes wider as the current capital x grows at a fixed t, which leads to that the interval contains the point 0.4587 – the zero of derivative of
$E \ln[(1 - a)m_1 + a\xi_t^2]$ – for any t if x is 2.5 or 3 (see the last two rows of the Table 1).

Table 2 presents the optimal decisions $\tilde{a}_t^1(x)$ in the myopic strategy for problem (41) under the second kind of the discount factor, where, as before, t runs over $\{1, \ldots, 4\}$ and x runs over $\{1.5, \ldots, 3\}$. Similar to problem (39), we have at $t = 0$ and with the prescribed initial capital $x_0 = 1$ that the optimal decision in this case is $\tilde{a}_0^1(1) = 0.01$.

By construction (see (32)), the optimal decisions $\tilde{a}_t^1(x)$ coincide with the right boundaries of the intervals $A_{t,x}$ for $t = 0, \ldots, 3$. When $t = 4$, $\tilde{a}_4^1(x)$ is a solution to the problem of maximizing $E \ln[(1 - a)m_1 + a\xi_4^2]$ subject to $a \in A_{4,x}$. Similar to

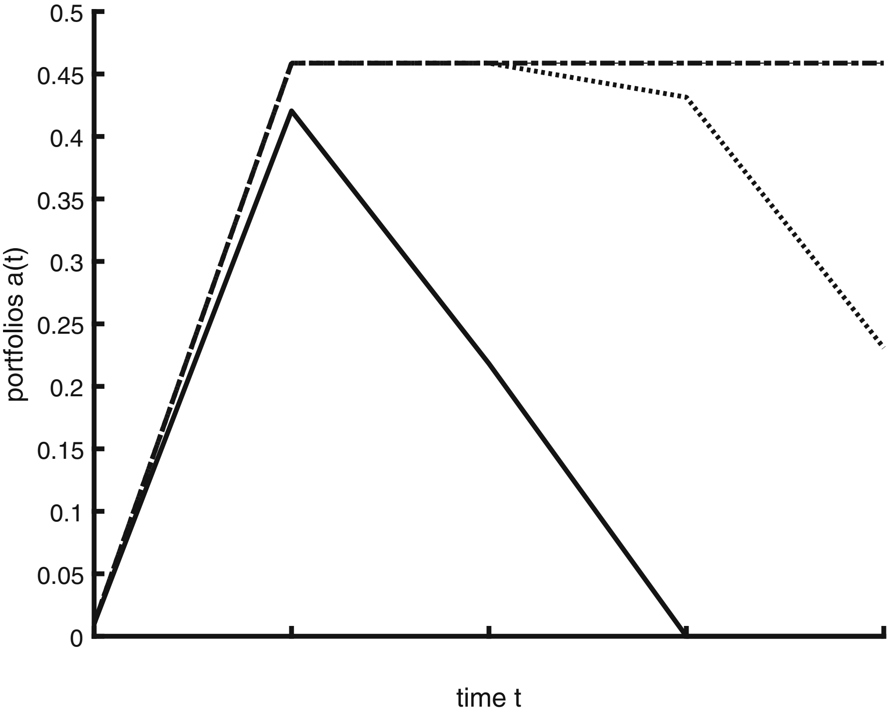

Fig. 1 Optimal shares $\hat{a}_t^1(x)$ of capital invested to the risky asset under a varying current capital x: $\hat{a}_t^1(1.5)$ – solid line, $\hat{a}_t^1(2)$ – point line, $\hat{a}_t^1(2.5)$ – dashed line, $\hat{a}_t^1(3)$ – dash-dot line

Table 1, if x equals 2.5 or 3, the intervals $A_{4,x}$ are wide enough to include the point 0.4587, the zero of the derivative of the goal function (see the last column of Table 2).

The obtained results in Tables 1-2 for problems (40)-(41) are illustrated on Fig. 1-2 respectively.

As is shown on Fig. 1, the investment to the risky asset with the capital $x = 1.5$ falls to zero if $t \geq 3$ (see the solid line). Also, $\hat{a}_t^1(2)$ (point line) decreases as t grows from $t = 3$. It can by explained by a relatively small capital value $x = 1.5$ *or* $x = 2$ that leads to investing more capital into the risky-free asset in order to satisfy the constraint in (40). In the other cases with $x = 2.5, 3$, one can see that at these "large" capitals $x \geq 2.5$ the investments to the risky asset remains the same.

