ファジィ動的計画法

岩本　誠一

1. はじめに

ファジィ動的計画法（fuzzy dynamic programming）についてはいくつかの考え方，定式化が考えられるが，その原型はベルマン・ザデー[2]が提案した『ファジィ環境下の意思決定』に遡る．ここでファジィ環境（fuzzy environment）とは制約とゴールのうち少なくとも一方がファジィ集合のときである．動的計画全体の視点から見れば，この制約はいわゆる制約式（条件式，状態推移など）の一つの雛型と考えられ，ゴールも評価式の一つの型を提示していると考えられる．ここにファジィ動的計画法を含む動的計画法などの再帰的方法の多様性が見られる[1,6,20]．

彼らの論文では多段決定過程における推移法則として(1)確定的（deterministic），(2)確率的（stochastic），(3)ファジィ（fuzzy）の三つのシステムを導入している．特に(1)確定的システムと(2)確率的システムについては動的計画法の再帰式を提示している．しかし，(3)ファジィシステムについては何も解析していない．(1)，(2)ではファジィ制約（fuzzy constraint）とファジィゴール（fuzzy goal）を同時に満たすという意味で最小型評価（minimum criterion）を考えている．(1)では動的計画法による「逐次解（sequential solution）」と本来の多変数問題の（列挙法などによる）「同時解（simultaneous solution）」は確かに一致している[12]．しかし，(2)では「逐次最適化≠同時最適化」になっている[12]．これは最小型評価関数の期待値最大化のためである．加法型評価（additive criterion）の期待値最大化では「逐次最適化＝同時最適化」である．非加法型評価（non-additive criterion）の期待値最大化においても「逐次最適解＝原ファジィ問題の最適解」を保証する一つの方法がファジィ動的計画法である[7〜9,11〜13,15,16]．

いわもと　せいいち
九州大学 大学院経済学研究院
〒812-8581 福岡市東区箱崎 6-19-1

以下，Bellman and Zadeh が提案した(2)確率的システム上の再帰式を中心にファジィ動的計画法の全体像を述べる．

2. 確定的ファジィ動的計画

まず，Bellman and Zadeh が提案した(1)確定的システム上のファジィ決定過程を考える．これは次の最小型評価系をもつ多段決定過程である：

$$\begin{aligned}&\text{Max.}\\ &\quad \mu_0(x_0, u_0)\wedge\cdots\wedge\mu_{N-1}(x_{N-1}, u_{N-1})\wedge\mu_G(x_N)\\ &\text{s.t.}\\ &\quad \text{(i)}\ \ x_{n+1}=f(x_n, u_n) \qquad n=0,\cdots,N-1\\ &\quad \text{(ii)}\ \ u_n\in U\end{aligned} \tag{1}$$

確定的システムでは状態 x_n で決定 u_n をとると，状態変換 $f: X\times U\to X$ によって次の状態 x_{n+1} が唯一 $x_{n+1}=f(x_n, u_n)$ で定まる．μ_n: $X\times U\to[0,1]$ は $X\times U$ 上のあるファジィ集合 R_n のメンバーシップ関数である：

$$\mu_n(x, u)=\mu_{R_n}(x, u).$$

μ_G: $X\to[0,1]$ は X 上のファジィ・ゴール（ファジィ集合）G のメンバーシップ関数である．多段決定過程ではシステム全体を通じてのファジィ共通集合 $R_0\cap R_1\cap\cdots\cap R_{N-1}\cap G$ のメンバーシップ関数

$$\begin{aligned}&\mu_{R_0\cap\cdots\cap R_{N-1}\cap G}(x_0, u_0, \cdots, x_{N-1}, u_{N-1}, x_N)\\ &\quad=\mu_0(x_0, u_0)\wedge\cdots\wedge\mu_{N-1}(x_{N-1}, u_{N-1})\\ &\qquad\wedge\mu_G(x_N)\end{aligned}$$

の最大化を次のように考える．一般に，第 n 段での決定の採り方は**マルコフ決定関数** π_n: $X\to U$ で記述される．この列 $\{\pi_0, \pi_1, \cdots, \pi_{N-1}\}$ を**マルコフ政策**という．問題(1)の最大化はマルコフ政策クラス（全体）上でとる．

