Remarks on risk neutral and risk sensitive portfolio optimization

G.B. Di Masi
Università di Padova
Dipartimento di Matematica
Pura ed Applicata
Via Belzoni 7, 35131 Padova
and CNR-LADSEB

Ł. Stettner
Institute of Mathematics
Polish Academy of Sciences
Sniadeckich 8, 00-950 Warsaw

Abstract

In this note it is shown that risk neutral optimal portfolio strategy is nearly optimal for risk sensitive portfolio cost functional with negative risk factor that is close to 0.

Key words: risk sensitive control, discrete time Markov processes, splitting, Poisson equation, Bellman equation
AMS subject classification: primary: 93E20 secondary: 60J05, 93C55

1 Introduction

Assume we are given a market with m risky assets. Denote by $S_i(t)$ the price of the i-th asset at time t. We shall assume that the prices of assets depend on k economical factors $x_i(n)$, $i = 1, \ldots, k$ with dynamics changing in discrete time moments denoted for simplicity by $n = 0, 1, \ldots,$ in the following way:
for $t \in [n, n+1)$

$$\frac{dS_i(t)}{S_i(t)} = a_i(x(n))dt + \sum_{j=1}^{k+m} \sigma_{ij}(x(n))dw_j(t), \tag{1}$$

where $(w(t) = (w_1(t), w_2(t), \ldots, w_{k+m}(t))$ is a $k+m$ dimensional Brownian motion defined on a given probability space $(\Omega, (\mathcal{F}_t), \mathcal{F})$. Economical factors $x(n) = (x_1(n), \ldots, x_k(n))$, satisfy the equation

$$x_i(n+1) = x_i(n) + b_i(x(n)) + \sum_{j=1}^{k+m} d_{ij}(x(n))(w_j(n+1) - w_j(n)) = g(x(n), W(n)), \quad (2)$$

where $W(n) := (w_1(n+1) - w_1(n), \ldots, w_{k+m}(n+1) - w_{k+m}(n))$.

We assume that a, b are bounded continuous vector functions, and σ, d are bounded continuous matrix functions of suitable dimensions. Additionally we shall assume that the matrix dd^T (^{T} stands for transponse) is nondegenerate. Notice that equation (2) corresponds to discretization of a diffusion process. The set of factors may include divident yields, price - earning rations, short term interest rates, the rate of inflation see e.g. [1]. The dynamics of such factors is usually modeled using diffusion, frequently linear equations eg. in the case when we assume following [1] that a is a function of spot interest rate governed by the Vasicek process. Our assumptions concerning boundedness of vector functions a and b may be relaxed allowing linear growth, however is such case we shall need more complicated assumptions to obtain analogs of assertions in Lemmas 4,5 and Corollary 3 which are important in the proof of Proposition 3.

Assume that starting with an initial capital $V(0)$ we invest in assets. Let $h_i(n)$ be the part of the wealth process located in the i-th asset at time n, which is assumed to be nonnegative. The choice of $h_i(n)$ depends on our observation of the asset prices and economical factors up to time n. Denoting by $V(n)$ the wealth process at time n and by $h(n) = (h_1(n), \ldots, h_m(n))$ our investment strategy at time n, we have that $h(n) \in U = \{(h_1, \ldots, h_m), h_i \geq 0, \sum_{i=1}^m h_i = 1\}$ and

$$\frac{V(n+1)}{V(n)} = \sum_{i=1}^{m} h_i(n)\xi_i(x(n), W(n)), \quad (3)$$

where

$$\xi_i(x(n), W(n)) = \exp\left\{a_i(x(n)) - \frac{1}{2}\sigma_{ij}^2(x(n)) \right.$$
$$\left. + \sum_{j=1}^{k+m} \sigma_{ij}(x(n))(w_j(n+1) - w_j(n))\right\}$$

We are interested in the following investment problems:
maximize *risk neutral cost functional*

$$J_x^0((h(n)) = \liminf_{n\to\infty} \frac{1}{n} E_x\left\{\ln V(n)\right\} \quad (4)$$

and maximize *risk sensitive cost functional*

$$J_x^\gamma((h(n)) = \frac{1}{\gamma} \limsup_{n\to\infty} \frac{1}{n} \ln E_x \left\{ V(n)^\gamma \right\} \tag{5}$$

with $\gamma < 0$. Using (3) we can rewrite the cost functionals (4) and (5) in the more convenient forms. Namely,

$$\begin{aligned} J_x^0((h(n)) &= \liminf_{n\to\infty} \frac{1}{n} E_x \left\{ \sum_{t=0}^{n-1} \ln \left(\sum_{i=1}^m h_i(t)\xi_i(x(t), h(t)) \right) \right\} \\ &= \liminf_{n\to\infty} \frac{1}{n} E_x \left\{ \sum_{t=0}^{n-1} c(x(t), h(t)) \right\}, \end{aligned} \tag{6}$$

with $c(x,h) = E\left\{\ln\left(\sum_{i=1}^m h_i \xi_i(x, W(0))\right)\right\}$. It is clear that risk neutral cost functional J^0 depends on uncontrolled Markov process $(x(n))$ and we practically maximize the cost function c itself. Consequently an optimal control is of the form control $(\hat{u}(x(n)))$, where $\sup_h c(x,h) = c(x, \hat{u}(x))$ and function Borel measurable $\hat{u} : R^k \mapsto U$ exists by continuity of c for fixed $x \in R^k$. This control does not depend on asset prices and is a time independent function of current values of the factors x only. The Bellman equation corresponding to the risk neutral control problem is of the form

$$w(x) + \lambda = \sup_h \left(c(x,h) + Pw(x) \right) \tag{7}$$

where $Pf(x) := E_x \left\{ f(x(1)) \right\}$ for $f \in b\mathcal{B}(R^k)$ - the space of bounded Borel measurable functions on R^k, is a transition operator corresponding to $(x(n))$. In section 2 we shall show that there are solutions w and λ to the equation (7) and λ is an optimal value of the cost functional J^0.
Letting

