Math Meth Oper Res (1999) 50:167–188

Mathematical Methods
of Operations Research

Risk sensitive control of finite state Markov chains in discrete time, with applications to portfolio management*

Tomasz Bielecki[1], Daniel Hernández-Hernández[2], Stanley R. Pliska[3]

[1] Department of Mathematics, The Northeastern Illinois University, 5500 N. St. Louis Ave., Chicago, IL 60625-4699, USA (e-mail: T-Bielecki@neiu.edu)
[2] Departamento de Matematicas, CINVESTAV-IPN, Apartado Postal 14-740, 07000 Mexico, D.F., Mexico (e-mail: dher@math.cinvestav.mx)
[3] Department of Finance, University of Illinois at Chicago, 601 S. Morgan St., Chicago, IL 60607-7124, USA (e-mail: srpliska@uic.edu)

Abstract. In this paper we extend standard dynamic programming results for the risk sensitive optimal control of discrete time Markov chains to a new class of models. The state space is only finite, but now the assumptions about the Markov transition matrix are much less restrictive. Our results are then applied to the financial problem of managing a portfolio of assets which are affected by Markovian microeconomic and macroeconomic factors and where the investor seeks to maximize the portfolio's risk adjusted growth rate.

Key words: Risk sensitive Markov decision processes, portfolio optimization, factor modeling

1 Introduction

In a recent series of papers (see [4], [5], and [6]), Bielecki and Pliska developed a new, control-theoretical approach for the optimal management of a portfolio of assets. The assets are modeled as lognormal processes, but there are also macroeconomic and financial factors which are Gaussian processes that directly affect the mean returns of the assets, as in Merton [13], for example. The innovative feature of their approach is the employment of methods of risk sensitive control theory, thereby leading to the infinite horizon objective of maximizing the portfolio's risk adjusted growth rate. This criterion is natural, for it is essentially a trade off between the long run expected growth rate and the asymptotic variance (i.e., the average squared volatility), analogous to the mean return and variance, respectively, in the single period Markowitz model.

Bielecki and Pliska [4] used continuous time risk sensitive control theory (see [3], [16], and [20]) to show that the optimal strategy is a simple function

* The research of the three authors was partially supported by NSF Grant DMS-9971307, CONACYT grant 27933E, and NSF Grant DMS-9971424, respectively.

of the factor levels. Moreover, with an assumption that the residuals of the assets are uncorrelated with the residuals of the factors, and even with constraints on the portfolio proportions, they showed that the optimal strategy can easily be computed by solving a parametric quadratic program. Explicit formulas can be obtained, as illustrated by an example in [6] where the only factor is a Vasicek-type interest rate and where there are two assets: cash and a stock index.

While their assumption about the uncorrelated residuals is reasonable for some financial applications, it is not for others. Unfortunately, when the residuals are correlated, and especially when there are also constraints on the portfolio proportions, the computational difficulties become formidable. The resulting Hamilton-Jacobi-Bellman equation cannot be solved directly, and so interest turns to the popular computational approach of making discrete time approximations of the underlying continuous time processes. In particular, interest turns to the idea of first building discrete time stochastic process models of the assets, factors, and trading strategies, then specifying the risk adjusted growth rate criterion in this discrete time context, and finally using discrete time risk sensitive control theory to develop a procedure for computing optimal strategies.

There is no difficulty with the first step. In Section 3 below we show that it is easy to formulate a discrete time risk sensitive portfolio management model that is in perfect analogy to the continuous time model in Bielecki and Pliska [4], [5]. In particular, the factor process is an exogenous Markov chain, and the set of possible factor values becomes the state space for the Markov control model. Each period the action taken is the allocation of wealth among the assets. The asset returns, which depend not only on the factor's state at the beginning of the period but also on its state at the end of the period (this additional dependence is important for many financial applications), combine with the chosen action to determine the portfolio's return. Portfolio wealth grows in a multiplicative fashion, but log wealth grows in a linear fashion, analogous to the cumulative reward (or cumulative cost) in a more conventional Markov control model.

The second step is also straightforward. Since log wealth is analogous to the cumulative reward, the discrete time risk adjusted growth rate is seen to be identical to the criterion studied in the risk sensitive Markov control literature. Hence the third step, the application of the discrete time risk sensitive control theory (see [8], [9], [10], and [11]) would also seem to be straightforward. Indeed it is. With suitable assumptions (especially about the transition matrix for the underlying Markov factor process), it can be shown that the optimal trading strategy can be characterized in terms of a dynamic programming equation which resembles that for conventional Markov control problems with the average reward (or average cost) criterion. Moreover, the maximum risk adjusted growth rate corresponds to the maximum average reward (or minimum average cost) as well as to the spectral radius of a certain matrix. Finally, by virtue of a contraction argument, a value iteration algorithm can be used to compute optimal solutions.

Hence it would seem that the theory of discrete time risk sensitive portfolio management falls neatly in place, but there is still one major problem: a critical assumption needed by the discrete time risk sensitive Markov control theory is much too severe when applied to the portfolio management situation. This assumption virtually requires every element of the factor process' Markov transition matrix to be strictly positive, and it results in the one-step value iteration operator being a contraction. This assumption is too severe for financial applications, because it is unreasonable to suppose that an economic factor can swing from one extreme value to another in a short period of time such as a day or even a month.

Fortunately, however, the discrete time risk sensitive control literature also offers a potential path to a solution of our problem. The literature assumes the state space of the underlying controlled Markov chain is countable, whereas for our financial applications (since we are seeking to compute results!) it suffices for the state space (of the factor process) to be finite. As will be demonstrated in the following section, by restricting the state space while simultaneously relaxing the assumption about the underlying Markov transition matrix to, roughly, only irreducibility, the same kinds of dynamic programming results found in the existing discrete time risk sensitive control theory literature still hold true. In particular, and consistent with well known results for classical Markov control problems with the average reward criterion and only finitely many states, while the one-stage value iteration operator is not necessarily a contraction, with an irreducibility assumption the N-stage operator is a contraction, where N is the number of states.

The control theoretic results in Section 2 stand alone, without any reference to financial applications. Hence this section is a contribution to the discrete time risk sensitive control theory literature such as [8], [9], [10], and [11]. In Section 3 we formulate our discrete time risk sensitive optimal portfolio problem and then reformulate it in terms of the Markov control model of Section 2. Finally, in Section 4 we apply the theory of Section 2 in order to obtain various fundamental results for our financial application.

After this paper was submitted for publication we became aware of very recent works by Balaji and Meyn [2] and Balaji, Borkar and Meyn [1]. The former one develops multiplicative ergodic theory for irreducible Markov chains. The latter one applies this theory to a risk sensitive control problem for a Markov chain on a denumerable state space. One of the underlying assumptions made there is that the one step cost function is norm-like. Overall, the approach of Balaji, Borkar and Meyn is different from ours.

2 Optimal risk sensitive control of finite state Markov chains

The control model. Let $(\mathbb{E}, A, P, c)$ be a Markov control model as in [12] satisfying the following. The finite set $\mathbb{E} = \{1, 2 \ldots, N\}$ is the state space, endowed with the discrete topology, while A is a Borel space, called the action or control space. For every $x \in \mathbb{E}$, there is a nonempty set $A(x) \subset A$, which represents the set of admissible actions when the system is in state x. The set of admissible pairs is $\mathbf{K} := \{(x, a) : x \in \mathbb{E}, a \in A(x)\}$, and is assumed to be a Borel subspace of $\mathbb{E} \times A$. The transition law $P_{x,y}(a)$ is a stochastic kernel on $\mathbb{E}$ given $\mathbf{K}$. Finally, $c : \mathbf{K} \to \mathbb{R}$ is a bounded, continuous function, not necessarily nonnegative, which represents the one stage cost.