このとき，任意のステージ $n(0\leq n\leq N)$ と状態 $x_n(\in X)$ からなる対 (n, x_n) を固定する．この対から始まって，ステージ N で終了する部分過程：

$$\begin{aligned}&\text{Max.}\\ &\quad \mu_n(x_n, u_n)\wedge\cdots\wedge\mu_{N-1}(x_{N-1}, u_{N-1})\wedge\mu_G(x_N)\end{aligned}$$

s.t. (2)

(i) $x_{m+1}=f(x_m, u_m)$
(ii) $u_m \in U$ $\quad m=n, \cdots, N-1$

の族$(x_n \in X, 0 \leq n \leq N)$を考える．特に，対$(0, x_0)$から始まる部分過程は，与えられた過程になっている．このとき，所与の過程は**埋め込まれた**という．対(n, x_n)からの始まる過程の最大値を$v_n(x_n)$とする．すなわち，$v_n(x_n)$は(n, x_n)からの始まる部分過程に対するマルコフ政策クラス上での最大値である．ただし，$v_N(x_N):=\mu_G(x_N)$．このとき，**最適値関数列**$\{v_0, v_1, \cdots, v_N\}$はいわゆる**後向きの再帰式**を満たす：

定理2.1 (Bellman and Zadeh[2])

$$v_n(x)=\max_{u \in U}[\mu_n(x, u) \wedge v_{n+1}(f(x, u))] \quad x \in X, 0 \leq n \leq N-1 \quad (3)$$
$$v_N(x)=\mu_G(x).$$

式(3)で最大値に到達する点（最大点）を$\pi_n^*(x)$とする．このとき，マルコフ決定関数$\pi_n^*: X \rightarrow U$からなる列$\{\pi_0^*, \pi_1^*, \cdots, \pi_{N-1}^*\}$は最適な**マルコフ政策**である．すなわち，再帰式を後向きに解くことによって，**マルコフ政策クラス**における最適政策が得られる．

これが確定系に対するファジィ動的計画法である．この動的計画法は，最小型評価系のみならず，単調性（単調非減少性）(monotonicity) と可分性 (separability) の下での確定的システムの最適化に対して広く成り立つ[1, 3〜5, 20]．

3. 確率的ファジィ動的計画法

一般に，リスク，リターン，メンバーシップなどが昨日・今日・明日といった時間的流れの中で変動しながら，集積する状況を考えよう．このような状況では，変動の展開は確率システムで表され，対象となる量は確率変数ととらえられる．そこでは，システム全体を通して集積する方法としては通常加法が考えられる．ヒト・モノ・カネなどは加法的に累積する．しかし，メンバーシップはシステム全体としては最小化演算で表現される．ここに，ファジィ動的計画法が直面する問題がある[2]．確率的システムは確定的システム上のダイナミックス$f: X \times U \rightarrow X$を一般化した状態推移確率$p=\{p(y|x, y)\}$で表される．ここでは，現在の状態が$x(\in X)$のとき，決定$u(\in U)$をとると，次期には確率$p(y|x, u)$で状態$y(\in X)$に推移する．確率的な展開においては，システム全体を通じてのファジィ共通集合$R_0 \cap R_1 \cap \cdots \cap R_{N-1} \cap G$のメンバーシップ関数

$$\mu_0 \wedge \mu_1 \wedge \cdots \wedge \mu_{N-1} \wedge \mu_G$$

は確率変数である．すなわち，メンバーシップ関数$\mu_n=\mu_n(X_n, U_n)$, $\mu_G=\mu_G(X_N)$は状態確率変数列$\{X_n\}$，決定確率変数列$\{U_n\}$を通じて定まる確率変数である．ここでは確率変動は始発の状態，マルコフ推移法則pおよび決定構造（政策）から定まる直積状態空間上の直積確率測度で規定される[12, 13, 17]．

4. 多段確率決定樹表

さて，確率系上でのファジィ動的計画法を次の2段確率決定樹表（図1-1, 図1-2）で説明しよう．**決定樹表** (decision tree-table) とは，多段決定過程の問題から計算過程を経て最適解が得られるまでを決定樹 (decision tree) と決定表 (decision table) を組み合わせて1枚の図表にまとめたものである[18, 19]．上段の決定樹は問題のデータ（メンバーシップと推移確率）と履歴を表している．下段の決定表は計算過程と各段での最適選択を表している．この2-2-2（2状態2決定2段）モデルのメンバーシップ関数列と推移行列は図2，図3で与えられている．

