

A first-passage problem with multiple costs

長岡工業高等専門学校 涌田和芳 (WAKUTA Kazuyoshi)

1. はじめに

Sancho(1985)は、完全エルゴード性の条件の下で、first-passage 問題の政策反復アルゴリズムを示した。Bertsekas & Tsitsiklis(1991) は、この問題を確率的最短経路問題と呼び、確率1で目的地に到達しない政策が存在する場合を扱った。そして、最短経路問題のいわゆる“コストは正”という仮定を弛めて、“improper な政策の期待コストは無限大”という仮定に置き換えた。一方、実際上の問題のために、多目的最短経路問題が研究されている(Sancho(1986), Sniedovich(1988))。そこで、本論では多目的 first-passage 問題を考える。確率1で目的地に到達しない政策の存在を認める。また、多目的コストを扱うために、コストは非負であるという仮定をおく。そして、最適な確定的定常政策を特徴付け、それを求める政策反復アルゴリズムを示す。

2. 多目的 first-passage 問題(FPPMC)

$S = \{1, 2, \dots, N\}$: 状態空間. $A(i)$ = 有限集合 : 行動空間.

$p(j|i, a)$, $i, j \in S, a \in A(i)$: 推移確率. $c(i, a) = (c^1(i, a), \dots, c^m(i, a))$: コスト関数.

N : 目標状態 (目的地). 任意の $a \in A(N)$ に対して, $p(N|N, a) = 1, c(N, a) = 0$.

[仮定 1] 任意の $i \in S, a \in A(i)$ に対して, $c(i, a) \geq 0$ (ベクトルに対する不等号 $\geq, \leq, >$ は、通常の意味で用いる) .

$f: i_1$ - proper $\Leftrightarrow P_f\{X_t = N \text{ for some } t \geq 1 | X_1 = i_1\} = 1$.

f : proper $\Leftrightarrow f$: 任意の $i_1 \in S'$ に対して, i_1 - proper. ただし, $S' = S \setminus N$.

[仮定 2] 少なくとも1つの properな確定的定常政策が存在する.

Π : すべての政策の集合, Π_D : すべての確定的定常政策の集合

$$I_\pi^k(i_1) = E_\pi \left[\sum_{t=1}^{\infty} c^k(X_t, Y_t) | X_1 = i_1 \right]$$

$I_\pi(i_1) = (I_\pi^1(i_1), \dots, I_\pi^m(i_1))$: π のコスト

$V(i_1) = \bigcup_{\pi \in \Pi} \{I_\pi(i_1)\}$, $V_D(i_1) = \bigcup_{f^\infty \in \Pi_D} \{I_{f^\infty}(i_1)\}$, $i_1 \in S$, とおく.

$\bar{R} = R \cup \{\infty\}$. $U \subset \bar{R}^m$ に対して, $e(U) = \{x \in U | y \leq x \text{ ならば } y = x\}$.

$\pi : i_1$ - 最適 $\Leftrightarrow I_\pi(i_1) \in e(V(i_1))$. π : 最適 $\Leftrightarrow \pi$: 任意の $i_1 \in S'$ に対して i_1 - 最適.

$$S(f, i_1) = \{j \in S \mid P_f(X_t = j \mid X_1 = i_1) > 0 \text{ for some } t \geq 1\}, f \in \Pi_D, i_1 \in S'$$

$$S(i_1) = \bigcup_{f \in \Pi_D} S(f, i_1), f \in \Pi_D, \text{ とおく.}$$

[補題 1] $f \in \Pi_D$ が i_1 -proper ならば, $I_f^k(i_1) < \infty, k = 1, \dots, m$. さもないければ, ある k に対して, $I_f^k(i_1) = \infty$.

[証明]

• f は i_1 -proper とする. このとき, 状態空間が $S(f, i_1)$ の吸収マルコフ連鎖ができる.
 $m_f(i_1)$: i_1 から N への平均 first passage 時間, とする. このとき

$$I_f^k(i_1) = E_f \left[\sum_{t=1}^{\infty} c^k(X_t, Y_t) \mid X_1 = i_1 \right]$$

$$\leq \left(\max_{j, a} c^k(j, a) \right) m_f(i_1) < \infty, k = 1, 2, \dots, m.$$

• f は i_1 -proper ではないとする. このとき, N を含まない再帰クラスがある.

