数理解析研究所講究録
1015 巻 1997 年 150-162

ファジィ環境下における条件付き意思決定について

九大・経済　岩本誠一
九工大・工　藤田敏治
九大・経済　津留崎和義

1　はじめに

マルコフ決定過程において結合型評価の期待値最大化問題を考える。期待値としては通常の期待値のほか、二つの条件付き期待値 — 事後条件付き期待値と事前条件付き期待値を考える。これらの期待値を評価とする決定過程をそれぞれ事後条件付き決定過程および事前条件付き決定過程と呼ぶ。事後条件付き決定過程は、Bellman と Zadeh が [1] でファジィ環境下における確率的意思決定過程に対し与えた再帰式に関する逆問題に解を与える。というのも、彼らの導いた再帰式は本来この論文における通常の決定過程に対するものであったが、実際は不正確なものだったことを我々は [2] で示しており、その際に「では、彼らの再帰式が意味するところは？」という疑問が残っていた。そして今回、彼らの再帰式が本論文における事後条件付き決定過程に対するものである、という結論を得たのである。

2　決定過程

本節では通常の期待値利得を考えた場合のいわゆる（通常の）決定過程問題を考える。なお、以後全節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 $(n = 1, 2, \ldots, N+1)$
$u_n \in U$	時刻 n における決定 $(n = 1, 2, \ldots, N)$
$\mu_n : X \times U \to [0, 1]$	時刻 n における利得 $(n = 1, 2, \ldots, N)$
$\mu_{N+1} : X \to [0, 1]$	終端利得
p	マルコフ推移法則　$p(y\|x, u) \geq 0 \ \ {}^\forall(x, u, y) \in X \times U \times X$ $\sum_{y \in X} p(y\|x, u) = 1 \ \ {}^\forall(x, u) \in X \times U$
$\circ : [0, 1] \times [0, 1] \to [0, 1]$	左単位元 ι をもつ結合型二項演算子 $\lambda \circ (\mu \circ \nu) = (\lambda \circ \mu) \circ \nu \ \ {}^\forall(\lambda, \mu, \nu) \in [0, 1]^3, \ \ \iota \circ \lambda = \lambda \ \ {}^\forall\lambda \in [0, 1]$

このとき、次の問題を考える。

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

ただし $\mu_t = \mu_t(x_t, u_t)$, $\mu_{N+1} = \mu_{N+1}(x_{N+1})$ で、$y \sim p(\cdot\,|x, u)$ 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 $X \times X \times \cdots \times X$ (N-times) 上の期待値を表す。具体的には次のような n 重和になる：

$$\begin{aligned} & E^\sigma_{x_1}[\,\mu_1 \circ \mu_2 \circ \cdots \circ \mu_{N+1} \mid \text{(i)}_{\text{n}} \ \ 1 \leq n \leq N\,] \\ = & \sum_{(x_2, \ldots, x_{N+1}) \in X \times \cdots \times X} \!\!\!\cdots \sum [\mu_1(x_1, u_1) \circ \mu_2(x_2, u_2) \circ \cdots \circ \mu_{N+1}(x_{N+1})] \times p(x_2|x_1, u_1) \cdots p(x_{N+1}|x_N, u_N) \end{aligned}$$

問題 (1) に対し再帰式を導く際は、不変埋没原理の手法を用いなければならない（[2],[3] 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 $\lambda_n \in [0,1]$ を導入した次の部分問題群：

$$v_n(x_n; \lambda_n) = \operatorname*{Max}_{\pi} E^{\pi}_{x_n,\lambda_n}[\lambda_n \circ \mu_n \circ \cdots \circ \mu_N \circ \mu_{N+1} | \text{(i)}_m \quad n \leq m \leq N] \qquad 1 \leq n \leq N \tag{2}$$

$$v_{N+1}(x_{N+1}; \lambda_{N+1}) = \lambda_{N+1} \circ \mu_{N+1}(x_{N+1}) \tag{3}$$

を考える。これらの部分過程は、状態空間が 1 次元拡大されたもので $(x_0, \lambda_0) \in X \times [0,1]$ を初期状態とし、(x_{N+1}, λ_{N+1}) で終了する。そして、最大化は部分過程に対するすべてのマルコフ政策 π に関して取られ、そのマルコフ政策 $\pi = \{\pi_n, \pi_{n+1}, \ldots, \pi_N\}$ は決定関数列：

$$\pi_m : X \times [0,1] \to U \quad n \leq m \leq N \tag{4}$$

からなる。なお (2) における期待値は次のように定義される。

$$\begin{aligned} & E^{\pi}_{x_n,\lambda_n}[\lambda_n \circ \mu_n \circ \cdots \circ \mu_N \circ \mu_{N+1} | \text{(i)}_m \quad n \leq m \leq N] \\ = & \sum_{(x_{n+1},\ldots,x_{N+1}) \in X \times \cdots \times X} \sum \cdots \sum \{[\lambda_n \circ \mu_n(x_n, u_n) \circ \cdots \circ \mu_N(x_N, u_N) \circ \mu_{N+1}(x_{N+1})] \\ & \qquad \times p(x_{n+1}|x_n, u_n) \cdots p(x_{N+1}|x_N, u_N)\} \end{aligned}$$

ここで、決定と拡大された状態の交互列：

$$\{u_n,\ (x_{n+1}, \lambda_{n+1}),\ u_{n+1},\ (x_{n+2}, \lambda_{n+2}),\ \ldots,\ u_N,\ (x_{N+1}, \lambda_{N+1})\}$$