Assumption A.1.

(i) For each $x, y \in \mathbb{E}$, the mapping $a \to P_{x,y}(a)$ with $a \in A(x)$ is continuous.
(ii) For each $x \in \mathbb{E}$, $A(x)$ is a compact subset of A.

Define $H_0 = \mathbb{E}$ and $H_t = \mathbf{K} \times H_{t-1}$ if $t = 1, 2, \ldots$. A control policy, or strategy, is a sequence $\pi = \{\pi_t\}$ of stochastic kernels on A given H_t that satisfy the constraint

$$\pi_t(A(x_t)|h_t) = 1 \quad \forall h_t \in H_t, t \geq 0.$$

The set of policies is denoted by $\mathscr{P}$. A policy $\pi \in \mathscr{P}$ is called a Markov policy if there exists a sequence of functions $\{\pi_t\}$, with $\pi_t : \mathbb{E} \to P(A)$, where $P(A)$ is the set of probability measures on A, such that $\pi_t(x)(A(x)) = 1$. We denote by $\mathscr{P}_M$ the set of Markov policies, and by $\mathscr{P}_{DM}$ the set of deterministic Markov policies, i.e., $\pi \in \mathscr{P}_{DM}$ if $\pi = \{\pi_t\} \in \mathscr{P}_M$ and $\pi_t(x)$ is a Dirac measure concentrated on some point of $A(x)$. We denote by $\mathbb{F}$ the set of functions $f : \mathbb{E} \to A$ such that $f(x) \in A(x)$ for all $x \in \mathbb{E}$. A policy $\pi \in \mathscr{P}_{DM}$ is stationary if there exists $f \in \mathbb{F}$ such that $\pi_t(f(x_t)|h_t) = 1$ for all $h_t \in H_t, t \geq 0$; this policy will also be denoted by $f \in \mathbb{F}$.

If the initial state $x \in \mathbb{E}$ and $\pi \in \mathscr{P}$ are given, there exists a unique probability measure P_x^π on $(\Omega, \mathscr{F})$, the canonical measurable space that consists of the sample space $\Omega := (\mathbb{E} \times A)^\infty$ and the corresponding product σ-algebra $\mathscr{F}$. Further, a stochastic process $\{(x_t, a_t), t = 0, 1, \ldots\}$ is defined in a canonical way, where x_t and a_t denote the state and action at time t, respectively. The expectation operator with respect to P_x^π is denoted by $\mathbb{E}_x^\pi$.

Assumption A.2.

(i) Under the action of any policy $\pi \in \mathscr{P}_{DM}$ the state space is irreducible, i.e., given $x, y \in \mathbb{E}$ and $\pi \in \mathscr{P}_{DM}$, there exists $m = m(x, y, \pi) < N$ such that $P_x^\pi[x_m = y] > 0$.
(ii) For all $x \in \mathbb{E}$ and $a \in A(x)$, $P_{x,x}(a) > 0$.

Risk sensitive optimality criterion. Given $x \in \mathbb{E}$, the risk sensitive average cost under policy $\pi \in \mathscr{P}$ is defined by

$$J(x, \pi) = \limsup_{T \to \infty} \frac{1}{T} \cdot \frac{1}{\gamma} \ln E_x^\pi \exp\left\{ \gamma \sum_{t=0}^{T-1} c(x_t, a_t) \right\},$$

where $\gamma > 0$ is a risk aversion parameter, and the corresponding value function is given by

$$J(x) := \inf_{\pi \in P} J(x, \pi).$$

The optimal control problem is to find a policy π^* such that $J(x) = J(x, \pi^*)$.

The following verification theorem was proved in [9].

Theorem 2.1. *Suppose that there exist a number λ and a function $W : \mathbb{E} \to \mathbb{R}$ such that*

$$e^{\lambda + W(x)} = \min_{a \in A(x)} \left\{ e^{\gamma c(x,a)} \sum_{y \in \mathbb{E}} e^{W(y)} P_{x,y}(a) \right\}. \tag{1}$$

Then $J(x) = \frac{1}{\gamma}\lambda$, *and the control* $f^* \in \mathbb{F}$ *with* $f^*(x)$ *achieving the minimum on the r.h.s. is optimal.*

In this section we shall prove that equation (1) has a unique (up to a constant added to W) solution and that the value iteration algorithm can be implemented.

Span contraction operator. Throughout we identify the set of functions from $\mathbb{E}$ to $\mathbb{R}$ with $\mathbb{R}^N$. Now, given $g \in \mathbb{R}^N$, we define the operator $T : \mathbb{R}^N \to \mathbb{R}^N$ by

$$Tg(x) := \inf_{a \in A(x)} \left\{ \gamma c(x,a) + \ln \sum_{y \in \mathbb{E}} e^{g(y)} P_{x,y}(a) \right\}. \tag{2}$$

Further, for each $k \geq 1$ we define $g_k \in \mathbb{R}^N$ as

$$g_k := Tg_{k-1}, \tag{3}$$

with $g_0 = g$.

Using the Markov property and the definition of g_k, it can be seen that for each $k \geq 1$

$$g_k(x) = \inf_{\pi \in \mathscr{P}} \ln E_x^{\pi} e^{\sum_{t=0}^{k-1} \gamma c(x_t, a_t) + g(x_k)}, \tag{4}$$

and, in fact, there exists $\pi^g = \{\pi_k^g\} \in \mathscr{P}_{DM}$ such that

$$g_k(x) = \ln E_x^{\pi^g} e^{\sum_{t=0}^{k-1} \gamma c(x_t, a_t) + g(x_k)}. \tag{5}$$

With $\pi^g \in P_{DM}$ as above, for each $t = 0, \ldots, k-1$ define the $N \times N$ matrix

$$Q_{x,y}^g(t) := e^{\gamma c(x, \pi_t^g(x))} P_{x,y}(\pi_t^g(x)).$$

Then

$$\begin{aligned} E_x^{\pi^g} e^{\sum_{t=0}^{k-1} \gamma c(x_t, a_t) + g(x_k)} &= Q^g(0) \cdots Q^g(k-1) e^g(x) \\ &= H_k^g e^g(x), \end{aligned} \tag{6}$$

where $H_k^g := Q^g(0) \cdots Q^g(k-1)$.

Remark 2.1. We observe that from Assumption (A.2) the matrix $H_k^g > 0$ for $k \geq N$.

The proof of the main results in this section requires some preliminary notation and results from positive matrices (see [19]). Given $g, h \in \mathbb{R}^n$ with strictly positive entries, define the projective pseudometric

$$d(g,h) := \ln\left[\max_x(g(x)/h(x)) \Big/ \min_x(g(x)/h(x))\right].$$

Observe that $sp(g-h) := \max_x(g(x)-h(x)) - \min_x(g(x)-h(x)) = d(e^g, e^h)$. Paralleling the theory for square matrices, we can define the Birkhoff's contraction coefficient τ_B for rectangular matrices in the following way: Given a $m \times n$ positive matrix $H > 0$, let

$$\tau_B(H) := \sup_{g,h>0} \frac{d(Hg, Hh)}{d(g,h)}. \tag{7}$$

The proof of the next proposition follows from adapting the arguments given in [19] for square matrices, and we omit it.

Proposition 2.1. *For a $m \times n$ positive matrix $H = \{H_{ij}\}$*

$$\tau_B(H) = \frac{1 - [\phi(H)]^{1/2}}{1 + [\phi(H)]^{1/2}},$$

where

$$\phi(H) := \min_{\substack{0 \le i,j \le m \\ 0 \le k,\ell \le n}} \frac{H_{ik}H_{j\ell}}{H_{jk}H_{i\ell}}.$$

The following result is fundamental to obtain the main results of this paper.

