数理解析研究所講究録 1337 巻 2003 年 32-43

When To Stop Accumulating Reward/Cost

九州大学・経済学研究院　岩本　誠一 (Seiichi Iwamoto)
Graduate School of Economics,
Kyushu University

Abstract

This paper studies an optimal stopping problem for three reward accumulation processes: terminal process, additive process and minimum process. The terminal process together with its optimal structure is well known. We show through dynamic programming that both additive process and minimum process have an optimal stopping time. The additive process admits the linearity of expectation operator. However, the minimum process does not admit the linearity. We apply an invariant imbedding approach, which expands the original state space by one dimension. A basic idear is a minimal Markovization of non-Markov process.

1 Introduction

In this paper, we consider the optimal stopping problem where the reward accumulation is terminal, additive and minimum. The theory of optimal stopping of terminal process has been studied both by dynamic programming [1] and by Snell's envelop method [4,15, 17]. It is difficult to discrimate between both approaches. The dynamic programming is methodorogical, and Snell's envelop is characteristic. In fact, both are equivalent. Here we rather consider dynamic programming approach.

2 Terminal Process

Let $\{X_n\}_0^N$ be a Markov chain on a finite state space X with a transition law $p = \{p(\cdot|\cdot)\}$. Let $g_n : X \to R^1$ be a *stop reward* for $0 \leq n \leq N$. We call $g = \{g_n\}$ a *reward sequence* (or *stopping-reward sequence*). Then a sequence of random rewards $\{g_n(X_n)\}_0^N$ is specified. The *reward process* (or *stopping-reward process*) $\{g_n(X_n)\}_0^N$ is called *terminal*. When a decision-maker stops the terminal reward process at state x_n on n-th stage, he/she will get the reward $g_n(x_n)$. His/her problem is when to stop it. This is an *optimal stopping*

Let $X^k := \overbrace{X \times X \times \cdots \times X}^{k \text{ times}}$ be the direct product of k state spaces X. We take $\Omega := X^{N+1}$; the set of all paths $\omega = x_0 x_1 \cdots x_N$:

$$\Omega = \{\omega = x_0 x_1 \cdots x_N \mid x_n \in X,\ 0 \le n \le N\}.$$

Let $\mathcal{F}_m^n$ be the set of all subsets in Ω which are determined by random variables $\{X_m, X_{m+1}, \ldots, X_n\}$, where $X_k : \Omega \to X$ is the projection, $X_k(\omega) = x_k$. Strictly, $\mathcal{F}_m^n$ is the σ-field on Ω generated by the set of all subsets of the form

$$\{X_m = x_m, X_{m+1} = x_{m+1}, \ldots, X_n = x_n\}\ (\subset\ \Omega)$$

where $x_m, x_{m+1}, \ldots, x_n$ are all elements in state space X. Let us take $\mathbb{N} = \{0, 1, \ldots, N\}$. A mapping $\tau : \Omega \to \mathbb{N}$ is called a *stopping time* if

$$\{\tau = n\} = \{x_0 x_1 \ldots x_N \mid \tau(x_0 x_1 \ldots x_N) = n\} \in \mathcal{F}_0^n \quad \forall n \in \mathbb{N}.$$

The stopping time τ is called $\{\mathcal{F}_0^n\}_0^N$*-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* (random variable) $X_\tau : \Omega \to X$:

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

and a *stopped reward* (random 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 :

$$\begin{aligned} E_{x_0}[g_\tau] &= \sum_{n=0}^{N} \sum_{\{\tau=n\}} g_n(x_n) P_{x_0}(X_n = x_n) \\ &= \sum_{n=0}^{N} \sum_{\{\tau=n\}} g_n(x_n) p(x_1|x_0) p(x_2|x_1) \cdots p(x_n|x_{n-1}). \end{aligned}$$

Now we consider the problem of maximizing an expected value of stopped process with *terminal* criterion [4, 14, 15, 17]:

$$\mathrm{T}_0(x_0) \qquad \text{Max} \quad E_{x_0}[g_\tau] \quad \text{s.t.} \quad \tau \in \mathcal{T}_0^N.$$

