Univerzita Karlova v Praze

Matematicko-fyzikální fakulta

DIPLOMOVÁ PRÁCE

Pavol Stanek

Exponenciální řízení homogenních markovských procesů

Katedra pravděpodobnosti a matematické statistiky

Vedoucí diplomové práce: Mgr. Peter Dostál Ph.D.
Studijní program: Matematika
Studijní obor: Pravděpodobnost, matematická statistika a ekonometrie

Praha 2012

Rád by som na tomto mieste poďakoval Mgr. Peterovi Dostálovi Ph.D za cenné rady, čas a ochotu, ktorú mi venoval pri tvorbe tejto práce.

Prohlašuji, že jsem tuto diplomovou práci vypracoval samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona v platném znění, zejména skutečnost, že Univerzita Karlova v Praze má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle § 60 odst. 1 autorského zákona.

V Praze dne 3.8.2012 Pavol Stanek

Název práce: Exponenciální řízení homogenních markovských procesů
Autor: Pavol Stanek
Katedra/Ústav: Katedra pravděpodobnosti a matematické statistiky MFF UK
Vedoucí diplomové práce: Mgr. Peter Dostál Ph.D., Katedra pravděpodobnosti a matematické statistiky MFF UK

Abstrakt: Táto diplomová práca sa zaoberá exponenciálnym riadením markovských reťazcov. V práci je odvodený iteračný algoritmus na nájdenie riadenia, ktoré maximalizuje mieru rastu očakávaného úžitku. Úžitok je meraný exponenciálnou úžitkovou funkciou. Algoritmus je odvodený pre reťazce v diskrétnom aj spojitom čase. Výsledný algoritmus je následne aplikovaný na problém optimálneho riadenia portfólia pri proporcionálnych transakčných nákladoch. Je odvodená dynamika vývoja investorovej pozície. Výsledný proces je approximovaný markovským reťazcom. Použitím iteračného algoritmu je numericky nájdená optimálna obchodná stratégia.

Klíčová slova: exponenciálne riadenie, markovský reťazec, optimalizácia porfólia, proporcionálne transakčné náklady

Title: Exponential control of homogeneous Markov processes
Author: Pavol Stanek
Department: Department of Probability and Mathematical Statistics, MFF UK
Supervisor: Mgr. Peter Dostál Ph.D., Department of Probability and Mathematical Statistics, MFF UK

Abstract: This master thesis concerns exponential control of Markov decision chains. An iterative alghorithm for finding a control, that maximizes a long term growth rate of expected utility is developed. The utility is measured by exponential utility function. The algorithm is derived for both discrete time and continuous time chain. Subsequently, the results are applied on the problem of optimally managing portfolio with proportional transaction costs. The dynamics of the investor's position is derived and the consequent process is approximated by Markov chain. Using the iterative alghorithm, the optimal trading strategy is numerically found.

Keywords: exponential control, Markov chain, portfolio optimization, proportional transaction costs

Contents

Introduction

This master thesis concerns with optimal control of Markov decision chainsWe develop iterative alghorithm for finding optimal control and later we apply this algorithm to a portfolio management problem.

In first chapter we first introduce the concept of utility. We focus mainly on an exponential utility function which plays crucial role in our perfomance criterion. The peformance criterion is set as a long term growth rate of a certainty equivalent of the chain's payoff. The rest of the chapter is devoted to the derivation of an iterative alghorithm for finding an optimal control. The derivation is based on Perron-Frobenius theory concerning non-negative matrices. Resulting algorithm works similarly to the standard policy iteration procedure. We derive the algorithm for both discrete time and continuous time Markov decision chain.

In the second chapter we focus on the optimal portfolio allocation problem. We consider two assets. The first one represents a bank account and the other one represents some risky asset following a Geometric Brownian motion. Moreover proportional transaction costs are paid for trading the risky asset. We analytically derive the dynamics of the value of the portfolio consisting of these two assets. For the resulting process we construct approximation by both discrete time and continuous time Markov chain. Algorithms from the Chapter 1 is then used to numerically find the strategy maximizing the long run performance criteria. Finally we compare results of both algorithms with analytical solution.

1. Exponential Control of Markov Decision Chain

In this chapter we are interested in Markov decision chains. Initial research in this field was done by Bellman [2]. He considered a Markov chain which pays a reward for moving from one state to another. Decision, affecting both transition probabilities and rewards, can by made in each state. Within this framework Howard [11] derive iterative algorithm for finding a policy which maximizes a long time expected reward. Our aim in this chapter is to derive an algorithm similar to that of Howard, but under different optimality criteria. Our optimality criteria will incorporate risk-sensitivity. To be specific, we will maximize long time expected utility from reward represented by exponential utility function.

1.1 Exponential utility function

In this section we introduce the concept of utility function. Especially we focus on the exponential utility function. We will show its basic properties and motivation for its usage. For more insight about utility, see the Chapter 2 in [8].

Imagine a lottery that pays certain prizes with certain probabilities. The outcome of the lottery is unknown, but what is known is its probability distribution. Interesting question is how much one should be willing to pay for opportunity to play such a lottery. Or, in other words, what is the fair price of a game. The first idea that crosses our mind is to take expected value as a fair price of a game. However according to this approach one is indifferent between lottery with a certain average pay-off and getting this amount of money straight away. In fact the majority of people would prefer the second alternative. We see that the answer is not so simple. Certainly, it depends on players aversion to risk. Personal attitude to risk can be modeled by a utility function.

Definition 1.1. Let S be an interval in $\mathbb{R}$. A function $u : S \to \mathbb{R}$ is called a *utility function* if u is strictly increasing, strictly concave and continuous on S.

Note that since concavity of a function on an open set implies its continuity, continuity requirement in the above definition concerns only boundary points of S.

Denote the set of all probability measures on $(S, \mathcal{B}(S))$ with finite expectation by $\mathcal{P}$. We interpret $\mathcal{P}$ as the set of all possible random pay-offs. In this sense every $\mu \in \mathcal{P}$ can be considered as a *lottery*. Denote the *expected utility* of the lottery μ by $U(\mu)$. That is

$$U(\mu) = \int_{\mathbb{R}} u \, d\mu.$$

The expected utility defines the preference relation $\succ$ on $\mathcal{P}$ by

$$\mu \succ \nu \Leftrightarrow U(\mu) > U(\nu).$$

The individual is willing to pay for the lottery with pay-off μ at most the amount of money $\mathtt{CE}(\mu)$ which provides him the same utility. In other words, the individual is indifferent between lottery μ and a certain amount of money $\mathtt{CE}(\mu)$. That means

$$U(\mu) = U(\delta_{\mathtt{CE}(\mu)}) = u(\mathtt{CE}(\mu))$$

must be satisfied. The number $\mathtt{CE}(\mu)$ is called *certainty equivalent.* Since S is an interval, $U(\mu) \in S$ and due to monotonicity of u, we can write $\mathtt{CE}(\mu) = u^{-1}(U(\mu))$. So the certainty equivalent is always defined and unique.

The requirement of strict concavity in the definition of utility function incorporates kind of risk averseness. Suppose the lottery ν that pays the amount a with probability p and the amount b with probability $1-p$. If $a \neq b$, then by concavity

$$p\,u(a) + (1-p)\,u(b) < u(p\,a + (1-p)\,b),$$

so $U(\nu) < u(\mathbb{E}(\nu))$. In words, receiving the amount of money equal to expected value of ν is preferred to playing the lottery ν.

We say that a utility function fulfils the *delta property* if for $\Delta > 0$ and $\mu \in \mathcal{P}$

$$\mathtt{CE}(\mu + \Delta) = \mathtt{CE}(\mu) + \Delta.$$

If the pay-off of a lottery is increased by the constant $\Delta > 0$, certainty equivalent also increases by Δ. It can be shown that the only utility function that possesses delta property is of the form

$$u(x) = a - be^{-\gamma x}, \quad \gamma < 0, \tag{1.1}$$

where $b > 0$ and $a \in \mathbb{R}$. A utility function of this type is called an *exponential utility function.* Since any increasing affine transformation of a utility function leads to the same preference relation on $\mathcal{P}$ and the same certainty equivalent, it is sufficient to consider exponential utility function with parameters $a = 0$ and $b = 1$.

If we relax the concavity condition in the definition of an utility function, then utility functions possessing delta property can be characterized as follows: It is either a linear function or a function of the type (1.1) with $\gamma \in \mathbb{R}$. Preferences described by a linear utility function coincides with the approach where the fair price of a lottery equals to its expected value.

The exponential utility function is also characterized by so called constant absolute risk aversion.

Definition 1.2. Suppose that a utility function u is twice differentiable. Then the number

$$\alpha(x) = -\frac{u''(x)}{u'(x)}$$

is called the *Arrow-Pratt coefficient of absolute risk aversion* at the level x.

The exponential utility function is characterized by the property $\alpha(x) = -\gamma > 0$. This can be easily shown by solving the respective differential equation. Due to this property the exponential utility function is often referred to as a constant absolute risk aversion (CARA) utility function.

The concept of expected utility can be naturally extended from lotteries to random variables with values in S. In the following sections we will examine the expected utility from a Markov reward chain under the exponential utility function.

1.2 Exponential utility and Markov reward chain

Prior to the introduction of formal definition of a Markov chain, let us make a note on the definition that we will use. Throughout this thesis we will use a more general concept of Markov chain than can be found in introductory textbooks. We define a Markov chain not only by terms of probability distribution, but also by a filtration, which could be different from the one generated by the process. This approach is widely used in the general theory of Markov processes which considers continuous time and continuous state domain. See the monograph [7] for more insight.

Fixing the filtration in the definition has such a consequence that we cannot consider different process when changing initial distribution. Instead the underlying probability measure changes and the process remains the same. With this approach we gain, beside a more general theory, a significant simplification of notation. So the computations will be more tractable.

Let $\mathcal{X}$ be a finite set. Each $x \in \mathcal{X}$ is called a *state* and $\mathcal{X}$ is called the *state space*. Let $\boldsymbol{P} = \{p_{xy}\}_{x,y\in\mathcal{X}}$ be a stochastic matrix, that is a matrix with non-negative entries and with each row summing up to 1. The matrix $\boldsymbol{P}$ is called the *transition matrix*. Denote entries of its n-th power by $p_{xy}^{(n)}$.

Definition 1.3. A system $(\Omega, \{\mathcal{F}_n\}_{n\in\mathbb{N}_0}, X, \{\mathbb{P}_x\}_{x\in\mathcal{X}})$ is called a *(homogenous) Markov chain* with the transition matrix $\boldsymbol{P}$ if

(i) $\mathcal{F}_n$, $n \in \mathbb{N}_0$ are σ-fields on Ω satisfying $\mathcal{F}_m \subset \mathcal{F}_n$ for all $m \le n$,

(ii) $X = \{X_n, n \in \mathbb{N}_0\}$ is $\mathcal{X}$-valued process defined on Ω and X_m is $\mathcal{F}_n$ measurable for all $m \le n$,

(iii) for all $x \in \mathcal{X}$, $\mathbb{P}_x$ is a probability measure on $\mathcal{F}_\infty = \sigma\left(\bigcup_{n=0}^\infty \mathcal{F}_n\right)$ satisfying for all $y \in \mathcal{X}$ and all $m, n \in \mathbb{N}_0$, $m \le n$

$$\mathbb{P}_x(X_0 = x) = 1 \quad \text{and} \quad \mathbb{P}_x(X_n = y|\mathcal{F}_m) = p_{X_m,y}^{(n-m)}, \quad \mathbb{P}_x - a.s.$$

The property (iii) implies that the transition probability

$$\mathbb{P}_x(X_n = z|X_m = y) = p_{y,z}^{(n-m)} \quad \mathbb{P}_x - a.s. \tag{1.2}$$

almost surely does not depend on the initial state x. Consider a function f defined on $\mathcal{X}$ and put

$$E_x^{(n)} f \triangleq \sum_{y \in \mathcal{X}} f(y)\, p_{x,y}^{(n)} = \boldsymbol{e}_x^{\mathrm{T}}\, \boldsymbol{P}^n\, f.$$

Then, using (1.2), for all $x \in \mathcal{X}$ we have

$$\mathbb{E}[f(X_n)|X_m = y] \stackrel{\text{as}}{=} \sum_{z \in \mathcal{X}} f(z) \mathbb{P}_x(X_n = z|X_m = y) \stackrel{\text{as}}{=} E_y^{(n-m)} f.$$

Consequently

$$\mathbb{E}_x[f(X_n)|\mathcal{F}_m] = E_{X_m}^{(n-m)} f, \quad \mathbb{P}_x - a.s. \quad \text{for all } x \in \mathcal{X}. \tag{1.3}$$

We see that the conditional expectation $\mathbb{E}_x[f(X_n)|\mathcal{F}_m]$ is almost surely equal to the same random variable regardless of initial distribution. Thus we will write $\mathbb{E}$ instead of $\mathbb{E}_x$ in this case. Of course its distribution typically depends on the initial state, since change of the initial state means change of the underlying measure. The equation (1.3) is often referred to as the Markov property.

Our definition treats explicitly only situation when the chain starts from arbitrary state x. However we can easily construct a probability measure that covers any initial distribution. Consider a probability distribution μ on $\mathcal{X}$. Define

$$\mathbb{P}_\mu(A) = \sum_{x \in \mathcal{X}} \mu(x)\, \mathbb{P}_x(A) \quad A \in \mathcal{F}_\infty. \tag{1.4}$$

Then $\mathbb{P}_\mu$ is a probability measure on $\mathcal{F}_\infty$ satisfying

$$\mathbb{P}_x(X_n = y|\mathcal{F}_m) = \sum_{x \in \mathcal{X}} \mu(x)\, p_{X_m,y}^{(n-m)} = p_{X_m,y}^{(n-m)}, \quad \mathbb{P}_x - a.s.$$

So the Markov property holds for $\mathbb{P}_\mu$. Moreover (1.4) implies

$$\mathbb{P}_\mu(X_0 = x) = \sum_{y \in \mathcal{X}} \mu(y)\, \mathbb{P}_y(X_0 = x) = \mu(x).$$

In words, the chain has initial distribution μ under the measure $\mathbb{P}_\mu$.

As one expects the n-th power of $\boldsymbol{P}$ describes transition probabilities in n periods. Indeed, for any states x and y we have

$$\mathbb{P}_x(X_n = y) = \mathbb{E}_x \mathbb{P}_x(X_n = y|\mathcal{F}_0) = \mathbb{E}_x\, p_{X_0,y}^{(n)} = p_{xy}^{(n)}, \quad \mathbb{P}_x - a.s.$$

We will restrict our consideration on Markov chains that have a agreeable behaviour described in a following couple of definitions.

Definition 1.4. We say that a state x is *accessible* from state y if the n-step transition probability $p_{xy}^{(n)}$ is strictly positive for at least one n. A Markov chain is called *irreducible* if all states are accessible from each other.

Definition 1.5. Let $x \in \mathcal{X}$ and let d_x be the greatest common divisor of those $n \geq 1$ for which $p_x^{(n)}$ is strictly positive. If $d_x > 1$ then the state x is called *periodic*. If $d_x = 1$ then the state x is called *aperiodic*. A Markov chain is called *aperiodic* if all states are aperiodic.

Now we add rewards to the chain. Suppose that the transition from state x to state y pays a reward $r(x, y) \in \mathbb{R}$ (sometimes denoted by r_{xy}). The matrix $\boldsymbol{R} = \{r_{xy}\}_{x,y \in \mathcal{X}}$ is called a *reward matrix* and a Markov chain enriched with a reward matrix is called *Markov reward chain*.

