利得関数のとる値が擬順序集合の要素であるときの確定的動的計画

弘前大学 大学院 理工学研究科　小野 皓多 (Ono KOTA)
Graduate School of Science and Technology, Hirosaki University

弘前大学 大学院 理工学研究科　金 正道 (Masamichi KON)
Graduate School of Science and Technology, Hirosaki University

概要

利得関数のとる値が擬順序集合の要素であるときの確定的動的計画を考察する。ここで、利得関数は結合的二項演算によって集計される。まず、二項演算が単調性のもとで再帰式を導く。そして、マルコフ政策がこの問題を解くのに十分な政策であり、この問題のスカラー化問題を解くことによってこの問題の (弱) 非劣マルコフ政策が得られることを示す。

1　はじめに

動的計画法は Bellman [1] によって始められ、非常に多く研究されている (Sniedovich [6] 参照)。有限期間の各期において、利得または費用 (実数) が状態と決定によって値が定まる。動的計問題は、すべての期間での利得または費用の集計値を最大または最小にするような最適な政策を求める問題である。各期において、政策はその期の状態またはそれ以前の期の状態から決定を定める。動的計画問題は、状態の推移が確定的であるとき確定的動的計画問題といい、確率的であるとき確率的動的計画問題という。大半の研究では、利得または費用 (実数) が、加法かあるいは他の演算によって集計される。本稿では二種類の政策を考える。一つは一般政策である。一般政策は各期にそれ以前の期の状態からその期の決定を定める政策である。もう一つはマルコフ政策である。マルコフ政策は各期にその期の状態からその期の決定を定める政策である。

Iwamoto [3] における確率的動的計画問題においては、利得または費用 (実数) が加法、乗法、加法と乗法、最小演算、除法によって集計されている。Fujita [2] においては、利得 (実数) が加法によって集計されるとき、マルコフ政策が確率的または確定的動的計画問題に対して十分であることが示された。Iwamoto, Tsurusaki and Fujita [4] では、利得 (実数) が最小演算によって集計されるとき、マルコフ政策が確率的または確定的動的

計画問題に対して十分でないことが示された。

本稿では、利得関数のとる値が擬順序集合の要素であるときの確定的動的計画を考察する。ここで、利得関数は結合的二項演算によって集計される。そして利得関数が実数であるとき、二項演算の単調の仮定の下で再帰式を導き、マルコフ政策のみを考えれば十分であることを示す。さらに、元の問題 (利得が実数である必要はない) のスカラー化問題を解くことによって元の問題の (弱) 非劣マルコフ政策が得られることを示す。

2　順序と二項演算

本稿を通して、$\mathcal{L} \neq \emptyset$ を集合とし、$\preceq$ を $\mathcal{L}$ 上の擬順序（反射的, 推移的）とし、$\prec$ を $\mathcal{L}$ 上の狭義半順序（非反射的, 推移的）とする。$x, y \in \mathcal{L}$ に対して、$x \preceq y, x \prec y$ をそれぞれ $y \succeq x, y \succ x$ とも表す。さらに

$$x, y \in \mathcal{L}, x \prec y \Rightarrow x \preceq y, x \not\succeq y \tag{2.1}$$

であると仮定する。$\mathcal{L} \subset \mathbb{R}$ であるときは、$\preceq$ は $\leq$ であり、$\prec$ は $<$ であるとする。また、以降の例で用いるため

$$\mathcal{K} = \{A \subset \mathbb{R} : \min A, \max A \text{ が存在する }\} \tag{2.2}$$

