

2つの評価基準を考慮したたまご落とし問題

九州工業大学・大学院工学研究院 藤田 敏治

Toshiharu Fujita

Graduate School of Engineering,

Kyushu Institute of Technology

九州工業大学・大学院工学府 近藤 匠

Takumi Kondou

Graduate School of Engineering,

Kyushu Institute of Technology

1 はじめに

本研究では、S. Wagon が提出した問題 “The Egg Drop” (<http://mathforum.org/wagon/fall00/p921.html>, 以下たまご落とし問題と呼ぶ) を拡張し、動的計画法による解法を与える。もとの問題は M. Sniedovich が [3] において扱っているが、ここでは非決定性動的計画法による独自の定式化を与え、さらには加法型評価 (コスト評価) を加えた多目的版に拡張し、再帰式を導く。

もとのたまご落とし問題は、いわゆる落下試験回数の最小化を行う問題である。たまごというのはあくまで象徴的なものであり、たまごに課せられた仮定:

- ・すべてのたまごは、落下に対し、同じ結果を生じる。
- ・一度割れた (割れなかった) たまごは、再度使用できない (できる)。
- ・ある高さから落としてたまごが割れれば (割れなければ)、それ以上 (以下) の高さから落としても割れる (割れない)。

をみたく対象の耐久性を、その破壊される高さにより知りたいのである。そして、限られた数の対象 (たまご) を実際に落下させて調べる際、最悪何回の落下試験が必要となるか、またその落とし方 (最適政策) はどうすればよいか、を求めることを目的としている。この問題は、非決定性動的計画法を用いて最大加法型評価の最小化問題として扱われるが、本研究では、加法型評価を同時に扱い、辞書式順序に関する最小化問題を考える。というのも、最大回数最小化には一般に多数の解 (最適政策) が存在するため、その中でもコスト最小のものを求めることができれば大いに役立つと考えられるからである。

以下、第2節では、元の問題を非決定性動的計画問題 ([1]) として定式化し、導かれる再帰式を紹介し、第3節で、落下コストを想定した加法型評価問題 ([4]) について述べる。そして、第4節で落下回数・落下コストを同時に扱った場合の結果を示す。

2 たまご落とし問題

耐久性 (壊れない最大の高さ) を知るための落下試験の進め方について概略を説明する。高さは離散的とし、単位は “階” で表すことにする。まず、ある階 (これを u 階とする) からたまごを落とす。もし割れてしまえば、その階より上から落としても割れることがわかるので、あとは u 階より下の階について調べる。割れなければ、その階より下から落としても割れないことがわかるので、あとは u 階より上の階について調べる。これを、たまごが足りなくならないような手順

で、調べる必要がある階が無くなるまで続けていくのである。たとえば、最初に5階からたまごを落とすと割れ、次に4階から落とすと割れなかった場合、割れない最大の高さは4階とわかる。この場合、2回の落下試験で結果が判明したが、どの階を落とす高さとして採用するかによって、割れない最大階が判明するまでの落下試験回数は変わる。しかも、割れない最大階は不明なので、たとえそれがどの階であっても対応できなければならない。リスク回避型で考えたとき、この問題は落下回数の最大値の最小化問題として扱われる。

ここでは、一般に M 個のたまごで T 階のビル ($M > 0, T > 0$) を調べる場合について、非決定性動的計画の枠組みで定式化する。状態は、その時点でのたまごの残り数が m 個で、残りの調べる階が s 階から t 階である場合に (m, s, t) であらわし、 u 階から落とすことを、決定 u とあらわす。よって、状態空間は

$$X = \{0, 1, 2, \dots, M\} \times \{0, 1, 2, \dots, T\}^2$$

決定空間および、可能決定集合はそれぞれ

$$U = \{0, 1, 2, \dots, T\}$$

$$U(x) = \begin{cases} \{s, s+1, \dots, t\} & m \geq 2, s, t \geq 1 \\ \{s\} & m = 1, s, t \geq 1 \\ \{0\} & \text{その他} \end{cases}, x = (m, s, t) \in X$$

となる。ただし、状態空間の階をあらわす第2, 3要素における0は、調べる階が無くなった状態を表す際に使用し、決定の0は調べる階が無い場合に、便宜上与える決定で、何も行わないということの意味するものとする。なお、たまごが残り1個になった際は、1番下の階から落とさなければ、結果が判明しない場合が生じる。

次に、状態 $x = (m, s, t) \in X$ において決定 $u \in U(x)$ をとった際の次状態全体をあらわす非決定性状態推移 T を定める。

$$T(x, u) = \begin{cases} \{(m, u+1, t), (m-1, s, u-1)\} & m > 0, s > 0, s \neq t, u \neq s, t \\ \{(m, u+1, t), (m-1, 0, 0)\} & m > 0, s > 0, s \neq t, u = s \\ \{(m, 0, 0), (m-1, s, u-1)\} & m > 0, s > 0, s \neq t, u = t \\ \{(m, 0, 0), (m-1, 0, 0)\} & m > 0, s > 0, s = t \\ \{x\} & \text{その他} \end{cases}$$

