

逐次決定過程と動的計画について

広島大 総合科学 岩本誠一

§1. 概要

有限段逐次決定過程において各段利得関数が時刻・状態・決定からなる各組に対して狭義増加または狭義減少のどちらかであるような最適化問題を考える。すなわち第  $n$  段利得関数  $J_n(x_n, a_n; \cdot)$  が  $n, x_n, a_n$  に応じて  $\cdot$  の狭義増加関数になったり狭義減少関数になったりしている。いわゆる動的計画においてはこの関数は、時刻・状態・決定に無関係に常に  $\cdot$  の増加関数になっている。したがって単調性が、常に増加性というのではなく、増加性と減少性が時刻・状態・決定に依存しているという意味で、動的計画の一般化としての逐次決定過程を対象にしている。

この決定過程に対して最適値を求めるための再帰式を導く。また逆過程を導入して逆定理が成り立つことを示す。反転過程は常に定義されるが、反転定理は特別な場合に成立する。特に、常に増加性が満たされているときは、動的計画に対する再帰式ならびに逆定理・反転定理になっている。

この報告では、単調性は増加性か減少性かのいずれか一

方のみを意味している。

## §2. 基本定理

単調性と再帰性があれば、2変数最適化問題が2段階の1変数最適化問題に分離できることを示す。この基本定理により、N段過程に対して再帰式が成り立つことがわかる。

$R^n$  :  $n$ 次元ユークリッド空間

$\times$  : 集合の直積

$+$  : 集合の直和

$X, Y$  : 空でない集合\*

$2^Y$  :  $Y$ の空でない部分集合の全体

$Y : X \rightarrow 2^Y$  すなわち  $\phi \neq Y(x) \subset Y \quad x \in X$  \*

$G_r(Y) \equiv \{(x, y) | y \in Y(x), x \in X\} \subset X \times Y$  :  $Y$ のグラフ

$g : Y \rightarrow R^1$

$f : X \times R^1 \rightarrow R^1$

\* 記号  $Y$  を集合と点対集合値写像の両方に用いている。

$X = X_1 + X_2$

単調性  $f(x; \cdot) : R^1 \rightarrow R^1$  は  $x \in X_1$  のとき

狭義増加,  $x \in X_2$  のとき 狭義減少。

$Q_{pt} = \text{Max}$  または  $\text{Min}$  のいずれか一方

$\overline{Q_{pt}} = \text{Min}(\text{Max})$ ,  $Q_{pt} = \text{Max}(\text{Min})$  のとき

このとき、次の関係式が成立する。

定理 1 (基本定理)

$$\begin{aligned} \text{Opt}_{(x,y) \in G_r(Y)} f(x; g(y)) &= \text{Opt}_{x \in X_1} \left( \text{Opt}_{y \in Y(x)} f(x; g(y)) \right) \\ &\quad , \text{Opt}_{x \in X_2} \left( \overline{\text{Opt}_{y \in Y(x)} f(x; g(y))} \right) \end{aligned} \quad (1)$$

ただし

$$\text{Max}_{z \in \phi} h(z) = -\infty, \quad \text{Min}_{z \in \phi} h(z) = \infty.$$

(証明)  $\text{Opt} = \text{Max}$ ,  $X_1 \neq \phi$ ,  $X_2 \neq \phi$  のときのみを示す。

$$\text{左辺} = \text{Max}_{(x,y) \in G_r(Y)} f(x; g(y)) = f(x^*; g(y^*)) \quad y^* \in Y(x^*), x^* \in X$$

$$\text{右辺} = \text{Max}_{x \in X_1} \left( \text{Max}_{y \in Y(x)} f(x; g(y)) \right), \text{Max}_{x \in X_2} \left( \text{Min}_{y \in Y(x)} f(x; g(y)) \right)$$

$$= \text{Max} \left( f(x_1; g(y_1)), f(x_2; g(y_2)) \right)$$

$$y_1 \in Y(x_1), x_1 \in X_1, y_2 \in Y(x_2), x_2 \in X_2$$

とすれば、 $(x_i, y_i) \in G_r(Y)$   $i=1, 2$  と  $\text{Max}$  の性質より

$$f(x^*; g(y^*)) \geq f(x_i; g(y_i)) \quad i=1, 2.$$

すなわち 左辺  $\geq$  右辺。

逆に、(1)  $x^* \in X_1$  ならば、 $f(x^*; \cdot)$  が増加だから

$$\text{左辺} = f(x^*; g(y^*))$$

$$\leq f(x^*; \text{Max}_{y \in Y(x^*)} g(y))$$

$$\leq \operatorname{Max}_{x \in X_1} f(x; \operatorname{Max}_{y \in Y(x)} g(y))$$

$$\leq \text{右辺.}$$

(ロ)  $x^* \in X_2$  ならば,  $f(x^*; \cdot)$  が減少だから

$$\text{左辺} = f(x^*; g(y^*))$$

$$\leq f(x^*; \operatorname{Min}_{y \in Y(x^*)} g(y))$$

$$\leq \operatorname{Max}_{x \in X_2} f(x; \operatorname{Min}_{y \in Y(x)} g(y))$$

$$\leq \text{右辺.}$$

ゆえに 左辺  $\leq$  右辺 となり, 等号が成立する。■

### §3. 逐次決定過程

N段確定的逐次決定過程  $\mathcal{S}$  は次の6つの要素で決

まる:

$$\mathcal{S} = ( \text{Opt}, \{S_n\}_1^{N+1}, \{A_n\}_1^N, \{g_n\}_1^N, k, \{T_n\}_1^N ) \quad (2)$$

ただし

(i)  $S_n$ : 空でない集合, 第n状態空間,  $S_n \ni \Delta_n$

(ii)  $A_n$ : 空でない集合, 第n決定空間,  $A_n \ni a_n$

$$A_n: S_n \longrightarrow 2^{A_n} \quad (\emptyset \neq A_n(\Delta_n) \subset A_n) \quad A_n(\Delta_n): \text{状態 } \Delta_n \text{ における}$$

$$\underline{A_n(\Delta_n)} = \underline{A_n'(\Delta_n) + A_n''(\Delta_n)} \quad \Delta_n \in S_n, 1 \leq n \leq N$$

(iii)  $g_n: S_n \times A_n \times R^1 \rightarrow R^1$ , 第  $n$  利得関数

単調性  $g_n(s_n, a_n; \cdot)$  は  $a_n \in A'_n(s_n)$  のとき 狭義増加,  
 $a_n \in A''_n(s_n)$  のとき 狭義減少