From now on we will consider an aperiodic irreducible Markov reward chain X with the state space $\mathcal{X} = \{1, \ldots, N\}$ and CARA utility function of the form

$$\mathcal{U}_\gamma^{\mathsf{C}}(x) = -e^{\gamma x} \quad \gamma < 0. \tag{1.5}$$

In Section 1.1 we explained that for fixed γ exponential utility function of any form represents the same preference order. Therefore the specific choice (1.5) does not impose any restriction on our consideration. The aim is to compute the expected utility of the total reward of X over the time and examine its asymptotic properties. For a given vector $\boldsymbol{b} \in \mathbb{R}^N$ define

$$U_n(\boldsymbol{b}) \triangleq \mathcal{U}_\gamma^{\mathsf{C}}\Big(\sum_{k=1}^{n} r(X_{k-1}, X_k)\Big) \cdot \boldsymbol{b}(X_n). \tag{1.6}$$

Here the symbol $\boldsymbol{b}(X_n)$ simply means the X_n-th component of the vector $\boldsymbol{b}$. Note that $U_n \triangleq U_n(\mathbf{1})$ is the total utility from reward over times $0, \ldots, n$. Considering this quantity with arbitrary factor $\boldsymbol{b}(X_n)$ turns out to be useful. In order to avoid confusion, where it is desirable, we write the dot symbol indicating multiplication. Finally note that $U_n(\boldsymbol{b})$ can be equally expressed as

$$U_n(\boldsymbol{b}) = U_n(\mathbf{1}) \cdot \boldsymbol{b}(X_n) = U_n \cdot \boldsymbol{b}(X_n). \tag{1.7}$$

First look at the conditional expectation of $U_n(\boldsymbol{b})$.

$$\begin{aligned}
\mathbb{E}[U_n(\boldsymbol{b})|\mathcal{F}_{n-1}] &= \mathbb{E}\left[-\exp\left\{\gamma \sum_{k=1}^{n-1} r(X_{k-1}, X_k)\right\} \exp\{\gamma\, r(X_{n-1}, X_n)\}\, \boldsymbol{b}(X_n) \Big| \mathcal{F}_{n-1}\right] \\
&= U_{n-1} \cdot \mathbb{E}[\exp\{\gamma\, r(X_{n-1}, X_n)\}\, \boldsymbol{b}(X_n)|\mathcal{F}_{n-1}] \\
&\overset{\text{as}}{=} U_{n-1} \cdot \sum_{i=1}^{N} \mathbb{I}_{[X_{n-1}=i]} \sum_{j=1}^{N} p_{ij}\, e^{\gamma\, r_{ij}}\, \boldsymbol{b}(j)
\end{aligned} \tag{1.8}$$

Define the matrix $\boldsymbol{S}$ by

$$\boldsymbol{S} = \{s_{ij}\}_{i,j=1}^{N} \quad \text{with} \quad s_{ij} \triangleq p_{ij}\, e^{\gamma\, r_{ij}}, \tag{1.9}$$

Equally we can write the definition of $\boldsymbol{S}$ in matrix notation as

$$\boldsymbol{S} \triangleq \boldsymbol{P} * \mathbf{exp}\{\gamma\, \boldsymbol{R}\}. \tag{1.10}$$

where the symbol **exp** represents exponential function applied entrywise and the symbol $*$ represents entrywise product, also known as Hadamard product. Contrary to the ordinary matrix multiplication, the Hadamard product is commutative. Now we can rewrite the second factor of the last term in (1.8) as

$$\mathbb{E}[-\mathcal{U}^{\mathsf{C}}_{\gamma}(r(X_{n-1},X_n))\,\boldsymbol{b}(X_n)|\mathcal{F}_{n-1}] = \boldsymbol{e}^{\mathsf{T}}_{X_{n-1}}\,\boldsymbol{S}\,\boldsymbol{b} = (\boldsymbol{S}\,\boldsymbol{b})(X_{n-1}), \tag{1.11}$$

where $\boldsymbol{e}_k$ is the k-th canonical vector. Thus, using the notation (1.7), we have

$$\mathbb{E}[U_n(\boldsymbol{b})|\mathcal{F}_{n-1}] = U_{n-1}\cdot(\boldsymbol{S}\,\boldsymbol{b})(X_{n-1}) = U_{n-1}(\boldsymbol{S}\,\boldsymbol{b}). \tag{1.12}$$

We make the following conclusion.

Proposition 1.6. *For any given vector $\boldsymbol{b}\in\mathbb{R}^N$, the variable $U_n(\boldsymbol{b})$ defined by* (1.6) *and $n \geq k \geq 0$ we have*

$$\mathbb{E}[U_n(\boldsymbol{b})|\mathcal{F}_{n-k}] = U_{n-k}(\boldsymbol{S}^k\,\boldsymbol{b}),$$

where the matrix $\boldsymbol{S}$ is defined by (1.10)*.*

Proof. Using induction the result immediately follows from (1.12). □

As a consequence we get the lemma which will be useful when we move to continuous time set-up. First generalize the definition (1.6) by putting

$$U_{m,n}(\boldsymbol{b}) \triangleq \mathcal{U}^{\mathsf{C}}_{\gamma}\Big(\sum_{k=m+1}^{n} r(X_{k-1},X_k)\Big)\cdot\boldsymbol{b}(X_n)\in L_1, \tag{1.13}$$

for $0 \leq m < n$. Variable $U_{m,n}(\boldsymbol{b})$ is integrable because it attains only a finite number of values. Note that $U_{0,n}(\boldsymbol{b}) = U_n(\boldsymbol{b})$ and

$$U_n(\boldsymbol{b}) = U_{n-k}\cdot(-U_{n-k,n}(\boldsymbol{b})).$$

Lemma 1.7. *For any given vector $\boldsymbol{b}\in\mathbb{R}^N$, the variable $U_{m,n}(\boldsymbol{b})$ defined by* (1.13) *and $0 < k < n$ we have*

$$\mathbb{E}[-U_{n-k,n}(\boldsymbol{b})|\mathcal{F}_{n-k}] = (\boldsymbol{S}^k\,\boldsymbol{b})(X_{n-k}).$$

Proof. Using Proposition 1.6 and the fact that $U_{n-k,n}$ is integrable we get

$$\mathbb{E}[U_n(\boldsymbol{b})|\mathcal{F}_{n-k}] = U_{n-k}\cdot\mathbb{E}[-U_{n-k,n}(\boldsymbol{b})|\mathcal{F}_{n-k}] = U_{n-k}\cdot(\boldsymbol{S}^k\,\boldsymbol{b})(X_{n-k}).$$ □

Now we can easily compute the expected utility of the reward up to time n.

Corollary 1.8. *Denote $\boldsymbol{u}_n(\boldsymbol{b}) = (u_{n,1}(\boldsymbol{b}),\ldots,u_{n,N}(\boldsymbol{b}))^{\mathsf{T}}$, where $u_{n,i}(\boldsymbol{b}) = \mathbb{E}_i[U_n(\boldsymbol{b})]$. Then*

$$u_n(\boldsymbol{b}) = -\boldsymbol{S}^n\boldsymbol{b}.$$

Proof. Using Proposition 1.6 with $n = k$ we get

$$\mathbb{E}_i[U_n(\boldsymbol{b})] = \mathbb{E}[U_n(\boldsymbol{b})|X_0 = i] = U_0(\boldsymbol{S}^n\boldsymbol{b}) = -(\boldsymbol{S}^n\boldsymbol{b})(i) = -\boldsymbol{e}_i^{\mathsf{T}}\boldsymbol{S}^n\boldsymbol{b} \quad \mathbb{P}_i - a.s. \qquad \square$$

Note that the i-th component of the vector $\boldsymbol{u}_n = \boldsymbol{u}_n(\mathbf{1})$ is equal to the total expected utility from reward over times $0, \ldots, n$ when starting in state i at time 0. Corollary 1.8 says that the time development of the total expected utility is described by the power of the matrix $\boldsymbol{S}$ defined by (1.9). Because the matrix $\boldsymbol{S}$ has non-negative entries we can apply Perron-Frobenius theory. This theory concerns the maximal eigenvalue of a non-negative matrix and consequently asymptotic properties of its power. For here used matrix results see Appendix A.

Because we assume that X is irreducible and aperiodic, the matrix $\boldsymbol{P}$ is also irreducible and aperiodic (see Appendix A). The matrix $\boldsymbol{S}$ is constructed from $\boldsymbol{P}$ by multiplying each element by the positive number. Thus the matrix $\boldsymbol{S}$ also possesses mentioned properties. By the Perron-Frobenius theorem A.3 the maximal eigenvalue of $\boldsymbol{S}$, that is the one with maximal Euclidean norm, is real and positive. Denote this eigenvalue by $\lambda > 0$. Moreover the eigenvector $\boldsymbol{u}$ respective to eigenvalue λ can be chosen, such that all its entries are positive, which we denote as $\boldsymbol{u} > 0$. Further by Theorem A.4 we have

$$\lim_{n\to\infty} \lambda^{-n}\,\boldsymbol{u}_n = \lim_{n\to\infty} -\lambda^{-n}\boldsymbol{S}^n\,\mathbf{1} = -k\,\boldsymbol{u}, \tag{1.14}$$

where $\boldsymbol{u} > 0$ is an eigenvector respective to λ and k is a positive constant. Consequently for any initial state i

$$\lim_{n\to\infty} \frac{\boldsymbol{u}_{n+1}(i)}{\boldsymbol{u}_n(i)} = \lim_{n\to\infty} \lambda\,\frac{\lambda^{-(n+1)}\boldsymbol{u}_{n+1}(i)}{\lambda^{-n}\boldsymbol{u}_n(i)} = \lambda.$$

So for large times the utility is multiplied by λ each time we move to the next state. This property holds regardless of initial state. That is the relation (1.14) can be equivalently expressed as

$$\boldsymbol{u}_n = \lambda^n(k\,\boldsymbol{u} + o(\mathbf{1})) \quad n\to\infty. \tag{1.15}$$

Now look what the preceding limiting property (1.15) means for the certainty equivalent. Denote the certainty equivalent of U_n under CARA utility when starting in state i by $\mathtt{CE}_n^{\mathtt{C}}(i)$. That is $\mathtt{CE}_n^{\mathtt{C}}(i)$ is equal to $(\mathcal{U}_\gamma^{\mathtt{C}})^{-1}(\boldsymbol{u}_n(i))$, where

$$(\mathcal{U}_\gamma^{\mathtt{C}})^{-1}(x) = \tfrac{1}{\gamma}\log(-x).$$

Then we can express a limiting relation for certainty equivalent as follows

$$\begin{aligned} \mathtt{CE}_n^{\mathtt{C}} &= \tfrac{1}{\gamma}\,\mathbf{log}(-\boldsymbol{u}_n) \\ &= \tfrac{1}{\gamma}\,n\,\log\lambda\,\mathbf{1} + \tfrac{1}{\gamma}\,\mathbf{log}(-k\,\boldsymbol{u} + o(\mathbf{1})), \quad n\to\infty, \end{aligned} \tag{1.16}$$

where $\mathbf{log}$ represents logarithm function applied entrywise. Denote $\boldsymbol{z} \triangleq \frac{1}{\gamma}\mathbf{log}(-k\boldsymbol{u})$ and

$$g \triangleq \tfrac{1}{\gamma}\log\lambda. \tag{1.17}$$

We can rewrite (1.16) as

$$\mathtt{CE}_n^{\mathtt{C}} = n\,g\,\mathbf{1} + \boldsymbol{z} + o(\mathbf{1}), \quad n \to \infty. \tag{1.18}$$

The interpretation is that for large times the certainty equivalent increases by g when the Markov chain moves to another state. The vector $\boldsymbol{z}$ is the correction reflecting the initial state. When time is large initial state becomes less important, which is expressed by the following limit holding for every initial state i

$$\frac{\mathtt{CE}_n^{\mathtt{C}}(i)}{n} \to g, \quad n \to \infty. \tag{1.19}$$

In the next section we will look for optimal control of a Markov chain which maximizes certain equivalent growth rate g. Note that due to negativity of γ, (1.17) implies that maximization of g is equivalent to the minimization of λ. This is a bit counter-intuitive and it is caused by the fact that our utility function attains only negative values. Finally note that the limit to be maximized, given by (1.19), can be equivalently expressed by previously used symbols as

$$g = \lim_{n\to\infty} \tfrac{1}{\gamma}\, n^{-1} \, \log(-\mathbb{E}[U_n]). \tag{1.20}$$

Before we proceed let us discuss an important special case. Sometimes the reward is not related with transition from one state to another but is paid when staying in a particular state. In this case the reward is described by a *reward vector* $\boldsymbol{r} = \{r_i\}_{i\in\mathcal{X}}$ determining a reward paid in each state. This is equivalent to considering the reward matrix $\boldsymbol{R} = \{r_{ij}\}_{i,j\in\mathcal{X}}$, with $r_{ij} = r_i$. Then the matrix $\boldsymbol{S}$ can be expressed as

$$\boldsymbol{S} = \boldsymbol{P} * \mathbf{exp}\{\gamma\,\boldsymbol{R}\} = \mathbf{exp}\{\gamma\,\mathrm{diag}(\boldsymbol{r})\} \cdot \boldsymbol{P}. \tag{1.21}$$

1.3 Optimal risk sensitive control of Markov decision chain

We introduce the concept where a decision maker can effect the behaviour of the chain. In each state an action can be chosen. Vector of all actions selected in each state determines the rewards and the transition probabilities of the chain. Formally consider a state space $\mathcal{X}$ and let A_x be a finite set representing a set of all admissible actions in a state $x \in \mathcal{X}$. Define an *action space* $\mathcal{A}$ as

$$\mathcal{A} = \prod_{x \in \mathcal{X}} A_x.$$

Further for any *policy* $a \in \mathcal{A}$ let $\boldsymbol{P}^a$ be a transition matrix and let $\boldsymbol{R}^a$ be a reward matrix corresponding to this policy. That is for every policy $a \in \mathcal{A}$ we have a Markov reward chain X^a. The whole system $(\Omega, \{\mathcal{F}_n\}_{n \in \mathbb{N}_0}, \{\mathrm{X}^a\}_{a \in \mathcal{A}})$ is called *a Markov decision chain.*

Consider a Markov decision chain with state space $\mathcal{X} = \{1, \ldots, N\}$, transition matrices

$$\boldsymbol{P}^a = (\boldsymbol{p}_1(a_1), \ldots, \boldsymbol{p}_N(a_N))^{\intercal}$$

and reward matrices

$$\boldsymbol{R}^a = (\boldsymbol{r}_1(a_1), \ldots, \boldsymbol{r}_N(a_N))^{\intercal}.$$

We assume that for all policies $a \in \mathcal{A}$ the transition matrix $\boldsymbol{P}^a$ is aperiodic and irreducible. For any policy a we have the variable $U_n^a(\cdot)$ defined like in (1.6) and the non-negative matrix $\boldsymbol{S}^a$ defined like in (1.10). That is

$$\begin{aligned} \boldsymbol{S}^a = (\boldsymbol{s}_1(a_1), \ldots, \boldsymbol{s}_N(a_N))^{\intercal}, \quad \boldsymbol{s}_i(a_i) &= \{s_{ij}(a_i)\}_{j=1}^N, \\ s_{ij}(a_i) &= p_{ij}(a_i) \exp\{\gamma \, r_{ij}(a_i)\}. \end{aligned}$$

By virtue of the Perron-Frobenius theorem A.3 there exists the maximal eigenvalue $\lambda^a > 0$ and respective eigenvector $\boldsymbol{v}^a > 0$ of $\boldsymbol{S}^a$.

$$M_n^a \triangleq (\lambda^a)^{-n} U_n^a(\boldsymbol{v}^a) \quad n \in \mathbb{N}_0, \tag{1.22}$$

where the variable $U_n^a(\cdot)$ defined like in (1.6).

