

On Dynamic Programming

九州大学 経済学部

岩本 誠一 Seichi IWAMOTO

1 はじめに

この報告では、動的計画法(Dynamic Programming)を含むより広い逐次最適化手法として両的計画法(Bynamic Programming)を提唱する。動的計画法は単調性(Monotonicity)と可分性(Separability)の下で用いられる[1-3, 8, 10, 12, 16-21]。目的関数の可分性はまた再帰性(Recursiveness)とも呼ばれている[4-13]。単調性は通常、目的関数における「非減少性」を意味している[3-5, 8-13, 16-21]。

この報告では、いわゆる単調性を「非減少性または非増加性のいずれか」と広義に解釈して、これを両調性(Bitonicity)と呼ぶ。すなわち、両的計画法は両調性と再帰性の下での逐次最適化法である。両的計画法は動的計画法の自然な2(両)方向への拡張であり、それを含むものである。両的計画法の例としては、これまで僅かにR. Bellman[1, p104]に一例見られるだけで、本格的な研究はまだないようである。

まず、第2節で両的計画法の基礎定理であるバイマックス定理を用意する。第3節では逐次決定過程として両的計画を導入する。これは、両調性の仮定の下で再帰型利得系をもつ決定過程として定義される。バイマックス定理に基づいて、相隣る最大値関数と最小値関数の対が満たす2元連立再帰式(これを両帰式Bicursive Formulaという)を導く。さらに、戦略、最大値関数列と最小値関数列の対、および最大行動と最小行動の対の三つの対から成る最適解の概念を与える。

第4節では、両的計画法の典型的な一例として乗法型計画法を紹介する。更に、第5節では、もう一つの具体的な応用例として乗加法型計画法を導入する。二つの計画法とも最適解まで例示する。

最後の第6節では、無限段両的計画法に対する2元連立関数方程式(これを両帰方程式Bicusive Equationという)を導き、その解析解を具体的に与える。

2 バイマックス定理

この節では次節以降での両帰式導出の基礎になるバイマックス定理を述べる。このバイマックス定理はマクシマックス定理[13]とミニマックス定理[11]の自然

な拡張になっている。

さて、 X, Y を空でない集合とし、各 $x \in X$ に対して $Y(x)$ を Y の空でない部分集合とする。すなわち、 $Y(\cdot) : X \rightarrow 2^Y$ を点対集合値写像とする。ここでは、記号 Y を集合と写像の両方に用いているが、混乱は起こらないであろう。このとき、 $Y(\cdot) : X \rightarrow 2^Y$ のグラフ G を

$$G = \{(x, y) \mid y \in Y(x), x \in X\} \subset X \times Y$$

で定義する。以下、実数 a, b に対して

$$a \vee b = \max(a, b), \quad a \wedge b = \min(a, b)$$

とする。

定理 2.1 (バイマックス定理) 関数 $g : X \times R^1 \rightarrow R^1$ は次の3条件を満たすとする。

- (i) X は X^-, X^+ の排反な和集合 : $X = X^- + X^+$.
- (ii) 各 $x \in X^-$ に対して $g(x; \cdot) : R^1 \rightarrow R^1$ は非増加。
- (iii) 各 $x \in X^+$ に対して $g(x; \cdot) : R^1 \rightarrow R^1$ は非減少。

関数 $h : G \rightarrow R^1$ は実数値とする。このとき、次の(1), (2)の右辺の2段階最適値が存在すれば、左辺の同時最適値 $\max_{(x,y) \in G} g(x; h(x, y))$, $\min_{(x,y) \in G} g(x; h(x, y))$ も存在して、それぞれ等しい:

$$\max_{(x,y) \in G} g(x; h(x, y)) = \max_{x \in X^-} g(x; \min_{y \in Y(x)} h(x, y)) \vee \max_{x \in X^+} g(x; \max_{y \in Y(x)} h(x, y)) \quad (1)$$

$$\min_{(x,y) \in G} g(x; h(x, y)) = \min_{x \in X^-} g(x; \max_{y \in Y(x)} h(x, y)) \wedge \min_{x \in X^+} g(x; \min_{y \in Y(x)} h(x, y)). \quad (2)$$

3 両的計画法

一般に、有限段逐次決定過程 (Sequential Decision Process, SDP) S は次の5つ組で与えられる [14]。

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

ただし各構成要素は次のように指定される。

- (i) N は全段数を表わす正整数。添字 n は $1 \leq n \leq N$ (または $N+1$) を動く。
 n は現在の段を示す。
- (ii) S_n は空でない集合。第 n 状態空間という。要素 $s_n, s_n^* \in S_n$ を第 n 状態という。 s_1 は初期状態、 s_{N+1} は終端状態である。
- (iii) A_n は空でない集合。第 n 決定態空間という。各 $s_n \in S_n$ に対して $A_n(s_n) \subset A_n$ を空でない部分集合とする。すなわち、 $A_n(\cdot) : S_n \rightarrow 2^{A_n}$, $A_n(\cdot)$

s_n)を状態 s_n における可能な決定空間という。 $A_n(\cdot)$ のグラフ G_n を

$$G_n = \{(s_n, a_n) \mid a_n \in A_n(s_n), s_n \in S_n\} \subset S_n \times A_n$$

