数理解析研究所講究録
第358巻 1979年 193-218

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 $g_n(s_n, a_n; \cdot)$ が n, s_n, a_n に応じて・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・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

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

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

$g : Y \longrightarrow R^1$

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

* 記号 Y を集合と点対集合値写像の両方に用いている。

$\underline{X = X_1 + X_2}$

<u>単調性</u>　$f(x; \cdot) : R^1 \longrightarrow R^1$ は $x \in X_1$ のとき狭義増加，$x \in X_2$ のとき狭義減少。

$Opt = Max$ または Min のいずれか一方

$\overline{Opt} = Min\ (Max)$, $Opt = Max\ (Min)$ のとき

このとき，次の関係式が成立する。

定理1 （基本定理）

$$\underset{(x,y)\in G_r(Y)}{Opt} f(x\,;\,g(y)) = Opt\left(\underset{x\in X_1}{Opt} f(x\,;\,\underset{y\in Y(x)}{Opt}\, g(y)),\ \underset{x\in X_2}{Opt} f(x\,;\,\underset{y\in Y(x)}{\overline{Opt}}\, g(y))\right) \tag{1}$$

ただし

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

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

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

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

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

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

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

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

すなわち 左辺 $\geqq$ 右辺．

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

$$\text{左辺} = f(x^*\,;\,g(y^*))$$

$$\leqq f(x^*\,;\,\underset{y\in Y(x^*)}{Max}\, g(y))$$

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

$\leqq$ 右辺。

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

$$左辺 = f(x^* ; g(y^*))$$

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

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

$\leqq$ 右辺。

ゆえに　左辺 $\leqq$ 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：

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

ただし

(i) S_n：空でない集合，第 n 状態空間，$S_n \ni s_n$

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

$A_n : S_n \longrightarrow 2^{A_n}$　$(\ \phi \neq A_n(s_n) \subset A_n\ )$　$A_n(s_n)$：状態 s_n における可能な決定空間

$A_n(s_n) = A_n'(s_n) + A_n''(s_n)$　$s_n \in S_n$，$1 \leqq n \leqq N$

(iii) $g_n : S_n \times A_n \times R^1 \longrightarrow R^1$ ，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) $k : S_{N+1} \longrightarrow R^1$，終端利得関数

(v) $T_n : S_n \times A_n \longrightarrow S_{n+1}$，第 n 状態変換

(vi) 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：

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

$$\text{s.t.}\ \ \text{(i)}\ \ T_n(s_n, a_n) = s_{n+1} \quad 1 \leqq n \leqq N \tag{2.2}$$

$$\text{(ii)}\ \ a_n \in A_n(s_n) \quad 1 \leqq n \leqq N$$

$H_N \equiv S_1 \times A_1 \times \cdots \times S_N \times A_N \times S_{N+1}$ を行動空間，$H_N \ni h_N = (s_1, a_1, \cdots, s_N, a_N, s_{N+1})$ を（s_1 からの）行動という。$\mu_n : S_n \longrightarrow A_n$ ，$\mu_n(s_n) \in A_n(s_n)$ ，$s_n \in S_n$，$1 \leqq n \leqq N$ から成る関数列 $\{\mu_1, \mu_2, \cdots, \mu_N\}$ を戦略という。戦略 $\{\mu_1^*, \mu_2^*, \cdots, \mu_N^*\}$ と各 s_1 から定まる行動 $h_N^* = (s_1, a_1^*, s_2^*, \cdots, s_N^*, a_N^*, s_{N+1}^*)$ 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，$\{\mu_1^*, \mu_2^*, \cdots, \mu_N^*\}$ を $\mathcal{S}$ の最適戦略という。ここに $a_1^* = \mu_1^*(s_1)$, $s_2^* = T_1(s_1, a_1^*)$, $\cdots$, $s_N^* = T_{N-1}(s_{N-1}^*, a_{N-1}^*)$ ，$a_N^* = \mu_N^*(s_N^*)$, $s_{N+1}^* = T_N(s_N^*, a_N^*)$. このと

き h_N^* を s_1 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 $(1 \leqq n \leqq N)$ 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：

$$\mathcal{S}_{N-n+1} = \left(\mathcal{O}pt, \{S_m\}_n^{N+1}, \{A_m\}_n^N, \{g_m\}_n^N, k, \{T_m\}_n^N \right) \quad (3)$$

$$\overline{\mathcal{S}}_{N-n+1} = \left(\overline{\mathcal{O}pt}, \{S_m\}_n^{N+1}, \{A_m\}_n^N, \{g_m\}_n^N, k, \{T_m\}_n^N \right) \quad (4)$$

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

$$\mathcal{O}pt \quad g_n(s_n, a_n; \cdots; g_N(s_N, a_N; k(s_{N+1}))\cdots) \quad (3.1)$$