Proposition 1.9. *The process M_n^a is a $\mathcal{F}_n$-supermartingale if $\boldsymbol{S}^a \boldsymbol{v}^a \geq \lambda^a \boldsymbol{v}^a$ entry-wise. The process M_n^a is a $\mathcal{F}_n$-martingale if $\boldsymbol{S}^a \boldsymbol{v}^a = \lambda^a \boldsymbol{v}^a$.*

Proof. As the Proposition concerns only one fixed policy, we will omit the upper index a throughout the proof. Using Proposition 1.6 and the fact that M_{n-1} is $\mathcal{F}_{n-1}$ measurable we get

$$\begin{aligned} \mathbb{E}[M_n - M_{n-1} | \mathcal{F}_{n-1}] &\overset{\text{as}}{=} (\lambda)^{-n} \, U_{n-1}(\boldsymbol{S}\boldsymbol{v}) - (\lambda)^{-n+1} \, U_{n-1}(\boldsymbol{v}) \\ &= (\lambda)^{-n} [U_{n-1}(\boldsymbol{S}\boldsymbol{v}) - \lambda \, U_{n-1}(\boldsymbol{v})] \\ &= (\lambda)^{-n} \, U_{n-1} \cdot [(\boldsymbol{S}\,\boldsymbol{v})(X_{n-1}) - (\lambda\,\boldsymbol{v})(X_{n-1})]. \end{aligned}$$

If $\boldsymbol{S}\,\boldsymbol{v} = \lambda\,\boldsymbol{v}$ the process M_n is clearly an $\mathcal{F}_n$-martingal. Finally, from assumption $\boldsymbol{S}\,\boldsymbol{v} - \lambda\,\boldsymbol{v} \geq 0$ and from facts that λ is positive and U_{n-1} is negative follows $\mathbb{E}[M_n|\mathcal{F}_{n-1}] \overset{\text{as}}{\leq} M_{n-1}$. □

Theorem 1.10. *Let $\widehat{a} \in \mathcal{A}$. If the inequality $\boldsymbol{S}^a \boldsymbol{v}^{\widehat{a}} \geq \lambda^{\widehat{a}}\,\boldsymbol{v}^{\widehat{a}}$ holds for every policy $a \in \mathcal{A}$, than $\widehat{a}$ is an optimal policy in sense that $g^{\widehat{a}} \geq g^a$ for all $a \in \mathcal{A}$.*

Proof. For all i we have

$$\frac{\sum_j s_{ij}(a_i)\, v_j^{\widehat{a}}}{v_i^{\widehat{a}}} \geq \lambda^{\widehat{a}}, \quad v_i^{\widehat{a}} > 0.$$

Using the inequality from Theorem C.1 we get

$$\lambda^a \geq \min_i \frac{\sum_j s_{ij}(a_i)\, v_j^{\widehat{a}}}{v_i^{\widehat{a}}} \geq \lambda^{\widehat{a}}.$$

Remind that according to (1.17) the growth rate of certainty equivalent g is equal to $\frac{1}{\gamma}\log\lambda$. Since γ is negative we have $\widehat{g} \geq g^a$. □

Theorem 1.11 (Policy iteration)**.** *Let $a_0 \in \mathcal{A}$ be the initial policy. Define the sequence $\{a_n\}$ recursively by*

$$a_{n+1}(i) \triangleq \operatorname*{argmin}_{\alpha\in\mathcal{A}_i} \boldsymbol{s}_i(\alpha)^{\mathsf{T}}\, \boldsymbol{v}^{a_n}. \tag{1.23}$$

If the minimum is attained for more the one policy and $a_n(i)$ is one of then, always make a conservative choice $a_{n+1}(i) = a_n(i)$. The resulting sequence a_n converges to an optimal policy $\widehat{a}$.

Proof. Suppose that a_n converges to $\widehat{a}$. Then by (1.23) for all i we have

$$\boldsymbol{e}_i^{\mathsf{T}}\, \boldsymbol{S}^a\, \boldsymbol{v}^{\widehat{a}} \geq \boldsymbol{e}_i^{\mathsf{T}}\, \boldsymbol{S}^{\widehat{a}}\, \boldsymbol{v}^{\widehat{a}} = \lambda^{\widehat{a}}\, v_i^{\widehat{a}}, \quad a \in \mathcal{A}.$$

Thus the inequality $\boldsymbol{S}^a\, \boldsymbol{v}^{\widehat{a}} \geq \lambda^{\widehat{a}}\, \boldsymbol{v}^{\widehat{a}}$ holds for every $a \in \mathcal{A}$ and the policy $\widehat{a}$ is optimal by virtue of Theorem 1.10.

It now remains to show that a_n converges. Because the action space is a discrete finite set, the sequence a_n converges if and only if it is eventually constant. Thus it is sufficient to show that every time the procedure moves to a new policy it is better than the previous one. So suppose that $a_n \neq a_{n+1}$. Denote $a_n \triangleq a$ and $a_{n+1} \triangleq b$. Then by (1.23) the inequality

$$\boldsymbol{e}_i^{\mathsf{T}}\, \boldsymbol{S}^b\, \boldsymbol{v}^a = \boldsymbol{s}_i(b_i)^{\mathsf{T}}\, \boldsymbol{v}^a \overset{(1.23)}{\leq} \boldsymbol{s}_i(a_i)^{\mathsf{T}}\, \boldsymbol{v}^a = \boldsymbol{e}_i^{\mathsf{T}}\, \boldsymbol{S}^a\, \boldsymbol{v}^a = \lambda\, \boldsymbol{v}_i^a$$

holds for every i. In addition for some i the inequality must be sharp, i.e.

$$\boldsymbol{S}^b\, \boldsymbol{v}^a \lneqq \lambda^a\, \boldsymbol{v}^a.$$

Otherwise we would have $a = b$ due to the conservative choice.

Here we distinguish two cases. If $\boldsymbol{v}^a$ is an eingenvector of $\boldsymbol{S}^b$ respective to λ^a, then $\lambda^b \boldsymbol{v}^a < \lambda^a \boldsymbol{v}^a$. Because $\boldsymbol{v}^a$ has positive entries we have $\lambda^b < \lambda^a$ and consequently $g^b > g^a$.

Now suppose that $\boldsymbol{v}^a$ is not an eingenvector of $\boldsymbol{S}^b$ respective to λ^a. Then for all i

$$\frac{\sum_j s_{ij}(b_i)\, v_j^a}{v_i^a} \leq \lambda^a$$

with sharp inequality for some i. Using inequality from theorem C.1 we get

$$\lambda^b < \max_i \frac{\sum_j s_{ij}(b_i)\, v_j^a}{v_i^a} \leq \lambda^a$$

and consequently $g^b > g^a$. Note that the inequality is sharp, because $\boldsymbol{v}^a$ is not an eingenvector of $\boldsymbol{S}^b$. $\square$

The Theorem 1.11 gives us a method for finding the optimal policy. The algorithm can by summarized by the following steps:

1. For every policy a construct the matrix $\boldsymbol{S}^a$ according to (1.10).
2. Choose an arbitrarily initial policy a_n, $n = 0$.
3. Having the policy a_n, find an eigenvector $\boldsymbol{v}^{a_n} > 0$ of the matrix $\boldsymbol{S}^{a_n}$ corresponding to its maximal eigenvalue.
4. Find the improved policy a_{n+1} by entrywise minimization $\boldsymbol{S}^a \boldsymbol{v}^{a_n}$ over all policies a.
5. Repeat steps 3. and 4. until $a_n = a_{n+1}$. You have just find an optimal policy.

1.4 Continuous time Markov reward chain

In order to define a continuous time Markov chain we start with a transition probabilities. Let $\mathcal{X}$ be a finite set and let $\{\boldsymbol{P}_t, t \geq 0\}$ with $\boldsymbol{P}_t = \{p_{xy}(t)\}_{x,y\in\mathcal{X}}$ be a system of stochastic matrixes satisfying

(B-1) $\boldsymbol{P}_0 = \boldsymbol{I}$,

(B-2) $\boldsymbol{P}_{s+t} = \boldsymbol{P}_s\, \boldsymbol{P}_t, \quad s, t \geq 0$,

(B-3) $\lim_{t\to 0_+} \boldsymbol{P}_t = \boldsymbol{I}$.

Property (B-2) is known as *Chapman-Kolmogorov relation*. Condition (B-1) together with (B-3) assert that the system is right continuous at $t = 0$. Moreover all three conditions together imply both sided continuity of the system $\{\boldsymbol{P}_t, t \geq 0\}$ in all $t \geq 0$. Right continuity follows from computation

$$\lim_{h \to 0_+} \boldsymbol{P}(t+h) = \boldsymbol{P}(t) \lim_{h \to 0_+} \boldsymbol{P}(h) = \boldsymbol{P}(t)\boldsymbol{I} = \boldsymbol{P}(t). \tag{1.24}$$

On the other hand, for $t > h > 0$ Chapman-Kolmogorov relation implies

$$\boldsymbol{P}_t = \boldsymbol{P}_{t-h}\boldsymbol{P}_h.$$

The matrix $\boldsymbol{P}_h$ is near to identity matrix for h small enough and thus the inverse $\boldsymbol{P}_h^{-1}$ exists and also converges to identity matrix. So we have

$$\boldsymbol{P}_t = \boldsymbol{P}_t \lim_{h \to 0_+} \boldsymbol{P}_h^{-1} = \lim_{h \to 0_+} \boldsymbol{P}(t-h). \tag{1.25}$$

Relations (1.24) and (1.25) says that $\{\boldsymbol{P}_t, t \geq 0\}$ is continuous in all $t \geq 0$.

Remind that every matrix represents a linear operator and composition of linear operators corresponds to matrix multiplication. Thus the above defined system together with matrix multiplication constitutes a continuous semigroup of linear operators on $\mathbb{R}^{|\mathcal{X}|}$.

Definition 1.12. A system $(\Omega, \{\mathcal{F}_t\}_{t \in \mathbb{R}_0}, X, \{\mathbb{P}_x\}_{x \in \mathcal{X}})$ is called a *(homogenous) continuous time Markov chain* with the transition probabilities $\{\boldsymbol{P}_t, t \geq 0\}$ if

(i) $\mathcal{F}_t$, $t \geq 0$ are σ-fields on Ω satisfying $\mathcal{F}_s \subset \mathcal{F}_t$ for all $s \leq t$,

(ii) $X = \{X_t, t \geq 0\}$ is $\mathcal{X}$-valued process on Ω and X_s is $\mathcal{F}_t$ measurable for all $s \leq t$,

(iii) for all $x \in \mathcal{X}$ $\mathbb{P}_x$ is a probability measure on $\mathcal{F}_\infty = \sigma\left(\bigcup_{t \geq 0} \mathcal{F}_t\right)$ satisfying for all $y \in \mathcal{X}$, $s, t \in \mathbb{R}_0$, $s \leq t$

$$\mathbb{P}_x(X_0 = x) = 1 \quad \text{and} \quad \mathbb{P}_x(X_t = y | \mathcal{F}_s) = p_{X_s, y}(t-s) \quad \mathbb{P}_x - a.s.$$

Similarly to the discussion after Definition 1.3 we can show that the process possesses the Markov property and the matrix $\boldsymbol{P}(t)$ describes transition in time t meaning that

$$p_{xy}(t) = \mathbb{P}_x(X_t = y).$$

In discrete time the distribution of a Markov chain is fully described by initial distribution (in our case represented by underlying measure) and one step transition matrix. In continuous time there is no smallest transition step. In order to find similar simple description in continuous time we need to examine derivative of transition functions $p_{xy}(t)$, which is, loosely speaking, infinitesimal transition step. We have already seen that these functions are continuous, but it can be shown that they have much stronger properties. The key results are summarized in the following theorem.

Theorem 1.13. *Let $\mathcal{X}$ be a countable set. Assume that the system $\boldsymbol{P}_t = \{p_{xy}(t)\}_{x,y\in\mathcal{X}}$ satisfies conditions (B-1)-(B-3) then*

(i) For every $x, y \in \mathcal{X}$ the function $p_{xy}(t)$ is uniformly continuous in t.

(ii) For all $x \in \mathcal{X}$ the derivative

$$q_x = -q_{xx} = -\frac{d}{dt_+} p_{xx}(t)\Big|_{t=0}$$

exists and $q_x \in [0, \infty]$.

(iii) For all $x, y \in \mathcal{X}$, $y \neq x$ the derivative

$$q_{xy} = \frac{d}{dt_+} p_{xy}(t)\Big|_{t=0}$$

exists and $q_{xy} \in [0, \infty)$.

(iv) If $q_x < \infty$ then $p_{xy}(t)$ is continuously differentiable in $t \geq 0$ for that x and every $y \in \mathcal{X}$, and satisfies the Kolmogorov backward differential equation

$$\frac{d}{dt} p_{xy}(t) = \sum_{z\in\mathcal{X}} q_{xz}\, p_{zy}(t).$$

Proof. The proof can be found in [15], theorems 2.13 and 2.14. □

The number q_{xy} is called the *transition rates* from state x to state y. The number q_x is called the *total rate out of state* x. The larger the total rate is, the shorter time period the chain remains in state x. If $q_x = 0$, the chain remains in x permanently and we say that the state x is *absorbing*. The state x with $q_x < \infty$ is called *stable*. If $q_x = \infty$ the chain can not remain in x for positive amount of time and we say that the state x is *instantaneous*. However instantaneous states can occur only in case of infinite state space. In this thesis we work only with Markov chain with finite state space, so all states are stable. In this case

$$\sum_{y\neq x} q_{xy} = q_x. \tag{1.26}$$

Indeed

$$0 = \frac{1 - \sum_y p_{xy}(t)}{t} = \frac{1 - p_{xx}(t) - \sum_{y\neq x} p_{xy}(t)}{t} \longrightarrow q_x - \sum_{y\neq x} q_{xy}, \quad t \longrightarrow 0_+.$$

The relation (1.26) justifies the name total rate.

Define the matrix $\boldsymbol{Q}$ by $\boldsymbol{Q} = \{q_{xy}\}_{x,y\in\mathcal{X}}$. Because all states are stable according to Theorem 1.13 (iv) the *Kolmogorov backward differential equation* (1.27) must hold.

$$\frac{d}{dt} \boldsymbol{P}_t = \boldsymbol{Q}\, \boldsymbol{P}_t \tag{1.27}$$

From the properties of matrix exponential follows that under initial condition $\boldsymbol{P}_0 = \boldsymbol{I}$ the equation (1.27) has the unique solution of the form

$$\boldsymbol{P}_t = \exp\{t\,\boldsymbol{Q}\}. \tag{1.28}$$

So all transition probabilities can be expressed by means of the matrix $\boldsymbol{Q}$. In this sense $\boldsymbol{Q}$ is a continuous time counter-part of the one step probability matrix. The matrix $\boldsymbol{Q}$ is called the *infinitesimal generator* or simply the *generator matrix*.

Now we add rewards to the process. Unlike in the case of a discrete time Markov chain, here the chain stays in every state generally different amount of time. Thus we need to evaluate separately staying in a particular state and transition from one state to another. Reward for staying in a particular state will be described by vector $\boldsymbol{r}$, where r_i is a reward paid for being in state i for one unit of time. Reward for a transition from one state to another will be described by matrix $\boldsymbol{R}$, meaning that moving from state i to state j pays a reward r_{ij}. Because the transition from state i to state i simply means staying in state i and the corresponding reward is included in the vector $\boldsymbol{r}$, we assume that $r_{ii} = 0$. A continuous time Markov chain with the vector $\boldsymbol{r}$ and the matrix $\boldsymbol{R}$ is called *a continuous time Markov reward chain.*

Consider a continuous time Markov reward chain X with the state space $\mathcal{X} = \{1, \dots, N\}$. We try to do analogue of preceding sections in the continuous time set-up. For a given vector $\boldsymbol{b} \in \mathbb{R}^N$, similarly to (1.6) and (1.13), define

$$U_{s,t}(\boldsymbol{b}) = \mathcal{U}^{\mathsf{C}}_{\gamma}\Big(\int_s^t r(X_v)\,dv + \sum_{s<v\le t} r(X_{v-}, X_v)\Big) \cdot \boldsymbol{b}(X_t) \tag{1.29}$$

and put $U_{0,t}(\boldsymbol{b}) = U_t(\boldsymbol{b})$. Note that the number of $v \le t$, for which $X_{v-} \ne X_v$ is almost surely finite. Thus the sum on the right hand side is almost surely well defined. Again $U_t = U_t(\mathbf{1})$ is the total utility from the reward over the time interval $[0, t]$. The aim is to derive a formula for the conditional expectation of $U_t(\boldsymbol{b})$ similar to the one from Proposition 1.6.

$$\begin{aligned} \mathbb{E}[U_t(\boldsymbol{b})|\mathcal{F}_s] &\overset{\text{as}}{=} U_s \cdot \mathbb{E}\Big[-\mathcal{U}^{\mathsf{C}}\Big(\int_s^t r(X_v)\,dv + \sum_{s<v\le t} r(X_{v-}, X_v)\Big) \cdot \boldsymbol{b}(X_t)\Big|\mathcal{F}_s\Big] \\ &\overset{\text{as}}{=} U_s \cdot \mathbb{E}[-U_{s,t}(\boldsymbol{b})|\mathcal{F}_s] \end{aligned} \tag{1.30}$$

The idea is to approximate the continuous time Markov chain between s and t by the discrete time process obtained by observing the original one in discrete time instances. Such a process turns out be again a Markov chain and we can use Lemma 1.7 to express the second factor of the equation (1.30).

Denote $t_{n,k} = s + \frac{k}{n}(t-s)$ for $k = 0, \dots, n$ and consider the partition $\Delta_n = \{t_{n,k}\}_{k=0}^n$ of the interval $[s, t]$. Note that the norm of the partition Δ_n is equal to $\frac{t-s}{n}$. Define the process

$$X^n_v = X_{\lfloor v \rfloor_{\Delta_n}} \quad v \in [s, t],$$

where $\lfloor v \rfloor_{\Delta_n}$ means the value v rounded down with respect to the set Δ_n. The process X^n has piecewise constant trajectories with jumps in points of partition Δ_n.