An invariant imbedding approach begins with taking a subprocess which starts at state $x_n (\in X)$ on n-th stage and terminates on the final N-th stage :

$$\mathrm{T}_n(x_n) \qquad \text{Max} \quad E_{x_n}[g_\tau] \quad \text{s.t.} \quad \tau \in \mathcal{T}_n^N$$

where $\mathcal{T}_n^N$ is the set of all stopping times which take values in $\{n, n+1, \ldots, N\}$. Let $v_n(x_n)$ be the maximum value of $\mathrm{T}_n(x_n)$, where

$$v_N(x_N) \triangleq g_N(x_N) \qquad x_N \in X.$$

Then we have the the backward recursive equation :

Theorem 2.1

$$\begin{cases} v_N(x) = g_N(x) \quad & x \in X \\ v_n(x) = \text{Max}\left[\, g_n(x),\, E_x[v_{n+1}(X_{n+1})]\,\right] \\ \qquad\qquad\qquad x \in X, \quad 0 \le n \le N-1 \end{cases}$$

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

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

Proof. The proof is done through an equivalent Markov decision problem in the following section. □

2.1 Optimal stopping time

Theorem 2.2 *The stopping time τ^* :*

$$\tau^*(\omega) = \min\{n : v_n(x_n) = g_n(x_n)\} \qquad \omega = x_0 x_1 \cdots x_N \in \Omega$$

is optimal :

$$E_{x_0}[g_{\tau^*}] \ge E_{x_0}[g_\tau] \qquad \forall \tau \in \mathcal{T}_0^N.$$

Let two sequences of functions $\{f_n\}_0^N$, $\{h_n\}_0^N$ on X be given. Then the process $\{f_n(X_n)\}_0^N$ is said to be *supermartingale* (resp. *martingale, submartingale*) if $f_n(x) \ge$ (resp. $=$, $\le$) $Tf_{n+1}(x)$ $x \in X$, $0 \le n \le N-1$, where

$$Tf_{n+1}(x) = E_x[f_{n+1}(X_{n+1})].$$

In this case, we say that the sequence of functions $\{f_n\}$ is *excessive.*

The process $\{f_n(X_n)\}$ is said to *dominate* the process $\{h_n(X_n)\}$ if $f_n(x) \ge h_n(x)$ $x \in$, $0 \le n \le N$. We also say that $\{f_n\}$ is *majorant* of $\{h_n\}$.

A supermartingale $\{f_n(X_n)\}$ which dominates $\{h_n(X_n)\}$ is said to be *minimal* if every supermartingale which dominates $\{h_n(X_n)\}$ dominates $\{f_n(X_n)\}$. In this case, we say that $\{f_n\}$ is the *smallest excessice majorant* of $\{h_n\}$.

Theorem 2.3 *(Characterization) The value process $\{v_n(X_n)\}$ is the minimal supermartingale which dominates the stopping-reward process $\{g_n(X_n)\}$.*

It is also said that the value functions $\{v_n\}$ is the *smallest excessive majorant* of $\{g_n\}$. Let a stopping time τ and a reward sequence $f = \{f_n\}$ be given. Then we define a *stopped process* $f^\tau = \{f_n^\tau\}$ by

$$f_n^\tau := f_{\tau \wedge n} \;\text{ or }\; f_{\tau \wedge n}(X_{\tau \wedge n})(\omega) := f_{\tau(\omega) \wedge n}(X_{\tau(\omega) \wedge n}(\omega)) \qquad \omega \in \Omega,\ 0 \le n \le N.$$

We note that

$$f_n^{\tau}(\omega) = f_{\tau(\omega)\wedge n}(X_{\tau(\omega)\wedge n}(\omega)) = \begin{cases} f_n(X_n(\omega)) & \tau(\omega) \geq n \\ f_{\tau(\omega)}(X_{\tau(\omega)}(\omega)) & \tau(\omega) < n. \end{cases}$$

Then $\{f_{\tau\wedge n}(X_{\tau\wedge n})\}$ is called the *stopped process* for process $\{f_n(X_n)\}$ by the stopping time τ. Thus the stopped process $\{f_{\tau\wedge n}(X_{\tau\wedge n})\}$ is supermatingale if and only if for each n $(0 \leq n \leq N-1)$

$$f_{\tau(\omega)\wedge n}(x_{\tau(\omega)\wedge n}) \geq E_{x_{\tau(\omega)\wedge n}}[f_{\tau(\omega)\wedge(n+1)}(X_{\tau(\omega)\wedge(n+1)})] \qquad \text{a.e.}$$

