A VALUE ITERATION IN CONTROLLED DIFFUSION PROCESSES ASSOCIATED WITH STOPPING

Yoshida, Yuji
Electron Optics Technical and Engineering Division, Japan Electron Optics Laboratory Ltd.

https://doi.org/10.5109/13377

出版情報：Bulletin of informatics and cybernetics. 22 (1/2), pp.59-69, 1986-03. Research Association of Statistical Sciences
バージョン：
権利関係：

Bulletin of Informatics and Cybernetics Vol. 22, No. 1~2, 1986

A VALUE ITERATION IN CONTROLLED DIFFUSION PROCESSES ASSOCIATED WITH STOPPING

By

Yuji YOSHIDA*

Abstract

The present paper deals with an optimal control problem in controlled diffusion processes with stopping times.

In this paper we shall present a value iteration for the optimization problem associated with controls and stopping times. Furthermore we shall investigate the relations among the value interation, optimal stopping times and Bellman's equation.

0. Introduction

Stopped decision problems in discrete time case have been studied by Dubins and Savage [2], Furukawa [3, 4], Hordijk [5] and etc.. On the other hand Krylov [6], Nisio [7], Ohtsubo [8] and etc. have studied optimization problems in controlled Markov processes with stopping. Especially [4] has studied a value iteration to find an optimal reward function in decision problems. And Doshi [1] has treated it in controlled Markov processes. This paper investigates the property of the iteration and its relation to Bellman's equation in the case of controlled diffusion processes with stopping times in infinite horizon.

In Section 1 we shall introduce diffusion processes associated with stochastic differential equations. In Section 2 we shall consider an optimal control problem associated with stopping times. In Section 3 we shall give and study a value iteration to find solutions of the problem. In Section 4 we shall investigate the relation between the value iteration and Bellman's equation.

1. Controlled Diffusion Processes

Let R_+ be the set of all nonnegative real numbers, the time space. Let R^d be d-dimensional Euclidean space and let $E=R_+\times R^d$. $\mathcal{B}$ denotes the field of Borel subsets of E. Let Ω be the set of all continuous mappings $\omega: R_+\to R^d$. A mapping $x(t)$ is given by $\omega\to x(t,\omega)=\omega(t)$ for $t\in R_+$ and $\omega\in\Omega$. $\mathcal{F}$ denotes the smallest σ-algebra generated by $\{x(t): t\in R_+\}$. Let P be a probability measure on $(\Omega, \mathcal{F})$. G is a compact subset of a separable metric space, the action space.

Now we consider a controlled stochastic differential equation. For $(\gamma, t, x)\in G\times E$,

* Electron Optics Technical and Engineering Division, Japan Electron Optics Laboratry Ltd.- 1418 Nakagami Akishima, Tokyo 196, Japan.

$\sigma(\gamma, t, x)(b(\gamma, t, x))$ is a $d\times d$-dimentional matrix (resp. a d-dimensional vector) for which there exist numbers $\delta>0$, $K>0$ and a $d\times d$-dimensional matrix valued measurable function $\tilde{C}$ on R_+ satisfying (i)$\sim$(iv):

(i) σ and b are bounded and continuous.

(ii) $(\sigma(\gamma, t, x)\xi, \xi)\geqq\delta|\xi|_d^2$ for all $\xi\in R^d$ and $(\gamma, t, x)\in G\times E$,

where $(y, z)=\sum_{i=1}^{d} y_i z_i$ and $|y|_d=(y, y)^{1/2}$ for $y=(y_i)\in R^d$ and $z=(z_i)\in R^d$.

(iii) $\delta^2|\xi|_d^2\leqq(\xi, \tilde{C}(t)\xi)\leqq K^2|\xi|_d^2$ for all $\xi\in R^d$ and $t\in R_+$.

(iv) $\|a(\gamma, t, x)-\tilde{C}(t)\|\leqq K$ for all $(\gamma, t, x)\in G\times E$, where $\|\cdot\|$ denotes the operator norm and $a(\gamma, t, x)=\sigma(\gamma, t, x)\sigma^*(\gamma, t, x)$.

REMARK. Conditions (iii) and (iv) are from Strook and Varadhan [10] and are clearly satisfied in the time-homogeneous case.

Let U be the set of all measurable mappings $u: E\to G$. In this paper we treat only Markov controls, therefore we shall call them only controls here.

The following existence theorem of controlled processes is due to Theorem 2.6.1 of Krylov [6].