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

$$\overline{\mathcal{O}pt} \quad g_n(s_n, a_n; \cdots; g_N(s_N, a_N; k(s_{N+1}))\cdots) \quad (4.1)$$

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

を表現している。これらの最適値をそれぞれ $u^{N-n+1}(s_n)$，$v^{N-n+1}(s_n)$ で表す。特に $u^0(s_{N+1}) = v^0(s_{N+1}) = k(s_{N+1})$ 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 $u^N(s_1)$ で与えられる。

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

（または $\overline{\mathcal{S}}$）の最適利得関数列という。2つの戦略 $\{\pi_1, \pi_2, \cdots, \pi_N\}$, $\{\sigma_1, \sigma_2, \cdots, \sigma_N\}$ を併せた $\{\pi_1, \pi_2, \cdots, \pi_N ; \sigma_1, \sigma_2, \cdots, \sigma_N\}$ を $\mathcal{S}$（または $\overline{\mathcal{S}}$）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）

$$u^{N-n+1}(s_n) = \mathrm{Opt}\Big(\mathop{\mathrm{Opt}}_{a_n \in A'_n(s_n)} g_n(s_n, a_n ; u^{N-n}(T_n(s_n, a_n))) , \mathop{\mathrm{Opt}}_{a_n \in A''_n(s_n)} g_n(s_n, a_n ; v^{N-n}(T_n(s_n, a_n)))\Big) \tag{5}$$

$$v^{N-n+1}(s_n) = \overline{\mathrm{Opt}}\Big(\mathop{\overline{\mathrm{Opt}}}_{a_n \in A'_n(s_n)} g_n(s_n, a_n ; v^{N-n}(T_n(s_n, a_n))) , \mathop{\overline{\mathrm{Opt}}}_{a_n \in A''_n(s_n)} g_n(s_n, a_n ; u^{N-n}(T_n(s_n, a_n)))\Big) \tag{6}$$

$$s_n \in S_n, \quad 1 \leqq n \leqq N$$

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

（証明）基本定理を用いる。[2],[4] の一般化になる。■

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

分とする政策 $\{\pi_1^*, \pi_2^*, \cdots, \pi_N^*;\ \hat{\sigma}_1, \hat{\sigma}_2, \cdots, \hat{\sigma}_N\}$ 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 $\pi_n^*(s_n)$, $\hat{\sigma}_n(s_n)$ を次のように表す。

$$\pi_n^*(s_n) = \mathrm{Arg}\,[(5)\text{の右辺}] \tag{8}$$

$$s_n \in S_n \quad 1 \leqq n \leqq N$$

$$\hat{\sigma}_n(s_n) = \mathrm{Arg}\,[(6)\text{の右辺}] \tag{9}$$

すなわち

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

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

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

$$\pi_n^{*\prime}(s_n) = \mathrm{Arg}\ \underset{a_n \in A_n'(s_n)}{\mathrm{Opt}}\ g_n(s_n, a_n;\ u^{N-n}(T_n(s_n, a_n))) \tag{10}$$

$$\pi_n^{*\prime\prime}(s_n) = \mathrm{Arg}\ \underset{a_n \in A_n''(s_n)}{\mathrm{Opt}}\ g_n(s_n, a_n;\ v^{N-n}(T_n(s_n, a_n))) \tag{11}$$

$$\hat{\sigma}_n'(s_n) = \mathrm{Arg}\ \underset{a_n \in A_n'(s_n)}{\overline{\mathrm{Opt}}}\ g_n(s_n, a_n;\ v^{N-n}(T_n(s_n, a_n))) \tag{12}$$

$$\hat{\sigma}_n''(s_n) = \mathrm{Arg}\ \underset{a_n \in A_n''(s_n)}{\overline{\mathrm{Opt}}}\ g_n(s_n, a_n;\ u^{N-n}(T_n(s_n, a_n))) \tag{13}$$

$$s_n \in S_n \quad 1 \leqq n \leqq N$$

を用いて

$$\pi_n^*(s_n) = \mathcal{O}pt\,(\pi_n^{*\prime}(s_n),\ \pi_n^{*\prime\prime}(s_n)) \qquad (14)$$

$$s_n \in S_n \quad 1 \leqq n \leqq N$$

$$\hat{\sigma}_n(s_n) = \overline{\mathcal{O}pt}\,(\hat{\sigma}_n^{\prime}(s_n),\ \hat{\sigma}_n^{\prime\prime}(s_n)) \qquad (15)$$