Theorem 2.2. *There exists a constant $0 < \tau < 1$ such that for any $g, h \in \mathbb{R}^N$,* $k = rN + j$, $r \ge 1$, and $0 \le j < N$ one has

$$sp(g_k - h_k) \le \tau^r sp(g-h),$$

with g_k, h_k defined by (3).

Proof. We consider first the case when $r = 1$. So, let $k = N + j$, with $0 \le j < N$. From (5) we know that there exist π^g, $\pi^h \in \mathscr{P}_{DM}$ such that

$$g_k(x) = \ln E_x^{\pi^g} e^{\sum_{t=0}^{k-1} \gamma c(x_t, a_t) + g(x_k)}$$

and

$$h_k(x) = \ln E_x^{\pi^h} e^{\sum_{t=0}^{k-1} \gamma c(x_t, x_t) + h(x_k)}.$$

Defining $\tilde{g}_j(x) := \ln(Q^g(N)\cdots Q^g(N+j-1)e^g(x))$, for $N > j \geq 1$ with $\tilde{g}_0 = g$, and similarly for $\tilde{h}_j(x)$, the following identities hold:

$$g_k(x) = \ln E_x^{\pi^g} e^{\sum_{t=0}^{N-1} \gamma c(x_t,a_t) + \tilde{g}_j(x_N)}$$

and

$$h_k(x) = \ln E_x^{\pi^h} e^{\sum_{t=0}^{N-1} \gamma c(x_t,x_t) + \tilde{h}_j(x_N)}.$$

Let $x^* \in \underset{x}{\operatorname{argmax}}\{g_k(x) - h_k(x)\}$ and $x_* \in \underset{x}{\operatorname{argmin}}\{g_k(x) - h_k(x)\}$. Then

$$g_k(x^*) - h_k(x^*) \leq \ln E_{x^*}^{\pi^h} e^{\sum_{t=0}^{N-1} \gamma c(x_t,a_t) + \tilde{g}_j(x_N)} - \ln E_{x^*}^{\pi^h} e^{\sum_{t=o}^{N-1} \gamma c(x_t,a_t) + \tilde{h}_j(x_N)}$$

and

$$g_k(x_*) - h_k(x_*) \geq \ln E_{x_*}^{\pi^g} e^{\sum_{t=0}^{N-1} \gamma c(x_t,a_t) + \tilde{g}_j(x_N)} - \ln E_{x_*}^{\pi^g} e^{\sum_{t=0}^{N-1} \gamma c(x_t,a_t) + \tilde{h}_j(x_N)}.$$

Hence

$$\begin{aligned} sp(g_k - h_k) &\leq \ln \left[\frac{E_{x^*}^{\pi^h} e^{\sum_{t=0}^{N-1} \gamma c(x_t,a_t) + \tilde{g}_j(x_N)} \cdot E_{x_*}^{\pi^g} e^{\sum_{t=0}^{N-1} \gamma c(x_t,a_t) + \tilde{h}_j(x_N)}}{E_{x^*}^{\pi^h} e^{\sum_{t=0}^{N-1} \gamma c(x_t,a_t) + \tilde{h}_j(x_N)} \cdot E_{x_*}^{\pi^g} e^{\sum_{t=0}^{N-1} \gamma c(x_t,a_t) + \tilde{g}_j(x_N)}} \right] \\ &\leq \max_{0 \leq n,m \leq N} \ln \left[\frac{H_N^h e^{\tilde{g}_j}(x_n) \cdot H_N^g e^{\tilde{h}_j}(x_m)}{H_N^h e^{\tilde{h}_j}(x_n) \cdot H_N^g e^{\tilde{g}_j}(x_m)} \right] \\ &\leq \max_{0 \leq n,m \leq 2N} \ln \left[\frac{(H_N^h/H_N^g) e^{\tilde{g}_j}(x_n) / (H_N^h/H_N^g) e^{\tilde{h}_j}(x_n)}{(H_N^h/H_N^g) e^{\tilde{g}_j}(x_m) / (H_N^h/H_N^g) e^{\tilde{h}_j}(x_m)} \right] \\ &= d((H_N^h/H_N^g)e^{\tilde{g}_j}, (H_N^h/H_N^g)e^{\tilde{h}_j}), \end{aligned} \tag{8}$$

where (H_N^h/H_N^g) is the $2N \times N$ matrix in which the rows of H_N^g follow the rows of H_N^h. Observe that $H_N^h, H_N^g > 0$ and so $(H_N^h/H_N^g) > 0$. Further, paralleling (8) it is easy to see that

$$sp(\tilde{g}_j - \tilde{h}_j) \leq sp(g - h). \tag{9}$$

Now we define

$$\phi^* := \min_{\pi_2, \pi_2 \in P_{DM}} \phi^{1/2}(H_N^{\pi_1}/H_N^{\pi_2}). \tag{10}$$

Then, from Proposition 2.1,

$$\begin{aligned}\tau_B(H_N^{\pi_1}/H_N^{\pi_2}) &= \frac{1-\phi^{1/2}(H_N^{\pi_1}/H_N^{\pi_2})}{1+\phi^{1/2}(H_N^{\pi_1}/H_N^{\pi^2})}\\ &\le \frac{1-\phi^*}{1+\phi^*}\\ &=: \tau.\end{aligned}$$

Therefore, from (8)–(10) we conclude that

$$sp(g_k - h_k) \le \tau sp(g-h). \tag{11}$$

Thus, for $r > 1$ the theorem follows from (11). □

From the above result it follows that the operator T has a span fixed point $W \in \mathbb{R}^N$; that is, there exist $W \in \mathbb{R}^N$ and a constant λ such that

$$\lambda + W(x) = TW(x) \quad \text{for all } x \in \mathbb{E}, \tag{12}$$

which is the dynamic programming equation (1).

Value iteration. Given arbitrary $g \in \mathbb{R}^N$ the value iteration functions g_k are defined recursively by (3). The relative value functions are given by $G_k(x) := g_k(x) - g_k(z)$, with $z \in \mathbb{E}$ being arbitrary but fixed. Also, $\lambda_k(x) := g_k(x) - g_{k-1}(x)$ with $x \in \mathbb{E}$ and $k \in \mathbb{N}$ represents the k-differential cost at state $x \in \mathbb{E}$ (see [12]).

Theorem 2.3. *Let λ and W be as in (12). Then, for each $x \in \mathbb{E}$, $\lambda_k(x) \to \lambda$ and $G_k(x) \to W(x) - W(z)$ when $k \to \infty$. Moreover,* $\sup_{x\in\mathbb{E}}\left|J(x,\pi_k^*) - \frac{\lambda}{\gamma}\right| \to 0$ *as $k \to \infty$, with $\pi_k^* \in \mathbb{F}$ achieving the minimum on the r.h.s. of (3).*

Proof. Let $k > N$ and $\pi_k^* \in \mathbb{F}$ be as in the statement of the theorem. Then

$$g_k(x) = \gamma c(x, \pi_k^*(x)) + \ln \sum_{y\in\mathbb{E}} e^{g_{k-1}(y)} P_{x,y}(\pi_k^*(x)),$$

which implies that

$$\begin{aligned}\lambda_k(x) &= g_k(x) - g_{k-1}(x)\\ &\ge \ln \sum_{y\in\mathbb{E}} e^{g_{k-1}(y)} P_{x,y}(\pi_k^*(x)) - \ln \sum_{y\in\mathbb{E}} e^{g_{k-2}(y)} P_{x,y}(\pi_k^*(x))\\ &\ge \inf_x \lambda_{k-1}(x).\end{aligned}$$

