

A fuzzy treatment of uncertain Markov decision processes: Average case

M. KURANO[§], M. YASUDA[†], J. NAKAGAMI[†] and Y. YOSHIDA[‡]

(§)Department of Mathematics, Faculty of Education, Chiba University, Chiba 263-8522 Japan, Tel: 81-43-290-2669, Fax: 81-43-290-2664, E-mail: kurano@math.e.chiba-u.ac.jp,

(†)Department of Mathematics & Informatics, Faculty of Science, Chiba University, Chiba 263-8522 Japan, Tel: 81-43-290-3662(Yasuda), Tel: 81-43-290-2718(Nakagami), Fax: 81-43-290-2733, E-mails: yasuda@math.s.chiba-u.ac.jp, nakagami@math.s.chiba-u.ac.jp,

(‡)Faculty of Economics & Business Administration, Kitakyushu University, Kitakyushu 802-8577 Japan, Tel: 81-93-964-4103, Fax: 81-93-964-4000, E-mail: yoshida@kitakyu-u.ac.jp.

Abstract In this paper, the uncertain transition matrices for inhomogeneous Markov decision processes are described by use of fuzzy sets. Introducing a ν -step contractive property, called a minorization condition, for the average case, we find a Pareto optimal policy maximizing the average expected fuzzy rewards under some partial order. The Pareto optimal policies are characterized by maximal solutions of an optimal equation including efficient set-functions.

Keywords Fuzzy analysis, Markov decision process, Pareto optimal, average reward criterion.

1. Introduction and notation

In modeling of Markov decision processes (MDPs for short), we often encounter the following case:

- (i) The information on the state-transition probabilities include imprecision or ambiguity.
- (ii) The state-transition matrix fluctuates at each step in time and its fluctuation is unknown or unobservable.

In order to deal with uncertain data and flexible requirements, we can use a fuzzy set representation ([12]). In our previous paper [9], we have developed a fuzzy treatment for inhomogeneous MDPs with uncertain transition matrices. The transition matrices are described by the use of fuzzy sets and a Pareto optimal policy for the discounted reward problem has found and characterized by an optimality equation. In this paper, the average case is considered in the same framework as that in our previous work [9]. That is, a Pareto optimal policy maximizing the average expected fuzzy reward(AEFR) under some partial order is found. In order to insure the ergodicity of the process, we introduce a ν -step contractive property for the average case([5,7]), called a minorization condition, which is often used in the study of Markov chains([11]). Using this property, a Pareto optimal periodic stationary policies are characterized as a maximal solution of optimality equation including efficient set functions. Recently, applying Hartfiel's interval method([3,4]) for

Markov chains, Kurano et al. [8] have introduced a decision model, called a controlled Markov set-chain, which is robust for rough approximation of transition matrices in MDPs. Also, under a contractive property for the average case, Hosaka et al. [6] treated the average reward problem for a controlled Markov set-chain. Our fuzzy decision model examined in this paper includes a controlled Markov set-chain as a special case. So, the results obtained here can be thought of as a fuzzy extension of those in [6].

We adopt the notation in [3,4,9]. Let \mathbb{R} , \mathbb{R}^n and $\mathbb{R}^{n \times n}$ be set of real numbers, real n -dimensional column vectors and real $n \times n$ matrices, respectively. Also denote by \mathbb{R}_+ , \mathbb{R}_+^n and $\mathbb{R}_+^{n \times n}$, the subsets of entrywise non-negative elements in \mathbb{R} , \mathbb{R}^n and $\mathbb{R}^{n \times n}$, respectively. We provide \mathbb{R} , \mathbb{R}^n and $\mathbb{R}^{n \times n}$ with the componentwise relation \leq and $<$. For any set X , we will denote a fuzzy set \tilde{a} on X by its membership function $\tilde{a} : X \rightarrow [0, 1]$. Denote by $\mathcal{F}(X)$ the set of all fuzzy sets on X . For the theory of fuzzy sets, refer to Zadeh [14]. The α -cut ($\alpha \in [0, 1]$) of the fuzzy set $\tilde{a} \in \mathcal{F}(X)$ is defined as

$$\begin{aligned}\tilde{a}_\alpha &:= \{x \in X \mid \tilde{a}(x) \geq \alpha\} \quad (\alpha > 0), \\ \tilde{a}_0 &:= \text{cl}\{x \in X \mid \tilde{a}(x) > 0\},\end{aligned}$$

where cl denote the closure of the set. For any interval Y in \mathbb{R} , $\tilde{a} \in \mathcal{F}(Y)$ is called a fuzzy number on Y if \tilde{a} is normal, fuzzy convex, upper semi-continuous and has a compact support. Denote by $\mathcal{F}_c(Y)$ the set of all fuzzy numbers on Y . Let $\mathcal{C}(Y)$ be the set of all closed and bounded intervals in Y . We note that $\tilde{a} \in \mathcal{F}_c(Y)$

means $\tilde{a}_\alpha \in \mathcal{C}(Y)$ for all $\alpha \in [0, 1]$. Let $\mathcal{F}_c(Y)^n$ be the set of all n -dimensional column vectors whose elements are in $\mathcal{F}_c(Y)$, i.e.,

$$\mathcal{F}_c(Y)^n := \{\tilde{\mathbf{u}} = (\tilde{u}_1, \tilde{u}_2, \dots, \tilde{u}_n)' \mid \tilde{u}_i \in \mathcal{F}_c(Y) \ (\forall i)\},$$

where d' denotes the transpose of a vector d .