This implies for each n

$$\text{on } \{\tau \geq n+1\}, \quad f_n(x_n) \geq E_{x_n}[f_{n+1}(X_{n+1})]$$

and

$$\text{on } \{\tau = n\}, \quad f_n(x_n) \geq E_{x_{\tau(\omega)}}[f_{\tau(\omega)}(X_{\tau(\omega)})]$$

$$\text{on } \{\tau \leq n-1\}, \quad f_{\tau(\omega)}(x_{\tau(\omega)}) \geq E_{x_{\tau(\omega)}}[f_{\tau(\omega)}(X_{\tau(\omega)})].$$

The latter two inequalities are satisfied with the equality. The supermartingaleness is equivalent to the first inequality.

Theorem 2.4 *(Martingale) The stopped process* $v^{\tau^*} = \{v_{\tau^*\wedge n}(X_{\tau^*\wedge n})\}$ *of process* $\{v_n(X_n)\}$ *by the optimal stopping time* τ^* *is a martingale.*

Proof. We see that for each n

$$\text{on } \{\tau^* \geq n+1\}, \quad v_n(x_n) = E_{x_n}[v_{n+1}(X_{n+1})]. \qquad \square$$

Theorem 2.5 *(Optimality) A stopping time* τ *is optimal if and only if* (i) *the stopped rewards are equal :* $v_\tau = g_\tau$ *a.e., and* (ii) *the stopped process* $v^\tau = \{v_{\tau\wedge n}(X_{\tau\wedge n})\}$ *is a martingale.*

3 Additive Process

In this section, we assume, in addition, that $r_n : X \to R^1$ be a *continuation reward* for $0 \leq n \leq N-1$.

We consider the problem of maximizing an expected value of stopped process with *additive* criterion [5]:

$$\mathrm{A}_0(x_0) \qquad \begin{array}{ll} \text{Max} & E_{x_0}[r_0 + r_1 + \cdots + r_{\tau-1} + g_\tau] \\ \text{s.t.} & \tau \in \mathcal{T}_0^N. \end{array}$$

We note that the expected value of additive reward is the following sum of multiple sums :

$$
\begin{aligned}
&E_{x_0}[\, r_0 + \cdots + r_{\tau-1} + g_\tau] \\
&= \sum_{n=0}^{N} \sum_{\{\tau=n\}} \left[\sum_{k=0}^{n} r_k(x_k) + g_n(x_n) \right] p(x_1|x_0)p(x_2|x_1)\cdots p(x_n|x_{n-1}).
\end{aligned}
$$

Then we have the corresponding recursive equation :

Theorem 3.1

$$
\begin{cases}
v_N(x) = g_N(x) & x \in X \\
v_n(x) = \text{Max}\,[\, g_n(x),\, E_x[r_n(x) + v_{n+1}(X_{n+1})]\,] & \\
 & x \in X, \quad 0 \le n \le N-1
\end{cases}
\tag{1}
$$

Here we remark that the linearity of expectation operator admits

$$
E_x[r_n(x) + v_{n+1}(X_{n+1})] = r_n(x) + E_x[v_{n+1}(X_{n+1})].
$$

Theorem 3.2 *The stopping time τ^* :*

$$
\tau^*(\omega) = \min\{n : v_n(x_n) = g_n(x_n)\} \qquad \omega = x_0 x_1 \cdots x_N
$$

is optimal :

$$
E_{x_0}[\, r_0 + \cdots + r_{\tau^*-1} + g_{\tau^*}\,] \ge E_{x_0}[\, r_0 + \cdots + r_{\tau-1} + g_\tau\,] \qquad \forall \tau \in \mathcal{T}_0^N.
$$

Let two sequences $\{f_n\}_0^N$, $\{h_n\}_0^N$ be given. Then the process $\{f_n(X_n)\}_0^N$ is said to be *r-supermartingale* if $f_n(x) \ge Rf_{n+1}(x) \quad x \in X,\ 0 \le n \le N-1$, where

$$
Rf_{n+1}(x) = E_x[r_n(x) + f_{n+1}(X_{n+1})].
$$

We also say that the sequence $\{f_n\}$ is *r-excessive.*

Theorem 3.3 *(Characterization) The value process $\{v_n(X_n)\}$ is the minimal r-supermartingale which dominates the stopping-reward process $\{g_n(X_n)\}$.*

It is also said that the value function $\{v_n\}$ is the *smallest r-excessive majorant* of $\{g_n\}$.