はマルコフ政策 π と初期状態 (x_n, λ_n) により次のように確率的に生成される。

$$\begin{aligned} & \pi_n(x_n, \lambda_n) = u_n \ \to \left\{ \begin{array}{l} p(\cdot\,|x_n, u_n) \sim x_{n+1} \\ \lambda_n \circ \mu_n(x_n, u_n) = \lambda_{n+1} \end{array} \right. \\ \to\ & \pi_{n+1}(x_{n+1}, \lambda_{n+1}) = u_{n+1} \ \to \left\{ \begin{array}{l} p(\cdot\,|x_{n+1}, u_{n+1}) \sim x_{n+2} \\ \lambda_{n+1} \circ \mu_{n+1}(x_{n+1}, u_{n+1}) = \lambda_{n+2} \end{array} \right. \\ \to\ & \quad \cdots \\ \to\ & \pi_N(x_N, \lambda_N) = u_N \ \to \left\{ \begin{array}{l} p(\cdot\,|x_N, u_N) \sim x_{N+1} \\ \lambda_N \circ \mu_N(x_N, u_N) = \lambda_{N+1} \end{array} \right. \end{aligned}$$

この時、値 $v_n(x; \lambda)$ と 2 変数関数 $v_{n+1}(\cdot\,;\cdot)$ の間に次の再帰式を得る。

定理 2.1

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

証明： [3] の Theorem 6.2 において 2 項関係 $\wedge$ を $\circ$ に置き換えることにより、同様な議論でこの定理は示される。 □

定理 2.1の再帰式を解くことにより、拡大状態空間上の最適値関数 $v_1(x_1; \lambda_1)$ および最適マル

コフ政策 $\pi^* = \{\pi_1^*, \pi_2^*, ..., \pi_N^*\}$ を得る。そしてこのとき ι が演算 $\circ$ に関する単位元であることから、

$$v_1(x_1; \iota)$$

が与問題 (1) の最適値を与えることが分かる。さらに、初期状態を (x_1, ι) とし π^* を元の状態空間 $X \times X \times \cdots \times X$ (N-times) に射影することにより、与問題 (1) に対する最適一般政策

$$\sigma^* = \{\sigma_1^*, \sigma_2^*, ..., \sigma_N^*\}$$

を得る。

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。一つは事後条件付き決定過程、もう一つは事前条件付き決定過程である。その際、マルコフ政策 $\pi = \{\pi_1, \pi_2, ..., \pi_N\}$ は状態空間 X 上で考える。すなわち各決定関数は

$$\pi_n : X \to U \quad 1 \le n \le N.$$

で与えられる。なお、本節を通して、二項演算子 $\circ$ は単調であること：

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

を仮定する。

3.1 事後条件付き決定過程

事後条件付き決定過程における期待値は、各段において、決定を行った後に残りの過程に対しそれぞれ取られる（図 1)。この期待値を事後条件付き期待値と呼び、事後条件付き期待値を評価とする次の事後条件付き決定過程問題を考える。

$$\begin{array}{ll} \text{Maximize} & \mu_1(x_1, u_1) \circ E_{x_1}^{u_1}[\mu_2(x_2, u_2) \circ \cdots \circ E_{x_{N-1}}^{u_{N-1}}[\mu_N(x_N, u_N) \circ E_{x_N}^{u_N}\mu_{N+1}] \cdots] \\ \text{subject to} & (\mathrm{i})_n \;\; x_{n+1} \sim p(\cdot | x_n, u_n), \; u_n \in U \quad 1 \le n \le N \end{array} \tag{5}$$

ただし、

$$E_x^u \mu = \sum_{y \in X} \mu(y) p(y|x, u) \quad \text{for } \mu = \mu(\cdot)$$

である。以後、簡単のため次の簡略化された記号を用いる。

$$\begin{aligned} E^n \mu &:= E_{x_n}^{u_n} \mu \\ \mu_n \circ E^n \mu &:= \mu_n(x_n, u_n) \circ E^n \mu \quad 1 \le n \le N \end{aligned}$$

このとき、問題 (5) における目的関数は次のように表わされる。

$$\begin{aligned} & \mu_1 \circ E^1[\mu_2 \circ \cdots \circ E^{N-1}[\mu_N \circ E^N \mu_{N+1}] \cdots] \\ := \; & \mu_1(x_1, u_1) \circ E_{x_1}^{u_1}[\mu_2(x_2, u_2) \circ \cdots \circ E_{x_{N-1}}^{u_{N-1}}[\mu_N(x_N, u_N) \circ E_{x_N}^{u_N}\mu_{N+1}] \cdots] \end{aligned} \tag{6}$$

ここで注意しておきたいのは、マルコフ政策 π が記号 E^n の中に陰に含まれているという点である：

$$E^n \mu = E_{x_n}^{u_n} \mu, \quad u_n = \pi_n(x_n) \quad 1 \le n \le N.$$

そうして、マルコフ政策 π に依存して定まる事後条件付き期待値は一つの反復和：

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

を構成するのである。一方で、いわゆる通常の期待値は多重和：

$$
\begin{aligned}
& E^{\pi}_{x_1}[\mu_1 \circ [\mu_2 \circ \cdots \circ [\mu_N \circ \mu_{N+1}] \cdots]] \\
= \ & \sum \sum \cdots \sum_{(x_2, \ldots, x_{N+1}) \in X \times \cdots \times X} \{[\mu_1(x_1, u_1) \circ [\mu_2(x_2, u_2) \circ \cdots \circ [\mu_N(x_N, u_N) \circ \mu_{N+1}(x_{N+1})] \cdots]] \\
& \qquad \times p(x_2|x_1, u_1) p(x_3|x_2, u_2) \cdots p(x_{N+1}|x_N, u_N)\} \qquad (7) \\
& \qquad (u_n = \pi_n(x_n) \quad 1 \leq n \leq N).
\end{aligned}
$$