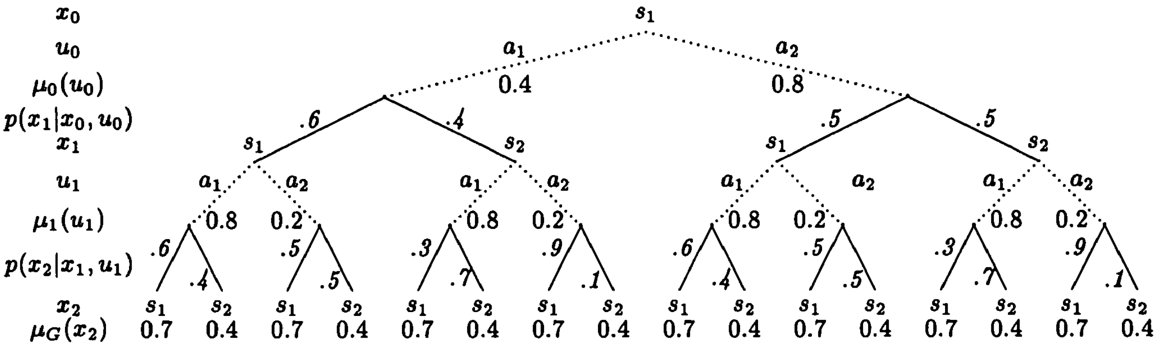
図1-1を考える．状態x_nと決定u_nの交互列を**履歴**という：$h_2:=(x_0, u_0, x_1, u_1, x_2)$．この履歴は

$$v^2(h_2)=\mu_0(x_0, u_0) \wedge \mu_1(x_1, u_1) \wedge \mu_G(x_2)$$

によって評価される．たとえば，最左端は，履歴$h_2=(s_1, a_1, s_1, a_1, s_1)$のメンバーシップを$\min=\mu_0(a_1) \wedge \mu_1(a_1) \wedge \mu_G(s_1)=0.4 \wedge 0.8 \wedge 0.7=0.4$で評価している．第1段終了時点までに，状態と決定の交互列(s_1, a_1, s_1, a_1)が実現していたとき，履歴h_2が0.4と評価される確率は$p_1=p(s_1|s_1, a_1)=0.6$である．また，右隣の履歴$h_2'=(s_1, a_1, s_1, a_1, s_2)$はこの時点で確率$p_1'=p(s_2|s_1, a_1)=0.4$で$\min'=0.4$と評価される．したがって，履歴$(s_1, a_1, s_1, a_1)$の期待評価値は

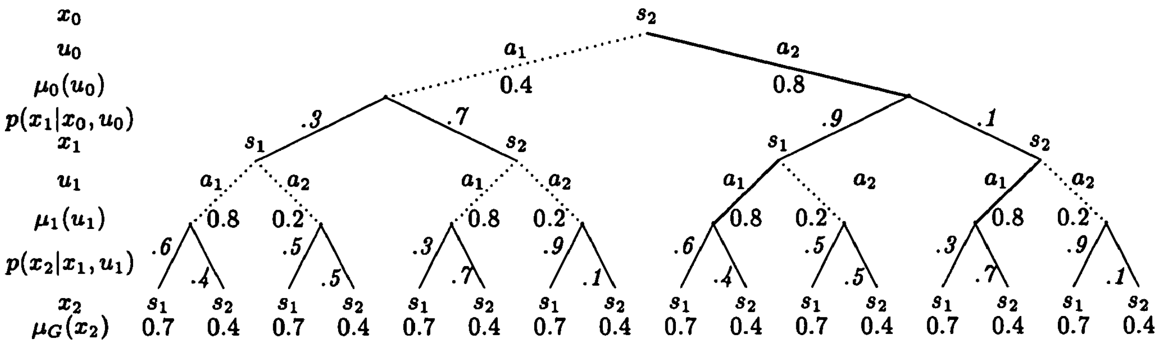
$$\begin{aligned} & v^2(h_2)p(s_1|s_1, a_1)+v^2(h_2')p(s_2|s_1, a_1) \\ & =v^2(h_2)p_1+v^2(h_2')p_1' \\ & =0.4 \times 0.6+0.4 \times 0.4 \\ & =0.24+0.16 \\ & =0.40 \end{aligned}$$