$\sum_{t=1}^{\infty} P_f\{X_t = j \mid X_1 = i_1\} = \infty$ なる $j \in S'$ が存在する. $c(j, f(j)) \geq 0$ なので, ある k に対して,
 $I_f^k(i_1) = \infty$.

$$\tilde{\Pi}(i_1) = \{\pi \in \Pi \mid I_{\pi}^k(i_1) < \infty, k = 1, \dots, m\}, \tilde{\Pi}_D(i_1) = \{f \in \Pi_D \mid I_f^k(i_1) < \infty, k = 1, \dots, m\}$$

$$\tilde{V}(i_1) = \bigcup_{\pi \in \tilde{\Pi}(i_1)} \{I_{\pi}(i_1)\}, \tilde{V}_D(i_1) = \bigcup_{f \in \tilde{\Pi}_D(i_1)} \{I_f(i_1)\}, i_1 \in S', \text{ とおく.}$$

$$[\text{補題 2}] e(\text{co}\tilde{V}_D(i_1)) = e(\text{co}\tilde{V}_D(i_1) + R_+^m) = e(\tilde{V}(i_1)) = e(V(i_1)), i_1 \in S'.$$

3. 最適政策

$B^m(S)$: S から R^m へのすべての関数の集合.

$\lambda \in B^m(S)$, $\lambda > 0$ に対して, $c^{\lambda}(i_1, i_t, a_t) = \langle \lambda(i_1), c(i_t, a_t) \rangle$ をコスト関数にもつ非定常動的計画 NDP(λ)を考える. 補題 2 より, 各 $i_1 \in S'$ に対して, 政策は $\tilde{\Pi}(i_1)$ に制限してよい.

$$J_{\pi}^{\lambda}(i_1) = E_{\pi} \left[\sum_{t=1}^{\infty} c^{\lambda}(X_1, X_t, Y_t) \mid X_1 = i_1 \right], i_1 \in S'.$$

$$\pi^* : \text{NDP}(\lambda) \text{ で最適} \Leftrightarrow \text{任意の } i_1 \in S' \text{ と } \pi \in \tilde{\Pi}(i_1) \text{ に対して } J_{\pi^*}^{\lambda}(i_1) \geq J_{\pi}^{\lambda}(i_1).$$

[命題 1] π が FPPMC で最適であるための必要十分条件は, π が NDP(λ), $\lambda > 0$, で最適であることである.

[命題 2] FPPMC に, 最適な proper 確定的定常政策が存在する.

$$T_a u(i) = c(i, a) + \sum_{j \in S} p(j|i, a) u(j), u \in B^m(S)$$

$$T_f u(i) = T_{f(i)} u(i), L_f(i, a) = T_a I_f(i)$$

$$H_c = \{x \in R^m | \langle c, x \rangle \leq 0\}, c \in R^m, \text{ とおく.}$$

次の定理は, Wakuta & Togawa(1998)と同様に証明できる.

[定理 1] f^* が最適ならば, f^* は proper, かつ各 $i_1 \in S'$ に対して

$$L_{f^*}(i_t, a_t) - I_{f^*}(i_t) \in H_{\lambda(i_1)}, i_t \in S(f^*, i_1), a_t \in A(i_t)$$

なる $\lambda \in B^m(S'), \lambda > 0$, が存在する.

[定理 2] f^* が proper, かつ各 $i_1 \in S'$ に対して

$$L_{f^*}(i_t, a_t) - I_{f^*}(i_t) \in H_{\lambda(i_1)}, i_t \in S(i_1), a_t \in A(i_t)$$

なる $\lambda \in B^m(S'), \lambda > 0$, が存在すれば, f^* は最適である.

[系] $f^* \in \Pi_D$ は proper, かつ

$$L_{f^*}(i, a) - I_{f^*}(i) \in H_\lambda, i \in S, a \in A(i)$$

なる $\lambda > 0$ が存在すれば, f^* は最適である.