と書くことにする。

§ 3.1 最適政策からの $\mathcal{S}$ の最適戦略の構成

再帰式 (5), (6), (7) を u^0, v^0 から $u^1, v^1, \cdots, u^N, v^N$ というように解くと，最適利得関数列 $\{u^0, u^1, \cdots, u^N; v^0, v^1, \cdots, v^N\}$ と最適政策 $\{\pi_1^*, \pi_2^*, \cdots, \pi_N^*; \hat{\sigma}_1, \hat{\sigma}_2, \cdots, \hat{\sigma}_N\}$ が求まる。このとき (5) ― (15) が成立している。$\mathcal{S}$ が表現している問題 (1.1), (1.2) の optimum value は $u^N(s_1)$ で与えられる。$\mathcal{S}$ の最適戦略 $\{\mu_1^*, \mu_2^*, \cdots, \mu_N^*\}$ は次のように決定すればよい：

$$\mu_1^*(s_1) \equiv \pi_1^*(s_1) \qquad s_1 \in S_1 \qquad (17)$$

$$\mu_2^*(s_2) \equiv \begin{cases} \pi_2^*(s_2) & \text{if } s_2 = T_1(s_1, \pi_1^{*\prime}(s_1)) \quad \text{for } s_1 \in S_1 \\ \hat{\sigma}_2(s_2) & \text{if } s_2 = T_1(s_1, \pi_1^{*\prime\prime}(s_1)) \quad \text{for } s_1 \in S_1 \\ \text{any} \in A_2(s_2) & \text{otherwise} \end{cases} \qquad (18)$$

$$\mu^*_{n+1}(s_{n+1}) \equiv \begin{cases} \pi^*_{n+1}(s_{n+1}) & \text{if } s_1 \xrightarrow{\mu^*_1 \cdots \mu^*_{n-1}\ \mu^*_n = \pi^{*\prime}_n \vee \hat{\sigma}^{\prime\prime}_n} s_{n+1}^{\;\star} \quad \text{for } s_1 \in S_1 \\ \hat{\sigma}_{n+1}(s_{n+1}) & \text{if } s_1 \xrightarrow{\mu^*_1 \cdots \mu^*_{n-1}\ \mu^*_n = \hat{\sigma}^{\prime}_n \vee \pi^{*\prime\prime}_n} s_{n+1} \quad \text{for } s_1 \in S_1 \\ \text{any} \in A_{n+1}(s_{n+1}) & \text{otherwise} \end{cases} \tag{19}$$

ただし ☆ は，s_1 から出発して第 n 段まで $\{\mu^*_1, \cdots, \mu^*_{n-1}, \mu^*_n = \pi^{*\prime}_n$ or $\hat{\sigma}^{\prime\prime}_n\}$ に従って行動した結果，第 $(n+1)$ 段では s_{n+1} になったことを示す。すなわち

$$T_1(s_1, \mu^*_1(s_1)) = s^*_2, \quad T_2(s^*_2, \mu^*_2(s^*_2)) = s^*_3, \quad \cdots,$$

$$T_{n-1}(s^*_{n-1}, \mu^*_{n-1}(s^*_{n-1})) = s^*_n, \quad T_n(s^*_n, \pi^{*\prime}_n(s^*_n)) = s_{n+1} \tag{20}$$

$$\text{or} \quad T_n(s^*_n, \hat{\sigma}^{\prime\prime}_n(s^*_n)) = s_{n+1}$$

である。

他方 $\overline{\mathcal{S}}$ が表現する問題の optimum value は $v^N(s_1)$ で与えられ，$\overline{\mathcal{S}}$ の最適戦略 $\{\hat{\nu}_1, \hat{\nu}_2, \cdots, \hat{\nu}_N\}$ は，(17)—(20) を用いて $\{\mu^*_1, \mu^*_2, \cdots, \mu^*_N\}$ を決定したのと，同様に構成される。

$\mathcal{S}$, $\overline{\mathcal{S}}$ の最適戦略 $\{\mu^*_1, \mu^*_2, \cdots, \mu^*_N\}$, $\{\hat{\nu}_1, \hat{\nu}_2, \cdots, \hat{\nu}_N\}$ が求められれば，$\mathcal{S}$, $\overline{\mathcal{S}}$ における s_1 からの最適行動

$$h^*_N = (s_1, a^*_1, s^*_2, \cdots, s^*_N, a^*_N, s^*_{N+1}) \tag{21}$$

$$\hat{h}_N = (s_1, \hat{a}_1, \hat{s}_2, \cdots, \hat{s}_N, \hat{a}_N, \hat{s}_{N+1}) \tag{22}$$

は最適戦略を用いてそれぞれ次のように与えられる。

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

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

これらの構成方法の妥当性は再帰式より明らかであろう。

§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 $\mathcal{S} = (\mathrm{Opt}, \{S_n\}_1^{N+1}, \{A_n\}_1^N, \{g_n\}_1^N, k, \{T_n\}_1^N)$ における第 n 利得関数 $g_n : S_n \times A_n \times R^1 \longrightarrow R^1$ は状態 $s_n \in S_n$ にも依存している。しかし，状態 s_n に依存しない第 n 利得関数 g_n' 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：

$$\mathcal{S}' = (\mathrm{Opt}, \{S_n\}_1^{N+1}, \{A_n'\}_1^N, \{g_n'\}_1^N, k, \{T_n'\}_1^N) \tag{25}$$