とする。要素 $a_n, \hat{a}_n \in A_n(s_n)$ を状態 s_n における決定という。

- (iv) $f_n : S_n \times A_n \times R^1 \rightarrow R^1$ は第 n 利得関数である。
- (v) $k : S_{N+1} \rightarrow R^1$ は終端利得関数である。対 $(\{f_n\}_1^N, k)$ を利得系という。
- (vi) $T_n : G_n \rightarrow S_{n+1}$ は第 n 状態変換である。
- (vii) Opt は Max か min のいずれかである。最適子という。これは SDP S が最適化問題:

$$\begin{aligned} & \text{Optimize } f_1(s_1, a_1; f_2(s_2, a_2; \dots; f_N(s_N, a_N; k(s_{N+1}))) \dots) \\ & \text{subject to (i) } T_n(s_n, a_n) = s_{n+1} \quad 1 \leq n \leq N \\ & \quad \quad \quad \text{(ii) } a_n \in A_n(s_n) \end{aligned} \quad (3)$$

を表現していることを意味する。

さて、われわれの目的は、SDP S が与えられたとき、最適化問題(3)を解くことである。しかし、設定 (i) — (vii) の下で解くことは不十分である。そこで、利得系 $(\{f_n\}_1^N, k)$ に次のように一般化された単調性すなわち両調性を課す。

両調性の仮定: 各 n ($1 \leq n \leq N$)、各 $(s_n, a_n) \in G_n$ に対して、 $f_n(s_n, a_n; \cdot) : R^1 \rightarrow R^1$ は非増加($\downarrow$)か非減少($\uparrow$)のいずれかである。

N 段逐次決定過程 S は、その利得系が両調性の仮定を満たすとき、両的計画(Dynamic Program, BP)といい、 S の代わりに B で表わす。

$$B = (\text{Opt}, \{S_n\}_1^{N+1}, \{A_n\}_1^N, (\{f_n\}_1^N, k), \{T_n\}_1^N).$$

さらに、 N 段両的計画 B は、その利得系が単調性の仮定;

単調性の仮定: 各 n ($1 \leq n \leq N$)、各 $(s_n, a_n) \in G_n$ に対して、 $f_n(s_n, a_n; \cdot) : R^1 \rightarrow R^1$ は常に非減少($\uparrow$)である。

を満たすとき、動的計画(Dynamic Program, DP)といい、 B の代わりに D で表わす。

$$D = (\text{Opt}, \{S_n\}_1^{N+1}, \{A_n\}_1^N, (\{f_n\}_1^N, k), \{T_n\}_1^N).$$

列 $(s_1, a_1, s_2, a_2, \dots, s_N, a_N, s_{N+1})$ は、(3)の条件(i), (ii)を満たすとき、すなわち、

$$T_n(s_n, a_n) = s_{n+1}, \quad a_n \in A_n(s_n) \quad 1 \leq n \leq N$$

のとき、 s_1 からの行動という。列 $\{\pi_n\}_1^N$ は、各 π_n が $\pi_n(s_n) \in A_n(s_n)$, $s_n \in S_n$ を満たすとき、政策という。政策の対 (π, σ) を戦略という。任意の戦略 $(\pi$

, σ) は初期状態 s_1 と共に、次のような s_1 からの二つの特定の行動を生む。すなわち、一つは上方行動 $(s_1, \bar{a}_1, \bar{s}_2, \bar{a}_2, \dots, \bar{s}_N, \bar{a}_N, \bar{s}_{N+1})$ である。ただし

$$\bar{a}_1 = \pi_1(s_1), \quad \bar{s}_2 = T_1(s_1, \bar{a}_1), \quad \bar{a}_2 = \begin{cases} \pi_2(\bar{s}_2) & \text{for } f_1(s_1, \bar{a}_1; \cdot): \uparrow \\ \sigma_2(\bar{s}_2) & \downarrow \end{cases}$$

$$\bar{s}_3 = T_2(\bar{s}_2, \bar{a}_2),$$

$$\bar{a}_3 = \begin{cases} \pi_3(\bar{s}_3) & \text{for } f_2(\bar{s}_2, \bar{a}_2; \cdot): \uparrow \\ \sigma_3(\bar{s}_3) & \uparrow \\ \sigma_3(\bar{s}_3) & \downarrow \\ \pi_3(\bar{s}_3) & \downarrow \end{cases} \quad \bar{a}_2 = \begin{cases} \pi_2(\bar{s}_2) \\ \sigma_2(\bar{s}_2) \\ \pi_2(\bar{s}_2) \\ \sigma_2(\bar{s}_2) \end{cases}, \dots$$