である。一般に

$$
\begin{aligned}
& E^{\pi}_{x_1}[\mu_1 \circ [\mu_2 \circ \cdots \circ [\mu_N \circ \mu_{N+1}] \cdots]] \\
= \ & \mu_1 \circ E^1[\mu_2 \circ \cdots \circ E^{N-1}[\mu_N \circ E^N \mu_{N+1}] \cdots] \qquad (8)
\end{aligned}
$$

は成り立たない。しかし、二つの典型的な過程では (8) が成り立つ。加法型 ($\circ = +$) の場合、と乗法型 ($\circ = \times$) の場合である。当然ながら我々の関心は、(8) が成り立たない場合にある。

では、問題 (5) に対し再帰式を導く。まず、任意の $n = 1, 2, \ldots, N+1$ 及び $x_n \in X$ について部分問題：

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

を考える。Max はすべてのマルコフ政策 $\pi = \{\pi_n, ..., \pi_N\}$ に関して取られる。このとき，次の再帰式が導かれる。

定理 3.1

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

証明： 2項演算子 $\circ$ の単調性より、

$$\begin{aligned} & \operatorname*{Max}_{\pi}[\mu_1 \circ E^1[\mu_2 \circ \cdots \circ E^{N-1}[\mu_N \circ E^N \mu_{N+1}]\cdots]] \\ = & \operatorname*{Max}_{\pi_1}[\mu_1 \circ E^1 \operatorname*{Max}_{\pi_2}[\mu_2 \circ \cdots \circ E^{N-1} \operatorname*{Max}_{\pi_N}[\mu_N \circ E^N \mu_{N+1}]\cdots]] \\ & \qquad\qquad (u_n = \pi_n(x_n) \quad 1 \le n \le N) \end{aligned}$$

が成り立つことより明らか。 □

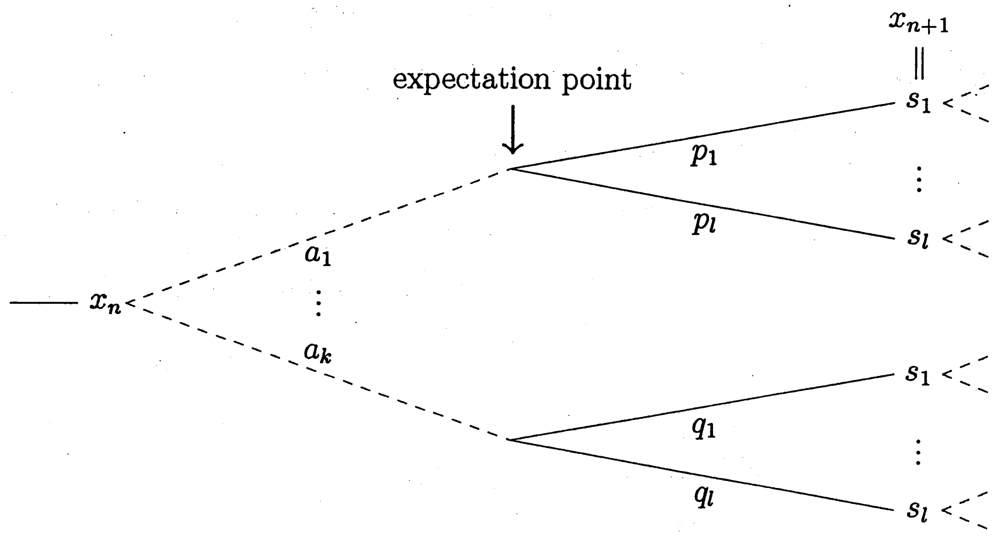


図 1 : Conditional expectation after take-action

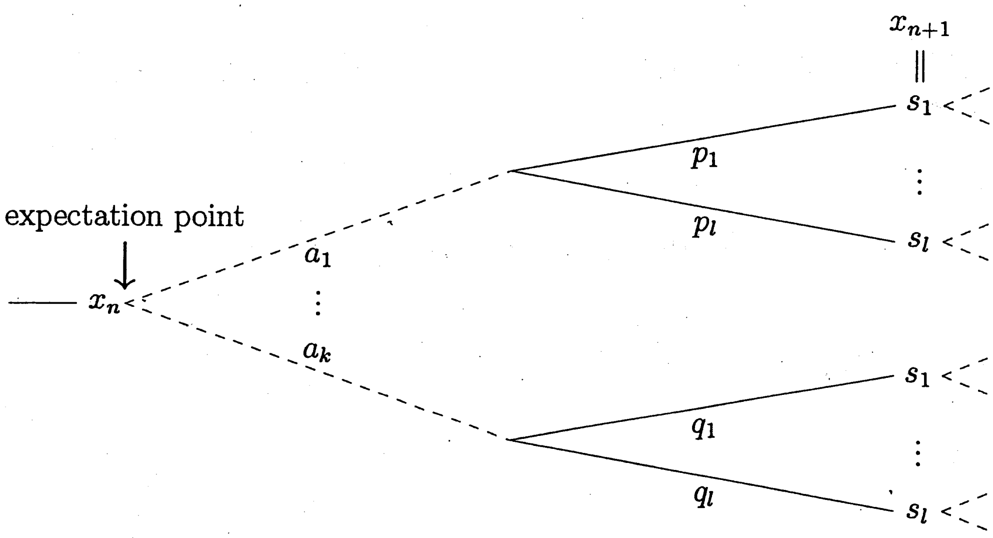


図 2 : Conditional expectation before take-action

3.2 事前条件付き決定過程

事前条件付き決定過程における期待値は、各段において、決定を行う前の時点で以後の過程に対しそれぞれ取られる（図 2)。この期待値を事前条件付き期待値と呼び、事前条件付き期待値を評価とする次の問題を考える。

