Bulletin of Informatics and Cybernetics, Vol. 28, No. 1, 1996

UTILITY-OPTIMAL STOPPING IN A DENUMERABLE MARKOV CHAIN

By

Yoshinobu KADOTA*, Masami KURANO† and Masami YASUDA‡

Abstract

This paper is concerned with a general utility of the optimal stopping problem for denumerable Markov chains. The validity of the one-step look ahead (OLA) stopping time is shown under a general utility criteria. It is developed from the view points of the optimality and a "risk-averse" or "risk-seeking" characterization. The results are applied to the case of a exponential utility function and illustrated by a simple example.

Keywords : Optimal stopping, utility function, OLA stopping time, denumerable Markov chain.

AMS(MOS) subject classification : 90C39, 90C47, 90A06.

1. Introduction

We are concerned with general utility treatment of optimal stopping problems for a denumerable Markov chain. As for the utility theory, an exponential utility has many attractive properties as are seen in Fishburn[4] and Pratt[8]. So many authors analyzed decision processes with it. For such examples of Markov decision processes, see Howard and Matheson[5] and Chung and Sobel[2].

The analysis under a general utility criterion has been done, for example, in Rieder[7] and our previous paper[6], which is expected to enlarge the practical applications of the utility. To our knowledge, Denardo and Rothblum[3] is the only work related with utility treatment of optimal stopping problem. They analyze the problem in a finite Markov decision chain with the exponential utility and give a linear programming corresponding to an optimal stopping time.

Here, we consider the stopping problem for a general utility function with the random variable of total rewards. The validity of one-step look ahead (OLA) stopping time is discussed. For the OLA stopping time, refer to Ross[9]. A sufficient condition is given for the stopping time to be optimal. A property of the stopping time is characterized by

* Faculty of Education, Wakayama University, Wakayama 640, Japan

† Faculty of Education, Chiba University, Chiba 263, Japan

‡ Faculty of Science, Chiba University, Chiba 263, Japan

being "risk-averse" or "risk-seeking". Those results are applied to the case of exponential utility functions and an illustrative example is given.

2. Optimal stopping problem

In this section, we define the optimal stopping problem with general utility criteria for a denumerable Markov chain. Let $S = \{0, 1, 2, \cdots\}$ be a state space and $Q = (q_{ij}; i, j \in S)$ is the transition probability matrix on S, where $q_{ij} \geq 0$ for any $i, j \in S$ and $\sum_{j \in S} q_{ij} = 1$ for any $i \in S$. The sample space is the product space $\Omega = S^{\infty}$. The projection X_t on Ω to the t-th factor S describes the state of the t-th time of the process. For any Borel measurable set M, $\mathcal{P}(M)$ denotes the set of all probability measures on M. Then, for any $\nu \in \mathcal{P}(S)$, we define the probability measure $P_\nu \in \mathcal{P}(\Omega)$ by the usual way, where $\{X_t\}$ becomes a stationary Markov chain on S with the initial distribution ν and the the transition matrix Q.

We call a random variable $\sigma : \Omega \to \{0, 1, 2, \cdots\}$ a stopping time if $P_\nu(\sigma < \infty) = 1$ and $\{\sigma = t\}$ is measurable with respect to the σ- algebra induced by $\{X_0, X_1, \cdots, X_t\}$ for $t = 0, 1, 2, \cdots$. Let denote by Σ_ν the set of all stopping times starting with the initial distribution $\nu \in \mathcal{P}(S)$. Let $\mathcal{R}$ be the set of all real numbers. The terminal reward at the state $i \in S$, $r_i = r(i)$, is a function from S to $\mathcal{R}$ and the observation cost per unit time is a constant $c > 0$. The total reward when the system is stopped at time t is given by the random variable

$$\mathcal{B}_t := c_t + r(X_t),$$

where $c_t := -ct$.

A utility g is a Borel measurable function from $\mathcal{R}$ to itself. Let denote by $E_\nu[Y]$ the expectation of a random variable Y with respect to P_ν. For any utility g and the initial distribution $\nu \in \mathcal{P}(S)$, our optimal stopping problem is to maximize the expected utility

$$G(g, \nu, \sigma) := E_\nu[g(\mathcal{B}_\sigma)]$$

over all $\sigma \in \Sigma_\nu$.

