13

On the Optimal Stopping Problem of Markov Chains by Variational Inequalities

Masami YASUDA

Reptrinted from the Journal of the College of Arts and Sciences
Chiba University, B-17
November 28, 1984

Journal of the College of Arts and Sciences, Chiba University,
B-17, pp. 11-16, November 28, 1984

On the Optimal Stopping Problem of Markov Chains by Variational Inequalities

Masami YASUDA

Dedicated to Professor Masayasu MURAKAMI on his sixtieth birthday

Statistical Institute, College of Arts and Sciences

(Received September 29, 1984)

Abstract

An optimality equation of the stopping problem is considered from an analytical point of view. We describe a variational inequality, which is equivalent to the optimality equation, and discuss the existence and the uniqueness of the equation and apply it to the best choice problem, Stopped Decision Process and the stopping problem of the Markov Jump process.

1. Optimal Stopping Problem for Markov Chains

Let $(x_n, \mathfrak{F}_n)$, $n \geqq 0$ be a Markov chain on a state space $(\boldsymbol{E}, \nu)$ with a transition probability $P(x, dy)$ where $\boldsymbol{E}$ is a locally compact separable Hausdroff space and ν is a positive Radon measure. Real valued functions $\varphi(x)$ and $c(x)$ on $x \in \boldsymbol{E}$ are called a terminal reward and a cost function respectively, and satisfy the assumption:

Assumption 1.1. Let φ, $c \in L^2(\boldsymbol{E}, \nu)$ = all of the square integrable functions on the state space $(\boldsymbol{E}, \nu)$.

For a stopping time τ, that is, $\{\tau \leqq n\} \in \mathfrak{F}_n$ for each n, define

(1.1) $$u(x) = \sup_{\tau} E_x[\varphi(x_\tau) - \sum_{i=1}^{\tau-1} c(x_i)],\ x = x_0 \in \boldsymbol{E}.$$

The optimal stopping problem $(\boldsymbol{E}, P, \varphi, c)$ is to find a stopping time which attains $u(x)$ for each $x \in \boldsymbol{E}$. The followings are the well known result (Dynkin/Yushkevich [3] and Ross [8]).

(a) $u = u(x)$ is the least excessive majorant of φ, i.e., $u(x) \geqq \varphi(x)$, $u(x) \geqq Pu(x) - c(x)$ for each $x \in \boldsymbol{E}$ and, if $v = v(x)$ satisfies $v(x) \geqq \varphi(x)$, $v(x) \geqq Pv(x) - c(x)$, then $v(x) \geqq u(x)$.

(b) $u = u(x)$ satisfies, by Dynamic Programming, the optimal equation (abr. by OE):

(1.2) $$u(x) = \max\{\varphi(x),\ Pu(x) - c(x)\},\ x \in \boldsymbol{E}.$$

Markov Decision Processes (abr. by MDP) contain the optimal stopping problem as the two-action "continue or stop" problem (refer to Derman [2]). To solve OE in MDP, the sequential approximation method is used commonly, while it converges the minimal solution of OE. If we confirm ourselves that the solution exists and it is unique, a limit function by the iteration method consists with the optimal value surely and so the discussion in the explicit form by OLA or ILA policy (refer to Ross [8]) becomes significant.

In this note we consider the uniqueness of the solution of OE in the optimal stopping problem by the notion of variational inequalities (Kinderlehrer/Stampacchia [6]) in the functional analysis. Its definition and lemma are prepared in section 2. In section 3 the equivalence between OE in the optimal stopping problem and variational inequalities are shown. The main result is to discuss a sufficient condition on the transition probability so that OE has a unique solution and is to apply it to some examples. When the transition probability is symmetric or doubly stochastic, it includes these cases. As the application of the theorem we show some examples in section 4, which are the best choice problem by Gilbert/Mosteller [5] and by Dynkin/Yushkevich [3], Stopped Decision Process by Furukawa/Iwamoto [4] and a cotinuous time stopping problem with Markov Jump processes.