(iv)  $k: S_{N+1} \rightarrow R^1$ , 終端利得関数

(v)  $T_m: S_m \times A_m \rightarrow S_{m+1}$ , 第  $m$  状態変換

(vi)  $\text{Opt}$  は  $\mathcal{S}$  が 次の逐次最適化問題を表現している  
 ことを意味する 最適子 である:

$$\text{Opt } g_1(s_1, a_1; g_2(s_2, a_2; \dots; g_N(s_N, a_N; k(s_{N+1}))) \dots) \quad (2.1)$$

$$\text{s.t. (i) } T_m(s_m, a_m) = s_{m+1} \quad 1 \leq m \leq N$$

(2.2)

$$\text{(ii) } a_m \in A_m(s_m) \quad 1 \leq m \leq N$$

$H_N \equiv S_1 \times A_1 \times \dots \times S_N \times A_N \times S_{N+1}$  を 行動空間,  $H_N \ni$   
 $h_N = (s_1, a_1, \dots, s_N, a_N, s_{N+1})$  を ( $s_1$  からの) 行動 という。  
 $\mu_m: S_m \rightarrow A_m$ ,  $\mu_m(s_m) \in A_m(s_m)$ ,  $s_m \in S_m$ ,  $1 \leq m \leq N$  から成る  
 関数列  $\{\mu_1, \mu_2, \dots, \mu_N\}$  を 戦略 という。戦略  $\{\mu_1^*, \mu_2^*,$   
 $\dots, \mu_N^*\}$  と各  $s_1$  から定まる行動  $h_N^* = (s_1, a_1^*, s_2^*, \dots,$   
 $s_N^*, a_N^*, s_{N+1}^*)$  が, 各  $s_1$  について問題 (2.1), (2.2) の最適  
 値に到達するとき,  $\{\mu_1^*, \mu_2^*, \dots, \mu_N^*\}$  を  $\mathcal{S}$  の 最適戦略  
 という。ここに  $a_1^* = \mu_1^*(s_1)$ ,  $s_2^* = T_1(s_1, a_1^*)$ ,  $\dots$ ,  $s_N^* =$   
 $T_{N-1}(s_{N-1}^*, a_{N-1}^*)$ ,  $a_N^* = \mu_N^*(s_N^*)$ ,  $s_{N+1}^* = T_N(s_N^*, a_N^*)$ . このと

き  $k_n^*$  を  $\Delta_n$  からの 最適行動 という。

$N$  段過程  $\mathcal{S}$  が与えられたとき, 各  $n$  ( $1 \leq n \leq N$ ) に  
対して  $(N-n+1)$  部分過程  $\mathcal{S}_{N-n+1}$ ,  $\overline{\mathcal{S}}_{N-n+1}$  を次で定義する:

$$\mathcal{S}_{N-n+1} = (\text{Opt}, \{S_m\}_n^{N+1}, \{A_m\}_m^N, \{g_m\}_n^N, k, \{T_m\}_m^N) \quad (3)$$

$$\overline{\mathcal{S}}_{N-n+1} = (\overline{\text{Opt}}, \{S_m\}_n^{N+1}, \{A_m\}_m^N, \{g_m\}_n^N, k, \{T_m\}_m^N) \quad (4)$$

$\mathcal{S}_{N-n+1}$ ,  $\overline{\mathcal{S}}_{N-n+1}$  はそれぞれ最適化問題

$$\text{Opt} \quad g_n(\Delta_n, a_n; \dots; g_N(\Delta_N, a_N; k(\Delta_{N+1}))) \dots \quad (3.1)$$

$$\text{s.t. (1) } T_m(\Delta_m, a_m) = \Delta_{m+1} \quad (2) \quad a_m \in A_m(\Delta_m) \quad n \leq m \leq N \quad (3.2)$$

$$\overline{\text{Opt}} \quad g_n(\Delta_n, a_n; \dots; g_N(\Delta_N, a_N; k(\Delta_{N+1}))) \dots \quad (4.1)$$

$$\text{s.t. (1) } T_m(\Delta_m, a_m) = \Delta_{m+1} \quad (2) \quad a_m \in A_m(\Delta_m) \quad n \leq m \leq N \quad (4.2)$$

を表現している。これらの最適値をそれぞれ  $u^{N-n}(\Delta_n)$ ,  $v^{N-n+1}(\Delta_n)$  で表す。特に  $u^0(\Delta_{N+1}) = v^0(\Delta_{N+1}) = k(\Delta_{N+1})$  とする。他方  $n=1$  のとき,  $\mathcal{S}_N, \overline{\mathcal{S}}_N$  はそれぞれ  $\mathcal{S}, \overline{\mathcal{S}}$  に一致する (対応する 6 つの要素が一致する)。よって問題 (1.1), (1.2) の optimum value は  $u^N(\Delta_1)$  で与えられる。

列  $\{u^0, u^1, \dots, u^N\}$  と  $\{v^0, v^1, \dots, v^N\}$  を併せた  $\{u^0, u^1, \dots, u^N; v^0, v^1, \dots, v^N\}$  を逐次決定過程  $\mathcal{S}$

(または  $\bar{S}$ ) の 最適利得関数列 という。2つの戦略  $\{\pi_1, \pi_2, \dots, \pi_N\}$ ,  $\{\sigma_1, \sigma_2, \dots, \sigma_N\}$  を併せた  $\{\pi_1, \pi_2, \dots, \pi_N; \sigma_1, \sigma_2, \dots, \sigma_N\}$  を  $S$  (または  $\bar{S}$ ) の 政策 という。最適政策は再帰式を導いた後に定義する。再帰式は次のようになる:

定理 2 (再帰式)

$$u^{N-n+1}(s_n) = \underset{a_n \in A'_n(s_n)}{\text{Opt}} \left( \underset{a_n \in A'_n(s_n)}{\text{Opt}} g_n(s_n, a_n; u^{N-n}(T_n(s_n, a_n))) \right. \\ \left. , \underset{a_n \in A''_n(s_n)}{\text{Opt}} g_n(s_n, a_n; v^{N-n}(T_n(s_n, a_n))) \right) \quad (5)$$

$$v^{N-n+1}(s_n) = \overline{\underset{a_n \in A'_n(s_n)}{\text{Opt}}} \left( \overline{\underset{a_n \in A'_n(s_n)}{\text{Opt}}} g_n(s_n, a_n; v^{N-n}(T_n(s_n, a_n))) \right. \\ \left. , \overline{\underset{a_n \in A''_n(s_n)}{\text{Opt}}} g_n(s_n, a_n; u^{N-n}(T_n(s_n, a_n))) \right) \quad (6)$$

$$s_n \in S_n, 1 \leq n \leq N$$

$$u^0(s_{N+1}) = v^0(s_{N+1}) = k(s_{N+1}) \quad s_{N+1} \in S_{N+1}. \quad (7)$$

(証明) 基本定理を用いる。[2], [4] の一般化になる。■

(5) の  $\text{Opt}$  に到達する  $a_n$  を  $\pi_n^*(s_n)$  とし, (6) の  $\overline{\text{Opt}}$  に到達する  $a_n$  を  $\hat{\sigma}_n(s_n)$  とする  $s_n \in S_n, 1 \leq n \leq N$ . これらを

分とする政策  $\{\pi_1^*, \pi_2^*, \dots, \pi_N^*; \hat{\sigma}_1, \hat{\sigma}_2, \dots, \hat{\sigma}_N\}$  を  $\mathcal{S}$  (または  $\overline{\mathcal{S}}$ ) の 最適政策 という。このようにして (一意とは限らないが) 定義される  $\pi_n^*(\Delta_n)$ ,  $\hat{\sigma}_m(\Delta_n)$  を次のように表す。

$$\pi_n^*(\Delta_n) = \text{Arg}[(5)の右辺] \quad (8)$$

$$\Delta_n \in S_n \quad 1 \leq n \leq N$$

$$\hat{\sigma}_m(\Delta_n) = \text{Arg}[(6)の右辺] \quad (9)$$

すなわち

$$y^*(x) = \text{Arg} \underset{y \in B(x)}{\text{Opt}} f(x; y)$$

は,  $y^*(x) \in B(x)$  かつ  $B(x)$  上で  $f(x; \cdot)$  の optimum に attain することの意味するものとする。

