数理解析研究所講究録
1132巻 2000年 221-229

A fuzzy treatment of uncertain Markov decision processes: Average case

千葉大学教育学部 蔵野正美 (Masami KURANO)
千葉大学理学部 安田正實 (Masami YASUDA)
千葉大学理学部 中神潤一 (Jun-ichi NAKAGAMI)
北九州大学経済学部 吉田祐治 (Yuji YOSHIDA)

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 optimality equation including efficient set-functions. As a numerical example, the machine maintenance problem is considered.

1. Introduction and notation

In modeling in terms of Markov decision processes (MDPs for short, cf. [1, 7, 9, 16, 20]), we often encounter the following two cases: (i) The information on the state-transition probabilities includes 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 (cf. [21]).

In our previous paper [14], 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 [14]. 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 (cf. [6, 10]), called a minorization condition, which is often used in the study of Markov chains (ch. [19]). By use of this property, a Pareto optimal periodic stationary policies are characterized as a maximal solution of optimality equation including efficient set functions. As a numerical example, the machine maintenance problem is considered.

Recently, applying Hartfiel's[4, 5] interval method for Markov chains, Kurano et al.[12] 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.[8] treated the average reward problem for a controlled Markov set-chain. Another approach to the average case has been given in [13].

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 [8]. For the optimization of fuzzy dynamic system, refer to [11, 23].

In the remainder of this section, we shall give some notations and preliminary lemmas on fuzzy sets and interval arithmetics. 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.

We adopt the notation in [4, 5, 14, 17]. 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 $\widetilde{a}$ on X by its membership function $\widetilde{a} : X \to [0,1]$. Denote by $\mathcal{F}(X)$ the set of all fuzzy sets on X. For the theory of fuzzy sets, refer to Zadeh[24] and Novák[18]. The α-cut ($\alpha \in [0,1]$) of the fuzzy set $\widetilde{a} \in \mathcal{F}(X)$ is defined as

$$\widetilde{a}_\alpha := \{x \in X \mid \widetilde{a}(x) \geq \alpha\} \ (\alpha > 0) \quad \text{and} \quad \widetilde{a}_0 := \mathrm{cl}\{x \in X \mid \widetilde{a}(x) > 0\},$$

where cl denotes the closure of the set. For any interval Y in $\mathbb{R}$, $\widetilde{a} \in \mathcal{F}(Y)$ is called a fuzzy number on Y if $\widetilde{a}$ has the following properties (i) – (iv): (i) $\widetilde{a}$ is normal, i.e., there exists an $x_0 \in Y$ with $\widetilde{a}(x_0) = 1$; (ii) $\widetilde{a}$ is convex, i.e., $\widetilde{a}(\alpha x + (1-\alpha)y) \geq \widetilde{a}(x) \wedge \widetilde{a}(y)$ for all $x, y \in Y$ and $\alpha \in [0,1]$, where $a \wedge b = \min\{a, b\}$; (iii) $\widetilde{a}$ is upper semi-continuous; (iv) $\widetilde{a}_0$ is a compact subset of Y.

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 $\widetilde{a} \in \mathcal{F}_c(Y)$ means $\widetilde{a}_\alpha \in \mathcal{C}(Y)$ 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.,

$$\mathcal{F}_c(Y)^n := \{\widetilde{\boldsymbol{u}} = (\widetilde{u}_1, \widetilde{u}_2, \ldots, \widetilde{u}_n)' \mid \widetilde{u}_i \in \mathcal{F}_c(Y) \ (1 \leq i \leq n)\},$$

where d' denotes the transpose of a vector d.

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

$$\mathcal{P}(S) := \{p = (p_1, p_2, \ldots, p_n) \mid p_j \geq 0 \ (1 \leq j \leq n), \sum_{j=1}^{n} p_j = 1\}.$$

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

$$[\widetilde{p}](p) = \min_{1 \leq j \leq n} \{\widetilde{p}_j(p_j)\} \quad \text{for any} \quad p = (p_1, p_2, \ldots, p_n) \in \mathcal{P}(S). \tag{1.1}$$

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,

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

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

$$\widetilde{Q}(Q) := \min_{1 \leq i \leq n} \{[\widetilde{q}_i](q_i)\}, \tag{1.2}$$

where $Q = (q_1, q_2, \ldots, q_n)' \in \mathcal{P}(S/S)$, $q_i = (q_{i1}, q_{i2}, \ldots, q_{in}) \in \mathcal{P}(S)$ and $[\widetilde{q}_i]$ is the fuzzy set on $\mathcal{P}(S)$ defined by (1.1).