Thus $\inf_x \lambda_{k-1}(x) \le \inf_x \lambda_k(x)$, and similarly it can be seen that $\sup_x \lambda_k(x) \le \sup_x \lambda_{k-1}(x)$. On the other hand, using analogous arguments, it can be seen also that $k \to \max_x[g_k(x) - W_k(x)]$ is nonincreasing and $k \to \inf_x[g_k(x) - W_k(x)]$ is nondecreasing. Hence

$$\begin{aligned}\sup_x \lambda_k(x) &= \sup_x\{g_k(x) - W_k(x) - g_{k-1}(x) + W_{k-1}(x) + W_k(x) - W_{k-1}(x)\}\\ &\le \sup_x\{g_k(x) - W_k(x)\} - \inf_x\{g_{k-1}(x) - W_{k-1}(x)\} + \lambda\\ &\le \sup_x\{g_{k-1}(x) - W_{k-1}(x)\} - \inf_x\{g_{k-1}(x) - W_{k-1}(x)\} + \lambda\\ &= sp(g_{k-1}(x) - W_{k-1}(x)) + \lambda\\ &\le \tau^r sp(g - W) + \lambda, \end{aligned} \tag{13}$$

where r is such that $k - 1 = rN + j$, $0 \le j < N$.

Analogously it can be seen that

$$\inf_x \lambda_k(x) \ge -\tau^r sp(g - W) + \lambda. \tag{14}$$

Therefore, for all $x \in \mathbb{E}$,

$$-\tau^r sp(g - W) + \lambda \le \inf_x \lambda_k(x) \le \lambda_k(x) \le \sup_x \lambda_k(x) \le \tau^r sp(g - W) + \lambda,$$

and hence

$$\sup_x |\lambda_k(x) - \lambda| \le \tau^r sp(g - W),$$

which implies the first part of the theorem.

Now we shall prove the second part of the theorem. First, we observe that $g_k(x) - W_k(x) = g_k(x) - W(x) - k\lambda$, and so

$$g_k(x) - g_k(z) - W(x) + W(z) = g_k(x) - W_k(x) - g_k(z) + W_k(z). \tag{15}$$

Further, from Theorem 2.2 we have that

$$sp(g_k - W_k) \le \tau^r sp(g - W), \tag{16}$$

which together with the monotone property of $k \to \max_x[g_k(x) - W_k(x)]$ and $k \to \min_x[g_k(x) - W_k(x)]$ we get $\lim_{k\to\infty} \max_x[g_k(x) - W_k(x)] = \lim_{k\to\infty} \cdot \min_x[g_k(x) - W_k(x)]$. It follows that

$$\left| g_k(x) - W_k(x) - \lim_{k\to\infty}\left[\max_x(g_k(x) - W_k(x))\right]\right| \le sp(g_k - W_k) \le \tau^r sp(g - W).$$

Now, from (15),

$$\sup_x |g_k(x) - g_k(z) - W(x) + W(z)| \leq 2\tau^r sp(g - W),$$

which proves the second part of the theorem.

Finally, (13) and (14) imply that $\inf_x \lambda_k(x) \leq \lambda \leq \sup_x \lambda_k(x)$ for all $k \geq 2$, and so

$$\frac{1}{\gamma} \inf_x \lambda_k(x) \leq \frac{1}{\gamma} \lambda \leq J(x, \pi_k^*), \tag{17}$$

where π_k^* is as in the statement of the theorem. Moreover, for all $x \in \mathbb{E}$,

$$g_k(x) = \gamma c(x, \pi_k^*(x)) + \ln \sum_{y \in \mathbb{E}} e^{g_{k-1}(y)} P_{x,y}(\pi_k^*(x)).$$

Hence

$$\begin{aligned} E_x^{\pi_k^*} e^{\gamma \sum_{t=0}^{T-1} c(x_t, a_t)} &= E_x^{\pi_k^*} \prod_{t=0}^{T-1} \frac{e^{g_k(x_t)}}{\sum_{y \in \mathbb{E}} e^{g_{k-1}(y)} P_{x_t,y}(\pi_k^*(x_t))} \\ &\leq e^{\sup_x (g_k(x) - g_{k-1}(x)) T} E_x^{\pi_k^*} \prod_{t=0}^{T-1} \frac{e^{g_{k-1}(x_t)}}{\sum_{y \in \mathbb{E}} e^{g_{k-1}(y)} P_{x_t,y}(\pi_k^*(x_t))}, \end{aligned}$$

and using the Markov property, it follows that the expectation on the r.h.s. is bounded (see [9], eq. (2.21)). Then from the above we obtain

$$J(x, \pi_k^*) \leq \frac{1}{\gamma} \sup_x \lambda_k(x),$$

which together with (17) give

$$\frac{1}{\gamma} \inf_x \lambda_k(x) \leq \frac{1}{\gamma} \lambda \leq J(x, \pi_k^*) \leq \frac{1}{\gamma} \sup_x \lambda_k(x).$$

Finally, using (13) and (14), we obtain

$$\begin{aligned} -\frac{1}{\gamma} \tau^r sp(g - W) &\leq \frac{1}{\gamma} \inf_x \lambda_k(x) - \frac{1}{\gamma} \lambda \leq J(x, \pi_k^*) - \frac{1}{\gamma} \lambda \\ &\leq \frac{1}{\gamma} \sup_x \lambda_k(x) - \frac{1}{\gamma} \lambda \leq \frac{1}{\gamma} \tau^r sp(g - W). \end{aligned}$$

This concludes the proof of the theorem. □

In the following sections we shall apply the results obtained above in order to characterize and compute optimal investment strategies for a problem of risk sensitive portfolio management.

3 Discrete time risk sensitive portfolio management: Model formulation

To formulate the model, let the factor process X be a discrete time, stationary Markov chain having transition matrix $Q = (Q_{x,y})$ and finite state space $\mathbb{E}$. For example, suppose five macroeconomic variables are observed and each classified into four categories, thereby giving a total of $4^5 = 1024$ states. There is a bank account with constant interest rate r, so a deposit of one dollar becomes e^{rt} dollars after t periods (we find it more convenient to model the bank account as e^{rt} than as $(1+r)^t$, even though the latter is often used for discrete time models). There are m risky assets with the m-dimensional random variable Z representing the vector of one-period price relatives (i.e., the multiplicative factors by which the prices change; see Example 1 in Section 5). There is a conditional probability distribution of the form $\nu(x, y, dz)$ such that the price relative vector Z_{t+1} for the period between times t and $t+1$ has this distribution whenever $X_t = x$ and $X_{t+1} = y$. Note that having asset returns depend on factor values at the end of a period, and not just the beginning of the period, is important for realistic modeling, because asset returns are often correlated with factor value changes.

To model the trading strategies, let A denote a compact subset of $\mathbb{R}^m$ whose elements a represent admissible vectors of proportions for the risky assets. For example, the ith component of a is the proportion of wealth invested in the ith risky asset for the period, and $1 - a_1 - \cdots - a_m$ is the proportion of wealth invested for the period in the bank account. Using the notation of section 2 we assume that $A(x) = A$ for every x in $\mathbb{E}$. We now define the set of admissible pairs $\mathbf{K}$, the set of histories H_t, the sets of policies (i.e., trading strategies), and so forth exactly the same as in Section 2. In particular, we denote by $\mathbb{F}$ the set of functions $f : \mathbb{E} \to A$, and a trading strategy $\pi \in \mathscr{P}_{DM}$ is said to be stationary if there exists $f \in \mathbb{F}$ such that $\pi_t(f(x_t)|h_t) = 1$ for all $h_t \in H_t, t \geq 0$; this trading strategy will also be denoted by $f \in \mathbb{F}$.

