OPTIMAL STOPPING PROBLEM ON FINITE MARKOV CHAIN

Hisano, Hiroshi
Department of Management Engineering, Tohwa University

https://doi.org/10.5109/13512

出版情報：Bulletin of informatics and cybernetics. 34 (2), pp.97-104, 2002-12. Research Association of Statistical Sciences

Bulletin of Informatics and Cybernetics, Vol. 34, No. 2, 2002

OPTIMAL STOPPING PROBLEM ON FINITE MARKOV CHAIN

By

Hiroshi HISANO*

Abstract

This paper studies an optimal stopping problem over a finite-horizon Markov chain on a finite-state space. First of all, we derive a recursive formula for the total number of all stopping rules in m-state n-stage stopping problem. Next we show an optimal stopping rule and give a characterization of optimal value functions.

Key Words and Phrases: optimal stopping, finite horizon Markov chain, total number of stopping rules

1. Introduction

It is wel known that a general theory of optimal stopping has been established by Snell (1952), Chow, Robbins and Siegmund (1971), Shiryaev (1978) and others. The main topic was a class of infinite-horizon problems. Recently a finite-horizon problem has been well applied to mathematical finance, in particular to fair pricing of American option, e.g. Shiryaev (1999).

On the other hand, the theory of optimal stopping is closely related to dynamic programming, e.g. Bellman (1957) and to Markov decision process, e.g. Howard (1960). A construction of optimal stopping rule for finite horizon problem is performed through backward induction. This is a basic idea of dynamic programming/Markov decision process.

In this paper we consider a class of optimal stopping problems over a finite horizon Markov chain on finite state space. We direct our attention to the total number of stopping times, optimality and characterizaion. Our approach separates the underlying process (Markov chain) and sequence of gain functions. This separation is due to the fact that the underlying process is a Markov chain. The value process over the Markov chain constitutes a minimal supermartingale.

2. Stopping Times

Let two integers $m \geq 1$ and $n \geq 1$ be given in this paper. We consider an n-stage Markov chain on state space with m states. Let $\{X_t\}_0^n$ be a Markov chain on a finite state space $S = \{s_1, s_2, \ldots, s_m\}$ with a transition law $p = \{p(\cdot|\cdot)\}$:

$$P(X_{t+1} = s_j | X_t = s_i) = p(s_j|s_i).$$

* Department of Management Engineering, TOHWA University Fukuoka 815-8510, Japan. tel +81 92-541-1152 his@tohwa-u.ac.jp

We assume that the Markov chain starts at a *preassigned* state $s_0 \in S$; $X_0 = s_0$.

Let $S^i := \overbrace{S \times S \times \cdots \times S}^{i \text{ times}}$ be the direct product of i state spaces S. We take $\Omega := S^{n+1}$ and $N := \{0, 1, \ldots, n\}$. A mapping $\tau : \Omega \to N$ is called a *stopping time* if for any $t \in N$ the set $\{\tau = t\}$ is determined by random variables $\{X_0, X_1, \ldots, X_t\}$.The stopping time is called an m-state n-stage stopping time. Let $\mathcal{T} := \mathcal{T}_0^n := \mathcal{T}_0^n(m)$ be the set of all m-state n-stage stopping times. The first question is how many stopping times there are for m-state n-stage Markov chain.

Let $f_m(n)$ be the the total number of m-state n-stage stopping times. Then we have the following recursive formula :

THEOREM 2.1.

$$f_m(n) = 1 + \sum_{k=0}^{m} (f_m(n-1) - 1)^{m-k} \ {}_m\mathrm{C}_k \qquad m \geq 1, \ n \geq 2 \qquad (2.1)$$

$$f_m(1) = 2 \qquad m \geq 1. \qquad (2.2)$$

PROOF. We consider an m-state n-stage stopping time τ, which does not stop immediately at starting point (on stage 0). Let $s_{i_1}, s_{i_2}, \cdots s_{i_k}$ be the states, at which τ stops on stage 1, that is,

$$\begin{aligned} & \{\omega | \tau(\omega) = 1\} \\ = \ & \{s_0 s_{i_1} x_2 \cdots x_n, \ s_0 s_{i_2} x_2 \cdots {}_0 x_n, \ \cdots, s s_{i_k} x_2 \cdots x_n, \ x_i \in S, \ 2 \leqq i \leqq n\}. \end{aligned}$$

