Fuzzy Decision Processes with an Average Reward Criterion

Masami KURANO†, Masami YASUDA‡, Jun-ichi NAKAGAMI‡
and Yuji YOSHIDA§

†*Faculty of Education,* ‡*Faculty of Science, Chiba University, Chiba 263-8522, Japan.* §*Faculty of Economics and Business Administration, Kitakyushu University, Kitakyushu 802-8577, Japan.*

Abstract

As the same framework of Fuzzy decision processes with the discounted case we will specify an average fuzzy criterion model and develop its optimization by "fuzzy max order" under appropriate conditions. The average reward is characterized, by introducing a relative value function, as a unique solution of the associated equation. Also we derive the optimality equation using the "vanishing discount factor" approach.

Keywords: Fuzzy sets; Dynamic programming; Decision theory; Average reward criterion; Relative value function; Optimality equation.

1 Introduction and notations

In the previous paper [6], Markov-type fuzzy decision processes (FDP's, for short) with a bounded fuzzy reward are defined. We have developed its optimization under the discount reward criterion. Also, the long-run average reward of some dynamic fuzzy system has been specified in our another paper [7]. However, the optimization was not considered there. In this paper, we will specify the long-run average fuzzy reward from a fuzzy policy and develop its optimization by the so-called "fuzzy max order" on the convex fuzzy numbers under the ergodicity (contraction) condition for the fuzzy state transition and the continuity condition for the fuzzy reward relation.

By introducing the relative value functions, the average reward from any admissible stationary policy is characterized as a unique solution of the associated equation. It will be useful in the policy improvement. Moreover, using the "vanishing discount factor" approach which is well-known in the theory of Markov decision processes (for example, see [12, 14]), we derive the optimality equation under the average fuzzy reward criterion. In the reminder of this section, we will give notations and some mathematical facts.

Let $E,\ E_1,\ E_2$ be convex compact subsets of a given Banach space. Throughout the paper we will denote a fuzzy set and a fuzzy relation by their membership functions. For the theory of fuzzy sets, refer to Zadeh [20] and Novàk [11].

A fuzzy set $\tilde{u} : E \to [0,1]$ on E is called convex if

$$\tilde{u}(\lambda x + (1-\lambda)y) \geq \tilde{u}(x) \wedge \tilde{u}(y), \quad x, y \in E,\ \lambda \in [0,1],$$

where $a \wedge b := \min\{a, b\}$ (c.f. [15, 16]). Also, a fuzzy relation $\tilde{h} : E_1 \times E_2 \to [0,1]$ between E_1 and E_2 is called convex if

$$\tilde{h}(\lambda x_1 + (1-\lambda)x_2, \lambda y_1 + (1-\lambda)y_2) \geq \tilde{h}(x_1, y_1) \wedge \tilde{h}(x_2, y_2)$$

for $x_1, x_2 \in E_1,\ \ y_1, y_2 \in E_2$ and $\lambda \in [0,1]$.

The α-cut ($\alpha \in [0,1]$) of the fuzzy set $\tilde{u}$ is defined by

$$\tilde{u}_\alpha := \{x \in E \mid \tilde{u}(x) \geq \alpha\}\ (\alpha > 0) \quad \text{and} \quad \tilde{u}_0 := cl\,\{x \in E \mid \tilde{u}(x) > 0\},$$

where cl denotes the closure of a set.

Let $\mathcal{F}(E)$ be the set of all convex fuzzy sets, $\tilde{u}$, on E whose membership functions are upper semi-continuous and have compact supports and the normality condition : $\sup_{x \in E} \tilde{u}(x) = 1$. We denote by $\mathcal{C}(E)$ the collection of all compact convex subsets of E and ρ_E the Hausdorff metric on $\mathcal{C}(E)$. Clearly, $\tilde{u} \in \mathcal{F}(E)$ means $\tilde{u}_\alpha \in \mathcal{C}(E)$ for all $\alpha \in [0,1]$.

Let $\mathbf{R}$ be the set of real numbers. We see from the definition that $\mathcal{C}(\mathbf{R})$ and $\mathcal{F}(\mathbf{R})$ are the set of all bounded closed intervals in $\mathbf{R}$ and all upper semi-continuous and convex fuzzy numbers on $\mathbf{R}$ with compact supports, respectively.

For the interval $[0, M]$ with a fixed positive number M,

$$\mathcal{F}([0,M]) = \{\tilde{u} \in \mathcal{F}(R) \mid \tilde{u}_0 \subset [0,M]\},$$

and $\mathcal{C}([0,M])$ be the set of all closed convex subsets of $[0, M]$. For non-empty closed intervals, the Hausdorff metric on $\mathcal{C}([0,M])$ is represented by δ, i.e.,

$$\delta([a,b],[c,d]) := \mid a - c \mid \vee \mid b - d \mid \quad \text{for} \quad [a,b],[c,d] \in \mathcal{C}([0,M]),$$

where $a \vee b = \max\{a, b\}$ for real numbers a, b.

The definitions of addition and scalar multiplication on $\mathcal{F}(\mathbf{R})$ are as follows: For $\tilde{m}, \tilde{n} \in \mathcal{F}(R)$ and $\lambda \geq 0$,

$$(\tilde{m} + \tilde{n})(x) := \sup_{x_1, x_2 \in \mathbf{R}:\ x_1 + x_2 = x} \{\tilde{m}(x_1) \wedge \tilde{n}(x_2)\}$$

and, for the scalar multiplication, let us define

$$(\lambda \hat{m})(x) := \begin{cases} \hat{m}(x/\lambda) & \text{if } \lambda > 0 \\ I_{\{0\}}(x) & \text{if } \lambda = 0, \end{cases} \qquad (x \in \mathbf{R})$$

where $I_{\{\cdot\}}(\cdot)$ is an indicator. By using the set operation $A + B := \{x + y \mid x \in A, y \in B\}$, $A + \emptyset = A$ for any non-empty sets $A, B \subset \mathbf{R}$, and $\lambda A := \{\lambda x \mid x \in A\}$ for $A \subset \mathbf{R}$, the following holds immediately.

$$(\tilde{m} + \tilde{n})_\alpha = \tilde{m}_\alpha + \tilde{n}_\alpha \quad \text{and} \quad (\lambda \tilde{m})_\alpha = \lambda \tilde{m}_\alpha \ (\alpha \in [0, 1])$$

Lemma 1.1 ([15, Theorem 2.3]).

(i) For any $\tilde{n}, \tilde{m} \in \mathcal{F}(R)$ and $\lambda \in R$, $\tilde{n} + \tilde{m} \in \mathcal{F}(R)$ and $\lambda \tilde{n} \in \mathcal{F}(R)$.

(ii) For any $\tilde{s} \in \mathcal{F}(E_1)$ and $\tilde{p} \in \mathcal{F}(E_1 \times E_2)$, then $\sup_{x \in E_1} \tilde{s}(x) \wedge \tilde{p}(x, \cdot) \in \mathcal{F}(E_2)$.

In the next section 2, we give the basic definition of FDP's and describe the optimization problem by specifying the average fuzzy reward from any fuzzy policy. In Section 3, referring [6] we give several results for the discounted case which will be also used to derive the optimality equation by the "vanishing discount factor" method in Section 5. In Section 4, the average fuzzy reward from any stationary fuzzy policy satisfying some reasonable condition is characterized as a unique solution of the associated relational equation. The numerical example is given to illustrate the theoretical results.

2 Fuzzy Decision Processes

In this section, we will formulate FDP's. By introducing the fuzzy max order on the convex fuzzy numbers the optimization problem could be described.