LEMMA 1. *For each $u\in U$ and $(s, x)\in E$ there exist a probability space $(\Omega^u, \mathcal{F}^u, P^{u,s,x})$, a Wiener process $(W^u(t), \mathcal{F}_t^u)_{t\geqq s}$ and a continuous process $(X_t^u)_{t\geqq s}$ such that X_t^u is progressively measurable with respect to $\{\mathcal{F}_t^u\}$ and $(X_t^u)_{t\geqq s}$ satisfies* (1):

$$X_t^u=x+\int_s^t\sigma(u(r, X_r^u), r, X_r^u)dW^u(r)+\int_s^t b(u(r, X_r^u), r, X_r^u)dr \qquad (1)$$

almost surely for all $t\geqq s$.

Let $Z_t^u=(t, X_t^u)$ for $t\geqq s$. Then for each $(s, x)\in E$, $(Z_t^u, \mathcal{F}_t^u, P^{u,s,x})_{t\geqq s}$ is a continuous process.

D_A denotes the space of bounded continuous functions g on E such that $g(t, x)$ has the first order derivative with respect to the time t and has the first and the second order derivatives with respect to the space x, which are bounded and continuous. For $u\in U$, A^u is an operator such that

$$A^u k(t, x)=\frac{\partial}{\partial t}k(t, x)+\frac{1}{2}\sum_{i,j=1}^{d} a_{i,j}(u(t, x), t, x)\frac{\partial^2}{\partial x_i\partial x_j}k(t, x)$$

$$+\sum_{i=1}^{d} b_i(u(t, x), t, x)\frac{\partial}{\partial x_i}k(t, x)$$

for $k\in D_A$ and $(t, x)\in E$. Especially, if u is a constnat mapping $u=\gamma$ for some $\gamma\in G$, then we express A^u by A^γ.

Then for each $u\in U$ a process $(Z_t^u)_{t\geqq s}$ satisfies (2):

$$\{g(Z_t^u)-\int_s^t A^u g(Z_r^u)dr, \mathcal{F}_t^u, P^{u,s,x}\}_{t\geqq s} \text{ is a martingale for } (s, x)\in E \text{ and } g\in D_A. \qquad (2)$$

This is due to Theorem 4.2.1 and Corollary 5.3.1 in Strook and Varadhan [10].

On the other hand from Corollary 7.1.7 and Theorem 6.2.2. in [10] and conditions (i)$\sim$(iv), $\{P^{u,s,x}: (s, x)\in E\}$ is a unique solution of the martingale problem for (2), is

measurable and has strong Markov property. Consequently $(Z_t^u, \mathcal{F}_t^u, P^{u,s,x})_{t\geq s}$ is a continuous strong Markov process.

Let I be the identity operator on D_A. Then we can easily check that for a number $\beta\geqq 0$ and $u\in U$, a process $(Z_t^u)_{t\geq s}$ also satisfies (2′):

$$\{e^{-\beta(t-s)}g(Z_t^u)-\int_s^t e^{-\beta(r-s)}(A^u-\beta I)g(Z_r^u)dr,\ \mathcal{F}_t^u,\ P^{u,s,x}\}_{t\geq s} \quad (2')$$
is a martingale for $(s, x)\in R_+\times R^d$ and $g\in D_A$.

2. Optimization Problems

Notations defined in this section will be used through this paper. We shall fix a number $\alpha>0$, which is called a discounted rate. For $u\in U$ and $(s, x)\in E$, $\mathcal{M}^u$ denotes the set of all $\{\mathcal{F}_t^u\}$-adapted Markov times and $\mathcal{M}^{u,s,x}=\{\tau\in\mathcal{M}^u: P^{u,s,x}(\tau\geqq s)=1\}$. Let $b\mathcal{B}$ be the set of all bounded Borel measurable functions on E. For $u\in U$, $\{Q_t^{u,\beta}\}_{t\geq 0}$ denotes a semigroup of operators on $b\mathcal{B}$ which is defined by

$$Q_t^{u,\beta}g(s, x)=E^{u,s,x}[e^{-\beta(t-s)}g(Z_t^u)]$$

for $\beta\geqq 0$, $t\in R_+$, $g\in b\mathcal{B}$ and $(s, x)\in E$. Hence from (2′) of Section 1 we have

$$E^{u,s,x}\left[\int_s^\tau e^{-\beta(r-s)}(A^u-\beta I)g(Z_r^u)dr\right]=Q_\tau^{u,\beta}g(s, x)-g(s, x) \quad (3)$$
for $\beta>0$, $u\in U$, $(s, x)\in E$, $g\in D_A$ and $\tau\in\mathcal{M}^{u,s,x}$.