3.1 DP Solution

Let us condider an two-state two-stage model (2-2 model) for additive criterion

$$
\begin{aligned}
&\text{Max } E_{x_0}[r_0(X_0) + \cdots + r_{\tau-1}(X_{\tau-1}) + g_\tau(X_\tau)] \\
&\text{s.t. (i) } \tau \in \mathcal{T}_0^2
\end{aligned}
$$

where the continue/stop reward $\{r_0,\ r_1;\ g_0,\ g_1,\ g_2\}$ is given in Table 1 :

x_n		s_1		s_2	
$g_0(x_0)$	$r_0(x_0)$	0.9	*0.2*	0.6	*0.1*
$g_1(x_1)$	$r_1(x_1)$	0.7	*0.0*	0.8	*0.2*
$g_2(x_2)$		0.7		0.6	

Table 3 stop/continue reward

and the transition matrix is symmetric ($p = q = 1/2$).

Let us find an optimal stopping time by solving recursive equation. First, the backward recursion (1) yields an optimal solution in Markov class Π ; optimal value functions

$$v_0 = v_0(x_0),\ v_1 = v_1(x_1),\ v_2 = v_2(x_2)$$

and an optimal policy

$$\gamma^* = \{\gamma_0^*(x_0),\ \gamma_1^*(x_1)\}.$$

$$v_2(s_1) = 0.7$$
$$v_2(s_2) = 0.6$$

$$v_1(s_1) = \mathrm{Max}[0.7,\ 0.0 + \frac{1}{2}0.7 + \frac{1}{2}0.6] = 0.7$$
$$\gamma_1^*(s_1) = \text{stop}$$
$$v_1(s_2) = \mathrm{Max}[0.8,\ 0.2 + \frac{1}{2}0.7 + \frac{1}{2}0.6] = 0.85$$
$$\gamma_1^*(s_2) = \text{continue}$$

$$v_0(s_1) = \mathrm{Max}[0.9,\ 0.2 + \frac{1}{2}0.7 + \frac{1}{2}0.85] = 0.975$$
$$\gamma_0^*(s_1) = \text{continue}$$
$$v_0(s_2) = \mathrm{Max}[0.9,\ 0.1 + \frac{1}{2}0.7 + \frac{1}{2}0.85] = 0.90$$
$$\gamma_1^*(s_2) = \text{stop}.$$

The optimal solution is tabulated as

x_n	$v_2(x_2)$	$v_1(x_1)$	$\gamma_1^*(x_1)$	$v_0(x_0)$	$\gamma_0^*(x_0)$
s_1	0.7	0.7	s	0.975	c
s_2	0.6	0.85	c	0.9	s

where s and c denote stop and continue, respectively.

Finally, an optimal stpping time τ^* from $x_0 = s_1$ is described through Theorem 3.2. In fact, for any path $\omega = x_0x_1x_2$, $\tau^*(\omega)$ takes the following time :

$$\begin{aligned}
&\tau^*(s_1s_1x_2) = 1\\
&\quad 0.975 = v_0(s_1) > g_0(s_1) = 0.9\\
&\quad 0.7 = v_1(s_1) = g_1(s_1) = 0.7
\end{aligned}$$

$$\begin{aligned}
&\tau^*(s_1s_2s_1) = 2\\
&\quad 0.975 = v_0(s_1) > g_0(s_1) = 0.9\\
&\quad 0.85 = v_1(s_2) > g_1(s_2) = 0.8\\
&\quad 0.7 = v_2(s_1) = g_2(s_1) = 0.7
\end{aligned}$$

$$\begin{aligned}
&\tau^*(s_1s_2s_2) = 2\\
&\quad 0.975 = v_0(s_1) > g_0(s_1) = 0.9\\
&\quad 0.85 = v_1(s_2) > g_1(s_2) = 0.8\\
&\quad 0.6 = v_2(s_2) = g_2(s_2) = 0.6
\end{aligned}$$