$$\bar{s}_N = T_{N-1}(\bar{s}_{N-1}, \bar{a}_{N-1}),$$

$$\bar{a}_N = \begin{cases} \pi_N(\bar{s}_N) & \text{for } f_{N-1}(\bar{s}_{N-1}, \bar{a}_{N-1}; \cdot): \uparrow \\ \sigma_N(\bar{s}_N) & \uparrow \\ \sigma_N(\bar{s}_N) & \downarrow \\ \pi_N(\bar{s}_N) & \downarrow \end{cases} \quad \bar{a}_{N-1} = \begin{cases} \pi_{N-1}(\bar{s}_{N-1}) \\ \sigma_{N-1}(\bar{s}_{N-1}) \\ \pi_{N-1}(\bar{s}_{N-1}) \\ \sigma_{N-1}(\bar{s}_{N-1}) \end{cases}$$

$$\bar{s}_{N+1} = T_N(\bar{s}_N, \bar{a}_N).$$

もう一つの下方行動 $(s_1, \underline{a}_1, \underline{s}_2, \underline{a}_2, \dots, \underline{s}_N, \underline{a}_N, \underline{s}_{N+1})$ は π と σ を入れ換えて定義される。

戦略 $(\pi^*, \hat{\sigma})$ は、その s_1 からの上方行動 $(s_1, \bar{a}_1^*, \bar{s}_2^*, \bar{a}_2^*, \dots, \bar{s}_N^*, \bar{a}_N^*, \bar{s}_{N+1}^*)$ が最大化問題(3)の最大値に到達するとき、 s_1 において最大という。このとき、上方行動 $(s_1, \bar{a}_1^*, \bar{s}_2^*, \bar{a}_2^*, \dots, \bar{s}_N^*, \bar{a}_N^*, \bar{s}_{N+1}^*)$ を s_1 からの最大行動という。同様に、 s_1 において最小と s_1 からの最小行動が定義される。戦略 $(\pi^*, \hat{\sigma})$ は、各 $s_1 \in S_1$ において最大のと看、最大という。最小も同様である。戦略 $(\pi^*, \hat{\sigma})$ は、最大かつ最小のと看、最適という。

さて、 N 段 BP B

$$B = (0pt, \{S_n\}_1^{N+1}, \{A_n\}_1^N, (\{f_n\}_1^N, k), \{T_n\}_1^N)$$

が与えられたとする。このとき、各 n ($1 \leq n \leq N$) に対して二つの $(N-n+1)$ 部分

BP $\bar{B}_{N-n+1}$ と $\underline{B}_{N-n+1}$ を次で定義する。

$$\bar{B}_{N-n+1} = (\text{Max}, \{S_m\}_n^{N+1}, \{A_m\}_n^N, (\{f_m\}_n^N, k), \{T_m\}_n^N),$$

$$\underline{B}_{N-n+1} = (\text{min}, \{S_m\}_n^{N+1}, \{A_m\}_n^N, (\{f_m\}_n^N, k), \{T_m\}_n^N).$$

この $(N-n+1)$ 段 BP $\bar{B}_{N-n+1}$ と $\underline{B}_{N-n+1}$ は $s_n \in S_n$ から出発して $s_{N+1} \in S_{N+1}$ で終る。両計画はそれぞれ最大（最小）化問題

$$\begin{aligned} & \text{Max (min)} \quad f_n(s_n, a_n; f_{n+1}(s_{n+1}, a_{n+1}; \dots; f_N(s_N, a_N; k(s_{N+1}))) \dots) \\ & \text{s.t.} \quad (i) \quad T_m(s_m, a_m) = s_{m+1} \quad n \leq m \leq N \\ & \quad \quad (ii) \quad s_m \in S_m \end{aligned} \quad (4)$$

を表現している。 $U^{N-n+1}(s_n)$ ($u^{N-n+1}(s_n)$) をそれぞれ (4) の最大値（最小値）とする。すなわち、

$$\begin{aligned} U^{N-n+1}(s_n) &= \text{Max} [f_n(s_n, a_n; \dots; f_N(s_N, a_N; k(s_{N+1}))) \dots | (i), (ii)] \\ u^{N-n+1}(s_n) &= \text{min} [f_n(s_n, a_n; \dots; f_N(s_N, a_N; k(s_{N+1}))) \dots | (i), (ii)] \\ & \quad s_m \in S_m, \quad 1 \leq n \leq N. \end{aligned}$$

関数 $U^{N-n+1}(u^{N-n+1}) : S_n \rightarrow R^1$ をそれぞれ $\bar{B}_{N-n+1}$ の最大値関数 ($\underline{B}_{N-n+1}$ の最小値関数) という。特に、 $n = N+1$ については、形式的 0 部分 BP を

$$\bar{B}_0 = \underline{B}_0 = (\Lambda, S_{N+1}, \Lambda, \Lambda, k, \Lambda)$$