2. Variational Inequalities

In this section the definition and the problem of variational inequalities are described. Let $\boldsymbol{X}$ be a reflexive Banach space with its dual $\boldsymbol{X}'$ and $<,>$ denotes a pairing of $\boldsymbol{X}'\boldsymbol{X}$ to $\boldsymbol{R}$. On a closed convex subset $\boldsymbol{K}$ of $\boldsymbol{X}$ let us assume that a continuous mapping $\boldsymbol{A}$ from $\boldsymbol{K}$ to $\boldsymbol{X}'$ is given. The following definition is due to Kinderlehrer/Stampacchia [6].

Definition 2.1. (1) A mapping $\boldsymbol{A}: \boldsymbol{K} \to \boldsymbol{X}'$ is called *monotone* if $<\boldsymbol{A}u-\boldsymbol{A}v,\ u-v> \geqq 0$ for all $u,\ v \in \boldsymbol{K}$. (2) The monotone mapping $\boldsymbol{A}$ is called *strictly monotone* if $<\boldsymbol{A}u-\boldsymbol{A}v,\ u-v>=0$ implies $u=v$. (3) $\boldsymbol{A}$ is *coercive* on $\boldsymbol{K}$ if there exists $u \in \boldsymbol{K}$ such that $\frac{<\boldsymbol{A}v-\boldsymbol{A}u,\ v-u>}{\| v-u \|} \to \infty$ as $\| v \| \to \infty$.

The problem in the theory of variational inequalities (abr. by VI) is to

$$\text{find } u \in \boldsymbol{K}:\ <\boldsymbol{A}u,\ v-u> \geqq 0 \text{ for all } v \in \boldsymbol{K}.$$

Lemma 2.1. (Kinderlehrer/Stampacchia [6]) *If the mapping is (strictly) monotone, then there exists a (unique) solution of the problem for VI.*

Note that the property of coerciveness implies strict monotonicity but not vice versa.

3. The Optimality Equation and Variational Inequalities

Assume that the stopping problem $(\boldsymbol{E},\ p,\ \varphi,\ c)$ is given. Let $\boldsymbol{L}^2$ be a Hilbert space $\boldsymbol{L}^2(\boldsymbol{E},\ \nu)$ and let $\boldsymbol{X}=\boldsymbol{L}^2$ in section 2 so that the pairing is considered as a inner product (,).

Definition 3.1. A function $u=u(x),\ x \in \boldsymbol{E}$ is a *solution* of OE if $u,\ Pu \in \boldsymbol{L}^2$ and it satisfies (1.2) for a. e. $x \in \boldsymbol{E}$.

For the terminal reward φ in the stopping problem let

(3.1) $\quad \boldsymbol{K}=\{v \in \boldsymbol{L}^2;\ v(x) \geqq \varphi(x),\ x \in \boldsymbol{E}\}$,

which is clearly a closed convex set in $\boldsymbol{L}^2$.

Definition 3.2. A function $u=u(x),\ x \in \boldsymbol{E}$ is a *solution* of VI if $u \in \boldsymbol{K}$ satisfies

(3.2) $\quad (u-Pu+c,\ v-u) \geqq 0$ for any $v \in \boldsymbol{K}$.

Theorem 3.1. *The solutions of OE (1.2) and VI (3.2) in each of these two equations are equivalent.*

Proof. Rewriting the equation (1.2), we have

$$u(x) \geqq \varphi(x),\ u(x) \geqq Pu(x)-c(x) \quad \text{and}$$
$$\{u(x)-\varphi(x)\}\{u(x)-Pu(x)+c(x)\}=0 \quad \text{for a.e. } x \in \boldsymbol{E}.$$

This is a complementarity form for VI of (3.2). Therefore the proof is immediately obtained. //

By Theorem 3.1 and Lemma 2.1, if the mapping $I-P: \boldsymbol{L}^2 \to \boldsymbol{L}^2$ is coercive then the solution of OE (2.1) exists uniquely. In order to apply this result, we define the next one of slightly strong version, which is used in the next section.

Definition 3.3. The function $u=u(x),\ x \in \boldsymbol{E}$ is *mean-square-subregular* (abr. by MSS) with respect to P if $\int_E \int_E u^2(y)P(x,\ dy)dx \leqq \beta \int_E u^2(x)dx$ for some $0<\beta \leqq 1$. When $0<\beta<1$, we call it strictly MSS (abr. by s-MSS).

A sufficient condition for the function $u=u(x)$ to be MSS is that its square function is subregular with respect to (abr. by w.r.t.) the transition probability P.