ここに

$$A_n' \equiv S_n \times A_n, \quad A_n'(s_n) \equiv \{(s_n, a_n) \mid a_n \in A_n(s_n)\} \qquad a_n' = (s_n, a_n)$$

$$g_n' : A_n' \times R^1 \longrightarrow R^1$$

$$g_n'(a_n'; r) \equiv g_n(s_n, a_n; r)$$

$$T_n' : S_n \times A_n' \longrightarrow S_{n+1}$$

$$T_n'(s_n, a_n') \equiv T_n(s_n, a_n)$$

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.

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

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

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

A：空でない（パラメーター）集合

補題 3（逆命題）

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

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

$\underset{a\in A''}{Opt}(f^a \circ k \circ T_a)$ が $(0, \infty)$ から $(0, \infty)$ の上への関数ならば，関係式

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

$$\left(\underset{a\in A}{Opt}(f^a \circ k \circ T_a)\right)^{-1} = \overline{Opt}\left(\underset{a\in A'}{\overline{Opt}}\left((T_a)^{-1} \circ k^{-1} \circ (f^a)^{-1}\right), \underset{a\in A''}{\overline{Opt}}\left((T_a)^{-1} \circ k^{-1} \circ (f^a)^{-1}\right)\right) \quad (27)$$

が成立する。ただし ※ は

$$u(s) = \underset{\substack{s'=T(s,a)\\ a\in A'}}{Opt} f(a, k(s'))$$

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

$$f(a, k(s')) = f^a(k(s')) \quad , T(s, a) = T_a(s).$$

（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。<u>N段逐次決定過程 $\mathcal{S}$</u> は７つの要素

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

で決定される。ただし

(i) S_n : $\underline{\underline{R^1 \text{ の区間}}}$ (29)

(ii) R_n : $\underline{\underline{R^1 \text{ の区間}}}$, $\underline{\text{第 } n \text{ 利得空間}}$, $R_n \ni r_n$ (30)

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

(iv) $f_n : A_n \times R_{n+1} \longrightarrow R_n$

$$f_n^{a_n} \equiv f_n(a_n ; \cdot) \begin{cases} \in \mathcal{J}(R_{n+1}, R_n) & \text{if} \quad a_n \in A'_n \quad (32) \\ \in \mathcal{D}(R_{n+1}, R_n) & \text{if} \quad a_n \in A''_n \quad (33) \end{cases}$$

ただし

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

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

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

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

(vi) $T_n : S_n \times A_n \longrightarrow S_{n+1}$

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

(vii) $\mathcal{O}pt$ は最適化問題

$$\mathcal{O}pt \quad f_1(a_1 ; f_2(a_2 ; \cdots ; f_N(a_N ; k(s_{N+1})) \cdots)) \tag{28.1}$$

$$\begin{aligned} \text{s.t.} \quad &(1) \quad T_n(s_n, a_n) = s_{n+1} \qquad 1 \leqq n \leqq N \\ &(2) \quad a_n \in A_n \qquad 1 \leqq n \leqq N \end{aligned} \tag{28.2}$$

この問題は次のようにも書ける：

$$\mathrm{Opt}\ f_1^{a_1}\circ f_2^{a_2}\circ\cdots\circ f_N^{a_N}\circ k\circ T_{Na_N}\circ\cdots\circ T_{2a_2}\circ T_{1a_1}(s_1) \tag{28.3}$$

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

$\mathcal{S}$ 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）

$$u^{N-n+1}(s_n) = \mathrm{Opt}\Big(\mathop{\mathrm{Opt}}_{a_n\in A'_n}\big(f_n^{a_n}\circ u^{N-n}\circ T_{na_n}(s_n)\big),\ \mathop{\mathrm{Opt}}_{a_n\in A''_n}\big(f_n^{a_n}\circ v^{N-n}\circ T_{na_n}(s_n)\big)\Big) \tag{29}$$

$$v^{N-n+1}(s_n) = \overline{\mathrm{Opt}}\Big(\mathop{\overline{\mathrm{Opt}}}_{a_n\in A'_n}\big(f_n^{a_n}\circ v^{N-n}\circ T_{na_n}(s_n)\big),\ \mathop{\overline{\mathrm{Opt}}}_{a_n\in A''_n}\big(f_n^{a_n}\circ u^{N-n}\circ T_{na_n}(s_n)\big)\Big) \tag{30}$$

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

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

§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：

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

ここに

$$A \equiv A_1 \times A_2 \times \cdots \times A_N \qquad a = (a_1, a_2, \cdots, a_N) \tag{33}$$