Key properties of the constructed process are that the Markov property preserves and X^n approximates well the original process X.

Lemma 1.14. *(i) The process $\{X_{t_{n,k}}\}_{k=0}^n$ is a Markov chain with filtration $\{\mathcal{F}_{t_{n,k}}\}_{k=0}^n$ and the transition matrix $\boldsymbol{P}_{\frac{t-s}{n}}$.*

(ii) Considering the processes $X^n = (X^n_v, v \in [s,t])$ and $X = (X_v, v \in [s,t])$ as random variables with values in the metric space $D[s,t]$ with Skorokhod metric, X^n converges to X almost surely.

Proof. (i) All three required conditions from the Definition 1.3 follow from their continuous-time couterparts from the Defintion 1.12. First two are obvious. The third one is shown by following computation. For any $0 \le k < l \le n$

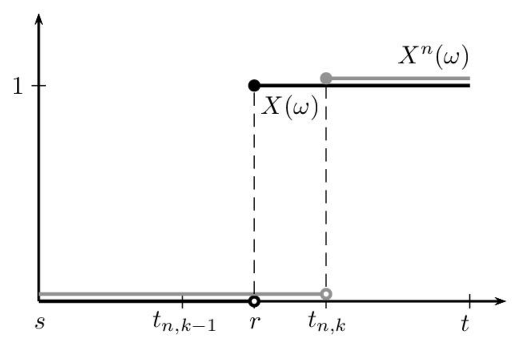
$$\begin{aligned}\mathbb{P}(X_{t_{n,k}} = j|\mathcal{F}_{t_{n,l}}) &= \mathbb{P}\left[X(s + \tfrac{k}{n}(t-s)) = j|\mathcal{F}(s + \tfrac{l}{n}(t-s))\right] \\ &= p_{X_{t_{n,l}},j}(\tfrac{k-l}{n}t - s) = \widetilde{p}^{(k-l)}_{X_{t_{n,l}},j},\end{aligned} \tag{1.31}$$

where $\widetilde{p}^{(k-l)}_{i,j}$ is the entry of the $(k-l)$-th power of the matrix $\boldsymbol{P}_{\frac{t-s}{n}}$ on i-th row and j-th column. Note that the probability measures $\widetilde{\mathbb{P}}_i$ can be defined by putting $\widetilde{\mathbb{P}}_i(X_{t_{n,0}} = i) = 1$. However the specific choice of measure does not influence the computation (1.31).

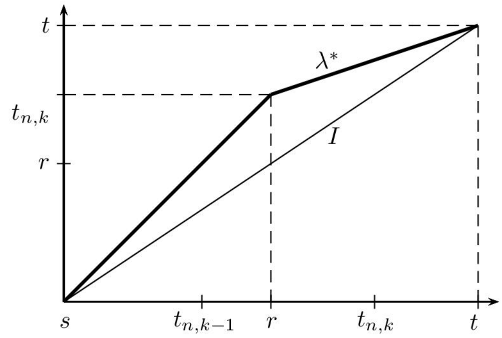
(ii) Remind that the Skorokhod metric is given by

$$\rho(f,g) = \inf_{\lambda \in \Lambda} \Big(\|\lambda - I\| \vee \|f - g \circ \lambda\| \Big),$$

where $\|\cdot\|$ is the uniform norm on $[s,t]$ and Λ is the set of all strictly increasing continuous bijection on $[s,t]$. We want to show that $\rho(X^n, X)$ almost surely goes to zero.



(a) Trajectories



(b) Time transformation

Figure 1.1: Reasoning for $X^n(\omega) \circ \lambda^* = X(\omega)$

For almost every $\omega \in \Omega$ trajectory $X(\omega)$ has only finite number jumps in the interval $[s,t]$. Choose one such ω. Because the state space is finite, size of each jump is finite. For clarity we will consider a trajectory $X(\omega)$ with only one jump of unit size. The general case with finite number of jumps is left as an exercise for the reader. So assume that $X(\omega)$ is of the form

$$X(\omega, v) = \begin{cases} 0 & \text{if } v \in [s, r) \\ 1 & \text{if } v \in [r, t]. \end{cases}$$

For every partition Δ_n there exists k for which $r \in (t_{n,k-1}, t_{n,k}]$. As we can see from Figure 1.1(a) the only difference between trajectories $X^n(\omega)$ and $X(\omega)$ is that $X^n(\omega)$ is shifted right by $t_{n,k} - r$. If we speed up the time appropriately on the interval $[0, r]$, the trajectory $X^n(\omega)$ would be equal to $X(\omega)$. Appropriate time transformation $\lambda^* \in \Lambda$ could be piecewise linear on intervals $[s, r]$, $[r, t]$ satisfying $\lambda^*(r) = t_{n,k}$. The function is plotted on Figure 1.1(b). Then evidently $X^n(\omega) \circ \lambda^* = X(\omega)$ and thus

$$\rho(X(\omega), X^n(\omega)) \le \|\lambda^* - I\| \vee \|X(\omega) - X^n(\omega) \circ \lambda^*\| = \|\lambda^* - I\| \le \|\Delta_n\|.$$

The result follows from the fact that the norm of the partition Δ_n goes to zero as n goes to infinity. □

Like at the end of the previous section Section 1.2, define the matrix $\widetilde{\boldsymbol{R}}$ with entries $\widetilde{r}_{ij} = r_i$. When we replace the process X by process the X^n in equation (1.30), we can express the second factor on the right hand side as

$$\begin{aligned}
\mathbb{E}[-U^n_{s,t}(\boldsymbol{b})|\mathcal{F}_s] &\overset{\text{as}}{=} \mathbb{E}\Big[-\mathcal{U}^{\mathsf{c}}_\gamma\Big(\int_s^t r(X^n_v)\,dv + \sum_{s<v\le t} r(X^n_{v-}, X^n_v)\Big)\cdot \boldsymbol{b}(X_t)\Big|\mathcal{F}_s\Big] \\
&\overset{\text{as}}{=} \mathbb{E}\Big[-\mathcal{U}^{\mathsf{c}}_\gamma\Big(\sum_{k=1}^n \tfrac{t-s}{n}\, r(X_{t_{n,k-1}}) + \sum_{k=1}^n r(X_{t_{n,k}-}, X_{t_{n,k}})\Big)\cdot \boldsymbol{b}(X_t)\Big|\mathcal{F}_s\Big] \\
&\overset{\text{as}}{=} \mathbb{E}\Big[-\mathcal{U}^{\mathsf{c}}_\gamma\Big(\sum_{k=1}^n \tfrac{t-s}{n}\widetilde{r}(X_{t_{n,k-1}}, X_{t_{n,k}}) + r(X_{t_{n,k-1}}, X_{t_{n,k}})\Big)\cdot \boldsymbol{b}(X_{t_{n,n}})\Big|\mathcal{F}_s\Big].
\end{aligned} \tag{1.32}$$

Using Lemma 1.7 we can further simplify this term.

$$\begin{aligned}
\mathbb{E}[-U^n_{s,t}(\boldsymbol{b})|\mathcal{F}_s] &\overset{\text{as}}{=} \Big(\Big[\boldsymbol{P}_{\frac{t-s}{n}} * \mathbf{exp}\{\gamma\, \tfrac{t-s}{n}\, \widetilde{\boldsymbol{R}}\} * \mathbf{exp}\{\gamma\, \boldsymbol{R}\}\Big]^n \boldsymbol{b}\Big)(X_s) \\
&= \Big(\Big[\big(\exp\{\tfrac{t-s}{n}\,\gamma\,\mathrm{diag}(\boldsymbol{r})\}\cdot \exp\{\tfrac{t-s}{n}\,\boldsymbol{Q}\}\big) * \mathbf{exp}\{\gamma\, \boldsymbol{R}\}\Big]^n \boldsymbol{b}\Big)(X_s).
\end{aligned} \tag{1.33}$$

Using the Taylor expansion, it can be shown that last term converges to a finite limit. In spite of the simplicity of idea behind the proof, the proof itself is quite technical and tedious. Thus we move this computation to Appendix C. Now we can make the following conclusion.

Proposition 1.15. *Let $0 \le s \le s+h$. Then for any given vector $\boldsymbol{b} \in \mathbb{R}^N$ and the variable $U_{s+h}(\boldsymbol{b})$ defined by (1.29) we have*

$$\mathbb{E}[U_{s+h}(\boldsymbol{b})|\mathcal{F}_s] = U_s(\boldsymbol{S}^h\, \boldsymbol{b}),$$

where $\boldsymbol{S}^h$ *has non-negative entries and is of the form*

$$\boldsymbol{S}^h = \exp\{h\,\boldsymbol{T}\}, \quad h \geq 0, \tag{1.34}$$

where

$$\boldsymbol{T} = (\boldsymbol{Q} + \gamma\,\mathrm{diag}(\boldsymbol{r})) * \mathbf{exp}\{\gamma\,\boldsymbol{R}\}. \tag{1.35}$$

Proof. Define the functional f on $D[s,t]$ by

$$f(y) = -\mathcal{U}_\gamma^{\mathrm{C}}\Big(\int_s^{s+h} r(y_v)\,dv + \sum_{s<v\leq s+h} r(y_{v-}, y_v)\Big) \cdot \boldsymbol{b}(y_t).$$

Because f is continuous, according to Lemma 1.14 $f(X^n)$ converges to the $f(X)$ almost surely. Moreover the sequence $f(X^n)$ is uniformly bounded, because the state space is finite. Thus we also have convergence in L^1 which implies convergence of conditional expectations. Now we can complete computation (1.33) by taking the limit. The matrix $\boldsymbol{R}$ has zeros on its main diagonal. So according to the Lemma C.2

$$\begin{aligned}\lim_{n\to\infty}\Big[\exp\{\tfrac{h}{n}\,\gamma\,\mathrm{diag}(\boldsymbol{r})\} \cdot \exp\{\tfrac{h}{n}\,\boldsymbol{Q}\} * \mathbf{exp}\{\gamma\,\boldsymbol{R}\}\Big]^n \\ = \exp\{(h)(\boldsymbol{Q} + \gamma\mathrm{diag}(\boldsymbol{r})) * \mathbf{exp}\{\gamma\boldsymbol{R}\}\} = \boldsymbol{S}^h.\end{aligned}$$

The left hand side of the above term can be also expressed as

$$\lim_{n\to\infty}\Big[\boldsymbol{P}_{\frac{h}{n}} * \mathbf{exp}\{\gamma\,\tfrac{h}{n}\,\widetilde{\boldsymbol{R}}\} * \mathbf{exp}\{\gamma\,\boldsymbol{R}\}\Big]^n.$$

So $\boldsymbol{S}^h$, $h \geq 0$ is a limit of non-negative matrices and thus it is itself non-negative. □

1.5 Optimal risk sensitive control of continuous time Markov decision chain

Similarly to the discrete time set-up, also here we add decisions to the process. Consider an action space

$$\mathcal{A} = \prod_{x\in\mathcal{X}} A_x,$$

with A_x representing the set of all admissible action in state x. For any policy $a \in \mathcal{A}$ let $\boldsymbol{Q}^a$ be an intensity matrix and $\boldsymbol{R}^a, \boldsymbol{r}^a$ be a reward matrix and a reward vector corresponding to the policy a. That is for every policy $a \in \mathcal{A}$ we have a continuous time Markov reward chain X^a. The whole system $(\Omega, \{\mathcal{F}_n\}_{n\in\mathbb{N}_0}, \{\mathrm{X}^a\}_{a\in\mathcal{A}})$ is called *a continuous time Markov decision chain.*

Consider a Markov decision chain with state space $\mathcal{X} = \{1, \ldots, N\}$, intensity matrices

$$\boldsymbol{Q}^a = (\boldsymbol{q}_1(a_1), \ldots, \boldsymbol{q}_N(a_N))^{\mathrm{T}},$$

reward matrices and reward vectors

$$\boldsymbol{R}^a = (\boldsymbol{r}_1(a_1), \ldots, \boldsymbol{r}_N(a_N))^{\mathsf{T}}, \quad \boldsymbol{r}^a = (r_1(a_1), \ldots, r_N(a_N))^{\mathsf{T}}.$$

We assume that for all policies $a \in \mathcal{A}$ the chain is aperiodic and irreducibile. For any policy a we have the variable $U_n^a(\cdot)$ defined by (1.29) and the non-negative matrix $\boldsymbol{S}^a = \exp\{\boldsymbol{T}^a\}$ defined by (1.34) and (1.35). That is

$$\boldsymbol{T}^a = (\boldsymbol{t}_1(a_1), \ldots, \boldsymbol{t}_N(a_N))^{\mathsf{T}}, \quad \boldsymbol{t}_i(a_i) = \{t_{ij}(a_i)\}_{j=1}^N,$$
$$t_{ij}(a_i) = (q_{ij}(a_i) + \gamma\, \delta_{ij}\, r_i(a_i))\, e^{\gamma r_{ij}(a_i)},$$

where δ_{ij} is the Kronecker delta. By virtue of the Perron-Frobenius theorem A.3 the matrix $\boldsymbol{S}^a$ has its maximal eigenvalue $\lambda^a > 0$ and respective eigenvector $\boldsymbol{v}^a > 0$. In addition, according to the results of Appendix B, $\boldsymbol{v}^a$ is also an eigenvector of $\boldsymbol{T}^a$ coressponding to its maximal eigenvalue $\kappa^a \triangleq \log \lambda^a$.

For fixed any $a \in \mathcal{A}$ consider a process

$$M_t^a = (\lambda^a)^{-t} U_t^a(\boldsymbol{v}^a) \quad t \geq 0. \tag{1.36}$$

Proposition 1.16. *The process M_t^a is a $\mathcal{F}_t$-supermartingale if $\boldsymbol{T}^a\, \boldsymbol{v}^a \geq \kappa^a\, \boldsymbol{v}^a$. The process M_t^a is a $\mathcal{F}_t$-martingale if $\boldsymbol{T}^a\, \boldsymbol{v}^a = \kappa^a\, \boldsymbol{v}^a$.*

Proof. As the Proposition concerns only one fixed policy, we will omit the upper index a throughout the proof. First we want to show that the assumption $\boldsymbol{T}\,\boldsymbol{v} = \kappa\,\boldsymbol{v}$ implies the inequality

$$\exp\{s\,\boldsymbol{T}\}\,\boldsymbol{v} = \boldsymbol{S}^s\,\boldsymbol{v} \geq \lambda^s\,\boldsymbol{v} = e^{s\,\kappa}\,\boldsymbol{v} \quad \text{for all } s \geq 0.$$

Using the fact $\exp\{-s\kappa\,\boldsymbol{I}\} = e^{-sk}\,\boldsymbol{I}$ we compute

$$\begin{aligned}\boldsymbol{S}^s\,\boldsymbol{v} - \lambda^s\,\boldsymbol{v} &= \exp\{s\,\boldsymbol{T}\}\,\boldsymbol{v} - e^{s\,\kappa}\,\boldsymbol{v} \\ &= e^{s\,\kappa}\,(e^{-s\,\kappa}\,\exp\{s\,\boldsymbol{T}\}\,\boldsymbol{v} - \boldsymbol{v}) \\ &= e^{s\,\kappa}\,(\exp\{s(\underbrace{\boldsymbol{T} - \kappa\,\boldsymbol{I}}_{\triangleq\widetilde{T}})\} - \boldsymbol{I})\,\boldsymbol{v}.\end{aligned}$$

The last equality is true due to the fact that matrices $\boldsymbol{T}$ and $\kappa\,\boldsymbol{I}$ commute. Denote $\widetilde{\boldsymbol{T}} \triangleq \boldsymbol{T} - \kappa\,\boldsymbol{I}$. Then by assumption $\widetilde{\boldsymbol{T}}\,\boldsymbol{v} \geq 0$. We continue with computation as follows

$$\begin{aligned}\boldsymbol{S}^s\,\boldsymbol{v} - \lambda^s\,\boldsymbol{v} &= e^{s\,\kappa}\,(\exp\{s\widetilde{\boldsymbol{T}}\} - \boldsymbol{I})\,\boldsymbol{v} \\ &= e^{s\,\kappa}\left(\int_0^s \frac{\mathrm{d}}{\mathrm{d}u}\exp\{u\widetilde{\boldsymbol{T}}\}\mathrm{d}u\right)\boldsymbol{v} = \underbrace{e^{s\,\kappa}}_{\geq 0}\left(\int_0^s \exp\{u\widetilde{\boldsymbol{T}}\}\mathrm{d}u\right)\underbrace{(\widetilde{\boldsymbol{T}}\boldsymbol{v})}_{\geq 0}.\end{aligned}$$

The integrant $\exp\{u\widetilde{\boldsymbol{T}}\} = e^{-u\,\kappa}\,\boldsymbol{S}^u$ is also non-negative according to Proposition 1.15. Thus the whole term is nonnegative.