For any g and $\nu \in \mathcal{P}(S)$, the stopping time $\sigma^* \in \Sigma_\nu$ is called $(\nu,\ g)$-*optimal*, if $G(g, \nu, \sigma^*) \geq G(g, \nu, \sigma)$ for all $\sigma \in \Sigma_\nu$. The $\sigma^* \in \bigcap_{\nu \in \mathcal{P}(S)} \Sigma_\nu$ is called *g-optimal* if it is $(\nu,\ g)$-optimal for all $\nu \in \mathcal{P}(S)$.

In the subsequent discussion, it is convenient to rewrite the expected utility $G(g, \nu, \sigma)$ by using the distribution function of $\mathcal{B}_\sigma$ with respect to P_ν. For this purpose, we define

$$F_\nu^\sigma(x) := P_\nu(\mathcal{B}_\sigma \leq x) \quad \text{for each } x \in \mathcal{R}.$$

Then, clearly it holds $G(g, \nu, \sigma) = \displaystyle\int_{-\infty}^{\infty} g(x)\, F_\nu^\sigma(dx)$. Let

$$\Phi(\nu) := \{F_\nu^\sigma(\cdot) | \sigma \in \Sigma_\nu\}$$

and

$$U\{g\}(a, i) := \sup_{F \in \Phi(i)} \int_{-\infty}^{\infty} g(a + x)\, F(dx)$$

for $a \in \mathcal{R}$ and $i \in S$, where $\Phi(i)$ denotes $\Phi(\nu)$ for ν such that $\nu(\{i\}) = 1$.

3. g-optimality and OLA stopping times

In this section, the validity of the OLA stopping times is discussed in the optimal stopping problem under the general utility.

In order to characterize the optimal stopping time, we consider the following set and its hitting time. For each $t = 0, 1, 2, \cdots$, let

$$S_t\{g\} := \{\, i \in S \,|\, g(c_t + r_i) \geq \sum_{j \in S} q_{ij} U\{g\}(c_{t+1}, j) \,\} \tag{3.1}$$

and

$$\sigma^* := \{ \text{ the first time } t \text{ such that } X_t \in S_t\{g\}\}.$$

Denote $x^+ = \max\{x, 0\}$. The following assumption is the statement for each $\nu \in \mathcal{P}(S)$.

ASSUMPTION A(ν). $E_\nu[\sup_{\{t \geq 0\}} g(\mathcal{B}_t)^+] < \infty$.

The next theorem is easily proved by applying the result in Chow, Robbins and Siegmund[1] to the sequence of random variables $\{g(\mathcal{B}_t)\}_{t=0,1,2,\cdots}$.

THEOREM 3.1. (refer Theorem 4.5 in [1])

(i) *For any $\nu \in \mathcal{P}(S)$, suppose that* Assumption A(ν) *and $P_\nu(\sigma^* < \infty) = 1$ hold. Then, $\sigma^* \in \Sigma_\nu$ and σ^* is (ν, g)-optimal.*

(ii) *Suppose that* Assumption A(ν) *holds for any $\nu \in P(S)$ and that $\lim_{t\to\infty} g(\mathcal{B}_t) = -\infty$ P_ν-a.e.. Then, $\sigma^* \in \bigcap_{\nu \in \mathcal{P}(S)} \Sigma_\nu$ and σ^* is g-optimal.*

Now, using an idea of the OLA stopping time for optimal stopping problems with additive utility functions (for example, see [9]), we derive some results on the general utility case. For each t, let

$$S_t^*\{g\} := \{\, i \in S \,|\, g(c_t + r_i) \geq \sum_{j \in S} q_{ij}\, g(c_{t+1} + r_j)\}. \tag{3.2}$$

Notice from (3.1) that $S_t\{g\} \subset S_t^*\{g\}$ for all t. The OLA stopping time is a stopping time whose value is determined by the first hitting time t such that $X_t \in S_t^*\{g\}$. Here, we introduce an assumption to get useful results on the validity of the OLA stopping time.