で定義して、 $U^0(u^0) : S_{N+1} \rightarrow R^1$ を

$$U^0(s_{N+1}) = u^0(s_{N+1}) = k(s_{N+1})$$

としておく。ただし、 Λ は空記号。

さて、われわれの目的は与えられた BP B の次のような最適解の対の 3 点セットを求めることである。すなわち、第一は最大値関数列 $\{U^0, U^1, \dots, U^N\}$ と最小値関数列 $\{u^0, u^1, \dots, u^N\}$ の対である。第二は最適戦略 $(\pi^*, \hat{\sigma})$ である。第三は最大行動 $(s_1, \hat{a}_1, \hat{s}_2, \hat{a}_2, \dots, \hat{s}_N, \hat{a}_N, \hat{s}_{N+1})$ と最小行動 $(s_1, \hat{a}_1, \hat{s}_2, \hat{a}_2, \dots, \hat{s}_N, \hat{a}_N, \hat{s}_{N+1})$ の対である。しかしながら、両的計画 $\bar{B}, \underline{B}$ に対する $\{U^n\}$ と $\{u^n\}$ の間の連立再帰式（両帰式という）を解けば、動的計画の場合と同様に、一挙に最適解の 3 点セットが得られる [4-8, 10, 12, 13, 15-21]。

Theorem 3.1 (BP の両帰式 [14])

$$\begin{aligned} U^{N-n+1}(s) &= \text{Max}_{a; -} f_n(s, a; u^{N-n}(T_n(s, a))) \\ & \quad \vee \text{Max}_{a; +} f_n(s, a; U^{N-n}(T_n(s, a))) \end{aligned} \quad (5)$$

$$s \in S_n$$

$$u^{N-n+1}(s) = \text{min}_{a; -} f_n(s, a; U^{N-n}(T_n(s, a)))$$

$$\bigwedge_{a;+} \min f_n(s, a; u^{N-n}(T_n(s, a))) \quad (6)$$

$$U^0(s) = u^0(s) = k(s) \quad s \in S_{N+1}.$$

ただし

$$A_n^-(s) = \{ a \in A_n(s) \mid f_n(s_n, a_n; \cdot) \text{ は非増加}(\downarrow) \}$$

$$A_n^+(s) = \{ a \in A_n(s) \mid f_n(s_n, a_n; \cdot) \text{ は非減少}(\uparrow) \}$$

とし、 $a; -$, $a; +$ はそれぞれ $a \in A_n^-(s)$, $a \in A_n^+(s)$ で最適化することとする。

さらに、

$$\pi_n^*(s_n) = (5) \text{ の右辺の最大値に到達する } a,$$

$$\hat{\sigma}_n(s_n) = (6) \text{ の右辺の最小値に到達する } a,$$

とすると、最適戦略 $(\pi^*, \hat{\sigma})$ とこれから生じる最大行動 $(s_1, a_1^*, s_2^*, a_2^*, \dots, s_N^*, a_N^*, s_{N+1}^*)$ と最小行動 $(s_1, \hat{a}_1, \hat{s}_2, \hat{a}_2, \dots, \hat{s}_N, \hat{a}_N, \hat{s}_{N+1})$ の対が得られる。

4 乗法型計画法

N段 BP

$$B = (0pt, \{S_n\}_1^{N+1}, \{A_n\}_1^N, (\{f_n\}_1^N, k), \{T_n\}_1^N)$$

が乗法型利得系 $(\{f_n\}_1^N, k)$ を持つとき、すなわち

$$f_n(s_n, a_n; g) = f_n(s_n, a_n) \times g$$

のとき、最適化問題：

$$\begin{aligned} & \text{Optimize} \quad f_1(s_1, a_1) \times f_2(s_2, a_2) \times \dots \times f_N(s_N, a_N) \times k(s_{N+1}) \\ & \text{subject to} \quad (i)_n \quad T_n(s_n, a_n) = s_{n+1} \quad 1 \leq n \leq N \\ & \quad \quad \quad (ii)_n \quad a_n \in A_n(s_n) \end{aligned} \quad (7)$$

を表現している。このときの最適値関数

$$U^{N-n+1}(s_n) = \max [f_n(s_n, a_n) \dots f_N(s_N, a_N) k(s_{N+1}) \mid (i)_m, (ii)_m \quad n \leq m \leq N]$$

$$u^{N-n+1}(s_n) = \min [f_n(s_n, a_n) \dots f_N(s_N, a_N) k(s_{N+1}) \mid (i)_m, (ii)_m \quad n \leq m \leq N]$$

$$s_m \in S_m, \quad 1 \leq m \leq N$$

は、両者が存在すれば、次の両帰式を満たす。

$$\begin{aligned} U^{N-n+1}(s) &= \max_{a; -} f_n(s, a) u^{N-n}(T_n(s, a)) \\ &\quad \vee \max_{a; +} f_n(s, a) u^{N-n}(T_n(s, a)) \end{aligned} \quad (8)$$

$$s \in S_n$$

$$u^{N-n+1}(s) = \min_{a;-} f_n(s, a) u^{N-n}(T_n(s, a)) \quad \wedge \quad \min_{a;+} f_n(s, a) u^{N-n}(T_n(s, a)) \quad (9)$$