From now on we write $A^u-\alpha I$ by B^u for simplicity and we put C be a bounded measurable function. We consider the following problem (4):

Find an control $u\in U$ and $\tau\in\mathcal{M}^{u,s,x}$ which maximizing
$$E^{u,s,x}\left[e^{-\alpha(\tau-s)}f(Z_\tau^u)+\int_s^\tau e^{-\alpha(r-s)}C(Z_r^u)dr\right] \text{ for each } (s, x)\in E. \quad (4)$$

We use the knowledge concerning optimal stopping problem to investigate Problem (4). For each control $u\in U$ we define a semigroup $\{S_t^u\}$ of operators on $b\mathcal{B}$ as follows:

$$S_t^u g(s, x)=E^{u,s,x}\left[e^{-\alpha(t-s)}g(Z_t^u)+\int_s^t e^{-\alpha(r-s)}C(Z_r^u)dr\right] \text{ for } t\geqq s \text{ and } x\in R^d.$$

For each $u\in U$ an optimal stopping reward function V^u is defined by $V^u(s, x)=\sup S_\tau^u f(s, x)$ for $(s, x)\in E$, where the supremum is taken over all $\tau\in\mathcal{M}^{u,s,x}$. We fix $f\in b\mathcal{B}$.

LEMMA 2. (Shiryayev [9]) *Let $u\in U$ be arbitrary but fixed. V^u has properties (i) and (ii):*

(i) *V^u is the smallest excessive majorant of f with respect to $\{S_t^u\}_{t\geq 0}$.*
(ii) *For $s\in R_+$ we define a Markov time $\tau\in\mathcal{M}^u$ by*

$$\tau=\inf\{t\geqq s: Z_t^u\in\{V^u=f\}\}.$$

Then it holds that

$$V^u(s, x)=S_\tau^u V^u(s, x)=S_\tau^u f(s, x) \quad \text{for } x\in R^d.$$

DEFINITION. (i) $V^*(s, x)=\sup_{u\in U} V^u(s, x)$, $(s, x)\in E$, is called an optimal reward function for Problem (4).

(ii) A control $u^*\in U$ is called an optimal control for Problem (4), if $V^*(s, x)=V^{u^*}(s, x)$ holds for all $(s, x)\in E$.

(iii) Let an optimal control u^* and $(s, x)\in E$ be fixed. A Markov time τ^* is an optimal for Problem (4), if $P^{u^*, s, x}(\tau^*\geqq s)=1$ and $V^*(s, x)=S^{u^*}_{\tau^*}f(s, x)$ for all $(s, x)\in E$ and the optimal control u^*.

3. A Value Iteration

First we suppose the following assumption throughout this paper

ASSUMPTION (I). It holds that $V^u\in D_A$ for each $u\in U$.

Then we consider the following iteration. Take an arbitrary control $u^1\in U$. For each natural number n, take a control $u^n\in U$ and a function V^n in the following iterative manner.

Step (1.n). For a control $u^n\in U$ we find a function $V^n\in D_A$ which satisfied (i), (ii) and (iii):

(i) $V^n\geqq f$ everywhere.

(ii) $B^{u^n}V^n+C\leqq 0$ a.e. on E.

(iii) $B^{u^n}V^n+C=0$ a.e. on $\{V^n>f\}$.

And we go to Step (2.n).

Step (2.n). (a) If there exists $(s, x)\in E$ such that $\sup_{\gamma\in G} B^\gamma V^n(s, x)+C(s, x)>0$, then we select a control $u^{n+1}\in U$ such that

$$B^{u^{n+1}}V^n(s, x)=\sup_{\gamma\in G} B^\gamma V^n(s, x)$$

for all $(s, x)\in E$. And we return to Step (1.n+1).

(b) If $\sup_{\gamma\in G} B^\gamma V^n(s, x)+C(s, x)\leqq 0$ for all $(s, x)\in E$, then we stop this iteration.

The existence of a control u^{n+1} in Step (2.n) is due to Lemma 3.

LEMMA 3. *For each $g\in D_A$ there exists a control $u\in U$ such that $B^u g(s, x)=\sup_{\gamma\in G} B^\gamma g(s, x)$ for all $(s, x)\in E$.*

PROOF. $B^\gamma g(s, x)$ is continuous with respect to $(\gamma, s, x)\in G\times E$, therefore this lemma follows from Borel selection theorems.

Next the existence of a function V^n in Step (1.n) is due to Lemma 4.

LEMMA 4. *Let $u\in U$ be arbitrary but fixed. The* the *following free boundary problem* (5) *has a unique solution V^u:*

Find a function $\overline{V}\in D_A$ which satisfies (i), (ii) and (iii):

(i) $\overline{V}\geqq f$ *everywhere.*

(ii) $B^u\overline{V}+C\leqq 0$ *a.e. on* E. (5)

(*iii*) $B^u\overline{V}+C=0$ *a.e. on* $\{\overline{V}>f\}$.