$$\begin{aligned} \zeta_n^{h,\gamma}(\omega) := \prod_{t=0}^{n-1} &\exp[\gamma \ln \left(\sum_{i=1}^m h_i(t)\xi_i(x(t), W(t)) \right)] \\ &\left(E\left\{ \exp[\gamma \ln \left(\sum_{i=1}^m h_i(t)\xi_i(x(t), W(t)) \right)] | \mathcal{F}_t \right\} \right)^{-1} \end{aligned}$$

consider a probability measure $P^{h,\gamma}$ defined by its restrictions $P^{h,\gamma}$ to the first n time moments given by the formula

$$P_{|n}^{h,\gamma}(d\omega) = \zeta_n^{h,\gamma}(\omega) P_{|n}(d\omega).$$

Then

$$
\begin{aligned}
J_x^\gamma((h(n)) &= \frac{1}{\gamma} \limsup_{n\to\infty} \frac{1}{n} \ln E_x \left\{ \exp \left\{ \gamma \sum_{t=0}^{n-1} \ln \left(\sum_{i=1}^m h_i(t)\xi_i(x(t), W(t)) \right) \right\} \right\} \\
&= \frac{1}{\gamma} \limsup_{n\to\infty} \frac{1}{n} \ln E_x^{h,\gamma} \left\{ \exp \left\{ \sum_{t=0}^{n-1} c_\gamma(x(t), h(t)) \right\} \right\}, \qquad (8)
\end{aligned}
$$

with

$$
c_\gamma(x,h) := \ln \left(E \left\{ \left(\sum_{i=1}^m h_i \xi_i(x, W(0)) \right)^\gamma \right\} \right). \qquad (9)
$$

Risk sensitive Bellman equation corresponding to the cost functional J^γ is of the form

$$
e^{w_\gamma(x)} = \inf_h [e^{(c_\gamma(x,h)-\lambda_\gamma)} \int_E e^{w_\gamma(y)} P^{h,\gamma}(x, dy)]. \qquad (10)
$$

where for $f \in b\mathcal{B}(R^k)$

$$
P^{h,\gamma} f(x) = E \left\{ \left(\sum_{i=1}^m h_i \xi_i(x, W(0)) \right)^\gamma \exp\left\{-c_\gamma(x,h)\right\} f(g(x, W(0))) \right\}, \qquad (11)
$$

and where $\frac{1}{\gamma}\lambda_\gamma$ corresponds to optimal value of the cost functional (8). Notice that under measure $P^{h,\gamma}$ the process $(x(n))$ is still Markov but with controlled transition operator $P^{h,\gamma}(x, dy)$. Following [6] we shall show that

$$
\frac{1}{\gamma}\lambda_\gamma \to \lambda \qquad (12)
$$

whenever $\gamma \uparrow 0$.

In what follows we shall distinguish the following special classes of controls (h_n): *Markov controls* $\mathcal{U}_M = \{(h(n)) : h(n) = u_n(x(n))\}$, where $u_n : R^k \mapsto U$, is a sequence of Borel measurable functions, and *stationary controls* $\mathcal{U}_s = \{(h_n) : h(n) = u(x(n))\}$, where $u : R^k \mapsto U$ is a Borel measurable function. We shall denote by $\mathcal{B}(R^k)$ the set of Borel subsets of R^k and by $\mathcal{P}(R^k)$ the set of probability measures on R^k.

The study of risk sensitive portfolio optimization has been originated in [1] and then continued in a number of papers in particular in [16]. Risk sensitive cost functional was studied in papers [13], [6], [7], [3], [4], [12] [2], [8] and references therein. In this paper using splitting of Markov processes arguments (see [15]) we study Poisson equation for additive cost functional the solution of which is also a solution to risk neutral Bellman equation. We consider then risk sensitive portfolio optimization with risk factor close to 0. We generalize the result of [16], where uniform ergodicity of factors was required and using

[8] show the existence of Bellman equation for small risk in a more general ergodic case. The proof of that nearly optimal continuous risk neutral control function is also nearly optimal for risk sensitive cost functional with risk factor close to 0 is based on modification of the arguments of [6] using some results from the theory of large deviations.

2 Risk neutral Bellman equation

By nondegeneracy of the matrix dd^T there exist: a compact set $C \subset R^k$, for which we can take a closed ball in R^k, $\beta > 0$ and $\nu \in \mathcal{P}(R^k)$ such that $\nu(C) = 1$ and $\forall_{A \in \mathcal{B}(R^k)}$

$$\inf_{x \in C} P(x, A) \geq \beta \nu(A). \tag{13}$$

We fix a compact set C, $\beta > 0$ and $\nu \in \mathcal{P}(R^k)$ satisfying the above minorization property. Additionally assume that the set C is ergodic, i.e. $\forall_{x \in R^k}\ E_x\{\tau_C\} < \infty$ and $\sup_{x \in C} E_x\{\tau_C\} < \infty$, where $\tau_C = \inf\{i > 0 : x_i \in C\}$.
Consider a splitting of the Markov process $(x(n))$ (see [15]).
Let $\hat{R^k} = \left\{C \times \{0\} \cup C \times \{1\} \cup R^k \setminus C \times \{0\}\right\}$ and $\hat{x}(n) = (x^1(n), x^2(n))$ be a Markov process defined on $\hat{R^k}$ such that

(i) when $(x^1(n), x^2(n)) \in C \times \{0\}$, $x^1(n)$ moves to y accordingly to $(1-\beta)^{-1}(P(x^1(n), dy)$ $-\beta\nu(dy))$ and whenever $y \in C$, $x^2(n)$ is changed into $x^2(n+1) = \beta_{n+1}$, where β_n is i.i.d. $P\{\beta_n = 0\} = 1 - \beta$, $P\{\beta_n = 1\} = \beta$,

(ii) when $(x^1(n), x^2(n)) \in C \times \{1\}$, $x^1(n)$ moves to y accordingly to ν and $x^2(n+1) = \beta_{n+1}$,

(iii) when $(x^1(n), x^2(n)) \in R^k \setminus C \times \{0\}$, $x^1(n)$ moves to y accordingly to $P(x^1(n), dy)$ and whenever $y \in C$, $x^2(n)$ is changed into $x^2(n+1) = \beta_{n+1}$.