$$\begin{aligned} &\text{Maximize} \quad E^{u_1}_{x_1}[\mu_1(x_1,u_1)\circ E^{u_2}_{x_2}[\mu_2(x_2,u_2)\circ\cdots \\ &\qquad\qquad\qquad\qquad \circ E^{u_N}_{x_N}[\mu_N(x_N,u_N)\circ\mu_{N+1}]\cdots]] \\ &\text{subject to} \quad (\mathrm{i})_\mathrm{n} \quad x_{n+1}\sim p(\cdot\,|\,x_n,u_n),\ u_n\in U \quad 1\le n\le N \end{aligned} \tag{9}$$

ただし、

$$E^u_x[\mu_n(x,u)\circ\mu]=\sum_{y\in X}[\mu_n(x,u)\circ\mu(y)]p(y|x,u) \quad \text{for } \mu=\mu(\cdot)$$

である。以後、簡単のため次の簡略化された記号を用いる。

$$E^n[\mu_n\circ\mu]:=E^{u_n}_{x_n}[\mu_n(x_n,u_n)\circ\mu] \quad 1\le n\le N$$

このとき、問題 (9) における目的関数は次のように表される。

$$\begin{aligned} &E^1[\mu_1\circ E^2[\mu_2\circ\cdots\circ E^N[\mu_N\circ\mu_{N+1}]\cdots]] \\ :=\ &E^{u_1}_{x_1}[\mu_1(x_1,u_1)\circ E^{u_2}_{x_2}[\mu_2(x_2,u_2)\circ\cdots\circ E^{u_N}_{x_N}[\mu_N(x_N,u_N)\circ\mu_{N+1}]\cdots]] \end{aligned} \tag{10}$$

上記において、やはりマルコフ政策 π が記号 E^n に包含されている。

$$E^n[\mu_n\circ\mu]=E^{u_n}_{x_n}[\mu_n(x_n,u_n)\circ\mu], \quad u_n=\pi_n(x_n) \quad 1\le n\le N$$

従って、事前条件付き期待値はもう一つの反復和を形成する。

$$\begin{aligned} &E^1[\mu_1\circ E^2[\mu_2\circ\cdots\circ E^N[\mu_N\circ\mu_{N+1}]\cdots]] \\ =\ &\sum_{x_2\in X}[\mu_1(x_1,u_1)\circ\sum_{x_3\in X}[\mu_2(x_2,u_2)\circ\cdots\circ\sum_{x_N\in X}[\mu_{N-1}(x_{N-1},u_{N-1}) \\ &\quad\circ\sum_{x_{N+1}\in X}[\mu_N(x_N,u_N)\circ\mu_{N+1}(x_{N+1})]p(x_{N+1}|x_N,u_N) \\ &\quad]p(x_N|x_{N-1},u_{N-1})\cdots p(x_3|x_2,u_2)]p(x_2|x_1,u_1)] \\ &\qquad (u_n=\pi_n(x_n) \quad 1\le n\le N) \end{aligned} \tag{11}$$

ここで、先に挙げた事後条件付き期待値 (7) とこの事前条件付き期待値を比較してみると、両者は必ずしも一致しない。当然、通常の期待値 (8) とも異なる。ただし、三つの期待値 (8), (7), (10) は加法型決定過程および乗法型決定過程においては一致する。

では、問題 (9) に対し再帰式を導く。任意の $n=1,2,\ldots,N+1$ 及び $x_n\in X$ について部分問題：

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

を考える。このとき、次の再帰式が成り立つ。

定理 3.2

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

証明： 等式

$$\begin{aligned} & \operatorname*{Max}_{\pi} E^1[\mu_1 \circ E^2[\mu_2 \circ \cdots \circ E^N[\mu_N \circ \mu_{N+1}] \cdots]] \\ = \; & \operatorname*{Max}_{\pi_1} E^1[\mu_1 \circ \operatorname*{Max}_{\pi_2} E^2[\mu_2 \circ \cdots \circ \operatorname*{Max}_{\pi_N} E^N[\mu_N \circ \mu_{N+1}] \cdots]] \end{aligned}$$

が成り立つことより明らか。 □

4 例題

数値例として次の 3 状態 2 決定の 2 段問題を扱う。2 項演算子としては最小型演算子 $\wedge$ （$a \wedge b := \min(a, b)$）を考える。なお、この数値は Bellman and Zadeh [1] の例題 (p. B154) の数値である。

$$\begin{gathered} \mu_3(s_1) = 0.3 \qquad \mu_3(s_2) = 1.0 \qquad \mu_3(s_3) = 0.8 \\ \mu_2(a_1) = 1.0 \qquad \mu_2(a_2) = 0.6 \\ \mu_1(a_1) = 0.7 \qquad \mu_1(a_2) = 1.0 \end{gathered} \tag{12}$$

$u_t = a_1$

$x_t \setminus x_{t+1}$	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

$u_t = a_2$

$x_t \setminus x_{t+1}$	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

(13)