In order to describe the structural properties on the fuzzy sets defined in (1.1) and (1.2), we need the concept of intervals of matrices. For the detail, refer to [5, 12, 17]. For any nonnegative vector $\underline{q} = (\underline{q}_1, \underline{q}_2, \ldots, \underline{q}_n)$ and $\overline{q} = (\overline{q}_1, \overline{q}_2, \ldots, \overline{q}_n) \in \mathbb{R}^n_+$ with $\underline{q} \leq \overline{q}$, we define the interval $\langle \underline{q}, \overline{q} \rangle \subset \mathcal{P}(S)$ by

$$\langle \underline{q}, \overline{q} \rangle := \{p = (p_1, p_2, \ldots, p_n) \in \mathcal{P}(S) \mid \underline{q} \leq p \leq \overline{q}\}. \tag{1.3}$$

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

$$\langle \underline{Q}, \overline{Q} \rangle := \{Q \in \mathcal{P}(S/S) \mid \underline{Q} \leq Q \leq \overline{Q}\}. \tag{1.4}$$

For any $\widetilde{a} \in \mathcal{F}_c([0,1])$, noting $\widetilde{a}_\alpha \in \mathcal{C}([0,1])$ $(0 \le \alpha \le 1)$, it will be denoted by $\widetilde{a}_\alpha = [\min \widetilde{a}_\alpha, \max \widetilde{a}_\alpha]$. The structural property of the fuzzy sets defined in (2.1) and (2.2) is given, whose proof is done by using the following Lemma 1.1.

Lemma 1.1 ([5, 14]).

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

(ii) *For any $\widetilde{q}_i \in \mathcal{F}_c([0,1])^n$ $(1 \le i \le n)$, let $\widetilde{Q} = [\widetilde{q}_1, \widetilde{q}_2, \dots, \widetilde{q}_n]'$ be a fuzzy set on $\mathcal{P}(S/S)$ defined by (1.2). Then, the α-cut of $\widetilde{Q}$ $(0 \le \alpha \le 1)$ is a polyhedral convex subset of $\mathcal{P}(S/S)$ and given by*

$$\widetilde{Q}_\alpha = \langle \underline{Q}_\alpha, \overline{Q}_\alpha\rangle, \quad \text{where } \underline{Q}_\alpha = \big(\min(\widetilde{q}_{ij})_\alpha\big) \text{ and } \overline{Q}_\alpha = \big(\max(\widetilde{q}_{ij})_\alpha\big). \tag{1.5}$$

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

$$\langle \underline{Q}, \overline{Q}\rangle \boldsymbol{u} = \{Qu \mid Q \in \langle \underline{Q}, \overline{Q}\rangle, u \in \boldsymbol{u}\}. \tag{1.6}$$

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

$$(\widetilde{Q}\widetilde{\boldsymbol{u}})(x) = \max_{\substack{x=Qu \\ Q\in\mathcal{P}(S/S), u\in\mathbb{R}_+^n}} \{\widetilde{Q}(Q) \wedge \widetilde{\boldsymbol{u}}(u)\}, \quad \text{for } x \in \mathbb{R}_+^n, \quad \text{where} \tag{1.7}$$

$$\widetilde{\boldsymbol{u}}(u) = \min_{1\le i\le n} \{\widetilde{u}_i(u_i)\} \quad \text{with} \quad u = (u_1,u_2,\dots,u_n) \in \mathbb{R}_+^n. \tag{1.8}$$

Lemma 1.2([14]). *For any $\widetilde{\boldsymbol{u}} = (\widetilde{u}_1, \widetilde{u}_2, \dots, \widetilde{u}_n)' \in \mathcal{F}_c(\mathbb{R}_+)^n$*
(i) $(\widetilde{Q}\widetilde{\boldsymbol{u}})_\alpha = \widetilde{Q}_\alpha \widetilde{\boldsymbol{u}}_\alpha$ for $\alpha \in [0,1]$; (ii) $\widetilde{Q}\widetilde{\boldsymbol{u}} \in \mathcal{F}_c(\mathbb{R}_+)^n$.