Let $C_0 = C \times \{0\}$, $C_1 = C \times \{1\}$.
Following [8] and [15] we have

Proposition 1 *For $n = 1, 2 \ldots$ we have P a.e.*

$$P\{\hat{x}(n) \in C_0 | \hat{x}(n) \in C_0 \cup C_1, \hat{x}(n-1), \ldots, \hat{x}(0)\} = 1 - \beta \tag{14}$$

The process $(\hat{x}(n) = (x^1(n), x^2(n)))$ is Markov with transition operator $\hat{P}(\hat{x}(n), dy)$ defined by (i)-(iii). Its first coordinate $(x^1(n))$ is also a Markov process with transition operator

$P(x^1(n), dy)$. Furthermore for any bounded Borel measurable function $f : (R^k)^{n+1} \mapsto R$ we have

$$E_x\left\{f(x(1), x(2), \ldots, x(n))\right\} = \hat{E}_{\delta_x^*}\left\{f(x^1(1), x^1(2), \ldots, x^1(n))\right\} \tag{15}$$

where $\delta_x^* = \delta_{(x,0)}$ for $x \in E \setminus C$ and $\delta_x^* = (1-\beta)\delta_{(x,0)} + \beta\delta_{(x,1)}$ for $x \in C$ and $\hat{E}_\mu$ stands for conditional law of Markov process $(\hat{x}(n))$ with initial law $\mu \in \mathcal{P}(\hat{R^k})$.

Proof. Since the Markov property of $(x^1(n))$ is fundamental in this paper we recall this proof from [8] leaving the proof of other statements to the reader. For $A \in R^k$ we have

$$\begin{aligned}
&P\left\{x^1(n+1) \in A | x^1(n), x^1(n-1), \ldots, x^1(0)\right\} \\
&= P\left\{x^1(n+1) \in A | x^1(n), x^2(n) = 0, x^1(n-1), \ldots, x^1(0)\right\} \\
&P\left\{x^2(n) = 0 | x^1(n), x^1(n-1), \ldots, x^1(0)\right\} \\
&+P\left\{x^1(n+1) \in A | x^1(n), x^2(n) = 1, x^1(n-1), \ldots, x^1(0)\right\} \\
&P\left\{x^2(n) = 1 | x^1(n), x^1(n-1), \ldots, x^1(0)\right\}.
\end{aligned}$$

In the case when $x^1(n) \in C$, the right hand side of the last equation is equal to

$$\frac{P^{a_n}(x^1(n), A) - \beta\nu(A)}{1-\beta}(1-\beta) + \beta\nu(A) = P^{a_n}(x^1(n), A).$$

For $x^1(n) \notin C$, it is equal to $P^{a_n}(x^1(n), A)$, which completes the proof of Markov property of $(x^1(n))$.

□

By the assumption on C and the construction of the split Markov process we immediately have

Corollary 1 *$\hat{E}_x\left\{\tau_{C_1}\right\} < \infty$ for $x \in \hat{R^k}$ and $\sup_{x \in C_1} \hat{E}_x\left\{\tau_{C_1}\right\} < \infty$.*

Lemma 1 *Given $(h(n)) \in \mathcal{U}_M$ there is a unique $\lambda((h(n)))$ such that for $x \in C_1$*

$$\hat{E}_x\left\{\left\{\sum_{t=1}^{\tau_{C_1}}\left(c(x^1(t), h(t)) - \lambda((h(n)))\right)\right\}\right\} = 0 \tag{16}$$

Proof. Notice that for $x \in C_1$ the mapping

$$D : \lambda \mapsto \hat{E}_x\left\{\left\{\sum_{t=1}^{\tau_{C_1}} c(x^1(t), h(t)) - \lambda\right\}\right\}$$

is continuous and strictly decreasing. Since the values of this mapping for $\|c\|$ and $-\|c\|$ are respectively nonpositive and nonnegative there is a unique λ for which the mapping attains 0. $\square$

For Borel measurable $u : R^k \mapsto U$ let

$$\hat{w}^u(x) = \hat{E}_x \left\{ \sum_{t=0}^{\tau_{C_1}} \Big(c(x^1(t), u(x^1(t))) - \lambda(u) \Big) \right\}, \tag{17}$$

where we use notation $\lambda(u) = \lambda(u(x(n))$.

Lemma 2 *Function $\hat{w}^u$ defined in (17) is a unique up to an additive constant solution to the additive Poisson equation (APE) for the split Markov process $(\hat{x}(n))$:*

$$\hat{w}^u(x) = c(x^1, u(x^1)) - \lambda(u) + \int_{\hat{R^k}} \hat{w}^u(y) \hat{P}(x, dy) \tag{18}$$

Furthermore, if $\hat{w}$ and λ satisfy the equation

$$\hat{w}(x) = c(x^1, u(x^1)) - \lambda + \int_{\hat{R^k}} \hat{w}(y) \hat{P}(x, dy) \tag{19}$$

then $\lambda = \lambda(u)$ (defined in Lemma 1) and $\hat{w}$ differs from $\hat{w}^u$ by an additive constant.