4.1 通常決定過程

上記の数値例に対する通常の決定過程問題は次で与えられる。

$$\begin{array}{ll} \text{Maximize} & E[\, v_1(u_1) \wedge v_2(u_2) \wedge v_3(x_3)\,] \\ \text{subject to} & \text{(i)}_n \;\; x_{n+1} \sim p(\,\cdot\,|x_n, u_n) \quad u_n \in \{a_1, a_2\} \quad n = 1, 2 \end{array}$$

通常の決定過程に対する不変埋没原理によるアプローチは [2] において詳しく行っているので、ここでは結果のみを述べる。

まず、定理 2.1の再帰式を計算した結果を挙げる。

$$v_1(s_1; \lambda) = \begin{cases} \lambda & \text{for } 0 \le \lambda \le 0.3 \\ 0.99\lambda + 0.003 & \text{for } 0.3 \le \lambda \le 0.6 \\ 0.9\lambda + 0.057 & \text{for } 0.6 \le \lambda \le 0.8 \\ 0.09\lambda + 0.705 & \text{for } 0.8 \le \lambda \le 1 \end{cases}$$

$$v_1(s_2;\lambda)=\begin{cases}\lambda & \text{for } 0\le\lambda\le 0.3\\ 0.91\lambda+0.027 & \text{for } 0.3\le\lambda\le 0.6\\ 0.1\lambda+0.513 & \text{for } 0.6\le\lambda\le 0.7\\ 0.1\lambda+0.513 & \text{for } 0.7\le\lambda\le 0.8\\ 0.01\lambda+0.585 & \text{for } 0.8\le\lambda\le 1\end{cases}$$

$$v_1(s_3;\lambda)=\begin{cases}\lambda & \text{for } 0\le\lambda\le 0.3\\ 0.91\lambda+0.027 & \text{for } 0.3\le\lambda\le 0.6\\ 0.1\lambda+0.513 & \text{for } 0.6\le\lambda\le 0.7\\ 0.583 & \text{for } 0.7\le\lambda\le 1\end{cases}$$

これより、この場合の $\wedge$ に関する単位元 1 を λ に代入して、初期状態 s_1,s_2,s_3 に対する最適値：

$$v_1(s_1;1)=0.795 \quad v_1(s_2;1)=0.595 \quad v_1(s_3;1)=0.583$$

をそれぞれ得る。また最適政策 $\sigma^*=\{\sigma_1^*,\sigma_2^*\}$ は次で与えられる。

$$\sigma_1^*(s_1)=a_2,\quad \sigma_1^*(s_2)=a_2,\quad \sigma_1^*(s_3)=a_1$$

$$\begin{array}{lll}\sigma_2^*(s_1,s_1)=a_2, & \sigma_2^*(s_2,s_1)=a_2, & \sigma_2^*(s_3,s_1)=a_2\\ \sigma_2^*(s_1,s_2)=a_1, & \sigma_2^*(s_2,s_2)=a_1, & \sigma_2^*(s_3,s_2)=a_1\\ \sigma_2^*(s_1,s_3)=a_1,a_2 & \sigma_2^*(s_2,s_3)=a_2, & \sigma_2^*(s_3,s_3)=a_2\end{array}$$

なお、この最適政策はマルコフ政策としても与えられるが、一般に通常決定過程に対する最適政策は必ずしもマルコフ政策とは成り得ず（[3] 参照)、上記の形で与えられる。

4.2 事後条件付き決定過程

数値データ (12), (13) に対する事後条件付き決定過程問題は次で与えられる。

$$\begin{array}{ll}\text{Maximize} & [\mu_1(u_1)\wedge E_{x_1}^{u_1}[\mu_2(u_2)\wedge E_{x_2}^{u_2}\mu_3\,]\,]\\ \text{subject to} & (\mathrm{i})_{\mathrm{n}}\ \ x_{n+1}\sim p(\cdot|\,x_n,u_n),\ u_n\in\{a_1,a_2\}\quad n=1,2\end{array}\tag{14}$$

このとき、定理 3.1より再帰式は次で与えられる。

$$\begin{aligned}w_3(x_3) &= \mu_3(x_3)\\ w_2(x_2) &= \operatorname*{Max}_{u_2}[\,\mu_2(u_2)\wedge\sum_{x_3}w_3(x_3)p(x_3|x_2,u_2)\,]\\ w_1(x_1) &= \operatorname*{Max}_{u_1}[\,\mu_1(u_1)\wedge\sum_{x_2}w_2(x_2)p(x_2|x_1,u_1)\,]\end{aligned}\tag{15}$$

冒頭述べたように、Bellman と Zadeh は [1] においてファジィ環境下における確率的意思決定過程（本論文における通常決定過程）を扱っているが、実際に彼らの与えた再帰式は上記の再帰式であった。すなわち彼らの与えた再帰式は、事後条件付き決定過程問題に対するものだったのである。また、彼らは再帰式 (15) を計算し、次の結果を与えている。

$$w_2(s_1)=0.6,\quad w_2(s_2)=0.82,\quad w_2(s_3)=0.6$$

$$\pi_2(s_1)=a_1,\quad \pi_2(s_2)=a_1,\quad \pi_2(s_3)=a_2$$

$$w_1(s_1) = 0.8, \quad w_1(s_2) = 0.62, \quad w_1(s_3) = 0.62$$

$$\pi_1(s_1) = a_1, \quad \pi_1(s_2) = a_1 \text{ or } a_2, \quad \pi_1(s_3) = a_1$$