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

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

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

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

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

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 $\widetilde{\boldsymbol{u}} = (\widetilde{u}_1, \widetilde{u}_2, \dots, \widetilde{u}_n)'$, $\widetilde{\boldsymbol{v}} = (\widetilde{v}_1, \widetilde{v}_2, \dots, \widetilde{v}_n)' \in \mathcal{F}_c(\mathbb{R})^n$,

$$\widetilde{\boldsymbol{u}} + \widetilde{\boldsymbol{v}} = (\widetilde{u}_1 + \widetilde{v}_1, \widetilde{u}_2 + \widetilde{v}_2, \dots, \widetilde{u}_n + \widetilde{v}_n)' \quad \text{and} \quad \lambda\widetilde{\boldsymbol{u}} = (\lambda\widetilde{u}_1, \lambda\widetilde{u}_2, \dots, \lambda\widetilde{u}_n)'.$$

For $a = (a_1,a_2,\dots,a_n)' \in \mathbb{R}^n$, $I_{\{a\}} = (I_{\{a_1\}}, I_{\{a_2\}}, \dots, I_{\{a_n\}}) \in \mathcal{F}_c(\mathbb{R})^n$ and writing I_a simply by a, $I_{\{a\}} + \widetilde{\boldsymbol{u}}$ is described by $a + \widetilde{\boldsymbol{u}}$. Also, $\widetilde{\boldsymbol{u}} - I_{\{a\}}$ is defined by $\widetilde{\boldsymbol{u}} + I_{\{-a\}}$, whose arithmetic is used in the sequel. The Hausdorff metric on $\mathcal{C}(\mathbb{R})$ is denoted by δ, i.e.,

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

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

$$\delta(\widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{v}}) = \max_{1 \le i \le n} \sup_{\alpha \in [0,1]} \delta((\widetilde{u}_i)_\alpha, (\widetilde{v}_i)_\alpha)$$

for $\widetilde{\boldsymbol{u}} = (\widetilde{u}_1, \widetilde{u}_2, \dots, \widetilde{u}_n)'$, $\widetilde{\boldsymbol{v}} = (\widetilde{v}_1, \widetilde{v}_2, \dots, \widetilde{v}_n)' \in \mathcal{F}_c(\mathbb{R})^n$. Then, it is known(c.f.[15]) that the metric space $(\mathcal{F}_c(\mathbb{R})^n, \delta)$ is complete.

2. The model with fuzziness

In this section, we formulate a fuzzy model for nonhomogenuous MDPs with uncertain transition matrices.

Let S and A be finite sets denoted by $S = \{1, 2, \dots, n\}$ and $A = \{1, 2, \dots, k\}$. Our sequential decision model consists of four objects:

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

where $r = r(i, a)$ is a function on $S \times A$ with $r \ge 0$. 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 $\widetilde{Q}(f)$ on $\mathcal{P}(S/S)$ as follows:

$$\widetilde{Q}(f) := [\widetilde{q}_1(f), \widetilde{q}_2(f), \dots, \widetilde{q}_n(f)]' \quad \text{where} \tag{2.1}$$

$$\widetilde{q}_i(f) := \left(\widetilde{q}_{i1}(f(i)), \widetilde{q}_{i2}(f(i)), \dots, \widetilde{q}_{in}(f(i))\right) \quad (1 \le i \le n). \tag{2.2}$$

Note that the basic notations of (2.1) and (2.2) are defined in (1.1) and (1.2).

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 $\nu (\nu \ge 1)$, a policy $\pi = (f_1, f_2, \dots)$ is called ν-periodic stationary or simply ν-periodic (cf. [10]) if $f_{\nu t + k} = f_k$ for each $t = 1, 2, \dots$ and $k (1 \le k \le \nu - 1)$. Such a policy will be denote by $\boldsymbol{f}^\infty$ simply by $\boldsymbol{f}$, where $\boldsymbol{f} = (f_1, f_2, \dots, f_\nu) \in F^\nu$. 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(cf.[18]), the fuzzy expected total reward up to time T from a policy π is a element of $\mathcal{F}(\mathbb{R}_+)^n$ and defined as follows:

$$\widetilde{\phi}_T(\pi) := (\widetilde{\phi}_T(1, \pi), \widetilde{\phi}_T(2, \pi), \dots, \widetilde{\phi}_T(n, \pi))' \quad \text{and} \tag{2.3}$$

$$\widetilde{\phi}_T(i, \pi)(x) := \max\{\min_{1 \le t \le T} \widetilde{Q}(f_t)(Q_t)\} \quad \text{for all } x \in \mathbb{R}_+,\ 1 \le i \le n, \tag{3.4}$$

where the maximum is taken over

$$\{Q_1, Q_2, \dots, Q_T \mid x = (r(f_1) + Q_1 r(f_2) + \dots + Q_1 Q_2 \cdots Q_T r(f_{T+1}))_i, \\ Q_t \in \mathcal{P}(S/S)\ (1 \le t \le T)\}. \tag{2.5}$$