4 Minimum Process

We consider the problem of maximizing an expected value of stopped process with *minimum* criterion ([3,9,16,19]. As for nonstopping but control problems, see [6–8,10–12]):

$$\mathrm{M}_0(x_0) \qquad \begin{array}{ll} \text{Max} & E_{x_0}[r_0 \wedge r_1 \wedge \cdots \wedge r_{\tau-1} \wedge g_\tau] \\ \text{s.t.} & \tau \in \mathcal{T}_0^N. \end{array}$$

The expected value of minimum reward is the sum of multiple sums as follows :

$$\begin{aligned}
&E_{x_0}[r_0 \wedge \cdots \wedge r_{\tau-1} \wedge g_\tau]\\
&= \sum_{n=0}^{N} \sum_{\{\tau=n\}} [r_0(x_0) \wedge \cdots \wedge r_{n-1}(x_{n-1}) \wedge g_n(x_n)]\, p(x_1|x_0)p(x_2|x_1)\cdots p(x_n|x_{n-1}).
\end{aligned}$$

Here we mention that the linearity of expectation operator does *not* admit the equality

$$E[c \wedge Z] = c \wedge E[Z]$$

where c is a constant and Z is a random variable.

So, we imbed $\mathrm{M}_0(x_0)$ into a new class of additional parametric subproblems [2, 13]. First we define the *past-valued (cumulative) random variables* $\{\tilde{\Lambda}_n\}$ up to n-th stage and the *past-value sets* $\{\Lambda_n\}$ they take :

$$\begin{aligned}
\tilde{\Lambda}_0 &\triangleq \tilde{\lambda}_0 \quad \text{where } \tilde{\lambda}_0 \text{ is larger than or equal to } g_n(x), r_n(x)\\
\tilde{\Lambda}_n &\triangleq r_0(X_0) \wedge \cdots \wedge r_{n-1}(X_{n-1}) \quad 1 \le n \le N,
\end{aligned}$$

$$\begin{aligned}
\Lambda_0 &\triangleq \{\tilde{\lambda}_0\}\\
\Lambda_n &\triangleq \left\{ \lambda_n \,\middle|\, \begin{array}{c} \lambda_n = r_0(x_0) \wedge \cdots \wedge r_{n-1}(x_{n-1}), \\ (x_0, \ldots, x_{n-1}) \in X \times \cdots \times X \end{array} \right\} \qquad 1 \le n \le N.
\end{aligned}$$

The minimum criterion is terminal now :

$$r_0(X_0) \wedge \cdots \wedge r_{\tau-1}(X_{\tau-1}) \wedge g_\tau(X_\tau) \;=\; \tilde{\Lambda}_\tau \wedge g_\tau(X_\tau).$$

We have

Lemma 4.1 *(Forward recursive formulae)*

$$\begin{aligned}
&\tilde{\Lambda}_0 = \tilde{\lambda}_0 \\
&\tilde{\Lambda}_{n+1} = \tilde{\Lambda}_n \wedge r_n(X_n) \qquad 0 \le n \le N\text{–}1, \\
&\Lambda_0 = \{\tilde{\lambda}_0\} \\
&\Lambda_{n+1} = \{\lambda \wedge r_n(x) \mid \lambda \in \Lambda_n,\ x \in X\} \qquad 0 \le n \le N\text{–}1.
\end{aligned} \tag{2}$$

Let us now expand the original state space X to a direct product space :

$$Y_n \triangleq X \times \Lambda_n \qquad 0 \le n \le N.$$

We define a sequence of *stopping reward functions* $\{G_n\}_0^N$ by

$$G_n(x;\lambda) \triangleq \lambda \wedge g_n(x) \qquad (x;\lambda) \in Y_n$$

and a *nonstationary Markov transition law* $q = \{q_n\}_0^{N-1}$ by

$$q_n(y;\mu \mid x;\lambda) \;\triangleq\; \begin{cases} p(y|x) & \text{if } \lambda \wedge r_n(x) = \mu \\ 0 & \text{otherwise.} \end{cases}$$

Then $\{(X_n, \tilde{\Lambda}_n)\}_0^N$ is a Markov process on state spaces $\{Y_n\}$ with transition law q. We consider the *terminal* criterion $\{G_n\}_0^N$ on the expanded process :

$$\overline{\mathrm{T}}_0(y_0) \qquad \text{Max } \mathbb{E}_{y_0}[G_\tau] \quad \text{s.t. } \tau \in \tilde{\mathfrak{T}}_0^N$$

where $y_0 = (x_0; \tilde{\lambda}_0)$, and $\tilde{\mathfrak{T}}_n^N$ is the set of all stopping times which take values in $\{n, n+1, \ldots, N\}$ on the new Markov chain.