A fuzzy decision process, in this paper, is a controlled dynamic fuzzy system defined by four objects $(S, A, \tilde{q}, \tilde{r})$ as follows:

(i) Let S and A be a state space and an action space, which are given as convex compact subsets of some Banach space respectively. The decision process is assumed to be fuzzy itself, so that both the state of the system and the action taken at each stage are denoted by the element of $\mathcal{F}(S)$ and $\mathcal{F}(A)$, called the fuzzy state and the fuzzy action respectively.

(ii) The law of motion for the system and the fuzzy reward can be characterized by time invariant fuzzy relations $\tilde{q} \in \mathcal{F}(S \times A \times S)$ and $\tilde{r} \in \mathcal{F}(S \times A \times [0, M])$, where M is a given positive number. Explicitly, if the system is in a fuzzy state $\tilde{s} \in \mathcal{F}(S)$ and the fuzzy action $\tilde{a} \in \mathcal{F}(A)$ is chosen, then it transfers to a new fuzzy state $Q(\tilde{s}, \tilde{a})$ and a fuzzy reward $R(\tilde{s}, \tilde{a})$ is earned, where Q, R are defined by the following:

$$Q(\tilde{s}, \tilde{a})(y) := \sup_{(x,a) \in S \times A} \tilde{s}(x) \wedge \tilde{a}(a) \wedge \tilde{q}(x, a, y) \quad (y \in S) \tag{2.1}$$

and

$$R(\tilde{s}, \tilde{a})(u) := \sup_{(x,a) \in S \times A} \tilde{s}(x) \wedge \tilde{a}(a) \wedge \tilde{r}(x, a, u) \quad (0 \leq u \leq M). \tag{2.2}$$

Note that, by Lemma 1.1, it holds that $Q(\tilde{s}, \tilde{a})(\cdot) \in \mathcal{F}(S)$ and $R(\tilde{s}, \tilde{a})(\cdot) \in \mathcal{F}([0, M])$ for all $\tilde{s} \in \mathcal{F}(S), \tilde{a} \in \mathcal{F}(A)$.

Firstly we will define a policy based on the fuzzy state and fuzzy action as follows. Let $\Pi := \{\pi | \pi : \mathcal{F}(S) \mapsto \mathcal{F}(A)\}$ be the set of all maps from $\mathcal{F}(S)$ to $\mathcal{F}(A)$. Any element $\pi \in \Pi$ is called a strategy. A policy, $\breve{\pi} = (\pi_1, \pi_2, \pi_3, \cdots)$, is a sequence of strategies such that $\pi_t \in \Pi$ for each t. Especially, the policy $(\pi, \pi, \pi, \cdots)$ is a stationary policy and is denoted by π^∞.

A fuzzy strategy $\pi \in \Pi$ is called admissible if the α-cut $\pi(\tilde{s})_\alpha$ of π depends only on the scalar α and the sets $\tilde{s}_\alpha$, that is, it would be written as

$$\pi(\tilde{s})_\alpha = \pi(\alpha, \tilde{s}_\alpha) \quad \text{for } \tilde{s} \in \mathcal{F}(S). \tag{2.3}$$

For the admissible fuzzy strategy $\pi = \{\pi(\cdot, \cdot)\}$, if $\pi(\alpha, D)$ is continuous in $(\alpha, D) \in [0, 1] \times \mathcal{C}(S)$, π is called continuous. We denote by Π_A and Π_C, respectively, the collections of all admissible and continuous fuzzy strategies. A policy $\breve{\pi} = (\pi_1, \pi_2, \cdots)$ is called admissible (continuous resp.) if $\pi_t \in \Pi_A$ (Π_C) for all $t \geq 0$. To specify a performance index (or an objective function) and describe an optimal decision problem, let us consider the convergence of a sequence of fuzzy numbers and a partial order.

Definition 2.1 (c.f. [5, 10]). For $\tilde{u}_t, \tilde{u} \in \mathcal{F}(E)$,

$$\lim_{t\to\infty} \tilde{u}_t = \tilde{u}$$

iff $\lim_{t\to\infty} \sup_{\alpha\in[0,1]} \rho(\tilde{u}_{t,\alpha}, \tilde{u}_\alpha) = 0$, where $\tilde{u}_{t,\alpha}$ and $\tilde{u}_\alpha$ are respectively the α-cut of $\tilde{u}_t$ and $\tilde{u}$.

For any closed interval $D \in \mathcal{C}([0, M])$, we put $D = [\underline{D}, \overline{D}]$, where $\underline{D}$ and $\overline{D}$ are the left and right end points of D respectively.

The partial order $\preceq$ on $\mathcal{C}([0, M])$ is defined as follows : For any $D_1, D_2 \in \mathcal{C}([0, M])$. $D_1 \succeq D_2$ means that $\overline{D_1} \geq \overline{D_2}$ and $\underline{D_1} \geq \underline{D_2}$. Then, $(\mathcal{C}([0, M]), \succeq)$ becomes a complete lattice (see [2]) and the following lemma holds obviously.

Lemma 2.1. *For any sequence* $\{D_n\}_{n=1}^\infty \subset \mathcal{C}([0, M])$, *it holds that*

(i) $\sup_{n\geq1} D_n = [\sup_{n\geq1} \underline{D_n}, \sup_{n\geq1} \overline{D_n}]$, *and*

(ii) *if* $\sum_{n\geq1}^\infty D_n$ *converges,* $\sum_{n\geq1}^\infty D_n = [\sum_{n\geq1}^\infty \underline{D_n}, \sum_{n\geq1}^\infty \overline{D_n}]$.

Using the order on $\mathcal{C}([0, M])$, let us define the partial order $\succeq$ on $\mathcal{F}([0, M])$ which is called a max fuzzy order.

Definition 2.2. For any $\tilde{n}, \tilde{m} \in \mathcal{F}([0, M])$, $\tilde{n} \succeq \tilde{m}$ iff $\tilde{n}_\alpha \succeq \tilde{m}_\alpha$ for all $\alpha \in [0, 1]$.

For any admissible policy $\breve{\pi} = (\pi_1, \pi_2, \cdots)$ and an initial fuzzy state $\tilde{s} \in \mathcal{F}(S)$, we can define a sequence of fuzzy rewards on $[0, M]$,

$$\{R(\tilde{s}_t, \pi_t(\tilde{s}_t))\}_{t=1}^\infty,$$

where

$$\tilde{s}_1 = \tilde{s} \quad \text{and} \quad \tilde{s}_{t+1} = Q(\tilde{s}_t, \pi_t(\tilde{s}_t)) \text{ for } t \geq 1. \tag{2.4}$$

Here we are concerned with two performance criteria. The first one is the total discounted fuzzy reward with a discount factor β $(0 < \beta < 1)$;

$$\psi_\beta(\breve{\pi}, \tilde{s}) := \sum_{t=1}^\infty \beta^{t-1} R(\tilde{s}_t, \pi_t(\tilde{s}_t)) \in \mathcal{F}([0, M/(1-\beta)]) \tag{2.5}$$

for $\tilde{s} \in \mathcal{F}(S)$ and $\check{\pi} = (\pi_1, \pi_2, \cdots)$.

The problem in the discounted case is to maximize $\psi_\beta(\check{\pi}, \tilde{s})$ over all admissible policy $\check{\pi}$ with respect to the order $\succeq$ on $\mathcal{F}([0, M])$, which has been investigated in [6]. The second performance criteria is the long-run average fuzzy reward per unit time, which is formally defined by

$$\Psi(\tilde{s}, \check{\pi}) := \lim_{T \to \infty} \frac{R_T}{T}, \tag{2.6}$$

where

$$R_T := \sum_{t=1}^{T} R(\tilde{s}_t, \pi_t(\tilde{s}_t)) \quad (T \geq 1). \tag{2.7}$$

Our problem in this paper is to show the convergence property of $\Psi(\tilde{s}, \check{\pi})$ and maximize $\Psi(\tilde{s}, \check{\pi})$ over some class of continuous policies $\check{\pi}$ with respect to the order $\succeq$ on $\mathcal{F}([0, M])$, which is given in Sections 4 and 5.