である．同様に，この隣りの履歴(s_1, a_1, s_1, a_2)の期待評価値は$0.2 \times 0.5+0.2 \times 0.5=0.10+0.10=0.20$である．したがって，部分履歴$h_1=(s_1, a_1, s_1)$が実現したとき，意思決定者はこの二つの値の大きい方が実現する決定を選択し，大きい方の値を得ると考えられる：



min	0.4	0.4	0.2	0.2	0.4	0.4	0.2	0.2	0.7	0.4	0.2	0.2	0.7	0.4	0.2	0.2
$\times p_1$	.24	.16	.10	.10	.12	.28	.18	.02	.42	.16	.10	.10	.21	.28	.18	.02
sub-e	**.40**		.20		**.40**		.20		**.58**		.20		**.49**		.20	
$\times p_0$	.240				.160				.290				.245			
ttl-e	.400								**.535**							

図 1-1 s_1 からの決定樹表



min	0.4	0.4	0.2	0.2	0.4	0.4	0.2	0.2	0.7	0.4	0.2	0.2	0.7	0.4	0.2	0.2
$\times p_1$	.24	.16	.10	.10	.12	.28	.18	.02	.42	.16	.10	.10	.21	.28	.18	.02
sub-e	**.40**		.20		**.40**		.20		**.58**		.20		**.49**		.20	
$\times p_0$	.120				.280				.522				.049			
ttl-e	.400								**.571**							

図 1-2 s_2 からの決定樹表

ただし　min$=\mu_0(u_0)\wedge\mu_1(u_1)\wedge\mu_G(x_2)$,　　sub-e$=\sum_{x_2}$ min$\times p_1(x_2|x_1,u_1)$,

$prob=p_0(x_1|x_0,u_0)p_1(x_2|x_1,u_1)$　　ttl-e$=\sum_{x_1}\sum_{x_2}$ min$\times prob$

u_n	a_1	a_2
$\mu_0(u_0)$	0.4	0.8
$\mu_1(u_1)$	0.8	0.2

x_2	$\mu_G(x_2)$
s_1	0.7
s_2	0.4

図 2　メンバーシップ関数列 $\{\mu_0,\mu_1,\mu_G\}$

$p(x_{n+1}|x_n,a_1)$

$x_n\backslash x_{n+1}$	s_1	s_2
s_1	*0.6*	*0.4*
s_2	*0.3*	*0.7*

$p(x_{n+1}|x_n,a_2)$

$x_n\backslash x_{n+1}$	s_1	s_2
s_1	*0.5*	*0.5*
s_2	*0.9*	*0.1*

図 3　推移行列 $\{p(x_{n+1}|x_n,u_n)\}$

$$v^1(h_1)=v^1(s_1,a_1,s_1)=0.40\vee0.20=\mathbf{0.40}$$

このとき，最適な期待値を選択したことを明示して sub-e の行ではボールド体 **0.40** にしている．同様にして，sub-e の行に

$$v^1(s_1,a_1,s_2)=0.40\vee0.20=\mathbf{0.40}$$
$$v^1(s_1,a_2,s_1)=0.58\vee0.20=\mathbf{0.58}$$
$$v^1(s_1,a_2,s_2)=0.49\vee0.20=\mathbf{0.49}$$

が得られる．以上が再帰式を解く第 1 段である．

次の第 2 段では，これら 4 つのボールド体の最適値を用いて，$v^0(h_0)=v^0(s_1)$ を計算する．まず，第 0 段で状態（第 0 段までの履歴）$h_0=x_0=s_1$ にあるとき，決定 $u_0=a_1$ をとれば，次の第 1 段では確率 $p_0=p(s_1|s_1,a_1)=0.6$ で状態 $x_1=s_1$ に行き，評価 $v^1(s_1,a_1,s_2)=0.40$ を得，確率 $p_0'=p(s_2|s_1,a_1)=0.4$ で状態 x_1

$=s_2$ に行き，評価 $v^1(s_2, a_1, s_2)=0.40$ を得る．したがって，履歴 $(x_0, u_0)=(s_1, a_1)$ の期待評価値は

$$
\begin{aligned}
&v^1(s_1, a_1, s_2)p(s_1|s_1, a_1)+v^1(s_2, a_1, s_2)p(s_2|s_1, a_1)\\
&=0.40p_0+0.40p_0'\\
&=0.40\times 0.6+0.40\times 0.4\\
&=0.240+0.160\\
&=0.400
\end{aligned}
$$