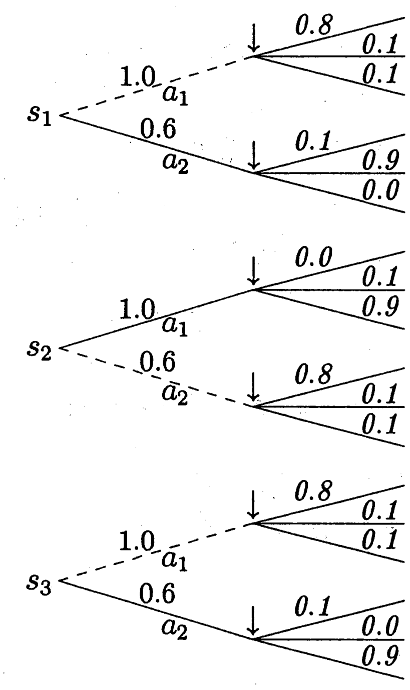
しかし、$w_1(x_1)$, $\pi_1(x_1)$ に関しては間違いで、正確には

$$w_1(s_1) = 0.798, \quad w_1(s_2) = 0.622, \quad w_1(s_3) = 0.622$$

$$\pi_1(s_1) = a_2, \quad \pi_1(s_2) = a_1 \text{ or } a_2, \quad \pi_1(s_3) = a_1$$

となることが、[2] で示されている。この結果は、図 3, 4, 5, 6（図 5, 6 は省略）によっても確認される。

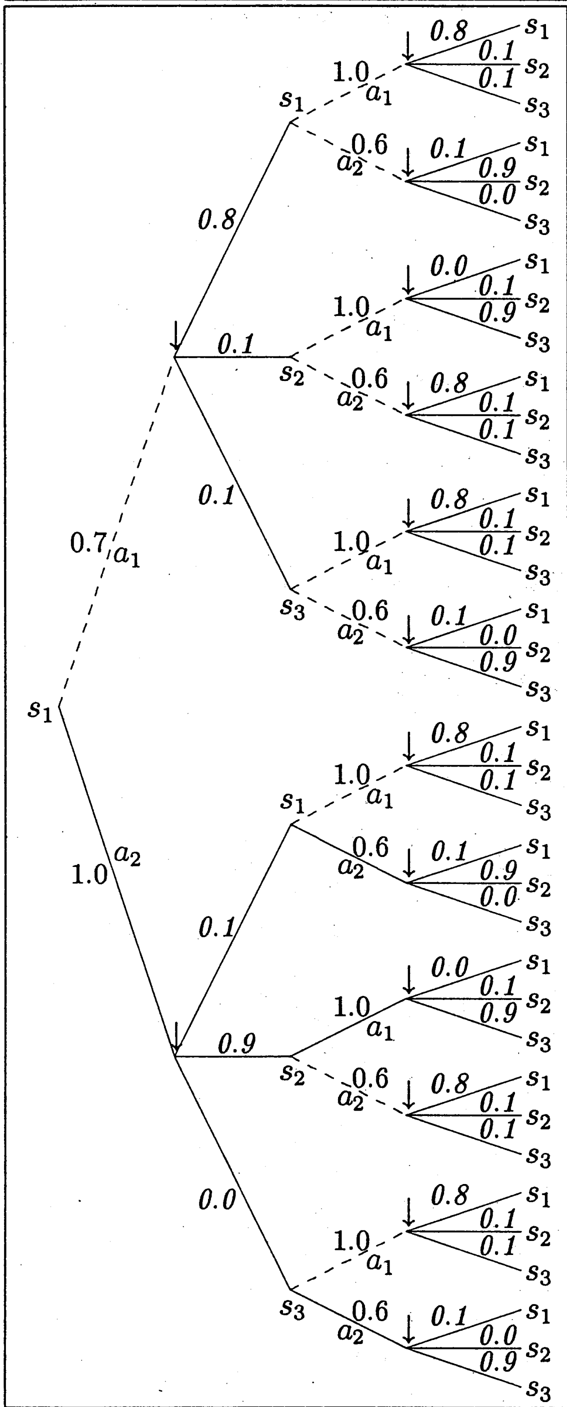
$$\operatorname*{Max}_{u_2}\left[\mu_2(u_2) \wedge \sum_{x_3} \mu_3(x_3) p(x_3|x_2, u_2)\right]$$



history	ter.	$\times p_1$	E^2	min
s_1	0.3	0.24		
s_2	1.0	0.1	0.42	0.42
s_3	0.8	0.08		
s_1	0.3	0.03		
s_2	1.0	0.9	0.93	**0.6**
s_3	0.8	0.0		
s_1	0.3	0.0		
s_2	1.0	0.1	0.82	**0.82**
s_3	0.8	0.72		
s_1	0.3	0.24		
s_2	1.0	0.1	0.42	0.42
s_3	0.8	0.08		
s_1	0.3	0.24		
s_2	1.0	0.1	0.42	0.42
s_3	0.8	0.08		
s_1	0.3	0.03		
s_2	1.0	0.0	0.75	**0.6**
s_3	0.8	0.72		