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

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

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

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

しかも　$A = A' + A''$ なる A', A'' も定まって

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

となるためには，

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

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

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<u>逆逐次決定過程 $\mathcal{S}^{-1}$</u> を次の7つの要素で定める：

$$\mathcal{S}^{-1} = (\overline{\mathrm{Opt}}, \{R_n\}_1^{N+1}, \{S_n\}_1^{N+1}, \{A_n\}_1^N, \{g_n\}_1^N, l, \{U_n\}_1^N) \tag{38}$$

ここに

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

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

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

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

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

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

逆過程 $\mathcal{S}^{-1}$ は 最適化問題

$$\overline{\mathrm{Opt}} \quad g_1(a_1\,;\,g_2(a_2\,;\,\cdots\,;\,g_N(a_N\,;\,l(r_{N+1}))\cdots)) \tag{38.1}$$

$$\text{s.t.} \quad (1)\ U_n(r_n, a_n) = r_{n+1} \quad 1 \leqq n \leqq N$$
$$(2)\ a_n \in A_n \quad 1 \leqq n \leqq N \tag{38.2}$$

すなわち

$$\overline{\mathrm{Opt}} \quad g_1^{a_1} \circ g_2^{a_2} \circ \cdots \circ g_N^{a_N} \circ l \circ U_{N a_N} \circ \cdots \circ U_{2 a_2} \circ U_{1 a_1}(r_1) \tag{38.3}$$

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

を表現している。

$\mathcal{S}^{-1}$（または $\overline{\mathcal{S}^{-1}}$）の最適利得関数列を $\{p^0, p^1, \cdots, p^N\,;\,q^0, q^1, \cdots, q^N\}$，最適政策を $\{\hat{\lambda}_1, \hat{\lambda}_2, \cdots, \hat{\lambda}_N\,;\,\xi_1^*, \xi_2^*, \cdots, \xi_N^*\}$ 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）

$$p^{N-n+1}(r_n) = \overline{\mathrm{Opt}}\Big(\overline{\mathrm{Opt}}_{a_n \in A'_n}\big(g_n^{a_n} \circ p^{N-n} \circ U_{n a_n}(r_n)\big), \ \overline{\mathrm{Opt}}_{a_n \in A''_n}\big(g_n^{a_n} \circ q^{N-n} \circ U_{n a_n}(r_n)\big)\Big) \tag{42}$$

$$g^{N-n+1}(r_n) = \mathop{Opt}\Big(\mathop{Opt}_{a_n \in A'_n}\big(g_n^{a_n} \circ g^{N-n} \circ U_{na_n}(r_n)\big), \mathop{Opt}_{a_n \in A''_n}\big(g_n^{a_n} \circ p^{N-n} \circ U_{na_n}(r_n)\big)\Big) \tag{43}$$

$$r_n \in R_n \quad 1 \leqq n \leqq N$$

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

$\hat{\lambda}_n(r_n)$, $\xi_n^*(r_n)$ は それぞれ (42), (43) の $\overline{\text{optimum}}$, optimum に到達する a_n を表す：

$$\hat{\lambda}_n(r_n) = \mathrm{Arg}\,[\,(42)\text{の右辺}\,] \tag{45}$$

$$r_n \in R_n \quad 1 \leqq n \leqq N$$

$$\xi_n^*(r_n) = \mathrm{Arg}\,[\,(43)\text{の右辺}\,] \tag{46}$$

<u>定理 7</u> （逆定理） (i) 主逐次決定過程 $\mathcal{S}$ が上への最適利得関数列 $\{u^0, u^1, \cdots, u^N ; v^0, v^1, \cdots, v^N\}$ と最適政策 $\{\pi_1^*, \pi_2^*, \cdots, \pi_N^* ; \hat{\sigma}_1, \hat{\sigma}_2, \cdots, \hat{\sigma}_N\}$ 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 $\{(u^0)^{-1}, (u^1)^{-1}, \cdots, (u^N)^{-1} ; (v^0)^{-1}, (v^1)^{-1}, \cdots, (v^N)^{-1}\}$ と最適政策 $\{\pi_1^* \circ (u^N)^{-1}, \pi_2^* \circ (u^{N-1})^{-1}, \cdots, \pi_N^* \circ (u^1)^{-1} ; \hat{\sigma}_1 \circ (v^N)^{-1}, \hat{\sigma}_2 \circ (v^{N-1})^{-1}, \cdots, \hat{\sigma}_N \circ (v^1)^{-1}\}$ 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