とし、各 $A \in \mathcal{K}$ に対して

$$A^L = \min A, \quad A^R = \max A \tag{2.3}$$

とする。

例 2.1　$\mathcal{L} = \mathcal{K}$ とする。$A, B \in \mathcal{L}$ に対して

$$\begin{aligned}
A \preceq B &\overset{\mathrm{def}}{\Leftrightarrow} A^L \leq B^L,\ A^R \leq B^R \\
A \prec_1 B &\overset{\mathrm{def}}{\Leftrightarrow} A^L < B^L,\ A^R < B^R \\
A \prec_2 B &\overset{\mathrm{def}}{\Leftrightarrow} A \preceq B,\ A \not\succeq B \\
A \prec_3 B &\overset{\mathrm{def}}{\Leftrightarrow} A \preceq B,\ A \neq B
\end{aligned}$$

とする。このとき、$\preceq$ は $\mathcal{L}$ 上の擬順序（反射的, 推移的）になり、$\prec_1, \prec_2$ は $\mathcal{L}$ 上の狭義半順序（非反射的, 推移的）になる。$\prec_3$ は非反射的であるが推移的にならないので、$\mathcal{L}$ 上の狭義半順序にはならない。また

$$A, B \in \mathcal{L},\ A \prec_1 B \ \Rightarrow\ A \preceq B,\ A \not\succeq B$$

となる。 $\square$

定義 2.1 $\mathcal{S} \subset \mathcal{L}$ とし、$x_0 \in \mathcal{S}$ とする。

(i) x_0 が $\mathcal{S}$ の最大（小）要素であるとは

$$x_0 \succeq (\preceq) x, \quad x \in \mathcal{S}$$

となるときをいう。

(ii) x_0 が $\mathcal{S}$ の極大（小）要素であるとは

$$x \in \mathcal{S},\ x \succeq (\preceq) x_0 \ \Rightarrow\ x \preceq (\succeq) x_0$$

となるときをいう。

(iii) x_0 が $\mathcal{S}$ の弱極大（小）要素であるとは

$$\nexists x \in \mathcal{S} \text{ s.t. } x \succ (\prec) x_0$$

となるときをいう。

$\mathcal{S} \subset \mathcal{L}$ に対して、$\preceq\text{-}\max \mathcal{S}$, $\preceq\text{-}\min \mathcal{S}$, $\prec\text{-}\max \mathcal{S}$, $\prec\text{-}\min \mathcal{S}$ をそれぞれ $\mathcal{S}$ のすべての極大要素, 極小要素, 弱極大要素, 弱極小要素の集合とする。

命題 2.1 $\mathcal{S} \subset \mathcal{L}, \mathcal{S} \neq \emptyset$ は有限集合であるとする。このとき、$\preceq\text{-}\max \mathcal{S} \neq \emptyset$, $\preceq\text{-}\min \mathcal{S} \neq \emptyset$ となる。

命題 2.2 $\mathcal{S} \subset \mathcal{L}$ とする。このとき、$\preceq\text{-}\max \mathcal{S} \subset \prec\text{-}\max \mathcal{S}$, $\preceq\text{-}\min \mathcal{S} \subset \prec\text{-}\min \mathcal{S}$ となる。

本稿を通して、$\circ : \mathcal{L} \times \mathcal{L} \to \mathcal{L}$ を $\mathcal{L}$ 上の結合的な二項演算とする。$x \circ y \circ z$ や $x_1 \circ x_2 \circ \cdots \circ x_n$ などと表す。

定義 2.2

$\circ$ が右（左）単調であるとは、

$$x, y, z \in \mathcal{L},\ y \preceq z \ \Rightarrow\ x \circ y \preceq x \circ z \ (y \circ x \preceq z \circ x)$$