Now we take a subprocess which starts at state $y_n = (x_n; \lambda_n) (\in Y_n)$ on n-th stage :

$$\overline{\mathrm{T}}_n(y_n) \qquad \text{Max } \mathbb{E}_{y_n}[G_\tau] \quad \text{s.t. } \tau \in \tilde{\mathfrak{T}}_n^N.$$

Let $v_n(y_n)$ be the maximum value of $\overline{\mathrm{T}}_n(y_n)$, where

$$v_N(y_N) \triangleq G_N(y_N) \qquad y_N \in Y_N.$$

Then we have the the backward recursive equation :

Corollary 4.1

$$\begin{cases} v_N(y) = G_N(y) \quad y \in Y_N \\ v_n(y) = \text{Max}\left[G_n(y),\, \mathbb{E}_y[v_{n+1}(Y_{n+1})]\right] \\ \qquad\qquad\qquad y \in Y_n, \quad 0 \le n \le N-1 \end{cases}$$

where $\mathbb{E}_y$ is the one-step expectation operator induced from the Markov transition probabilities $q_n(\cdot|\cdot)$:

$$\mathbb{E}_y[h(Y_{n+1})] = \sum_{z \in Y_{n+1}} h(y) q_n(z|y).$$

Corollary 4.2 *The stopping time τ^* :*

$$\tau^*(\omega) = \min\{n : v_n(y_n) = G_n(y_n)\} \qquad \omega = y_0 y_1 \cdots y_N$$

is optimal :

$$\mathbb{E}_{y_0}[G_{\tau^*}] \ge \mathbb{E}_{y_0}[G_\tau] \qquad \forall \tau \in \tilde{\mathfrak{T}}_0^N.$$

Then we have the corresponding recursive equation for the original process with minimum reward :

Theorem 4.1

$$\begin{cases} v_N(x,\lambda) = \lambda \wedge g_N(x) \qquad\qquad x \in X, \quad \lambda \in \Lambda_N \\ v_n(x,\lambda) = \text{Max}\left[\lambda \wedge g_n(x),\, E_x[v_{n+1}(X_{n+1}, \lambda \wedge r_n(x))]\right] \\ \qquad\qquad\qquad\qquad x \in X, \quad \lambda \in \Lambda_n, \quad 0 \le n \le N-1 \end{cases} \tag{3}$$

Theorem 4.2 *The stopping time τ^* :*

$$\tau^*(\omega) = \min\{n : v_n(x_n, \lambda_n) = \lambda_n \wedge g_n(x_n)\} \qquad \omega = (x_0, \tilde{\lambda}_0)(x_1, \lambda_1) \cdots (x_N, \lambda_N)$$

is optimal :

$$E_{x_0}[r_0 \wedge \cdots \wedge r_{\tau^*-1} \wedge g_{\tau^*}] \ge E_{x_0}[r_0 \wedge \cdots \wedge r_{\tau-1} \wedge g_\tau] \qquad \forall \tau \in \mathfrak{T}_0^N.$$

4.1 DP Solution

A 2-2 model is specified by :

$$\begin{aligned} &\text{Max } E_{x_0}[r_0(X_0) \wedge \cdots \wedge r_{\tau-1}(X_{\tau-1}) \wedge g_\tau(X_\tau)] \\ &\text{s.t. (i) } \tau \in \mathfrak{T}_0^2 \end{aligned}$$

where the continue/stop reward $\{r_0,\ r_1;\ g_0,\ g_1,\ g_2\}$ is given in Table 2 :

x_n		s_1		s_2	
$g_0(x_0)$	$r_0(x_0)$	0.5	*0.8*	0.6	*0.9*
$g_1(x_1)$	$r_1(x_1)$	0.7	*0.8*	0.4	*0.6*
$g_2(x_2)$		0.7		0.6	

Table 2 stop/continue reward

and the transition matrix is symmetric ($p = q = 1/2$).