B : 最適でないことがわかっている確定的定常政策の集合

$$\Pi'_D = \Pi_D \setminus B$$

$$V'_D(i_1) = \bigcup_{f \in \Pi'_D} \{I_f(i_1)\}, i_1 \in S', \text{ とおく.}$$

[定理 3] $f^* \in \Pi_D$ が最適であるための必要十分条件は, f^* は proper, かつ各 $i_1 \in S'$ に対して

$$I_{f^*}(i_1) - I_{f^*}(i_1) \in H_{\lambda(i_1)}, f \in \Pi'_D$$

なる $\lambda \in B^m(S'), \lambda > 0$, が存在することである.

$$q_f(i, a) = L_f(i, a) - I_f(i), (i, a) \in GrA$$

$$Q_f: q_f(i, a), i \in S', a \in A(i) \text{ を行にもつ行列}$$

$$\overline{Q}_f(i_1), i_1 \in S': q_f(i_t, a_t), i_t \in S(f, i_1) \setminus \{N\}, a_t \in A(i_t) \text{ を行にもつ行列}$$

$$\overline{\overline{Q}}_f(i_1), i_1 \in S': q_f(i_t, a_t), i_t \in S(i_1) \setminus \{N\}, a_t \in A(i_t) \text{ を行にもつ行列, とおく.}$$

[定理 4] $f^* \in \Pi_D$ が最適ならば, f^* は proper, かつ次の各線形不等式系が解をもつ.

$$(S_1): \begin{cases} x > 0 \\ \overline{Q}_{f^*}(1)x \geq 0 \end{cases}, \dots, (S_{N-1}): \begin{cases} x > 0 \\ \overline{Q}_{f^*}(N-1)x \geq 0 \end{cases}$$

[定理 5] $f^* \in \Pi_D$ が proper, かつ次の各線形不等式系が解をもてば, f^* は最適である.

$$(T_1): \begin{cases} x > 0 \\ \overline{Q}_{f^*}(1)x \geq 0 \end{cases}, \dots, (T_{N-1}): \begin{cases} x > 0 \\ \overline{Q}_{f^*}(N-1)x \geq 0 \end{cases}$$

[系] $f^* \in \Pi_D$ が proper, かつ次の線形不等式系が解をもてば, f^* は最適である.

$$(S_0): \begin{cases} x > 0 \\ Q_{f^*}x \geq 0 \end{cases}$$

$$d_f^g(i_1) = I_g(i_1) - I_f(i_1), g \in \Pi'_D, i_1 \in S'$$

$D_f(i_1), i_1 \in S': d_f^g(i_1), g \in \Pi'_D$ を行にもつ行列, とおく.

[定理 6] $f^* \in \Pi_D$ が最適であるための必要十分条件は, f^* が proper, かつ次の各線形不等式系が解をもつことである.

$$(U_1): \begin{cases} x > 0 \\ D_f(1)x \geq 0 \end{cases}, \dots, (U_{N-1}): \begin{cases} x > 0 \\ D_f(N-1)x \geq 0 \end{cases}$$

以上の線形不等式系の問題は, 次のような LP 問題として定式化できる.

$$\begin{aligned} P(S): \quad & \text{Maximize } y \\ & \text{subject to} \\ & \begin{cases} x_1 \geq y, \dots, x_m \geq y \\ Bx \geq 0 \\ x_1 \geq 0, \dots, x_m \geq 0, y \geq 0, \end{cases} \end{aligned}$$

4. 政策反復アルゴリズム

$f, g \in \Pi_D$ に対して, $I_f^g(i) = T_g I_f(i), i \in S$ とおく.

[補題 3] f は proper とする. このとき,

- (i) $I_f^g - I_f^f \leq 0$ ならば, g も proper で, $I_g \leq I_f$.
- (ii) $I_f^g - I_f^f = 0$ ならば, $I_g = I_f$.
- (iii) $I_f^g - I_f^f \geq 0$ かつ g が proper ならば, $I_g \geq I_f$.

Policy iteration algorithm

E_n : 政策反復で支配されない政策の集合

F_n : 政策反復で支配される政策の集合, とおく.

Phase I

1. $E_0 = F_0 = \phi$ とおき, 任意の $f_1 \in \Pi_D$ を選ぶ.
2. $f_n, n \geq 1$ に対して, $I_{f_n}(i), i \in S'$ を求 proper かどうか判定する.
3. f_n が proper でなければ, $E_n = E_{n-1}, F_n = F_{n-1} \cup \{f_n\}$ とおく.
そして, $f_{n+1} \in (\Pi_D \setminus (E_n \cup F_n))$ を選び, 2 へ行く.
4. f_n が proper ならば,

$$A_{f_n} = \{g \in (\Pi_D \setminus (E_{n-1} \cup F_{n-1})) \mid I_{f_n}^g - I_{f_n}^{f_n} \leq 0\}$$