Note that π_t represents the vector of proportions that will be in place between times t and $t+1$. Thus with V_t representing the time-t value of the portfolio under a particular trading strategy, it follows that

$$V_{t+1} = V_t[e^r + \pi_t \circ (Z_{t+1} - e^r\mathbf{1})],$$

where '$\circ$' here is meant to represent an inner product of vectors, and $\mathbf{1}$ represents the vector of 1's.

If π denotes an admissible trading strategy, then the risk sensitive measure of performance (i.e., the risk adjusted growth rate) is

$$J(x, \pi) = \liminf_{T \to \infty} \left(-\frac{2}{\theta}\right) \frac{1}{T} \ln E_x^\pi \exp\left(-\frac{\theta}{2} \ln V_T\right), \tag{18}$$

where E_x^π denotes conditional expectation given policy π and $x_0 = x$. The aim, of course, is to choose an admissible, Markov trading strategy π that maximizes this expression, which has exactly the same form as in the continuous-time formulation of Bielecki and Pliska [4]. The parameter θ here is a nonnegative scalar capturing the investor's attitudes about risk aversion. As explained in Bielecki, Pliska, and Sherris [6], the bigger the value, the more risk averse is the investor, and the "risk null" case $\theta = 0$ is the well-known

Kelly criterion case where the objective is simply to maximize the portfolio's long run growth rate.

We now seek to reformulate our optimal portfolio model in the terms of the preceding section. To do this we first introduce the expected value

$$\begin{aligned}\mu^{\theta}(x,y,a) :&= E_{x,y}\exp\left(-\frac{\theta}{2}\ln[e^{r}+a\circ(Z-e^{r}\mathbf{1})]\right)\\ &= \int\exp\left(-\frac{\theta}{2}\ln[e^{r}+a\circ(z-e^{r}\mathbf{1})]\right)\nu(x,y,dz).\end{aligned} \tag{19}$$

In order for this to make good sense, it is necessary to make the following:

Assumption A.3.

(i) The action set A is such that, for each $a\in A$, $e^{r}+a\circ(Z-e^{r}\mathbf{1})>0$ almost surely.
(ii) For each $x,y\in\mathbb{E}$ and each $a\in A$, the conditional expectation $\mu^{\theta}(x,y,a)$ exists and is finite.

For assumption A.3(i) it may be necessary to impose constraints on the portfolio proportions. For example, if a asset can become worthless, then you cannot invest all your money in this asset. And there may need to be limits on short selling.

We also define the "transition probability"

$$P^{\theta}_{x,y}(a)=\frac{Q_{x,y}\mu^{\theta}(x,y,a)}{\sum_{s}Q_{x,s}\mu^{\theta}(x,s,a)},$$

and the "one period cost"

$$c^{\theta}(x,a)=\frac{2}{\theta}\ln\left(\sum_{s}Q_{x,s}\mu^{\theta}(x,s,a)\right).$$

Then, by some calculations involving iterated conditional expectations, it is straightforward to show that our measure of performance $J(x,\pi)$ can be expressed as

$$J(x,\pi)=\liminf_{T\to\infty}\left(-\frac{2}{\theta}\right)\frac{1}{T}\ln E^{\pi,\theta}_{x}\exp\left(\frac{\theta}{2}\sum_{t=0}^{T-1}c^{\theta}(X_t,\pi_t)\right), \tag{20}$$

where $E^{\pi,\theta}_{x}$ is the conditional expectation with respect to the probability measure generated by the transition kernels $P^{\theta}_{y,y'}(a)$ and by the strategy π on the canonical (i.e. trajectory) space, given that the trajectory of the Markov chain X_t originates from x (that is $X_0=x$). Substituting $\gamma=\theta/2$, this becomes

$$J(x,\pi)=-\limsup_{T\to\infty}\frac{1}{T}\cdot\frac{1}{\gamma}\ln E^{\pi,\gamma}_{x}\exp\left(\gamma\sum_{t=0}^{T-1}c^{\gamma}(X_t,\pi_t)\right), \tag{21}$$

where the change of notation to $E_x^{\pi,\gamma}$ and $c^\gamma(X_t, \pi_t)$ is done to keep things nice and tidy.

We thus have arrived at the discrete time, risk sensitive, Markov control model that was presented in the preceding section, except that the measure of performance there is the negative of the measure of performance here. Thus minimizing the risk sensitive cost measure of Section 2 is the same as maximizing the risk adjusted growth rate for this section's portfolio management model. Note also that the cost function $c^\gamma(x, a)$ here explicitly depends on the risk aversion parameter γ, whereas it does not in Section 2; this difference is not important for what we intend to do.

In the following section we shall apply the theoretical results of Section 2 in order to draw some fundamental conclusions about risk sensitive portfolio management in discrete time.

4 Discrete time risk sensitive portfolio management: Main results

With Assumption A.3 holding throughout this section and with the risk adjusted growth rate $J(x, \pi)$ for any trading strategy π as in (18), the optimal portfolio problem is to find a trading strategy π^* such that $J(x) = J(x, \pi^*)$, where $J(x)$ is the value function given by

$$J(x) := \sup_{\pi \in \mathscr{P}} J(x, \pi).$$

We intend to solve this problem using the results of Section 2. In order to be sure that Assumptions A.1 and A.2 hold, we shall make the following:

Assumption A.4.

(i) The set A of admissible portfolio proportion vectors is compact.
(ii) For each $x, y \in \mathbb{E}$ and each $a \in A$, the conditional expectation $\mu(x, y, a)$ is strictly positive.
(iii) The Markov transition matrix Q for the factor process is irreducible.
(iv) For each $x \in \mathbb{E}$, $Q_{x,x} > 0$.

Clearly $a \to \mu^\theta(x, y, a)$ is continuous, so $a \to P^\theta_{x,y}(a)$ and $a \to c^\theta(x, a)$ are too. In particular, with A.4(i), it follows that Assumption A.1 holds. Assumption A.4(ii) is meant to rule out pathological cases, so we can be assured that $Q_{x,y} > 0$ if and only if $P^\theta_{x,y}(a) > 0$ for all $a \in A$. Thus Assumption A.2(i) holds. Similarly, Assumption A.4(iv), which is very reasonable for financial economic applications, together with A.4(ii) imply Assumption A.2(ii).

Making some simple substitutions (and keeping in mind that the value function in Section 2 is the negative of the value function here), the Verification Theorem 2.1 can be restated as follows:

Theorem 4.1. *Suppose that there exist a number λ and a function $W : \mathbb{E} \to \mathbb{R}$ such that*

$$e^{W(x)+\lambda} = \min_{a \in A(x)} \left\{ \sum_{y \in \mathbb{E}} Q_{x,y} \mu^{\theta}(x, y, a) e^{W(y)} \right\}. \tag{22}$$

Then $J(x) = -\frac{2}{\theta}\lambda$, *and the stationary trading strategy* $f^* \in \mathbb{F}$, *with* $f^*(x)$ *achieving the minimum on the r.h.s., is optimal.*

Of course we know by Theorem 2.2 in Section 2 that the solution (λ, W) of this dynamic programming equation exists and is unique up to a constant added to W. The theory of non-negative matrices can be exploited to provide an alternative and direct proof of this theorem, at least for the case where all admissible trading strategies are deterministic and Markov. Since this will also provide a better understanding of this theorem, we now digress to make these arguments. We begin by defining for each $f \in \mathbb{F}$ the square matrix

$$B^{\theta}(f) := (Q_{x,y} \mu^{\theta}(x, y, f(x))).$$

This matrix enables us to recast Theorem 4.1 as follows:

