1998年度日本オペレーションズ・リサーチ学会
春季研究発表会

ファジィ環境下の事前・事後条件付き意思決定

01003676 九州大学 岩本誠一 IWAMOTO Seiichi
02501655 九州工業大学 藤田敏治 FUJITA Toshiharu
02302176 九州大学 津留崎和義 TSURUSAKI Kazuyoshi

1 はじめに

マルコフ決定過程上において結合型評価の期待値最大化問題を考える。多段決定過程の期待値としてはいわゆる期待値のほか、あらたに事前、事後の二つの条件付き期待値をそれぞれ導入する。Bellman & Zadeh([1]) の「ファジィ環境における確率的意思決定」における再帰式は通常の期待値最適化という意味では理論的整合性を欠いているが ([2])、実は事後条件付き決定過程の再帰式に他ならないことを示す。すなわち、Bellman & Zadeh の「決定過程」は事後条件付き決定過程であることが分かる。

2 決定過程

本節では、不確実性の下で結合型評価の決定過程を考える ([3])。なお、以後全体を通して次のデータが与えられているものとする：

$N \geq 2$	終端時刻
$X = \{s_1, \cdots, s_p\}$	状態集合
$U = \{a_1, \cdots, a_k\}$	決定集合
$x_n \in X$	時刻 n における状態
$u_n \in U$	時刻 n における決定
$\mu_n : X \times U \to [0,1]$	時刻 n における利得
$\mu_{N+1} : X \to [0,1]$	終端利得
p	マルコフ推移法則 $p(y\|x,u) \geq 0$, $\sum_{y \in X} p(y\|x,u) = 1$
$\circ : [0,1] \times [0,1] \to [0,1]$	左単位元 ι をもつ結合型二項演算子 $\lambda \circ (\mu \circ \nu) = (\lambda \circ \mu) \circ \nu$

このとき、次の問題を考える：

$$
\begin{aligned}
&\text{Maximize } E^{\sigma}_{x_1}[\,\mu_1 \circ \mu_2 \circ \cdots \circ \mu_N \circ \mu_{N+1}] \\
&\text{subject to (i)}_{\mathrm{n}} \quad x_{n+1} \sim p(\cdot \mid x_n, u_n) \\
&\qquad\qquad\qquad u_n \in U \quad 1 \leq n \leq N
\end{aligned}
$$

ただし $\mu_t = \mu_t(x_t, u_t)$, $\mu_{N+1} = \mu_{N+1}(x_{N+1})$ で、$y \sim p(\cdot \mid x, u)$ は現時刻の状態が x、決定が u であるとき、次の時刻で状態 y へ確率 $p(y|x,u)$ で推移することをあらわす。また $E^{\sigma}_{x_1}$ は条件付き確率 $p(x_{n+1} \mid x_n, u_n)$、政策 $\sigma = \{\sigma_1, \sigma_2, ..., \sigma_N\}$ 及び初期状態 $x_1 \in X$ に依存して定まる $X \times U \times X \times U \times \cdots \times U \times X$ 上の期待値を表す：

$$
\begin{aligned}
&E^{\sigma}_{x_1}[\,\mu_1 \circ \cdots \circ \mu_{N+1} \mid \text{(i)}_{\mathrm{n}} \ \ 1 \leq n \leq N] \\
&= \sum_{(x_2,\ldots,x_{N+1}) \in X \times \cdots \times X} \cdots \sum [\mu_1(x_1, u_1) \circ \cdots \circ \mu_{N+1}(x_{N+1})] \\
&\qquad \times p(x_2|x_1,u_1) \cdots p(x_{N+1}|x_N,u_N)
\end{aligned}
$$

この決定過程の再帰式は不変埋没原理によって導かれる ([2] 参照)。すなわち、任意の $n = 1, 2, \ldots, N+1$ 及び $x_n \in X$ に対し、新たにパラメーター $\lambda_n \in [0,1]$ を導入した部分問題：

$$
\begin{aligned}
v_n(x_n; \lambda_n) &:= \operatorname*{Max}_{\pi} \ E[\,\lambda_n \circ \mu_n \circ \cdots \circ \mu_{N+1} \\
&\qquad\qquad \mid \text{(i)}_{\mathrm{m}} \ \ n \leq m \leq N] \\
v_{N+1}(x_{N+1}; \lambda_{N+1}) &:= \lambda_{N+1} \circ \mu_{N+1}(x_{N+1})
\end{aligned}
$$