For each $s \in S$, we define

$$\mu(x_1 x_2 \cdots x_n) = \tau(s x_1 x_2 \cdots x_n).$$

Then μ is a stopping rule of m-state $(n-1)$-stage stopping problem. Thus for any

$$s \in S - \{s_{i_1}, s_{i_2}, \cdots, s_{i_k}\}$$

the total number of m-state n-stage stopping rules which start at state s is

$$(f_m(n-1) - 1)^{m-k}.$$

Since the total number of state $x_1 \in S$ is ${}_m\mathrm{C}_k$, it holds that the total number of τ, which do not stop immediately at start point is

$$\sum_{k=0}^{m} (f_m(n-1) - 1)^{m-k} \ {}_m\mathrm{C}_k.$$

Adding one stopping time which stops immediately on stage 0, we have the formula (2.1).

Let us take $n = 1$. There exists two stopping rules: one is immidiate stop and the other is not. Thus (2.2) is valid. □

THEOREM 2.2.

$$f_m(n+1) = 1 + (f_m(n))^m \qquad n \geq 1, \ \ f_m(1) = 2. \qquad (2.3)$$

PROOF. We prove this by induction of n. From (2.1) and (2.2), substituting $n = 2$ we have

$$\begin{aligned} f_m(2) &= 1 + \sum_{k=0}^{m} 1^{m-k} \ {}_m\mathrm{C}_k \\ &= 1 + 2^m. \end{aligned}$$

Thus theorem holds for $n = 1$. Now, we suppose that it hold for $n = k$. Then, by theorem 2.1, we have

$$\begin{aligned} f_m(k+2) &= 1 + \sum_{k=0}^{m} (f_m(k+1) - 1)^{m-k} \ {}_m\mathrm{C}_k \\ &= 1 + \sum_{k=0}^{m} ((f_m(k))^m)^{m-k} \ {}_m\mathrm{C}_k \\ &= 1 + (1 + (f_m(k))^m)^m \\ &= 1 + (f_m(k+1))^m. \end{aligned}$$

This last equation shows that it holds for $n = k + 1$. This completes the proof. □

The recursive formula generates the following :

$$\begin{aligned} f_m(3) &= 1 + (1 + 2^m)^m \\ f_m(4) &= 1 + (1 + (1 + 2^m)^m)^m \\ f_m(n) &= \underbrace{1 + (1 + (1 + (\cdots (1 + (1}_{n-1 \text{ elements}} + 2^m)^m)^m \cdots)^m. \end{aligned}$$

Table 1 shows a list of explicit numbers for two-state ($m = 2$) and three-state ($m = 3$) models.

$n \backslash m$	2	3
0	1	1
1	2	2
2	5	9
3	26	730
4	677	389,017,001
5	458,330	$5.887\ldots \times 10^{25}$

Table 1 Total numbers of m-stage n-state stopping times

3. Optimal Stopping Problem

Let $\{X_t\}_0^n$ be an n-stage Markov chain on m-state space $S = \{s_1, s_2, \ldots, s_m\}$ with a transition law $p = \{p(\cdot|\cdot)\}$ and a *preassigned* initial state $x_0 \in S$; $X_0 = x_0$. Let $\mathcal{F}_t^u$ be the set of all subsets in Ω which are determined by random variables $\{X_t, X_{t+1}, \ldots, X_u\}$. Let us take $\mathbb{N} = \{0, 1, \ldots, n\}$. A mapping $\tau : \Omega \to \mathbb{N}$ is called a *stopping time* if

$$\{\tau = t\} = \{x_0 x_1 \ldots x_n \,|\, \tau(x_0 x_1 \ldots x_n) = t\} \in \mathcal{F}_0^t \quad \forall t \in \mathbb{N}.$$