$\{p^0, p^1, \cdots, p^N ; q^0, q^1, \cdots, q^N\}$ と最適政策 $\{\hat{\lambda}_1, \hat{\lambda}_2, \cdots, \hat{\lambda}_N ; \xi_1^*, \xi_2^*, \cdots, \xi_N^*\}$ 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 $\{(p^0)^{-1}, (p^1)^{-1}, \cdots, (p^N)^{-1} ; (q^0)^{-1}, (q^1)^{-1}, \cdots, (q^N)^{-1}\}$ と最適政策 $\{\hat{\lambda}_1 \circ (p^N)^{-1}, \hat{\lambda}_2 \circ (p^{N-1})^{-1}, \cdots, \hat{\lambda}_N \circ (p^1)^{-1} ; \xi_1^* \circ (q^N)^{-1}, \xi_2^* \circ (q^{N-1})^{-1}, \cdots, \xi_N^* \circ (q^1)^{-1}\}$ 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

$$u^{N-n+1} : S_n \longrightarrow R_n, \quad v^{N-n+1} : S_n \longrightarrow R_n$$
$$p^{N-n+1} : R_n \longrightarrow S_n, \quad q^{N-n+1} : R_n \longrightarrow S_n \qquad 1 \leq n \leq N+1 \quad (47)$$

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。

$$\mathcal{S} = (\mathrm{Opt}, \{S_n\}_1^{N+1}, \{R_n\}_1^{N+1}, \{A_n\}_1^N, \{f_n\}_1^N, k, \{T_n\}_1^N) \qquad (28)$$

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

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

で定義する。ただし $u^N \in \mathcal{V}(S_1, R_1)$ とし，残りの要素を次で与える：

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

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

$$\bar{f}_n : \bar{A}_n \times \bar{R}_{n+1} \longrightarrow \bar{R}_n \quad (49)$$

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

$$\bar{T}_n : \bar{S}_n \times \bar{A}_n \longrightarrow \bar{S}_{n+1}$$

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

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

$$\mathcal{S}_{-1} = (Opt, \{S_n\}_{N+1}^1, \{R_n\}_{N+1}^1, \{A_n\}_N^1, \{(f_n^{a_n})^{-1}\}_N^1, u^N, \{(T_{na_n})^{-1}\}_N^1)$$

で表すことにする。$\mathcal{S}_{-1}$ は s_{N+1} を初期状態とし s_1 を終端状態とする次の最適化問題を表現している：

$$\overline{Opt} \; (f_N^{a_N})^{-1} \circ (f_{N-1}^{a_{N-1}})^{-1} \circ \cdots \circ (f_1^{a_1})^{-1} \circ u^N \circ (T_{1a_1})^{-1} \circ \cdots \circ (T_{N-1\,a_{N-1}})^{-1} \circ (T_{Na_N})^{-1}(s_{N+1}) \quad (48.1)$$

$$\text{s.t.} \quad (1) \quad a_n \in A_n \quad N \geqq n \geqq 1 \quad (48.2)$$

$N \geqq n \geqq 1$ は時間が $N+1, N, \cdots, 1$ の順に逆進することを意味する。

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

<u>系 8</u>　（反転過程に対する再帰式）

$$x^{N-n+1}(s_{N-n+2}) = \overline{\mathcal{O}pt}\Big(\underset{a_{N-n+1}\in A'_{N-n+1}}{\overline{\mathcal{O}pt}} \big((f_{N-n+1}^{a_{N-n+1}})^{-1}\circ x^{N-n}\circ(T_{N-n+1\,a_{N-n+1}})^{-1}(s_{N-n+2})\big), \underset{a_{N-n+1}\in A''_{N-n+1}}{\overline{\mathcal{O}pt}} \big((f_{N-n+1}^{a_{N-n+1}})^{-1}\circ y^{N-n}\circ(T_{N-n+1\,a_{N-n+1}})^{-1}(s_{N-n+2})\big)\Big) \tag{50}$$

$$y^{N-n+1}(s_{N-n+2}) = \mathcal{O}pt\Big(\underset{a_{N-n+1}\in A'_{N-n+1}}{\mathcal{O}pt} \big((f_{N-n+1}^{a_{N-n+1}})^{-1}\circ y^{N-n}\circ(T_{N-n+1\,a_{N-n+1}})^{-1}(s_{N-n+2})\big), \underset{a_{N-n+1}\in A''_{N-n+1}}{\mathcal{O}pt} \big((f_{N-n+1}^{a_{N-n+1}})^{-1}\circ x^{N-n}\circ(T_{N-n+1\,a_{N-n+1}})^{-1}(s_{N-n+2})\big)\Big) \tag{51}$$