このとき  $\pi_n^*(\Delta_n)$ ,  $\hat{\sigma}_m(\Delta_n)$  を,

$$\pi_n^{*'}(\Delta_n) = \text{Arg} \underset{a_n \in A_n'(\Delta_n)}{\text{Opt}} g_n(\Delta_n, a_n; u^{N-n}(T_n(\Delta_n, a_n))) \quad (10)$$

$$\pi_n^{*''}(\Delta_n) = \text{Arg} \underset{a_n \in A_n''(\Delta_n)}{\text{Opt}} g_n(\Delta_n, a_n; v^{N-n}(T_n(\Delta_n, a_n))) \quad (11)$$

$$\hat{\sigma}_m'(\Delta_n) = \text{Arg} \overline{\text{Opt}} \underset{a_n \in A_n'(\Delta_n)}{g_n(\Delta_n, a_n; v^{N-n}(T_n(\Delta_n, a_n)))} \quad (12)$$

$$\hat{\sigma}_m''(\Delta_n) = \text{Arg} \overline{\text{Opt}} \underset{a_n \in A_n''(\Delta_n)}{g_n(\Delta_n, a_n; u^{N-n}(T_n(\Delta_n, a_n)))} \quad (13)$$

$$\Delta_n \in S_n \quad 1 \leq n \leq N$$



を用いて

$$\pi_n^*(\Delta_n) = \text{Opt}(\pi_n^{*'}(\Delta_n), \pi_n^{*''}(\Delta_n)) \quad (14)$$

$$\Delta_n \in S_n \quad 1 \leq n \leq N$$

$$\hat{\sigma}_n^*(\Delta_n) = \overline{\text{Opt}}(\hat{\sigma}_n^{*'}(\Delta_n), \hat{\sigma}_n^{*''}(\Delta_n)) \quad (15)$$

と書くことにする。

### § 3.1 最適政策からの $\mathcal{J}$ の最適戦略の構成

再帰式 (5), (6), (7) を  $u^0, v^0$  から  $u^1, v^1, \dots, u^N, v^N$  というように解くと、最適利得関数列  $\{u^0, u^1, \dots, u^N; v^0, v^1, \dots, v^N\}$  と最適政策  $\{\pi_1^*, \pi_2^*, \dots, \pi_N^*; \hat{\sigma}_1, \hat{\sigma}_2, \dots, \hat{\sigma}_N\}$  が求まる。このとき (5) - (15) が成立している。 $\mathcal{J}$  が表現している問題 (1.1), (1.2) の optimum value は  $u^N(\Delta_1)$  で与えられる。 $\mathcal{J}$  の最適戦略  $\{\mu_1^*, \mu_2^*, \dots, \mu_N^*\}$  は次のように決定すればよい:

$$\mu_1^*(\Delta_1) \equiv \pi_1^*(\Delta_1) \quad \Delta_1 \in S_1 \quad (17)$$

$$\mu_2^*(\Delta_2) \equiv \begin{cases} \pi_2^*(\Delta_2) & \text{if } \Delta_2 = T_1(\Delta_1, \pi_1^{*'}(\Delta_1)) \text{ for } \Delta_1 \in S_1 \\ \hat{\sigma}_2^*(\Delta_2) & \text{if } \Delta_2 = T_1(\Delta_1, \pi_1^{*''}(\Delta_1)) \text{ for } \Delta_1 \in S_1 \\ \text{any } \in A_2(\Delta_2) & \text{otherwise} \end{cases} \quad (18)$$

$$\mu_{n+1}^*(\delta_{n+1}) \equiv \begin{cases} \pi_{n+1}^*(\delta_{n+1}) & \text{if } \delta_1 \xrightarrow{\mu_1^* \cdots \mu_{n-1}^* \mu_n^* = \pi_n^* \vee \hat{\sigma}_n''} \delta_{n+1} & \text{for } \delta_1 \in S_1 \\ \hat{\sigma}_{n+1}(\delta_{n+1}) & \text{if } \delta_1 \xrightarrow{\mu_1^* \cdots \mu_{n-1}^* \mu_n^* = \hat{\sigma}_n' \vee \pi_n''} \delta_{n+1} & \text{for } \delta_1 \in S_1 \\ \text{any } \in A_{n+1}(\delta_{n+1}) & \text{otherwise} \end{cases} \quad (19)$$

ただし \* は,  $\delta_1$  から出発して第  $n$  段まで  $\{\mu_1^*, \dots, \mu_{n-1}^*, \mu_n^* = \pi_n^* \text{ or } \hat{\sigma}_n''\}$  に従って行動した結果, 第  $(n+1)$  段では  $\delta_{n+1}$  になったことを示す。可なりゆち

$$T_1(\delta_1, \mu_1^*(\delta_1)) = \delta_2^*, \quad T_2(\delta_2^*, \mu_2^*(\delta_2^*)) = \delta_3^*, \quad \dots, \\ T_{n-1}(\delta_{n-1}^*, \mu_{n-1}^*(\delta_{n-1}^*)) = \delta_n^*, \quad T_n(\delta_n^*, \mu_n^*(\delta_n^*)) = \delta_{n+1} \quad (20)$$

$$\text{or } T_n(\delta_n^*, \hat{\sigma}_n''(\delta_n^*)) = \delta_{n+1}$$

である。

他方  $\bar{S}$  が表現する問題の optimum value は  $v^N(\delta_1)$  で与えられ,  $\bar{S}$  の最適戦略  $\{\hat{\nu}_1, \hat{\nu}_2, \dots, \hat{\nu}_N\}$  は, (17) - (20) を用いて  $\{\mu_1^*, \mu_2^*, \dots, \mu_N^*\}$  を決定したのと, 同様に構成される。

$S, \bar{S}$  の最適戦略  $\{\mu_1^*, \mu_2^*, \dots, \mu_N^*\}, \{\hat{\nu}_1, \hat{\nu}_2, \dots, \hat{\nu}_N\}$  が求められれば,  $S, \bar{S}$  における  $\delta_1$  からの最適行動

$$h_N^* = (\delta_1, a_1^*, \delta_2^*, \dots, \delta_N^*, a_N^*, \delta_{N+1}^*) \quad (21)$$

$$\hat{h}_N = (\delta_1, \hat{a}_1, \hat{\delta}_2, \dots, \hat{\delta}_N, \hat{a}_N, \hat{\delta}_{N+1}) \quad (22)$$

は最適戦略を用いてそれぞれ次のように与えられる。

$$a_1^* = \mu_1^*(s_1), s_2^* = T_1(s_1, a_1^*), \dots, a_N^* = \mu_N^*(s_N^*), s_{N+1}^* = T_N(s_N^*, a_N^*) \quad (23)$$

$$\hat{a}_1 = \hat{\nu}_1(s_1), \hat{s}_2 = T_1(s_1, \hat{a}_1), \dots, \hat{a}_N = \hat{\nu}_N(\hat{s}_N), \hat{s}_{N+1} = T_N(\hat{s}_N, \hat{a}_N) \quad (24)$$