The stopping time τ is called $\{\mathcal{F}_0^t\}_0^n$*-adapted.* Let $\mathcal{T}_0^n$ be the set of all such stopping times. Any stopping time $\tau \in \mathcal{T}_0^n$ generates a *stopped state* (ramdom variable) $X_\tau : \Omega \to R^1$:

$$X_\tau(\omega) = X_{\tau(\omega)}(\omega)$$

and a *stopped reward* (ramdom variable) $g_\tau : \Omega \to R^1$:

$$g_\tau(\omega) = g_{\tau(\omega)}(X_\tau(\omega)).$$

We remark that the expected value $E_{x_0}[g_\tau]$ is expressed by sum of multiple sums :

$$E_{x_0}[g_\tau] = \sum_{t=0}^{n} \sum_{\{\tau=t\}} g_t(x_t)p(x_1|x_0)p(x_2|x_1)\cdots p(x_t|x_{t-1}).$$

Now we consider the problem of maximizing an expected value of stopped process :

$$\mathrm{T}_0(x_0) \qquad \text{Maximize} \quad E_{x_0}[g_\tau] \quad \text{subject to} \quad \tau \in \mathcal{T}_0^n.$$

3.1. Optimality

Let us define the sequence of functions $\{v_t\}_0^n$ backwardly as follows :

$$\begin{aligned} v_n(x_n) &= g_n(x_n) \\ v_{n-1}(x_{n-1}) &= \mathrm{Max}[g_{n-1}(x_{n-1}), E_{x_{n-1}}(v_n(X_n))] \\ &\vdots \\ v_1(x_1) &= \mathrm{Max}[g_1(x_1), E_{x_1}(v_2(X_2))] \\ v_0(x_0) &= \mathrm{Max}[g_0(x_0), E_{x_0}(v_1(X_1))] \end{aligned}$$

where E_x is the one-step expectation operator induced from the Markov transition matrix $p(\cdot|\cdot)$:

$$E_x(h(X_{t+1})) = \sum_{y \in X} h(y)p(y|x).$$

We define τ^* for Markov chain $\{X_t\}_0^n$ which starts at state x_0 on stage 0 : $X_0 = x_0$. For $\omega = x_0x_1x_2x_3\cdots x_{n-1}x_n$, let

$$\tau^*(\omega) \textit{ be the first } n \textit{ such that } g_n(x_n) = v_n(x_n).$$

Then we see that

$$\tau^* \in \mathcal{T}_0^n.$$

This is based upon the observation that

$$\tau^*(\omega) = t$$

if and only if

$$\begin{aligned} g_0(x_0) &< v_0(x_0) \\ g_1(x_1) &< v_1(x_1) \\ &\vdots \\ g_{t-1}(x_{t-1}) &< v_{t-1}(x_{t-1}) \\ g_t(x_t) &= v_t(x_t). \end{aligned}$$

Theorem 3.1.

$$E_{s_0}[g_\tau] \leqq v_0(s_0) \qquad \forall \tau \in \mathcal{T}_0^n. \tag{3.1}$$

Proof. We note that

$$\begin{aligned} E_{s_0}[g_\tau] \;=\; & \sum_{\{\tau=0\}} g_0(s_0) + \sum_{\{\tau=1\}} p(x_1|s_0)g_1(x_1) + \cdots \\ & + \sum_{\{\tau=n-1\}} p(x_1|s_0)p(x_2|x_1)\cdots p(x_{n-1}|x_{n-2})g_{n-1}(x_{n-1}) \\ & + \sum_{\{\tau=n\}} p(x_1|s_0)p(x_2|x_1)\cdots p(x_n|x_{n-1})g_n(x_n). \end{aligned} \tag{3.2}$$

From the definitions of v_n and v_{n-1}, we have

the sum of the last two terms

$$\begin{aligned} = & \sum_{\{\tau=n-1\}} p(x_1|s_0)p(x_2|x_1)\cdots p(x_{n-1}|x_{n-2})g_{n-1}(x_{n-1}) \\ & + \sum_{\{\tau>n-1\}} p(x_1|s_0)p(x_2|x_1)\cdots\cdots p(x_n|x_{n-1})\left[\sum_{x_n\in S} p(x_n|x_{n-1})v_n(x_n)\right] \\ \leqq & \sum_{\{\tau=n-1\}} p(x_1|s_0)p(x_2|x_1)\cdots p(x_{n-1}|x_{n-2})v_{n-1}(x_{n-1}) \\ & + \sum_{\{\tau>n-1\}} p(x_1|s_0)p(x_2|x_1)\cdots\cdots p(x_n|x_{n-1})v_{n-1}(x_{n-1}) \\ = & \sum_{\{\tau\geq n-1\}} p(x_1|s_0)p(x_2|x_1)\cdots p(x_{n-1}|x_{n-2})v_{n-1}(x_{n-1}). \end{aligned}$$