図 3：状態 s_1, s_2 および s_3 からの一段事後条件付き決定ツリー

$$\underset{u_1}{\mathrm{Max}}\left[\mu_1(u_1)\wedge\sum_{x_2}\underset{u_2}{\mathrm{Max}}\left\{\mu_2(u_2)\wedge\sum_{x_3}\mu_3(x_3)p(x_3|x_2,u_2)\right\}p(x_2|s_1,u_1)\right]$$



history	ter.	$\times p_2$	E^2	min	$\times p_1$	E^1	Min
	0.3	0.24					
	1.0	0.1	0.42	0.42			
	0.8	0.08					
	0.3	0.03			0.48		
	1.0	0.9	0.93	**0.6**			
	0.8	0.0					
	0.3	0.0					
	1.0	0.1	0.82	**0.82**			
	0.8	0.72					
	0.3	0.24			0.082	0.622	0.622
	1.0	0.1	0.42	0.42			
	0.8	0.08					
	0.3	0.24					
	1.0	0.1	0.42	0.42			
	0.8	0.08					
	0.3	0.03			0.06		
	1.0	0.0	0.75	**0.6**			
	0.8	0.72					
	0.3	0.24					
	1.0	0.1	0.42	0.42			
	0.8	0.08					
	0.3	0.03			0.06		
	1.0	0.9	0.93	**0.6**			
	0.8	0.0					
	0.3	0.0					
	1.0	0.1	0.82	**0.82**			
	0.8	0.72					
	0.3	0.24			0.738	0.798	**0.798**
	1.0	0.1	0.42	0.42			
	0.8	0.08					
	0.3	0.24					
	1.0	0.1	0.42	0.42			
	0.8	0.08					
	0.3	0.03			0.0		
	1.0	0.0	0.75	**0.6**			
	0.8	0.72					

図 4：状態 s_1 からの二段事後条件付き決定ツリー

4.3 事前条件付き決定過程

数値データ (12), (13) に対する事前条件付き決定過程問題は次で与えられる。

$$\begin{aligned} &\text{Maximize} \quad E^{u_1}_{x_1}[\mu_1(u_1) \wedge E^{u_2}_{x_2}[\mu_2(u_2) \wedge \mu_3\,]\,] \\ &\text{subject to} \quad (\mathrm{i})_{\mathrm{n}} \quad x_{n+1} \sim p(\cdot|\, x_n, u_n), \ u_n \in \{a_1, a_2\} \quad n = 1, 2 \end{aligned} \tag{16}$$

このとき、定理 3.2より再帰式は次で与えられる。

$$\begin{aligned} W_3(x_3) &= \mu_3(x_3) \\ W_2(x_2) &= \underset{u_2}{\mathrm{Max}} \sum_{x_3} [\,\mu_2(u_2) \wedge W_3(x_3)\,] p(x_3|x_2, u_2) \\ W_1(x_1) &= \underset{u_1}{\mathrm{Max}} \sum_{x_2} [\,\mu_1(u_1) \wedge W_2(x_2)\,] p(x_2|x_1, u_1) \end{aligned} \tag{17}$$

これらを計算することにより、最適値関数 W_3, W_2, W_1 及び最適政策 $\pi^* = \{\pi^*_2, \pi^*_1\}$ が得られる。

$$W_3(s_1) = 0.3, \quad W_3(s_2) = 1.0, \quad W_3(s_3) = 0.8$$

$$W_2(s_1) = 0.57, \quad W_2(s_2) = 0.82, \quad W_2(s_3) = 0.57$$

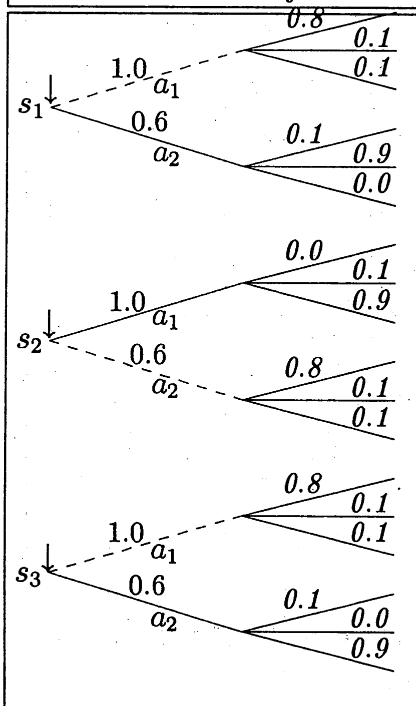
$$\pi^*_2(s_1) = a_2, \quad \pi^*_2(s_2) = a_1, \quad \pi^*_2(s_3) = a_2$$