これらの構成方法の妥当性は再帰式より明らかである。

### § 3.2 状態に依存しない利得関数をもつ決定過程

(2) で与えた逐次決定過程  $\mathcal{S} = (\mathcal{O}_{pt}, \{S_n\}_1^{N+1}, \{A_m\}_1^N, \{g_n\}_1^N, k, \{T_m\}_1^N)$  における第  $m$  利得関数  $g_m: S_m \times A_m \times R^1 \rightarrow R^1$  は状態  $s_m \in S_m$  にも依存している。しかし、状態  $s_m$  に依存しない第  $m$  利得関数  $g'_m$  を用いて ( $\mathcal{S}$  が表現する問題 (2.1), (2.2) と同じ問題を表現するという意味で) 同値な  $\mathcal{S}'$  を次のように構成することができる:

$$\mathcal{S}' = (\mathcal{O}_{pt}, \{S_n\}_1^{N+1}, \{A'_m\}_1^N, \{g'_m\}_1^N, k, \{T_m\}_1^N) \quad (25)$$

ここに

$$A'_m \equiv S_m \times A_m, A'_m(s_m) \equiv \{(s_m, a_m) \mid a_m \in A_m(s_m)\} \quad a'_m = (s_m, a_m)$$

$$g'_m: A'_m \times R^1 \rightarrow R^1$$

$$g'_m(a'_m; r) \equiv g_m(s_m, a_m; r)$$

$$T'_m: S_m \times A'_m \rightarrow S_{m+1}$$

$$T'_m(s_m, a'_m) \equiv T_m(s_m, a_m)$$

このとき  $\mathcal{S}'$  は  $\mathcal{S}$  と同じ問題を表現している。したがって、§3 の  $N$  段逐次決定過程  $\mathcal{S}$  の 第  $n$  利得関数は状態  $s_n$  に依存しない と仮定してよい。このことは一般に、確定的であれ確率的であれまた有限段であれ無限段であれ、すべての逐次決定過程についてもいえる。この考え方の基本は 対(状態, 決定)をこの状態における新しい決定と見做すことにある。しかし計算量の観点からすれば、この reduction を行っても最適解(政策・利得関数列・行動)を求める計算量は不変である。

#### §4. 逆過程

まず次の記号を導入して、逆命題を準備する。

$$\mathcal{V} \equiv \{f \mid f: (0, \infty) \rightarrow (0, \infty) \text{ 上への狭義増加関数}\}$$

$$\mathcal{D} \equiv \{f \mid f: (0, \infty) \rightarrow (0, \infty) \text{ 上への狭義減少関数}\}$$

$$\mathcal{M} \equiv \mathcal{V} \cup \mathcal{D} = \mathcal{V} + \mathcal{D}$$

$A$ : 空でない(パラメータ)集合

#### 補題 3 (逆命題)

$$A = A' + A'' \text{ とし, } k \in \mathcal{V}, f^a, T_a \in \mathcal{V} \text{ for } a \in A'$$

$f^a, T_a \in \mathcal{D}$  for  $a \in A''$  とする。このとき、 $\text{Opt}_{a \in A'}(f^a \circ k \circ T_a)^*$

$\text{Opt}_{a \in A''}(f^a \circ k \circ T_a)$  から  $(0, \infty)$  から  $(0, \infty)$  の上への関数とならば、関係式

$$\text{Opt}_{a \in A}(f^a \circ k \circ T_a) = \text{Opt}_{a \in A'}(\text{Opt}_{a \in A'}(f^a \circ k \circ T_a), \text{Opt}_{a \in A''}(f^a \circ k \circ T_a)) \quad (26)$$

$$\begin{aligned} \left(\text{Opt}_{a \in A}(f^a \circ k \circ T_a)\right)^{-1} &= \overline{\text{Opt}_{a \in A'}\left(\overline{\text{Opt}_{a \in A''}\left((T_a)^{-1} \circ k^{-1} \circ (f^a)^{-1}\right)}\right)}, \\ &\quad \overline{\text{Opt}_{a \in A''}\left((T_a)^{-1} \circ k^{-1} \circ (f^a)^{-1}\right)} \end{aligned} \quad (27)$$

が成立する。ただし \* は

$$u(x) \equiv \text{Opt}_{\substack{a' = T(x, a) \\ a \in A'}} f(a, k(x))$$

で定義される関数  $u: (0, \infty) \rightarrow (0, \infty)$  を意味する。ここに

$$f(a, k(x)) = f^a(k(x)) \quad , \quad T(x, a) = T_a(x).$$

(証明) [3] の Lemma 2 の一般化である。■

次に逆過程を導入するために、一次元状態空間上の  $N$  段過程を利得空間を明示しながら定義する。 $N$  段逐次決定過程  $\mathcal{S}$  は 7 つの要素

$$\mathcal{S} = (\text{Opt}, \{S_m\}_1^{N+1}, \{R_n\}_1^{N+1}, \{A_n\}_1^N, \{f_m\}_1, k, \{T_m\}_1^N) \quad (28)$$

で決定される。ただし

$$(i) \quad S_m : \underline{R^1 \text{ の 区 間}} \quad (29)$$

$$(ii) \quad R_n : \underline{R^1 \text{ の 区 間}}, \quad \underline{\text{第 } n \text{ 利得空間}}, \quad R_m \ni r_m \quad (30)$$

$$(iii) \quad \underline{A_m = A'_n + A''_m} \quad \underline{A_m(\lambda_n) \equiv A_n} \quad (31)$$

$$(iv) \quad f_m : A_m \times R_{n+1} \longrightarrow R_m$$

$$f_m^{a_n} \equiv f_m(a_n; \cdot) \begin{cases} \in \mathcal{J}(R_{n+1}, R_m) & \text{if } a_n \in A'_n \\ \in \mathcal{D}(R_{n+1}, R_m) & \text{if } a_n \in A''_n \end{cases} \quad (32)$$

$$\begin{cases} \in \mathcal{D}(R_{n+1}, R_m) & \text{if } a_n \in A''_n \end{cases} \quad (33)$$

ただし

$$\mathcal{J}(X, Y) \equiv \{f \mid f: X \longrightarrow Y \text{ 上への狭義増加関数}\}$$

$$\mathcal{D}(X, Y) \equiv \{f \mid f: X \longrightarrow Y \text{ 上への狭義減少関数}\}$$

$$(v) \quad k : S_{N+1} \longrightarrow R_{N+1}$$

$$k \in \mathcal{J}(S_{N+1}, R_{N+1})$$

$$(vi) \quad T_m : S_m \times A_n \longrightarrow S_{n+1}$$

$$T_{ma_n} \equiv T_m(\cdot, a_n) \begin{cases} \in \mathcal{J}(S_n, S_{n+1}) & \text{if } a_n \in A'_n \\ \in \mathcal{D}(S_n, S_{n+1}) & \text{if } a_n \in A''_n \end{cases} \quad (34)$$