となるときをいう。

例 2.2 $L=\mathbb{R}$, $L\neq\emptyset$ とし、

$$\mathcal{L}=\{A\in\mathcal{K}:A\subset L\}$$

とする。また、L 上の二項演算 $\bullet:L\times L\to L$ は可換、結合的かつ右単調であるとする。二項演算 $\circ:\mathcal{L}\times\mathcal{L}\to\mathcal{L}$ を各 $A,B\in\mathcal{L}$ に対して

$$A\circ B=\{x\bullet y:x\in A,y\in B\}$$

と定義し、$A,B\in\mathcal{L}$ に対して

$$A\preceq B\overset{\text{def}}{\Leftrightarrow}A^L\leq B^L,\ A^R\leq B^R$$

と定義する。このとき、$\circ$ は可換、結合的かつ右単調である。

$[0,1]$ 上の二項演算である t-ノルム（または t-劣ノルム）および t-コノルム（または t-劣コノルム）は可換、結合的かつ右単調になる。t-ノルム（または t-劣ノルム）および t-コノルム（または t-劣コノルム）について、詳しくは [5] 参照。

3　スカラー化

$\psi:\mathcal{L}\to\mathbb{R}$ とし、ψ を**スカラー化関数**とよぶ。$\bullet$ を $\psi(\mathcal{L})$ 上の二項演算とし、任意の $x',y'\in\psi(\mathcal{L})$ および $x'=\psi(x)$, $y'=\psi(y)$ となる任意の $x,y\in\mathcal{L}$ に対して

$$x'\bullet y'=\psi(x\circ y) \tag{3.1}$$

であると仮定する。このとき、任意の有限個の $x,y,\cdots,z\in\mathcal{L}$ に対して

$$\psi(x)\bullet\psi(y)\bullet\cdots\bullet\psi(z)=\psi(x\circ y\circ\cdots\circ z) \tag{3.2}$$

となる。

定義 3.1 ψ が狭義単調増加であるとは

$$x,y\in\mathcal{L},\ x\prec y\ \Rightarrow\ \psi(x)<\psi(y)$$

となるときをいう。

例 3.1 $L=\mathbb{R}$, $L\neq\emptyset$ とし、

$$\mathcal{L}=\{A\in\mathcal{K}:A\subset L\}$$

とする。また、L 上の二項演算 $\bullet : L \times L \to L$ は可換かつ結合的かつ狭義右単調であるとする。$\bullet$ が狭義右単調であるとは、$x, y, z \in L$ に対して、$y < z$ ならば $x \bullet y < x \bullet z$ が成り立つことである。二項演算 $\circ : \mathcal{L} \times \mathcal{L} \to \mathcal{L}$ を各 $A, B \in \mathcal{L}$ に対して

$$A \circ B = \{x \bullet y : x \in A, y \in B\}$$

と定義し、$A, B \in \mathcal{L}$ に対して

$$A \preceq B \overset{\text{def}}{\Leftrightarrow} A^L \leq B^L,\ A^R \leq B^R$$
$$A \prec B \overset{\text{def}}{\Leftrightarrow} A \preceq B,\ A \not\succeq B$$

と定義する。また、スカラー化関数 $\psi : \mathcal{L} \to \mathbb{R}$ を各 $A \in \mathcal{L}$ に対して

$$\psi(A) = A^L \bullet A^R$$

と定義する。このとき、ψ は狭義単調増加であり、$x, y \in \psi(\mathcal{L})$ と $x = \psi(A), y = \psi(B)$ となるような $A, B \in \mathcal{L}$ に対して

$$x \bullet y = \psi(A \circ B)$$

となる。

$(0, 1]$ に制限した狭義単調 t-ノルムや、$[0, 1)$ に制限した狭義単調 t-コノルムは可換、結合的かつ狭義右単調である。詳しくは [5] 参照。

4 確定的動的計画

本稿を通して、次のように表記を定義する。

$N \geq 2$	: 段の総数
$X = \{s_1, s_2, \cdots, s_p\}$	: 有限状態空間
$U = \{a_1, a_2, \cdots, a_k\}$	: 有限決定空間
$r_n : X \times U \to \mathcal{L}$	: 第 n 利得関数 $(1 \leq n \leq N)$
$r_G : X \to \mathcal{L}$	: 終端利得関数
$f : X \times U \to X$	: 推移関数（$f(s, a)$ は第 n 段で状態 s, 決定 a のときの第 $n+1$ 段での状態）