PROOF. Let an arbitrary $u\in U$ be fixed. It is sufficient for the existence of $\overline{V}$ to prove that V^u satisfies (ii) and (iii), because of the definition of V^u. From Lemma 2(i) we have $V^u(s,x)\geqq S_\tau^u V^u(s,x)$ for each $(s,x)\in E$ and $\tau\in\mathcal{M}^{u,s,x}$. Then using (3), we obtain

$$E^{u,s,x}\left[\int_s^\tau e^{-\alpha(r-s)}(B^uV^u(Z_r^u)+C(Z_r^u))dr\right]\leqq 0.$$

Namely the set $\{B^uV^u+C>0\}$ is of potential zero with respect to (Z_t^u). Therefore for all number $\lambda>0$, $(s,x)\in E$ and $\tau\in\mathcal{M}^{u,s,x}$ it holds that

$$E^{u,s,x}\left[\int_s^\tau e^{-(\alpha+\lambda)(r-s)}(B^uV^u(Z_r^u)+C(Z_r^u))dr\right]\leqq 0.$$

Now we fix a bounded open subset D of E and we let $(s,x)\in D$ be arbitrary but fixed. Then we put τ be the first exit time from D after starting at (s,x). Hence from Theorem 2.4.6(a) of Krylov [6] there exists a sequeuce $\{\lambda_n\}_n$ in positive numbers which satisfies $\lambda_n\to\infty(n\to\infty)$ and

$$\lambda_n E^{u,s,x}\left[\int_s^\tau e^{-(\alpha+\lambda_n)(r-s)}(B^uV^u(Z_r^u)+C(Z_r^u))dr\right]$$
$$\to B^uV^u(s,x)+C(s,x)\qquad \text{a.a.}(s,x)\in D \text{ as } n\to\infty.$$

Consequently we have $B^uV^u+C\leqq 0$ a.e. on D. Since D is arbitrary bounded open, V^u satisfies the contition (ii). Next we let D be a bounded open subset of $\{V^u>f\}$ and we take $(s,x)\in D$ and $\tau\in\mathcal{M}^{u,s,x}$ similarly. Then from Lemma 2 we have

$$E^{u,s,x}\left[\int_s^\tau e^{-\alpha(r-s)}(B^uV^u+C)(Z_r^u)dr\right]=S_\tau^uV^u(s,x)-V^u(s,x)=0.$$

Therefore from Corollary 2.5.8 of [6] we have $B^uV^u+C=0$ a.e. on D. Since D is arbitrary bounded open subset of $\{V^u>f\}$, V^u also satisfies the condition (iii).

Next we show the uniqueness of solutions of (5). Let $\overline{V}$ be a solution of (5). For $(s,x)\in E$ we define a Markov time σ by

$$\sigma=\inf\{t\geqq s: Z_t^u\in\{\overline{V}=f\}\}$$

Then, by using the condition (iii) for $\overline{V}$ and Theorem 2.2.4 of [6], we have

$$S_\sigma^u\overline{V}(s,x)-\overline{V}(s,x)=E^{u,s,x}\left[\int_s^\sigma e^{-\alpha(r-s)}(B^u\overline{V}+C)(Z_r^u)dr\right]=0.$$

Therefore it holds that

$$\overline{V}(s,x)=S_\sigma^u\overline{V}(s,x)=S_\sigma^uf(s,x)\leqq V^u(s,x)$$

for all $(s,x)\in E$. On the other hand, by using the condition (ii) for $\overline{V}$ and Theorem 2.2.4 of [6], we have $\overline{V}(s,x)\geqq S_\tau^u\overline{V}(s,x)$ for $(s,x)\in E$ and $\tau\in\mathcal{M}^{u,s,x}$. Since $\overline{V}\geqq f$, it holds that $\overline{V}(s,x)\geqq S_\tau^u\overline{V}(s,x)\geqq S_\tau^uf(s,x)$ for each $(s,x)\in E$ and $\tau\in\mathcal{M}^{u,s,x}$. Namely we have $\overline{V}\geqq V^u$ everywhere. Consequently V^u is a unique solution of (5). This completes Lemma 4.

A sequence $\{V^n\}_n$, which is constructed in the iteration, has the following pro-

perty.

LEEMA 5. *$\{V^n\}_n$ is a nondecreasing sequence of functions on E.*

PROOF. Let n be arbitrary but fixed. From the definition of u^{n+1} in Step (2.n), we have

$$B^{u^{n+1}}V^n+C\geqq B^{u^n}V^u+C \text{ everywhere.}$$

Hence from Lemma 4 we also have

$$B^{u^{n+1}}V^n+C=0 \quad \text{a.e. on } \{V^n>f\}.$$

Since $B^{u^{n+1}}V^n+C$ is continuous from Lemma 3, it holds that

$$B^{u^{n+1}}V^n+C\geqq 0 \quad \text{on } \{V^n>f\}.$$

Let τ^n be a Mrkov time defined by

$$\tau^n=\inf\{t\geqq s: Z_t^{u^{n+1}}\in\{V^n=f\}\}$$

for $s\in R_+$. Then, by using (3), we obtain $V^n(s, x)\leqq S_{\tau^n}u^{n+1}V^n(s, x)=S_{\tau^n}u^{n+1}f(s, x)\leqq V^{n+1}(s, x)$ for all $(s, x)\in E$. Therefore $\{V^n\}_n$ is nondecreasing.