Let $S := \{1, 2, \dots, n\}$ and $\mathcal{P}(S)$ the set of all probability distributions on S , that is,

$$\mathcal{P}(S) := \{p = (p_1, \dots, p_n) \mid p_j \geq 0 \ (\forall j), \sum_{j=1}^n p_j = 1\}.$$

From any $\tilde{p} = (\tilde{p}_1, \tilde{p}_2, \dots, \tilde{p}_n)' \in \mathcal{F}_c([0, 1])^n$, we will construct the fuzzy set $[\tilde{p}] = [\tilde{p}_1, \tilde{p}_2, \dots, \tilde{p}_n]$ on $\mathcal{P}(S)$ by the following:

$$[\tilde{p}](p) = \min_{1 \leq j \leq n} \{\tilde{p}_j(p_j)\} \quad (1)$$

for any $p = (p_1, p_2, \dots, p_n) \in \mathcal{P}(S)$. The above definition will be extended to the case of stochastic matrices. Let $\mathcal{P}(S/S)$ be the set of all stochastic matrices on S , that is,

$$\mathcal{P}(S/S) := \{Q = (q_{ij}) \mid q_{ij} \geq 0, \sum_{j=1}^n q_{ij} = 1 \ (\forall i)\}.$$

For any $\tilde{q}_i = (\tilde{q}_{i1}, \tilde{q}_{i2}, \dots, \tilde{q}_{in}) \in \mathcal{F}_c([0, 1])^n$ ($1 \leq i \leq n$), we define the fuzzy set $\tilde{Q} = [\tilde{q}_1, \tilde{q}_2, \dots, \tilde{q}_n]'$ on $\mathcal{P}(S/S)$ as follows:

$$\tilde{Q}(Q) := \min_{1 \leq i \leq n} \{\tilde{q}_i(q_i)\}, \quad (2)$$

where $Q = (q_1, q_2, \dots, q_n)' \in \mathcal{P}(S/S)$, $q_i = (q_{i1}, q_{i2}, \dots, q_{in}) \in \mathcal{P}(S)$ and $[\tilde{q}_i]$ is the fuzzy set on $\mathcal{P}(S)$ defined by (1).

In order to describe the structural properties on the fuzzy sets defined in (1) and (2), we need the concept of intervals of matrices. For the detail, refer to [4, 8]. For any nonnegative vector $\underline{q} = (q_1, q_2, \dots, q_n)$ and $\bar{q} = (\bar{q}_1, \bar{q}_2, \dots, \bar{q}_n) \in \mathbb{R}_+^n$ with $\underline{q} \leq \bar{q}$, we define the interval $\langle \underline{q}, \bar{q} \rangle \subset \mathcal{P}(S)$ by

$$\langle \underline{q}, \bar{q} \rangle := \{p = (p_1, p_2, \dots, p_n) \in \mathcal{P}(S) \mid \underline{q} \leq p \leq \bar{q}\}.$$

And, for $\underline{Q} = (q_{ij}), \bar{Q} = (\bar{q}_{ij}) \in \mathbb{R}_+^{n \times n}$ with $\underline{Q} \leq \bar{Q}$,

$$\langle \underline{Q}, \bar{Q} \rangle := \{Q \in \mathcal{P}(S/S) \mid \underline{Q} \leq Q \leq \bar{Q}\}.$$

For any $\tilde{a} \in \mathcal{F}_c([0, 1])$, noting $\tilde{a}_\alpha \in \mathcal{C}([0, 1])$ ($0 \leq \alpha \leq 1$), it will be denoted by $\tilde{a}_\alpha = [\min \tilde{a}_\alpha, \max \tilde{a}_\alpha]$.

Lemma 1 ([4, 9]).

- (i) For any $\underline{Q}, \bar{Q} \in \mathbb{R}_+^{n \times n}$ with $\underline{Q} \leq \bar{Q}$ and $\langle \underline{Q}, \bar{Q} \rangle \neq \emptyset$, $\langle \underline{Q}, \bar{Q} \rangle$ is a polyhedral convex set in the vector space $\mathbb{R}^{n \times n}$.