(vii)  $\text{Opt}$  は最適化問題

$$\text{Opt } f_1(a_1; f_2(a_2; \dots; f_N(a_N; k(S_{N+1})) \dots)) \quad (28.1)$$

$$\text{s.t. } (1) \quad T_m(\lambda_n, a_n) = \lambda_{n+1} \quad 1 \leq n \leq N$$

$$(2) \quad a_n \in A_m \quad 1 \leq m \leq N \quad (28.2)$$

この問題は次のようにも書ける:

$$\text{Opt } f_1^{a_1} \circ f_2^{a_2} \circ \dots \circ f_N^{a_N} \circ k \circ T_{Na_N} \circ \dots \circ T_{2a_2} \circ T_{1a_1}(\Delta_1) \quad (28.3)$$

$$\text{s.t. (1) } a_n \in A_n \quad 1 \leq n \leq N \quad (28.4)$$

$\mathcal{S}$  を 主逐次決定過程 という。これに対する再帰式は次のようになる。

系 5 (主過程に対する再帰式)

$$u^{N-n+1}(\Delta_n) = \text{Opt}_{a_n \in A'_n} \left( \text{Opt}_{a_n \in A''_n} (f_n^{a_n} \circ u^{N-n} \circ T_{ma_n}(\Delta_n)) \right) \quad (29)$$

$$, \text{Opt}_{a_n \in A''_n} (f_n^{a_n} \circ v^{N-n} \circ T_{ma_n}(\Delta_n))$$

$$v^{N-n+1}(\Delta_n) = \overline{\text{Opt}}_{a_n \in A'_n} \left( \overline{\text{Opt}}_{a_n \in A''_n} (f_n^{a_n} \circ v^{N-n} \circ T_{ma_n}(\Delta_n)) \right) \quad (30)$$

$$, \overline{\text{Opt}}_{a_n \in A''_n} (f_n^{a_n} \circ u^{N-n} \circ T_{ma_n}(\Delta_n))$$

$$\Delta_n \in S_n, \quad 1 \leq n \leq N$$

$$u^0(\Delta_{N+1}) = v^0(\Delta_{N+1}) = k(\Delta_{N+1}) \quad (31)$$

#### §4.1 1段過程への帰着

与えられた  $\mathcal{S}$  は  $N$  段過程であるが、これを同値な 1段過程  $\mathcal{S}''$  に帰着させることができる:

$$\mathcal{S}'' = (\text{Opt}, \{S_1, S_{N+1}\}, \{R_1, R_{N+1}\}, A, f, k, T) \quad (32)$$

ここに

$$A \equiv A_1 \times A_2 \times \cdots \times A_N \quad a = (a_1, a_2, \dots, a_N) \quad (33)$$

$$f: A \times R_{N+1} \rightarrow R_1$$

$$f(a, r_{N+1}) \equiv f_1^{a_1} \circ f_2^{a_2} \circ \cdots \circ f_N^{a_N}(r_{N+1}) \quad (34)$$

$$T: S_1 \times A \rightarrow S_{N+1}$$

$$T(s_1, a) \equiv T_{Na_N} \circ \cdots \circ T_{2a_2} \circ T_{1a_1}(s_1) \quad (35)$$

しかも  $A = A' + A''$  なる  $A', A''$  も定まると

$$T_a \begin{cases} \in \mathcal{N}(S_1, S_{N+1}) & \text{if } a \in A' \\ \in \mathcal{D}(S_1, S_{N+1}) & \text{if } a \in A'' \end{cases}$$

となるためには,

$$A' \equiv \{a \in A \mid f^a \in \mathcal{N}(R_{N+1}, R_1)\} \quad (36)$$

$$A'' \equiv \{a \in A \mid f^a \in \mathcal{D}(R_{N+1}, R_1)\} \quad (37)$$

とすればよい。このとき  $\mathcal{S}''$  は  $\mathcal{S}$  と同じ最適化問題を表現している。

## §4.2 逆過程の構成と逆定理

$N$  段主過程  $\mathcal{S}$  が与えられたとき、 $\mathcal{S}$  の 逆逐次決定過程  $\mathcal{S}^{-1}$  を次の 7 つの要素で定める:

$$\mathcal{S}^{-1} = (\overline{\mathcal{Q}_{pt}}, \{R_n\}_1^{N+1}, \{S_n\}_1^{N+1}, \{A_n\}_1^N, \{g_n\}_1^N, l, \{U_n\}_1^N) \quad (38)$$

ここに

$$g_n: A_n \times S_{n+1} \rightarrow S_n$$



$$g_n(a_n; s_{n+1}) \equiv (T_{na_n})^{-1}(s_{n+1}) \quad (39)$$

$$l: R_{N+1} \longrightarrow S_{N+1}$$

$$l(r_{N+1}) \equiv k^{-1}(r_{N+1}) \quad (40)$$

$$U_n: R_n \times A_n \longrightarrow R_{n+1}$$

$$U_n(r_n, a_n) \equiv (f_n^{a_n})^{-1}(r_{n+1}). \quad (41)$$

逆過程  $\mathcal{J}^{-1}$  は最適化問題

$$\overline{\text{Opt}} \quad g_1(a_1; g_2(a_2; \dots; g_N(a_N; l(r_{N+1}))) \dots)) \quad (38.1)$$

$$\text{s.t. (1) } U_n(r_n, a_n) = r_{n+1} \quad 1 \leq n \leq N \quad (38.2)$$

$$(2) \quad a_n \in A_n \quad 1 \leq n \leq N$$

可なり

$$\overline{\text{Opt}} \quad g_1^{a_1} \circ g_2^{a_2} \circ \dots \circ g_N^{a_N} \circ l \circ U_{Na_n} \circ \dots \circ U_{2a_2} \circ U_{1a_1}(r_1) \quad (38.3)$$

$$\text{s.t. (1) } a_n \in A_n \quad 1 \leq n \leq N \quad (38.4)$$

を表現している。

$\mathcal{J}^{-1}$  (または  $\overline{\mathcal{J}^{-1}}$ ) の最適利得関数列を  $\{p^0, p^1, \dots, p^N; g^0, g^1, \dots, g^N\}$ , 最適政策を  $\{\hat{\lambda}_1, \hat{\lambda}_2, \dots, \hat{\lambda}_N; \xi_1^*, \xi_2^*, \dots, \xi_N^*\}$  とする。次の再帰式が成立する:

系 6 (逆過程に対する再帰式)

$$p^{N-n}(r_n) = \overline{\text{Opt}}_{a_n \in A_n} \left( \overline{\text{Opt}}_{a_n \in A_n} (g_n^{a_n} \circ p^{N-n} \circ U_{na_n}(r_n)) \right. \\ \left. , \overline{\text{Opt}}_{a_n \in A_n} (g_n^{a_n} \circ g^{N-n} \circ U_{na_n}(r_n)) \right) \quad (42)$$

$$\begin{aligned}
 g^{N-n+1}(r_n) &= \underset{a_n \in A'_n}{\text{Opt}} \left( \underset{a_n \in A'_n}{\text{Opt}} (g_n^{a_n} \circ g^{N-n} \circ U_{na_n}(r_n)) \right. \\
 &\quad \left. , \underset{a_n \in A''_n}{\text{Opt}} (g_n^{a_n} \circ p^{N-n} \circ U_{na_n}(r_n)) \right) \\
 r_n &\in R_n \quad 1 \leq n \leq N
 \end{aligned} \tag{43}$$

$$p^0(r_{N+1}) = f^0(r_{N+1}) = l(r_{N+1}) \quad r_{N+1} \in R_{N+1} \tag{44}$$

$\hat{\lambda}_m(r_m), \hat{\xi}_m^*(r_m)$  はそれぞれ (42), (43) の optimum, optimum に到達する  $a_m$  を表す:

$$\hat{\lambda}_m(r_m) = \text{Arg}[(42) \text{の右辺}] \quad r_m \in R_m \quad 1 \leq m \leq N \tag{45}$$