$$s_{N-n+2}\in S_{N-n+2}, \quad 1\leqq n\leqq N$$

$$x^0(s_1) = y^0(s_1) = u^N(s_1) \qquad s_1\in S_1. \tag{52}$$

最適政策は次を満たしている：

$$\hat{\lambda}_n(s_{N-n+1}) = \mathrm{Arg}\,[(50)\text{の右辺}] \tag{53}$$

$$\xi_n^*(s_{N\,n+2}) = \mathrm{Arg}\,[(51)\text{の右辺}] \tag{54}$$

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

$$u^{N-n+1}(s_n) = \underset{a_n \in A'_n}{Max}\left(f_n^{a_n} \circ u^{N-n} \circ T_{n a_n}(s_n)\right) \vee \underset{a_n \in A''_n}{Max}\left(f_n^{a_n} \circ v^{N-n} \circ T_{n a_n}(s_n)\right) \quad (55)$$

$$v^{N-n+1}(s_n) = \underset{a_n \in A'_n}{Min}\left(f_n^{a_n} \circ v^{N-n} \circ T_{n a_n}(s_n)\right) \wedge \underset{a_n \in A''_n}{Min}\left(f_n^{a_n} \circ u^{N-n} \circ T_{n a_n}(s_n)\right) \quad (56)$$

$$s_n \in S_n \quad 1 \leqq n \leqq N$$

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

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

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

$$u^{N-n}(s_{n+1}) \leqq \underset{a_n \in A'_n}{Min}\left(\left(f_n^{a_n}\right)^{-1} \circ u^{N-n+1} \circ \left(T_{n a_n}\right)^{-1}(s_{n+1})\right)$$
$$\wedge \underset{a_n \in A''_n}{Min}\left(\left(f_n^{a_n}\right)^{-1} \circ v^{N-n+1} \circ \left(T_{n a_n}\right)^{-1}(s_{n+1})\right) \quad (59)$$

$$v^{N-n}(s_{n+1}) \geqq \underset{a_n \in A'_n}{Max}\left(\left(f_n^{a_n}\right)^{-1} \circ v^{N-n+1} \circ \left(T_{n a_n}\right)^{-1}(s_{n+1})\right)$$
$$\vee \underset{a_n \in A''_n}{Max}\left(\left(f_n^{a_n}\right)^{-1} \circ u^{N-n+1} \circ \left(T_{n a_n}\right)^{-1}(s_{n+1})\right) \quad (60)$$

$$s_{N-n+1} \in S_{N-n+1} \quad 1 \leqq n \leqq N$$

は成立するが, <u>一般に (59), (60) における等式は成立しない。</u>

<u>反例</u>　次の1段過程 $\mathcal{S} = (\mathrm{Max}, \{S_1, S_2\}, \{R_1, R_2\}, A_1, f_1, k, T_1)$ を考える。ただし

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

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

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

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

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

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

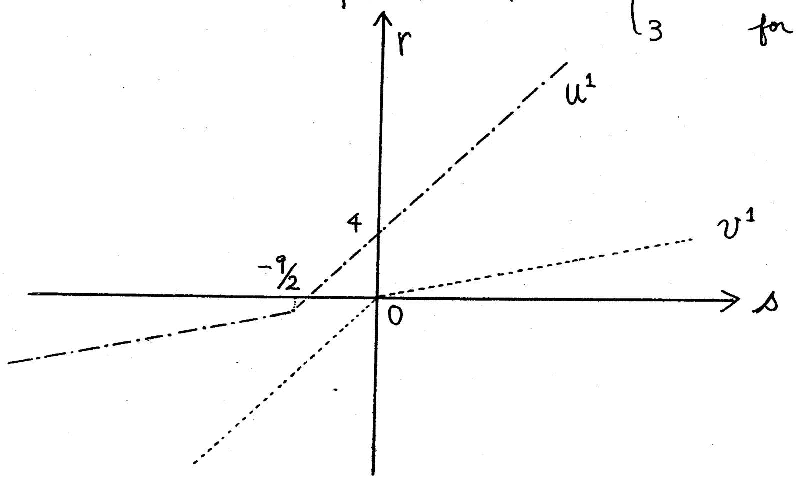
$$T(s, 2) = -2 - s, \qquad T(s, 4) = -4s.$$