Then, for any policy $\pi \in \Pi$, it holds from Lemma 3.1 in [14] that

$$\widetilde{\phi}_T(\pi) \in \mathcal{F}_c(\mathbb{R}_+)^n \text{ for all } T \ge 1.$$

Here, applying the definition of the supremum of fuzzy numbers in Congxin and Cong [2], 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

$$\overline{\phi}_\alpha(i, \pi) = \liminf_{T \to \infty} \frac{1}{T} \widetilde{\phi}_{T,\alpha}(i, \pi), \tag{2.6}$$

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

Now, let $\phi_\alpha(i,\pi) = \bigcap_{0\le\alpha'<\alpha} \overline{\phi}_{\alpha'}(i,\pi)$ for each $\alpha \in (0,1]$. Then, since $\overline{\phi}_\alpha(i,\pi) \in \mathcal{C}(\mathbb{R}_+)$ and $\overline{\phi}_\alpha(i,\pi) \subset \overline{\phi}_{\alpha'}(i,\pi)$ for $\alpha' < \alpha$, the following holds obviously.

Lemma 2.1 *For $i \in S$ and $\pi \in \Pi$, we have:*

(i) $\phi_\alpha(i,\pi) \in \mathcal{C}(\mathbb{R}_+)$.

(ii) $\phi_\alpha(i,\pi) \subset \phi(i,\pi)$ *for* $0 \le \alpha' < \alpha \le 1$.

(iii) $\lim_{\alpha'\uparrow\alpha} \phi_{\alpha'}(i,\pi) = \phi_\alpha(i,\pi)$.

Using the representative theorem (cf. [18]), we can define a fuzzy number

$$\widetilde{\phi}(i,\pi)(x) := \sup_{\alpha\in[0,1]} \{\alpha \wedge I_{\phi_\alpha(i,\pi)}(x)\}, \ x \in \mathbb{R}_+. \tag{2.7}$$

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

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

$$\begin{aligned} [a,b] \preccurlyeq [c,d] \quad & \text{if } \ a \le c \text{ and } \ b \le d, \\ [a,b] \prec [c,d] \quad & \text{if } \ [a,b] \preccurlyeq [c,d] \text{ and } \ [a,b] \ne [c,d]. \end{aligned}$$

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

$$\begin{aligned} \widetilde{u} \preccurlyeq \widetilde{v} \quad & \text{if } \ \widetilde{u}_\alpha \preccurlyeq \widetilde{v}_\alpha \text{ for all } \ \alpha \in [0,1], \\ \widetilde{u} \prec \widetilde{v} \quad & \text{if } \ \widetilde{u} \preccurlyeq \widetilde{v} \text{ and } \ \widetilde{u} \ne \widetilde{v}. \end{aligned}$$

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

$$\begin{aligned} \widetilde{\boldsymbol{u}} \preccurlyeq \widetilde{\boldsymbol{v}} \quad & \text{if } \ \widetilde{u}_i \preccurlyeq \widetilde{v}_i \text{ for all } \ i = 1,2,\dots,n, \\ \widetilde{\boldsymbol{u}} \prec \widetilde{\boldsymbol{v}} \quad & \text{if } \ \widetilde{\boldsymbol{u}} \preccurlyeq \widetilde{\boldsymbol{v}} \text{ and } \ \widetilde{\boldsymbol{u}} \ne \widetilde{\boldsymbol{v}}. \end{aligned}$$

The following lemma is used in the sequel whose proof is easily done.