ASSUMPTION B. For each $t = 0, 1, 2, \cdots$, $Q = (q_{ij})$ and $S_t^*\{g\}$ satisfy that

$$q_{ij} = 0 \text{ if } i \in S_t^*\{g\} \text{ and } j \notin S_{t+1}^*\{g\}.$$

If $S_0^*\{g\} = S_t^*\{g\}$ for all t, Assumption B assures the *closedness* of $S_0^*\{g\}$. Notice that if $S_t^*\{g\} \neq \emptyset$ for some t, then $S_{t+n}^*\{g\} \neq \emptyset$ for $n = 1, 2, \cdots$.

Let denote E_ν and Σ_ν for $\nu \in \mathcal{P}(S)$ such that $\nu(\{i\}) = 1$ by E_i and Σ_i, respectively. In the following lemma, (3.3) is sometimes called the monotone property.

LEMMA 3.2. *Suppose that* $E_i[\sup_{\{k\geq 0\}} g(c_t + \mathcal{B}_k)^+] < \infty$ *for any* $i \in S_t^*\{g\}$ *and* $t = 0, 1, 2, \cdots$. *Then, Assumption B implies that*

$$U\{g\}(c_t, j) = g(c_t + r_j) \quad \text{for any } j \in S_t^*\{g\} \text{ and } t = 0, 1, 2, \cdots. \tag{3.3}$$

PROOF. For any $\varepsilon > 0$ there exists a $\sigma \in \Sigma_j$ such that

$$U\{g\}(c_t, j) \leq E_j[g(c_t + \mathcal{B}_\sigma)] + \varepsilon. \tag{3.4}$$

by the definition of $U\{g\}(c_t, j)$. Let $j \in S_t^*\{g\}$ and $i \in S_{t+k-1}^*\{g\}$ for $k = 1, 2, \cdots$. Then we have the following:

$$\begin{aligned} & E_j[\, g(c_t + \mathcal{B}_\sigma) \,|\, \sigma = k,\, X_{k-1} = i\,] \\ = \; & \textstyle\sum_{\ell \in S} q_{i\ell} g(c_{t+k} + r_\ell) \\ \leq \; & g(c_{t+k-1} + r_i). \end{aligned}$$

This implies from Assumption B that

$$E_j[\, g(c_t + \mathcal{B}_\sigma) \,|\, \sigma,\, X_{\sigma-1}\,] \leq g(\, c_t + \mathcal{B}_{(\sigma-1)\vee 0}\,)$$

with probability one, where $x \vee y = \max\{\, x,\, y\}$. Thus we get

$$E_j[\, g(c_t + \mathcal{B}_\sigma)\,] \leq E_j[\, g(c_t + \mathcal{B}_{(\sigma-1)\vee 0})\,].$$

By repeating the above discussion, we have for $\ell = 1, 2, \cdots$ that

$$E_j[\, g(c_t + \mathcal{B}_\sigma)\,] \leq E_j[\, g(c_t + \mathcal{B}_{(\sigma-\ell)\vee 0})\,]. \tag{3.5}$$

Evaluating the right-hand side of (3.5), we have

$$\begin{aligned} & E_j[\, g(c_t + \mathcal{B}_{(\sigma-\ell)\vee 0})\,] \\ \leq \; & P_j(\sigma \leq \ell) g(c_t + r_j) + P_j(\sigma > \ell) E_j[\sup_{\{k\geq 0\}} g(c_t + \mathcal{B}_k)^+\,]. \end{aligned}$$

Since $P_j(\sigma > \ell) \to 0$ as $\ell \to \infty$, together with (3.4) the above inequalities imply that

$$U\{g\}(c_t, j) \leq g(c_t + r_j) + \varepsilon. \tag{3.6}$$

As $\varepsilon \to 0$ in (3.6), we get $U\{g\}(c_t, j) \leq g(c_t + r_j)$. This completes the proof. □

The next theorem gives a sufficient condition for the OLA stopping time to be optimal under the general utility.