$$W_1(s_1) = 0.795, \quad W_1(s_2) = 0.595, \quad W_1(s_3) = 0.583$$

$$\pi^*_1(s_1) = a_2, \quad \pi^*_1(s_2) = a_2, \quad \pi^*_1(s_3) = a_1$$

この解は図 7, 8, 9, 10（図 9, 10 は省略）によっても確認できる。

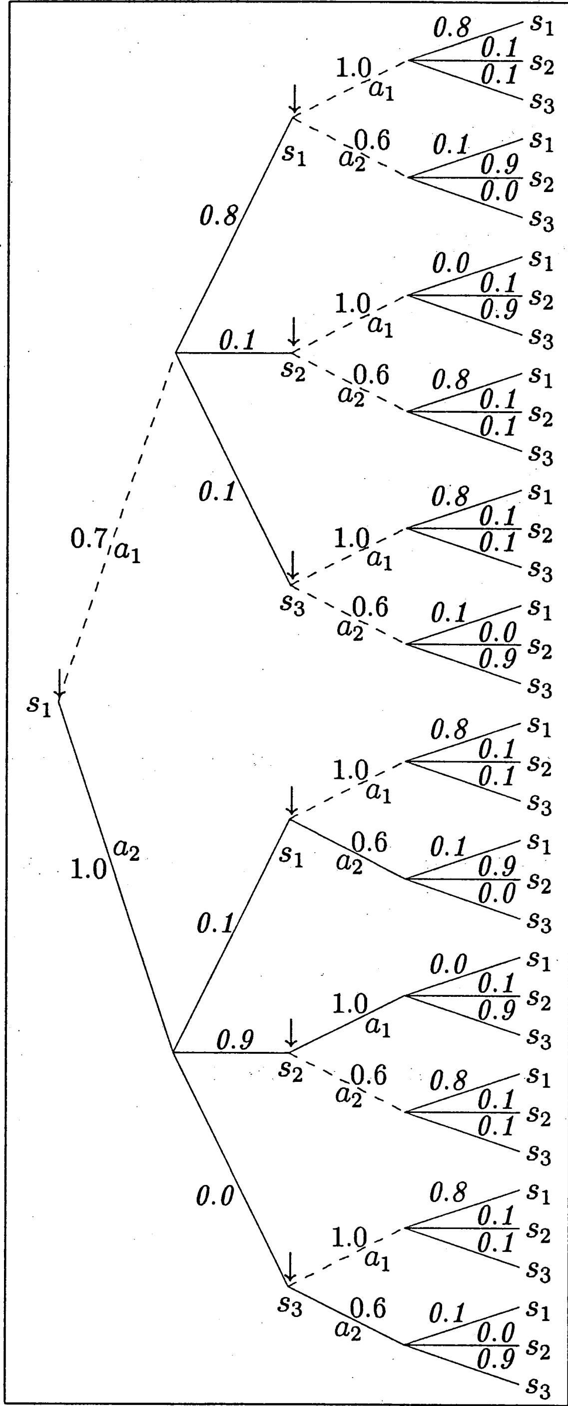
$$\underset{u_2}{\mathrm{Max}} \left[\sum_{x_3} \{\mu_2(u_2) \wedge \mu_3(x_3)\} p(x_3|x_2, u_2) \right]$$



history	ter.	min	$\times p$	E^2
s_1	0.3	0.3	0.24	
s_2	1.0	1.0	0.1	0.42
s_3	0.8	0.8	0.08	
s_1	0.3	0.3	0.03	
s_2	1.0	0.6	0.54	**0.57**
s_3	0.8	0.6	0.0	
s_1	0.3	0.3	0.0	
s_2	1.0	1.0	0.1	**0.82**
s_3	0.8	0.8	0.72	
s_1	0.3	0.3	0.24	
s_2	1.0	0.6	0.06	0.36
s_3	0.8	0.6	0.06	
s_1	0.3	0.3	0.24	
s_2	1.0	1.0	0.1	0.42
s_3	0.8	0.8	0.08	
s_1	0.3	0.3	0.03	
s_2	1.0	0.6	0.0	**0.57**
s_3	0.8	0.6	0.54	

図 7 : 状態 s_1, s_2 および s_3 からの一段事前条件付き決定ツリー

$$\mathop{\mathrm{Max}}_{u_1}\left[\sum_{x_2}\left\{\mu_1(u_1)\wedge \mathop{\mathrm{Max}}_{u_2}\sum_{x_3}\{\mu_2(u_2)\wedge\mu_3(x_3)\}p(x_3|x_2,u_2)\right\}p(x_2|s_1,u_1)\right]$$



history	ter.	min	$\times p_2$	E^2	min	$\times p_1$	E^1
	0.3	0.3	0.24				
	1.0	1.0	0.1	0.42			
	0.8	0.8	0.08				
	0.3	0.3	0.03		0.57	0.456	
	1.0	0.6	0.54	**0.57**			
	0.8	0.6	0.0				
	0.3	0.3	0.0				
	1.0	1.0	0.1	**0.82**			
	0.8	0.8	0.72				
	0.3	0.3	0.24		0.7	0.07	0.583
	1.0	0.6	0.06	0.36			
	0.8	0.6	0.06				
	0.3	0.3	0.24				
	1.0	1.0	0.1	0.42			
	0.8	0.8	0.08				
	0.3	0.3	0.03		0.57	0.057	
	1.0	0.6	0.0	**0.57**			
	0.8	0.6	0.54				
	0.3	0.3	0.24				
	1.0	1.0	0.1	0.42			
	0.8	0.8	0.08				
	0.3	0.3	0.03		0.57	0.057	
	1.0	0.6	0.54	**0.57**			
	0.8	0.6	0.0				
	0.3	0.3	0.0				
	1.0	1.0	0.1	**0.82**			
	0.8	0.8	0.72				
	0.3	0.3	0.24		0.82	0.738	**0.795**
	1.0	0.6	0.06	0.36			
	0.8	0.6	0.06				
	0.3	0.3	0.24				
	1.0	1.0	0.1	0.42			
	0.8	0.8	0.08				
	0.3	0.3	0.03		0.57	0.0	
	1.0	0.6	0.0	**0.57**			
	0.8	0.6	0.54				

図 8：状態 s_1 からの二段事前条件付き決定ツリー

5 さいごに

一般に、事後条件付き決定過程および事前条件付き決定過程に対する値関数 w_n, W_n に対し、2項演算子が $\circ = \wedge$ の場合

$$W_n(x) \le w_n(x) \quad 1 \le n \le N$$

という関係が成り立ち、$\circ = \vee$ の場合

$$W_n(x) \ge w_n(x) \quad 1 \le n \le N$$

という関係が成り立つ。

また、実定数 λ, 関数 $g : X \to R^1$ および確率関数 p に対し

$$\sum_{x \in X} [\lambda \circ g(x)] p(x) = \lambda \circ \sum_{x \in X} g(x) p(x) \quad \circ = +, \times$$

は成り立つが、任意の2項演算子 $\circ$ に関して

$$\sum_{x \in X} [\lambda \circ g(x)] p(x) = \lambda \circ \sum_{x \in X} g(x) p(x)$$

は必ずしも成立しない。

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., Tsurusaki, Y. and Fujita, T.:On Markov policies for minimax decision processes, under submission.

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