Proof. In fact, we have using (16)

$$\hat{E}_x \left\{ w(\hat{x}(1)) \right\} = \hat{E}_x \left\{ \chi_{\hat{x}(1) \in C_1} \hat{E}_{x(1)} \left\{ \sum_{t=0}^{\tau_{C_1}} c(x^1(t), u(x^1(t))) - \lambda(u) \right\} \right\}$$

$$+ \hat{E}_x \left\{ \chi_{\hat{x}(1) \notin C_1} \hat{E}_{x(1)} \left\{ \sum_{t=0}^{\tau_{C_1}} c(x^1(t), u(x^1(t))) - \lambda(u) \right\} \right\} = \hat{E}_x \Big\{ \chi_{\hat{x}(1) \in C_1}$$

$$\Big\{ c(x^1(1), u(x^1(1))) - \lambda(u) \Big\} \Big\} + \hat{E}_x \left\{ \chi_{\hat{x}(1) \notin C_1} \sum_{t=0}^{\tau_{C_1}} c(x^1(t), u(x^1(t))) - \lambda(u) \right\}$$

$$= \hat{E}_x \left\{ \sum_{t=0}^{\tau_{C_1}} c(x^1(t), u(x^1(t))) - \lambda(u) \right\} - (c(x^1, u(x^1)) - \lambda(u))$$

from which (18) follows. If $\hat{w}^u$ is a solution to (18) then by iteration we obtain that

$$\hat{w}^u(x) = \hat{E}_x \left\{ \sum_{t=0}^{\tau_{C_1}} \Big(c(x^1(t), u(x^1(t))) - \lambda(u) \Big) + \hat{E}_{\hat{x}_{\tau_{C_1}}} \left\{ \hat{w}^u(\hat{x}(1)) \right\} \right\}, \tag{20}$$

where by the construction of the split Markov process

$$\hat{E}_{x_{\tau_{C_1}}}\left\{\hat{w}^u(\hat{x}(1))\right\} = (1-\beta)\int_{R^k}\hat{w}^u(z,0)\nu(dz) + \beta\int_{R^k}\hat{w}^u(z,1)\nu(dz).$$

Consequently $\hat{w}^u$ differs from $\hat{w}^u$ defined in (17) only by an additive constant. Similarly, if $\hat{w}$ and λ are solutions to (19) then $\hat{w}$ differs from

$$\tilde{w}(x) = \hat{E}_x\left\{\sum_{t=0}^{\tau_{C_1}}\Big(c(x^1(t),u(x^1(t))) - \lambda\Big)\right\}$$

by an additive constant $\hat{E}_z\left\{\hat{w}(\hat{x}(1))\right\}$ with $z \in C_1$. Since $\tilde{w}$ itself is a solution to (19) we have that $\hat{E}_z\left\{\tilde{w}(\hat{x}(1))\right\} = 0$ for $z \in C_1$. Therefore for $z \in C_1$

$$\begin{aligned}
0 = \hat{E}_z\left\{\tilde{w}(\hat{x}(1))\right\} &= \hat{E}_z\Bigg\{\chi\hat{R^k}\setminus C_1(\hat{x}(1))\sum_{t=1}^{\tau_{C_1}}\Big(c(x^1(t),u(x^1(t))) - \lambda\Big) \\
&+\chi_{C_1}(\hat{x}(1))\hat{E}_{\hat{x}(1)}\left\{\sum_{t=0}^{\tau_{C_1}}\Big(c(x^1(t),u(x^1(t))) - \lambda\Big)\right\}\Bigg\} \\
&= \hat{E}_z\left\{\sum_{t=1}^{\tau_{C_1}}\Big(c(x^1(t),u(x^1(t))) - \lambda\Big)\right\}
\end{aligned}$$

and by Lemma 1 we have $\lambda = \lambda(u)$ which completes the proof.

□

Corollary 2 *Given solution $\hat{w}^u : \hat{R^k} \mapsto R$ to APE (18) we have that w^u defined by*

$$w^u(x) := \hat{w}^u(x,0) + 1_C(x)\beta(\hat{w}^u(x,1) - \hat{w}^u(x,0)) \tag{21}$$

is a solution to APE for the original Markov process $(x(n))$

$$w^u(x) = c(x,u(x)) - \lambda(u) + \int_{R^k} w^u(y)P(x,dy). \tag{22}$$

Furthermore if w^u is a solution to (22) then $\hat{w}^u$ defined by

$$\hat{w}^u(x^1,x^2) = c(x^1,u(x^1)) - \lambda(u) + \hat{E}_{x^1,x^2}\left\{w^u(x^1(1))\right\} \tag{23}$$

is a solution to (18).

Proof. By (14) we have

$$\begin{aligned}&\hat{E}_x\left\{\hat{w}^u(\hat{x}(1))\right\} = \hat{E}_x\left\{\hat{E}_x\left\{\hat{w}^u(\hat{x}(1))|x^1(1)\right\}\right\}\\&= \hat{E}_x\Big\{\chi_C(x^1(1))((1-\beta)\hat{w}^u(x^1(1),0) + \beta\hat{w}^u(x^1(1),1))\\&+\chi_{E\setminus C}(x^1(1))\hat{w}^u(x^1(1),0)\Big\} = \hat{E}_x\left\{w^u(x^1(1))\right\}\end{aligned} \tag{24}$$

Therefore by (18) we obtain that w^u defined in (21) is a solution to (22). Assume now that w^u is a solution to (22). Then by (15)

$$\hat{E}_{\delta_x^*}\left\{w^u(x^1(1))\right\} = E_x\left\{w^u(x(1))\right\}$$

and for $\hat{w}^u$ given in (23) we obtain (21). From (21) we obtain (24) which in turn by (23) shows that $\hat{w}^u$ is a solution to (18).