We have shown that $\boldsymbol{S}^s\,\boldsymbol{v} \geq \lambda^s\,\boldsymbol{v}$, $s \geq 0$. Employing the Proposition 1.15 and the fact that M_s is $\mathcal{F}_s$ measurable we get

$$\begin{aligned}\mathbb{E}[M_{s+h} - M_s|\mathcal{F}_s] &= \lambda^{-(s+h)}\,U_s(\boldsymbol{S}^h\boldsymbol{v}) - \lambda^{-s}\,U_s(\boldsymbol{v})\\ &= \lambda^{-s-h}[U_s(\boldsymbol{S}^h\boldsymbol{v}) - \lambda^h\,U_s(\boldsymbol{v})]\\ &= \underbrace{\lambda^{-s-h}}_{>0}\,\underbrace{U_s}_{<0}\cdot[(\boldsymbol{S}^h\,\boldsymbol{v})(X_s) - (\lambda^s\,\boldsymbol{v})(X_s)].\end{aligned}$$

Because λ is positive and U_s is negative we have

$$\lambda^s\,\boldsymbol{v} \leq \boldsymbol{S}^s\,\widehat{\boldsymbol{v}} \quad\Longrightarrow\quad \mathbb{E}[M_{s+h}|\mathcal{F}_s] \overset{\text{as}}{\leq} M_s, \quad 0 \leq s \leq s+h. \qquad \square$$

Theorem 1.17. *If the inequality $\boldsymbol{T}^a\,\boldsymbol{v}^{\widehat{a}} \geq \kappa^{\widehat{a}}\,\boldsymbol{v}^{\widehat{a}}$ holds for every policy $a \in \mathcal{A}$, then $\widehat{a}$ is an optimal policy.*

Proof. In the proof of the Lemma 1.16 we show that the assumption implies

$$\boldsymbol{S}^a\,\boldsymbol{v}^{\widehat{a}} \geq \lambda^{\widehat{a}}\,\boldsymbol{v}^{\widehat{a}} \quad \text{for all } a \in \mathcal{A}.$$

The rest is a direct analogue of the proof of the Theorem 2.4. $\square$

Theorem 1.18 (Policy iteration)**.** *Let $a_0 \in \mathcal{A}$ be the initial policy. Let $a_0 \in \mathcal{A}$ be the initial policy. Define sequence $\{a_n\}$ recursively by*

$$a_{n+1}(i) = \operatorname*{argmin}_{a\in\mathcal{A}} \boldsymbol{t}_i^a\,\boldsymbol{v}^{a_n}. \tag{1.37}$$

If the minimum is attained for more the one policy and a_n is one of then, always make conservative choice $a_{n+1}(i) = a_n(i)$. Then a_n converges to an optimal policy $\widehat{a}$.

Proof. Suppose that a_n converges to $\widehat{a}$. Then by (1.37) for all i we have

$$\boldsymbol{e}_i^{\intercal}\,\boldsymbol{T}^a\,\boldsymbol{v}^{\widehat{a}} \geq \boldsymbol{e}_i^{\intercal}\,\boldsymbol{T}^{\widehat{a}}\,\boldsymbol{v}^{\widehat{a}} = \kappa^{\widehat{a}}\,v_i^{\widehat{a}} \quad a \in \mathcal{A}$$

Thus the inequality $\boldsymbol{T}^a\,\boldsymbol{v}^{\widehat{a}} \geq \kappa^{\widehat{a}}\,\boldsymbol{v}^{\widehat{a}}$ holds for every $a \in \mathcal{A}$ and the policy $\widehat{a}$ is optimal by virtue of Theorem 1.10. The rest is a direct analogue of the proof of the Theorem 1.11. $\square$

The Theorem 1.18 gives us a method for finding the optimal policy for continuous time Markov Chain. The algorithm works exactly as for discrete time case, which is described at the end of section 1.3. The only difference is that the matrix $\boldsymbol{S}$ is replaced by the matrix $\boldsymbol{T}$, which is given by (1.35).

2. Optimal Investment with Proportional Transaction Costs

The problem of optimal portfolio management of securities was first formulated by Merton [17]. It is known as Merton's porfolio problem or the consumption-investment problem. It concerns a question how to allocate wealth between consumption and investment in order to maximize expected utility over a time horizon. This general problem can be considered under numerous different specific formulations. For example, Merton derived analytical solution for the problem with two assets under logarithmic utility function over both finite and infinite horizon.

In this section we consider the problem with proportional transaction costs as formulated in [6]. We summarize the dynamics of the model and propose its approximation by Markov chain. Then we use the algorithm from the Chapter 1 to solve the problem. The results are compared with analytical solution.

2.1 Model description

Suppose that a market consists of two assets. The first one is assumed to be riskless and the second one is assumed to be risky. Time development of the riskless asset, denoted by S^0, is deterministic and is given by

$$\mathrm{d}S_t^0 = r\, S_t^0\, \mathrm{d}t. \tag{2.1}$$

This asset represents a bank account with constant interest rate r and with continuous compounding. Starting with a certain deposit the account grows exponentially in time with growth rate equal to r.

The second asset, denoted by S^1, represents a stock or a stock index. We assume that it's price follows geometric Brownian motion with drift μ and volatility σ. That is

$$\mathrm{d}S_t^1 = \mu\, S_t^1 \mathrm{d}t + \sigma\, S_t^1\, \mathrm{d}W_t, \tag{2.2}$$

where W_t is a Brownian motion. We assume that no other assets are available.

Further we assume that both assets, the stock and the money, are infinitely divisible. So the investor can possess any non-negative real volumes of these assets.

As mentioned in the introduction, the transaction costs are paid when trading the risky asset. These transaction costs are proportional to the size of the deal. In case of buying a $(1 + c_+)$ - multiple of the stock price is paid. On the other hand, in case of selling, $(1 - c_-)$ - multiple of the stock price is received. We consider $c_+ \in (0, \infty)$ and $c_- \in (0, 1)$.

We will consider a utility function with hyperbolic absolute risk aversion (HARA) of the form

$$\mathcal{U}_\gamma^{\mathtt{H}}(x) = \tfrac{1}{\gamma}\, x^\gamma, \quad \gamma < 0. \tag{2.3}$$

The aim of the investor is to maximize the growth rate of the certainty equivalent of the value of the portfolio. That is to maximize

$$\lim_{t\to\infty} t^{-1}\log(\mathtt{CE}^{\mathtt{H}}_{\gamma}(V_t)), \quad \mathtt{CE}^{\mathtt{H}}_{\gamma}(V_t) := (\mathcal{U}^{\mathtt{H}}_{\gamma})^{-1}\,\mathbb{E}\,\mathcal{U}^{\mathtt{H}}_{\gamma}(V_t), \tag{2.4}$$

where V_t is the value of the portfolio at time t.

There is an important relation (2.5) between HARA utility function and CARA utility function from the first Chapter, which can be easily shown by direct computation.

$$\mathtt{CE}^{\mathtt{C}}_{\gamma}(\log(\cdot)) = \log(\mathtt{CE}^{\mathtt{H}}_{\gamma}(\cdot)). \tag{2.5}$$

According to (2.5), the criteria (2.4) can be reformulated as

$$\lim_{t\to\infty} t^{-1}\log(\mathtt{CE}^{\mathtt{H}}_{\gamma}(V_t)) = \lim_{t\to\infty} t^{-1}\mathtt{CE}^{\mathtt{C}}_{\gamma}(\log(V_t)). \tag{2.6}$$

The relation (2.6) coverts the problem of maximizing (2.4) to the problem we considered in the chapter one. The main benefit of this conversion is that we can work with $\log(V_t)$ instead of V_t. As we will see further, $\log(V_t)$ behaves better for purpose of approximation.

Now we analytically derive the dynamics of the portfolio. First we look at the dynamics of the value of the portfolio investment. Denote the number of riskless and risky assets in the portfolio at time t by N^0_t and N^1_t respectively. The value of the portfolio at time t is given by

$$V_t = N^{\mathtt{T}}_t S_t = N^0_t\, S^0_t + N^1_t\, S^1_t.$$

If the investor does not trade, number of both assets N^0_t, N^1_t remains constant. Then we can write

$$\begin{aligned}
\mathrm{d}V_t &= N^0_t\,\mathrm{d}S^0_t + N^1_t\mathrm{d}S^1_t \\
&= N^0_t\, r\, S^0_t\mathrm{d}t + N^1_t(\,\mu\, S^1_t\mathrm{d}t + \sigma\, N^0_t\, r\, S^1_t\mathrm{d}W_t) \\
&= r\, V_t\,\frac{N^0_t\, S^0_t}{N^{\mathtt{T}}_t S_t}\,\mathrm{d}t + \mu\, V_t\,\frac{N^1_t\, S^1_t}{N^{\mathtt{T}}_t S_t}\,\mathrm{d}t + \sigma\, V_t\,\frac{N^1_t\, S^1_t}{N^{\mathtt{T}}_t S_t}\,\mathrm{d}W_t.
\end{aligned}$$

Defining

$$G_t = \frac{N^1_t\, S^1_t}{N^0_t\, S^0_t + N^1_t\, S^1_t}, \tag{2.7}$$

we gain

$$\frac{\mathrm{d}V_t}{V_t} = [r + (\mu - r)\, G_t]\,\mathrm{d}t + \sigma\, G_t\,\mathrm{d}W_t. \tag{2.8}$$

Note that the variable G_t represents the portion of the investor's wealth that is hold in the risky asset. We would refer to this quantity as an investor's *position*. The process $(G_t, t \geq 0)$ attains only values from the interval $[0, 1]$. This is particularly important, because later we would like to approximate this process by a Markov chain with finite state space. This would be problematic in the case of unbounded state domain.

Using Ito's lemma we compute dynamics of the logarithm of the value of the portfolio. According to (2.8)

$$V_t^{-1}\mathrm{d}\,\langle V\rangle_t = \sigma^2\, G_t^2\,\mathrm{d}t,$$

and we have

$$\mathrm{d}\log V_t = \frac{\mathrm{d}V_t}{V_t} - \frac{\langle \mathrm{d}V\rangle_t}{2\,V_t^2} = [r + (\mu - r)\,G_t - \tfrac{1}{2}\sigma^2\,G_t^2]\,\mathrm{d}t + \sigma\,G_t\,\mathrm{d}W_t. \tag{2.9}$$

For technical details of the above computation see Appendix D. Denote

$$q_0(x) = r + (\mu - r)\,x - \tfrac{1}{2}\sigma^2\,x^2.$$

Then we can write

$$\mathrm{d}\log V_t = q_0(G_t)\,\mathrm{d}t + \sigma\,G_t\,\mathrm{d}W_t. \tag{2.10}$$

We managed to derive the dynamics of $\log V_t$ in terms of G_t. Now we would like to compute the dynamics of G_t. Using Ito's formula we gain

$$V_t\,\mathrm{d}V_t^{-1} = -\frac{\mathrm{d}V_t}{V_t} + \frac{(\mathrm{d}V_t)^2}{V_t^2} = [-r - (\mu - r)\,G_t + \sigma^2\,G_t^2]\,\mathrm{d}t - \sigma\,G_t\,\mathrm{d}W_t.$$

Once again, the details of the computation can be found in Appendix D. Realizing that $G_t = N_t^1\,S_t^1\,V_t^{-1}$, we compute

$$\begin{aligned}
\mathrm{d}G_t &= N_t^1\,V_t^{-1}\,\mathrm{d}S_t^1 + N_t^1\,S_t^1\,\mathrm{d}V_t^{-1} + N_t^1\,(\mathrm{d}S_t^1)\,(\mathrm{d}V_t^{-1})\\
&= N_t^1\,V_t^{-1}\,(\mu\,S_t^1\,\mathrm{d}t + \sigma\,S + t^1\mathrm{d}W_t) + N_t^1\,S_t^1\,V_t^{-1}[(-r(\mu - r)\,G_t + \sigma^2\,G_t^2)\mathrm{d}t\\
&\quad - \sigma^2\,G_t\,\mathrm{d}W_t] - N_t^1\,V_t^{-1}\,\sigma^2\,G_t\,S_t^1\,\mathrm{d}t\\
&= G_t(\mu\mathrm{d}t + \sigma\mathrm{d}W_t) + G_t[(-r - (\mu - r)\,G_t + \sigma^2\,G_t^2)\,\mathrm{d}t - \sigma\,G_t\,\mathrm{d}W_t] - \sigma^2\,G_t^2\,\mathrm{d}t\\
&= G_t(\mu - r - (\mu - r)\,G_t - \sigma^2\,G_t^2 - \sigma^2\,G_t)\,\mathrm{d}t + G_t\,(\sigma - \sigma\,G_t)\,\mathrm{d}W_t\\
&= G_t\,(1 - G_t)[(\mu - r - \sigma^2\,G_t)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t].
\end{aligned}$$

Define functions describing drift and volatility

$$b(x) = x\,(1 - x)\,(\mu - r - \sigma^2\,x),$$

$$s(x) = \sigma\,x\,(1 - x).$$

Now we can express the dynamics of G_t as

$$\mathrm{d}G_t = b(G_t)\mathrm{d}t + s(G_t)\mathrm{d}W_t. \tag{2.11}$$

In order to add trading to the dynamics, denote the sum of stocks bought and sold on the interval $[0, t)$ by N_t^+ and N_t^- respectively. So the total number of stocks

N_t^1 is equal to $N_t^+ - N_t^-$. Using Ito's lemma we can compute the dynamics of G_t and $\log(V_t)$. The technical details are similar to the computations above. Here we only state the resulting differentials. For G_t we get

$$\mathrm{d}G_t = b(G_t)\mathrm{d}t + s(G_t)\mathrm{d}W_t + \mathrm{d}^+G_t - \mathrm{d}^-G_t, \tag{2.12}$$

where

$$\mathrm{d}^+G_t = \frac{(1 + c_+ G_t) S_t}{V_t} \mathrm{d}N_t^+, \quad \mathrm{d}^-G_t = \frac{(1 - c_- G_t) S_t}{V_t} \mathrm{d}N_t^-.$$

For $\log(V_t)$ we get

$$\mathrm{d}\log V_t = q_0(G_t)\,\mathrm{d}t + \sigma\, G_t\,\mathrm{d}W_t - \nu_+(G_t)\mathrm{d}^+G_t - \nu_-(G_t)\mathrm{d}^-G_t, \tag{2.13}$$

where

$$\nu_+(x) = \frac{c_+}{1 + c_+ x}, \quad \nu_-(x) = \frac{c_-}{1 + c_- x}. \tag{2.14}$$

Finally we look at dynamics of the utility $\mathcal{U}_\gamma^{\mathtt{H}}(V_t)$, which should determine the rewards for our approximating chain. For sake of simplicity we will consider the dynamics without trading. First remind that

$$\mathcal{U}_\gamma^{\mathtt{H}}(x) = \tfrac{1}{\gamma}\, x^\gamma = \gamma^{-1} \exp\{\gamma \log(V_t)\}. \tag{2.15}$$

Using Ito's Lemma with $f(x) = \exp\{\gamma x\}$ and (2.10) we get

$$\begin{aligned}
\mathrm{d}\gamma^{-1} \exp\{\gamma \log(V_t)\} &= V_t^\gamma\, \mathrm{d}\log(V_t) + \tfrac{1}{2}\gamma\, V_t^\gamma \mathrm{d}\,\langle \log(V)\rangle_t \\
&= (V_t^\gamma\, q_0(G_t) + \tfrac{1}{2}\gamma\sigma^2 G_t^2)\mathrm{d}t + V_t^\gamma \sigma \mathrm{d}W_t \\
&= V_t^\gamma\, q_\gamma(G_t)\mathrm{d}t + V_t^\gamma \sigma \mathrm{d}W_t,
\end{aligned} \tag{2.16}$$

where

$$q_\gamma(x) = q_0(x) + \tfrac{1}{2}\gamma\sigma^2 x^2.$$

Thus the expected utility of the value of the portfolio is

$$\mathbb{E}\,\mathcal{U}_\gamma(V_t) = \mathbb{E} \int_0^t V_s^\gamma\, q_\gamma(G_s)\mathrm{d}s.$$