とする。このとき、次の確定的動的計画問題を考える。

$$\left|\begin{array}{ll} \max & r_1(x_1, u_1) \circ r_2(x_2, u_2) \circ \cdots \\ & \qquad \circ\, r_N(x_N, u_N) \circ r_G(x_{N+1}) \\ \text{s.t.} & f(x_n, u_n) = x_{n+1}, \quad 1 \leq n \leq N \\ & u_n \in U, \quad 1 \leq n \leq N \end{array}\right. \tag{4.1}$$

ここで、$x_1 \in X$ である。問題 (4.1) は、各 $x_1 \in X$ に対して、一般政策（定義は後述）の範囲で最適一般政策または（弱）非劣一般政策（定義は後述）を求める問題であると解釈する場合と、マルコフ政策（定義は後述）の範囲で最適マルコフ政策または（弱）非劣マルコフ政策（定義は後述）を求める問題であると解釈する場合がある。それらの場合を区別して、一般政策の範囲で考えた問題 (4.1) を**一般問題**とよび、マルコフ政策の範囲で考えた問題 (4.1) を**マルコフ問題**とよぶ。

$$\sigma_1 : X \to U, \quad \sigma_2 : X \times X \to U, \quad \cdots, \quad \sigma_N : \underbrace{X \times \cdots \times X}_{N\text{ 個}} \to U \tag{4.2}$$

のとき、$\sigma = (\sigma_1, \sigma_2, \cdots, \sigma_N)$ を問題 (4.1) に対する**一般政策**という。問題 (4.1) に対するすべての一般政策の集合を $\mathcal{G}$ とする。

$$\pi_1 : X \to U, \quad \pi_2 : X \to U, \quad \cdots, \quad \pi_N : X \to U \tag{4.3}$$

のとき、$\pi = (\pi_1, \pi_2, \cdots, \pi_N)$ を問題 (4.1) に対する**マルコフ政策**という。問題 (4.1) に対するすべてのマルコフ政策の集合を $\mathcal{M}$ とする。

問題 (4.1) において、$x_1 \in X$ を任意に固定する。このとき、一般政策 $\sigma \in \mathcal{G}$ またはマルコフ政策 $\pi \in \mathcal{M}$ を与えると

$$u_1, u_2, \cdots, u_n \quad \text{および} \quad x_2, \cdots, x_N, x_{N+1}$$

が定まり、目的関数の値

$$g(x_1; \sigma) = r_1(x_1, u_1) \circ r_2(x_2, u_2) \circ \cdots \circ r_N(x_N, u_N) \circ r_G(x_{N+1})$$

または

$$g(x_1; \pi) = r_1(x_1, u_1) \circ r_2(x_2, u_2) \circ \cdots \circ r_N(x_N, u_N) \circ r_G(x_{N+1})$$

が定まる。一般問題を考えるときは $g : X \times \mathcal{G} \to \mathcal{L}$ と解釈し、マルコフ問題を考えるときは $g : X \times \mathcal{M} \to \mathcal{L}$ と解釈する。

定義 4.1

(i) $\sigma^* \in \mathcal{G}$ が**最適一般政策**であるとは、任意の $x_1 \in X$ に対して

$$\forall \sigma \in \mathcal{G}, \qquad g(x_1, \sigma^*) \succeq g(x_1, \sigma)$$

となるときをいう。

(ii) $\sigma^* \in \mathcal{G}$ が**非劣一般政策**であるとは、任意の $x_1 \in X$ に対して

$$\sigma \in \mathcal{G}, g(x_1, \sigma) \succeq g(x_1, \sigma^*) \ \Rightarrow \ g(x_1, \sigma) \preceq g(x_1, \sigma^*)$$

となるときをいう。

(iii) $\sigma^* \in \mathcal{G}$ が**弱非劣一般政策**であるとは、任意の $x_1 \in X$ に対して

$$\nexists \sigma \in \mathcal{G} \text{ s.t. } g(x_1, \sigma) \succ g(x_1, \sigma^*)$$