- (ii) For any $\tilde{q}_i \in \mathcal{F}_c([0, 1])^n$ ($1 \leq i \leq n$), let $\tilde{Q} = [\tilde{q}_1, \tilde{q}_2, \dots, \tilde{q}_n]'$ be a fuzzy set on $\mathcal{P}(S/S)$ defined by (2). Then, the α -cut of \tilde{Q} ($0 \leq \alpha \leq 1$) is a polyhedral convex subset of $\mathcal{P}(S/S)$ and given by

$$\tilde{Q}_\alpha = \langle \underline{Q}_\alpha, \bar{Q}_\alpha \rangle,$$

where

$$\underline{Q}_\alpha = (\min(\tilde{q}_{ij})_\alpha) \text{ and } \bar{Q}_\alpha = (\max(\tilde{q}_{ij})_\alpha).$$

If $\mathbf{u} = ([a_1, b_1], [a_2, b_2], \dots, [a_n, b_n])' \in \mathcal{C}(\mathbb{R}_+)^n$, \mathbf{u} will be denoted by $\mathbf{u} = [a, b]$, where $a = (a_1, a_2, \dots, a_n)'$, $b = (b_1, b_2, \dots, b_n)'$ and $[a, b] = \{x \in \mathbb{R}_+^n \mid a \leq x \leq b\}$. For any $\mathbf{u} \in \mathcal{C}(\mathbb{R}_+)^n$ and $\underline{Q}, \bar{Q} \in \mathbb{R}_+^{n \times n}$ with $\underline{Q} \leq \bar{Q}$ and $\langle \underline{Q}, \bar{Q} \rangle \neq \emptyset$, we define their product by

$$\langle \underline{Q}, \bar{Q} \rangle \mathbf{u} = \{Q\mathbf{u} \mid Q \in \langle \underline{Q}, \bar{Q} \rangle, \mathbf{u} \in \mathbf{u}\}.$$

The following arithmetical notation is used in the sequel. Let $\tilde{Q} = [\tilde{q}_1, \tilde{q}_2, \dots, \tilde{q}_n]'$ be a fuzzy set on $\mathcal{P}(S/S)$ with $\tilde{q}_i \in \mathcal{F}_c([0, 1])^n$ ($1 \leq i \leq n$). Then, for $\tilde{\mathbf{u}} = (\tilde{u}_1, \tilde{u}_2, \dots, \tilde{u}_n)' \in \mathcal{F}_c(\mathbb{R}_+)^n$, $\tilde{Q}\tilde{\mathbf{u}} \in \mathcal{F}(\mathbb{R}_+^n)$ is defined as follows:

$$(\tilde{Q}\tilde{\mathbf{u}})(x) = \max_{x=Q\mathbf{u}, Q \in \mathcal{P}(S/S), \mathbf{u} \in \mathbb{R}_+^n} \{\tilde{Q}(Q) \wedge \tilde{\mathbf{u}}(\mathbf{u})\} \quad (3)$$

for $x \in \mathbb{R}_+^n$, where $\tilde{\mathbf{u}}(\mathbf{u}) = \min_{1 \leq i \leq n} \{\tilde{u}_i(u_i)\}$ with $\mathbf{u} = (u_1, u_2, \dots, u_n) \in \mathbb{R}_+^n$.

Lemma 2 ([9]). For any $\tilde{\mathbf{u}} = (\tilde{u}_1, \tilde{u}_2, \dots, \tilde{u}_n)' \in \mathcal{F}_c(\mathbb{R}_+)^n$, the following (i) and (ii) hold:

$$(i) \ (\tilde{Q}\tilde{\mathbf{u}})_\alpha = \tilde{Q}_\alpha \tilde{\mathbf{u}}_\alpha \text{ for } \alpha \in [0, 1];$$

$$(ii) \ \tilde{Q}\tilde{\mathbf{u}} \in \mathcal{F}_c(\mathbb{R}_+)^n.$$

The addition and the scalar multiplication on $\mathcal{F}_c(\mathbb{R})$ are defined as follows: For $\tilde{a}, \tilde{b} \in \mathcal{F}_c(\mathbb{R})$ and $\lambda \in \mathbb{R}_+$, define

$$(\tilde{a} + \tilde{b})(x) := \sup_{x_1, x_2 \in \mathbb{R}_+ : x_1 + x_2 = x} \{\tilde{a}(x_1) \wedge \tilde{b}(x_2)\},$$

$$\lambda \tilde{a}(x) := \begin{cases} \tilde{a}(x/\lambda) & \text{if } \lambda > 0 \\ I_{\{0\}}(x) & \text{if } \lambda = 0 \end{cases} \quad (x \in \mathbb{R}_+),$$

where I_A is the indicator of a set A . It is easily shown that, for $\alpha \in [0, 1]$,

$$(\tilde{a} + \tilde{b})_\alpha = \tilde{a}_\alpha + \tilde{b}_\alpha \quad \text{and} \quad (\lambda \tilde{a})_\alpha = \lambda \tilde{a}_\alpha,$$

where the operation on sets is defined ordinary as $A+B := \{x+y \mid x \in A, y \in B\}$ and $\lambda A = \{\lambda x \mid x \in A\}$ for $A, B \subset \mathbb{R}$. The above operations are extended to those on $\mathcal{F}_c(\mathbb{R})^n$ as follows: For $\tilde{\mathbf{u}} = (\tilde{u}_1, \tilde{u}_2, \dots, \tilde{u}_n)'$, $\tilde{\mathbf{v}} = (\tilde{v}_1, \tilde{v}_2, \dots, \tilde{v}_n)' \in \mathcal{F}_c(\mathbb{R})^n$,

$$\tilde{\mathbf{u}} + \tilde{\mathbf{v}} = (\tilde{u}_1 + \tilde{v}_1, \tilde{u}_2 + \tilde{v}_2, \dots, \tilde{u}_n + \tilde{v}_n)',$$

$$\lambda \tilde{\mathbf{u}} = (\lambda \tilde{u}_1, \lambda \tilde{u}_2, \dots, \lambda \tilde{u}_n)'.$$