The process G is the one we would like to approximate by Markov chain. However, in order to determine the rewards for G, we need the expected utility $\mathbb{E}\,\mathcal{U}_\gamma(V_t)$ to be dependent only on G_t. The Girsanov theorem (see Appendix D) can help us to get rid of dependence on V_t. Assume without loss of generality that $V_0 = 1$. Than by equation (2.10) we can express $\mathcal{U}_\gamma(V_t)$ as

$$\begin{aligned}
\mathcal{U}_\gamma(V_t) &= \tfrac{1}{\gamma} \exp\{\gamma \log(V_t)\} \\
&= \gamma^{-1} \exp\left\{\gamma \int_0^t q_0(G_s)\mathrm{d}s + \sigma\gamma \int_0^t G_s \mathrm{d}W_s\right\} \\
&= \gamma^{-1} \exp\left\{\sigma\gamma \int_0^t G_s \mathrm{d}W_s - \tfrac{1}{2}\sigma^2\gamma \int_0^t G_s^2 \mathrm{d}s\right\} \exp\left\{\gamma \int_0^t q_\gamma(G_s)\mathrm{d}s\right\}.
\end{aligned}$$

Defining the stochastic exponential

$$\mathcal{E}(X)_t = \exp\{X_t - \tfrac{1}{2}\langle X\rangle_t\}, \quad X_t = \sigma\gamma\int_0^t G_s \mathrm{d}W_s,$$

we can write

$$\mathcal{U}_\gamma(V_t) = \gamma^{-1}\,\mathcal{E}(X)_t\,\exp\Big\{\gamma\int_0^t q_\gamma(G_s)\mathrm{d}s\Big\}. \tag{2.17}$$

Since G_t attains only values in $[0,1]$, it satisfies Novikov condition,

$$\mathbb{E}\left[\exp\Big(\frac{1}{2}\int_0^t G_s^2\,\mathrm{d}s\Big)\right] < e^{\frac{t}{2}} < \infty, \quad t \geq 0. \tag{2.18}$$

Thus, we can use the Girsanov theorem D.2 which gives us existence of the measure $\mathbb{Q}_t$, absolutely continuous to the underlying measure $\mathbb{P}$, for which

$$\begin{aligned}\mathbb{E}_\mathbb{P}\,\mathcal{U}_\gamma(V_t) &= \mathbb{E}_\mathbb{P}\,\mathcal{E}(X)_t\exp\Big\{\gamma\int_0^t q_\gamma(G_s)\mathrm{d}s\Big\}\\ &= \mathbb{E}_{\mathbb{Q}_t}\,\exp\Big\{\gamma\int_0^t q_\gamma(G_s)\mathrm{d}s\Big\}.\end{aligned} \tag{2.19}$$

We see that by moving to the measure $\mathbb{Q}_t$ we get rid of dependence on V. The only thing that remains to derive is the dynamics of G under the measure $\mathbb{Q}_t$. According to D.2

$$\mathrm{d}W_s = \mathrm{d}\widetilde{W}_s + \langle W, X\rangle_s = \mathrm{d}\widetilde{W}_s + \gamma\sigma G_t\mathrm{d}t, \tag{2.20}$$

where $\widetilde{W}_s$ is a standard Brownian motion under $\mathbb{Q}_t$ on the interval $[0,t]$. Putting (2.12) and (2.20) together we get

$$\begin{aligned}\mathrm{d}G_t &= b(G_t)\mathrm{d}t + s(G_t)\mathrm{d}W_t + \mathrm{d}^+G_t - \mathrm{d}^-G_t\\ &= b(G_t)\mathrm{d}t + s(G_t)(\mathrm{d}\widetilde{W}_t + \gamma\sigma G_t\mathrm{d}t) + \mathrm{d}^+G_t - \mathrm{d}^-G_t\end{aligned}$$

So the dynamics of G_t under the $\mathbb{Q}$ is

$$\mathrm{d}G_t = \widetilde{b}(G_t)\mathrm{d}t + s(G_t)\mathrm{d}\widetilde{W}_t + \mathrm{d}^+G_t - \mathrm{d}^-G_t, \tag{2.21}$$

where

$$\begin{aligned}\widetilde{b}(x) &= b(x) + s(x)\,\gamma\,\sigma\,x\\ &= (1-\gamma)\,\sigma^2\,x\,(1-x)\,[\tfrac{(\mu-r)\,\sigma^2}{1-\gamma} - x].\end{aligned} \tag{2.22}$$

2.2 Analytical Solution

Dostál [6] derived analytical solution for the problem of maximizing (2.4). The optimal strategy is not to trade if the position G_t is in certain interval $[\alpha, \beta]$, and to buy or sell the stock in order to keep the position within the interval $[\alpha, \beta]$. Dostál shows that, considering a *interval strategy* $[\alpha, \beta]$, the certainty equivalent growth rate can be computed as

$$\lim_{t\to\infty} t^{-1} \log(\mathtt{CE}^{\mathtt{H}}(V_t)) = \tfrac{\sigma^2}{2} u(\alpha, \beta), \tag{2.23}$$

where the function $u(\alpha, \beta)$ is defined as the unique solution to the equation

$$\log \frac{1/\alpha - 1}{1/\beta - 1} = \int_{\xi_-(\beta)}^{\xi_+(\alpha)} \frac{dz}{\gamma z^2 + 2\rho z - u(\alpha, \beta)}, \tag{2.24}$$

where

$$\rho = \frac{\mu}{\sigma^2} - \frac{1}{2}, \qquad \xi_-(x) = x\frac{1 - c_-}{1 - c_- x}, \qquad \xi_+(x) = x\frac{1 + c_+}{1 + c_+ x}.$$

The optimal interval strategy can be found by maximizing the function $u(\alpha, \beta)$ over $\{(\alpha, \beta) : \alpha, \beta \in (0, 1), \alpha < \beta\}$. Unfortunately, the equation 2.24 do not have a closed-form solution, so the optimization must be carried out numerically.

2.3 Discrete Time Approximation

We will approximate the continuous time model from previous section by discrete time Markov chain. Methods we use are based on monograph [13]. For further detail see chapters 4 and 5 in this monograph. The key requirement for approximating Markov chain is that it must preserve the basic characteristics of original process, namely conditional mean and conditional variance.

As we mentioned earlier we will approximate the position process G under the measure $\mathbb{Q}$, which is given by the equation (2.21). First we look at the model without trading. Investor's trading decisions or, in terminology of the first chapter, investor's actions, will be added later. Approximation in both time and state domain are needed. In order to do this we create a lattice on which the approximating chain will be allowed to move. Let us start with time domain. Let $0 = t_0 < t_1 < t_2 \ldots$ be the sequence with step $t_k - t_{k-1} = \Delta t$, which is the same for all k. The change of characteristics of the original process in time step Δt can be approximated according to (2.21) as follows

$$\mathbb{E}[G_{t_k+1} - G_{t_k} | G_{t_k} = g] \sim \widetilde{b}(g)\Delta t, \tag{2.25}$$

$$\operatorname{var}[G_{t_k+1} - G_{t_k} | G_{t_k} = g] \sim \mathbb{E}[(G_{t_k+1} - G_{t_k})^2 | G_{t_k} = g] \sim s^2(g)\Delta t. \tag{2.26}$$

In case of conditional variance we neglect the second order term

$$(\mathbb{E}[(G_{t_k+1} - G_{t_k}) | G_{t_k} = g])^2 \sim (\widetilde{b}(g))^2\, (\Delta t)^2,$$

which is acceptable for small enough Δt.

Now we move to approximation in state domain of G_t, which is the interval $[0,1]$. Let $\{g_i\}_{i=1}^n$ be the equidistant partition of the interval $[0,1]$ with the norm equal to $\delta = \frac{1}{n}$. The set $\{g_i\}_{i=1}^n$ will be the state space of the approximating Markov chain. We denote this chain by $\widehat{G} = (\widehat{G}_k, k \in \mathbb{N})$. Suppose the $\widehat{G}$ is in state g. Denote the probability that it moves to the state $g+h$ by $P_+(g)$ and the probability that it moves to the state $g-h$ by $P_-(g)$. The probability that the chain stays in g in the next step is denoted $P_0(g)$. We will not allow probability of moving to the other states to be positive. Of course these probabilities must satisfy

$$P_0(g) = 1 - P_+(g) - P_-(g).$$

First we need to compute moment characteristics of the chain in term of these probabilities. Then we can determine values $P_+(g)$, $P_-(g)$ and $P_0(g)$ by comparison with characteristics of the original process. We obtain

$$\begin{aligned}\mathbb{E}[\widehat{G}_{k+1}|\widehat{G}_k = g] &= P_-(g)\,(g-\delta) + P_0(g)\,g + P_+(g)\,(g+\delta)\\ &= g - P_-(g)\,\delta + P_+(g)\,\delta.\end{aligned}$$

Subtracting $\widehat{G}_k = g$ we gain

$$\mathbb{E}[\widehat{G}_{k+1} - \widehat{G}_k|\widehat{G}_k = g] = \delta(P_+(g) - P_-(g)), \tag{2.27}$$

$$\begin{aligned}\operatorname{var}[\widehat{G}_{k+1} - \widehat{G}_k|\widehat{G}_k = g] &\sim \mathbb{E}[(\widehat{G}_{k+1} - \widehat{G}_k)^2|\widehat{G}_k = g]\\ &\sim \delta^2(P_+(g) + P_-(g)).\end{aligned} \tag{2.28}$$

When we compare (2.25) with (2.27) and (2.26) with (2.28) we gain the following equations

$$\begin{aligned}b(g)\Delta t &= \delta(P_+(g) - P_-(g)),\\ s^2(g)\Delta t &= \delta^2(P_+(g) + P_-(g)).\end{aligned} \tag{2.29}$$

Solving the equations (2.29) we get

$$P_+(g) = \frac{\delta\,\widetilde{b}(g) + s^2(g)}{2\delta^2}\,\Delta t,$$

$$P_-(g) = \frac{-\delta\,\widetilde{b}(g)t + s^2(g)}{2\delta^2}\,\Delta t.$$

From (2.29) also immediately follows that

$$P_0(g) = 1 - P_+(g) - P_-(g) = 1 - \frac{s^2(g)\Delta t}{\delta^2}.$$

Nonnegativity of $P_0(g)$ gives us the constraint $s^2(g)\Delta t \le \delta^2$. We can choose the time step Δt such that

$$\Delta t = \frac{\delta^2}{k},$$

where $k \ge s^2(g)$ holds for every state g.

Now we add actions to the chain, meaning trading decisions. We consider three types of actions: action $+$ means to buy stocks, action $-$ means to sell stocks and action 0 means to do nothing. For the states close to extreme points, not all decisions are allowed. The set of all admissible actions depending on state A_g are as following:

- $A_0 = \{+\}$, $A_1 = \{-\}$: in the extreme points only one decision is allowed in order to push back inside the interval $(0, 1)$,
- $A_h = \{+, 0\}$, $A_{1-h} = \{0, -\}$: only decisions that keep the position inside the interval $(0, 1)$ are allowed,
- $A_g = \{+, 0, -\}$, $g = 2h, 3h \dots, 1-2h$: in the rest of the states all alternatives are possible.

As we derived above, in case of action 0, the transition probabilities are following

$$(0): \qquad \begin{aligned} p_{i,i-1} &= P_-(g_i), \\ p_{i,i} &= P_0(g_i), \\ p_{i,i+1} &= P_+(g_i). \end{aligned}$$

In case of buying or selling, we left the small probability $\epsilon > 0$ that the decision is not realized, in order to keep the chain irreducible. So if the decision to buy is realized the chain $\widehat{G}$ is shifted up. So, the transition probabilities looks like

$$(+): i \to i+1 \qquad \begin{aligned} &(1-\epsilon)\, P_-(g_{i+1}) \text{ in case of transition } i+1 \to i, \\ &(1-\epsilon)\, P_0(g_{i+1}) \text{ in case of transition } i+1 \to i+1, \\ &(1-\epsilon)\, P_+(g_{i+1}) \text{ in case of transition } i+1 \to i+2. \end{aligned}$$

In case that decision to buy is not realized the transition probabilities looks like

$$(+): i \to i \qquad \begin{aligned} &\epsilon\, P_-(g_i) \text{ in case of transition } i \to i-1, \\ &\epsilon\, P_0(g_i) \text{ in case of transition } i \to i, \\ &\epsilon\, P_+(g_i) \text{ in case of transition } i \to i+1. \end{aligned}$$

Putting altogether, we have

$$(+): \qquad \begin{aligned} p_{i,i-1} &= \epsilon\, P_-(g_i), \\ p_{i,i} &= \epsilon\, P_0(g_i) + (1-\epsilon)\, P_-(g_{i+1}), \\ p_{i,i+1} &= \epsilon\, P_+(g_i) + (1-\epsilon)\, P_0(g_{i+1}), \\ p_{i,i+2} &= (1-\epsilon)\, P_+(g_{i+1}). \end{aligned}$$

Similarly for decision to sell we get

$$(-): \quad \begin{aligned} p_{i,i-2} &= (1-\epsilon)\, P_-(g_{i-1}), \\ p_{i,i-1} &= (1-\epsilon)\, P_0(g_{i-1}) + \epsilon\, P_-(g_i), \\ p_{i,i} &= (1-\epsilon)\, P_+(g_{i-1}) + \epsilon\, P_0(g_i), \\ p_{i,i+1} &= \epsilon\, P_+(g_i). \end{aligned}$$

For any given control we have defined Markov chain $\{\widehat{G}_n\}_{n=0}^{\infty}$ that approximates the original process $\{G_t, t \geq 0\}$. Denote rewards in case of actions 0, + and − by r_0, r_+ and r_- respectively. According to discrete time analogue of the equation (2.19), we define rewards in case of no trading

$$r_0(g_i) = q_\gamma(g_i)\,\Delta t = (r - (\mu - r)\, g_i + \tfrac{1}{2}(1-\gamma)\sigma^2\, g_i^2)\,\Delta t$$

Trading bears a costs given by (2.14). Thus we define the rewards in case of trading as

$$\begin{aligned} r_+(g_i) &= r_0(g_i) - \delta\, \nu_+(g_i) = r_0(g_i) - \delta\, c_+\,(1 + c_+ g_i)^{-1}, \\ r_-(g_i) &= r_0(g_i) - \delta\, \nu_-(g_i) = r_0(g_i) - \delta\, c_-\,(1 - c_- g_i)^{-1}. \end{aligned}$$

2.4 Continuous Time Approximation

In this section, we approximate the model given in (2.21) by continuous time Markov Chain. In that case time domain of the original process and the approximating chain are the same, thus only the of approximation in the state domain is needed. Similarly to previous section, we start with derivation of change in moment characteristics. Due to continuity of the time domain, we focus on the infinitesimal change.

$$\mathbb{E}[\underbrace{G_{t+\mathrm{d}t} - G_t}_{\mathrm{d}G_t} | G_t = g] = b(g)\,\mathrm{d}t \tag{2.30}$$

$$\mathrm{var}[\mathrm{d}G_t | G_t = g] \sim \mathbb{E}[\mathrm{d}G_t^2 | G_t = g] = s^2(g)\,\mathrm{d}t. \tag{2.31}$$

Here we neglect the second order term

$$(\mathbb{E}[\mathrm{d}G_t | G_t = g])^2 = (b(g))^2\,(\mathrm{d}t)^2.$$

Consider the same approximation of the state space $[0,1]$ as in previous section. That is the partition $\{g_i\}_{i=1}^n$ with the norm equal to δ. We denote the continuous time approximating chain by $\widetilde{G} = (\widetilde{G}_t, t \geq 0)$. Further we denote the transition rate from any state g to $g + \delta$ by Q_+ and the transition rate from state g to $g - \delta$ by Q_-. We will not allow the other rates to be positive. So, for the total rate out of state g, denoted by Q_0, the equation $Q_0 + Q_+ + Q_- = 0$ holds. Note that, according to Kolmogorov backward differential equation (1.27), for t close to zero, we have the following differential relation between transition rates and transition probabilities

$$\mathrm{d}\boldsymbol{P}_t = \boldsymbol{Q}\,\boldsymbol{P}_0\,\mathrm{d}t = \boldsymbol{Q}\,\mathrm{d}t.$$