を考えると、次の再帰式を得る。

定理 2.1

$$
\begin{aligned}
v_n(x; \lambda) &= \operatorname*{Max}_{u \in U} \sum_{y \in X} v_{n+1}(y; \lambda \circ \mu_n(x,u)) p(y|x,u) \\
&\qquad x \in X, \quad \lambda \in [0,1] \quad n = 1, 2, \ldots, N \\
v_{N+1}(x; \lambda) &= \lambda \circ \mu_{N+1}(x) \qquad x \in X, \quad 0 \leq \lambda \leq 1.
\end{aligned}
$$

したがって、与問題の最適値は $v_0(x_0; \iota)$ で与える。

3 条件付き決定過程

前節で述べた通常の決定過程に対し、本節では二つの条件付き決定過程を提案する。なお本節を通して、二項演算子 $\circ$ は単調であること：

$$
\mu < \nu \implies \lambda \circ \mu \leq \lambda \circ \nu
$$

を仮定する。

3.1 事後条件付き決定過程

事後条件付き期待値を評価とする次の問題を考える：

$$
\begin{aligned}
&\text{Max } \mu_1 \circ E^{u_1}_{x_1}[\mu_2 \circ \cdots \circ E^{u_{N-1}}_{x_{N-1}}[\mu_N \circ E^{u_N}_{x_N} \mu_{N+1}] \cdots] \\
&\quad \text{subject to (i)}_{\mathrm{n}} \ \ x_{n+1} \sim p(\cdot \mid x_n, u_n) \\
&\qquad\qquad\qquad u_n \in U \quad 1 \leq n \leq N.
\end{aligned}
$$

ここに、 $E_x^u \mu = \sum_{y \in X} \mu(y) p(y \mid x, u).$

したがって、ここでの期待値は次のような後ろ向き反復和になる：

$$\begin{aligned}
&\mu_1 \circ E^1[\mu_2 \circ \cdots \circ E^{N-1}[\mu_N \circ E^N \mu_{N+1}] \cdots] \\
= \;& \mu_1(x_1, u_1) \circ \sum_{x_2 \in X} [\mu_2(x_2, u_2) \circ \cdots \\
&\circ \sum_{x_N \in X} [\mu_N(x_N, u_N) \circ \sum_{x_{N+1} \in X} \mu_{N+1}(x_{N+1}) \\
&\quad p(x_{N+1}|x_N, u_N)] p(x_N|x_{N-1}, u_{N-1})] \cdots \\
&\quad] p(x_2|x_1, u_1)
\end{aligned}$$

$$(\text{ただし}, E^n = E_{x_n}^{u_n})$$

この問題に対し、任意の $n = 1, 2, \ldots, N+1$ 及び $x_n \in X$ について部分問題：

$$\begin{aligned}
w_n(x_n) &:= \operatorname*{Max}_{\pi} \; [\mu_n \circ E^n[\mu_{n+1} \circ \cdots \circ E^{N-1}[\mu_N \circ \\
&\qquad E^N \mu_{N+1}] \cdots] \mid (\mathrm{i})_{\mathrm{m}} \quad n \le m \le N] \\
w_{N+1}(x_{N+1}) &:= \mu_{N+1}(x_{N+1})
\end{aligned}$$

を考えることにより、次の再帰式が導かれる。

定理 3.1

$$\begin{aligned}
w_n(x) &= \operatorname*{Max}_{u \in U} [\mu_n(x, u) \circ \sum_{y \in X} w_{n+1}(y) p(y|x, u)] \\
&\qquad x \in X, \quad n = 1, 2, \ldots, N \\
w_{N+1}(x) &= \mu_{N+1}(x) \qquad x \in X.
\end{aligned}$$

この再帰式において特に $\circ = \wedge$ の場合が、Bellman & Zadeh が [1] が導いた再帰式に他ならない。すなわち、彼らの再帰式を導く問題は、いわゆる決定過程に対してではなく、この事後条件付き決定過程に対してである。