Theorem 3.2. *If each $u \in \boldsymbol{K}$ is s-MSS w.r.t. P then the solution of OE (1.2) exists uniquely.*

Proof. By the result in Kinderlehrer/Stampacchia [6], the coerciveness is equivalently that, for some α; $0<\alpha$, $((I-P)u,u) \geqq \alpha(u,u)$ in the bilinear form. Because $((I-P)u,u)=\int_E u^2(x)dx-\int_E \int_E u(x)u(y)P(x,dy)dx \geqq \left\{\int_E u^2(x)dx-\int_E \int_E u^2(y)P(x,dy)dx\right\}/2$, the assumption implies that the last is greater than or equal to $\frac{1-\beta}{2}\int_E u^2(x)dx$. Therefore, letting $\alpha=(1-\beta)/2$, this shows the coerciveness of mapping $I-P$. Hence the existence is immediate from Lemma 2.1. //

On the condition of MSS the following two cases are sufficient, and the existence of OE is assured certainly. But if P is asymmetric, there is a counter example.

Corollary 3.3. *If* (i) *$P(x,dy)$ is symmetric, i.e.,*

$$\int_E \int_E f(x,\ y)P(x,dy)dx=\int_E \int_E f(y,\ x)P(x,dy)dx$$

for any measurable function $f=f(x,\ y)$, or (ii) *$\boldsymbol{E}$ is denumerable and $P=(p(i,\ j),\ i,\ j \in \boldsymbol{E})$ is doubly stochastic, i.e.,*

$$\sum_{i \in E} p(i,j)=1 \text{ for each } j \in \boldsymbol{E},$$

then a solution of OE (2.1) exists.

We note here that a stopping problem for a variant of Matrix Game is also able to discuss by VI approach. The details are omitted.

4. Applications

In this section the result of Theorem 3.2 is applied to three problems.

4.1. The Best Choice Problem

The best choice problem is one of the typical stopping problem discussed by Gilbert/Mosteller [5]. The details are omitted and so it should be refered to these. Dynkin/Yushkevich [3] formulated the problem as Markov chain and it is described as follows. Consider the stopping problem $(\boldsymbol{E}, P, \varphi, c)$: $\boldsymbol{E}=\{1, 2, .., n\}$; $P=(p(i, j);\ i, j\in \boldsymbol{E})$, $p(i, j)=\frac{i}{j(j-1)}$ $(1\leq i<j\leq n)$, $=0$ (otherwise); $c(i)=0$ and $\varphi(i)=i/n$ for all $i\in \boldsymbol{E}$. Thus OE (1.2) of the best choice problem can be determined. To show the solution exists uniquely, we check the condition of s-MSS. Easily we have that, for all $v=v(j)$, $j\in \boldsymbol{E}$,

$$\sum_{k=1}^{n} Pv^2(k)=\sum_{k=1}^{n}\sum_{j=k+1}^{n}\frac{k}{j(j-1)}v^2(j)=\sum_{j=1}^{n}\sum_{k=1}^{j-1}\frac{k}{j(j-1)}v^2(j)=\frac{1}{2}\sum_{j=1}^{n}v^2(j)\geq 0.$$

Therefore Theorem 3.2 implies the conclusion.

We shall consider an extension of the choice problem, which is to select one of 1st, 2nd,.., m-th best object among all n objects, discussed by Gusein/Zade (refer to Dynkin/Yushkevich [3]). To simplify the description of corresponding OE, we take a scaling limit of OE tending i and n to infinity provided $i/n=x$ are fixed. Then OE with respect to $u_k=u_k(x)$, $k=1, 2,.., m$ is obtained as follows:

(4.1) $$u_k(x)=\max\{x^k, \int_x^1 \frac{x^m}{y^{m+1}}\sum_{l=1}^{m} u_l(y)dy\},\ x\in[0, 1],\ k=1,.., m$$

where $u_k(x)=\lim_{\substack{i/n=x \\ i,\ n\to\infty}} u_k(i)$. Let $u(x)=\sum_{k=1}^{m} u_k(x)$. Consider the next VI: Set $\boldsymbol{L}^2=\boldsymbol{L}^2$ ([0, 1], dx) where dx is the Lebesgue measure and $\boldsymbol{K}=\{v\in \boldsymbol{L}^2;\ v(x)\geq \sum_{k=1}^{m} x^k\}$.