$$\hat{\xi}_m^*(r_m) = \text{Arg}[(43) \text{の右辺}] \tag{46}$$

定理 7 (逆定理) (i) 主逐次決定過程  $\mathcal{S}$  が上への最適利得関数列  $\{u^0, u^1, \dots, u^N; v^0, v^1, \dots, v^N\}$  と最適政策  $\{\pi_1^*, \pi_2^*, \dots, \pi_N^*; \hat{\delta}_1, \hat{\delta}_2, \dots, \hat{\delta}_N\}$  をもつならば, 逆逐次決定過程  $\mathcal{S}^{-1}$  は上への最適利得関数列  $\{(u^0)^{-1}, (u^1)^{-1}, \dots, (u^N)^{-1}; (v^0)^{-1}, (v^1)^{-1}, \dots, (v^N)^{-1}\}$  と最適政策  $\{\pi_1^* \circ (u^N)^{-1}, \pi_2^* \circ (u^{N-1})^{-1}, \dots, \pi_N^* \circ (u^1)^{-1}; \hat{\delta}_1 \circ (v^N)^{-1}, \hat{\delta}_2 \circ (v^{N-1})^{-1}, \dots, \hat{\delta}_N \circ (v^1)^{-1}\}$  をもつ。

(ii) 逆逐次決定過程  $\mathcal{S}^{-1}$  が上への最適利得関数列

$\{p^0, p^1, \dots, p^N; q^0, q^1, \dots, q^N\}$  と最適政策  $\{\hat{\lambda}_1, \hat{\lambda}_2, \dots, \hat{\lambda}_N; \xi_1^*, \xi_2^*, \dots, \xi_N^*\}$  をもつならば, 主逐次決定過程  $\mathcal{S}$  は上への最適利得関数列  $\{(p^0)^{-1}, (p^1)^{-1}, \dots, (p^N)^{-1}; (q^0)^{-1}, (q^1)^{-1}, \dots, (q^N)^{-1}\}$  と最適政策  $\{\hat{\lambda}_1 \circ (p^N)^{-1}, \hat{\lambda}_2 \circ (p^{N-1})^{-1}, \dots, \hat{\lambda}_N \circ (p^1)^{-1}; \xi_1^* \circ (q^N)^{-1}, \xi_2^* \circ (q^{N-1})^{-1}, \dots, \xi_N^* \circ (q^1)^{-1}\}$  をもつ。

(証明) [3] の Th. 4 (Inverse Theorem) の一般化である。■

[注意] 最適利得関数列は狭義増加になっている。

上へのとは

$$\begin{aligned} u^{N-n+1}: S_n &\longrightarrow R_n, & v^{N-n+1}: S_n &\longrightarrow R_n \\ p^{N-n+1}: R_m &\longrightarrow S_m, & q^{N-n+1}: R_m &\longrightarrow S_m \end{aligned} \quad 1 \leq n \leq N+1 \quad (47)$$

が onto であることをいう。[1] では狭義単調になっている。

## §5. 反転過程

§4 の主逐次決定過程  $\mathcal{S}$  の反転過程は定義できるが, 一般に反転定理 (Reverse Theorem) は成立しない。しかし特別な 2 つの場合については成立する。

$$\mathcal{S} = (\mathcal{Q}_{pt}, \{S_n\}_1^{N+1}, \{R_n\}_1^{N+1}, \{A_n\}_1^N, \{f_n\}_1^N, k, \{T_n\}_1^N) \quad (28)$$

を (28) で定義される主過程とする。これに対する反転過程

$\mathcal{S}_{-1}$  を

$$\mathcal{S}_{-1} = (\overline{\mathcal{O}}_{pt}, \{\bar{S}_m\}_1^{N+1}, \{\bar{R}_m\}_1^{N+1}, \{\bar{A}_m\}_1^N, \{\bar{f}_m\}_1^N, u^N, \{\bar{T}_m\}_1^N) \quad (48)$$

で定義する。左だし  $u^N \in \mathcal{U}(\mathcal{S}_1, R_1)$  とし、残りの要素を次で与える：

$$\bar{S}_m \equiv S_{N-n+2}, \quad \bar{R}_m \equiv R_{N-n+2}$$

$$\bar{A}_m \equiv A_{N-n+1}, \quad \bar{A}'_m \equiv A'_{N-n+1}, \quad \bar{A}''_m \equiv A''_{N-n+1}$$

$$\bar{f}_m : \bar{A}_m \times \bar{R}_{m+1} \rightarrow \bar{R}_m \quad (49)$$

$$\bar{f}_m(\bar{a}_m; \bar{r}_{m+1}) \equiv (f_{N-n+1}^{a_{N-n+1}})^{-1}(r_{N-n+1})$$

$$\bar{T}_m : \bar{S}_m \times \bar{A}_m \rightarrow \bar{S}_{m+1}$$

$$\bar{T}_m(\bar{s}_m, \bar{a}_m) \equiv (T_{N-n+1}^{a_{N-n+1}})^{-1}(s_{N-n+2}).$$

反転過程  $\mathcal{S}_{-1}$  を 主過程  $\mathcal{S}$  の要素を用いて以下

$$\mathcal{S}_{-1} = (\overline{\mathcal{O}}_{pt}, \{S_m\}'_{N+1}, \{R_m\}'_{N+1}, \{A_m\}'_N, \{(f_m^{a_m})^{-1}\}'_N, u^N, \{(T_m^{a_m})^{-1}\}'_N)$$

で表すことにする。 $\mathcal{S}_{-1}$  は  $\mathcal{S}_{N+1}$  を初期状態とし  $\mathcal{S}_1$  を終端状態とする次の最適化問題を表現している：

$$\overline{\mathcal{O}}_{pt} (f_N^{a_N})^{-1} \circ (f_{N-1}^{a_{N-1}})^{-1} \circ \dots \circ (f_1^{a_1})^{-1} \circ u^N \circ (T_{1a_1})^{-1} \circ \dots \circ (T_{N-1a_{N-1}})^{-1} \circ (T_{Na_N})^{-1}(s_{N+1}) \quad (48.1)$$

$$s.t. \quad (i) \quad a_m \in A_m \quad N \geq m \geq 1 \quad (48.2)$$

$N \geq m \geq 1$  は時間から  $N+1, N, \dots, 1$  の順に逆進することを意味する。

反転過程  $S_{-1}$  の最適利得関数列を  $\{x^0, x^1, \dots, x^N;$   
 $y^0, y^1, \dots, y^N\}$  とし, 最適政策を  $\{\hat{\lambda}_1, \hat{\lambda}_2, \dots, \hat{\lambda}_N; \mu_1^*, \mu_2^*$   
 $, \dots, \mu_N^*\}$  とすれば, 次の再帰式が成立する。