Fig. 2 is somewhat similar to the graphs depicted on Fig. 1; however, after $t = 1$ all the investments to the risky asset fall, which is explained by the investor's desire to partially rid the portfolio of the risky asset in order to increase optimal utility value in problem (33) at the last stage $T - 1 = 4$ of decision making.

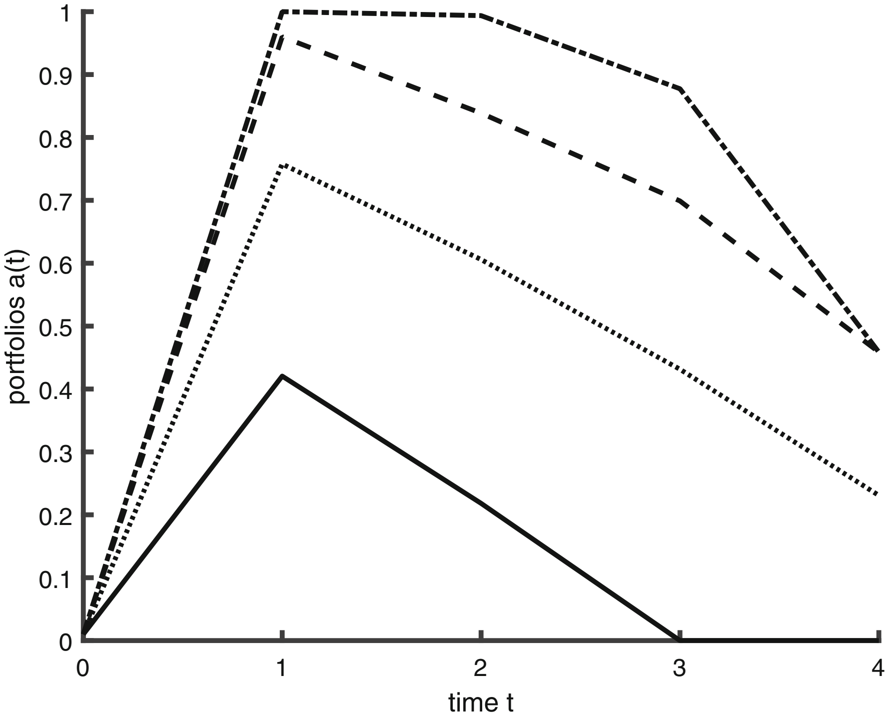

Fig. 2 The case of the second type of the discount factor (see section 8). Optimal shares $\tilde{a}_t^1(x)$ of capital x invested to the risky asset under the varying x: $\tilde{a}_t^1(1.5)$ – solid line, $\tilde{a}_t^1(2)$ – point line, $\tilde{a}_t^1(2.5)$ – dashed line, $\tilde{a}_t^1(3)$ – dash-dot line

10 Conclusions

This paper examines an issue of finding an optimal strategy in the controlled Markov chains with a finite horizon and state space $X \subseteq R$. The goal functional is the expected utility of the final state. It is supposed that the chains possess some monotonicity and concavity properties related to the transition function and the utility function of the decision maker. Introducing a stochastic discount factor based on the myopic optimal strategy allows for a relatively simple finding of an optimal strategy in the process modified by the stochastic discount. Moreover, the modified process turns out to be a super-martingale under any admissible strategy and converts into a martingale under the myopic strategy. Applications of the suggested discount approach to investment problems, including the call option, are presented. Also, in practice, this approach can be applied to inventory, investment, pension plans, and eco-finance network problems (see section 1 and references therein).

Further research may also include the investigation of more complicated option schemes, e.g., the so-called Asian or American options [8]. The subjects for studying may be the models of Markov decision processes in continuous time described in [2] (including two players, portfolio manager and Nature), and a Markov regime-switching jump-diffusion model with delay and an application to a problem of optimal consumption problem from a cash flow [13]. These problems may be modified by the discretization of time [9, 18] or by inventing an appropriate stochastic discount for the continuous time models. Another venue of research is to analyze a hypothesis that the myopic strategy is the first term of some expansion of an optimal strategy in the initial optimal control problem for non-discounted controlled Markov chains.