3 Assumptions and preliminary results

In this section, we introduce some assumptions for the fuzzy relation $\tilde{q}$ and fuzzy reward $\tilde{r}$. And several preliminary results are given for the discounted case, which guarantee the validity of the "vanishing discount factor" approach. In order to discuss the structure of the fuzzy state transition and fuzzy reward, let us introduce some notations.

A map $Q_\alpha : \mathcal{C}(S) \times \mathcal{C}(A) \mapsto \mathcal{C}(S)$ $(\alpha \in [0, 1])$ is defined by

$$Q_\alpha(D \times B) := \begin{cases} \{y \in S \mid \tilde{q}(x, a, y) \geq \alpha \text{ for some } (x, a) \in D \times B\}, & \alpha > 0, \\ cl\{y \in S \mid \tilde{q}(x, a, y) > 0 \text{ for some } (x, a) \in D \times B\}, & \alpha = 0, \end{cases}$$

and a map $R_\alpha : \mathcal{C}(S) \times \mathcal{C}(A) \mapsto \mathcal{C}([0, M])$ $(\alpha \in [0, 1])$ by

$$R_\alpha(D \times B) := \begin{cases} \{u \in R_+ \mid \tilde{r}(x, a, u) \geq \alpha \text{ for some } (x, a) \in D \times B\}, & \alpha > 0, \\ cl\{u \in R_+ \mid \tilde{r}(x, a, u) > 0 \text{ for some } (x, a) \in D \times B\}, & \alpha = 0. \end{cases}$$

Since $R_\alpha(D \times B)$ is a closed interval for each $\alpha \in [0, 1]$, we can write it as

$$R_\alpha(D \times B) := [\underline{R}_\alpha(D \times B), \overline{R}_\alpha(D \times B)].$$

In this paper, we need the following two assumptions.

Assumption A (Ergodicity or contraction). There exits γ_1 $(0 < \gamma_1 < 1)$ such that

$$\rho(Q_\alpha(D \times B), Q_\alpha(D' \times B)) \le \gamma_1 \rho(D, D') \quad \text{for all } B \in \mathcal{C}(A).$$

Assumption B (Lipschitz condition). There exists a constant C such that

$$|\underline{R}_\alpha(D \times B) - \underline{R}_\alpha(D' \times B)| \vee |\overline{R}_\alpha(D \times B) - \overline{R}_\alpha(D' \times B)| \le C\rho(D, D')$$

for any $D, D' \in \mathcal{C}(S)$ and $B \in \mathcal{C}(A)$.

We note that Assumption A has been given in [5, 17], under which several limit theorems for the sequence of fuzzy states have bee obtained.

We now derive the optimality equation for the discounted case. Let

$$V := \{v : \mathcal{C}(S) \mapsto \mathcal{C}([0, M])\}.$$

Define a metric d_V on V by

$$d_V(v, w) := \sup_{D \in \mathcal{C}(S)} \delta(v(D), w(D)) \quad \text{for } v, w \in V.$$

Then, (V, d_V) becomes a complete metric space. Define a map $U_\alpha^\beta : V \mapsto V$ $(\alpha \ge 0)$ by

$$U_\alpha^\beta v(D) := \sup_{B \in \mathcal{C}(A)} \{R_\alpha(D \times B) + \beta v(Q_\alpha(D \times B))\} \tag{3.1}$$

for $v \in V$ and $D \in \mathcal{C}(S)$, where β is a discount factor and $0 < \beta < 1$.

If we write $v(D)$ and $U_\alpha^\beta v(D)$ respectively by $v(D) = [\underline{v}(D), \overline{v}(D)]$ and $U_\alpha^\beta v(D) = [\underline{U}_\alpha^\beta v(D), \overline{U}_\alpha^\beta v(D)]$, (3.1) becomes, from Lemma 2.1,

$$\underline{U}_\alpha^\beta v(D) = \sup_{B \in \mathcal{C}(A)} \{\underline{R}_\alpha(D \times B) + \beta \underline{v}(Q_\alpha(D \times B))\}, \tag{3.2}$$

$$\overline{U}_\alpha^\beta v(D) = \sup_{B \in \mathcal{C}(A)} \{\overline{R}_\alpha(D \times B) + \beta \overline{v}(Q_\alpha(D \times B))\}. \tag{3.3}$$

In [6], it is shown that the operator U_α^β is a contraction with modulus β. Thus, there exists a unique map $v_{\alpha,\beta} \in V$ such that

$$v_{\alpha,\beta} = U_\alpha^\beta v_{\alpha,\beta}. \tag{3.4}$$

Let $v_{\alpha,\beta}(D) := [\underline{v}_{\alpha,\beta}(D), \overline{v}_{\alpha,\beta}(D)]$ for all $D \in \mathcal{C}(S)$. The property of $v_{\alpha,\beta}$ is given in the following lemma.

Lemma 3.1. *Suppose that Assumptions A and B hold. Then, we have*

$$|\underline{v}_{\alpha,\beta}(D) - \underline{v}_{\alpha,\beta}(D')| \vee |\overline{v}_{\alpha,\beta}(D) - \overline{v}_{\alpha,\beta}(D')| \leq \frac{C}{1-\beta\gamma_1}\rho(D, D') \tag{3.5}$$

for all $D, D' \in \mathcal{C}(S)$.

Proof. Putting $\underline{v}^0_{\alpha,\beta} \equiv \{0\} \in V$, we define the iterates $\underline{v}^t_{\alpha,\beta}$ by

$$\underline{v}^{t+1}_{\alpha,\beta} = \underline{U}^{\beta}_{\alpha}\underline{v}^t_{\alpha,\beta}, \quad (t \geq 0) \tag{3.6}$$

Then, by the contractive property of $\underline{U}^{\beta}_{\alpha}$, $\underline{v}^t_{\alpha,\beta}(D) \to \underline{v}_{\alpha,\beta}(D)$ uniformly for $D \in \mathcal{C}(S)$ as $n \to \infty$. Now, we show by induction on t that

$$|\underline{v}^t_{\alpha,\beta}(D) - \underline{v}^t_{\alpha,\beta}(D')| \leq C\sum_{l=0}^{t-1}(\beta\gamma_1)^l\rho(D, D') \tag{3.7}$$

for all $D, D' \in \mathcal{C}(S)$ and $t \geq 1$. It holds, from Assumption B, that

$$|\underline{v}^1_{\alpha,\beta}(D) - \underline{v}^1_{\alpha,\beta}(D')| \leq \sup_{B\in\mathcal{C}(A)} |\underline{R}_{\alpha}(D \times B) - \underline{R}_{\alpha}(D' \times B)| \leq C\rho(D, D'),$$

which implies that (3.7) holds for $t = 1$. Suppose that (3.7) holds for t. Then, we have

$$\begin{aligned}
&|\underline{v}^{t+1}_{\alpha,\beta}(D) - \underline{v}^{t+1}_{\alpha,\beta}(D')| \\
&\leq \sup_{B\in\mathcal{C}(A)} |\underline{R}_{\alpha}(D \times B) - \underline{R}_{\alpha}(D' \times B)| \\
&+ \beta \sup_{B\in\mathcal{C}(A)} |\underline{v}^t_{\alpha,\beta}(Q_{\alpha}(D \times B)) - \underline{v}^t_{\alpha,\beta}(Q_{\alpha}(D' \times B))| \\
&\leq C\rho(D, D') + \beta\sum_{l=0}^{t-1}(\gamma_1\beta)^l \sup_{B\in\mathcal{C}(A)} \rho(Q_{\alpha}(D \times B), Q_{\alpha}(D' \times B)) \\
&\qquad \text{, from Assumption B and the hypothesis of induction} \\
&\leq \sum_{l=0}^{t}(\gamma_1\beta)^l\rho(D, D') \\
&\qquad \text{, from Assumption A.}
\end{aligned}$$