(4.2) Find $u\in \boldsymbol{K}$ such that $\int_0^1 (u(x)-\int_0^1 \frac{x^m}{y^{m+1}} u(y)dy)\,(v(x)-u(x))dx\geq 0$ for all $v\in \boldsymbol{K}$.

Similarly as before, we check the condition of s-MSS with $\beta=\frac{1}{m+1}$ for each $u\in \boldsymbol{K}$ and it implies the existence of the solution of OE (4.1).

4.2. Stopped Decision Process

Furukawa/Iwamoto [4] discussed a paired problem of MDP and Stopping problem, which they call as Stopped Decision Process. The problem is to decide a transition probability and a stopping time which maximize an objective function. That is, for a class $\boldsymbol{P}$ of a transition probability $P=(P_1, P_2,..)$ and an associated stopping time τ,

$$u(x)=\sup_{\tau,P} E_{x,P}\Big[\varphi(x_\tau)-\sum_{k=1}^{\tau-1} c(x_k)\Big]$$

where $E_{x,P}$, $P\in \boldsymbol{P}$ is an expectaion of non-stationary Markov chain $(x_n;\ n=1, 2,..)$. Denote this problem by $(\boldsymbol{E}, \boldsymbol{P}, \varphi, c)$. The corresponding OE becomes

(4.3) $$u(x)=\max\{\varphi(x), \sup_P Pu(x)-c(x)\},\ x\in \boldsymbol{E}.$$

Let define an operator $\boldsymbol{A}$ by $I-\sup_P P$. Assuming that the domain of $A\subset \boldsymbol{L}^2$, VI is follows:

(4.4) find $u\in \boldsymbol{K}$, $(\boldsymbol{A}u+c,\ v-u)\geq 0$ for $v\in \boldsymbol{K}=\{v;\ v(x)\geq\varphi(x)\}$.

Hence we have that, if each u in $\boldsymbol{K}$ is MSS (s-MSS) w.r.t. $\sup_P P$, the solution of (4.3) in $(\boldsymbol{E}, \boldsymbol{P}, \varphi, c)$ exists (uniquely exists).

To assure the existence, one of the conditions is the reward in future is discounted and it is reduced to be contractive.

4.3. A Stopping Problem of Markov Jump Process

In case of a continuous time parameter for a homogenuous Markov process $(x_t,\ t\geq 0)$ on state space $\boldsymbol{E}$, consider an infinitesimal generator $\mathfrak{A}$: $\mathfrak{A}f(x)=\lim_{t\to 0}\frac{E_x f(x_t)-f(x)}{t}$, $x=x_0\in \boldsymbol{E}$ in stead of the difference $-(I-P)$ in the discrete time parameter case. Let τ be a stopping time and let

$$u(x)=\sup_\tau E_x[e^{-\alpha\tau}\varphi(x_\tau)-\int_0^\tau e^{-\alpha t}c(x_t)dt]$$

with a discount factor $\alpha>0$. For this stopping problem $(\boldsymbol{E}, \mathfrak{A}, \alpha, \varphi, c)$, the corresponding OE is as follows:

(4.5) $$u(x)=\max\{\varphi(x), -(\alpha\boldsymbol{I}-\mathfrak{A})u(x)-c(x)+u(x)\},\ x\in \boldsymbol{E},$$

provided that all of the functions are in $\boldsymbol{L}^2$. Also VI becomes that

(4.6) find $u\in \boldsymbol{K}$ such that $((\alpha\boldsymbol{I}-\mathfrak{A})u+c,\ v-u)\geq 0$ for $v\in \boldsymbol{K}=\{v;\ v(x)\geq\varphi(x)\}$.

By Theorem 3.2, we obtain that if each $u\in \boldsymbol{K}$ is MSS w.r.t. $\mathfrak{A}$ and $\alpha>0$, then there exists a unique solution of OE (4.5).

There are many research works concerned with the partial differential equations for a local type of diffusion processes which the system is described by a stochastic differential equation. Refer to Bensousal/Friedman [1], Nagai [7] and else. Here we pickup a jump type of a Markov Jump process as a concrete example.