The Hausdorff metric on $\mathcal{C}(\mathbb{R})$ is denoted by δ , i.e.,

$$\delta([a, b], [c, d]) := |a - c| \vee |b - d|$$

for $[a, b], [c, d] \in \mathcal{C}(\mathbb{R})$, where $x \vee y = \max\{x, y\}$ for $x, y \in \mathbb{R}$. This metric can be extended to $\mathcal{F}_c(\mathbb{R})^n$ by

$$\delta(\tilde{\mathbf{u}}, \tilde{\mathbf{v}}) = \max_{1 \leq i \leq n} \sup_{\alpha \in [0, 1]} \delta((\tilde{u}_i)_\alpha, (\tilde{v}_i)_\alpha)$$

for $\tilde{\mathbf{u}} = (\tilde{u}_1, \tilde{u}_2, \dots, \tilde{u}_n)'$, $\tilde{\mathbf{v}} = (\tilde{v}_1, \tilde{v}_2, \dots, \tilde{v}_n)'$ $\in \mathcal{F}_c(\mathbb{R})^n$. Then, it is known that the metric space $(\mathcal{F}_c(\mathbb{R})^n, \delta)$ is complete.

In Section 2, we describe a nonhomogeneous MDPs by the use of fuzzy sets and specify the optimization problem under average reward criteria. In Section 3, the AEFr from a periodic stationary policy is characterized by a fixed point of a corresponding operator, whose results are applied to derive the optimality equation in Section 4.

2. The model with fuzziness

In this section, we formulate a fuzzy model for nonhomogeneous MDPs with uncertain transition matrices. Let finite sets $S = \{1, 2, \dots, n\}$ and $A = \{1, 2, \dots, k\}$. Our sequential decision model consists of four objects:

$$(S, A, \{\tilde{q}_{ij}(a) \in \mathcal{F}_c([0, 1]), i, j \in S, a \in A\}, r),$$

where $r = r(i, a)$ is a function on $S \times A$ with $r \geq 0$. We interpret S as the set of states of some system and A as the set of actions available at each state. We denote by F the set of all functions from S to A . For any $f \in F$, we define the fuzzy set $\tilde{Q}(f)$ on $\mathcal{P}(S/S)$ as follows:

$$\tilde{Q}(f) := [\tilde{q}_1(f), \tilde{q}_2(f), \dots, \tilde{q}_n(f)]', \quad (4)$$

where for $1 \leq i \leq n$,

$$\tilde{q}_i(f) := (\tilde{q}_{i1}(f(i)), \tilde{q}_{i2}(f(i)), \dots, \tilde{q}_{in}(f(i))). \quad (5)$$

The structural property of the fuzzy sets defined in (4) and (5) is given by Lemma 1.

A policy π is a sequence (f_1, f_2, \dots) of functions with $f_t \in F$ ($t = 1, 2, \dots$). Let Π denote the class of policies. For an integer ν ($\nu \geq 1$), a policy $\pi = (f_1, f_2, \dots)$ is called ν -periodic stationary or simply ν -periodic ([7]) if $f_{\nu t + k} = f_k$ for each $t = 1, 2, \dots$ and k ($1 \leq k \leq \nu - 1$). Such a policy will be denoted by \mathbf{f}^∞ simply by \mathbf{f} , where $\mathbf{f} = (f_1, f_2, \dots, f_\nu) \in F^\nu$. Let Π_ν denote the class of ν -periodic policies. Any $\pi = (f, f, \dots) \in \Pi_1$ is called stationary.

For any $f \in F$, let $r(f)$ be an n -dimensional column vector whose i -th element is $r(i, f(i))$. Applying Zadeh's extension principle ([14]), the fuzzy expected total discounted reward up to time T from a policy π is a element of $\mathcal{F}(\mathbb{R}_+)^n$ and defined as follows:

$$\tilde{\phi}_T(\pi) := (\tilde{\phi}_T(1, \pi), \tilde{\phi}_T(2, \pi), \dots, \tilde{\phi}_T(n, \pi))'$$

with $\tilde{\phi}_T(i, \pi)(x) := \max\{\min_{1 \leq t \leq T} \tilde{Q}(f_t)(Q_t)\}$ for all $x \in \mathbb{R}_+$, $1 \leq i \leq n$, where the maximum is taken over

$$\{Q_1, Q_2, \dots, Q_T \mid Q_t \in \mathcal{P}(S/S) \ (1 \leq t \leq T), \\ x = (r(f_1) + Q_1 r(f_2) + \dots + Q_1 Q_2 \dots Q_T r(f_{T+1}))_i\}.$$

Then, for any policy $\pi \in \Pi$, it holds from [9, Lemma 3.1] that $\tilde{\phi}_T(\pi) \in \mathcal{F}_c(\mathbb{R}_+)^n$ for all $T \geq 1$.