This shows that (3.6) holds for $t+1$. Letting $t \to \infty$ in (3.7), we get

$$|\underline{v}_{\alpha,\beta}(D) - \underline{v}_{\alpha,\beta}(D')| \le \frac{C}{1-\beta\gamma_1}\rho(D,D'). \tag{3.8}$$

Similarly as the case of $\underline{v}_{\alpha,\beta}$, we obtain that

$$|\overline{v}_{\alpha,\beta}(D) - \overline{v}_{\alpha,\beta}(D')| \le \frac{C}{1-\beta\gamma_1}\rho(D,D'). \tag{3.9}$$

These inequalities (3.8) and (3.9) implies (3.5). $\square$

4 Characterization of the average fuzzy reward

This section concerns the convergence and characterization of the average fuzzy reward $\Psi(\tilde{s},\pi^\infty)$, which formally given in (2.6).

Let $\pi \in \Pi_C$. For simplicity, we put

$$R_T(\tilde{s},\pi^\infty) := \sum_{t=1}^{T} R(\tilde{s}_t, \pi(\tilde{s}_t)), \tag{4.1}$$

where

$$\tilde{s}_1 := \tilde{s} \quad \text{and} \quad \tilde{s}_{t+1} = Q(\tilde{s}_t, \pi(\tilde{s}_t)) \ (t \ge 1)$$

By using $\tilde{q}$ and $\tilde{r}$, we define maps $Q_\alpha^\pi : \mathcal{C}(S) \mapsto \mathcal{C}(S)$ and $R_\alpha^\pi : \mathcal{C}(S) \mapsto \mathcal{C}([0,M])$ $(\pi \in \Pi_A, \alpha \in [0,1])$ by

$$\begin{array}{ll} Q_\alpha^\pi(D) & := Q_\alpha(D \times \pi(\alpha, D)) \\ R_\alpha^\pi(D) & := R_\alpha(D \times \pi(\alpha, D)) \end{array}$$

for $D \in \mathcal{C}(S)$. For any admissible fuzzy policy $\pi \in \Pi_A$, $Q_{t,\alpha}^\pi$ $(t \ge 1)$ is defined inductively by using the composition of maps as follows :

$$Q_{1,\alpha}^\pi(D) := Q_\alpha^\pi(D)$$

and

$$Q_{t+1,\alpha}^\pi(D) := Q_{t,\alpha}^\pi Q_\alpha^\pi(D)$$

for $t \geq 1$ and $D \in \mathcal{C}(S)$.

The following lemma can be proved analogously to those of [7, Lemma 2.1].

Lemma 4.1. *Let $\pi \in \Pi_C$. Then :*

(i) $\tilde{s}_{t+1,\alpha} = Q^{\pi}_{t,\alpha}(\tilde{s}_\alpha)$ *for $t \geq 1$,*

(ii) $R_T(\tilde{s}, \pi^\infty) \in \mathcal{F}([0, TM])$ *for $T \geq 1$,*

(iii) $(R_T(\tilde{s}, \pi^\infty))_\alpha = \sum_{t=1}^{T} R_\alpha(\tilde{s}_{t,\alpha}, \pi(\alpha, \tilde{s}_{t,\alpha}))$ *for $T \geq 1$.*

In order to insure the ergodicity of the process and the uniform continuity of the fuzzy reward, we introduce the condition $L(\pi)$ depending on $\pi \in \Pi_C$.

For any continuous strategy $\pi \in \Pi_C$, we shall say that $L(\pi)$ holds if there exist constant $\gamma\,(0 < \gamma < 1)$, $C > 0$ and a positive integer t_0 satisfying the following (i) and (ii) :

(i) $\rho(Q^{\pi}_{t_0,\alpha}(D), Q^{\pi}_{t_0,\alpha}(D')) \leq \gamma\rho(D, D')$,

(ii) $\delta(R^{\pi}_\alpha(D), R^{\pi}_\alpha(D')) \leq C\rho(D, D')$.

Note that the conditions (i) and (ii) are corresponding to Assumption A and B in Section 3 respectively.

Let us denote by Π^*_{SC} the set of all continuous strategies $\pi \in \Pi_C$ satisfying the above condition $L(\pi)$. The convergence of $\Psi(\tilde{s}, \pi^\infty)$ is given in the following theorem. The proof is analogous to those of [7].

Theorem 4.1. *For any $\pi \in \Pi^*_{SC}$, $\Psi(\tilde{s}, \pi^\infty)$ in (2.6) converges and satisfies the following :*

$$\Psi(\tilde{s}, \pi^\infty) = R(\tilde{p}^\pi, \pi(\tilde{p}^\pi)),$$

where $\tilde{p}^\pi \in \mathcal{F}(S)$ is a limiting fuzzy state satisfying that

(i) $\lim_{t\to\infty} \tilde{s}_t = \tilde{p}^\pi$, *and* $Q^{\pi}_\alpha(\tilde{p}^\pi_\alpha) = \tilde{p}^\pi_\alpha$ *for all $\alpha \in [0, 1]$,*

(ii) $\tilde{p}^\pi$ *is independent of the initial fuzzy state $\tilde{s}$, and*

(iii) $\rho(\tilde{s}_{t,\alpha}, \tilde{p}^\pi_\alpha) \leq C^*\gamma^t$ *with γ in Assumption $L(\pi)$ and some $C^* > 0$.*

Recently, Yoshida [19] has given the notion of α-recurrent set for the fuzzy relation and shown that the α-cut of the limiting fuzzy set $\tilde{p}^\pi$ in the above lemma is characterized as the maximum α-recurrent set. Theorem 4.1 says that for any $\pi \in \Pi^*_{SC}$, $\Psi(\tilde{s}, \pi^\infty)$ is independent of $\tilde{s}$, so we write it by $\Psi(\pi^\infty)$.

For simplicity, let, for each $\pi \in \Pi^*_{SC}$ and $D \in \mathcal{C}(S)$,

$$R^{\pi^\infty}_{T,\alpha}(D) = \sum_{t=1}^{T} R^\pi_\alpha(Q^\pi_{t,\alpha}(D)).$$