$\square$

Remark 1 *APE has been a subject of intensive studies in [14] (together with so called multiplicative Poisson equation). As is shown above the use of splitting techniques gives an explicit form for a solution to this equation.*

The value of $\lambda(u)$ has another important characterization. Namely, we have

Proposition 2 *For Borel measurable $u : R^k \to U$ the value $\lambda(u)$ defined in Lemma 1 is equal to*

$$\lambda(u) = \lim_{n\to\infty} \frac{1}{n} E_x\left\{\sum_{t=0}^{n-1} c(x(t), u(x(t)))\right\} \tag{25}$$

Proof. Let $\lambda > \lambda(u)$. For $z \in C_1$ we have

$$\hat{E}_z\left\{\sum_{t=1}^{\tau_{C_1}}\Big(c(x^1(t), u(x^1(t))) - \lambda\Big)\right\} < 0$$

and consequently for $N \geq N_0$

$$\hat{E}_z\left\{\sum_{t=1}^{\tau_{C_1}\wedge N}\Big(c(x^1(t), u(x^1(t))) - \lambda\Big)\right\} \leq 0. \tag{26}$$

Let

$$w_N^u(x) = \hat{E}_x\left\{\sum_{t=0}^{\sigma_{C_1}\wedge N-1}\Big(c(x^1(t), u(x^1(t))) - \lambda\Big)\right\} \tag{27}$$

with $\sigma_{C_1} = \inf\{t \geq 0 : \hat{x}(t) \in C_1\}$.
For $x \notin C_1$

$$w^u_{N+1}(x) = \hat{E}_x \left\{ c(x^1(0), u(x^1(0))) - \lambda + \hat{E}_{\hat{x}(1)} \left\{ \sum_{t=0}^{\sigma_{C_1} \wedge N-1} \Big(c(x^1(t), u(x^1(t))) - \lambda \Big) \right\} \right\}$$
$$= \hat{E}_x \left\{ c(x^1(0), u(x^1(0))) - \lambda + w^u_N(\hat{x}(1)) \right\} \tag{28}$$

and $x \in C_1$ by (26) we have

$$w^u_{N+1}(x) = c(x^1(0), u(x^1(0))) - \lambda \geq$$
$$\geq \hat{E}_x \left\{ c(x^1(0), u(x^1(0))) - \lambda + \hat{E}_{\hat{x}(1)} \left\{ \sum_{t=0}^{\sigma_{C_1} \wedge N-1} \Big(c(x^1(t), u(x^1(t))) - \lambda \Big) \right\} \right\}$$
$$= \hat{E}_x \left\{ c(x^1(0), u(x^1(0))) - \lambda + w^u_N(\hat{x}(1)) \right\} \tag{29}$$

Consequently

$$w^u_{N+1}(x) \geq \hat{E}_x \left\{ c(x^1(0), u(x^1(0))) - \lambda + w^u_N(\hat{x}(1)) \right\} \tag{30}$$

and by iteration for $N \geq N_0$

$$w^u_{N+k}(x) \geq \hat{E}_x \left\{ \sum_{t=0}^{k-1} \Big(c(x^1(t), u(x^1(t))) - \lambda \Big) + w^u_N(\hat{x}(k)) \right\}$$
$$\geq \hat{E}_x \left\{ \sum_{t=0}^{k-1} c(x^1(t), u(x^1(t))) - \lambda - \|c\| N \right\}.$$

Therefore

$$\frac{1}{k} \hat{E}_x \left\{ \sum_{t=0}^{k-1} c(x^1(t), u(x^1(t))) \right\} \leq \frac{1}{k} \|c\| N$$
$$+ \frac{1}{k} \sup_N \hat{E}_x \left\{ \sum_{t=1}^{\sigma_{C_1} \wedge N-1} \Big(c(x^1(t), u(x^1(t))) - \lambda(u) \Big) \right\} + \lambda$$

and consequently

$$\limsup_{k \to \infty} \frac{1}{k} \hat{E}_x \left\{ \sum_{t=0}^{k-1} c(x^1(t), u(x^1(t))) \right\} \leq \lambda$$

Letting λ decreasing to $\lambda(u)$ we obtain

$$\limsup_{k \to \infty} \frac{1}{k} \hat{E}_x \left\{ \sum_{t=0}^{k-1} c(x^1(t), u(x^1(t))) \right\} \leq \lambda(u) \tag{31}$$

Assume now that $\lambda < \lambda(u)$. For $z \in C_1$ we have

$$\hat{E}_z \left\{ \sum_{t=1}^{\tau_{C_1}} \Big(\gamma c(x^1(t), u(x^1(t))) - \lambda \Big) \right\} > 0$$

and consequently for $N \geq N_0$

$$\hat{E}_z \left\{ \sum_{t=1}^{\tau_{C_1} \wedge N} \Big(c(x^1(t), u(x^1(t))) - \lambda \Big) \right\} \geq 0. \tag{32}$$

Therefore for w_N^u defined as in (27) similarly to (28)-(29) we have

$$w_{N+1}^u(x) \leq \hat{E}_x \left\{ c(x^1(0), u(x^1(0))) - \lambda + w_N^u(\hat{x}(1)) \right\} \tag{33}$$

and by iteration for $N \geq N_0$

$$w_{N+k}^u(x) \leq \hat{E}_x \left\{ \sum_{t=0}^{k-1} \Big(c(x^1(t), u(x^1(t))) - \lambda \Big) + w_N^u(\hat{x}(k)) \right\}$$

$$\leq \hat{E}_x \left\{ \sum_{t=0}^{k-1} \Big(c(x^1(t), u(x^1(t))) - \lambda \Big) + \|c\| N \right\}.$$

Therefore

$$\frac{1}{k} \hat{E}_x \left\{ \sum_{t=0}^{k-1} c(x^1(t), u(x^1(t))) \right\} \geq -\frac{1}{k} \|c\| N$$

$$+ \frac{1}{k} \inf_N \hat{E}_x \left\{ \sum_{t=1}^{\sigma_{C_1} \wedge N - 1} \Big(c(x^1(t), u(x^1(t))) - \lambda(u) \Big) \right\} + \lambda$$

and

$$\liminf_{k \to \infty} \frac{1}{k} \hat{E}_x \left\{ \sum_{t=0}^{k-1} c(x^1(t), u(x^1(t))) \right\} \geq \lambda$$

and finally

$$\liminf_{k \to \infty} \frac{1}{k} \hat{E}_x \left\{ \sum_{t=0}^{k-1} c(x^1(t), u(x^1(t))) \right\} \geq \lambda(u) \tag{34}$$

which together with (31) completes the proof.