$$B_{f_n} = \{g \in (\Pi_D \setminus (E_{n-1} \cup F_{n-1})) \mid I_{f_n}^g - I_{f_n}^{f_n} \geq 0\}$$

$$C_{f_n} = \{g \in (\Pi_D \setminus (E_{n-1} \cup F_{n-1})) \mid I_{f_n}^g - I_{f_n}^{f_n} = 0\}, \text{ とおく.}$$
 - (i) $A_{f_n} \neq \phi$ ならば, $E_n = E_{n-1}, F_n = F_{n-1} \cup B_{f_n} \cup C_{f_n}$ とおく. そして, 2 へ行く.
 - (ii) $A_{f_n} = \phi$ ならば,
 $E_n = E_{n-1} \cup C_{f_n}, F_n = F_{n-1} \cup B_{f_n}$ とおく. そして, 2 へ行く.
5. $E_n \cup F_n = \Pi_D$ となったら止める.

Phase II

$E^* = E_n, F^* = F_n, E$: 最適なすべての政策, F : 最適でないすべての政策, とおく.

1. 各 $f \in E^*$ に対して, LP 問題 $P(S_0)$ を解いて, $E' \subset E$ を求める.

Phase III

1. 各 $f \in (E^* \setminus E')$ に対して, LP 問題 $P(U_1), \dots, P(U_N)$ を解いて, 最適かどうか判定する. E と F が求まる.

5. 数値例

$S = \{1, 2, 3, 4\}$, ここで, 4 が目標状態

$$A(1) = \{1, 2, 3\}, A(2) = \{1\}, A(3) = \{1, 2, 3\}, A(4) = \{1\}$$

$$p(2|1, 1) = p(3|1, 2) = p(4|1, 3) = 1, p(3|2, 1) = 1$$

$$p(1|3, 1) = p(4|3, 2) = 1, p(1|3, 3) = p(4|3, 3) = \frac{1}{2}, p(4|4, 1) = 1$$

$$c(1, 1) = (0, 2), c(1, 2) = (2, 1), c(1, 3) = (6, 6), c(2, 1) = (0, 1)$$

$$c(3, 1) = (2, 1), c(3, 2) = (4, 1), c(3, 3) = (3, 1), c(4, 1) = (0, 0).$$

$$\alpha_1 : \alpha_1(1) = 1, \alpha_1(3) = 1; \alpha_2 : \alpha_2(1) = 1, \alpha_2(3) = 2; \alpha_3 : \alpha_3(1) = 1, \alpha_3(3) = 3$$

$$\beta_1 : \beta_1(1) = 2, \beta_1(3) = 1; \beta_2 : \beta_2(1) = 2, \beta_2(3) = 2; \beta_3 : \beta_3(1) = 2, \beta_3(3) = 3$$

$$\gamma_1 : \gamma_1(1) = 3, \gamma_1(3) = 1; \gamma_2 : \gamma_2(1) = 3, \gamma_2(3) = 2; \gamma_3 : \gamma_3(1) = 3, \gamma_3(3) = 3.$$

Phase I

(1) $f_1 = \alpha_1$ を選び, $I_{\alpha_1}(i), i=1,2,3$ を計算する. 補題 3 より α_1 は, proper ではない.

$$E_1 = \phi, \quad F_1 = \{\alpha_1\}$$

(2) $f_2 = \alpha_2$ を選び, $I_{\alpha_2}(i), i=1,2,3$ を計算する.

$$I_{\alpha_2}(1)=(4,4), I_{\alpha_2}(2)=(4,2), I_{\alpha_2}(3)=(4,1)$$

α_2 は, proper.

$$I_{\alpha_2}^{\alpha_3}(1)-I_{\alpha_2}^{\alpha_2}(1)=(0,0), I_{\alpha_2}^{\alpha_3}(2)-I_{\alpha_2}^{\alpha_2}(2)=(0,0), I_{\alpha_2}^{\alpha_3}(3)-I_{\alpha_2}^{\alpha_2}(3)=(1,2)$$

$$I_{\alpha_2}^{\beta_1}(1)-I_{\alpha_2}^{\alpha_2}(1)=(2,-2), I_{\alpha_2}^{\beta_1}(2)-I_{\alpha_2}^{\alpha_2}(2)=(0,0), I_{\alpha_2}^{\beta_1}(3)-I_{\alpha_2}^{\alpha_2}(3)=(2,4)$$