Proposition 4.1. *Suppose that there exist a strictly positive number* ρ, *a strictly positive column vector* v, *and some* $f^* \in \mathbb{F}$ *such that*

$$\rho v = B^{\theta}(f^*) v \leq B^{\theta}(f) v, \quad \forall f \in \mathbb{F}. \tag{23}$$

Then $J(x) = -\frac{2}{\theta} \ln \rho$, *and the stationary trading strategy* f^* *is optimal.*

In order to prove this proposition for the special case where "optimality" is with respect to just the deterministic Markov trading strategies, use will be made of the following:

Lemma 4.1. *Given any deterministic Markov trading strategy* $\pi = (f_0, f_1, \ldots) \in \mathscr{P}_{DM}$, *and assuming the initial capital* $V_0 = 1$, *then* $E_x^{\pi} \exp\left(-\frac{\theta}{2} \ln V_T\right)$ *equals the sum of the elements in row* x *of the matrix* $B^{\theta}(f_0) B^{\theta}(f_1) \cdots B^{\theta}(f_{T-1})$.

Proof. It suffices to show by induction that for an arbitrary deterministic Markov trading strategy $\pi = (f_0, f_1, \ldots) \in \mathscr{P}_{DM}$

$$E^{\pi}\left[\exp\left(-\frac{\theta}{2} \ln V_T\right) \middle| X_0 = i, X_T = j\right] P(X_0 = i, X_T = j) = B_{ij}^{\theta, T}$$

for all $i, j \in \mathbb{E}$ and all times T, where $B_{ij}^{\theta, T}$ denotes the element in row i and column j of the matrix $B^{\theta}(f_0) B^{\theta}(f_1) \cdots B^{\theta}(f_{T-1})$. The case $T = 1$ is left to the reader. For the induction step, suppose this is true for $T - 1$. Then

$$
\begin{aligned}
&E^{\pi}\left[\exp\left(-\frac{\theta}{2}\ln V_T\right)\,\middle|\, X_0 = i, X_T = j\right] \\
&\quad = E^{\pi}\left[\exp\left(-\frac{\theta}{2}\ln V_{T-1}\right)\right. \\
&\qquad \left.\times \exp\left(-\frac{\theta}{2}\ln[e^r + f_{T-1}(X_{T-1})(Z_{t+1} - e^r 1)]\right)\,\middle|\, X_0 = i, X_T = j\right] \\
&\quad = E^{\pi}\left[E^{\pi}\left[\exp\left(-\frac{\theta}{2}\ln V_{T-1}\right)\right.\right. \\
&\qquad \times \exp\left(-\frac{\theta}{2}\ln[e^r + f_{T-1}(X_{T-1})(Z_{t+1} - e^r 1)]\right) \\
&\qquad \left.\left.\times \middle|\, X_0 = i, X_{T-1}, X_T = j\right]\,\middle|\, X_0 = i, X_T = j\right] \\
&\quad = E^{\pi}\left[E^{\pi}\left[\exp\left(-\frac{\theta}{2}\ln V_{T-1}\right)\,\middle|\, X_0 = i, X_{T-1}\right]\right. \\
&\qquad \left.\times \mu^{\theta}(X_{T-1}, j, f_{T-1}(X_{T-1}))\,|\, X_0 = i, X_T = j\right] \\
&\quad = \sum_k E^{\pi}\left[\exp\left(-\frac{\theta}{2}\ln V_{T-1}\right)\,\middle|\, X_0 = i, X_{T-1} = k\right] \\
&\qquad \times \mu^{\theta}(k, j, f_{T-1}(k)) P(X_{T-1} = k \,|\, X_0 = i, X_T = j).
\end{aligned}
$$

But

$$
\begin{aligned}
\mathrm{Prob}(X_{T-1} = k \,|\, X_0 = i, X_T = j) &= \frac{\mathrm{Prob}(X_0 = i, X_{T-1} = k, X_T = j)}{\mathrm{Prob}(X_0 = i, X_T = j)} \\
&= \frac{\mathrm{Prob}(X_0 = i, X_{T-1} = k) Q_{kj}}{\mathrm{Prob}(X_0 = i, X_T = j)}
\end{aligned}
$$

and by the induction assumption

$$
E^{\pi}\left[\exp\left(-\frac{\theta}{2}\ln V_{T-1}\right)\,\middle|\, X_0 = i, X_{T-1} = k\right] \mathrm{Prob}(X_0 = i, X_{T-1} = k) = B_{ik}^{\theta,T-1},
$$

so substituting gives

$$E^{\pi}\left[\exp\left(-\frac{\theta}{2}\ln V_T\right) \,\middle|\, X_0 = i, X_T = j\right] = \frac{\sum_k B_{ik}^{\theta, T-1}\mu^{\theta}(k, j, f_{T-1}(k))Q_{kj}}{\text{Prob}(X_0 = i, X_T = j)}$$

$$= \frac{\sum_k B_{ik}^{\theta, T-1} B_{kj}^{\theta}(f_{T-1})}{\text{Prob}(X_0 = i, X_T = j)},$$

thereby completing this proof. □

We shall also need some theory of positive matrices. Recall that since each $B^{\theta}(f)$ is a nonnegative, irreducible, aperiodic matrix, we know (see [19]) that it has a positive spectral radius, which we shall denote ρ_f, with a corresponding eigenvector that is strictly positive. Moreover, if v is any other eigenvalue, then $|v| < \rho_f$, and the corresponding eigenvector is not strictly positive. Note, by the way, that the number ρ in equation (23) is precisely ρ_{f^*}, the spectral radius of $B(f^*)$.

Proof of Proposition 4.1. Let ρ, v, and f^* be as in the Proposition, and let $\pi = (f_0, f_1, \dots) \in \mathscr{P}_{DM}$ be an arbitrary deterministic Markov trading strategy. Thus (23) implies

$$\rho v = B^{\theta}(f^*)v \leq B^{\theta}(f_n)v, \quad n = 0, 1, \dots \tag{24}$$

In particular, taking $n = 1$ this gives

$$\rho v \leq B^{\theta}(f_1)v;$$

multiplying both sides of this on the left by $B^{\theta}(f_0)$ and then applying (24) with $n = 0$ gives

$$(\rho)^2 v \leq B^{\theta}(f_0)B^{\theta}(f_1)v.$$

It thus becomes apparent that an easy induction argument can be used to show that

$$(\rho)^n v \leq B^{\theta}(f_0)B^{\theta}(f_1)\cdots B^{\theta}(f_{n-1})v, \quad \forall n. \tag{25}$$

Now the eigenvector v is strictly positive, so without loss of generality we can assume $v \leq 1$. Hence (25) implies

$$(\rho)^n v \leq B^{\theta}(f_0)B^{\theta}(f_1)\cdots B^{\theta}(f_{n-1})\mathbf{1}, \quad \forall n.$$

By Lemma 4.1, therefore,

$$(\rho)^T v(x) \leq E_x^{\pi} \exp\left(-\frac{\theta}{2}\ln V_T\right),$$

so

$$T \ln \rho + \ln v(x) \leq \ln E_x^{\pi} \exp\left(-\frac{\theta}{2}\ln V_T\right).$$

Hence by the definition (18) of the risk adjusted growth rate, we conclude that

$$-\frac{2}{\theta}\ln \rho \geq J(x,\pi), \quad \forall x \in \mathbb{E}. \tag{26}$$

Since π was chosen arbitrarily, this inequality holds for all deterministic Markov trading strategies.