Note from Lemma 4.1 that $R_T(\tilde{s}, \pi^\infty)_\alpha = R^{\pi^\infty}_{T,\alpha}(\tilde{s}_\alpha)$ for all $T \geq 1$ and $\alpha \in [0,1]$. Let $R^{\pi^\infty}_{T,\alpha}(D) := [\underline{R}^{\pi^\infty}_{T,\alpha}(D), \overline{R}^{\pi^\infty}_{T,\alpha}(D)]$. Then, by Lemma 2.1, we have

$$\underline{R}^{\pi^\infty}_{T,\alpha}(D) = \sum_{t=1}^{T} \underline{R}^\pi_\alpha(Q^\pi_{t,\alpha}(D)) \tag{4.2}$$

and

$$\overline{R}^{\pi^\infty}_{T,\alpha}(D) = \sum_{t=1}^{T} \overline{R}^\pi_\alpha(Q^\pi_{t,\alpha}(D)), \tag{4.3}$$

where

$$R^\pi_\alpha(D') := [\underline{R}^\pi_\alpha(D'), \overline{R}^\pi_\alpha(D')]$$

for all $D' \in \mathcal{C}(S)$. By Theorem 4.1 and Assumption B, we observe that $R^\pi_\alpha(\tilde{s}_{t,\alpha}) \to R^\pi_\alpha(\tilde{p}^\pi_\alpha)$ exponentially first as $t \to \infty$. Thus, by (4.2) and (4.3),

$$\underline{h}^\pi_\alpha(D) := \lim_{T\to\infty} (\underline{R}^{\pi^\infty}_{T,\alpha}(D) - T \times \underline{R}^\pi_\alpha(\tilde{p}^\pi_\alpha)) \tag{4.4}$$

and

$$\overline{h}^\pi_\alpha(D) := \lim_{T\to\infty} (\overline{R}^{\pi^\infty}_{T,\alpha}(D) - T \times \overline{R}^\pi_\alpha(\tilde{p}^\pi_\alpha)) \tag{4.5}$$

converge for all $D \in \mathcal{C}(S)$. The function $\underline{h}_\alpha$ ($\overline{h}_\alpha$ resp.) is called lower (upper) relative value function, whose basic ideas are appearing in the theory of Markov processes (c.f. [12]).

Let us denote the α-cut of the discounted fuzzy reward :

$$\psi_\beta(\pi^\infty, \tilde{s})_\alpha := [\underline{\psi}_\beta(\pi^\infty, \tilde{s})_\alpha, \overline{\psi}_\beta(\pi^\infty, \tilde{s})_\alpha], \quad \alpha \in [0,1]. \tag{4.6}$$

Then, for any $\pi \in \Pi_{SC}^*$, the extremal points

$$\Psi(\pi^\infty)_\alpha := [\underline{\Psi}_\alpha(\pi^\infty), \overline{\Psi}_\alpha(\pi^\infty)]$$

are characterized in the following theorem, whose description is popular in the theory of Markov decision processes (cf. [1, 3, 12]).

Theorem 4.2. *For any $\pi \in \Pi_{SC}^*$, we have*

$$\underline{\psi}_\beta(\pi^\infty, \tilde{s})_\alpha = \underline{\Psi}_\alpha(\pi^\infty)/(1-\beta) + \underline{h}_\alpha^\pi(\tilde{s}_\alpha) + \underline{\varepsilon}(\beta, \alpha), \tag{4.7}$$

$$\overline{\psi}_\beta(\pi^\infty, \tilde{s})_\alpha = \overline{\Psi}_\alpha(\pi^\infty)/(1-\beta) + \overline{h}_\alpha^\pi(\tilde{s}_\alpha) + \overline{\varepsilon}(\beta, \alpha), \tag{4.8}$$

where

$$|\underline{\varepsilon}(\beta, \alpha)| \vee |\overline{\varepsilon}(\beta, \alpha)| \to 0$$

uniformly for $\alpha \in [0, 1]$ as $\beta \to 1$.

Proof. From Lemma 2.1, we have

$$\begin{aligned}
\underline{\psi}_\beta(\pi^\infty, \tilde{s})_\alpha &= \textstyle\sum_{t=1}^\infty \beta^{t-1} \underline{R}_\alpha^\pi(Q_{t,\alpha}^\pi(\tilde{s}_\alpha)) \\
&= \textstyle\sum_{t=1}^\infty \beta^{t-1} \underline{\Psi}_\alpha(\pi^\infty) + \sum_{t=1}^\infty \beta^{t-1}(\underline{R}_\alpha^\pi(Q_{t,\alpha}^\pi(\tilde{s}_\alpha)) - \underline{\Psi}_\alpha(\pi^\infty)) \\
&= \underline{\Psi}_\alpha(\pi^\infty)/(1-\beta) + \underline{h}_\alpha^\pi(\tilde{s}_\alpha) + \underline{\varepsilon}(\beta, \alpha),
\end{aligned}$$

where $\underline{\varepsilon}(\beta, \alpha) = \sum_{t=1}^\infty \beta^{t-1}(\underline{R}_\alpha^\pi(Q_{t,\alpha}^\pi(\tilde{s}_\alpha)) - \underline{\Psi}_\alpha(\pi^\infty)) - \underline{h}_\alpha^\pi(\tilde{s}_\alpha)$.
As $\underline{\Psi}_\alpha(\pi^\infty) = \underline{R}_\alpha^\pi(\tilde{p}_\alpha)$, by (4.4) and Abel theorem, it holds that $\underline{\varepsilon}(\beta, \alpha) \to 0$ uniformly for $\alpha \in [0, 1]$ as $\beta \uparrow 1$, which proves (4.7). By arguments similar to the above, we can prove (4.8). This completes the proof. □

The proof of the following theorem is analogous to that of [7, Theorem 3.1].

Theorem 4.3. *For any $\pi \in \Pi_{SC}^*$, let $\underline{h}_\alpha^\pi$ and $\overline{h}_\alpha^\pi$ be defined as (4.4) and (4.5). Then, the following equations hold :*

$$\underline{h}_\alpha^\pi(D) + \underline{\Psi}_\alpha(\pi^\infty) = \underline{R}_\alpha^\pi(D) + \underline{h}_\alpha^\pi(Q_\alpha^\pi(D)) \tag{4.9}$$

and

$$\overline{h}_\alpha^\pi(D) + \overline{\Psi}_\alpha(\pi^\infty) = \overline{R}_\alpha^\pi(D) + \overline{h}_\alpha^\pi(Q_\alpha^\pi(D)) \tag{4.10}$$

for all $D \in \mathcal{C}(S)$.

The following theorem is useful in policy improvement.

Theorem 4.4. *For any $\pi \in \Pi^*_{SC}$, let $\underline{h}^\pi_\alpha$ and $\overline{h}^\pi_\alpha$ be defined as in (4.4) and (4.5). Let $\pi' \in \Pi^*_{SC}$ be such that*

$$\underline{h}^\pi_\alpha(D) + \underline{\Psi}_\alpha(\pi^\infty) \le \underline{R}^{\pi'}_\alpha(D) + \underline{h}^\pi_\alpha(Q^{\pi'}_\alpha(D)) \tag{4.11}$$

and

$$\overline{h}^\pi_\alpha(D) + \overline{\Psi}_\alpha(\pi^\infty) \le \overline{R}^{\pi'}_\alpha(D) + \overline{h}^\pi_\alpha(Q^{\pi'}_\alpha(D)) \tag{4.12}$$

for all $D \in \mathcal{C}(S)$ and $\alpha \in [0,1]$. Then $\Psi(\pi^\infty) \preceq \Psi(\pi'^\infty)$.