$$I_{\alpha_2}^{\beta_2}(1)-I_{\alpha_2}^{\alpha_2}(1)=(2,-2), I_{\alpha_2}^{\beta_2}(2)-I_{\alpha_2}^{\alpha_2}(2)=(0,0), I_{\alpha_2}^{\beta_2}(3)-I_{\alpha_2}^{\alpha_2}(3)=(0,0)$$

$$I_{\alpha_2}^{\beta_3}(1)-I_{\alpha_2}^{\alpha_2}(1)=(2,-2), I_{\alpha_2}^{\beta_3}(2)-I_{\alpha_2}^{\alpha_2}(2)=(0,0), I_{\alpha_2}^{\beta_3}(3)-I_{\alpha_2}^{\alpha_2}(3)=(1,2)$$

$$I_{\alpha_2}^{\gamma_1}(1)-I_{\alpha_2}^{\alpha_2}(1)=(2,2), I_{\alpha_2}^{\gamma_1}(2)-I_{\alpha_2}^{\alpha_2}(2)=(0,0), I_{\alpha_2}^{\gamma_1}(3)-I_{\alpha_2}^{\alpha_2}(3)=(2,4)$$

$$I_{\alpha_2}^{\gamma_2}(1)-I_{\alpha_2}^{\alpha_2}(1)=(2,2), I_{\alpha_2}^{\gamma_2}(2)-I_{\alpha_2}^{\alpha_2}(2)=(0,0), I_{\alpha_2}^{\gamma_2}(3)-I_{\alpha_2}^{\alpha_2}(3)=(0,0)$$

$$I_{\alpha_2}^{\gamma_3}(1)-I_{\alpha_2}^{\alpha_2}(1)=(2,2), I_{\alpha_2}^{\gamma_3}(2)-I_{\alpha_2}^{\alpha_2}(2)=(0,0), I_{\alpha_2}^{\gamma_3}(3)-I_{\alpha_2}^{\alpha_2}(3)=(1,2)$$

$$A_{\alpha_2} = \phi, B_{\alpha_2} = \{\alpha_3, \gamma_1, \gamma_2, \gamma_3\}, C_{\alpha_2} = \{\alpha_2\}$$

$$E_2 = \{\alpha_2\}, \quad F_2 = \{\alpha_1, \alpha_3, \gamma_1, \gamma_2, \gamma_3\}$$

(3) $f_3 = \beta_1$ を選び, $I_{\beta_1}(i), i=1,2,3$ を計算する. 補題 3 より β_1 は, proper ではない.

$$E_3 = \{\alpha_2\}, \quad F_3 = \{\alpha_1, \alpha_3, \beta_1, \gamma_1, \gamma_2, \gamma_3\}$$

(4) $f_4 = \beta_2$ を選び, $I_{\beta_2}(i), i=1,2,3$ を計算する.

$$I_{\beta_2}(1)=(6,2), I_{\beta_2}(2)=(4,2), I_{\beta_2}(3)=(4,1)$$

β_2 は, proper.

$$I_{\beta_2}^{\beta_3}(1)-I_{\beta_2}^{\beta_2}(1)=(2,1), I_{\beta_2}^{\beta_3}(2)-I_{\beta_2}^{\beta_2}(2)=(0,0), I_{\beta_2}^{\beta_3}(3)-I_{\beta_2}^{\beta_2}(3)=(2,1)$$

$$A_{\beta_2} = \phi, B_{\beta_2} = \{\beta_3\}, \quad C_{\beta_2} = \{\beta_2\}$$

$$E_4 = \{\alpha_2, \beta_2\}, \quad F_4 = \{\alpha_1, \alpha_3, \beta_1, \beta_3, \gamma_1, \gamma_2, \gamma_3\}$$

(5) $E_2 \cup F_2 = \Pi_D$ なので止める.

参考文献

1. Bertsekas & Tsitsiklis (1991) An analysis of stochastic shortest path problems, MOR 16, 580-595.
2. Sancho (1985) Routing problems and Markov decision processes, JMAA 105, 76-85.
3. Sancho (1986) A multi-objective routing problem, Eng. Opt. 10, 71-76.
4. Sniedovich (1988) A multi-objective routing problems revisited, Eng. Opt. 13, 99-108.
5. Wakuta & Togawa (1998) Solution procedures for multi-objective Markov decision processes, Optimization 43, 29-46.