Using this, the infinitesimal change in mean value of the chain $\widetilde{G}$ can be expressed as

$$\begin{aligned}\mathbb{E}[\mathrm{d}\widetilde{G}_t|\widetilde{G}_t = g] &= \mathrm{d}P_-(g)\,(g-\delta) + \mathrm{d}P_0(g)\,g + \mathrm{d}P_+(g)\,(g+\delta)\\ &= [Q_-(g)\,(g-\delta) + Q_0(g)\,g + Q_+(g)\,(g+\delta)]\,\mathrm{d}t\\ &= [\delta(Q_-(g) + Q_+(g))]\,\mathrm{d}t\end{aligned} \tag{2.32}$$

Neglecting the second order term, we can express change in variance

$$\begin{aligned}\mathrm{var}[\mathrm{d}\widetilde{G}_t|\widetilde{G}_t = g] &\sim \mathbb{E}[(\mathrm{d}\widetilde{G}_t)^2|\widetilde{G}_t = g]\\ &= [Q_-(g)\,(g-\delta)^2 + Q_0(g)\,g^2 + Q_+(g)\,(g+\delta)^2]\,\mathrm{d}t\\ &\sim [\delta^2(Q_+(g) + Q_-(g))]\,\mathrm{d}t.\end{aligned} \tag{2.33}$$

Comparing (2.30) with (2.32), and (2.31) with (2.33) we gain the following equations

$$\begin{aligned}b(g) &= \delta(Q_+(g) - Q_-(g)),\\ s^2(g) &= \delta^2(Q_+(g) + Q_-(g)).\end{aligned} \tag{2.34}$$

Solving the equations (2.34) we get

$$Q_+(g) = \frac{\delta\,b(g) + s^2(g)}{2\delta^2} \geq 0,$$

$$Q_-(g) = \frac{-\delta\,b(g) + s^2(g)}{2\delta^2} \geq 0,$$

From (2.34) immediately follows that

$$Q_0(g) = -Q_+(g) - Q_-(g) = -\frac{s^2(g)}{\delta^2}.$$

Again consider the same set of admissible actions $\mathcal{U} = \{+,-,0\}$ as in previous section. As we derived above, in case of non-trading (action 0), the transition rates as follows

$$(0): \qquad \begin{aligned} q_{i,i-1} &= Q_-(g_i),\\ q_{i,i} &= Q_0(g_i),\\ q_{i,i+1} &= Q_+(g_i).\end{aligned}$$

In case of trading we would like the chain to move immediately from the current state. The immediate shift would be accomplished by infinite transition rate. However, for the finite state continuous chain the infinite rates are not allowed. Thus, we are forced to choose some large $K > 0$, which will represent the intensity out of a state in case of trading. So, for a decision to buy we have

$$(+): \qquad \begin{aligned} q_{i,i-1} &= Q_-(g_i),\\ q_{i,i} &= Q_0(g_i) - K,\\ q_{i,i+1} &= Q_+(g_i) + K,\end{aligned}$$

and for decision to sell we have

$$(-): \qquad \begin{aligned} q_{i,i-1} &= Q_-(g_i) + K, \\ q_{i,i} &= Q_0(g_i) - K, \\ q_{i,i+1} &= Q_+(g_i). \end{aligned}$$

For continuous time Markov chain we distinguish reward for staying in a state and reward for transition from one state to another. The reward for staying in a state g_i is described by (2.19). Thus we define

$$r(g_i) = q_\gamma(g_i) = (r - (\mu - r)\, g_i + \tfrac{1}{2}(1-\gamma)\sigma^2\, g_i^2).$$

This reward does not depend on decision taken in state g_i. The reward for transition from one state to another is related to a transaction cost. So these rewards occur only in case of trading assets, that is only in case of decision + or −. According to (2.14) we have

$$\begin{aligned} r_+(g_i) &= \nu_+(g_i) = -c_+\,(1 + c_+ g_i)^{-1}, \\ r_-(g_i) &= \nu_-(g_i) = -c_-\,(1 - c_- g_i)^{-1}. \end{aligned}$$

All the other transition rewards are equal to zero.

2.5 Numerical results

We implemented both discrete time and continuous time approximation of the model. Using the policy iteration algorithm developed in the first chapter we derive the optimal interval strategy according to the performance criteria (2.4). We compare the results with analytical solution.

Remind that the analytical solution is not of the closed-form. It is given as the maximum of the function $u(\alpha, \beta)$, which is defined by integral equation (2.24). We find the maximum approximately by evaluating the function u on the discrete lattice $\{(a,b) : 0 < a < b < 1, a = n\,\delta, b = m\,\delta\}$ with the step $\delta = 0.005$.

For both discrete time and continuous time approximation we choose the step in the state domain approximation $\delta = 0.005$. In case of discrete time approximation we choose the step in the time space approximation $\Delta t = 10^{-4}$.

Table 2.1: Results comparison

(a) $\mu = 0.5$

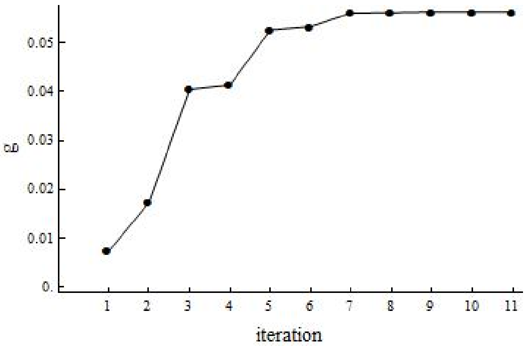
	g	interval
Cont. model	0.056373	$(0.170, 0.335)$
Discr. appr.	0.056364	$(0.165, 0.335)$
Cont. appr.	0.056375	$(0.165, 0.335)$

(b) $\mu = 0.9$

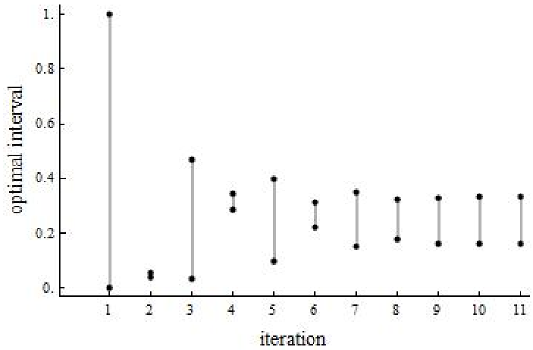
	g	interval
Cont. model	0.193513	$(0.350, 0.550)$
Discr. appr.	0.193490	$(0.355, 0.550)$
Cont. appr.	0.193510	$(0.355, 0.550)$

The table 2.1 shows the results of continuous model and both of its approximations. The certainty equivalent growth rates g and corresponding optimal interval strategies are compared. We consider transaction costs $c_+ = c_- = 0.02$, interest rate $r = 0$, volatility of the stock $\sigma = 1$ and $\gamma = -1$. The cases (a) and (b) differs by choice drift of the stock μ. We observed a sufficient consistency of both approximations with analytical solution.

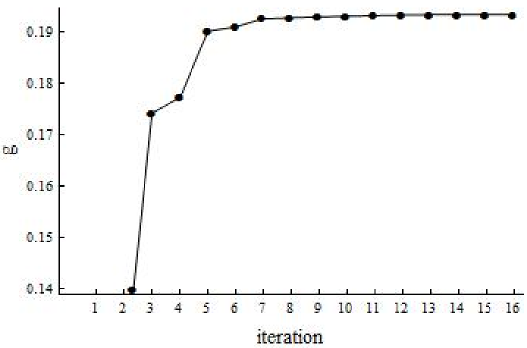
The figures 2.1 and 2.2 shows the iterations for discrete time approximation and continuous time approximation respectively. The left-hand side figures show the development of certainty equivalent growth rate. We see that each iteration improves the strategy. The right-hand figures show the development of the strategy.



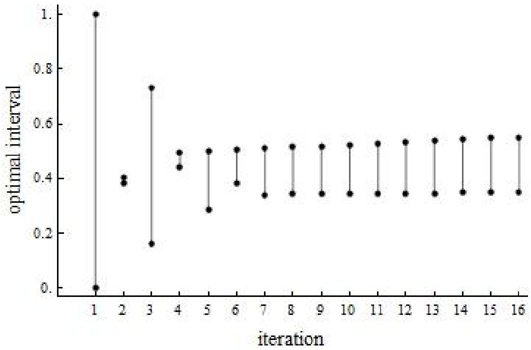
(a) Growth rate of `CE`, $\mu = 0.5$.



(b) Optimal strategy, $\mu = 0.5$.

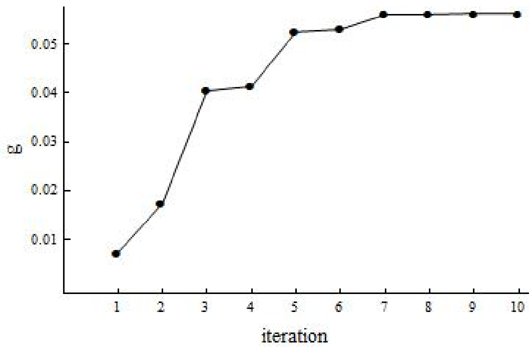


(c) Growth rate of `CE`, $\mu = 0.9$.

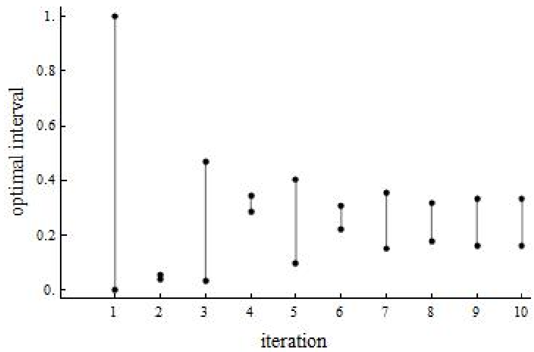


(d) Optimal strategy, $\mu = 0.9$

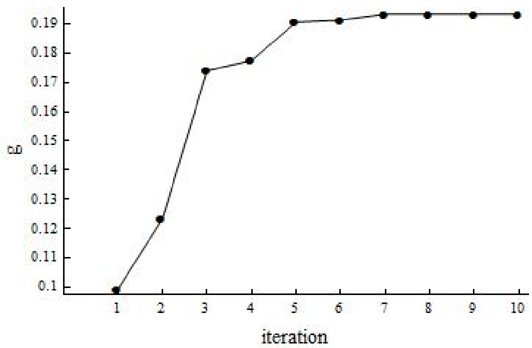
Figure 2.1: Discrete time approximation, process of finding the optimal strategy.



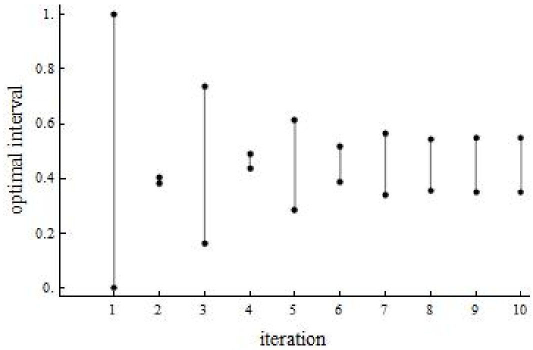
(a) Growth rate of `CE`, $\mu = 0.5$.



(b) Optimal strategy, $\mu = 0.5$.



(c) Growth rate of `CE`, $\mu = 0.9$.



(d) Optimal strategy, $\mu = 0.9$

Figure 2.2: Continuous time approximation, process of finding the optimal strategy.

Conclusion

For both discrete time and continuous time Markov decision chains we developed the iterative algorithm for finding a control that maximizes

$$\lim_{t\to\infty} \frac{1}{\gamma}\, t^{-1} \log(-\mathbb{E}[U_t]),$$

where U_t is utility from reward over the time horizon $[0,t]$. The algorithm is numerically tractable.

Both discrete time and continuous time version of algorithm were applied on a particular problem in portfolio optimization theory. It presents how a continuos time problem of optimal stochastic control can be solved numerically via Markov chain approximation. The method can be possibly applied for different problems. The results provide a sufficient consistency with the analytical solution, which demonstrate that they work properly.

A. Perron-Frobenius Theory

In this section we summarize the results of Perron-Frobenius theory about non-negative matrices that is used throughout the thesis. Comprehensive explanation of the topic can be found in monograph by Lancaster and Tismenetsky [14] in Chapter 15.

Let $\boldsymbol{A}$ be a square $n \times n$ matrix with non-negative entries. We denote this fact by $A \geq 0$.

Definition A.1. Let d_j be the greatest common divisor of those $m \geq 1$ for which $\boldsymbol{A}_{jj}^m > 0$. If $d_j = 1$ for all $j = 1, \dots, n$ then the matrix $\boldsymbol{A}$ is called *aperiodic.*

Definition A.2. The square matrix $\boldsymbol{A}$ of order n is called *irreducible* if for every permutation matrix $\boldsymbol{R}$

$$\boldsymbol{R}\,\boldsymbol{A}\,\boldsymbol{R}^{-1} \neq \begin{pmatrix} \boldsymbol{C} & \boldsymbol{0} \\ \boldsymbol{D} & \boldsymbol{F} \end{pmatrix},$$

where $\boldsymbol{C}$ and $\boldsymbol{F}$ are square matrices of order 2 at least.

These terms are directly related to the same terms defined for Markov chains. If X is a Markov chain with transition matrix $\boldsymbol{P}$, then X is aperiodic iff $\boldsymbol{P}$ is aperiodic, and X is irreducible iff $\boldsymbol{P}$ is irreducible.

Theorem A.3 (Perron-Frobenius)**.** *Let $\boldsymbol{A}$ be a nonnegative, irreducible and aperiodic square matrix. Then there exists a real eigenvalue $\lambda_1 > 0$ of $\boldsymbol{A}$, such that $\lambda_1 > |\lambda|$ for all other eigenvalues of $\boldsymbol{A}$. Moreover, (right) eigenvector $\boldsymbol{v}$ respective to λ_1 can be chosen entrywise positive and ker$(\boldsymbol{A} - \lambda \boldsymbol{I})$ is onedimensional. The same holds for a left eigenvalue $\boldsymbol{w}$, that is for a eigenvector of $\boldsymbol{A}^T$ respective to λ_1.*

Theorem A.4. *Let $\boldsymbol{A}$ be a nonnegative, irreducible and aperiodic square matrix. Let $\lambda > 0$ be its maximal eigenvalue given by Perron-Frobenius theorem. Then*

$$\lim_{k \to \infty} \lambda^{-k} \boldsymbol{A}^k = \frac{\boldsymbol{v}\,\boldsymbol{w}^T}{\boldsymbol{v}^T\,\boldsymbol{w}}$$

where $\boldsymbol{v}$ is any right eigenvector and $\boldsymbol{w}$ is any left eigenvector of $\boldsymbol{A}$ respective to λ.

Note that if $\boldsymbol{v}$ is chosen positive, then

$$\lim_{k \to \infty} \lambda^{-k} \boldsymbol{A}^k\,\boldsymbol{1} = \frac{\sum_i w_i}{\sum_i v_i\,w_i}\,\boldsymbol{v} = k(\boldsymbol{v}, \boldsymbol{w})\,\boldsymbol{v},$$

where $k(\boldsymbol{v}, \boldsymbol{w})$ is a positive constant.

Theorem A.5. *Let $\boldsymbol{A}$ be a nonnegative, ireducible square matrix. Let $\lambda > 0$ be its maximal eigenvalue and let $\boldsymbol{v}$ be a positive vector. Then following inequality holds*

$$\min_i \frac{\sum_j a_{ij}\,v_j}{v_i} \leq \lambda \leq \max_i \frac{\sum_j a_{ij}\,v_j}{v_i}$$

and the inequality with equality sign holds if and only if $\boldsymbol{v}$ is an eigenvector of $\boldsymbol{A}$ respective to λ.

B. Matrix Exponential Function

Here we summarize some facts about the matrix exponential that we used in Section 1.5. Particularlly, we are interested in its eigenvalues and eigenvectors.

Let $\boldsymbol{A}$ be a square $n \times n$ matrix. Define a matrix function on $[0, \infty)$ by

$$\exp\{u\,\boldsymbol{A}\} \triangleq \sum_{k=0}^{\infty} \frac{(u\,\boldsymbol{A})^k}{k!}. \tag{B.1}$$

We will assume that $\boldsymbol{A}$ has n distinct eigenvalues $\lambda_1, \ldots, \lambda_n$, i.e. the Jordan canonical form of $\boldsymbol{A}$ is of the form

$$\boldsymbol{D} \triangleq \operatorname{diag}(\{\lambda_i\}).$$

This simplifying assumtion hepls us to clearly demonstrate the idea behind. However, note that all the results that will be stated hold also for general case. The methods used for general case are similar, but considering the general Jordan canonical form the notation becomes more technical. For fully rigorous treatment see [1].