The following proposition is designed for Step (2.n) (b).

PROPOSITION 1. *If there exists a number $\bar{n}$ such that*

$$\sup_{\gamma\in G} B^\gamma V^{\bar{n}}(s, x)+C(s, x)\leqq 0$$

for all $(s, x)\in E$, then $V^{\bar{n}}$ is an optimal reward function and $u^{\bar{n}}$ is an optimal control.

PROOF. Let arbitrary $u\in U$ be fixed. We have $B^uV^{\bar{n}}+C\leqq 0$ everywhere. By using (3) and $V^{\bar{n}}\geqq f$, we obtain

$$V^{\bar{n}}(s, x)\geqq S_\tau^u V^{\bar{n}}(s, x)\geqq S_\tau^u f(s, x)$$

for each $(s, x)\in E$ and $\tau\in\mathcal{M}^{u, s, x}$. Therefore it holds that $V^{\bar{n}}\geqq V^u$ everywhere. Since $u\in U$ is arbitrary and $V^{\bar{n}}=V^{u^{\bar{n}}}$ from Lemman 4, V^n is an optimal reward function and $u^{\bar{n}}$ is an optimal control.

Now we consider a limit of $\{V^n\}_n$. Since f and C is bounded and $\alpha>0$, $\{V^n\}_n$ has upper bounds. Therefore from Lemma 5 we can define the limit funciton

$$V(s, x)=\lim_{n\to\infty} V^n(s, x) \quad \text{for each } (s, x)\in E.$$

Hence V is a lower semi-continuous function on E.

Hence we need one more assumption.

ASSUMPTION (II). Let any $i, j\in\{1, 2, \cdots, d\}$ and any compact subset K of E be fixed. Sequences $\left\{\frac{\partial^2}{\partial x_i\partial x_j}V^n\Big|_K\right\}_n$, $\left\{\frac{\partial}{\partial x_i}V^n\Big|_{\bar{K}}\right\}_n$ and $\left\{\frac{\partial}{\partial t}V^n\Big|_K\right\}_n$ are uniformly bounded where $\bar{K}=K\{(t, x)\in E: b(\gamma, t, x)\neq 0 \text{ for some } \gamma\in G\}$.

REMARK. In the time-homogeneous case when $d=1$, $b\equiv 0$ and $f\in D_A$, Assumption (II) is satisfied under Assumption (I). Because from Lemma 4 we can easily check

$$\left|\frac{d^2}{dx^2}V^n(x)\right|\leqq\frac{2}{\delta^2}\cdot(\alpha\cdot\|f\|+\|C\|)$$

for a.a. $x\in\{V^n>f\}$ and all n. Moreover since $V^n\in D_A$ and $f\in D_A$, we obtain that

$\left\{\frac{d^2}{dx^2}V^n\right\}_n$ is uniformaly bounded.

THEOREM 1. *Let Assumptions (I) and (II) be satisfied. Then V is an optimal reward function.*

PROOF. Let a compact subset K of E be arbitrary but fixed. From Assumptions (I) and (II) we can extract a subsequence $\{n'\}$ from the set of all natural numbers such that $\left\{\frac{\partial^2}{\partial x_i \partial x_j}V^{n'}\Big|_K\right\}_{n'}$, $\left\{\frac{\partial}{\partial x_i}V^{n'}\Big|_K\right\}_{n'}$ and $\left\{\frac{\partial}{\partial t}V^{n'}\Big|_K\right\}_{n'}$ are weakly convergent for all i, $j\in\{1, 2, \cdots, d\}$. Let a control $u\in U$ be arbitrary but fixed. Then a sequence $\{(B^uV^{n'}+C)|_K\}_{n'}$ converges pointwise to some function $g\in b\mathcal{B}$. Now we define $N_1=\{n'\}$. First we shall show $g<0$ a.e. on E. We consider the following sets and Markov times:

$$C_\beta{}^m=\bigcap_{\substack{k\geqq m\\ k\in N_1}}\{B^{u^{k+1}}V^k+C>\beta\}.$$

$$\sigma_\beta{}^{n,m}=\inf\{t\geqq s: Z_t{}^{u^{n+1}}\in C_\beta{}^m\},$$

$$\tau_\beta{}^n=\inf\{t\geqq s: Z_t{}^{u^{n+1}}\in\{B^{u^{n+1}}V^n+C\leqq\beta\}\}$$