Lemma 2.2 *Let a sequence $\{\widetilde{\boldsymbol{u}}_n\} \subset \mathcal{F}_c(\mathbb{R}_+)^n$ be such that $\widetilde{\boldsymbol{u}}_1 \preccurlyeq \widetilde{\boldsymbol{u}}_2 \preccurlyeq \cdots$, and $\lim_{k\to\infty} \widetilde{\boldsymbol{u}}_k = \widetilde{\boldsymbol{u}}$ for some $\widetilde{\boldsymbol{u}} \in \mathcal{F}_c(\mathbb{R}_+)^n$. Then, it holds that $\widetilde{\boldsymbol{u}}_1 \preccurlyeq \widetilde{\boldsymbol{u}}$.*

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)(cf. [6, 10])
There exists an integer $\nu (\nu \ge 1)$ and $\varepsilon > 0$ such that

$$\underline{Q}(f_1) \cdots \underline{Q}(f_\nu) \ge \varepsilon E \quad \textit{for all} \ \ f_1, f_2, \cdots, f_\nu \in F,$$

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

Our problem is to maximize the $\widetilde{\phi}(\pi)$ over all $\pi \in \Pi$ with respect to the partial order $\preccurlyeq$ 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 $U(f) : \mathcal{F}_c(\mathbb{R}_+)^n \to \mathcal{F}_c(\mathbb{R}_+)^n$ defined as follows: For $\widetilde{\boldsymbol{u}} \in \mathcal{F}_c(\mathbb{R}_+)^n$ and $f \in F$,

$$U(f)\widetilde{\boldsymbol{u}} = r(f) + \widetilde{Q}(f)\widetilde{\boldsymbol{u}}, \tag{3.1}$$

where the arithmetics in (3.1) are defined in (1.7). Note that from Lemma 1.2 $U(f)$ is well-defined. The following holds obviously.

Lemma 3.1 *For $\widetilde{\boldsymbol{u}} \in \mathcal{F}_c(\mathbb{R}_+)^n$ and $\widetilde{v} \in \mathcal{F}_c(\mathbb{R}_+)$, it holds*

$$U(f)(\widetilde{\boldsymbol{u}} + \widetilde{v}\boldsymbol{e}) = U(f)\widetilde{\boldsymbol{u}} + \widetilde{v}\boldsymbol{e}, \quad \textit{where } \boldsymbol{e} = (1, 1, \ldots, 1)' \in \mathbb{R}^n_+.$$

For any policy $\pi = (f_1, f_2, \ldots)$, let $\pi^{-l} = (f_{l+1}, f_{l+2}, \ldots)$ for each $l \geq 1$. The sequence $\{\widetilde{\phi}_T(\pi)\}_{T=1}^\infty$ is recursively described, whose proof is the same as that of Lemma 4.1 in [14].

Lemma 3.2 *For any policy $\pi = (f_1, f_2, \ldots)$, it holds*

$$\widetilde{\phi}_T(\pi) = U(f_1)U(f_2)\cdots U(f_l)\widetilde{\phi}_{T-l}(\pi^{-l}) \quad \textit{for each } \; l \geq 1. \tag{3.2}$$

From Lemma 3.2, we have that for $\boldsymbol{f} = (f_1, f_2, \ldots, f_\nu) \in F^\nu$,

$$\widetilde{\phi}_{\nu k}(\boldsymbol{f}) = U(\boldsymbol{f})^k 0 \quad (k \geq 1) \tag{3.3}$$

where 0 means $I_{\{0\}} \in \mathcal{F}_c(\mathbb{R}_+)^n$ and

$$U(\boldsymbol{f}) = U(f_1)\cdots U(f_\nu). \tag{3.4}$$

Applying the minorization condition (L_ν), for each ν-periodic policy $\boldsymbol{f} = (f_1, f_2, \ldots, f_\nu) \in F^\nu$ we introduce the corresponding operator $V(\boldsymbol{f}) : \mathcal{F}_c(\mathbb{R}_+)^n \to \mathcal{F}_c(\mathbb{R}_+)^n$ defined as follows: For $\widetilde{\boldsymbol{u}} = (\widetilde{u}_1, \widetilde{u}_2, \ldots, \widetilde{u}_n)' \in \mathcal{F}_c(\mathbb{R}_+)^n$,

$$V(\boldsymbol{f})\widetilde{\boldsymbol{u}}(x) = \max\{\min_{1\leq t\leq\nu} \widetilde{Q}(f_t)(Q_t) \wedge \widetilde{\boldsymbol{u}}(u)\} \quad (x \in \mathbb{R}^n_+), \tag{3.5}$$

where the maximum is taken over

$$\begin{aligned}\{Q_1, Q_2, \ldots Q_T, u \mid x = {} & r(f_1) + Q_1 r(f_2) + \cdots + Q_1\cdots Q_{\nu-1}r(f_{\nu-1}) \\ & + (Q_1\cdots Q_\nu - \varepsilon E)u,\; Q_t \in \mathcal{P}(S/S)\; (1 \leq t \leq \nu),\; u \in \mathbb{R}^n_+\}\end{aligned} \tag{3.6}$$

and

$$\widetilde{\boldsymbol{u}}(u) = \min_{1\leq i\leq n} \widetilde{u}_i(u_i) \quad \text{for } \; u = (u_1, u_2, \ldots, u_n)' \in \mathbb{R}^n_+. \tag{3.7}$$