□

We summarize the results of this section in the form of Theorem

Theorem 1 *There exist unique up to an additive constant function $w : R^k \mapsto R$ and unique constant λ which are solutions to the Bellman equation (7). Furthermore λ is an optimal value of the cost functional J^0.*

Proof. Notice for $\hat{u}$ we find w and λ as a solution to APE

$$w(x) = c(x, \hat{u}(x)) - \lambda + \int_{R^k} w(y)P(x, dy),$$

which exist by Lemma 1, 2 and Corollary 2. By Proposition 2, λ is an optimal value of the cost functional J^0. Uniqueness up to an additive constant of w follows from uniqueness of the solutions to APE for the split Markov process (Lemma 2) and Corollary 2.

□

3 Risk sensitive asymptotics

In what follows we shall assume that $\gamma \in (-1, 0)$. The following estimation will be useful in this section

Lemma 3 *We have*

$$e^{\gamma\|a\|} \leq E\left\{\left(\sum_{i=1}^{m} h_i \xi_i(x, W(0))\right)^{\gamma}\right\} \leq e^{|\gamma|\|a\| + \frac{1}{2}\gamma^2\|\sigma^2\|} \tag{35}$$

Proof. Since $r(z) = z^\gamma$ is convex by Jensen inequality we have

$$E\left\{\left(\sum_{i=1}^{m} h_i \xi_i(x, W(0))\right)^{\gamma}\right\} \leq \sum_{i=1}^{m} h_i E\left\{(\xi_i(x, W(0)))^{\gamma}\right\}.$$

Using Hölder inequality twice we have

$$E\left\{\left(\sum_{i=1}^{m} h_i \xi_i(x, W(0))\right)^{\gamma}\right\} \geq \frac{1}{E\left\{\left(\sum_{i=1}^{m} h_i \xi_i(x, W(0))\right)^{-\gamma}\right\}}$$
$$\geq \frac{1}{\left(\sum_{i=1}^{m} h_i E\left\{\left(\sum_{i=1}^{m} \xi_i(x, W(0))\right)\right\}\right)^{-\gamma}}.$$

Then using standard estimations for ξ_i we easily obtain (35).

□

Immediately from Lemma 3 we have

Corollary 3

$$\limsup_{\gamma\to 0} \sup_{x\in R^k} \sup_{h\in U} |E\left\{\left(\sum_{i=1}^{m} h_i\xi_i(x,W(0))\right)^{\gamma}\right\} - 1| = 0 \tag{36}$$

and

$$\lim_{\gamma\to 0} \sup_{x\in R^k} \sup_{h\in U} |c_\gamma(x,h) - 1| = 0. \tag{37}$$

We furthermore have

Lemma 4

$$\lim_{\gamma\to 0} \frac{1}{\gamma} c_\gamma(x,h) = c(x,h) \tag{38}$$

and the limit is increasing and uniform in x and h from compact subsets.

Proof. By Hölder inequality $\frac{1}{\gamma}c_\gamma(x,h)$ is increasing in γ. Using l'Hospital rule for $\gamma \to 0$ we identify the limit as $c(x,h)$. Since the functions $c(x,h)$ and $c_\gamma(x,h)$ are continuous by Dini's theorem the convergence is uniform on compact sets.

□

Lemma 5 *We have that*

$$\sup_{A\in\mathcal{B}(R^k)} \sup_{x\in R^k} \sup_{h\in U} |\frac{P^{h,\gamma}(x,A)}{P(x,A)} - 1| \to 0 \tag{39}$$

as $\gamma \to 0$.

Proof. Notice that by Hölder inequality we have

$$P^{h,\gamma}(x,A) \le e^{-c_\gamma(x,h)} e^{\frac{1}{2}c_{2\gamma}(x,h)} \sqrt{P(x,A)} \tag{40}$$

and

$$P(x,A) \le e^{\frac{1}{2}c_\gamma(x,h)} e^{-\frac{1}{2}\gamma\|a\|} \sqrt{P^{h,\gamma}(x,A)} \tag{41}$$

from which (39) easily follows.

□

In what follows we shall assume that for some $\gamma < 0$ we have

$$E_x\left\{e^{|\gamma|\tau_C}\right\} < \infty \tag{42}$$

for $x \in R^k$ and

$$\sup_{x\in C} E_x\left\{e^{|\gamma|\tau_C}\right\} < \infty. \tag{43}$$

where C is the same compact set as in section 2.
We recall the following fundamental result from [8]

Theorem 2 *For $\gamma < 0$ sufficiently close to 0 exist λ^γ and a continuous function $w_\gamma : R^k \mapsto R$ such that Bellman equation (10) is satisfied. Moreover $\frac{1}{\gamma}\lambda^\gamma$ is an optimal value of the cost functional J_x^γ and the control $\hat{u}(x_n)$, where $\hat{u}$ is a Borel measurable function for which the infimum in the right hand side of (10) is attained, is an optimal control within the class of all of controls from $\mathcal{U}_s$.*
Furthermore, if for admissible control (h_n) we have that

$$\limsup_{t\to\infty} E_x^{(h_n)} \left\{ \left(E_{x_t}^{h_t} \left\{ e^{w_\gamma(x_1)} \right\} \right)^\alpha \right\} < \infty$$

for every $\alpha > 1$, then $\frac{1}{\gamma}\lambda^\gamma \leq J_x^\gamma((a_n))$.