Consider a non-homogenuous Poisson process $N(t)$, $t\geq 0$ with a parameter $\mu(t)>0$ and, for a state space $\boldsymbol{E}=\{0, 1, 2,..\}$, a terminal reward $\varphi(n)$, a cost function $c(n)=0$, $n\in \boldsymbol{E}$ and a discount factor $\alpha>0$ are given. The objective of this stopping problem is

$$u(t,n)=\sup_{\tau\geq t} E_{N(t)=n}[e^{-\alpha(\tau-t)}\varphi(N(t))],\ t\geq 0,\ n\in \boldsymbol{E}.$$

To reduce the problem into the above discussion, we calculate an infinitesimal generator $\mathfrak{A}$ of a space time $X(t)=(t,N(t))$ and then check the condition is satisfied. Since

$$\mathfrak{A}u(t,n)=\frac{\partial u(t,n)}{\partial t}+\mu(t)\{u(t,n+1)-u(t,n)\}$$

and OE of the problem is

(4.7) $$u(t,n)=\max\{\varphi(n), -(\alpha\boldsymbol{I}-\mathfrak{A})u(t,n)+u(t,n)\},\ t\geq 0,\ n\in \boldsymbol{E},$$

the corresponding VI is determined similarly as before, by letting $\boldsymbol{L}^2 = \{u = u(t,n)\,;\ \sum_{n=0}^{\infty}\int_0^{\infty} u^2(t,n)\,dt < \infty\}$ be a Hilbert space with $\| u \|^2 = \sum_{n=0}^{\infty}\int_0^{\infty} u^2(t,n)dt$ and $\boldsymbol{K} = \{u = u(t,n)\,;\ u(t,n) \geqq \varphi(n)$ for all $t,n\} \subset \boldsymbol{L}^2$. Now we show that every function $u = u(t,n)$ satisfies s-MSS condition w.r.t. the mapping $\alpha \boldsymbol{I} - \mathfrak{A}$.

Because

$$(\mathfrak{A}u,\ u) = \sum_{n=0}^{\infty}\int_0^{\infty}\frac{\partial}{\partial t}u(t,n)\ u(t,n)dt + \int_0^{\infty}\mu(t)\sum_{n=0}^{\infty}\{u(t,n+1) - u(t,n)\}u(t,n)dt$$

$$\leqq -\frac{1}{2}\left\{\sum_{n=0}^{\infty} u^2(0,n) + \int_0^{\infty}\mu(t)u^2(t,0)dt\right\} \leqq 0,$$

it implies $(\alpha u - \mathfrak{A}u,\ u) \geqq \alpha(u,\ u)$. Hence OE (4.7) has a unique solution by Theorem 3.2 providing the discount factor $\alpha > 0$. As a variation for the problem, a control problem in the queueing theory, which varies the parameter $\mu = \mu(t)$ in a certain class, is able to discuss and, by combining with the previous Stopped Decision Process in a continuous version, some results would be obtained.

Acknowledgement

The author wishes to thank Dr. Masami KURANO and Dr. Junichi NAKAGAMI for their helpful discussion.

References

[1] Benssousan, A. and Friedman, A.; Nonlinear Variational inequalities and Differential Games with Stopping Times, *J. Func. Anal.,* **16** (1974), 305-352.

[2] Derman, C; *Finite State Markovian Decision Processes,* Academic Press, New York, 1970.

[3] Dynkin, E.B. and Yushkevich, A.A.; *Markov processes ; Theorems and problem,* Plenum Press, New York, 1969.

[4] Furukawa, N. and Iwamoto, S.; Stopped Decision Processes on Complete Separable Metric Spaces, *J. Math. Anal. Appl.,* **31** (1970), 615-658.

[5] Gilbert, J.P. and Mosteller, F.; Recognizing the Maximum of a Sequence, *J. Amer. Stat. Assoc.,* **61** (1966), 35-73.

[6] Kinderlehrer, D. and Stampacchia, G.; *An Introduction to Variational Inequalities and Their Applications,* Academic Press, New York, 1980.

[7] Nagai, H.; On a Optimal Stopping Problem and a Variational Inequality, *J. Math. Soc. Japan,* **30** (1978), 303-312.

[8] Ross, S.M.; Infinitesimal Look-ahead Stopping Rules, *Ann. Math. Stat.,* **42** (1971), 297-303.