Using the Jordan canonical form the matrix $\boldsymbol{A}$ can be decomposed to $\boldsymbol{P}\,\boldsymbol{D}\,\boldsymbol{P}^{-1}$ for some matrix $\boldsymbol{P}$. Note that

$$\boldsymbol{A}^2 = \boldsymbol{P}\,\boldsymbol{D}\,\boldsymbol{P}^{-1}\,\boldsymbol{P}\,\boldsymbol{D}\,\boldsymbol{P}^{-1} = \boldsymbol{P}\,\boldsymbol{D}^2\,\boldsymbol{P}^{-1}.$$

So by induction we have $\boldsymbol{A}^k = \boldsymbol{P}\,\boldsymbol{D}^k\,\boldsymbol{P}^{-1}$ for $k \in \mathbb{N}$. Then the power series in B.1 can be expressed as

$$\begin{aligned}
\exp\{u\,\boldsymbol{A}\} &= \sum_{k=0}^{\infty} \frac{1}{k!} u^k\,\boldsymbol{P}\,\boldsymbol{D}^k\,\boldsymbol{P}^{-1} \\
&= \boldsymbol{P} \left(\sum_{k=0}^{\infty} \frac{u^k\,\boldsymbol{D}^k}{k!} \right) \boldsymbol{P}^{-1} \\
&= \boldsymbol{P}\,\exp\{u\,\boldsymbol{D}\}\,\boldsymbol{P}^{-1},
\end{aligned}$$

where

$$\exp\{u\,\boldsymbol{D}\} \triangleq \operatorname{diag}(\{e^{u\,\lambda_i}\}).$$

This shows the existence of $\exp\{u\,\boldsymbol{A}\}$. As $\exp\{u\,\boldsymbol{D}\}$ is the Jordan canonical form of $\boldsymbol{D}$, the eigenvalues of $\boldsymbol{D}$ are $e^{u\,\lambda_1}, \ldots, e^{u\,\lambda_n}$. Moreover, if $\boldsymbol{v}$ is an eigenvector of $\boldsymbol{A}$ corresponding to λ_i, then

$$\exp\{u\,\boldsymbol{A}\}\boldsymbol{v} = \sum_{k=0}^{\infty} \frac{u^k\,\boldsymbol{A}^k\,\boldsymbol{v}}{k!} = \sum_{k=0}^{\infty} \frac{u^k\,\lambda_i^k\,\boldsymbol{v}}{k!} = e^{\lambda_i}\,\boldsymbol{v}.$$

Consequently, the eigenspace of $\boldsymbol{A}$ corresponding to λ_i is identical to the eigenspace of $\exp\{\boldsymbol{A}\}$ corresponding to e^{λ_i}.

C. A Matrix Result

Here is the technical Lemma that is used in the proof of the Proposition 1.15. The idea behind the proof is just a direct use of Taylor expansion.

We start with the definition of a matrix norm that we will use.

Definition C.1. Let $\boldsymbol{A}$ be a square $n \times n$ matrix. Define the matrix *operator norm* by

$$\|\boldsymbol{A}\| = \max\{\|\boldsymbol{A}\,\boldsymbol{x}\| : \|\boldsymbol{x}\| \leq 1\},$$

where $\boldsymbol{x} \in \mathbb{R}^n$ and $\|\boldsymbol{x}\| = \max_i |x_i|$.

For any non-zero vector $\boldsymbol{x}$ we have $\|\boldsymbol{A}\,\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}\| \leq \|\boldsymbol{A}\|$ and thus $\|\boldsymbol{A}\,\boldsymbol{x}\| \leq \|\boldsymbol{A}\|\,\|\boldsymbol{x}\|$. Employing this inequality we get

$$\|\boldsymbol{A}\,\boldsymbol{B}\,\boldsymbol{x}\| \leq \|\boldsymbol{A}\|\;\|\boldsymbol{B}\,\boldsymbol{x}\| \leq \|\boldsymbol{A}\|\;\|\boldsymbol{B}\|\;\|\boldsymbol{x}\|\,,$$

and consequently

$$\|\boldsymbol{A}\,\boldsymbol{B}\| \leq \|\boldsymbol{A}\|\;\|\boldsymbol{B}\|\,. \tag{C.1}$$

Lemma C.2. *Let $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$ be square $n \times n$ matrices. Let entries of the main diagonal of the matrix $\boldsymbol{C}$ be equal to 1, e.i. $c_{ii} = 1$. Then*

$$\lim_{n\to\infty} \left[\left(\exp\{\tfrac{1}{n}\,\boldsymbol{A}\}\cdot\exp\{\tfrac{1}{n}\,\boldsymbol{B}\}\right) * \boldsymbol{C}\right]^n = \exp\{(\boldsymbol{A}+\boldsymbol{B}) * \boldsymbol{C}\}.$$

Proof. 1) First we show that the exponential $\exp\{\frac{1}{n}\,\boldsymbol{A}\}$ can be well approximated by $\boldsymbol{I} + \frac{1}{n}\,\boldsymbol{A}$, meaning that

$$\exp\{\tfrac{1}{n}\,\boldsymbol{A}\} = \boldsymbol{I} + \tfrac{1}{n}\,\boldsymbol{A} + O\left(\tfrac{1}{n^2}\right), \quad n \longrightarrow \infty. \tag{C.2}$$

It is sufficient to show that $\boldsymbol{H}(t) = t^{-2}(\exp\{\boldsymbol{A}t\} - \boldsymbol{I} - \boldsymbol{A}\,t)$ converges to some finite matrix as t goes to zero. Using L' Hospital's rule twice we get

$$\frac{d\,\boldsymbol{H}(t)}{d^2\,t} = \tfrac{1}{2}\,\boldsymbol{A}^2\,\exp\{\boldsymbol{A}\,t\} \longrightarrow \tfrac{1}{2}\,\boldsymbol{A}^2, \quad t \longrightarrow 0.$$

2) The relation (C.2) implies

$$\exp\{\tfrac{1}{n}\,\boldsymbol{A}\}\cdot\exp\{\tfrac{1}{n}\,\boldsymbol{B}\} = \boldsymbol{I} + \tfrac{1}{n}\,(\boldsymbol{A}+\boldsymbol{B}) + O\left(\tfrac{1}{n^2}\right), \quad n \longrightarrow \infty. \tag{C.3}$$

3) Because the main diagonal of the matrix $\boldsymbol{C}$ consists of ones, we have $\boldsymbol{I} * \boldsymbol{C} = \boldsymbol{I}$. Thus, using the relation (C.3) for n going to infinity

$$\begin{aligned}\left(\exp\{\tfrac{1}{n}\,\boldsymbol{A}\}\cdot\exp\{\tfrac{1}{n}\,\boldsymbol{B}\}\right) * \boldsymbol{C} &= \left[\boldsymbol{I} + \tfrac{1}{n}\,(\boldsymbol{A}+\boldsymbol{B}) + O\left(\tfrac{1}{n^2}\right)\right] * \boldsymbol{C} \\ &= \boldsymbol{I} + \tfrac{1}{n}\,\boldsymbol{E} + O\left(\tfrac{1}{n^2}\right),\end{aligned} \tag{C.4}$$

where $\boldsymbol{E} = (\boldsymbol{A}+\boldsymbol{B}) * \boldsymbol{C}$.
4) Finally we show the convergence. Define sequence

$$\boldsymbol{D}_n = \left(\exp\{\tfrac{1}{n}\,\boldsymbol{A}\} \cdot \exp\{\tfrac{1}{n}\,\boldsymbol{B}\}\right) * \boldsymbol{C} - \boldsymbol{I} + [\tfrac{1}{n}\,[(\boldsymbol{A}+\boldsymbol{B}) * \boldsymbol{C}]],$$

which is according to (C.4) $O\left(\frac{1}{n^2}\right)$. Using the binomial theorem we get

$$\begin{aligned}\left[\left(\exp\{\tfrac{1}{n}\,\boldsymbol{A}\} \cdot \exp\{\tfrac{1}{n}\,\boldsymbol{B}\}\right) * \boldsymbol{C}\right]^n &= [(\boldsymbol{I} + \tfrac{1}{n}\,(\boldsymbol{A}+\boldsymbol{B}) * \boldsymbol{C}) + \boldsymbol{D}_n]^n \\ &= \left[\boldsymbol{I} + \tfrac{1}{n}\,\boldsymbol{E}\right]^n + \sum_{k=1}^{n} \binom{n}{k} \boldsymbol{D}_n^k \,\left(\boldsymbol{I} + \tfrac{1}{n}\boldsymbol{E}\right)^{n-k}.\end{aligned}$$

We tend to show that the second term is negligible. Then using the triangle inequality and (C.1) we get

$$\begin{aligned}\left\|\sum_{k=1}^{n} \binom{n}{k} \boldsymbol{D}_n^k \,\left(\boldsymbol{I} + \tfrac{1}{n}\boldsymbol{E}\right)^{n-k}\right\| &\leq \sum_{k=1}^{n} \binom{n}{k} \|\boldsymbol{D}_n\|^k \,\left(\|\boldsymbol{I}\| + \tfrac{1}{n}\,\|\boldsymbol{E}\|\right)^{n-k} \\ &= \|\boldsymbol{D}_n\| \sum_{l=0}^{n-1} \binom{n-1}{l+1} \|\boldsymbol{D}_n\|^l \,\left(1 + \tfrac{1}{n}\,\|\boldsymbol{E}\|\right)^{n-1-l} \\ &\leq n\,\|\boldsymbol{D}_n\| \sum_{l=0}^{n-1} \binom{n-1}{l} \|\boldsymbol{D}_n\|^l \,\left(1 + \tfrac{1}{n-1}\,\|\boldsymbol{E}\|\right)^{n-1-l} \\ &= (n\,\|\boldsymbol{D}_n\|)\,(1 + \|\boldsymbol{D}_n\| + \tfrac{1}{n-1}\,\|\boldsymbol{E}\|)^{n-1}.\end{aligned}$$

First factor converges to 0 as $\boldsymbol{D}_n$ is $O(\frac{1}{n^2})$. The second term converges to $\exp\{\|\boldsymbol{E}\|\}$, e.i. finite number. Thus the whole term converges to 0. We can conclude

$$\begin{aligned}\left[\left(\exp\{\tfrac{1}{n}\,\boldsymbol{A}\} \cdot \exp\{\tfrac{1}{n}\,\boldsymbol{B}\}\right) * \boldsymbol{C}\right]^n &= [(\boldsymbol{I} + \tfrac{1}{n}\,(\boldsymbol{A}+\boldsymbol{B}) * \boldsymbol{C})]^n + o(1) \\ &\longrightarrow \exp\{(\boldsymbol{A}+\boldsymbol{B}) * \boldsymbol{C}\}, \quad n \longrightarrow \infty. \qquad \square\end{aligned}$$

D. Stochastic calculus

In this section we only introduce basic stochastic calculus tools that we use in Chapter 2. These are Ito's formula for computing the differential of a process and Girsanov Theorem. The whole stochastic calculus is a complex mathematical theory while deeper insight is out of the scope of this thesis. Readable but sufficiently rigorous explanation of the topic can be found for instance in monograph [18].

Theorem D.1 (Ito's Formula). *Let $f : \mathbb{R} \to \mathbb{R}$ be a twice continuously differentiable function and let $X = (X_t, t \geq 0)$ be a real-valued continuous semimartingale. Then*

$$f(X_t) = f(X_0) + \int_0^t f'(X_s)\mathrm{d}X_s + \int_0^t f''(X_s)\mathrm{d}\langle X, X\rangle_s. \tag{D.1}$$

We can reformulate the equation (D.1) using the differential notation in a following way,

$$\mathrm{d}f(X_t) = f'(X_t)\mathrm{d}X_t + f''(X_t)\mathrm{d}\langle X\rangle_t. \tag{D.2}$$

Here are the two applications of (D.2) that are used in the Chapter 2.

1. For $f(x) = \ln(x)$ we have $f'(x) = x^{-1}$, $f''(x) = -x^{-2}$ and
$$\mathrm{d}\ln(X_t) = \frac{\mathrm{d}X_t}{X_t} - \frac{\mathrm{d}\langle X\rangle_t}{2X_t^2}.$$
2. For $f(x) = x^{-1}$ we have $f'(x) = -x^{-2}$ $f''(x) = 2x^{-3}$ and
$$\mathrm{d}X_t^{-1} = -\frac{\mathrm{d}X_t}{X_t^2} + \frac{\mathrm{d}\langle X\rangle_t}{X_t^2}.$$

We say that the process G satisfies *Novikov condition* on $[0, T]$ if

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^T G_s^2\,\mathrm{d}s\right)\right] < \infty. \tag{D.3}$$

Theorem D.2 (Girsanov Theorem). *Let $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ be a filtrated probability space satisfying usual conditions (UC). Let W be an $(\mathcal{F}_t)$-martingale and let G_t be an $(\mathcal{F}_t)$-progressively measeruble process satisfying Novikov condition on every interval $[0, T]$. Put*

$$X \triangleq \int_0^t G_s\,\mathrm{d}W_s.$$

Define stochastic exponential $\mathcal{E}(X)$ by

$$\mathcal{E}(X)_t \triangleq \exp\{X_t - \tfrac{1}{2}\langle X\rangle_t\}.$$

Then for any fixed $T > 0$ there exist a measure $\mathbb{Q}_T$ such that

$$\frac{\mathrm{d}\mathbb{Q}_T}{\mathrm{d}\mathbb{P}}|_{\mathcal{F}_T} = \mathcal{E}(X)_T.$$

Moreover

$$\widetilde{W}_t = W_t - \langle W, X\rangle_t\,, \quad 0 \leq t \leq T.$$

is a Brownian motion under the measure $\mathbb{Q}_T$.

Bibliography

[1] Baker A., (2006), “Matrix Groups: An Introduction to Lie Group Theory”, 3rd edition, Princeton Springer Great Britain.

[2] Bellman R.E., (1957), “Dynamic Programming”, Princeton University Press, NJ.

[3] Bielicky T.R., Chancelier J. P., Pliska S. R., Sulem A., (2003), “Risk sensitive portfolio optimization with transaction cost”, *Journal of Computational Finance, 8(1), 39-63.*

[4] Bielicky T.R., Pliska S. R., (2000), “Risk sensitive dynamic asset management in the presence of transaction costs”, *Finance and Stochastics*, vol. 4, pp. 1-33.

[5] Cadenilas A., (2000), “Consumption-investment problems with transaction cost: survey and open problems,”, *Math. Methods Operations Research, vol 51*, pp,43-68.

[6] Dostál P., “Investment strategies in the long run with proportional transaction costs and a HARA utility function”,*Quantitative Finance, Vol. 9, No. 2, March 2009.*

[7] Ethier N., Kurtz G., (2005), “Markov Processes”, John Wiley & Sons, New Jersey.

[8] Follmer H., Schied A., (2004), “Stochastic Finance. An Introduction in Discrete Time”, second edition, Walter de Gruyter, Berlin

[9] Howard R.A., (1972), “Dynamic Programming and Markov Processes”, *Management Science*, Vol. 18, No. 7, March 1972.

[10] Howard R.A., Matheson J.E., (1960), “Risk-Sensitive Markov Decision pocesses”, The M.I.T. Press, Princeton, NJ.

[11] Howard R.A, Metheson J.E., (1972), “Risk-Sensitive Markov Decision Processes”,. *Management Science*, Vol. 18, No. 7, March.

[12] Karlin S., Taylor H.M., (1981), “A Second Course in Stochastic Processes”, Academic Press, New York.

[13] Kushner H.J., Dupuis P., (2001), “Numerical Methods for Stochastic Control Problems in Continuous Time”, Springer-Verlag, New York .

[14] Lancaster P., Tismenetsky M., (1985), “The Theory of Matrices Second Edition”, Academic Press.

[15] Liggett T.M., (2010), “Continuous Time Markov Processes: An Introduction”, American Mathematical Society, Providence.

[16] Norris J.R., (1997), “Markov Chain”, Cambridge University Press, Cambridge.

[17] Merton R.C., (1992), “Continuous-Time Finance”, B. Blackwell.

[18] Shreve S.E., (1981), “Stochastic Calculus for Finance II”, Springer, New York.

[19] Staníková D. (2006), “Asymptotic Control of Portfolio”, Bachelor thesis, MFF UK.