First, the forward recursion (2) generates the following past-value sets :

$$\Lambda_0 = \{1.0\}, \quad \Lambda_1 = \{0.8,\ 0.9\}, \quad \Lambda_2 = \{0.6,\ 0.8\}.$$

Second, the backward recursion (3) yields an optimal solution in expanded Markov class $\tilde{\Pi}$; optimal value functions

$$v_0 = v_0(x_0; \lambda_0), \ v_1 = v_1(x_1; \lambda_1), \ v_2 = v_2(x_2; \lambda_2)$$

and an optimal policy

$$\gamma^* = \{\gamma_0^*(x_0; \lambda_0), \ \gamma_1^*(x_1; \lambda_1)\}.$$

$$v_2(s_1, 0.6) = 0.6 \wedge g_2(s_1) = 0.6 \wedge 0.7 = 0.6$$
$$v_2(s_2, 0.6) = 0.6 \wedge g_2(s_2) = 0.6 \wedge 0.6 = 0.6$$
$$v_2(s_1, 0.8) = 0.8 \wedge g_2(s_1) = 0.8 \wedge 0.7 = 0.7$$
$$v_2(s_2, 0.8) = 0.8 \wedge g_2(s_2) = 0.8 \wedge 0.6 = 0.6$$

$$v_1(s_1, 0.8) = \mathrm{Max}[0.8 \wedge 0.7, \ \frac{1}{2}0.7 + \frac{1}{2}0.6] = 0.7$$
$$\gamma_1^*(s_1, 0.8) = \text{stop}$$
$$v_1(s_2, 0.8) = \mathrm{Max}[0.8 \wedge 0.4, \ \frac{1}{2}0.6 + \frac{1}{2}0.6] = 0.6$$
$$\gamma_1^*(s_2, 0.8) = \text{continue}$$

$$v_1(s_1, 0.9) = \mathrm{Max}[0.9 \wedge 0.7, \ \frac{1}{2}0.7 + \frac{1}{2}0.6] = 0.7$$
$$\gamma_1^*(s_1, 0.9) = \text{stop}$$
$$v_1(s_2, 0.9) = \mathrm{Max}[0.9 \wedge 0.4, \ \frac{1}{2}0.6 + \frac{1}{2}0.6] = 0.6$$
$$\gamma_1^*(s_2, 0.9) = \text{continue}$$

$$v_0(s_1, 1.0) = \mathrm{Max}[1.0 \wedge 0.5, \ \frac{1}{2}0.7 + \frac{1}{2}0.6] = 0.65$$
$$\gamma_0^*(s_1, 1.0) = \text{continue}$$
$$v_0(s_2, 1.0) = \mathrm{Max}[1.0 \wedge 0.6, \ \frac{1}{2}0.7 + \frac{1}{2}0.6] = 0.65$$
$$\gamma_1^*(s_2, 0.8) = \text{continue}.$$

The optimal solution is tabulated as

$x_2 \backslash \lambda_2$	$v_2(x_2;\lambda_2)$ 0.6	0.8
s_1	0.6	0.7
s_2	0.6	0.6

$x_n \backslash \lambda_n$	$v_1(x_1;\lambda_1)$ 0.8	$\gamma_1^*(x_1;\lambda_1)$	0.9		$v_0(x_0;1)$ 1.0	$\gamma_0^*(x_0;1)$
s_1	0.7	s	0.7	s	0.65	s
s_2	0.6	c	0.6	c	0.65	s

Finally, an optimal stpping time σ^* from $(x_0, \lambda_0) = (s_1, 1.0)$ is described through Theorem 4.2. In fact, for any path $\tilde{\omega} = (x_0, 1.0)(x_1, \lambda_1)(x_2, \lambda_2)$, $\sigma^*(\tilde{\omega})$ takes the following time :

$$\sigma^*((s_1, 1.0)(s_1, 0.8)(x_2, \lambda_2)) = 1$$
$$0.65 = v_0(s_1, 1.0) > 1.0 \wedge g_0(s_1) = 1.0 \wedge 0.5 = 0.5$$
$$0.7 = v_1(s_1, 0.8) = 0.8 \wedge g_1(s_1) = 0.8 \wedge 0.7 = 0.7$$