Here, applying the definition of the supremum of fuzzy numbers in Congxin and Cong [1], we will define the average expected reward for the decision process operating over a long time horizon. For each $\alpha \in [0, 1]$ and $i \in S$, let

$$\bar{\phi}_\alpha(i, \pi) = \liminf_{T \rightarrow \infty} \frac{1}{T} \tilde{\phi}_{T, \alpha}(i, \pi), \quad (6)$$

where $\tilde{\phi}_{T, \alpha}(i, \pi)$ is the α -cut of $\tilde{\phi}_T(i, \pi)$ and for a sequence $\{D_k\} \subset \mathcal{C}(\mathbb{R}_+)$, $\liminf_{k \rightarrow \infty} D_k = \{x \in \mathbb{R}_+ \mid \limsup_{l \rightarrow \infty} \delta(x, D_l) = 0\}$ and $\delta(x, D_l) = \inf_{y \in D} |x - y|$ for $D \in \mathcal{C}(\mathbb{R}_+)$.

Now, let $\phi_\alpha(i, \pi) = \bigcap_{0 \leq \alpha' < \alpha} \bar{\phi}_{\alpha'}(i, \pi)$ for each $\alpha \in (0, 1]$. Then, since $\bar{\phi}_\alpha(i, \pi) \in \mathcal{C}(\mathbb{R}_+)$ and $\bar{\phi}_\alpha(i, \pi) \subset \bar{\phi}_{\alpha'}(i, \pi)$ for $\alpha' < \alpha$, using the representative theorem ([14]), we can define a fuzzy number

$$\tilde{\phi}(i, \pi)(x) := \sup_{\alpha \in [0, 1]} \{\alpha \wedge I_{\phi_\alpha(i, \pi)}(x)\}, \quad x \in \mathbb{R}_+.$$

Note that $\tilde{\phi}(\pi) := (\tilde{\phi}(1, \pi), \tilde{\phi}(2, \pi), \dots, \tilde{\phi}(n, \pi)) \in \mathcal{F}_c(\mathbb{R}_+)^n$. We call $\tilde{\phi}(\pi)$ an AEFr vector from a policy π .

Here, we will give a partial order \preceq on $\mathcal{C}(\mathbb{R}_+)$ by the definition: For $[a, b], [c, d] \in \mathcal{C}(\mathbb{R}_+)$,

$$[a, b] \preceq [c, d] \quad \text{if } a \leq c \text{ and } b \leq d, \\ [a, b] \prec [c, d] \quad \text{if } [a, b] \preceq [c, d] \text{ and } [a, b] \neq [c, d].$$

This partial order \preceq on $\mathcal{C}(\mathbb{R}_+)$, called a fuzzy max order, is defined as follows: For $\tilde{u}, \tilde{v} \in \mathcal{F}_c(\mathbb{R}_+)$,

$$\tilde{u} \preceq \tilde{v} \quad \text{if } \tilde{u}_\alpha \preceq \tilde{v}_\alpha \text{ for all } \alpha \in [0, 1], \\ \tilde{u} \prec \tilde{v} \quad \text{if } \tilde{u} \preceq \tilde{v} \text{ and } \tilde{u} \neq \tilde{v}.$$

Also, the partial order on $\mathcal{F}_c(\mathbb{R}_+)^n$ is given by the definition: For $\tilde{\mathbf{u}} = (\tilde{u}_1, \tilde{u}_2, \dots, \tilde{u}_n)'$, $\tilde{\mathbf{v}} = (\tilde{v}_1, \tilde{v}_2, \dots, \tilde{v}_n)'$ $\in \mathcal{F}_c(\mathbb{R}_+)^n$,

$$\tilde{\mathbf{u}} \preceq \tilde{\mathbf{v}} \quad \text{if } \tilde{u}_i \preceq \tilde{v}_i \text{ for all } i = 1, 2, \dots, n, \\ \tilde{\mathbf{u}} \prec \tilde{\mathbf{v}} \quad \text{if } \tilde{\mathbf{u}} \preceq \tilde{\mathbf{v}} \text{ and } \tilde{\mathbf{u}} \neq \tilde{\mathbf{v}}.$$

In order to insure the ergodicity of the process, we introduce the minorization condition (L_ν) which is assumed to remain operative throughout this paper.

Minorization Condition (L_ν) ([5, 7]). There exists an integer ν ($\nu \geq 1$) and $\varepsilon > 0$ such that

$$\underline{Q}(f_1) \dots \underline{Q}(f_\nu) \geq \varepsilon E \quad \text{for all } f_1, f_2, \dots, f_\nu \in F,$$

where $\underline{Q}(f) := (\min(\tilde{q}_{ij}(f))_0)$, $Q(f) = (\tilde{q}_{ij}(f))$ for $f \in F$ and $E = (e_{ij})$ with $e_{ij} = 1(1 \leq i, j \leq n)$.

Our problem is to maximize the $\tilde{\phi}(\pi)$ over all $\pi \in \Pi$ with respect to the partial order \preceq under the minorization condition (L_ν) .

3. Periodic policies and operators

In this section, under the minorization condition (L_ν) the AEFr vector from a ν -periodic policy will be characterized by the use of a unique fixed point of a corresponding operator.