THEOREM 3.3. *Suppose* $E_i[\sup_{\{k\geq 0\}} g(c_t + \mathcal{B}_k)^+] < \infty$ *for any* $i \in S_t^*\{g\}$ *and* $t = 0, 1, 2, \cdots$. *If Assumption B holds, then it holds that* $S_t^*\{g\} = S_t\{g\}$ *for* $t = 0, 1, 2, \cdots$.

PROOF. For simplicity of notation, put $S_t^* := S_t^*\{g\}$ and $S_t := S_t\{g\}$. Since $S_t \subset S_t^*$ is obvious, it is sufficient to prove $S_t^* \subset S_t$ for each t. Let $i \in S_t^*$. Using Assumption B and Lemma 3.1, we have

$$\sum_{j\in S} q_{ij}U\{g\}(c_{t+1}, j) = \sum_{j\in S_{t+1}^*} q_{ij}U\{g\}(c_{t+1}, j) = \sum_{j\in S_{t+1}^*} q_{ij}g(c_{t+1} + r_j) \leq g(c_t + r_i).$$

This implies $i \in S_t$. □

In case of a linear utility function $g(x) = x$, (3.2) is reduced to

$$S^* := S_t^*\{x\} = \{i \in S \,|\, c + r_i \geq \sum_{j\in S} q_{ij}r_j)\}, \tag{3.7}$$

which is independent of t and so we denote it as S^*. The next theorem shows a property of the OLA stopping times characterized by the non-decreasing utility.

THEOREM 3.4. *Let* g *be a non-decreasing function.*

(i) *If it is concave, then* $S_t^*\{g\} \supset S^*$ *for each* t.

(ii) *If it is convex, then* $S_t^*\{g\} \subset S^*$ *for each* t.

PROOF. We give only the proof for (i), since (ii) is proved similarly. For any $i \in S^*$ and t, we have from (3.7) that $c_t + r_i \geq c_{t+1} + \sum_{j\in S} q_{ij}r_j$. Therefore, using the monotonicity and the Jensen's inequality, we get

$$g(c_t + r_i) \geq g(c_{t+1} + \sum_{j\in S} q_{ij}r_j) \geq \sum_{j\in S} q_{ij}g(c_{t+1} + r_j).$$

This leads $i \in S_t^*\{g\}$, which implies $S^* \subset S_t^*\{g\}$. □

We note that the concave function is risk-averse (cf. Pratt[8]) and the convex one is risk-seeking. Theorem 3.4 says that the OLA stopping time of a risk-averse decision maker has a tendency to stop earlier than that of a risk seeking one.

4. Exponential utility functions

We consider the case of an exponential utility function

$$g_\lambda(x) = \text{sign}(-\lambda)\exp\{-\lambda x\} \tag{4.1}$$

for a non-zero constant λ in this section. This utility function is risk-averse when $\lambda > 0$ and risk-seeking when $\lambda < 0$. In order to describe the OLA stopping time for this utility, let

$$\eta_\lambda(i) := \sum_{j\in S} q_{ij}\exp\{\lambda(r_i - r_j)\} \quad \text{for} \quad i \in S.$$

ASSUMPTION C. For any $\lambda > 0$, $\eta_\lambda(\cdot)$ is non-decreasing on $S = \{0, 1, 2, \cdots\}$, and for $\lambda < 0$, $\eta_\lambda(\cdot)$ is non-increasing.

PROPOSITION 4.1. *Suppose $\eta_\lambda(i)$ is finite for any $i \in S$ and $g_\lambda(x)$ is given by* (4.1).

(i) *$S_t^*\{g_\lambda\}$ is independent of t, i.e. $S^*\{g_\lambda\} = S_t^*\{g_\lambda\}$ for $t = 0, 1, 2, \cdots$.*

(ii) *Suppose Assumption C holds. If $S^*\{g_\lambda\} \neq \emptyset$, there exists an integer $l_\lambda^* \in S$ such that*

$$S^*\{g_\lambda\} = \{ j \in S \,|\, j \geq l_\lambda^* \}. \tag{4.2}$$

(iii) *Let $l_{\lambda_1}^*$ and $l_{\lambda_2}^*$ be the numbers which satisfy* (4.2) *for $\lambda_1 < 0$ and $\lambda_2 > 0$, respectively. Then it holds that $l_{\lambda_1}^* \geq l_{\lambda_2}^*$.*