である．同様に，履歴 (s_1, a_2) の期待評価値は $0.58\times 0.5+0.49\times 0.5=0.535$ である．したがって，部分履歴 $h_0=x_0=s_1$ から出発したとき，意思決定者は大きい方が実現する決定を選択する：

$$v^0(h_0)=v^0(s_1)=0.400\vee 0.535=\mathbf{0.535}$$

このとき，ttl-e の行では最適な期待値をボールド体 **0.535** にしている．

以上が $x_0=s_1$ からの最適解（最適値と最適選択）である．

図 1-2 に対しても，同様にして $x_0=s_2$ からの最適解が得られる．図 1-2 では最適選択を実線で記入している．

2 段確率決定樹表（図 1-1, 図 1-2）の最適解に至るまでの 2 回の 1 段期待値演算と 1 段最適化演算を N 回で表すと，次になる：

定理 4.1（Iwamoto, Tsurusaki and Fujita[14]）

$$
\begin{aligned}
&v^n(h)=\max_{u\in U}\sum_{y\in X} v^{n+1}(h, u, y)p(y|x, u)\\
&\qquad\qquad h\in H_n, 0\leq n\leq N-1 \qquad (4)\\
&v^N(h)=r_0\wedge\cdots\wedge r_{N-1}\wedge\mu_G \quad h\in H_N.
\end{aligned}
$$

式(4)の最大点を $\nu_n^*(h)(h\in H_n)$ とすると，原始政策クラスの中での最適政策 $\nu^*=\{\nu_0^*, \nu_0^*, \cdots, \nu_{N-1}^*\}$ が得られる．

ここでは，時刻 n では，ファジィシステムの**実際の状態**は x_n にあるが，それまでの部分履歴

$$h_n:=(x_0, u_0, x_1, u_1, \cdots, x_{n-1}, u_{n-1}, x_n)$$

自身を**新しい**状態と見なしている．このとき，1 段期待値演算と 1 段最適化演算を繰り返して，全段にわたる期待値最適問題を解いている．これも動的計画である．これを原始政策クラス上での動的計画法という．この動的計画法は評価関数

$$g: H_N\rightarrow R^1$$

に依存しない．評価関数フリーである．上記の計算過程ではたまたま**最小型**

$$g(h)=r_0\wedge r_1\wedge\cdots\wedge r_{N-1}\wedge\mu_G$$

であった．この動的計画法では状態数が段数のべき乗で増える．

5. 埋め込み法

さて，以上の動的計画法に対して，以下のファジィ動的計画法では，状態空間を **1 次元だけ拡大**している．ここでは，実際の状態 x_n までの部分履歴

$$h_n=(x_0, u_0, x_1, u_1, \cdots, x_{n-1}, u_{n-1}, x_n)$$

に代わって，それまでの**累積メンバーシップ値**

$$\lambda_n:=\mu_0(x_0, u_0)\wedge\mu_1(x_1, u_1)\wedge\cdots\wedge\mu_{n-1}(x_{n-1}, u_{n-1})$$

を追加して，対 $(x_n; \lambda_n)$ を第 n 段での状態と考える．このとき，新状態 $(x_n; \lambda_n)$ は決定 $u_n\in U$ によって，次の状態 $(x_{n+1}; \lambda_{n+1})$ に確率 $p(x_{n+1}|x_n, u_n)$ で推移する．ただし，λ_n は次には確定的に $\lambda_n\wedge\mu_n(x_n, u_n)$ に推移する：

$$\lambda_{n+1}=\lambda_n\wedge\mu_n(x_n, u_n).$$

本来の状態は第 0 段の $x_0\in X$ から推移するが，拡大状態上の最小型決定過程では，十分大きい数 λ_0 を想定して，対 (x_0, λ_0) からスタートすると解釈できる（λ_0 は各段メンバーシップ値以上であればよい）．特に，ファジィ決定過程では $\lambda_0=1.0$ として，(x_0, λ_0) から始まると考えられる．最後の第 N 段では，$(x_{N-1}; \lambda_{N-1})$ から，決定 u_{N-1} によって確率 $p(x_N|x_{N-1}, u_{N-1})$ で，終端状態 $(x_N; \lambda_N)$ に推移する．このとき，**終端状態のみ**に依存して定まるメンバーシップ値

$$\lambda_N\wedge\mu_G(x_N)$$

がシステム全体にかかる，と考える．これは**終端型**評価系である．$\{\lambda_n\}_0^N$ の構成より，この終端メンバーシップ値はシステム全体のメンバーシップになっている：

$$
\begin{aligned}
&\lambda_N\wedge\mu_G(x_N)\\
&\quad=\mu_0(x_0, u_0)\wedge\cdots\wedge\mu_{N-1}(x_{N-1}, u_{N-1})\wedge\mu_G(x_N).
\end{aligned}
$$