Obviously, $V(\boldsymbol{f})\widetilde{\boldsymbol{u}} \in \mathcal{F}_c(\mathbb{R}_+)^n$ for $\widetilde{\boldsymbol{u}} \in \mathcal{F}_c(\mathbb{R}_+)^n$, so that $V(\boldsymbol{f})$ is well-defined.

Here are some basic properties of $V(\boldsymbol{f})$.

Lemma 3.3. *Let $\boldsymbol{f} \in F^\nu$. Then we have:*

(i) *$V(\boldsymbol{f})$ is a contraction with modulus $1 - n\varepsilon$.*

(ii) *$V(\boldsymbol{f})$ is monotone, i.e., $\widetilde{\boldsymbol{u}} \preccurlyeq \widetilde{\boldsymbol{v}}$ implies $V(\boldsymbol{f})\widetilde{\boldsymbol{u}} \preccurlyeq V(\boldsymbol{f})\widetilde{\boldsymbol{v}}$.*

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

$$\widetilde{\boldsymbol{h}}(\boldsymbol{f}) = V(\boldsymbol{f})\widetilde{\boldsymbol{h}}(\boldsymbol{f}). \tag{3.9}$$

Then, by (3.1), (3.4) and (3.5) to (3.7), we observe that

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

Noting $[a-c, b-d] + [c, d] = [a, b]$, we get from (3.9)

$$\widetilde{\boldsymbol{h}}(\boldsymbol{f}) + \varepsilon E\widetilde{\boldsymbol{h}}(\boldsymbol{f}) = U(\boldsymbol{f})\widetilde{\boldsymbol{h}}(\boldsymbol{f}). \tag{3.10}$$

Theorem 3.1. *For any ν-periodic policy $\boldsymbol{f} = (f_1, f_2, \dots, f_\nu) \in F^\nu$,*

$$\widetilde{\phi}(\boldsymbol{f}) = \frac{\varepsilon}{\nu} E\widetilde{\boldsymbol{h}}(\boldsymbol{f}) = \frac{\varepsilon}{\nu}\left(\sum_{j=1}^{n} \widetilde{h}_j(\boldsymbol{f})\right)\boldsymbol{e}, \tag{3.11}$$

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

As a simple example, we consider a fuzzy treatment for a machine maintenance problem dealt with in ([16], p.1, p.17–18).

An example (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 $0 \le a < b < c \le 1$, the fuzzy number (a, b, c) on $[0, 1]$ is defined by

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

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 rapid repair and f_2 a policy of the usual repair. We easily observe that

$$r(f_1) = \begin{pmatrix} 3 \\ -2 \end{pmatrix} \text{ and } \widetilde{Q}(f_1) = \begin{pmatrix} (0.6, 0.7, 0.8) & (0.2, 0.3, 0.4) \\ (0.5, 0.6, 0.7) & (0.3, 0.4, 0.5) \end{pmatrix},$$

Applying Theorem 3.1, we can obtain the AEFR $\widetilde{\phi}(f_1)$. After some calculations, we find

$$\widetilde{\boldsymbol{h}}(f_1)_\alpha = \left([\frac{85+25\alpha}{18}, \frac{135-25\alpha}{18}], [\frac{-15+25\alpha}{18}, \frac{35-25\alpha}{18}]\right)',$$

which leads to

$$\widetilde{\boldsymbol{h}}(f_1) = \left((\frac{85}{18}, \frac{110}{18}, \frac{135}{18}), (\frac{-15}{18}, \frac{10}{18}, \frac{35}{18})\right)'.$$

By (3.11),

$$\widetilde{\phi}(f_1) = \left((\frac{7}{9}, \frac{12}{9}, \frac{17}{9}), (\frac{7}{9}, \frac{12}{9}, \frac{17}{9})\right)'.$$

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 $\boldsymbol{f}^* \in \Pi_\nu$ is called Pareto optimal if there is no $\boldsymbol{f} \in \Pi_\nu$ such that $\widetilde{\phi}(\boldsymbol{f}^*) \prec \widetilde{\phi}(\boldsymbol{f})$. In this section, we derive the optimality equation, by which Pareto optimal policies are characterized.