for real number $\beta>0$, $s\in R_+$, n, $m\in N_1$ satisfying $n\geqq m$. From Theorem 2.2.4 of [6] and Lemma 2(3) we have $V^{n+1}(s, x)\geqq S^{u^{n+1}}_{\tau_\beta{}^n}V^{n+1}(s, x)$ for all $(s, x)\in E$, $n\in N_1$ and $\beta>0$. Then from Lemma 5 and (3) it holds that

$$\begin{aligned}
&V^{n+1}(s, x)-V^n(s, x)\\
&\geqq S^{u^{n+1}}_{\tau_\beta{}^n}V^n(s, x)-V^n(s, x)\\
&=E^{u^{n+1}, s, x}\left[\int_s^{\tau_\beta{}^n}e^{-\alpha(r-s)}(B^{u^{n+1}}V^n+C)(Z_r{}^{u^{n+1}})dr\right]\\
&\geqq\beta E^{u^{n+1}, s, x}\left[\int_s^{\tau_\beta{}^n}e^{-\alpha(r-s)}dr\right]\\
&\geqq\frac{\beta}{\alpha}E^{u^{n+1}, s, x}[1-e^{-\alpha(\tau_\beta{}^n-s)}]\\
&\geqq\frac{\beta}{\alpha}E^{u^{n+1}, s, x}[1-e^{-\alpha(\sigma_\beta{}^{n,m}-s)}]
\end{aligned}$$

for $(s, x)\in E$, $\beta>0$ and n, $m\in N_1(n\geqq m)$. By letting $n(\in N_1)$ infinite, we obtain

$$\lim_{\substack{n\in N_1\\ n\to\infty}}E^{u^{n+1}, s, x}[\sigma_\beta{}^{n,m}-s]=0$$

for $(s, x)\in E$, $\beta>0$ and $m\in N_1$. We also have and following inequalities. Let $\|\cdot\|_d$ be the matrix norm and define $\eta_1\wedge\eta_2=\min(\eta_1, \eta_2)$ for real numbers η_1 and η_2.

$$\begin{aligned}
&E^{u^{n+1}, s, x}[|Z^{u^{n+1}}_{\sigma^{n,m}_\beta\wedge T}-(s, x)|^2_{d+1}]\\
&\leqq 2\cdot E^{u^{n+1}, s, x}\left[\int_s^{\sigma_\beta{}^{n,m}\wedge T}\|\sigma(u^{m+1}(Z_r{}^{u^{n+1}}), Z_r{}^{u^{n+1}})\|^2_d dr\right]
\end{aligned}$$

$$+2\cdot E^{u^{n+1}s,x}\left[\int_s^{\tau_\beta^{n,m}\wedge T}|b(u^{n+1}(Z_r^{u^{n+1}}),\ Z_r^{u^{n+1}})|_d^2 dr\cdot((\sigma_\beta^{n,m}\wedge T)-s)\right]$$

$$\leqq 2K^2\{E^{u^{n+1},s,x}[(\sigma_\beta^{n,m}\wedge T)-s]+E^{u^{n+1},s,x}[((\sigma_\beta^{n,m}\wedge T)-s)^2]\}$$

for $(s,\ x)\in E$, $\beta>0$, numbers $T>s$ and $n,\ m\in N_1(n\geqq m)$. Therefore we obtain

$$\lim_{\substack{n\in N_1\\ n\to\infty}} E^{u^{n+1},s,x}[|Z_{\sigma_\beta^{n,m}}^{u^{n+1}}-(s,\ x)|_{d+1}^2]=0$$

for $(s,\ x)\in E$, $\beta>0$ and $m\in N_1$. From the definition of $\sigma_\beta^{n,m}$ and this equality we have that $E\subset\overline{(C_\beta^m)^c}$ for $\beta>0$ and $m\in N_1$, where $\overline{(C_\beta^m)^c}$ implies the closure of the complement of C_β^m in E. Moreover since $C_\beta^m\supset\bigcap_{\substack{k\geqq m\\ k\in N_1}}\{B^{u^{k+1}}V^k+C\geqq 2\beta\}$ for $\beta>0$ and $n\in N_1$, we obtain (6):

$$E\subset\bigcup_{\substack{k\geqq m\\ k\in N_1}}\{B^{u^{k+1}}V^k+C<2\beta\}\qquad\text{for } \beta>0 \text{ and } m\in N_1. \tag{6}$$

On the other hand we have

$$\{g>2\beta\}\subset\bigcup_{m\in N_1}\bigcap_{\substack{k\in N_1\\ k\geqq m}}\{B^{u^{k+1}}V^k+C\geqq 2\beta\}\cap K$$

for $\beta>0$. By putting $O_\beta^m=\bigcup_{\substack{k\geqq m\\ k\in N_1}}\{B^{u^{k+1}}V^k+C<2\beta\}$ for $\beta>0$ and $m\in N_1$, we obtain (7):

$$\{g>2\beta\}\subset\bigcup_{m\in N_1}(O_\beta^m)^c\cap K\qquad\text{for } \beta>0. \tag{7}$$

Since O_β^m is open, (6) and (7) imply that $g\leqq 2\beta$ a.e. on E. However since β is arbitrary and positive, we obtain $g\leqq 0$ a.e. on E.