このメンバーシップ値が実現する確率は推移確率の積

$$p(x_1|x_0, u_0)p(x_2|x_1, u_1)\cdots p(x_N|x_{N-1}, u_{N-1})$$

である．このように，各時点で最適な決定を下しながら全メンバーシップの期待値を最大化しようとするのが，ファジィ決定過程である．このときの決定 $u_n\in U$ はその時の状態対 $(x_n; \lambda_n)$ に依存して定まる．この決定関数を**拡大マルコフ決定関数**といい，この列を**拡大マルコフ政策**という[18, 19]．しかし，λ_n は過去の状態列 $(x_0, x_1, \cdots, x_{n-1})$ にも依存しているから，現在までの状態列 $(x_0, x_1, \cdots, x_n)$ に依存して決定 u_n を選択するのが自然な決定関数である．これを**一般決定関数**といい，その列を**一般政策**という[18, 19]．したがって，われわれの確率的ファジィ決定過程は**一般政**

策クラスの中で最適な政策を求めることになる．この点が，Bellman and Zadeh が提案した確率システム上のファジィ意思決定過程と本質的に異なる．彼らは，あたかも加法型評価の期待値最適化方法と同様に，しかも確定的システム上の最小型評価の（非期待値）最適化方法と同じ方法で，確率システム上のファジィ決定過程を解析しようとしている（文献[2]の§.5)．最小型評価の期待値最適化という非線形問題を加法型評価という線形的方法，ないしは非線形の確定的問題の方法を直接当てはめる訳にはいかない．非線形・確率的最適化問題独自のアプローチが求められる．ここではそれは**埋め込み法**(imbedding method）である．この埋め込みには**累積過去値**という新たな1次元パラメータを導入して，拡大状態空間上の**終端評価問題**に持ち込む必要がある．以上の考え方を以下で数学的に定式化して表現する．

このように考えて来ると，Bellman and Zadeh が提案した(2)確率的システム上のファジィ決定過程は次の最小型評価系をもつ期待値最大化と考えられる[11, 15]：

$$\begin{aligned}&\text{Max.}\\&\quad E^{\sigma}_{x_0}[\mu_0\wedge\mu_1\wedge\cdots\wedge\mu_{N-1}\wedge\mu_G]\\&\text{s.t.}\\&\quad \text{(i)}\ \ X_{n+1}\sim p(\cdot|x_n, u_n)\\&\quad \text{(ii)}\ \ u_n\in U \qquad n=0,\cdots,N-1.\end{aligned}\tag{5}$$

問題(5)の目的式は，一般政策 σ，初期状態 x_0 および推移法則 p で定まる（離散）確率測度による最小型関数の期待値である：

$$\begin{aligned}&E^{\sigma}_{x_0}[\mu_0\wedge\mu_1\wedge\cdots\wedge\mu_{N-1}\wedge\mu_G]\\&\quad=\sum\sum_{(x_1,\cdots,x_N)\in X\times\cdots\times X}\cdots\sum[\mu_0\wedge\cdots\wedge\mu_{N-1}\wedge\mu_G]\\&\qquad\qquad\times p_0p_1\cdots p_{N-1}\end{aligned}\tag{6}$$

ただし

$$\begin{aligned}&\mu_0\wedge\cdots\wedge\mu_{N-1}\wedge\mu_G\\&\quad=\mu_0(x_0, u_0)\wedge\cdots\wedge\mu_{N-1}(x_{N-1}, u_{N-1})\wedge\mu_G(x_N),\\&p_0p_1\cdots p_{N-1}\\&\quad=p(x_1|x_0, u_0)p(x_2|x_1, u_1)\cdots p(x_N|x_{N-1}, u_{N-1})\end{aligned}$$