The following important result is crucial to the development in the characterization of Pareto optimality.

Lemma 4.1. *For any $\boldsymbol{f}, \boldsymbol{g} \in \Pi_\nu$, let $\widetilde{\boldsymbol{h}}(\boldsymbol{f})$ and $\widetilde{\boldsymbol{h}}(\boldsymbol{g})$ be the fixed points of the corresponding operators $V(\boldsymbol{f})$ and $V(\boldsymbol{g})$. Suppose that*

$$\widetilde{\boldsymbol{h}}(\boldsymbol{f}) \ \begin{Bmatrix} \preccurlyeq \\ \prec \end{Bmatrix} \ V(\boldsymbol{g})\widetilde{\boldsymbol{h}}(\boldsymbol{f}). \tag{4.1}$$

Then, it holds that

$$\widetilde{\boldsymbol{h}}(\boldsymbol{f}) \ \begin{Bmatrix} \preccurlyeq \\ \prec \end{Bmatrix} \ \widetilde{\boldsymbol{h}}(\boldsymbol{g}). \tag{4.2}$$

Let $\mathcal{D}$ be an arbitrary subset of $\mathcal{F}_c(\mathbb{R}_+)^n$. A point $\widetilde{\boldsymbol{u}} \in \mathcal{D}$ is called an efficient element of $\mathcal{D}$ with respect to $\preccurlyeq$ on $\mathcal{F}_c(\mathbb{R}_+)^n$ if and only if it holds that there does not exist $\widetilde{\boldsymbol{v}} \in \mathcal{D}$ such that $\widetilde{\boldsymbol{u}} \prec \widetilde{\boldsymbol{v}}$. We denote by $\mathrm{eff}(\mathcal{D})$ the set of all elements of $\mathcal{D}$ efficient with respect to $\preccurlyeq$ on $\mathcal{F}_c(\mathbb{R}_+)^n$. For any $\widetilde{\boldsymbol{u}} \in \mathcal{F}_c(\mathbb{R}_+)^n$, let $\mathcal{V}(\widetilde{\boldsymbol{u}}) := \mathrm{eff}(\{V(\boldsymbol{f})\widetilde{\boldsymbol{u}} \mid \boldsymbol{f} \in F^\nu\})$. Note that $\mathcal{V}(\widetilde{\boldsymbol{u}}) \subset \mathcal{F}_c(\mathbb{R}_+)^n$. Here, we consider the following fuzzy equation including efficient set-functions $\mathcal{V}(\cdot)$ on $\mathcal{F}_c(\mathbb{R}_+)^n$:

$$\widetilde{\boldsymbol{u}} \in \mathcal{V}(\widetilde{\boldsymbol{u}}), \quad \widetilde{\boldsymbol{u}} \in \mathcal{F}_c(\mathbb{R}_+)^n. \tag{4.3}$$

The equation (4.3) is called an optimality equation, by which Pareto optimal policies are characterized. A solution of (4.3), $\widetilde{\boldsymbol{u}}$, is called maximal if there does not exist any solution $\widetilde{\boldsymbol{u}}'$ of (4.3) such that $E\widetilde{\boldsymbol{u}} \prec E\widetilde{\boldsymbol{u}}'$. Pareto optimal policies are characterized by maximal solutions of the optimality equation (4.3).

Theorem 4.1. *A policy $\boldsymbol{f} \in \Pi_\nu$ is Pareto optimal if and only if the fixed point of the corresponding $V(\boldsymbol{f})$, $\widetilde{\boldsymbol{h}}(\boldsymbol{f})$, is a maximal solution to the optimal equation (4.3).*

Remark. For vector-valued discounted MDPs, Furukawa[3] and White[22] had derived the optimality equation including efficient set-function on $\mathbb{R}^n$, by which Pareto optimal policies are characterized. The form of the optimal equation (4.3) is corresponding to the average case of MDPs with fuzziness.