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：

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

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

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

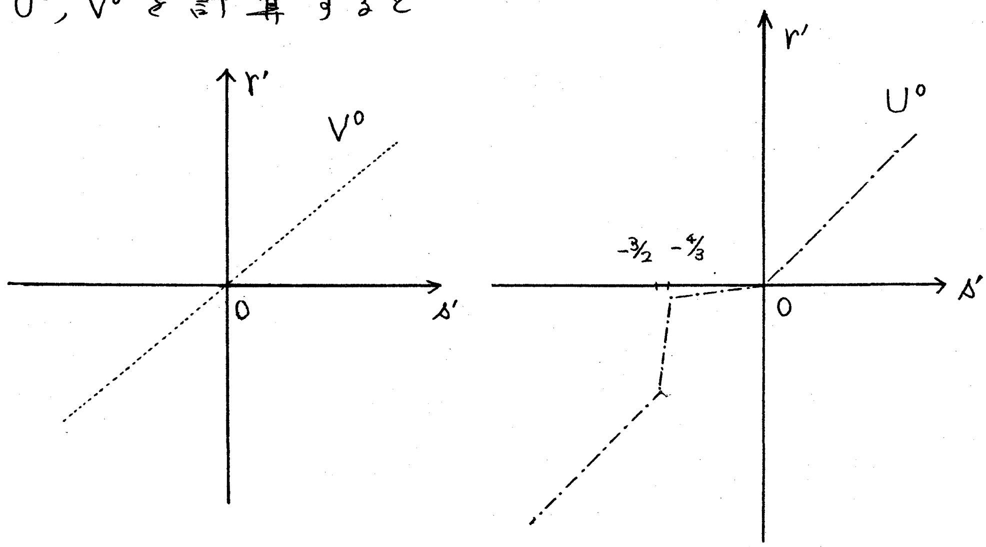


他方

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

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

とおいて，U^0，V^0 を計算すると



$$V^0(s') = s' \qquad U^0(s') = \begin{cases} s' & s' \geq 0 \\ \dfrac{s'}{9} & -\dfrac{4}{3} \leq s' < 0 \\ 12+s' & -\dfrac{3}{2} \leq s' < -\dfrac{4}{3} \\ s' & s' < -\dfrac{3}{2}. \end{cases}$$

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

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

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

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 $A'_n=\phi$ または $A''_n=\phi$ のときは成立する：

<u>定理 9</u>（反転定理）各 n $(1\leqq n\leqq N)$ 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 $\{u^0, u^1, \cdots, u^N; v^0, v^1, \cdots, v^N\}$ と最適政策 $\{\pi_1^*, \pi_2^*, \cdots, \pi_N^*; \hat{\sigma}_1, \hat{\sigma}_2, \cdots, \hat{\sigma}_N\}$ をもつならば，反転過程 $\mathcal{S}_{-1}$ は最適利得関数列 $\{u^N, u^{N-1}, \cdots, u^0; v^N, v^{N-1}, \cdots, v^0\}$ と最適政策 $\{\pi_N^*\circ(T_{N\pi_N^*})^{-1}, \pi_{N-1}^*\circ(T_{N-1\,\pi_{N-1}^*})^{-1}, \cdots, \pi_1^*\circ(T_{1\pi_1^*})^{-1}; \hat{\sigma}_N\circ(T_{N\hat{\sigma}_N})^{-1}, \hat{\sigma}_{N-1}\circ(T_{N-1\,\hat{\sigma}_{N-1}})^{-1}, \cdots, \hat{\sigma}_1\circ(T_{1\hat{\sigma}_1})^{-1}\}$ をもつ。ただし $u^N: S_1\to R_1$ は上への，$T_{n\pi_n^*}$, $T_{n\hat{\sigma}_n}: S_n\longrightarrow S_{n+1}$ 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 $\{x^0, x^1, \cdots, x^N; y^0, y^1, \cdots, y^N\}$ と最適政策 $\{\hat{\nu}_1, \hat{\nu}_2, \cdots, \hat{\nu}_N; \xi_1^*, \xi_2^*, \cdots, \xi_N^*\}$ をもつならば，主過程 $\mathcal{S}$ は最適利得関数列 $\{x^N, x^{N-1}, \cdots, x^0; y^N, y^{N-1}, \cdots, y^0\}$ と最適政策 $\{\hat{\nu}_N\circ(\bar{T}_{N\hat{\nu}_N})^{-1}, \hat{\nu}_{N-1}\circ(\bar{T}_{N-1\,\hat{\nu}_{N-1}})^{-1}, \cdots, \hat{\nu}_1\circ(\bar{T}_{1\hat{\nu}_1})^{-1}; \xi_N^*\circ(\bar{T}_{N\xi_N^*})^{-1}, \xi_{N-1}^*\circ(\bar{T}_{N-1\,\xi_{N-1}^*})^{-1}, \cdots, \xi_1^*\circ(\bar{T}_{1\xi_1^*})^{-1}\}$ をもつ。ただし $u^N: S_1\to R_1$ は上への，$\bar{T}_{n\hat{\nu}_n}$, $\bar{T}_{n\xi_n^*}: S_{N-n+2}\longrightarrow S_{N-n+1}$ は上への1対1写像とする。

(注) $T_{n\pi_n^*}(s_n)\equiv T_{n\,\pi_n^*(s_n)}(s_n)=T_n(s_n, \pi_n^*(s_n))$. 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 $X_2 = \phi$，§3 において $A''_n(s_n) \equiv \phi$，§4 において $A'' = \phi$，$A''_n \equiv \phi$ なる場合に対応して，それぞれの結果が得られる。こときき逐次決定過程 $\mathcal{S}$ は動的計画であるといい，$\mathcal{S}$ 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.