ここに，各時点での決定は一般政策およびそれ以前の状態列に依存して定まる：

$$u_0=\sigma_0(x_0),\ u_1=\sigma_1(x_0, x_1),\ \cdots,\ u_{N-1}=\sigma_{N-1}(x_0, x_1, \cdots, x_{N-1})$$

まず，**累積値集合列** $\{\Lambda_n\}_0^N$ を次で定義する：

$$\begin{aligned}&\Lambda_0:=\{1.0\}\\&\Lambda_{n+1}:=\{\lambda_{n+1}|\lambda_{n+1}=\mu_0(x_0, u_0)\wedge\cdots\wedge\mu_n(x_n, u_n),\\&\qquad (x_m, u_m)\in X\times U\quad 0\le m\le n\}\\&\qquad\qquad 1\le n\le N-1\end{aligned}$$

このとき，累積値集合列は前向きの再帰式

補題 5.1

$$\begin{aligned}&\Lambda_0=\{1.0\}\\&\Lambda_{n+1}=\{\lambda_n\wedge\mu_n(x_n, u_n)|\lambda_n\in\Lambda_n,\\&\qquad\qquad (x_n, u_n)\in X\times U\}\end{aligned}\tag{7}$$

を満たす．他方，第 n 段の拡大状態 $(x_n, \lambda_n)(\in X_n\times\Lambda_n)$ から始まる部分過程

$$\begin{aligned}&\text{Max.}\\&\quad E^{\gamma}_{x_n}[\lambda_n\wedge\mu_n\wedge\cdots\wedge\mu_{N-1}\wedge\mu_G]\\&\text{s.t.}\\&\quad \text{(i)}\ \ X_{m+1}\sim p(\cdot|x_m, u_m)\\&\quad \text{(ii)}\ \ u_m\in U \qquad m=n,\cdots,N-1\end{aligned}\tag{8}$$

の拡大マルコフ政策クラス上での最大値を $v_n(x_n;\lambda_n)$ とする．ただし，

$$v_N(x_N;\lambda_N):=\lambda_N\wedge\mu_G(x_N).$$

このとき，最大値関数列は次の後向き再帰式を満たす：

定理 5.1 (Iwamoto and Fujita[12], Iwamoto, Tsurusaki and Fujita[15])

$$\begin{aligned}&v_n(x;\lambda)\\&\quad=\underset{u\in U}{\text{Max}}\sum_{y\in X}v_{n+1}(y;\lambda\wedge\mu_n(x, u))p(y|x, u)\\&\qquad\qquad x\in X, \lambda\in\Lambda_n, 0\le n\le N-1\\&v_N(x;\lambda)=\lambda\wedge\mu_G(x)\quad x\in X, \lambda\in\Lambda_N.\end{aligned}\tag{9}$$

再帰式(9)の最大値を与える決定 $u\in U$ を $\gamma_n^*(x;\lambda)$ で表すと，決定関数 $\gamma_n^*: X\times\Lambda_n\to U$ が得られる．この列 $\gamma^*=\{\gamma_0^*, \gamma_1^*, \cdots, \gamma_{N-1}^*\}$ は拡大マルコフクラスの中での最適政策である．

一般に，任意の拡大マルコフ政策 $\gamma=\{\gamma_0, \gamma_1, \cdots, \gamma_{N-1}\}$ は以下の変換(10)によって（期待値が等しい，という意味で）**同値な一般政策** $\sigma=\{\sigma_0, \sigma_1, \cdots, \sigma_{N-1}\}$ を生成する．すなわち，任意の状態列 $(x_0, x_1, \cdots, x_n)$ が与えられたとき，中間のパラメータ値列 $\{\lambda_m\}_{0\le m\le n}$ を生成しながら $\sigma_n(x_0, x_1, \cdots, x_n)$ を次で定義する：

$$\begin{aligned}&\lambda_0:=1.0\\&u_0:=\gamma_1(x_0;\lambda_0),\ \lambda_1:=\lambda_0\wedge\mu_0(x_0, u_0)\\&u_1:=\gamma_1(x_1;\lambda_1),\ \lambda_2:=\lambda_1\wedge\mu_1(x_1, u_1)\\&\qquad\vdots\\&u_{n-1}:=\gamma_{n-1}(y_{n-1}),\ \lambda_n:=\lambda_{n-1}\wedge\mu_{n-1}(x_{n-1}, u_{n-1})\\&\sigma_n(x_0, x_1, \cdots, x_n):=\gamma_n(x_n;\lambda_n).\end{aligned}\tag{10}$$