Next let any $(s,\ x)\in K$ be fixed. We define a Markov time $\eta=\inf\{t\geqq s: Z_t^u\notin K\}$. From (3) we have

$$E^{u,s,x}\left[\int_s^{\eta\wedge\tau}e^{-\alpha(r-s)}(B^uV^n+C)(Z_r^u)dr\right]$$

$$=S_{\eta\wedge\tau}^uV^n(s,\ x)-V^n(s,\ x)$$

for all $n\in N_1$ and $\tau\in\mathcal{M}^{u,s,x}$. Since $g\leqq 0$ a.e., by using the bounded convergence theorem and Theorem 2.2.4 of [6], we obtain $V(s,\ x)\geqq S_{\eta\wedge\tau}^uV(s,\ x)$ for all $(s,\ x)\in E$ and $\tau\in\mathcal{M}^{u,s,x}$. Since K is arbitrary and $\alpha>0$, we have

$$V(s,\ x)\geqq S_\tau^uV(s,\ x)\geqq S_\tau^uf(s,\ x) \tag{8}$$

for all $(s,\ x)\in E$ and $\tau\in\mathcal{M}^{u,s,x}$. Therefore it holds that $V\geqq V^u$ everywhere for all $u\in U$. Therefore V is an optimal reward function.

A sequence $\{u^n\}$ of controls, which are constructed in the iteration, has the following property.

PROPOSITION 2. *Let Assumptions (I) and (II) be satisfied. If $\{u^n\}$ has a subsequence which converges almost surely to some $u^*\in U$, then u^* is an optimal control and $\tau^*=\inf\{t\geqq s: Z_t^{u^*}\in\{V=f\}\}$ is an optimal Markov time.*

PROOF. Let a compact subset K of E be arbitrary but fixed. Let $N_2=\{n'\}$ be a subsequence of all natural numbers for which $\{u^{n'}\}$ converges almost surely to u^*. In

the same line as the first part of proof of Theorem 1, we can extract a subsequence $N_3=\{u''\}$ of the set N_2 for which $\{(B^{u^*}V^{n'}+C)|_K\}_{n'}$ converges pointwise to some function $g\in b\mathcal{B}$. Then for each $n\in N_3$ and $(s, x)\in E$ we have

$$|B^{u^{n+1}}V^n(s, x)-B^{u^*}V^n(s, x)|$$

$$\leqq \sum_{i,j=1}^{d} |a_{i,j}(u^{n+1}(s, x), s, x)-a_{i,j}(u^*(s, x), s, x)|\cdot L$$

$$+\sum_{i=1}^{d} |b_i(u^{n+1}(s, x), s, x)-b_i(u^*(s, x), s, x)|\cdot L,$$

where L is a constant. Therefore we obtain (9):

$$g=\lim_{\substack{n\to\infty \\ n\in N_2}} (B^{u^{n+1}}V^n+C) \quad \text{a.e. on } E. \tag{9}$$

On the other hand from the difinition of u^{n+1} in the iteration we have

$$B^{u^{n+1}}V^n+C\geqq 0 \quad \text{on } \{V^n>f\}$$

for all n. Therefore from Lemma 5, (7) and this fact we have $g\geqq 0$ a.e. on $\{V>f\}$. In the same line as the first part of Theorem 1 we have

$$V(s, x)\leqq S^{u^*}_{\tau^*}V(s, x)=S^{u^*}_{\tau^*}f(s, x)$$

for all $(s, x)\in E$ and $\tau^*=\inf\{t\geqq s: Z_t^{u^*}\in\{V=f\}\}$. Consequently u^* is an optimal control and τ^* is an optimal Markov time.

4. Bellman's Equation

We need one more assumption to study the relation of the value iteration and Bellman's equation.

ASSUMPTION (III). $V\in D_A$.

THEOREM 2. *Let Assumptions (I), (II) and (III) be satisfied. An optimal reward function V, which is constructed in the iteration, satisfies the following equations (i), (ii) and (iii):*

(i) $V\geqq f$ *everywhere.*

(ii) $\sup_{\gamma\in G}(B^\gamma V+C)\leqq 0$ *everywhere.*

(iii) $\sup_{\gamma\in G}(B^\gamma V+C)=0$ on $\{V>f\}$.