Proof. Let $\pi' \in \Pi^*_{SC}$ be such that (4.11) and (4.12) hold. Then, we have

$$\underline{h}^\pi_\alpha(Q^{\pi'}_{t,\alpha}(\tilde{s}_\alpha)) + \underline{\Psi}_\alpha(\pi^\infty) \le \underline{R}^{\pi'}_\alpha(Q^{\pi'}_{t,\alpha}(\tilde{s}_\alpha)) + \underline{h}^\pi_\alpha(Q^{\pi'}_{t+1,\alpha}(\tilde{s}_\alpha)) \quad \text{for all } t \ge 1.$$

Summing up the above inequality for $1 \le t \le T$, we get

$$\underline{h}^\pi_\alpha(\tilde{s}_\alpha) + T \times \underline{\Psi}_\alpha(\pi^\infty) \le \underline{R}^{\pi'^\infty}_{T,\alpha}(\tilde{s}_\alpha) + \underline{h}^\pi_\alpha(Q^{\pi'}_{T+1,\alpha}(\tilde{s}_\alpha)) \quad \text{for all } t \ge 1.$$

Thus, since

$$\underline{\Psi}_\alpha(\pi'^\infty) = \lim_{T\to\infty} \frac{1}{T} \underline{R}^{\pi'^\infty}_{T,\alpha}(\tilde{s}_\alpha),$$

it holds that

$$\underline{\Psi}_\alpha(\pi^\infty) \le \underline{\Psi}_\alpha(\pi'^\infty).$$

Similarly

$$\overline{\Psi}_\alpha(\pi^\infty) \le \overline{\Psi}_\alpha(\pi'^\infty).$$

Therefore,

$$[\underline{\Psi}_\alpha(\pi^\infty), \overline{\Psi}_\alpha(\pi^\infty)] \preceq [\underline{\Psi}_\alpha(\pi'^\infty), \overline{\Psi}_\alpha(\pi'^\infty)],$$

which shows

$$\Psi(\pi^\infty) \preceq \Psi(\pi'^\infty).$$

Thus we get this theorem. $\square$

Applying the method of the proof in Theorem 4.4, we can easily prove the following corollary.

Corollary 4.1. *For any $\pi \in \Pi^*_{SC}$, let $\underline{h}^\pi_\alpha$ and $\overline{h}^\pi_\alpha$ be defined as in (4.4) and (4.5). Suppose that the following inequalities hold :*

$$\underline{h}^\pi_\alpha(D) + \underline{\Psi}_\alpha(\pi^\infty) \ge \underline{R}^{\pi'}_\alpha(D) + \underline{h}^\pi_\alpha(Q^{\pi'}_\alpha(D))$$

and

$$\overline{h}^{\pi}_{\alpha}(D) + \overline{\Psi}_{\alpha}(\pi^{\infty}) \geq \overline{R}^{\pi'}_{\alpha}(D) + \overline{h}^{\pi}_{\alpha}(Q^{\pi'}_{\alpha}(D))$$

for all $\pi' \in \Pi^*_{SC}$, $D \in \mathcal{C}(S)$ *and* $\alpha \in [0,1]$. *Then* π^{∞} *is absolutely optimal in* Π^*_{SC}, *i.e.,*

$$\Psi(\pi^{\infty}) \succeq \Psi(\pi'^{\infty}) \quad \textit{for all } \pi' \in \Pi^*_{SC}.$$

5 The optimality equation

In this section, we derive the optimality equation and consider its validity for optimization. The proof is done by the "vanishing discounted factor" method, using Ascoli-Arzela theorem (c.f. [4, 8, 13]).

Theorem 5.1. *Suppose that Assumptions A and B in Section 3 hold. Then, for any* $\alpha \in [0,1]$, *there exist constants* $\underline{\Psi}_{\alpha}, \overline{\Psi}_{\alpha}$ *and* $\underline{v}_{\alpha}, \overline{v}_{\alpha} : \mathcal{C}(S) \mapsto \mathcal{C}([0,M])$ *such that*

(i) $[\underline{\Psi}_{\alpha'}, \overline{\Psi}_{\alpha'}] \supset [\underline{\Psi}_{\alpha}, \overline{\Psi}_{\alpha}]$ (5.1)

for any α, α' $(\alpha' < \alpha)$ *belonging to some countable subset dense in* $[0,1]$, *and*

(ii) $\underline{v}_{\alpha}(D) + \underline{\Psi}_{\alpha} = \sup_{B \in \mathcal{C}(A)}\{\underline{R}_{\alpha}(D \times B) + \underline{v}_{\alpha}(Q_{\alpha}(D \times B))\}$ (5.2)

(iii) $\overline{v}_{\alpha}(D) + \overline{\Psi}_{\alpha} = \sup_{B \in \mathcal{C}(A)}\{\overline{R}_{\alpha}(D \times B) + \overline{v}_{\alpha}(Q_{\alpha}(D \times B))\}$ (5.3)

for all $D \in \mathcal{C}(S)$ *and* $\alpha \in [0,1]$.

Proof. Select $D_0 \in \mathcal{C}(S)$ arbitrary. For $\underline{v}_{\alpha,\beta}$ in Lemma 3.1, let

$$\underline{g}_{\alpha,\beta}(D) := \underline{v}_{\alpha,\beta}(D) - \underline{v}_{\alpha,\beta}(D_0) \quad \text{for } D \in \mathcal{C}(S).$$

Then, noting (3.4), it holds that

$$\begin{aligned} & \underline{g}_{\alpha,\beta}(D) + (1-\beta)\underline{v}_{\alpha,\beta}(D_0) \\ = \; & \sup_{B \in \mathcal{C}(A)}\{\underline{R}_{\alpha}(D \times B) + \beta \underline{g}_{\alpha,\beta}(Q_{\alpha}(D \times B))\}. \end{aligned} \tag{5.4}$$

Since $0 \leq (1-\beta)\underline{v}_{\alpha,\beta}(D_0) \leq M$ for all $\beta \in (0,1)$, there exists a sequence $\{\beta_n\}_{n=0}^{\infty}$ such that $\beta_n \uparrow 1$ and $(1-\beta_n)\underline{v}_{\alpha,\beta_n}(D_0) \to \underline{\Psi}_{\alpha}$. By Lemma 3.1,

$\underline{g}_{\alpha,\beta_n}(\cdot)$, $n \geq 1$, are equi-continuous and uniformly bounded, so that, applying the Ascoli-Arzela theorem ([13]), $\underline{g}_{\alpha,\beta_n}(\cdot) \to \underline{v}_\alpha(\cdot)$ uniformly along a subsequence (also called $\{\beta_n\}$ by abuse of notation). Thus, letting $n \to \infty$ in (5.4) for $\beta = \beta_n$, we get (5.2). Also, putting $\overline{g}_{\alpha,\beta}(D) := \overline{v}_{\alpha,\beta}(D) - \overline{v}_{\alpha,\beta}(D_0)$, the proof of (5.3) is given as same as the above. Note that we can assume, without loss of generality, that $(1-\beta_n)\underline{v}_{\alpha,\beta_n}(D_0) \to \underline{\Psi}_\alpha$ and $(1-\beta_n)\overline{v}_{\alpha,\beta_n}(D_0) \to \overline{\Psi}_\alpha$ for any rational number α in $[0,1]$. Since

$$[\underline{v}_{\alpha',\beta}(D_0), \overline{v}_{\alpha',\beta}(D_0)] \supset [\underline{v}_{\alpha,\beta}(D_0), \overline{v}_{\alpha,\beta}(D_0)]$$

for all α, α' $(\alpha' < \alpha)$ and $\beta \in [0,1]$ (c.f. [6]), (i) holds obviously. This completes the proof. □

It will be shown in Theorem 5.2 below that $[\underline{\Psi}_\alpha, \overline{\Psi}_\alpha]$ given in Theorem 5.1 is corresponding to the α-cut of the maximum average fuzzy reward, so that (5.2) and (5.3) are interpreted as the optimality equations for our average fuzzy decision model. For this purpose, we need the following lemmas.