Associated with each function $f \in F$ is a corresponding operator $U(f) : \mathcal{F}_c(\mathbb{R}_+)^n \rightarrow \mathcal{F}_c(\mathbb{R}_+)^n$ defined as follows: For $\tilde{\mathbf{u}} \in \mathcal{F}_c(\mathbb{R}_+)^n$ and $f \in F$,

$$U(f)\tilde{\mathbf{u}} = r(f) + \tilde{Q}(f)\tilde{\mathbf{u}}, \quad (7)$$

where the arithmetics in (7) are defined in (3).

Lemma 3. For $\tilde{\mathbf{u}} \in \mathcal{F}_c(\mathbb{R}_+)^n$ and $\tilde{\mathbf{v}} \in \mathcal{F}_c(\mathbb{R}_+)^n$, it holds

$$U(f)(\tilde{\mathbf{u}} + \tilde{\mathbf{v}}\mathbf{e}) = U(f)\tilde{\mathbf{u}} + \tilde{\mathbf{v}}\mathbf{e},$$

where $\mathbf{e} = (1, 1, \dots, 1)' \in \mathbb{R}_+^n$.

For any policy $\pi = (f_1, f_2, \dots)$, let $\pi^{-l} = (f_{l+1}, f_{l+2}, \dots)$ for each $l \geq 1$. The sequence $\{\tilde{\phi}_T(\pi)\}_{T=1}^\infty$ is recursively described.

Lemma 4. For any policy $\pi = (f_1, f_2, \dots)$ and each $l \geq 1$, it holds that

$$\tilde{\phi}_T(\pi) = U(f_1)U(f_2) \cdots U(f_l)\tilde{\phi}_{T-l}(\pi^{-l}).$$

From Lemma 4, for $\mathbf{f} = (f_1, f_2, \dots, f_\nu) \in F^\nu$,

$$\tilde{\phi}_{\nu k}(\mathbf{f}) = U(\mathbf{f})^k I_{\{0\}} \quad (k \geq 1),$$

where $I_{\{0\}} \in \mathcal{F}_c(\mathbb{R}_+)^n$ is the crisp set of zero vector $0 \in \mathbb{R}_+^n$ and $U(\mathbf{f}) = U(f_1) \cdots U(f_\nu)$. Applying the minorization condition (L_ν) , for each ν -periodic policy $\mathbf{f} = (f_1, f_2, \dots, f_\nu) \in F^\nu$ we introduce the corresponding operator $V(\mathbf{f}) : \mathcal{F}_c(\mathbb{R}_+)^n \rightarrow \mathcal{F}_c(\mathbb{R}_+)^n$ defined as follows: For $\tilde{\mathbf{u}} = (\tilde{u}_1, \tilde{u}_2, \dots, \tilde{u}_n)' \in \mathcal{F}_c(\mathbb{R}_+)^n$,

$$V(\mathbf{f})\tilde{\mathbf{u}}(x) = \max\left\{\min_{1 \leq t \leq \nu} \tilde{Q}(f_t)(Q_t) \wedge \tilde{\mathbf{u}}(u)\right\}$$

for $x \in \mathbb{R}_+^n$, where the maximum is taken over

$$\begin{aligned} &\{Q_1, Q_2, \dots, Q_T, u \mid Q_t \in \mathcal{P}(S/S) \ (1 \leq t \leq \nu), \\ &x = r(f_1) + Q_1 r(f_2) + \cdots + Q_1 \cdots Q_{\nu-1} r(f_\nu) \\ &+ (Q_1 \cdots Q_\nu - \varepsilon E)u, \ u \in \mathbb{R}_+^n\} \end{aligned}$$

and $\tilde{\mathbf{u}}(u) = \min_{1 \leq i \leq n} \tilde{u}_i(u_i)$ for $u = (u_1, u_2, \dots, u_n)' \in \mathbb{R}_+^n$. Here are some basic properties of $U(\mathbf{f})$.

Lemma 5. Let $\mathbf{f} \in F^\nu$. Then we have:

(i) $V(\mathbf{f})$ is a contraction with modulus $1 - n\varepsilon$;

(ii) $V(\mathbf{f})$ is monotone, i.e., $\tilde{\mathbf{u}} \preceq \tilde{\mathbf{v}}$ implies $V(\mathbf{f})\tilde{\mathbf{u}} \preceq V(\mathbf{f})\tilde{\mathbf{v}}$.