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

ただし

$$A_n^-(s) = \{ a \in A_n(s) \mid f_n(s_n, a_n) \leq 0 \}$$

$$A_n^+(s) = \{ a \in A_n(s) \mid f_n(s_n, a_n) > 0 \}$$

とし、 $a;-$, $a;+$ はそれぞれ $a \in A_n^-(s)$, $a \in A_n^+(s)$ とする。

例 4.1

Max and min xyz

$$\text{s. t. (i) } x - y + z \leq 2$$

$$(ii) \quad x - y - z \leq -1$$

$$(iii) \quad x \geq -1, y \leq 1, z \geq 0.$$

まず、同時不等式系 (i), (ii), (iii) を次のように $x \rightarrow y \rightarrow z$ の順に関する逐次不等式系 (1), (2), (3) に同値変形しておく。

$$(1) \quad -1 \leq x \leq 3/2$$

$$(2) \quad -1/2 + x \leq y \leq 1$$

$$(3) \quad 0 \vee (1 + x - y) \leq z \leq 2 - x + y.$$

つぎに、最大と最小の両最適値関数列を

$$U^1(x, y) = \text{Max}_{(3)} z,$$

$$u^1(x, y) = \text{min}_{(3)} z$$

$$U^2(x) = \text{Max}_{(2), (3)} yz,$$

$$u^2(x) = \text{min}_{(2), (3)} yz$$

$$U^3 = \text{Max}_{(1)-(3)} xyz,$$

$$u^3 = \text{min}_{(1)-(3)} xyz,$$

で定義すると、両帰式

$$U^1(x, y) = 2 - x + y,$$

$$z^*(x, y) = 2 - x + y,$$

$$u^1(x, y) = 0 \vee (1 + x - y),$$

$$\hat{z}(x, y) = 0 \vee (1 + x - y),$$

$$U^2(x) = \text{Max}_{-\frac{1}{2}+x \leq y \leq 0} [y \cdot u^1(x, y)] \vee \text{Max}_{0 \vee (-\frac{1}{2}+x) \leq y \leq 1} [y \cdot U^1(x, y)]$$

$$u^2(x) = \text{min}_{-\frac{1}{2}+x \leq y \leq 0} [y \cdot U^1(x, y)] \wedge \text{min}_{0 \vee (-\frac{1}{2}+x) \leq y \leq 1} [y \cdot u^1(x, y)]$$

$$U^3 = \text{Max}_{-1 \leq x \leq 0} [x \cdot u^2(x)] \vee \text{Max}_{0 \leq x \leq \frac{3}{2}} [x \cdot U^2(x)]$$

$$u^3 = \text{min}_{-1 \leq x \leq 0} [x \cdot U^2(x)] \wedge \text{min}_{0 \leq x \leq \frac{3}{2}} [x \cdot u^2(x)]$$

が成り立つ。これを解いて

$$U^2(x) = 3 - x,$$

$$y^*(x) = 1$$

$$\begin{aligned}
u^2(x) &= 3/2(x - 1/2), & \hat{y}(x) &= x - 1/2 \\
u^3 &= 9/4, & x^* &= -1 \text{ or } 3/2 \\
u^3 &= -4, & \hat{x} &= -1
\end{aligned}$$

を得る。したがって、 $(x^*, y^*, z^*) = (-1, -3/2, 3/2), (3/2, 1, 3/2)$ のとき、最大値 $9/4$ になり、 $(\hat{x}, \hat{y}, \hat{z}) = (-1, 1, 4)$ のとき、最小値 -4 が得られる。

5 乗加法型計画法

N段 BP

$$B = (Opt, \{S_n\}_1^{N+1}, \{A_n\}_1^N, (\{f_n\}_1^N, k), \{T_n\}_1^N)$$

が乗加法型利得系 $(\{f_n\}_1^N, \{\beta_n\}_1^N, k)$ を持つとき、すなわち

$$f_n(s_n, a_n; g) = f_n(s_n, a_n) + \beta_n(s_n, a_n) \times g$$

のとき、最適化問題：

$$\begin{aligned}
&\text{Optimize } f_1 + \beta_1 f_2 + \beta_1 \beta_2 f_3 + \dots + \beta_1 \beta_2 \dots \beta_{N-1} f_N + \beta_1 \beta_2 \dots \beta_N k \\
&\text{subject to (i)}_n \quad T_n(s_n, a_n) = s_{n+1} \quad 1 \leq n \leq N \\
&\quad \quad \quad \text{(ii)}_n \quad a_n \in A_n(s_n).
\end{aligned} \tag{10}$$

を表わす。ただし

$$\beta_n = \beta_n(s_n, a_n), \quad f_n = f_n(s_n, a_n), \quad k = k(s_{N+1}).$$

このとき、最適値関数

$$\begin{aligned}
U^{N-n+1}(s_n) &= \text{Max } [f_n + \beta_n f_{n+1} + \dots + \beta_n \beta_{n+1} \dots \beta_{N-1} f_N + \dots \\
&\quad + \beta_n \beta_{n+1} \dots \beta_N k \mid (i)_m, (ii)_m \quad n \leq m \leq N] \\
u^{N-n+1}(s_n) &= \min [f_n + \beta_n f_{n+1} + \dots + \beta_n \beta_{n+1} \dots \beta_{N-1} f_N + \dots \\
&\quad + \beta_n \beta_{n+1} \dots \beta_N k \mid (i)_m, (ii)_m \quad n \leq m \leq N] \\
&\quad \quad \quad s_n \in S_n, \quad 1 \leq n \leq N
\end{aligned}$$