It remains to show that equality holds in (26) when $\pi = f^*$. To do this we define $C := \rho^{-1}B^\theta(f^*)$, a matrix with spectral radius one. By Theorem 1.2 in Seneta [19] we know that C^T, the T-th power of C, converges to some strictly positive matrix as $T \to \infty$. So $\ln(C^T 1(x))$ also converges for all $x \in \mathbb{E}$, in which case $T^{-1}\ln(C^T 1(x))$ converges to zero. Since by Lemma 4.1

$$\begin{aligned}\ln E_x^{f^*}\exp\left(-\frac{\theta}{2}\ln V_T\right) &= \ln[B^\theta(f^*)]^T\mathbf{1}(x)\\ &= \ln[\rho C]^T\mathbf{1}(x) = T\ln\rho + \ln C^T\mathbf{1}(x),\end{aligned}$$

it follows that

$$\begin{aligned}J(x,f^*) &= \liminf_{T\to\infty}\left(-\frac{2}{\theta}\right)\frac{1}{T}[T\ln\rho + \ln C^T\mathbf{1}(x)]\\ &= -\frac{2}{\theta}\ln\rho + \liminf_{T\to\infty}\left(-\frac{2}{\theta}\right)\frac{1}{T}\ln C^T\mathbf{1}(x) = -\frac{2}{\theta}\ln\rho.\end{aligned}$$

This proof is done. □

With a similar kind of argument we obtain the following interesting characterization of the "optimal" spectral radius ρ.

Proposition 4.2. *The spectral radius ρ of Proposition 4.1 satisfies*

$$\rho = \inf_{f\in\mathbb{F}} \rho_f$$

Proof. Let $f \in \mathbb{F}$ be arbitrary, and let u denote the eigenvalue corresponding to spectral radius ρ_f and matrix $B^\theta(f)$. With ρ, v, and f^* as in Proposition 4.1, the fact that both u and v are strictly positive mean that we can assume without loss of generality that $v \leq u$. Hence by (23) we have

$$\rho v = B^\theta(f^*)v \leq B^\theta(f)v \leq B^\theta(f)u = \rho_f u.$$

It follows by a simple induction argument that we have, in fact,

$$\rho^n v \leq \rho_f^n u, \quad \forall n.$$

Hence $\rho_f < \rho$ gives rise to a contradiction, so this proof is completed. □

This ends our discussion of the optimality criteria Theorem 4.1 and Proposition 4.1. We now turn to the span contraction operator, which we denote by T^θ to emphasize its dependence on θ, and value iteration. According to definition (2), the operator T^θ for our portfolio management model given by (22) can be written as

$$T^\theta g(x) = \inf_{a \in A} \ln \sum_{y \in \mathbb{E}} Q_{x,y} \mu^\theta(x, y, a) e^{g(y)}.$$

Equivalently, we can write this in vector form as

$$T^\theta g = \inf_{f \in \mathbb{F}} \ln B^\theta(f) e^g.$$

Then as explained in Section 2, given arbitrary $g_0 : \mathbb{E} \to \mathbb{R}$ the value iteration functions are defined recursively by

$$g_k = T^\theta g_{k-1} = \inf_{f \in \mathbb{F}} \ln B^\theta(f) e^{g_{k-1}}.$$

The sequence defined by $\lambda_k(x) := g_k(x) - g_{k-1}(x)$ converges to $\lambda = \ln \rho$, part of the dynamic programming solution, as $k \to \infty$. With $z \in \mathbb{E}$ being arbitrary and fixed, the relative value functions given by $G_k(x) := g_k(x) - g_k(z)$ converge as $k \to \infty$ to a function W which is also part of the dynamic programming solution. This algorithm will be illustrated in the following section.

It is important to remark that as a consequence of the successive approximations and span contraction results, we now know that there exists a solution to the dynamic programming equation. Moreover, this solution is unique, up to an additive constant for W in the case of equation (22) or up to multiplication of v by a scalar in the case of relation (23).

We also remark that a policy improvement algorithm can be used to compute a dynamic programming equation solution, provided that the set A of admissible proportion vectors is finite. This procedure for a general risk sensitive Markov chain control problem is fully explained in Di Masi and Stettner [8].

We conclude this section by observing that the results presented here can be easily generalized to the case when the interest rate r is a function of the underlying Markov factor process. We illustrate this observation in the example 2 below, where the short rate r is a factor process itself.

5 Numerical examples

Example 1. In this toy example we suppose $\mathbb{E} = (1, 2)$, there is one risky asset, the interest rate is $r = 0$, and the risk aversion parameter is $\theta = 2$. The Markov transition matrix for the factor process is

$$Q = \begin{pmatrix} .8 & .2 \\ .3 & .7 \end{pmatrix}.$$

The conditional probability distribution for the risky asset's price relative is:

$$\nu(1, 1, 1.1) = 1, \quad \nu(1, 2, .7) = 1$$

$$\nu(2, 1, .6) = 1, \quad \nu(2, 2, 1.2) = 1.$$

Thus if the initial price is S, then after one period the price will be either $1.1S$, $0.7S$, $0.6S$, or $1.2S$, and these are perfectly correlated with the four possible transitions of the factor process. Note that the conditional expected one period returns are the same for both states (i.e., $E_1[Z] = E_2[Z] = 1.02$), but state 2 is more volatile than state 1.

The set A of admissible risky asset proportions must be specified with some care. In view of Assumption A.3(i) we must have

$$1 + a(Z - 1) > 0$$

for all $a \in A$ and all four possible values of Z. It follows with some simple algebra that we must have $-5 < a < 2.5$ for all $a \in A$. Since A must also be compact, we shall therefore assume that $A = [-1, 2]$.

Given our simplifying assumptions, the conditional expectations $\mu^\theta(x, y, a)$ have simple, explicit expressions:

$$\mu^\theta(1,1,a) = [1 + .1a]^{-1}, \quad \mu^\theta(1,2,a) = [1 - .3a]^{-1}$$

$$\mu^\theta(2,1,1) = [1 - .4a]^{-1}, \quad \mu^\theta(2,2,1) = [1 + .2a]^{-1}.$$

Consequently, the matrix $B^\theta(f)$ has the form:

$$B^\theta(f) = \begin{pmatrix} .8[1 + .1f(1)]^{-1} & .2[1 - .3f(1)]^{-1} \\ .3[1 - .4f(2)]^{-1} & .7[1 + .2f(2)]^{-1} \end{pmatrix}.$$

We implemented the value iteration algorithm starting with $g_0 = 1$ and using Solver on an Excel spreadsheet. After seven iterations the optimal risky asset proportions had converged to 0.3416 and 0.1301 for factor states 1 and 2, respectively. After eight iterations the value of λ had converged to -0.00264, which corresponds to a risk adjusted growth rate of $J(x) = -\frac{2}{\theta}\lambda = 0.00264$. We took fixed state $z = 1$ and thus $W(0) = 0$. After ten iterations the value of $W(2)$ had converged to 0.00428.

Example 2. This example is motivated by the continuous time "Vasicek" example in the asset allocation paper [6] by Bielecki, Pliska and Sherris. There is a single factor representing the short interest rate, and it takes values in one of three possible states: "low" (namely, 4 per cent per annum), "medium" (6 per cent), and "high" (8 per cent). The Markov transition matrix for the factor process is

$$Q = \begin{pmatrix} .96 & .04 & .0 \\ .02 & .96 & .02 \\ 0 & .04 & .96 \end{pmatrix}.$$

We are thinking of time periods being months, so the expected sojourn time in each state is 25 months. Note that this matrix is irreducible but not strictly positive; just seven kinds of one-step transitions can occur with positive probability: low-low, low-medium, medium-low, etc.