For the machine maintenance problem in Section 3, we find that

$$V(f_2)\widetilde{\boldsymbol{h}}(f_1) = \Big((\frac{85}{18}, \frac{110}{18}, \frac{135}{18}), (\frac{-17}{18}, \frac{8}{18}, \frac{32}{18}) \Big)',$$

Recall that

$$V(f_1)\widetilde{\boldsymbol{h}}(f_1) = \widetilde{\boldsymbol{h}}(f_1) = \Big((\frac{85}{18}, \frac{110}{18}, \frac{135}{18}), (\frac{-15}{18}, \frac{10}{18}, \frac{35}{18}) \Big)',$$

which satisfies $V(f_2)\widetilde{\boldsymbol{h}}(f_1) \prec \widetilde{\boldsymbol{h}}(f_1)$. Thus, $\widetilde{\boldsymbol{h}}(f_1) \in \mathcal{V}(\widetilde{\boldsymbol{h}}(f_1))$, so that from Theorem 4.1 f_1 is Pareto optimal in Π_1. In fact, we can find, by solving (3.9) for f_2, that

$$\widetilde{\phi}(f_2) = \Big((\frac{5}{7}, \frac{9}{7}, \frac{13}{7}), (\frac{5}{7}, \frac{9}{7}, \frac{13}{7}) \Big)', \text{ and } \widetilde{\phi}(f_2) \prec \widetilde{\phi}(f_1).$$

References

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

[2] Congxin, W. and Cong, C., The supremum and infimum of the set of fuzzy numbers and its application, *J. Math. Anal. Appl.*, **210** (1997), 499-511.

[3] Furukawa, N., Characterization of optimal policies in vector-valued Markovian decision processes, *Math. Oper. Res.* **5** (1980), 271-279.

[4] Hartfiel, D. J. and Seneta, E., On the theory of Markov Set-chains, *Adv. Appl. Prob.* **26** (1994), 947-964.

[5] Hartfiel, D. J., *Markov Set-chains*, (1998), Springer-Verlag, Berlin.

[6] Hernondex-Lerma, O., *Adaptive Markov Control Processes*, (1989), Springer-Verlag, New-York Inc.

[7] Hinderer, K., *Foundations of Non-Stationary Dynamic Programming with Discrete Time Parameter,* (1970), Springer-Verlag, New York.

[8] Hosaka, M., Horiguchi, M. and Kurano, M., Controlled Markov set-chains under average criteria, *submitted to Proceedings of 7th BC* , Santa Fe, (1999).

[9] Howard, R., *Dynamic Programming and Markov processes*, (1960), MIT Press, Cambridge MA.

[10] Kurano, M., Markov decision processes with a Borel measureable cost function – The average case, *Math. Oper. Res.*, **11** (1986), 309-320.

[11] Kurano, M., Yasuda, M., Nakagami, J. and Yoshida, Y., Markov-type fuzzy decision processes with a discounted reward on a closed interval, *European J. O. R.*, **92** (1996), 649-662.

[12] Kurano, M., Song, J., Hosaka, M. and Huang, Y., Controlled Markov Set-Chains with Discounting, *J. Appl. Prob.*, **35** (1998), 293-302.

[13] Kurano, M., Yasuda, M. and Nakagami, J., Interval methods for uncertain Markov decision processes, *submitted to the International Workshop on Markov Processes and Controlled Markov Chains*, Changsha, Husan, China on August 22-28, 1999.

[14] Kurano, M., Yasuda, M., Nakagami, J. and Yoshida, Y., A fuzzy treatment of uncertain Markov decision processes, *Preprint.*

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

[16] Mine, H. and Osaki, S., *Markov Decision Processes*, (1970), Elsevier, Amsterdam.

[17] Nenmaier, A., New techniques for the analyses of linear interval equations, *Linear Algebra Appli.*, **58** (1984), 273-325.

[18] Novák, V., *Fuzzy sets and their applications*, (1989), Adam Hilger, Bristole-Boston.

[19] Nummelin, E., *General irreducible Markov chains and non-negative operators*, (1984), Cambridge University Press, Cambridge.

[20] Puterman, M. L., *Markov decision processes: Discrete Stochastic Dynamic Programming*, (1994), John Wiley & Sons, Inc.

[21] Stowínski, R., (ed.), *Fuzzy Sets in Decision Analysis, Operation Research and Statistics*, (1998), Kluwer Academic Publishers.

[22] White, D. J., Multi-objective infinite-horizon discounted Markov Decision Processes, *J. Math. Anal. Appl.*, **89** (1982), 639-647.

[23] Yoshida, Y., A time-average fuzzy reward criterion in fuzzy decision processes, *Information Sciences*, **110** (1998), 103-112.

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