は、存在すれば、両帰式

$$\begin{aligned}
U^{N-n+1}(s) &= \text{Max}_{a; -} [f_n(s, a) + \beta_n(s, a) u^{N-n}(T_n(s, a))] \\
&\quad \quad \quad \vee \text{Max}_{a; +} [f_n(s, a) + \beta_n(s, a) U^{N-n}(T_n(s, a))] \\
&\quad \quad \quad s \in S_n \tag{11}
\end{aligned}$$

$$\begin{aligned}
u^{N-n+1}(s) &= \min_{a; -} [f_n(s, a) + \beta_n(s, a) U^{N-n}(T_n(s, a))] \\
&\quad \quad \quad \wedge \min_{a; +} [f_n(s, a) + \beta_n(s, a) u^{N-n}(T_n(s, a))]
\end{aligned}$$

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

を満たす。ただし

$$A_n^-(s) = \{ a \in A_n(s) \mid \beta_n(s_n, a_n) \leq 0 \}$$

$$A_n^+(s) = \{ a \in A_n(s) \mid \beta_n(s_n, a_n) > 0 \}$$

とし、 a ; $-$, a ; $+$ はそれぞれ $a \in A_n^-(s)$, $a \in A_n^+(s)$ とする。

例 5.1 次の最大化問題と最小化問題の対を解こう。

$$\text{Max } x_1 + x_2/x_1 + \dots + x_N/(x_1 x_2 \dots x_{N-1}) + k/(x_1 x_2 \dots x_N) \quad (12)$$

$$\text{s.t. } (i)_n \quad x_n \in \mathbb{Z} \quad 1 \leq n \leq N,$$

$$\text{min } x_1 + x_2/x_1 + \dots + x_N/(x_1 x_2 \dots x_{N-1}) + k/(x_1 x_2 \dots x_N) \quad (13)$$

$$\text{s.t. } (i)_n \quad x_n \in \mathbb{Z} \quad 1 \leq n \leq N$$

ただし

$$N \geq 2, \quad \mathbb{Z} = [-2, -1] \cup [1, 2], \quad k \in \mathbb{R}^1$$

とする。まず、後ろ向きの両帰式

を導こう。そのために

$$U^{N-n+1}(k) = \text{Max } [x_n + \dots + x_N/(x_n x_{n+1} \dots x_{N-1}) + k/(x_n x_{n+1} \dots x_N)] \\ | (i)_m, (ii)_m \quad n \leq m \leq N]$$

$$u^{N-n+1}(k) = \text{min } [x_n + \dots + x_N/(x_n x_{n+1} \dots x_{N-1}) + k/(x_n x_{n+1} \dots x_N)] \\ | (i)_m, (ii)_m \quad n \leq m \leq N]$$

$$U^0(k) = u^0(k) = k$$

とおくと、

$$U^{N-n+1}(k) = \text{Max}_{-2 \leq x \leq -1} [x + U^{N-n}(k)/x] \vee \text{Max}_{1 \leq x \leq 2} [x + U^{N-n}(k)/x] \quad (14)$$

$$u^{N-n+1}(k) = \text{min}_{-2 \leq x \leq -1} [x + U^{N-n}(k)/x] \wedge \text{min}_{1 \leq x \leq 2} [x + U^{N-n}(k)/x] \quad (15)$$

が得られる。両帰式 (14), (15) を解くと、次のようになる。

$N \geq 3$ のとき、最大値

$$U^N(k) = \begin{cases} N-1+U(k) \\ N+U(k)/2 \end{cases} \begin{cases} k \leq -3, \quad k \geq 0 & \text{のとき} \\ -3 \leq k \leq 0 \end{cases}$$

$$= \begin{cases} N-1+(1+k) \\ N-1+(2+k/2) \\ N+1/2(2+k/2) \\ N+1/2(-1-k) \end{cases} = \begin{cases} N+k \\ N+1+k/2 \\ N+1+k/4 \\ N-1/2-k/2 \end{cases} \begin{cases} k \geq 2 & \text{のとき} \\ 0 \leq k \leq 2 \\ -2 \leq k \leq 0 \\ -3 \leq k \leq -2 \end{cases}$$

$$\begin{cases} N-1+(-1-k) \\ N-2-k \end{cases} \quad \begin{cases} k \leq -3 \end{cases}$$

と対応する最大点

$$(x_1^*, x_2^*, \dots, x_N^*) = \begin{cases} (1, 1, \dots, 1, 1, 1) \\ (1, 1, \dots, 1, 1, 2) \\ (1, 1, \dots, 1, 2, 2) \\ (1, 1, \dots, 1, 2, -1) \\ (1, 1, \dots, 1, 1, -1) \end{cases}$$