PROOF. Observing (3.2), we have

$$S_t^*\{g_\lambda\} = \{ i \in S \,|\, \mathrm{sign}(-\lambda) e^{-\lambda r_i} \geq \mathrm{sign}(-\lambda) e^{\lambda c} \sum_{j \in S} q_{ij} \exp\{-\lambda r_j\} \}, \tag{4.3}$$

which shows that $S_t^*\{g_\lambda\}$ is independent of t. If $\lambda > 0$, (4.3) implies $S^*\{g_\lambda\} = \{ i \in S \,|\, \eta_\lambda(i) \geq \exp(-\lambda c) \}$. Let l_λ^* be the smallest integer in $S^*\{g_\lambda\}$. Then (4.2) follows from Assumption C. The case for $\lambda < 0$ is proved as similarly as the above. Also, (iii) follows clearly from Theorem 3.4. This completes the proof. □

Here we give a specific example to illustrate the theoretical results. Let $S = \{1, 2, 3, \cdots\}$ and

$$q_{ij} := \begin{cases} \dfrac{(\mu/i)^{j-i} \exp\{-\mu/i\}}{(j-i)!} & \text{if } 1 \leq i \leq j, \\ 0 & \text{if } j < i, \end{cases}$$

for $\mu > 0$. Let $r_i = i$ for all $i \in S$.

In case of a linear utility $g(x) = x$, S^* is given by

$$S^* = \{ i \in S \,|\, i \geq \lceil \mu/c \rceil \},$$

where $\lceil x \rceil$ is the smallest integer which is greater than or equal to x. By (4.1), it becomes

$$\eta_\lambda(i) = \sum_j q_{ij} \exp\{\lambda(r_i - r_j)\} = \exp\{\frac{\mu}{i}(e^{-\lambda} - 1)\}.$$

So $\eta_\lambda(\cdot)$ satisfies Assumption C. After a simple calculation, the integer l_λ^* of (4.2) is given by

$$l_\lambda^* = \left\lceil \frac{\mu}{c} \cdot \frac{1 - e^{-\lambda}}{\lambda} \right\rceil.$$

Since $(1-e^{-\lambda})/\lambda > 1$ if $\lambda < 0$ and $(1-e^{-\lambda})/\lambda < 1$ if $\lambda > 0$ respectively, $l_{\lambda_2}^* \leq \lceil \mu/c \rceil \leq l_{\lambda_1}^*$ for any $\lambda_1 < 0$ and $\lambda_2 > 0$ which is noted in Proposition 4.1(iii).

Acknowledgements

The authors would like to thank the referee for his valuable comments.

References

[1] Y. S. Chow, H. Robbins and D. Siegmund, The Theory of Optimal Stopping : Great Expectations, Houghton Mifflin Company, 1971.

[2] K.J. Chung and M.J. Sobel, Discounted MDP's: Distribution functions and exponential utility maximization, *SIAM J. Control and Optimization*, **25**(1987), 49-62.

[3] E.V. Denardo and U.G. Rothblum, Optimal stopping, exponential utility and linear programming, *Math. Prog.*, **16** (1979), pp.228–244.

[4] P.C. Fishburn, Utility Theory for Decision Making, John Wiley & Sons, New York, 1970.

[5] R.S. Howard and J.E. Matheson, Risk-sensitive Markov decision processes, *Manag. Sci.*, **8** (1972), pp.356–369.

[6] Y. Kadota, M. Kurano and M. Yasuda, Discounted Markov decision processes with general utility functions, Proceedings of APORS' 94, World Scientific, pp330–337, 1995.

[7] U. Rieder, Non-cooperative dynamic games with general utility functions, T. E. S. Raghavan et al.(eds). Stochastic Games and Related Topics, Klumer Academic publishers. 161–174, 1991.

[8] J.W. Pratt, Risk aversion in the small and in the large, *Econometrica*, **32** (1964) 122-136.

[9] S.M. Ross, Applied Probability Models with Optimization Applications, Holden-Day, 1970.

Received May 18, 1995
Revised December 15, 1995