系 8 (反転過程に対する再帰式)

$$\begin{aligned} x^{N-n+1}(\delta_{N-n+2}) &= \overline{\text{Opt}}_{a_{N-n+1} \in A'_{N-n+1}} \left( \overline{\text{Opt}}_{a_{N-n+1} \in A'_{N-n+1}} \left( (f_{N-n+1}^{a_{N-n+1}})^{-1} \circ x^{N-n} \circ (T_{N-n+1} a_{N-n+1})^{-1} (\delta_{N-n+2}) \right) \right. \\ &\quad \left. , \overline{\text{Opt}}_{a_{N-n+1} \in A''_{N-n+1}} \left( (f_{N-n+1}^{a_{N-n+1}})^{-1} \circ y^{N-n} \circ (T_{N-n+1} a_{N-n+1})^{-1} (\delta_{N-n+2}) \right) \right) \end{aligned} \quad (50)$$

$$\begin{aligned} y^{N-n+1}(\delta_{N-n+2}) &= \text{Opt}_{a_{N-n+1} \in A'_{N-n+1}} \left( \text{Opt}_{a_{N-n+1} \in A'_{N-n+1}} \left( (f_{N-n+1}^{a_{N-n+1}})^{-1} \circ y^{N-n} \circ (T_{N-n+1} a_{N-n+1})^{-1} (\delta_{N-n+2}) \right) \right. \\ &\quad \left. , \text{Opt}_{a_{N-n+1} \in A''_{N-n+1}} \left( (f_{N-n+1}^{a_{N-n+1}})^{-1} \circ x^{N-n} \circ (T_{N-n+1} a_{N-n+1})^{-1} (\delta_{N-n+2}) \right) \right) \end{aligned} \quad (51)$$

$$\delta_{N-n+2} \in S_{N-n+2}, \quad 1 \leq n \leq N$$

$$x^0(\delta_1) = y^0(\delta_1) = u^N(\delta_1) \quad \delta_1 \in S_1. \quad (52)$$

最適政策は次を満足している:

$$\hat{\lambda}_n(\delta_{N-n+1}) = \text{Arg} [(50) \text{の右辺}] \quad (53)$$

$$\xi_n^*(\delta_{N-n+2}) = \text{Arg} [(51) \text{の右辺}] \quad (54)$$

系5 すなわち 主過程に対する再帰式 (29), (30) において

Opt = Max の場合を考えると

$$u^{N-n+1}(\Delta_n) = \text{Max}_{a_n \in A'_n} (f_n^{a_n} \circ u^{N-n} \circ T_{na_n}(\Delta_n)) \vee \text{Max}_{a_n \in A''_n} (f_n^{a_n} \circ v^{N-n} \circ T_{na_n}(\Delta_n)) \quad (55)$$

$$v^{N-n+1}(\Delta_n) = \text{Min}_{a_n \in A'_n} (f_n^{a_n} \circ v^{N-n} \circ T_{na_n}(\Delta_n)) \wedge \text{Min}_{a_n \in A''_n} (f_n^{a_n} \circ u^{N-n} \circ T_{na_n}(\Delta_n)) \quad (56)$$

$$\Delta_n \in S_n \quad 1 \leq n \leq N$$

$$u^0(\Delta_{N+1}) = v^0(\Delta_{N+1}) = k(\Delta_{N+1}) \quad \Delta_{N+1} \in S_{N+1}$$

になる。ただし  $a \vee b = \max(a, b)$ ,  $a \wedge b = \min(a, b)$ 。

このとき,  $\{u^0, u^1, \dots, u^N; v^0, v^1, \dots, v^N\}$  が (55), (56) を満たす関数列ならば, 不等式

$$u^{N-n}(\Delta_{n+1}) \leq \text{Min}_{a_n \in A'_n} ((f_n^{a_n})^{-1} \circ u^{N-n+1} \circ (T_{na_n})^{-1}(\Delta_{n+1})) \wedge \text{Min}_{a_n \in A''_n} ((f_n^{a_n})^{-1} \circ v^{N-n+1} \circ (T_{na_n})^{-1}(\Delta_{n+1})) \quad (59)$$

$$v^{N-n}(\Delta_{n+1}) \geq \text{Max}_{a_n \in A'_n} ((f_n^{a_n})^{-1} \circ v^{N-n+1} \circ (T_{na_n})^{-1}(\Delta_{n+1})) \vee \text{Max}_{a_n \in A''_n} ((f_n^{a_n})^{-1} \circ u^{N-n+1} \circ (T_{na_n})^{-1}(\Delta_{n+1})) \quad (60)$$

$$\Delta_{N-n+1} \in S_{N-n+1} \quad 1 \leq n \leq N$$

は成立するが, 一般に (59), (60) における等式は成立しない。

反例 次の1段階過程  $\mathcal{S} = (\text{Max}, \{S_1, S_2\}, \{R_1, R_2\}, A_1, f_1, k, T_1)$  を考える。ただし

$$S_1 = S_2 = R_1 = R_2 = \mathbb{R}^1 (= (-\infty, \infty))$$

$$f_1 = f, S_1 = S, S_2 = S', R_1 = R, R_2 = R', A_1 = A,$$

$$T_1 = T, k(s') = s', A = A' + A'', A' = \{1, 3\}, A'' = \{2, 4\}$$

$$f(1; r') = 1 + r', \quad f(3; r') = r'/3$$

$$f(2; r') = 2 - r', \quad f(4; r') = -r'/4$$

$$T(s, 1) = -1 + s, \quad T(s, 3) = s/3$$

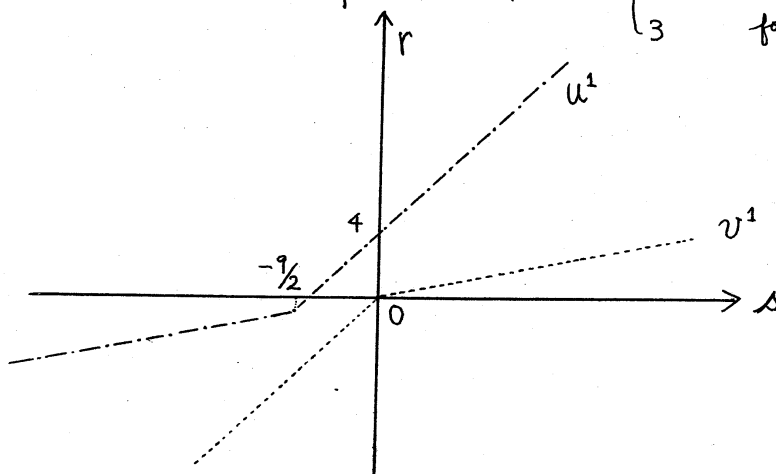
$$T(s, 2) = -2 - s, \quad T(s, 4) = -4s.$$

このとき最適利得関数列  $\{u^0, u^1; v^0, v^1\}$  と最適政策  $\{\pi_1^*; \hat{\sigma}_1\}$  は次のようになる:

$$u^0(s') = v^0(s') = s'$$

$$u^1(s) = \frac{s}{9} \vee (4 + s), \quad \pi_1^*(s) = \begin{cases} 3 & \text{for } s \leq -\frac{9}{2} \\ 2 & \text{for } s > -\frac{9}{2} \end{cases}$$

$$v^1(s) = s \wedge \left(\frac{s}{9}\right), \quad \hat{\sigma}_1(s) = \begin{cases} 1 \text{ or } 4 & \text{for } s \leq 0 \\ 3 & \text{for } s > 0 \end{cases}$$