$$\sigma^*((s_1, 1.0)(s_2, 0.8)(s_1, 0.6)) = 2$$
$$0.65 = v_0(s_1, 1.0) > 1.0 \wedge g_0(s_1) = 1.0 \wedge 0.5 = 0.5$$
$$0.7 = v_1(s_2, 0.8) > 0.8 \wedge g_1(s_2) = 0.8 \wedge 0.4 = 0.4$$
$$0.6 = v_2(s_1, 0.6) = 0.6 \wedge g_2(s_1) = 0.6 \wedge 0.7 = 0.6$$

$$\sigma^*((s_1, 1.0)(s_2, 0.8)(s_2, 0.6)) = 2$$
$$0.65 = v_0(s_1, 1.0) > 1.0 \wedge g_0(s_1) = 1.0 \wedge 0.5 = 0.5$$
$$0.7 = v_1(s_2, 0.8) > 0.8 \wedge g_1(s_2) = 0.8 \wedge 0.4 = 0.4$$
$$0.6 = v_2(s_2, 0.6) = 0.6 \wedge g_2(s_2) = 0.6 \wedge 0.6 = 0.6$$

Similarly the stopping time σ^* also turns out to be optimal from $x_0 = s_2$.

References

[1] R.E. Bellman, *Dynamic Programming*, Princeton Univ. Press, NJ, 1957.

[2] R.E. Bellman and E.D. Denman, *Invariant Imbedding*, Lect. Notes in Operation Research and Mathematical Systems, Vol. 52, Springer-Verlag, Berlin, 1971.

[3] R.E. Bellman and L.A. Zadeh, Decision-making in a fuzzy environment, *Management Sci.* **17**(1970), B141-B164.

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

[5] N. Furukawa and S. Iwamoto, Stopped decision processes on complete separable metric spaces, *J. Math. Anal. Appl.*, **31**(1970), 615–658.

[6] S. Iwamoto, Maximizing threshold probability through invariant imbedding, Ed. H.F. Wang and U.P. Wen, Proceedings of 8-th Bellman Continuum, National Tsing Hua University, Hsinchu, ROC, Dec., 2000, 17-22.

[7] S. Iwamoto, Fuzzy decision-making through three dynamic programming approaches, Ed. H.F. Wang and U.P. Wen, Proceedings of 8-th Bellman Continuum, National Tsing Hua University, Hsinchu, ROC, Dec., 2000, 23–27.

[8] S. Iwamoto, Recursive method in stochastic optimization under compound criteria, *Advances in Mathematical Economics* **3**(2001), 63–82.

[9] S. Iwamoto, Optimal stopping in fuzzy environment, Proceedings of the 9-th Bellman Continuum : Intl Workshop on Uncertain Systems and Soft Computing, July, 2002, Beijing, 264–269.

[10] S. Iwamoto, K. Tsurusaki and T. Fujita, "On Markov policies for minimax decision processes," *J. Math. Anal. Appl.*, **253**(2001), no.1, 58–78.

[11] S. Iwamoto, T. Ueno and T. Fujita, "Controlled Markov chains with utility functions," *Markov Processes and Controlled Markov Chains* : Proc. of Intl Workshop on Markov Processes and Controlled Markov Chains, Changsha, China, August, 1999, Chap. 8, Kluwer, 2002, pp.135–148.

[12] S. Iwamoto and T. Fujita, Stochastic decision-making in a fuzzy environment, *J. Operations Res. Soc. Japan* **38**(1995), 467–482.

[13] E.S. Lee, *Quasilinearization and Invariant Imbedding*, Academic Press, New York, 1968.

[14] A. Maitra and W. Sudderth, *Discrete Gambling and Stochastic Games*, Springer-Verlag, New York, 1996.

[15] A.N. Shiryaev, *Optimal Stopping Rules*, Springer-Verlag, New York, 1978.

[16] A.N. Shiryaev, *Essentials of Stochastic Finance*, World-Scientific, Singapore, 1999.

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

[18] M. Sniedovich, *Dynamic Programming*, Marcel Dekker, Inc. NY, 1992.

[19] W.E. Stein, Optimal stopping in a fuzzy environment, *Fuzzy Sets and Systems* **3**(1980), 253-259.