を得る。また、 $N = 1, 2$ のときの最大解は次のようになる。

$$U^1(k) = \begin{cases} 1+k \\ 2+k/2 \\ -1-k \end{cases} \quad x_1^* = \begin{cases} 1 \\ 2 \\ -1 \end{cases} \quad \begin{cases} k \geq 2 \\ -2 \leq k \leq 2 \\ k \leq -2. \end{cases} \quad \text{のとき}$$

$$U^2(k) = \begin{cases} 2+k \\ 3+k/2 \\ 3/2-k/2 \\ -k \end{cases} \quad (x_1^*, x_2^*) = \begin{cases} (1, 1) \\ (1, 2) \\ (2, -1) \\ (1, -1) \end{cases} \quad \begin{cases} k \geq 2 \\ -2 \leq k \leq 2 \\ -3 \leq k \leq -2 \\ k \leq -3. \end{cases} \quad \text{のとき}$$

したがって、最小解は次のように与えられる。

$$u^N(k) = -U^N(k) \\ \hat{x} = (\hat{x}_1, \hat{x}_2, \dots, \hat{x}_N) = (-x_1^*, x_2^*, \dots, x_N^*).$$

6 無限段問題

この節では乗加法型利得系をもつ無限段定常両的計画法を取り扱う。

さて、次の定常問題を考えよう。

Max and min

$$r_1 + \beta_1 r_2 + \beta_1 \beta_2 r_3 + \dots + \beta_1 \beta_2 \dots \beta_n r_{n+1} + \dots$$

s. t.

$$(i) \quad T(s_n, a_n) = s_{n+1} \quad n \geq 1$$

$$(ii) \quad a_n \in A(s_n),$$

ただし

$$\beta_n = \beta(s_n, a_n), \quad r_n = r(s_n, a_n)$$

とする。ここで、「定常」とは両的計画の構成要素 $S, A, A(s), r(s, a), \beta(s, a), T(s, a)$ がすべて定常であること、すなわち、現在の時刻(段) n に依存しないこととする。

いま、 $U(s)$ と $u(s)$ をそれぞれ求める最大値と最小値とする。これが存在すれば、連立再帰方程式(両帰方程式という)

$$U(s) = \begin{aligned} & \text{Max}_{a; -} [\gamma(s, a) + \beta(s, a)u(T(s, a))] \\ & \quad \vee \text{Max}_{a; +} [\gamma(s, a) + \beta(s, a)U(T(s, a))] \end{aligned} \quad (16)$$

$$u(s) = \begin{aligned} & \text{min}_{a; -} [\gamma(s, a) + \beta(s, a)U(T(s, a))] \\ & \quad \wedge \text{min}_{a; +} [\gamma(s, a) + \beta(s, a)u(T(s, a))] \end{aligned} \quad (17)$$

が成立する。ただし

$$A^-(s) = \{a \in A(s) \mid \beta(s, a) \leq 0\}$$

$$A^+(s) = \{a \in A(s) \mid \beta(s, a) > 0\}$$

とし、 $a; -, a; +$ はそれぞれ $a \in A^-(s), a \in A^+(s)$ で最適化することとする。

例 6.1 Max and min

$$1 + \beta(x_1) + \beta(x_1)\beta(x_2) + \dots + \beta(x_1)\beta(x_2)\dots\beta(x_n) + \dots$$

$$\text{s.t. (i) } -\infty < x_n < \infty \quad n \geq 1.$$

ただし

$$\beta: \mathbb{R}^1 \rightarrow (-1, 1) \text{ は連続、}$$

$$-1 < m = \min_{x \in \mathbb{R}^1} \beta(x) \leq \max_{x \in \mathbb{R}^1} \beta(x) = M < 1,$$

とし、 $\hat{x}, x^*$ はそれぞれ唯一の最小点、最大点とする:

$$\beta(\hat{x}) = m, \quad \beta(x^*) = M.$$

このとき、対応する次の両帰方程式を解かねばならない。

$$U = 1 + \left(\text{Max}_{\beta(x) \leq 0} [\beta(x)u] \vee \text{Max}_{\beta(x) \geq 0} [\beta(x)U] \right)$$

$$u = 1 + \left(\text{min}_{\beta(x) \leq 0} [\beta(x)U] \wedge \text{min}_{\beta(x) \geq 0} [\beta(x)u] \right).$$

ケース (i) $0 \leq m (\leq M)$

この場合は $\beta(x) \geq 0 \quad x \in \mathbb{R}^1$ である。したがって、動的計画法における互いに独立な二つの再帰方程式

$$U = 1 + \text{Max}_{\beta(x) \geq 0} [\beta(x)U] = 1 + MU$$

$$u = 1 + \min_{\beta(x) \geq 0} [\beta(x)u] = 1 + mu$$

を得る。それぞれは一意解

$$U = 1/(1-M), \quad u = 1/(1-m)$$

をもち、点 $\mathbf{x}^* = (x^*, x^*, \dots)$ が最大値を与え、点 $\hat{\mathbf{x}} = (\hat{x}, \hat{x}, \dots)$ が最小値を与える。