Lemma 5.1. *For any $\tilde{n}, \tilde{m} \in \mathcal{F}([0,M])$, if $\tilde{n}_\alpha \succeq (=) \tilde{m}_\alpha$ on some subset F dense in $[0,1]$, then $\tilde{n} \succeq (=) \tilde{m}$.*

Proof. From the denseness of F, we observe that for any $\alpha \in [0,1]$

$$\tilde{n}_\alpha = \lim_{\alpha' \uparrow \alpha \text{ with } \alpha' \in F} \tilde{n}_{\alpha'} \succeq (=) \lim_{\alpha' \uparrow \alpha \text{ with } \alpha' \in F} \tilde{m}_{\alpha'} = \tilde{m}_\alpha,$$

which implies $\tilde{n} \succeq (=) \tilde{m}$. □

Lemma 5.2 (c.f. [5, 11]). *Suppose that a family of subsets $\{D_\alpha, \alpha \in [0,1]\}$ satisfies the following* (i) *and* (ii) *:*

(i) *$D_{\alpha'} \subset D_\alpha$ for all $\alpha, \alpha' (0 \leq \alpha' \leq \alpha \leq 1)$,*

(ii) *$\lim_{\alpha' \uparrow \alpha} D_{\alpha'} = D_\alpha$ for all $\alpha \in [0,1]$.*

Then, $\tilde{s}(x) := \sup_{\alpha \in [0,1]} \{\alpha \wedge I_{D_\alpha}(x)\}$, $x \in S$, satisfies $\tilde{s} \in \mathcal{F}(S)$ and $\tilde{s}_\alpha = D_\alpha$ for all $\alpha \in [0,1]$.

Let denote by F a countable subset on which (i) in Theorem 5.1 holds. For $\underline{\Psi}_\alpha$ and $\overline{\Psi}_\alpha$ given in Theorem 5.1, let, for any $\alpha \in [0,1]$,

$$\Psi_\alpha^- := \lim_{\alpha' \uparrow \alpha \text{ with } \alpha' \in F} \underline{\Psi}_{\alpha'}, \quad \Psi_0^- := \underline{\Psi}_0,$$

$$\Psi_\alpha^+ := \lim_{\alpha' \uparrow \alpha \,\mathrm{with}\, \alpha' \in F} \overline{\Psi}_{\alpha'}, \quad \Psi_0^+ := \overline{\Psi}_0.$$

Since the conditions (i) and (ii) in Lemma 5.2 hold for the family $\{[\Psi_\alpha^-, \Psi_\alpha^+]; \alpha \in [0,1]\}$, we can construct the fuzzy set $\tilde{\Psi}$ by

$$\tilde{\Psi}(u) := \sup_{\alpha \in [0,1]} \{\alpha \wedge 1_{[\Psi_\alpha^-, \Psi_\alpha^+]}(u)\}, \quad u \in [0, \infty).$$

Now, we can state the following theorem.

Theorem 5.2.

(i) $\tilde{\Psi} \succeq \Psi(\pi^\infty)$ *for any* $\pi \in \Pi_{SC}^*$.

(ii) *If there exists a strategy* $\pi^* \in \Pi_{SC}^*$ *such that*

$$\underline{v}_\alpha(D) + \underline{\Psi}_\alpha = \underline{R}_\alpha(D \times \pi^*(\alpha, D)) + \underline{v}_\alpha(Q_\alpha(D \times \pi^*(\alpha, D)))$$

and

$$\overline{v}_\alpha(D) + \overline{\Psi}_\alpha = \overline{R}_\alpha(D \times \pi^*(\alpha, D)) + \overline{v}_\alpha(Q_\alpha(D \times \pi^*(\alpha, D)))$$

for all $D \in \mathcal{C}(S)$ *and* $\alpha \in [0,1]$, *then* $\pi^{*\infty}$ *is absolutely optimal in the family of continuous stationary policies, i.e.,*

$$\Psi(\pi^{*\infty}) \succeq \Psi(\pi^\infty) \quad \textit{for all } \pi \in \Pi_{SC}^*.$$

Proof. Let $\pi \in \Pi_{SC}^*$. By the optimality equations (5.2) and (5.3), we have

$$\underline{v}_\alpha(D) + \underline{\Psi}_\alpha \geq \underline{R}_\alpha^\pi(D) + \underline{v}_\alpha(Q_\alpha^\pi(D)) \tag{5.5}$$

and

$$\overline{v}_\alpha(D) + \overline{\Psi}_\alpha \geq \overline{R}_\alpha^\pi(D) + \overline{v}_\alpha(Q_\alpha^\pi(D)) \tag{5.6}$$

for all $D \in \mathcal{C}(S)$. By the same discussion as that in the proof of Theorem 4.4, we can prove from (5.5) and (5.6) that

$$[\underline{\Psi}_\alpha, \overline{\Psi}_\alpha] \succeq \Psi(\pi^\infty)_\alpha \quad \text{for all } \alpha \in [0,1]. \tag{5.7}$$

Since $[\Psi_\alpha^-, \Psi_\alpha^+] = [\underline{\Psi}_\alpha, \overline{\Psi}_\alpha]$ for all $\alpha \in F$, we obtain (i) of Theorem 5.2 from (5.7) and Lemma 5.1. For (ii), by the same method as the above, we can prove that $\tilde{\Psi} = \Psi(\pi^{*\infty})$. Thus, together with (i), (ii) follows. This completes the proof. $\square$

Here we give a numerical example to illustrate the theoretical results in this section. Let $S := [0,1]$, $A := [0,1/2]$ and $M := 1$. Take the fuzzy relation and the fuzzy reward by

$$\tilde{q}(x,a,y) = (1 - 3|y - ax|) \vee 0, \quad x, y \in [0,1], \ a \in [0,1/2] \tag{5.8}$$

and

$$\tilde{r}(x,a,z) = (1 - 6|x - z|) \vee 0, \quad x, z \in [0,1]. \tag{5.9}$$

$\tilde{r}$ is independent of $a \in [0,1/2]$. Obviously, they satisfy Assumption A for $t_0 = 1$ and Assumption B. Let $\alpha \in [0,1]$ and $\pi(a) := I_{\{a\}} \in \Pi^*_{SC}$ for $a \in A$. From (5.8) and (5.9), we get

$$Q_\alpha^{\pi(a)}(\{x\}) = [(ax - (1-\alpha)/3) \vee 0, ax + (1-\alpha)/3]$$

for $x \in [0,1]$, $a \in [0,1/2]$. So

$$Q_\alpha^{\pi(a)}([c,b]) = [T_1^a(c), T_2^a(b)] \tag{5.10}$$

for $0 \le c \le b \le 1$, where maps $T_i^a : [0,1] \mapsto [0,1]$ $(i = 1,2; a \in [0,1/2])$ are given by

$$T_1^a(x) := (ax - (1-\alpha)/3) \vee 0; \quad T_2^a(x) := ax + (1-\alpha)/3 \quad \text{for } x \in [0,1].$$

Similarly we observe that

$$R_\alpha^{\pi(a)}([c,b]) = [(c - (1-\alpha)/6) \vee 0, (b + (1-\alpha)/6) \wedge 1] \quad \text{for } 0 \le c \le b \le 1.$$

By (5.10), a unique fixed point $\tilde{p}_\alpha^{\pi(a)}$ of the map $Q_\alpha^{\pi(a)} : \mathcal{C}([0,1]) \mapsto \mathcal{C}([0,1])$ can be found to be

$$\tilde{p}_\alpha^{\pi(a)} = [0, \ (1-\alpha)/3(1-a)].$$

Applying Theorem 4.1, we get

$$\Psi(\tilde{s}, \pi(a)^\infty)_\alpha = [0, \ (1-\alpha)(3-a)/6(1-a)] \quad \text{for } \alpha \in [0,1]. \tag{5.11}$$