There are some limitations of the suggested approach, including: continuity of the transition function and its derivation $F'_x(x, a, z) > 0$, concavity and differentiability of the utility function $U(x)$, convexity of the sets $A_{t,x}$ of admissible decisions (see section 4). In the problems with the "hard" constraints (see sections 6-7), it is supposed that the risk aversion function is $R(x) \leq 2/x$.

Acknowledgements The author wishes to thank anonymous referees for many helpful comments. The research was supported by State program FFNR-2024-0003.

References

1. Arapostathis, V.S., Borkar, E., Fernandez-Gaucherand, E., Ghosh, M.K., Marcus, S.I.: Discrete-time controlled markov processes with average cost criterion: a survey. SIAM J. Control. Optim. **31**, 282–344 (1993)
2. Baltas, I.: Optimal investment in a general stochastic factor framework under model uncertainty. J. Dyn. Games **11**(1), 20–47 (2024). https://doi.org/10.3934/jdg.2023011
3. Dem'Yanov, V.F., Vasil'ev, L.V., Sasagawa, T.: Nondifferentiable Optimization. Springer-Verlag, New York (2012)
4. Felmer, H., Schied, A.: Stochastic Finance: An Introduction in Discrete Time. Walter de Gruyter GmbH, Berlin and Boston (2016)
5. Fujiwara-Greve, T.: Non-Cooperative Game Theory. Springer, Tokyo (2015). https://doi.org/10.1007/978-4-431-55645-9
6. Golubin, A.Y.: Optimal investment policy in a multi-stage problem with bankruptcy and stage-by-stage probability constraints. Optimization **70**(10), 2963–2977 (2022). https://doi.org/10.1080/02331934.2021.1892674
7. Howard, R.A.: Dynamic Programming and Markov Processes. The M.I.T Press, Cambridge (1960)
8. Hull, J.C.: Options, Futures, and Other Derivatives, 10th edn. Pearson, London (2017)
9. Kushner, H.J., Dupuis, P.: Controlled Markov Chains. In: Numerical Methods for Stochastic Control Problems in Continuous Time. Stochastic Modeling and Applied Probability. 24-th edn. Springer, New York (2001)
10. Pinar, M.C.: Static and dynamic VaR constrained portfolios with application to delegated portfolio management. Optimization **62**(11), 1419–1432 (2013)
11. Pratt, J.W.: Risk aversion in the small and in the large. Econometrica **32**, 122–136 (1964)
12. Rockafellar, R.T.: Convex Analysis. Princeton University Press, Princeton (2015)
13. Savku, E., Weber, G.W.: A stochastic maximum principle for a markov regime-switching jump-diffusion model with delay and an application to finance. J. Optim. Theory Appl. **179**(2), 696–721 (2018). https://doi.org/10.1007/s10957-017-1159-3
14. Still, G.: Lectures on Parametric Optimization: An Introduction. University of Twente, Twente (2018)
15. Temocin, B., Korn, R., Sevtap, K.: Constant proportion portfolio insurance in defined contribution pension plan management under discrete-time trading. Ann. Oper. Res. **260**(3), 1–30 (2018). https://doi.org/10.1007/s10479-017-2638-5
16. Weber, G.W., Ozlem, D., Kropat, E., Zeynep Alparslan, S.: Modeling, inference and optimization of regulatory networks based on time series data. Eur. J. Oper. Res. **211**(1), 1–14 (2011). https://doi.org/10.1016/j.ejor.2010.06.038
17. Wrobel, A.: On Markovian decision models with a finite skeleton. Oper. Res. Lett. **28**, 17–27 (1984). https://doi.org/10.1007/BF01919083

18. Yilmaz, F., Oz, H., Weber, G.W.: Weak second-order conditions of Runge-Kutta metod for stochastic optimal control problems. J. Optim. Theory Appl. **202**(11), 497–517 (2024)