Notice now that by Hölder inequality the value of the functional J^γ is increasing in $\gamma < 0$ and by Jensen inequality is dominated by the value of J^0. Consequently the same holds for the optimal values of the cost functionals i.e.

$$\frac{1}{\gamma}\lambda_\gamma \leq \lambda. \tag{44}$$

Furthermore there is a sequence u_n of continuous functions from R^k to U such that $c(x, u_n(x))$ converges uniformly in x from compact subsets to $\sup_{h\in U} c(x,h)$. By Lemma 1 and Theorem 1 we immediately have that $\lambda((u_n)) \to \lambda$ as $n \to \infty$. This means that for any $\varepsilon > 0$ there is a ε optimal continuous control function u_ε. We are going to show that for each $\varepsilon > 0$

$$J^\gamma(u_\varepsilon(x(n))) \to J^0(u_\varepsilon(x(n))) \tag{45}$$

as $\gamma \to 0$. Since the proof will be based following section 5 of [6] on the large deviation estimates we shall need the following assumption:

(A) there is a continuous function $f_0 : R^k \mapsto [1, \infty)$ such that for each positive integer n the set $K_n := \left\{ x \in R^k : \frac{f_0(x)}{Pf_0(x)} \leq n \right\}$ is compact.

Remark 2 *By direct calculation one can show that for a large class of ergodic processes $(x(n))$ function $f_0(x) = e^{c\|x\|^2}$ satisfies (A) for small c. To be more precise assume for simplicity that $k = 1$ and $|x + b(x)| \leq \beta|x|$ for a sufficiently large x with $0 < \beta < 1$. Then for $0 < c < \frac{1-\beta^2}{2dd^T}$ the assumption (A) holds.*

Proposition 3 *Under (A) for continuous control function $u : R^k \mapsto U$ we have*

$$J^\gamma(u(x(n))) \to J^0(u(x(n))) \tag{46}$$

as $\gamma \to 0$.

Proof. Under (A) using Lemma 5 we see that the set $K_n^{u,\gamma} := \left\{x \in R^k : \frac{f_0(x)}{P^{u,\gamma} f_0(x)} \le n\right\}$ is compact for each n. Therefore by Theorem 4.4 of [10] we have an upper large deviation estimate for empirical distributions of Markov process with transition operator $P^{u(x),\gamma}(x,\cdot)$. Using Theorem from Section 3 of [11] we also have a lower large deviation estimate. Consequently we have a large deviation principle satisfied with the rate function

$$I^{u,\gamma}(\nu) := \sup_{h\in H} \int_{R^k} \ln \frac{h(x)}{P^{u(x),\gamma} h(x)} \nu(dx), \tag{47}$$

where H is the set of all bounded functions $h : R^k \mapsto R$ such that $\frac{1}{h(x)}$ is also bounded and $\nu \in \mathcal{P}(R^k)$. By Varadhan theorem (Theorem 2.1.1 of [5]) we therefore obtain that

$$\frac{1}{\gamma} \lim_{n\to\infty} \frac{1}{n} \ln E_x^{h,\gamma} \left\{ \exp\left\{ \sum_{t=0}^{n-1} c_\gamma(x(t), h(t)) \right\} \right\} = \inf_{\nu \in \mathcal{P}(R^k)} \left(\int_{R^k} \frac{1}{\gamma} c_\gamma(z, u(z)) \nu(dz) - \frac{1}{\gamma} I^{u,\gamma}(\nu) \right) \tag{48}$$

There is a sequence of measures ν_{γ_i} with $\gamma_i \to 0$ as $i \to \infty$ such that

$$\begin{aligned} &\int_{R^k} \frac{1}{\gamma_i} c_{\gamma_i}(z, u(z)) \nu_{g_i}(dz) - \frac{1}{\gamma_i} I^{u,\gamma_i}(\nu_{\gamma_i}) \le \\ &\inf_{\nu \in \mathcal{P}(R^k)} \left(\int_{R^k} \frac{1}{\gamma_i} c_{\gamma_i}(z, u(z)) \nu(dz) - \frac{1}{\gamma_i} I^{u,\gamma_i}(\nu) \right) + \frac{1}{i} \end{aligned} \tag{49}$$

Since from (35)

$$\frac{1}{\gamma} \lim_{n\to\infty} \frac{1}{n} \ln E_x^{h,\gamma} \left\{ \exp\left\{ \sum_{t=0}^{n-1} c_\gamma(x(t), h(t)) \right\} \right\} \le \|a\| \tag{50}$$

we have that $I^{u,\gamma_i}(\nu_{\gamma_i}) \to 0$. We shall show that the sequence (ν_{γ_i}) is tight. Using Fatou Lemma to the sequence $\{f_0 \wedge N\}$ with $N \to \infty$ we obtain that

$$\int_{R^k} \ln \frac{f_0(x)}{P^{u(x),\gamma} f_0(x)} \nu_{\gamma_i}(dx) \le I^{u,\gamma_i}(\nu_{\gamma_i}). \tag{51}$$

By (39) for $\varepsilon > 0$ there is γ_0 such that for $\gamma \ge \gamma_0$

$$(1-\varepsilon) P f_0(x) \le P^{u(x),\gamma} f_0(x) \le (1+\varepsilon) P f_0(x). \tag{52}$$