ケース (ii) $m \leq 0 \leq M$

このときは $U \geq 0$ に注意すると、次になる。

$$U = 1 + (\max_{\beta(x) \geq 0} [\beta(x)u] \vee MU)$$

$$u = 1 + (mU \wedge \min_{\beta(x) \geq 0} [\beta(x)u]).$$

この系を解くと、

$$U = 1/(1-M), \quad u = (1-M+m)/(1-M)$$

になり、点 $\mathbf{x}^* = (x^*, x^*, \dots)$ で最大値 $U = 1/(1-M)$ をとり、点 $\hat{\mathbf{x}} = (\hat{x}, x^*, x^*, \dots)$ で最小値 $u = (1-M+1)/(1-M)$ をとることが分かる。

ケース (iii) $(m \leq) M \leq 0$

この場合は $-1 < \beta(x) \leq 0 \quad x \in R^1$ に同値である。 $0 \leq u \leq U$ に注意すると、

$$U = 1 + \max_{\beta(x) \geq 0} [\beta(x)u] = 1 + Mu$$

$$u = 1 + \min_{\beta(x) \geq 0} [\beta(x)U] = 1 + mU$$

が得られる。したがって、

$$U = (1+M)/(1-mM), \quad u = (1+m)/(1-mM)$$

となり、 x^* からの交互列 $\mathbf{x}^* = (x^*, x, x^*, x, \dots)$ が最大値 $U = (1+M)/(1-mM)$ を与え、 x からの交互列 $\hat{\mathbf{x}} = (\hat{x}, x^*, \hat{x}, x^*, \dots)$ が最小値 $u = (1+m)/(1-mM)$ を与える。

参考文献

1. R. BELLMAN, "Dynamic Programming," Princeton Univ. Press, Princeton, N.J., 1957.
2. E.V. Denardo, "Dynamic Programming : Models and Applications," Prentice-Hall, N.J., 1982.
3. E.V. Denardo and L.G. Mitten, Elements of sequential decision

- processes, J. Indust. Engrg. 18(1967), 106-111.
4. F. FURUKAWA AND S. IWAMOTO, Markovian decision processes with recursive reward functions, Bull. Math. Statist. 15(1973), 79-91.
 5. F. FURUKAWA AND S. IWAMOTO, Dynamic programming on recursive reward systems, Bull. Math. Statist. 17(1976), 103-126.
 6. S. IWAMOTO, The second principle of optimality, Bull. Math. Statist. 17(1977), 104-114.
 7. S. IWAMOTO, A class of inverse theorems on recursive programming with monotonicity, J. Operations Res. Soc. Japan 20(1977), 94-112.
 8. 岩本 誠一、逐次決定過程としての動的計画論(1)、(2)、オペレーションズ・リサーチ 22(1977)、427-434; 496-501。
 9. S. IWAMOTO, An inverse theorem between main and inverse dynamic programming: infinite-stage case, in: International Conference on Dynamic Programming, Vancouver, Canada, April, 1977; Dynamic Programming and Its Applications, (Ed. M.L. Puterman), 319-334, Academic Press, New York, 1978.
 10. 岩本 誠一、動的計画の理論と応用、数学 31(1979)、331-348。
 11. S. IWAMOTO, Sequential minimaximization under dynamic programming structure, J. Math. Anal. Appl. 108(1985), 267-282.
 12. 岩本 誠一、「動的計画論」、九州大学出版会、1987年。
 13. S. IWAMOTO, A three-mirror problem on dynamic programming, Modeling and Control of Systems in Engineering, Quantum Mechanics, Economics and Biosciences, Proceedings of the Third Bellman Continuum Workshop, 1988, June 13-14, Sophia, Antipolis, France, Ed. A. Blaquiere, Lecture Notes in Control and Information Sciences 121, p.363-382.
 14. 岩本 誠一、動的計画法の最近の進歩、第2回 RAMP シンポジウム論文集、1990、p.129-140。
 15. S. IWAMOTO, R.J. TOMKINS AND C.-L. WANG, Sequential control process for unknown influences, in: Congressus Numerantium Vol. 51 (D.S. Meek and G.H.J. van Rees Eds.) Proceedings of the Fifteenth Manitoba Conference on Numerical Mathematics and Computing, Winnipeg, 1985, Utilitas Mathematica Publ. Winnipeg, (1986), 193-207.
 16. J.L. Lauriere, "Elements de Programmation Dynamique", Gauthier-Villars, Paris, 1979.
 17. T.L. Morin, Monotonicity and the principle of optimality, J. Math.

- Anal. Appl. **86**(1982), 665-674.
18. M. Sniedovich, Dynamic programming and principle of optimality, J. Math. Anal. Appl. **65**(1978), 586-606.
 19. M. Sniedovich, A new look at Bellamn's principle of optimality, J. Opt. Theo. Appl. **49**(1986), 161-176.
 20. P. Whittle, "Optimization over Time — Dynamic Programming and Stochastic Control, Vol. I", John Wiley & Sons, New York, 1981.
 21. P. Whittle, "Optimization over Time — Dynamic Programming and Stochastic Control, Vol. II", John Wiley & Sons, New York, 1984.