このとき，両政策による期待値は等しい：

$$\begin{aligned}&E^{\sigma}_{x_0}[\mu_0\wedge\cdots\wedge\mu_{N-1}\wedge\mu_G]\\&\quad=E^{\gamma}_{x_0}[1.0\wedge\mu_0\wedge\cdots\wedge\mu_{N-1}\wedge\mu_G].\end{aligned}\tag{11}$$

したがって，拡大マルコフ政策として最適な γ^* から変換(10)によって一般政策として最適な σ^* が一般政策クラス Π_g の中で得られる．

6. 終りに

ファジィ動的計画法を(2)確率システム上で議論してきたが，ファジィ決定過程としては他に，(3)ファジィシステム上でも考えられる[9～11, 13]．さらに(4)条件つき（conditional）ファジィ決定過程として，(4.a)事後条件つき決定過程，(4.b)事前条件つき決定過程が考えられる．冒頭述べた，Bellman and Zadeh の確率的システム上の決定過程は他でもない，事後条件つき決定過程であることが分かる[14]．

参考文献

[1] Bellman R. E.: "Dynamic Programming", Princeton Univ. Press, NJ (1957).

[2] Bellman R. E. and Zadeh L.: "Decision-making in a fuzzy environment", *Management Science* 17(1970), B 141-B 164.

[3] 茨木俊秀："組合せ最適化の理論"，電子通信学会(1979).

[4] 岩本誠一："逐次決定過程としての動的計画論 I, II"，オペレーションズ・リサーチ 22(1977)，427-434，496-501.

[5] 岩本誠一："動的計画論"，九州大学出版会（1987).

[6] 岩本誠一："動的計画の最近の進歩"，第2回 RAMP シンポジウム論文集（1990)，129-140.

[7] Iwamoto S.: "Maximizing threshold probability through invariant imbedding", *Proceedings of The Eighth BELLMAN CONTINUUM*, ROC, 2000, 17-22.

[8] Iwamoto S.: "Fuzzy decision-making through three dynamic programming approaches", *Proceedings of The Eighth BELLMAN CONTINUUM*, ROC, 2000, 23-27; (Full Paper) *International Journal of Fuzzy Systems* 3(2001), No. 4, 520-526.

[9] 岩本誠一："ファジィ動的計画法"，石井・坂和・岩本（共編)，"ファジィ OR"，朝倉書店（2001)，112-157.

[10] Iwamoto S.: "A class of dual fuzzy dynamic programs", *Proceedings of The Seventh BELLMAN CONTINUUM*, The Santa Fe Institute, 1999; *Applied Mathematics and Computation* 120(2001), No. 1/3, 91-108.

[11] 岩本誠一："Fuzzy Dynamic Programming"，坂和（編)，日本 OR 学会第46回シンポジウム「ファジィ OR」(2001)，87-104.

[12] Iwamoto S. and Fujita T.: "Stochastic decision-making in a fuzzy environment", *J. Operations Res. Soc. Japan* 38(1995), No. 4, 467-482.

[13] Iwamoto S. and Sniedovich M.: "Sequential decision making in fuzzy environment", *J. Math. Anal. Appl.* 222(1998), No. 1, 208-224.

[14] Iwamoto S., Tsurusaki K. and Fujita T.: "Conditional decision-making in a fuzzy environment", *J. Operations Res. Soc. Japan* 42(1999), No. 2, 198-218.

[15] Iwamoto S., Tsurusaki K. and Fujita T.: "On Markov policies for minimax decision processes", *J. Math. Anal. Appl.* 253(2001), No. 1, 58-78.

[16] Iwamoto S. and Ueno T.: "Fuzzy decision-making under threshold membership criterion", *Frontiers in Artificial Intelligence and Applications* 69(2001), IOS Press, 797-801.

[17] Iwamoto S., Ueno T. and Fujita T.: "Controlled Markov chains with utility functions", *Markov Processes and Controlled Markov Chains, Chap.* 8, Kluwer, 2002, 135-148.

[18] 日本 OR 学会編："OR 用語辞典"，日科技連（2000).

[19] 日本 OR 学会編："OR 事典 2000"，CD-ROM 版，日本 OR 学会（2000).

[20] M. Sniedovich,: "Dynamic Programming", Marcel Dekker, Inc. NY, 1992.