Fuzzy perceptive values for stopping models and MDPs

KURANO Masami
YASUDA Masami
NAKAGAMI Jun-ichi
YOSHIDA Yuji

1. Introduction

In a real application of such a mathematical model as an optimal stopping or a Markov decision process (MDP), it often occurs that the required data is linguistically and roughly perceived (for example, the price of the asset is about $100, etc.). A possible way of handling such a perception-based information is to use the fuzzy set (cf. [3]), whose membership function describes the perception value of the required data. If the fuzzy perception of the required data is given, how can we estimate the future expected reward, called a fuzzy perceptive value, under the condition that we can know the true value of the required data immediately before our decision making. The problem formulation is inspired by Zadeh's paper [6], in which the perception-based-theory of probabilistic reasoning is developed and the idea of the perceptive value(possibility distribution) of the objective function under the possibility constraints is proposed by using a generalized extension principle.

Here, we formulate the perceptive models for optimal stopping problems and MDPs, and the corresponding fuzzy perceptive values are characterized and calculated by a fuzzy optimality equations. Numerical examples are given.

Let $\mathbb{R}$ be the set of all real numbers and $\tilde{\mathbb{R}}$ the set of all fuzzy numbers, i.e., $\tilde{s} \in \tilde{\mathbb{R}}$ means that $\tilde{s} : \tilde{\mathbb{R}} \to [0, 1]$ is normal, upper-semicontinuous and fuzzy convex and has a compact support, The $\alpha$-cut of $\tilde{s} \in \tilde{\mathbb{R}}$ is given by $\tilde{s}_\alpha := \{x \in \mathbb{R} \mid \tilde{s} \geq \alpha \} (\alpha \in (0, 1])$ and $\tilde{s}_0 := \{x \in \mathbb{R} \mid \tilde{s} > 0\}$, where $\text{cl} A$ is the closure of a set $A$. We write $\tilde{s} = [\tilde{s}_\alpha, \tilde{s}_0] \alpha \in [0, 1]$. For $\tilde{s}, \tilde{r} \in \tilde{\mathbb{R}}$,

$$\max\{\tilde{s}(x), \tilde{r}(x)\}(y) := \sup_{x_1, x_2 \in \mathbb{R} \atop y = \tilde{s}(x_1) \lor \tilde{r}(x_2)} \{\tilde{s}(x_1) \lor \tilde{r}(x_2)\} \quad (y \in \mathbb{R}),$$

then $\tilde{s} \preceq \tilde{r}$ (fuzzy max order) means $\tilde{r} = \max\{\tilde{s}, \tilde{r}\}$.

2. Perceptive stopping model

Let $\mathcal{X}$ be the set of all integrable random variables on the probability space $(\Omega, \mathcal{M}, P)$ and $\mathcal{X}^n = \{X = (X_1, \ldots, X_n) \mid X_t \in \mathcal{X} \ (1 \leq t \leq n)\}$. For each sequence of random variables $X = (X_1, \ldots, X_n) \in \mathcal{X}^n$, we denote by $\delta = \delta(X)$ the optimal stopping time for $X$ (cf. [1]) with the optimal expected reward $E(X_{\delta^*}) := E(X_{\delta^*(X)})$.

A measurable map $\tilde{\mathcal{X}} : \Omega \to \tilde{\mathbb{R}}$ is called a fuzzy perception on $\mathcal{X}$. For a sequence of fuzzy perceptions $\tilde{\mathcal{X}} = \tilde{X}_1, \ldots, \tilde{X}_n$, the problem is to characterize and compute the perceptive value $E \tilde{\mathcal{X}}_{\delta^*}$ where

$$(2.1) \quad E \tilde{\mathcal{X}}_{\delta^*}(x) = \sup_{x = E(\tilde{X}_{\delta^*})} \tilde{\mathcal{X}}(x),$$

$$(2.2) \quad \tilde{\mathcal{X}}(X) = \sup_{\omega \in \Omega} \tilde{X}_1(\omega)(X_1(\omega)) \land \cdots \land \tilde{X}_n(\omega)(X_n(\omega)) \quad (a \land b = \min\{a, b\}).$$

**Theorem 2.1** The following holds:

(i) $E(\tilde{\mathcal{X}}_{\delta^*}) \in \tilde{\mathbb{R}}$.