となるときをいう。

(iv) $\pi^* \in \mathcal{M}$ が**最適マルコフ政策**であるとは、任意の $x_1 \in X$ に対して

$$\forall \pi \in \mathcal{M}, \quad g(x_1, \pi^*) \succeq g(x_1, \pi)$$

となるときをいう。

(v) $\pi^* \in \mathcal{M}$ が**非劣マルコフ政策**であるとは、任意の $x_1 \in X$ に対して

$$\pi \in \mathcal{M}, g(x_1, \pi) \succeq g(x_1, \pi^*) \ \Rightarrow \ g(x_1, \pi) \preceq g(x_1, \pi^*)$$

となるときをいう。

(vi) $\pi^* \in \mathcal{M}$ が**弱非劣マルコフ政策**であるとは、任意の $x_1 \in X$ に対して

$$\nexists \pi \in \mathcal{M} \text{ s.t. } g(x_1, \pi) \succ g(x_1, \pi^*)$$

となるときをいう。

利得関数が実数であるならば、最適一般政策は必ず存在する。しかし、利得関数のとる値が擬順序集合の要素であるとき、最適一般政策が存在するとは限らない。

たとえ利得関数が実数であったとしても、最適一般政策でマルコフ政策になるようなものが存在することは明らかではない。この点について次の節で述べる。

5　スカラー確定的動的計画

本節では

$$\mathcal{L} \subset \mathbb{R}$$

であると仮定する。

問題 (4.1) に対応する次の部分問題群を考える。

$$\left|\begin{array}{ll} \max & r_n(x_n, u_n) \circ \cdots \circ r_N(x_N, u_N) \circ r_G(x_{N+1}) \\ \text{s.t.} & f(x_m, u_m) = x_{m+1}, \quad n \leq m \leq N \\ & u_m \in U, \quad n \leq m \leq N \end{array}\right. \tag{5.1}$$

ここで、$n \in \{1, 2, \cdots, N\}$ であり、$x_n \in X$ である。一般政策の範囲で考えた部分問題 (5.1) を**一般部分問題**とよび、マルコフ政策の範囲で考えた部分問題 (5.1) を**マルコフ部分問題**とよぶ。

$$\sigma_n : X \to U, \quad \sigma_{n+1} : X \times X \to U, \quad \cdots, \quad \sigma_N : \underbrace{X \times \cdots \times X}_{N-n+1\ \text{個}} \to U \tag{5.2}$$

のとき、$\sigma = (\sigma_n, \sigma_{n+1}, \cdots, \sigma_N)$ を部分問題問題 (5.1) に対する**一般政策**という。部分問題 (5.1) に対するすべての一般政策の集合を $\mathcal{G}_n$ とする。$\mathcal{G}_1 = \mathcal{G}$ である。

$$\pi_n : X \to U, \quad \pi_{n+1} : X \to U, \quad \cdots, \quad \pi_N : X \to U \tag{5.3}$$

のとき、$\pi = (\pi_n, \pi_{n+1}, \cdots, \pi_N)$ を部分問題 (5.1) に対する**マルコフ政策**という。部分問題 (5.1) に対するすべてのマルコフ政策の集合を $\mathcal{M}_n$ とする。$\mathcal{M}_1 = \mathcal{M}$ である。

部分問題 (5.1) において、$x_n \in X$ を任意に固定する。このとき、一般政策 $\sigma \in \mathcal{G}_n$ またはマルコフ政策 $\pi \in \mathcal{M}_n$ を与えると

$$u_n, u_{n+1}, \cdots, u_n \quad \text{および} \quad x_{n+1}, \cdots, x_N, x_{N+1}$$