There is a single risky security whose one-period return is a lognormal random variable of the form $\exp^{m-\sigma^2/2+\sigma M}$, where M is a standard normal random variable. The volatility σ always equals 0.06; this is on a monthly basis and thus equals approximately $20.78(=6\sqrt{12})$ per cent on an annual basis). The mean return parameter m varies with the type of transition. It equals 13%, 12%, and 11% (on an annual basis) if the transition is low-low, medium-medium, or high-high, respectively. This reflects the idea that if the interest rate remains unchanged, then the asset return will be fairly routine, although there will be a small advantage for low interest rates. An upward shift in the interest rate is bearish, so m equals 3% and 2% (on an annual basis) if the transition is low-medium and medium-high, respectively. Finally, a downward shift in the interest rate is bullish, so m equals 22% and 21% (again, on an annual basis) if the transition is medium-low and high-medium, respectively.

As with Example 1, we did all our computations taking $\theta = 2$, only this time we did our computations using Maple rather than a spreadsheet. The first step was to specify the seven μ^θ functions, one for each type of transition. We wanted these to be explicit functions of the decision parameter a, the proportion of wealth in the risky asset, in order to facilitate the minimization operation that is part of the dynamic programming equation and the operator T^θ. However, with each μ^θ function being the expectation of a function of a lognormal random variable, Maple was not able to produce the corresponding indefinite integral, and thus it could not produce the μ^θ functions in the form we wanted. Consequently, we took a discrete approximation of the underlying standard normal random variable M, and so Maple was able to produce each of the seven μ^θ's as an explicit function of the decision parameter a, as desired.

At this point it was straightforward to implement with Maple the successive approximations algorithm; we did so starting with $g = 0$ and conducting 27 iterations. Convergence was judged by looking at the three values of $g_k - g_{k-1}$, all three of which should have been converging to $\lambda = \ln \rho$. For instance, after 20 iterations these three values were $-0.00674, -0.00671$, and -0.00741, whereas after 27 iterations these values were $-0.00674, -0.00679$, and -0.00725. Convergence is somewhat slow, perhaps due to the fact that Q is not strictly positive. After 27 iterations the optimal proportions were computed to be 0.991009, 0.691934, and 0.392575 for states low, medium, and high, respectively.

We initially imposed the constraint $0 \leq a \leq 1$, but it soon became apparent from our successive approximation calculations that the optimal value of the risky asset proportion a would be in the interior of this interval for all three of the states. We then realized that we could set up a system of six equations in six unknowns that could be solved for an optimal solution of the dynamic programming equation. We fixed $W(m) = 0$, so the six variables were $W(l), W(h), \lambda, a_l, a_m$, and a_h (the last three are the optimal proportions in states low, medium, and high, respectively). There were three equations corresponding to the dynamic programming equation, one for each state in the form

$$W(x) + \lambda = \ln \sum_{y \in \mathbb{E}} Q_{x,y} \mu^\theta(x, y, a_x) e^{W(y)}.$$

The other three equations are the three necessary conditions, one for each state. Each such equation, which is easily provided by Maple, is simply the derivative with respect to a_x of the right hand side of the preceding equation, set equal to zero.

Maple solved this system within a few seconds, obtaining the solution $W(h) = -0.008228396$, $W(l) = 0.000131498$, $a_h = 0.392966$, $a_m = 0.692224$, $a_l = 0.991776$, and $\lambda = -0.00689525$. These results are consistent with the successive approximation calculations. The optimal objective value is $J(x) = 0.00689525$.

6 Concluding remarks

In the introduction we have already emphasized the importance of discrete time modeling as a computationally feasible approximation to a continuous time problem. Discrete time modeling provides other benefits as well. It allows for modeling of time delays (time lags) within a finite dimensional framework. This is an important feature as the econometric literature indicates that both linear and non-linear models involving time lags provide more accurate description of dynamic economic systems (see e.g. Pesaran and Timmermann [17]). Modeling time delays within the framework presented in this paper poses no difficulty: it can be easily achieved by (finitely) enlarging the state space of the underlying Markov chain and/or by introducing more than one random variable Z [possibly dependent on the underlying Markov chain]. On the other hand, the model presented here offers a potential of considering various probability distributions for the *random noise variable* Z. Heteroschedasticity (correlations between the random noises corresponding to different time epochs) can be easily built into the model as well [this could serve as a proxy for the so called long-range dependence that is reported in the literature to manifest itself in some financial markets (see e.g. Peters [18])]. In addition, our discrete time framework allows for considering probability distributions for Z for which the exponential moment in (19) may not exist. This is important as many recent studies (see e.g. Mittnik and Rachev [14], and McCulloch [15]) indicate that returns on numerous securities are distributed according to leptokurtic (e.g. stable non-Gaussian) distributions. In such case of course our techniques would need to be adapted to maximization of the risk-sensitized value of the so called location parameter of the distribution of the log return $\ln\left(\frac{V_{t+1}}{V_t}\right)$ (see e.g. Bielecki and Pliska [7]).

References

[1] Balaji S, Borkar VS, Meyn SP (1999) Multiplicative ergodicity for an irreducible Markov chain. preprint

[2] Balaji S, Meyn SP (1999) Risk sensitive dynamic programming with unbounded cost. preprint

[3] Bensoussan A, Frehse J (1992) On Bellman equations of ergodic control in R^n. J Reine Angew Math 429:125–160

[4] Bielecki TR, Pliska SR (1999) Risk sensitive dynamic asset management. Appl. Math. Optim. 39:337–360

[5] Bielecki TR, Pliska SR (1998) Risk-sensitive dynamic asset management in the presence of transaction costs. Finance & Stochastics, forthcoming

[6] Bielecki TR, Pliska SR, Sherris M (1998) Risk sensitive asset allocation. Journal of Economic Dynamics and Control, forthcoming
[7] Bielecki TR, Pliska SR (1998) Investing coherently: Risk-Sensitive vs value-at-risk approach. unpublished manuscript
[8] Di Masi GB, Stettner L (1997) Risk sensitive control of discrete time Markov processes with infinite horizon. SIAM J. Control Optim, forthcoming
[9] Fleming WH, Hernández-Hernández D (1997) Risk sensitive control of finite state machines on an infinite horizon I. SIAM J. Control Optim. 35:1790–1810
[10] Hernández-Hernández D, Marcus SI (1996) Risk sensitive control of Markov processes in countable state space. Systems & Control Letters 29:147–155, Corrigendum (1998), vol. 34:105–106
[11] Hernández-Hernández D, Marcus SI (1997) Existence of risk sensitive optimal stationary policies for controlled Markov processes. working paper
[12] Hernández-Lerma O (1989) Adaptive Markov control processes. Springer-Verlag, New York
[13] Merton RC (1973) An intertemporal capital asset pricing model. Econometrica 41:866–887
[14] Mittnik S, Rachev ST (1993) Modeling asset returns with alternative stable distributions. Econometric Review 12:261–330
[15] McCulloch JH (1996) Financial applications of stable distributions. In: Statistical Methods in Finance (Handbook of Statistics, vol. 14), Elsevier Science
[16] Nagai H (1996) Bellman equations of risk-sensitive control. SIAM J Control & Optimization 34:74–101
[17] Pesaran MH, Timmermann A (1995) Predictability of stock returns: robustness and economic significance. J Finance 50:1201–1228
[18] Peters EE (1997) Fractal market analysis. 2nd ed., John Wiley & Sons
[19] Seneta E (1973) Non-negative matrices and Markov Chains. Springer-Verlag, New York
[20] Whittle P (1990) Risk sensitive optimal control. John Wiley, New York