(ii) Suppose that $\tilde{\mathcal{X}} = (\tilde{X}_1, \ldots, \tilde{X}_n)$ is independent with each $\tilde{X}_t$ $(t = 1, 2, \ldots, n)$. Then $E \tilde{\mathcal{X}}$ is given by the backward recursive equation.

$$(2.3) \quad \tilde{\gamma}_n = E(\tilde{X}_n), \quad \tilde{\gamma}_k = E\max\{\tilde{X}_k, \tilde{\gamma}_{k+1}\} \quad (k = n - 1, \ldots, 2, 1) \quad \text{and} \quad \tilde{\gamma}_1 = E(\tilde{X}_{\delta^*}).$$

− 124 −
3. Perceptive MDPs

Consider finite state and action spaces, $S$ and $A$, containing $n < \infty$ and $k < \infty$ elements with $S = \{1, 2, \ldots, n\}$ and $A = \{1, 2, \ldots, k\}$. Let $\mathcal{P}(S) \subset \mathbb{R}^n$ and $\mathcal{P}(S|A) \subset \mathbb{R}^{n \times kn}$ be the sets of all probabilities on $S$ and conditional probabilities on $S$ given $S \times A$, that is,

$$
\mathcal{P}(S) := \{q = (q(1), q(2), \ldots, q(n))' \mid q(i) \geq 0, \sum_{i=1}^{n} q(i) = 1, \ i \in S\},
$$

$$
\mathcal{P}(S|A) := \{Q = (q_{ia}() : i \in S, a \in A) \mid q_{ia}() = (q_{ia}(1), q_{ia}(2), \ldots, q_{ia}(n))' \in \mathcal{P}(S), \ i \in S, a \in A\}.
$$

For any $Q = (q_{ia}()) \in \mathcal{P}(S|A)$, A MDP is specified by $\{S, A, Q, r\}$, where $r : S \times A \to \mathbb{R}_+$ is an immediate reward function. Denote by $F$ the set of functions from $S$ to $A$. A policy $\pi$ is a sequence $(f_1, f_2, \ldots)$ of functions with $f_t \in F$ $(t \geq 1)$. Let $\Pi$ denote the class of policies. We denote by $f$ the policy $(f_1, f_2, \ldots)$ with $f_t = f$ for all $t \geq 1$ and some $f \in F$. Such a policy is called stationary.

We associate with each $f \in F$, $Q \in \mathcal{P}(S|A)$ the column vector $r(f) = (r(1(f(1)), \ldots, r(n(f(n))))'$ and the $n \times n$ transition matrix $Q(f)$, whose $(i, j)$ element is $q_{ia}(j)$ $(1 \leq i, j \leq n)$. Then, the expected total discounted reward from $\pi = (f_1, f_2, \ldots)$ is the column vector $\psi(\pi|Q) = (\psi(1|Q), \ldots, \psi(n|Q))'$, which is defined, as a function of $Q \in \mathcal{P}(S|A)$, by

$$
(3.1) \quad \psi(\pi|Q) = \sum_{t=0}^{\infty} \beta^t Q(f_1)Q(f_2) \cdots Q(f_t) r(f_{t+1}),
$$

where $0 < \beta < 1$ is a discount factor.

It is well-known (cf. [4]) that for each $Q \in \mathcal{P}(S|A)$, an optimal stationary policy $f^* = f^*(Q)$ exists with

$$
(3.2) \quad \psi(i, f^*|Q) = \sup_{\pi \in \Pi} f(i, \pi|Q) := \psi^*(i, Q).
$$

The perception $\widetilde{Q}_{ia}$ on $\mathcal{P}(S)$ is supposed to be given for each $i \in S, a \in A$. Then the perception $\overline{Q}$ on $\mathcal{P}(S|A)$ is defined by

$$
(3.3) \quad \overline{Q}(Q) = \min_{i \in S, a \in A} \widetilde{Q}_{ia}(q_{ia})(\cdot), \text{ where } Q = (q_{ia} : i \in S, a \in A) \in \mathcal{P}(S|A).
$$

The problem is to find the perception value $\overline{\psi}(\cdot)$, where

$$
(3.4) \quad \overline{\psi}(i)(x) = \sup_{Q \in \mathcal{P}(S|A), \ x = \psi^*(i, Q)} \overline{Q}(Q).
$$

**Theorem 3.1** The following holds:

(i) $\overline{\psi}(i) \in \mathbb{R}$.

(ii) $\overline{\psi}(i) \ (i \in S)$ is given as a unique solution of the fuzzy optimality equation:

$$
(3.5) \quad \overline{\psi}(i) = \max \{1_{\{(\tau(i), a)\}} + \beta \overline{Q}_{ia}, \overline{\psi}\},
$$

where $\overline{Q} \cdot \overline{\psi}(x) = \sup \overline{Q}_{ia}(q) \wedge \overline{\psi}(\cdot)$ and the supremum is taken on the range $\{(q, \psi) \in \mathcal{P}(S) \times \mathbb{R}^n \mid x = \sum_{j=1}^{n} q(j)\psi_j \}$ and $\overline{\psi}(\cdot) = \psi(1(\psi_1) \wedge \cdots \wedge \psi(n)\psi_n)$ with $\psi = (\psi_1, \ldots, \psi_n) \in \mathbb{R}^n$.

**References**