Therefore by (51)

$$\int_{R^k} \ln \frac{f_0(x)}{P f_0(x)} \nu_{\gamma_i}(dx) \le I^{u,\gamma_i}(\nu_{\gamma_i}) + \ln(1+\varepsilon) \tag{53}$$

for $i > i_0$. Let $\rho_n := \inf_{x \in K_n} \ln \frac{f_0(x)}{Pf_0(x)}$. Then

$$\rho_n \nu_{\gamma_i}(K_n) + \ln n \nu_{\gamma_i}(K_n^c) \leq I^{u,\gamma_i}(\nu_{\gamma_i}) + \ln(1+\varepsilon) \tag{54}$$

where $K_n^c := R^k \setminus K_n$. Consequently

$$\ln n \nu_{\gamma_i}(K_n^c) \leq \frac{I^{u,\gamma_i}(\nu_{\gamma_i}) + \ln(1+\varepsilon) - \rho_n}{\ln n - \rho_n} \tag{55}$$

and since $\ln n \geq 1 + \rho_n$ for sufficiently large n, we have tightness of the measures ν_{γ_i}. By Prokhorov theorem there is a subsequence of ν_{γ_i}, for simplicity denoted by ν_{γ_i} and a probability measure $\bar{\nu}$ such that $\nu_{\gamma_i} \to \nu$ as $i \to \infty$. Since by (39) $I^{u,\gamma}(\nu)$ converges uniformly to $I^u(\nu) := \sup_{h \in H} \int_{R^k} \ln \frac{h(x)}{P^{u(x)}h(x)} \nu(dx)$ as $\gamma \to 0$ and I^u is nonnegative lowersemicontinous function we therefore have that $I^u(\bar{\nu}) = 0$. By Lemma 2.5 of [9] the measure $\bar{\nu}$ is invariant for the transition operator $P(x, \cdot)$. Therefore by Lemma 4

$$\begin{aligned}
&\lim_{i\to\infty} \frac{1}{\gamma_i} \lim_{n\to\infty} \frac{1}{n} \ln E_x^{h,\gamma_i} \left\{ \exp\left\{ \sum_{t=0}^{n-1} c_{\gamma_i}(x(t), h(t)) \right\} \right\} \geq \\
&\lim_{i\to\infty} \int_{R^k} \frac{1}{\gamma_i} c_{\gamma_i}(z, u(z)) \nu_{\gamma_i} = \int_{R^k} c(z, u(z)) \bar{\nu}(dz) = J^0(u(x(n)))
\end{aligned} \tag{56}$$

and using the fact that the cost functional J^γ is increasing in γ we obtain (46), which completes the proof.

$\square$

We are now in position to summarize the results of this section

Theorem 3 *Under (A) a continuous ε control function u_ε for J is also 2ε optimal control function for J^γ provided $0 > \gamma > \gamma_0$. Consequently convergence (12) holds.*

Remark 3 *On can expect that the convergence at least of a subsequence $\frac{1}{\gamma} w_\gamma(x)$ uniformly on compact subsets, as $\gamma \to 0$ to $w(x)$ holds, where w is a solution to the risk neutral Bellman equation (7). The authors unfortunately were not able to show it.*

References

[1] T.R. Bielecki, S. Pliska, "Risk sensitive dynamic asset management", *JAMO*, **39** (1999), pp. 337-360.

[2] V.S. Borkar, S.P. Meyn, "Risk-Sensitive Optimal Control for Markov Decision Processes with Monotone Cost", *Math. Oper. Res.*, **27** (2002), pp. 192-209.

[3] R. Cavazos-Cadena, “Solution to the risk-sensitive average cost optimality in a class of Markov decision processes with finite state space”, *Math. Meth. Oper. Res.*, **57** (2003), pp. 263-285.

[4] R. Cavazos-Cadena, D. Hernandez-Hernandez, “Solution to the risk-sensitive average optimality equation in communicating Markov decision chains with finite state space: An alternative approach”, *Math. Meth. Oper. Res.*, **56** (2002), pp. 473-479.

[5] J.D. Deuschel, D.W. Stroock, “Large Deviations”, Academic Press 1989,

[6] G.B. Di Masi, L. Stettner, “Risk sensitive control of discrete time Markov processes with infinite horizon”, *SIAM J. Control Optimiz.*, **38** (2000), pp. 61-78.

[7] G.B. Di Masi, L. Stettner, “Infinite horizon risk sensitive control of discrete time Markov processes with small risk, *Sys. Control Lett*, **40** (2000), pp. 305-321.

[8] G.B. Di Masi, L. Stettner, “Infinite horizon risk sensitive control of discrete time Markov processes under minorization property”, submitted.

[9] M.D. Donsker, S.R.S. Varadhan, “Asymptotic Evaluation of Certain Markov Process Expectations for Large Time - I”, *Comm. Pure Appl. Math.*, **28** (1975), 1-47.

[10] M.D. Donsker, S.R.S. Varadhan, “Asymptotic Evaluation of Certain Markov Process Expectations for Large Time - III”, *Comm. Pure Appl. Math.*, **29** (1976), 389-461.

[11] M. Duflo, “Formule de Chernoff pour des chaines de Markov”, in Grandes deviations et Applications Statistiques, *Asterisque* **68**, (1979), 99-124.

[12] W.H. Fleming, D. Hernandez-Hernandez, “Risk sensitive control of finite state machines on an infinite horizon”, *SIAM J. Control Optimiz.*, **35** (1997), pp. 1790–1810.

[13] D. Hernandez-Hernandez, S.J. Marcus, “Risk sensitive control of Markov processes in countable state space”, *Sys. Control Letters*, **29** (1996), pp. 147–155.

[14] I. Kontoyiannis, S.P. Meyn, “Spectral Theory and Limit Theorems for Geometrically Ergodic Markov Processes”, *Ann. Appl. Prob.* **13** (2003), 304-362.

[15] S.P. Meyn, R.L. Tweedie, “Markov Chains and Stochastic Stability”, Springer 1996

[16] L. Stettner, “Risk sensitive portfolio optimization”, *Math. Meth. Oper. Res.* **50** (1999), 463-474.