For any $\mathbf{f} \in F^\nu$, let $\tilde{\mathbf{h}}(\mathbf{f}) \in \mathcal{F}_c(\mathbb{R}_+)^n$ be a unique fixed point of $V(\mathbf{f})$, that is,

$$\tilde{\mathbf{h}}(\mathbf{f}) = V(\mathbf{f})\tilde{\mathbf{h}}(\mathbf{f}). \quad (8)$$

Then, we observe that

$$\begin{aligned} (V(\mathbf{f})\tilde{\mathbf{h}}(\mathbf{f}))_\alpha &= [\min(U(\mathbf{f})\tilde{\mathbf{h}}(\mathbf{f}))_\alpha - \min(\varepsilon E\tilde{\mathbf{h}}(\mathbf{f}))_\alpha, \\ &\max(U(\mathbf{f})\tilde{\mathbf{h}}(\mathbf{f}))_\alpha - \max(\varepsilon E\tilde{\mathbf{h}}(\mathbf{f}))_\alpha]. \end{aligned}$$

Noting $[a - c, b - d] + [c, d] = [a, b]$, we get from (8)

$$\tilde{\mathbf{h}}(\mathbf{f}) + \varepsilon E\tilde{\mathbf{h}}(\mathbf{f}) = U(\mathbf{f})\tilde{\mathbf{h}}(\mathbf{f}). \quad (9)$$

Theorem 1. For any ν -periodic policy $\mathbf{f} = (f_1, f_2, \dots, f_\nu) \in F^\nu$,

$$\tilde{\phi}(\mathbf{f}) = \frac{\varepsilon}{\nu} E\tilde{\mathbf{h}}(\mathbf{f}) = \frac{\varepsilon}{\nu} \left(\sum_{j=1}^n \tilde{h}_j(\mathbf{f}) \right) \mathbf{e}, \quad (10)$$

where $\tilde{\mathbf{h}}(\mathbf{f}) = (\tilde{h}_1(\mathbf{f}), \tilde{h}_2(\mathbf{f}), \dots, \tilde{h}_n(\mathbf{f}))'$ is a unique fixed point of $V(\mathbf{f})$.

4. Pareto optimal policy

Here, we confine our attention to the class of ν -periodic stationary policies, which simplifies our discussion under the minorization condition (L_ν) . A policy $\mathbf{f}^* \in \Pi_\nu$ is called Pareto optimal if there is no $\mathbf{f} \in \Pi_\nu$ such that $\tilde{\phi}(\mathbf{f}^*) \prec \tilde{\phi}(\mathbf{f})$. In this section, we derive the optimality equation, by which Pareto optimal policies are characterized.

Let \mathcal{D} be an arbitrary subset of $\mathcal{F}_c(\mathbb{R}_+)^n$. A point $\tilde{\mathbf{u}} \in \mathcal{D}$ is called an efficient element of \mathcal{D} with respect to \preceq on $\mathcal{F}_c(\mathbb{R}_+)^n$ if and only if it holds that there does not exist $\tilde{\mathbf{v}} \in \mathcal{D}$ such that $\tilde{\mathbf{u}} \prec \tilde{\mathbf{v}}$. We denote by $\text{eff}(\mathcal{D})$ the set of all elements of \mathcal{D} efficient with respect to \preceq on $\mathcal{F}_c(\mathbb{R}_+)^n$. For any $\tilde{\mathbf{u}} \in \mathcal{F}_c(\mathbb{R}_+)^n$, let $\mathbf{V}(\tilde{\mathbf{u}}) := \text{eff}\{V(\mathbf{f})\tilde{\mathbf{u}} \mid \mathbf{f} \in F^\nu\}$. Note that $\mathbf{V}(\tilde{\mathbf{u}}) \subset \mathcal{F}_c(\mathbb{R}_+)^n$.

Here, we consider the following fuzzy equation including efficient set-functions $\mathbf{V}(\cdot)$ on $\mathcal{F}_c(\mathbb{R}_+)^n$:

$$\tilde{\mathbf{u}} \in \mathbf{V}(\tilde{\mathbf{u}}), \quad \tilde{\mathbf{u}} \in \mathcal{F}_c(\mathbb{R}_+)^n. \quad (11)$$

The equation (11) is called an optimality equation, by which Pareto optimal policies are characterized. A solution of (11), $\tilde{\mathbf{u}}$, is called maximal if there does not exist any solution $\tilde{\mathbf{u}}'$ of (11) such that $E\tilde{\mathbf{u}} \prec E\tilde{\mathbf{u}}'$. Pareto optimal policies are characterized by maximal solutions of the optimality equation (11).

Theorem 2. A policy $\mathbf{f} \in \Pi_\nu$ is Pareto optimal if and only if the fixed point of the corresponding $V(\mathbf{f})$, $\tilde{\mathbf{h}}(\mathbf{f})$, is a maximal solution to the optimal equation (11).

The form of the optimal equation (11) is corresponding to a fuzzy extension of the average case of MDPs ([2,13]).

As a simple example, we consider a fuzzy treatment for a machine maintenance problem dealt with in [10,p.1,pp.17–18].