が定まり、目的関数の値

$$g_n(x_n; \sigma) = r_n(x_n, u_n) \circ r_{n+1}(x_{n+1}, u_{n+1}) \circ \cdots \circ r_N(x_N, u_N) \circ r_G(x_{N+1})$$

または

$$g_n(x_n; \pi) = r_n(x_n, u_n) \circ r_{n+1}(x_{n+1}, u_{n+1}) \circ \cdots \circ r_N(x_N, u_N) \circ r_G(x_{N+1})$$

が定まる。一般部分問題を考えるときは $g_n : X \times \mathcal{G}_n \to \mathcal{L}$ と解釈し、マルコフ部分問題を考えるときは $g_n : X \times \mathcal{M}_n \to \mathcal{L}$ と解釈する。$g_1 = g$ である。

一般部分問題群 (5.1) に対して

$$\begin{aligned} V^{N+1}(x_{N+1}) &= r_G(x_{N+1}), \quad x_{N+1} \in X \\ V^n(x_n) &= \max_{\sigma \in \mathcal{G}_n} g_n(x_n; \sigma), \quad x_n \in X,\ 1 \leq n \leq N \end{aligned} \tag{5.4}$$

とし、マルコフ部分問題群 (5.1) に対して

$$\begin{aligned} v^{N+1}(x_{N+1}) &= r_G(x_{N+1}), \quad x_{N+1} \in X \\ v^n(x_n) &= \max_{\pi \in \mathcal{M}_n} g_n(x_n; \pi), \quad x_n \in X,\ 1 \leq n \leq N \end{aligned} \tag{5.5}$$

とする。

このとき、$\circ$ が右単調ならば、次の 3 つの定理 5.1-5.3 が得られる。定理 5.1 および 5.2 は再帰式である。

定理 5.1 $\circ$ は右単調であるとする。

$$\begin{aligned} V^{N+1}(x) &= r_G(x), \quad x \in X \\ V^n(x) &= \max_{u \in U}\left[r_n(x,u) \circ V^{n+1}(f(x,u))\right], \quad x \in X,\ 1 \le n \le N \end{aligned} \tag{5.6}$$

定理 5.2 $\circ$ は右単調であるとする。

$$\begin{aligned} v^{N+1}(x) &= r_G(x), \quad x \in X \\ v^n(x) &= \max_{u \in U}\left[r_n(x,u) \circ v^{n+1}(f(x,u))\right], \quad x \in X,\ 1 \le n \le N \end{aligned} \tag{5.7}$$

定理 5.3 $\circ$ は右単調であるとする。

(i)

$$\exists \pi^* \in \mathcal{M} \text{ s.t. } g(x_1; \pi^*) = V^1(x_1),\ \forall x_1 \in X$$

実際、各 $n \in \{1, 2, \cdots, N\}$ および各 $x \in X$ に対して、(5.7)（または (5.6)）において最大値を与える $u \in U$ を $\pi_n^*(x)$ とすると、最適マルコフ政策 $\pi^* = (\pi_1^*, \pi_2^*, \cdots, \pi_n^*)$ が得られる。

(ii)

$$v^n(x) = V^n(x), \quad x \in X,\ 1 \le n \le N$$

6 確定的動的計画のスカラー化

本節では、利得が実数であるとは限らない場合、問題 (4.1) の (弱) 非劣マルコフ政策が、その問題のスカラー化問題を解くことによって得られることを示す。

確定的動的計画問題 (4.1) に対して、次のスカラー確定的動的計画問題を考える。

$$\left|\begin{array}{ll} \max & \psi(r_1(x_1,u_1)) \bullet \psi(r_2(x_2,u_2)) \bullet \cdots \\ & \qquad \bullet\, \psi(r_N(x_N,u_N)) \bullet \psi(r_G(x_{N+1})) \\ \text{s.t.} & f(x_n,u_n) = x_{n+1}, \quad 1 \le n \le N \\ & u_n \in U, \quad 1 \le n \le N \end{array}\right. \tag{6.1}$$