Substituting this inequality into (3.2), we have

$$\begin{aligned} & E_{s_0}[g_\tau] \\ \leqq & \sum_{\{\tau=0\}} g_0(s_0) + \sum_{\{\tau=1\}} p(x_1|s_0)g_1(x_1) + \cdots \\ & + \sum_{\{\tau=n-2\}} p(x_1|s_0)p(x_2|x_1)\cdots p(x_{n-2}|x_{n-3})g_{n-2}(x_{n-2}) \\ & + \sum_{\{\tau\geq n-1\}} p(x_1|s_0)p(x_2|x_1)\cdots p(x_{n-1}|x_{n-2})v_{n-1}(x_{n-1}) \\ = & \sum_{\{\tau=0\}} g_0(s_0) + \sum_{\{\tau=1\}} p(x_1|s_0)g_1(x_1) + \cdots \\ & + \sum_{\{\tau=n-2\}} p(x_1|s_0)p(x_2|x_1)\cdots p(x_{n-2}|x_{n-3})g_{n-2}(x_{n-2}) \\ & + \sum_{\{\tau>n-2\}} p(x_1|s_0)p(x_2|x_1)\cdots p(x_{n-2}|x_{n-3})\left[\sum_{x_{n-2}\in S} p(x_{n-1}|x_{n-2})v_{n-1}(x_{n-1})\right] \end{aligned} \tag{3.3}$$

The definition of v_{n-2} shows that the sum of last two terms in (3.3) is dominated by

$$\sum_{\{\tau\geq n-2\}} p(x_1|s_0)p(x_2|x_1)\cdots p(x_{n-2}|x_{n-3})v_{n-2}(x_{n-2}).$$

This in turn yields

$$\begin{aligned} E_{s_0}[g_\tau] \leqq & \sum_{\{\tau=0\}} g_0(s_0) + \sum_{\{\tau=1\}} p(x_1|s_0)g_1(x_1) + \cdots \\ & + \sum_{\{\tau=n-3\}} p(x_1|s_0)p(x_2|x_1)\cdots p(x_{n-3}|x_{n-4})g_{n-3}(x_{n-3}) \\ & + \sum_{\{\tau\geq n-2\}} p(x_1|s_0)p(x_2|x_1)\cdots p(x_{n-3}|x_{n-3})v_{n-2}(x_{n-2}). \end{aligned}$$

Repeating this argument, we have

$$E_{s_0}[g_\tau] \leqq \sum_{\{\tau=0\}} g_0(s_0) + \sum_{\{\tau\geq 1\}} p(x_1|s_0)v_1(x_1).$$

Let us assume that $\{\tau = 0\} = \Omega$. Then we have $v_0(s_0) = g_0(s_0)$ on Ω. This implies

$$E_{s_0}[g_\tau] \leqq v_0(s_0).$$

Otherwise, we have $\{\tau = 0\} = \phi$. Then we get

$$E_{s_0}[g_\tau] \leqq \sum_{x_1\in S} p(x_1|s_0)v_1(x_1) \leqq v_0(s_0).$$

Thus we have the desired inequality (3.1). □

THEOREM 3.2.

$$E_{s_0}[g_{\tau^*}] = v_0(s_0).$$

PROOF. We note that the defintion of τ^* keeps the equality for all inequalities in proof of Theorem 3.1. □

3.2. Characterization

Let two sequences of functions $\{f_t\}_0^n$, $\{h_t\}_0^n$ on S be given. Then the process $\{f_t(X_t)\}_0^n$ is said to be *supermartingale* (resp. *martingale, submartingale*) if $f_t(x) \geq$ (resp. $=$, $\leqq$) $Tf_{t+1}(x)$ $x \in X$, $0 \leqq t \leqq n-1$, where

$$Tf_{t+1}(x) = E_x[f_{t+1}(X_{t+1})] = \sum_{y\in X} f_{t+1}(y)p(y|x)$$

denotes the expected value of tomorrow's reward function f_{t+1} given today's sate $X_t = x$.