Example 1. (a machine maintenance problem). A machine can be operated synchronously, say, once an hour. At each period there are two states; one is operating (state 1), and the other is in failure (state 2). If the machine fails, it can be restored to perfect functioning by repair. At each period, if the machine is running, we earn the return of \$ 3.00 per period; the fuzzy set of probability of being in state 1 at the next step is (0.6, 0.7, 0.8) and that of the probability of moving to state 2 is (0.2, 0.3, 0.4), where for any $0 \leq a < b < c \leq 1$, the fuzzy number (a, b, c) on $[0, 1]$ is defined by

$$(a, b, c)(x) = \begin{cases} (x - a)/(b - a) \vee 0 & \text{if } 0 \leq x \leq b \\ (x - c)/(b - c) \vee 0 & \text{if } b \leq x \leq 1. \end{cases}$$

If the machine is in failure, we have two actions to repair the failed machine; one is a rapid repair, denoted by 1, that yields the cost of \$ 2.00 (that is, a return of -\$2.00) with the fuzzy set (0.5, 0.6, 0.7) of the probability moving in state 1 and the fuzzy set (0.3, 0.4, 0.5) of the probability being in state 2; another is a usual repair, denoted by 2, that requires the cost of \$1.00 (that is, a return of -\$1.00) with the fuzzy set (0.3, 0.4, 0.5) of the probability moving in state 1 and the fuzzy set (0.5, 0.6, 0.7) of the probability being in state 2.

For the model considered, $S = \{1, 2\}$ and there exists two stationary policies, $F = \{f_1, f_2\}$ with $f_1(2) = 1$ and $f_2(2) = 2$, where f_1 denotes a policy of the usual repair and f_2 a policy of the rapid repair. We easily observe that $r(f_1) = (3, -2)'$ and

$$\tilde{Q}(f_1) = \begin{pmatrix} (0.6, 0.7, 0.8) & (0.2, 0.3, 0.4) \\ (0.5, 0.6, 0.7) & (0.3, 0.4, 0.5) \end{pmatrix},$$

Now, applying Theorem 1, we can obtain the AEFR $\tilde{\phi}(f_1)$. Here, we observe that the minorization condition (L_ν) holds for $\nu = 1$ and $\varepsilon = 0.2$. After a simple calculation, we find $\tilde{\mathbf{h}}(f_1)$ and $\tilde{\phi}(f_1)$ are

$$\begin{aligned} & ((85/18, 110/18, 135/18), (-17/18, 8/18, 32/18))', \\ & ((7/9, 12/9, 17/9), (7/9, 12/9, 17/9))'. \end{aligned}$$

Next, we find that $V(f_2)\tilde{\mathbf{h}}(f_1)$ and $V(f_1)\tilde{\mathbf{h}}(f_1)$ are

$$\begin{aligned} & ((85/18, 110/18, 135/18), (-17/18, 8/18, 32/18))', \\ & ((85/18, 110/18, 135/18), (-15/18, 10/18, 35/18))', \end{aligned}$$

which satisfies $V(f_2)\tilde{\mathbf{h}}(f_1) \prec \tilde{\mathbf{h}}(f_1)$. Thus, $\tilde{\mathbf{h}}(f_1) \in \mathbf{V}(\tilde{\mathbf{h}}(f_1))$, so that from Theorem 2 f_1 is Pareto optimal in Π_1 . In fact, we can find, by solving (8) for f_2 , that $\tilde{\phi}(f_2) = ((5/7, 9/7, 13/7), (5/7, 9/7, 13/7))'$, and $\tilde{\phi}(f_2) \prec \tilde{\phi}(f_1)$.

References

- [1] W.Congxin and C.Cong, "The supremum and infimum of the set of fuzzy numbers and its application," *J. Math. Anal. Appl.*, Vol.210, pp.499-511, 1997.
- [2] N.Furukawa, "Characterization of optimal policies in vector-valued Markovian decision processes," *Math. Oper. Res.*, Vol.5, pp.271-279, 1980.
- [3] D.J.Hartfiel and E.Seneta, "On the theory of Markov Set-chains," *Adv. Appl. Prob.*, Vol.26, pp.947-964, 1994.
- [4] D.J.Hartfiel, *Markov Set-chains*, Berlin: Springer-Verlag, 1998.
- [5] O.Hernandez-Lerma, *Adaptive Markov Control Processes*, New York: Springer-Verlag, 1989.
- [6] M.Hosaka, M.Horiguchi and M.Kurano, "Controlled Markov set-chains under average criteria," *Applied Mathematics and Computation*, to appear.
- [7] M.Kurano, "Markov decision processes with a Borel measurable cost function - The average case," *Math. Oper. Res.*, Vol.11, pp.309-320, 1986.
- [8] M.Kurano, J.Song, M.Hosaka and Y.Huang, "Controlled Markov Set-Chains with Discounting," *J. Appl. Prob.*, Vol.35, pp.293-302, 1998.
- [9] M.Kurano, M.Yasuda, J.Nakagami and Y.Yoshida, "A fuzzy treatment of uncertain Markov decision processes," submitted.
- [10] H.Mine and S.Osaki, *Markov Decision Processes*, Amsterdam: Elsevier, 1970.
- [11] E.Nummelin, *General irreducible Markov chains and non-negative operators*, Cambridge: Cambridge University Press, 1984.
- [12] R.Stowinski, (ed.), *Fuzzy Sets in Decision Analysis, Operation Research and Statistics*, Kluwer Academic Publishers, 1998.
- [13] D.J.White, "Multi-objective infinite-horizon discounted Markov Decision Processes," *J. Math. Anal. Appl.*, Vol.89, pp.639-647, 1982.
- [14] L.A.Zadeh, "Fuzzy sets," *Inform. and Control*, Vol.8, pp.338-353, 1965.