In [7], the average fuzzy reward and the relative value functions $\underline{h}_\alpha$ and $\overline{h}_\alpha$ for $\pi(1/2)^\infty$ have been given as follows :

$$\Psi(\tilde{s}, \pi(1/2)^\infty)(x) = \begin{cases} 1 - 6x/5 & 0 \le x \le 5/6 \\ 0 & 5/6 < x \le 1, \end{cases}$$

$$\begin{aligned} \underline{h}_\alpha^{\pi(1/2)}(c) &:= \underline{h}_\alpha^{\pi(1/2)}([c, b]) \\ &= \begin{cases} 2\left(1 - (1/2)^{t^*}\right)(c + 2(1-\alpha)/3) - 5t^*(1-\alpha)/6 & \alpha < 1 \\ 2c & \alpha = 1, \end{cases} \end{aligned}$$

where t^* is the smallest non-negative integer such that

$$(1/2)^{t^*}(c + 2(1-\alpha)/3) - 5(1-\alpha)/6 < 0,$$

and

$$\overline{h}_\alpha^{\pi(1/2)}(b) := \overline{h}_\alpha^{\pi(1/2)}([c, b]) = \begin{cases} 2b - 4(1-\alpha)/3 & \text{if } 0 \le b < (5+\alpha)/6 \\ b + (3\alpha - 1)/2 & \text{if } (5+\alpha)/6 \le b \le 1. \end{cases}$$

The graph of $\underline{h}_\alpha^{\pi(1/2)}$ and $\overline{h}_\alpha^{\pi(1/2)}$ for $\alpha = 1/2$ is given in Figure 1, from which we conjecture the monotonicity of $\underline{h}_\alpha^{\pi(1/2)}$ and $\overline{h}_\alpha^{\pi(1/2)}$.

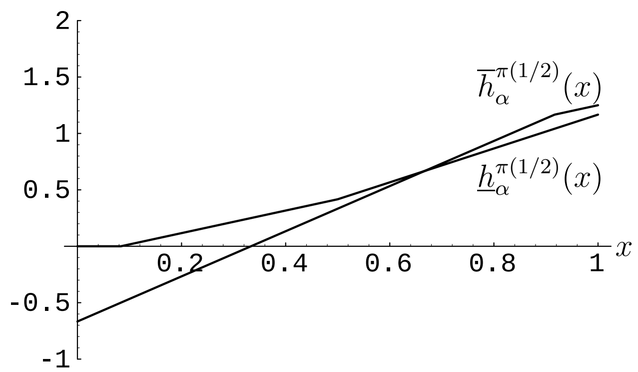

Fig. 1 : The relative value functions $\underline{h}_\alpha^{\pi(1/2)}$ and $\overline{h}_\alpha^{\pi(1/2)}$ for $\alpha = 1/2$.

By the definition of Section 3, we see that

$$Q_\alpha([c, b] \times [c', b']) = [(cc' - (1-\alpha)/3) \vee 0, (bb' + (1+\alpha)/3)]$$

and

$$R_\alpha([c,b]\times[c',b'])=[(c-(1-\alpha)/6)\vee 0,(b+(1-\alpha)/6)\wedge 1]$$

for $[c,b]\in\mathcal{C}(S)$, $[c',b']\in\mathcal{C}(A)$. Noting (5.11) for $a=1/2$, let us put $\underline{\Psi}_\alpha=0$ and $\overline{\Psi}_\alpha=5(1-\alpha)/6$. Then, by the monotonicity of Q_α, R_α, $\underline{h}_\alpha^{\pi(1/2)}$ and $\overline{h}_\alpha^{\pi(1/2)}$, we get the optimality equations described in (5.2) and (5.3) :

$$\begin{aligned}
&\underline{h}_\alpha^{\pi(1/2)}([c,b])+\underline{\Psi}_\alpha\\
&\quad=\sup_{[c',b']\in\mathcal{C}(A)}\{\underline{R}_\alpha([c,b]\times[c',b'])+\underline{h}_\alpha^{\pi(1/2)}(Q_\alpha([c,b]\times[c',b']))\}\\
&\quad=\underline{R}_\alpha([c,b]\times\{1/2\})+\underline{h}_\alpha^{\pi(1/2)}(Q_\alpha([c,b]\times\{1/2\}))
\end{aligned}\tag{5.12}$$

and

$$\begin{aligned}
&\overline{h}_\alpha^{\pi(1/2)}([c,b])+\overline{\Psi}_\alpha\\
&\quad=\sup_{[c',b']\in\mathcal{C}(A)}\{\overline{R}_\alpha([c,b]\times[c',b'])+\overline{h}_\alpha^{\pi(1/2)}(Q_\alpha([c,b]\times[c',b']))\}\\
&\quad=\overline{R}_\alpha([c,b]\times\{1/2\})+\overline{h}_\alpha^{\pi(1/2)}(Q_\alpha([c,b]\times\{1/2\}))
\end{aligned}\tag{5.13}$$

for all $[c,b]\in\mathcal{C}(S)$. From Theorem 5.2, $\pi(1/2)^\infty$ becomes absolutely optimal.

References

[1] Bertsekas,D.P. and Shreve,S.E., *Stochastic Optimal Control : the Discrete Time Case* (Academic Press, New York, 1978).

[2] Birkhoff,G., "Lattice Theory", *Amer. Math. Soc., Coll. Pub.*, **25** (1940).

[3] Blackwell,D., "Discrete dynamic programming", *Ann. Math. Stat.* **33** (1962) 719-726.

[4] Fernándes-Gaucherand,E., "A note on Ross-Taylor theorem", *Applied Math. and Computation* **64** (1994) 207-212.

[5] Kurano,M., Yasuda,M., Nakagami,J. and Yoshida,Y., "A limit theorem in some dynamic fuzzy systems", *Fuzzy Sets and Systems* **51** (1992) 83-88.

[6] Kurano,M., Yasuda,M., Nakagami,J. and Yoshida,Y., "Markov-type fuzzy decision processes with a discounted reward on a closed interval", *European J. of Oper. Res.,* **92** (1994) pp.649–662.

[7] Kurano,M., Yasuda,M., Nakagami,J. and Yoshida,Y., "The time average reward with some dynamic fuzzy systems", to appear in *Comp. Math. Appli.*

[8] Kuratowski,K., *Topology I* (Academic Press, New York, 1966).

[9] Puri,M.L. and Ralescu,D.A., "The concept of normality for fuzzy random variables", *Ann. Prob.* **13** (1985) 1373-1379.

[10] Nanda,S., "On sequences of fuzzy numbers", *Fuzzy Sets and Systems* **33** (1989) 123-126.

[11] Novàk,V., *Fuzzy Sets and Their Applications* (Adam Hilder, Bristol-Boston, 1989).

[12] Ross,S.M., *Applied probability models with optimization applications* (Holden-Say, San Francisco, 1970).

[13] Royden,H.L., *Real Analysis*, 2nd ed. (Macmillan, New York, 1968).

[14] Taylor,H.M., "Markovian sequential replacement processes", *Ann. Math. Statist.* **38** (1965) 1677-2122.

[15] Xu,Chen-wei, "On convexity of fuzzy sets and fuzzy relations", *Information Science* **59** (1992) 92-102.

[16] Yasuda,M., "A linear structure and convexity for relations in dynamic fuzzy systems", *Computers Math. Applic.* **27** (1994) 23-27.

[17] Yoshida,Y., Yasuda,M., Nakagami,J. and Kurano,M., "A potential of fuzzy relations with a linear structure : The contractive case", *Fuzzy Sets and Systems* **60** (1993) 283-294.

[18] Yoshida,Y., Yasuda,M., Nakagami,J. and Kurano,M., "A potential of fuzzy relations with a linear structure : The unbounded case", *Fuzzy Sets and Systems* **66** (1994) 83-95.

[19] Yoshida,Y., "The recurrence of dynamic fuzzy systems", to appear in *Fuzzy Sets and Systems.*

[20] Zadeh,L.A., "Fuzzy sets", *Inform. and Control* **8** (1965) 338-353.