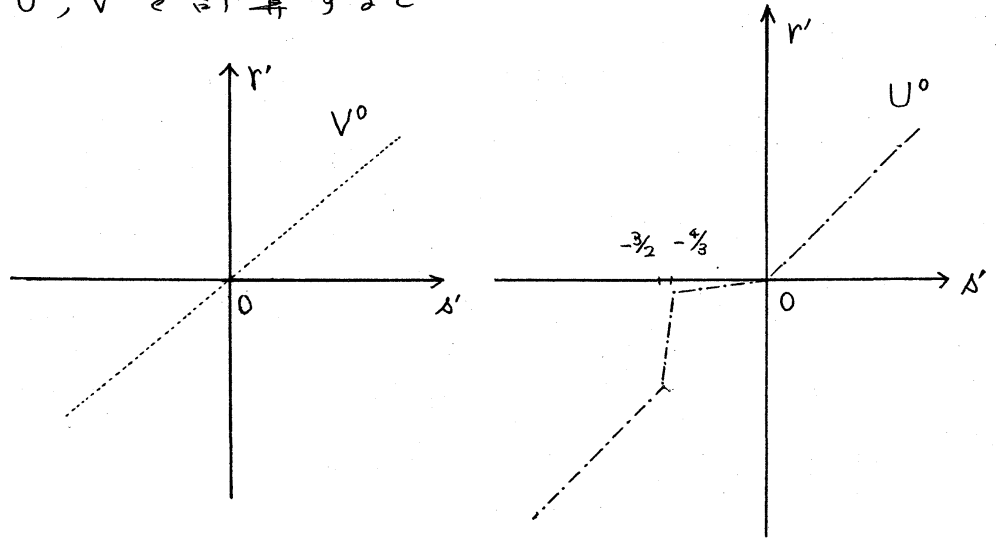


他方

$$M_{\min}^{a=1,3} (f^a)^{-1} \circ u^1 \circ (T_a)^{-1}(s') \wedge M_{\min}^{a=2,4} (f^a)^{-1} \circ v^1 \circ (T_a)^{-1}(s') \equiv U^0(s')$$

$$M_{\max}^{a=1,3} (f^a)^{-1} \circ v^1 \circ (T_a)^{-1}(s') \vee M_{\max}^{a=2,4} (f^a)^{-1} \circ u^1 \circ (T_a)^{-1}(s') \equiv V^0(s')$$

とあって、 $U^0, V^0$  を計算すると



$$V^0(s') = s' \quad U^0(s') = \begin{cases} s' & s' \geq 0 \\ \frac{s'}{4} & -\frac{4}{3} \leq s' < -\frac{3}{2} \\ 12 + s' & -\frac{3}{2} \leq s' < -\frac{4}{3} \\ s' & s' < -\frac{4}{3} \end{cases}$$

したがってすべての実数  $s'$  については

$$u^0(s') = U^0(s'), \quad v^0(s') = V^0(s')$$

は成立しない。この場合は

$$V^0 = v^0 = u^0 = U^0 \quad \text{for } s' \geq 0 \text{ or } s' \leq -\frac{3}{2}$$

$$V^0 = v^0 = u^0 < U^0 \quad \text{for } -\frac{3}{2} < s' < 0$$

になっている。■



上述の反例からもわかるように、 $\mathcal{S}, \mathcal{S}_1$  に対しては次に述べるような反転定理は一般に成立しない。しかし各  $n$  について  $A'_n = \emptyset$  または  $A''_n = \emptyset$  のときは成立する:

定理 9 (反転定理) 各  $n$  ( $1 \leq n \leq N$ ) について

$A_n = A'_n$  または  $A_n = A''_n$  とする。(i) 主過程  $\mathcal{S}$  が最適利得関数列  $\{u^0, u^1, \dots, u^N; v^0, v^1, \dots, v^N\}$  と最適政策  $\{\pi_1^*, \pi_2^*, \dots, \pi_N^*; \hat{\sigma}_1, \hat{\sigma}_2, \dots, \hat{\sigma}_N\}$  をもつならば, 反転過程  $\mathcal{S}_1$  は最適利得関数列  $\{u^N, u^{N-1}, \dots, u^0; v^N, v^{N-1}, \dots, v^0\}$  と最適政策  $\{\pi_N^* \circ (T_N \pi_N^*)^{-1}, \pi_{N-1}^* \circ (T_{N-1} \pi_{N-1}^*)^{-1}, \dots, \pi_1^* \circ (T_1 \pi_1^*)^{-1}; \hat{\sigma}_N^* \circ (T_N \hat{\sigma}_N^*)^{-1}, \hat{\sigma}_{N-1}^* \circ (T_{N-1} \hat{\sigma}_{N-1}^*)^{-1}, \dots, \hat{\sigma}_1^* \circ (T_1 \hat{\sigma}_1^*)^{-1}\}$  をもつ。ただし  $u^N: S_1 \rightarrow R_1$  は上への,  $T_m \pi_m^*, T_m \hat{\sigma}_m^*: S_m \rightarrow S_{m+1}$  は上への 1対1 写像とする。

(ii) 反転過程  $\mathcal{S}_1$  が最適利得関数列  $\{x^0, x^1, \dots, x^N; y^0, y^1, \dots, y^N\}$  と最適政策  $\{\hat{\nu}_1, \hat{\nu}_2, \dots, \hat{\nu}_N; \xi_1^*, \xi_2^*, \dots, \xi_N^*\}$  をもつならば, 主過程  $\mathcal{S}$  は最適利得関数列  $\{x^N, x^{N-1}, \dots, x^0; y^N, y^{N-1}, \dots, y^0\}$  と最適政策  $\{\hat{\nu}_N^* \circ (T_N \hat{\nu}_N^*)^{-1}, \hat{\nu}_{N-1}^* \circ (T_{N-1} \hat{\nu}_{N-1}^*)^{-1}, \dots, \hat{\nu}_1^* \circ (T_1 \hat{\nu}_1^*)^{-1}; \xi_N^* \circ (T_N \xi_N^*)^{-1}, \xi_{N-1}^* \circ (T_{N-1} \xi_{N-1}^*)^{-1}, \dots, \xi_1^* \circ (T_1 \xi_1^*)^{-1}\}$  をもつ。ただし  $u^N: S_1 \rightarrow R_1$  は上への,  $T_m \hat{\nu}_m^*, T_m \xi_m^*: S_{N-m+2} \rightarrow S_{N-m+1}$  は上への 1対1 写像とする。

(注)  $T_m \pi_m^*(\Delta_n) \equiv T_m \pi_m^*(\Delta_n)(\Delta_n) = T_m(\Delta_n, \pi_m^*(\Delta_n))$ . 他も同様.

(証明) [3] の Th. 5 (Reverse Theorem) の一般化である。■

### §5. 動的計画

§2 において  $X_2 = \emptyset$ , §3 において  $A''_n(x_n) \equiv \emptyset$ , §4 において  $A'' = \emptyset$ ,  $A''_n \equiv \emptyset$  なる場合に対応して, それぞれの結果が得られる。こゝき 逐次決定過程  $\mathcal{S}$  は 動的計画 であるといひ,  $\mathcal{S}$  の代りに  $\mathcal{D}$  で表す。したがって 逆動的計画  $\mathcal{D}^{-1}$ , 反転動的計画  $\mathcal{D}_{-1}$  が導入される ([3] 参照)。

### 参考文献

- [1] S. IWAMOTO, A class of inverse theorem on recursive programming with monotonicity, J. Operations Res. Soc. Japan, 20 (1977), 94-112.
- [2] S. IWAMOTO, The second principle of optimality, Bull. Math. Statist., 17(1977), 101-114.
- [3] S. IWAMOTO, Some operations on dynamic programming with one-dimensional state space, J. Math. Anal. Appl., to appear.
- [4]. G.L. NEMHAUSER, "Introduction to Dynamic Programming", Wiley, New York, 1966.