The process $\{f_t(X_t)\}$ is said to *dominate* the process $\{h_t(X_t)\}$ if $f_t(x) \geq h_t(x)$ $x \in$, $0 \leqq t \leqq n$.

A supermartingale $\{f_t(X_t)\}$ which dominates $\{h_t(X_t)\}$ is said to be *minimal* if every supermartingale which dominates $\{h_t(X_t)\}$ dominates $\{f_t(X_t)\}$.

THEOREM 3.3. *(Characterization) The value process $\{v_t(X_t)\}$ is the minimal supermartingale which dominates the stopping-reward process $\{g_t(X_t)\}$.*

PROOF. Let $\{u_t(X_t)\}$ be any supermartingale which dominates the stopping-reward process $\{g_t(X_t)\}$. Since

$$u_n(x_n) \geq g_n(x_n) = v_n(x_n) \qquad x_n \in S,$$

we have

$$u_{n-1}(x_{n-1}) \geq E_{x_{n-1}}(u_n(X_n)) = E_{x_{n-1}}(g_n(X_n)).$$

Further the domination impiles that

$$u_{n-1}(x_{n-1}) \geq g_{n-1}(x_{n-1}) \qquad x_{n-1} \in S.$$

Thus we have

$$u_{n-1}(x_{n-1}) \geq \mathrm{Max}[g_{n-1}(x_{n-1}), E_{x_{n-1}}(u_n(X_n))] = v_{n-1}(x_{n-1}).$$

Repeating these arguments, we have

$$u_t(x_t) \geq v_t(x_t) \qquad 0 \leqq t \leqq n, \ x_t \in S.$$

This shows that $\{v_t(X_t)\}$ is minimal, which completes the proof. □

Let $\mathfrak{M}$ be the set of all sequences of functions $\{u_t\}_0^n$ such that $\{u_t(X_t)\}$ is a supermartingale which dominates the stopping-reward process $\{g_t(X_t)\}$.

THEOREM 3.4. *If there exists $\{u_t\}_0^n \in \mathfrak{M}$ such that any $\{h_t\}_0^n \in \mathfrak{M}$ enjoys the property that $\{h_t(X_t)\}$ dominates $\{u_t(X_t)\}$. Then it holds that*

$$u_t(x_t) = v_t(x_t) \qquad 0 \leqq t \leqq n, \ \ \forall x_t \in S.$$

PROOF. Since $\{u_t\}_0^n \in \mathfrak{M}$, we have from Theorem 3.3

$$u_t(x_t) \geq v_t(x_t) \qquad 0 \leqq t \leqq n, \ \ \forall x_t \in S.$$

On the other hand, $\{v_t\}_0^n \in \mathfrak{M}$ enjoys the domination property. This implies that the reverse inequality. Thus we have the equality, which completes the proof. □

Acknowledgement

Author would like to thank Prof. Seiichi Iwamoto for helpful comments and suggestions.

References

Bellman, R.E. (1957). *Dynamic Programming*, NJ:Princeton Univ. Press.

Chow, Y.S., Robbins, H. and Siegmund, D. (1971). *Great Expectations: The Theory of Optimal Stopping*, Boston: Hunghton Mifflin Company.

Howard, R.A. (1960). *Dynamic Programming and Markov Processes*, Mass.: MIT Press.

Iwamoto, S. (2002). *Optimal stopping in fuzzy environment*, Proc. 9-th Bellman Continuum, Beijing, 264-269.

Shiryaev, A.N. (1978). *Optimal Stopping Rules*, New York: Springer-Verlag.

Shiryaev, A.N. (1999). *Essentials of Stochastic Finance*, Singapore: World Scientific.

Snell, J.L. (1952). *Applications of martingale system theorems*, Transactions of the American Mathematical Society,**73**, 171–176.

Received October 17, 2002
Revised April 17, 2003