PROOF. (i) is trivial. From the proof of Theorem 1 we have $S_\tau^\gamma V(s, x)\leqq V(s, x)$ for all $(s, x)\in E$, $\gamma\in G$ and $\tau\in\mathcal{M}^{u, s, x}$. Then we obtain (ii) in the same line as the first part of the proof of Lemma 4. Next from Lemma 3 we can choose a control $\bar{u}\in U$ such that

$$B^{\bar{u}}V+C=\sup_{\gamma\in G}(B^\gamma V+C) \text{ everywhere.}$$

Let a natural number m and a bounded open subset O of $\{V^m>f\}$ be fixed. Define a Markov time $\tau^n=\inf\{t\geqq s: Z_t^{u^n}\notin O\}$. By using Lemmas 2, 4, 5 and (8) we have

$$V^n(s, x)=S^{u^n}_{\tau^n}V^n(s, x)\leqq S^{u^n}_{\tau^n}V(s, x)\leqq V(s, x)$$

for $n \geqq m$ and $(s, x) \in O$. Therefore we obtain

$$\sup_{n \geqq m} \{S_\tau^{u^n} V^n(s, x) - V(s, x)\} = 0 \qquad \text{for all } (s, x) \in O.$$

Namely, it holds that

$$\sup_{n \geqq m} E^{u^n, s, x}\left[\int_s^{\tau^n} e^{-\alpha(r-s)}(B^{\bar{u}} V + C)(Z_r^{u^*}) dr\right] \geqq 0$$

for all $(s, x) \in O$. From this inequality and Assumption (III), by using Corollary 2.5.8 of [6], we have

$$B^{\bar{u}} V(s, x) + C(s, x) = 0 \qquad \text{for all } (s, x) \in O.$$

By considering the choice of O, we obtain (iii).

COROLLARY 1. *Let Assumptions (I), (II) and (III) be satisfied. Let u^* be a control which is defined in Proposition 2. Then it holds that*

$$B^{u^*} V + C = \sup_{\tau \in G} (B^\tau V + C) \qquad \text{a.e. on } \{V > f\}. \tag{10}$$

Conversely a control $u^* \in U$, which satisfies (10), is an optimal control and a Markov time $\tau^* = \inf\{t \geqq s : Z_t^{u^*} \in \{V = f\}\}$ is an optimal Markov time.

PROOF. From Proposition 2 we have $V^{u^*} = V$ everywhere. Therefore from Lemma 4 we obtain (11):

$$B^{u^*} V + C = 0 \qquad \text{a.e. on } \{V > f\}. \tag{11}$$

This inequality and Theorem 2 imply that u^* satisfies (10). Next let $u^* \in U$ satisfy (10). Then u^* satisfies (11) from Theorem 2. Define a Markov time $\tau^* = \inf\{t \geqq s : Z_t^{u^*} \in \{V = f\}\}$ for $(s, x) \in E$. Hence from (3) and (11) we obtain

$$V(s, x) = S_{\tau^*}^{u^*} V(s, x) = S_{\tau^*}^{u^*} f(s, x) \qquad \text{for all } (s, x) \in E.$$

This equality completes this corollary.

Acknowledgement

The auther would like to thank Prof. N. Furukawa for his helpful advices.

References

[1] DOSHI, B. T.: *Continuous time control of Markov processes on an arbitrary state space: discounted rate*, Ann. Stat., **6**, (1976), 1219-1235.

[2] DUBINS, L. and SAVAGE, L. J.: *How to gamble if you must*, McGraw-Hill, New York, (1965).

[3] FURUKAWA, N.: *Nearly optimal policies and stopping times in Markov decision processes with general rewards*, Bull. Math. Stat., **15**, 1~2, 89-109.

[4] FURUKAWA, N.: *Markov decision processes with compact action spaces*, Ann. Math. Stat., **43**, (1972), 1612-1622.

[5] HORDIJK, A.: *Dynamic programming and Markov potential theory*, Mathematisch Centrum, Ansterdum (1974).

[6] KRYLOV, N. V.: *Controlled diffusion processes*, Springer-Verlag, New York (1980).

[7] NISIO, M.: *Remarks on stochastic optimal controls*, Japan J. Math., 1 (1975), 159-183.

[8] OHTSUBO, Y.: *Controlled Markov jump processes associated with a choice of stopping rules*, Mem. Fac. Sci., **34**, (1980), 311-337.
[9] SHIRYAYEV, A.N.: *Optimal stopping rules*, Spring-Verlag, Berlin (1978).
[10] STROOK, D.W. and VARADHAN, S.R.S.: Springer-Verlag, Berlin (1979).

Communicated by N. Furukawa
Received March 30, 1985