3.2 事前条件付き決定過程

事前条件付き期待値を評価とする次の問題を考える：

$$\begin{aligned}
&\operatorname{Max} \; E_{x_1}^{u_1}[\mu_1 \circ E_{x_2}^{u_2}[\mu_2 \circ \cdots \circ E_{x_N}^{u_N}[\mu_N \circ \mu_{N+1}] \cdots] \\
&\quad \text{subject to } (\mathrm{i})_{\mathrm{n}} \quad x_{n+1} \sim p(\cdot \mid x_n, u_n) \\
&\qquad\qquad\qquad u_n \in U \quad 1 \le n \le N.
\end{aligned}$$

ここに、 $E_x^u[\mu_n(x, u) \circ \mu]$
$= \sum_{y \in X} [\mu_n(x, u) \circ \mu(y)] p(y \mid x, u).$

したがって、この期待値は次の後ろ向き反復和になる：

$$\begin{aligned}
&E^1[\mu_1 \circ \cdots \circ E^N[\mu_N \circ \mu_{N+1}] \cdots] \\
= \;& \sum_{x_2 \in X} [\mu_1(x_1, u_1) \circ \cdots \\
&\circ \sum_{x_{N+1} \in X} [\mu_N(x_N, u_N) \circ \mu_{N+1}(x_{N+1})] \\
&\quad p(x_{N+1}|x_N, u_N) \cdots] p(x_2|x_1, u_1)
\end{aligned}$$

$$(\text{ただし}, E^n[\mu_n \circ \mu] = E_{x_n}^{u_n}[\mu_n(x_n, u_n) \circ \mu])$$

この問題に対し、任意の $n = 1, 2, \ldots, N+1$ 及び $x_n \in X$ について部分問題：

$$\begin{aligned}
W_n(x_n) &:= \operatorname*{Max}_{\pi} \; [E^n[\mu_n \circ E^{n+1}[\mu_{n+1} \circ \cdots \\
&\qquad \circ E^N[\mu_N \circ \mu_{N+1}] \cdots]] \mid (\mathrm{i})_{\mathrm{m}} \quad n \le m \le N] \\
W_{N+1}(x_{N+1}) &:= \mu_{N+1}(x_{N+1})
\end{aligned}$$

を考える。このとき、次の再帰式が成り立つ。

定理 3.2

$$\begin{aligned}
W_n(x) &= \operatorname*{Max}_{u \in U} \sum_{y \in X} [\mu_n(x, u) \circ W_{n+1}(y)] p(y|x, u) \\
&\qquad x \in X, \quad n = 1, 2, \ldots, N \\
W_{N+1}(x) &= \mu_{N+1}(x) \qquad x \in X.
\end{aligned}$$

4 3・2・2モデル

数値例として、Bellman & Zadeh([1]) の3状態2決定2段問題を考え、三つの決定過程の最適構造を明らかにする。

【各段評価】

$\mu_3(s_1) = 0.3 \quad \mu_3(s_2) = 1.0 \quad \mu_3(s_3) = 0.8$

$\mu_2(a_1) = 1.0 \quad \mu_2(a_2) = 0.6$

$\mu_1(a_1) = 0.7 \quad \mu_1(a_2) = 1.0$

【推移確率】 $p(x_{t+1}|x_t, u_t), \ t = 1, 2$

$u_t = a_1$

$x_t \backslash x_{t+1}$	s_1	s_2	s_3
s_1	0.8	0.1	0.1
s_2	0.0	0.1	0.9
s_3	0.8	0.1	0.1

$u_t = a_2$

$x_t \backslash x_{t+1}$	s_1	s_2	s_3
s_1	0.1	0.9	0.0
s_2	0.8	0.1	0.1
s_3	0.1	0.0	0.9

References

[1] Bellman, R.E. and Zadeh, L.A.: *Decision-making in a fuzzy environment*, *Management Science*, **17**, 1970

[2] Iwamoto, S. and Fujita, T.: *Stochastic decision-making in a fuzzy environment*, *J. Operations Res. Soc. Japan*, **38**, 1995

[3] Iwamoto, S.: *Associative dynamic programs*, *J. Math. Anal. Appl.*, **201**, 1996