ここで、$x_1 \in X$ である。

問題 (6.1) において、、$x_1 \in X$ を任意に固定する。このとき、一般政策 $\sigma \in \mathcal{G}$ またはマルコフ政策 $\pi \in \mathcal{M}$ を与えると

$$u_1, u_2, \cdots, u_n \quad \text{および} \quad x_2, \cdots, x_N, x_{N+1}$$

が定まり、目的関数の値

$$h(x_1; \sigma) = \psi(r_1(x_1, u_1)) \bullet \psi(r_2(x_2, u_2)) \bullet \cdots \bullet \psi(r_N(x_N, u_N)) \bullet \psi(r_G(x_{N+1}))$$

または

$$h(x_1; \pi) = \psi(r_1(x_1, u_1)) \bullet \psi(r_2(x_2, u_2)) \bullet \cdots \bullet \psi(r_N(x_N, u_N)) \bullet \psi(r_G(x_{N+1}))$$

が定まる。一般問題を考えるときは $h : X \times \mathcal{G} \to \mathbb{R}$ と解釈し、マルコフ問題を考えるときは $h : X \times \mathcal{M} \to \mathbb{R}$ と解釈する。

定理 6.1 $x, y \in \mathcal{L}$ に対して

$$x \prec y \Leftrightarrow x \preceq y, x \not\succeq y$$

であると仮定し、ψ は狭義単調増加であるとする。また、$\pi^* \in \mathcal{M}$（または $\sigma^* \in \mathcal{G}$）をスカラー確定的動的計画問題 (6.1) の最適マルコフ（または一般）政策とする。このとき、$\pi^* \in \mathcal{M}$（または $\sigma^* \in \mathcal{G}$）は元の確定的動的計画問題 (4.1) の非劣マルコフ（または一般）政策になる。

定理 6.2 ψ は狭義単調増加であるとする。また、$\pi^* \in \mathcal{M}$（または $\sigma^* \in \mathcal{G}$）をスカラー確定的動的計画問題 (6.1) の最適マルコフ（または一般）政策とする。このとき、$\pi^* \in \mathcal{M}$（または $\sigma^* \in \mathcal{G}$）は元の確定的動的計画問題 (4.1) の弱非劣マルコフ（または一般）政策になる。

7　結論

本稿では、利得関数のとる値が擬順序集合の要素であるときの確定的動的計画を考察した。ここで、利得関数は結合的二項演算によって集計される。まず、利得が実数であるとき、二項演算の単調性の仮定に下で再帰式を導き、マルコフ政策のみを考えれば十分であることを示した。次に、利得が実数であるとは限らない場合においても、元の問題のスカラー化問題を解くことによって元の問題の (弱) 非劣マルコフ政策が得られることを示した。

しかし、利得関数のとる値が擬順序集合の要素であるとき、(弱) 非劣一般政策が存在するということは示していない。これは今後の課題の一つである。

本稿で扱った確定的動的計画は最大化問題であったが、最小化問題においても定理 5.1-5.3、6.1 および 6.2 と同様な結果が得られる。

参考文献

[1] T. Fujita, Re-examination of Markov policies for additive decision process, Bulletin of Informatics and Cybernetics, Vol.29, 1997, pp.51–65

[2] 岩本誠一, 動的計画論, 九州大学出版会, 1987

[3] S. Iwamoto, Associative dynamic programs, Journal of Mathematical Analysis and Applications, Vol.201, 1996, pp.195–211

[4] S. Iwamoto, K. Tsurusaki and T. Fujita, On Markov policies for minimax decision processes, Journal of Mathematical Analysis and Applications, Vol.253, 2001, pp.58–78

[5] E. P. Klement, R. Mesiar and E. Pap, Triangular norms, Kluwer Academic Publishers, 2000

[6] M. Sniedovich, Dynamic Programming: Foundations and Principles (Second Edition) CRC Press, 2011