さらに、たまごを落とす回数は建物の階数 T を超えることはない（同じ階から落とすことはない）ので、決定は最大でも T 回となり、本問題における記数は $N = T$ ととればよい。また、利得関数については、正の決定をとる（たまごを落とす）ごとに1をとるように定める。すなわち、利得関数、終端利得関数をそれぞれ

$$r(x, u) = \begin{cases} 1 & u > 0 \\ 0 & u = 0 \end{cases}, x = (m, s, t) \in X, u \in U(x)$$

$$k(x) = 0, x \in X$$

とし、初期状態 x_0 に対する目的関数は

$$V_1(x_0; \pi) = r_0 + \bigvee_{x_1 \in T_1} \left\{ r_1 + \bigvee_{x_2 \in T_2} \left\{ r_2 + \dots + \bigvee_{x_{N-1} \in T_{N-1}} \left\{ r_{N-1} + \bigvee_{x_N \in T_N} k \right\} \dots \right\} \right\}$$

$$x_0 \in X, \pi = \{\pi_0, \pi_1, \dots, \pi_{N-1}\} \in \Pi$$

で定める。ただし、 Π はマルコフ政策全体をあらわし、

$$r_n = r(x_n, \pi_n(x_n)), \quad k = k(x_N), \quad T_{n+1} = T(x_n, \pi_n(x_n))$$

である。この目的関数は、各期において、次期以降に必要となる落下試験回数の最大値をとり、(その期に落下試験を行えば) それに 1 を加える、という操作の繰り返しにより構成されている。これにより、マルコフ政策 π に従って落下試験を行った際の最大試験回数を表している。

以上より、最大落下回数最小化問題は

$$P_1(x_0) \quad \text{Minimize } V_1(x_0; \pi) \quad \text{subject to } \pi \in \Pi$$

とあらわされる。そして、 x_n ではじまる第 n 期以降の部分過程の最適値を $v_n(x_n)$ であらわすとき ($n = 0, 1, \dots, N$)、次の再帰式が成り立つ。

$$\begin{aligned} v_N(x) &= k(x) \quad x \in X \\ v_n(x) &= \min_{u \in U(x)} \left[r(x, u) + \bigvee_{y \in T(x, u)} v_{n+1}(y) \right] \quad x \in X, \quad 0 \leq n \leq N-1. \end{aligned}$$

$v_n(x)$ の最大値を与える決定 $u \in U(x)$ を $\pi_n^*(x)$ とすると ($n = 0, 1, \dots, N-1$)、マルコフ政策クラス Π の中で最適政策 $\pi^* = \{\pi_0^*, \pi_1^*, \dots, \pi_{N-1}^*\}$ が得られる。

3 コスト最小化問題

前節では、落下回数の最適化を行ったが、落とす階に応じたコストを考えた場合、落下回数よりもむしろコストが問題になることも考えられる。そこで本節では、利得関数が状態・決定に応じた何らかのコストを表す際に、その総和を最小化する問題を考える。なお、ここでは、非決定推移のもとで単純に利得の和を扱っているが、この種の目的関数によって、落下コストの期待値を表すことができることを補足しておく。

さて、コスト最小化問題における状態・決定・状態推移等は前節と同じとし、利得関数を

$$q(x, u) = \begin{cases} x, u \text{ に依存する何らかのコスト} & u > 0 \\ 0 & u = 0 \end{cases}, \quad x \in X, \quad u \in U(x)$$

とし (終端利得関数も前節と同じ)、目的関数を次で定める。

$$V_2(x_0; \pi) = q_0 + \sum_{x_1 \in T_1} \left\{ q_1 + \sum_{x_2 \in T_2} \left\{ q_2 + \dots + \sum_{x_{N-1} \in T_{N-1}} \left\{ q_{N-1} + \sum_{x_N \in T_N} k \right\} \dots \right\} \right\}$$

ただし、 $q_n = q(x_n, \pi_n(x_n))$ である。このとき、初期状態 x_0 に対するコスト最小化問題は

$$P_2(x_0) \quad \text{Minimize } V_2(x_0; \pi) \quad \text{subject to } \pi \in \Pi$$

とあらわされ、前節と同様な次の再帰式が導かれる。

$$\begin{aligned} v_N(x) &= k(x) \quad x \in X \\ v_n(x) &= \min_{u \in U(x)} \left[q(x, u) + \sum_{y \in T(x, u)} v_{n+1}(y) \right] \quad x \in X, \quad 0 \leq n \leq N-1. \end{aligned}$$

4 落下回数とコストの評価

本節では、落下回数が最小となる落とし方の中でも、コスト最小のものを求める。利得関数を

$$R(x, u) = \begin{pmatrix} r(x, u) \\ q(x, u) \end{pmatrix}, \quad x \in X, u \in U(x)$$

$$K(x) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad x \in X$$

とする。ただし、 $r(x, u)$, $q(x, u)$ はそれぞれ、第2節および第3節で定めた利得関数である。ここで次の順序（辞書式順序）を導入する。

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \lesssim \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \iff a_1 \leq b_1 \text{ または } (a_1 = b_1 \text{ かつ } a_2 \leq b_2)$$

この順序は全順序となり、以後、 $\text{Minimize}^{\lesssim}$ および \min^{\lesssim} は、順序 \lesssim に関する最小化（最小値）を意味するものとする。さらに、

$$a_i = \begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} \in \mathbf{R}^2, \quad i = 1, 2, \dots, n$$

に対して、次の集合演算子を新たに導入する。

$$\bigvee_{i \in \{1, 2, \dots, n\}} a_i = \begin{pmatrix} \bigvee \alpha_i \\ \sum \beta_i \end{pmatrix} = \begin{pmatrix} \alpha_1 \vee \alpha_2 \vee \dots \vee \alpha_n \\ \beta_1 + \beta_2 + \dots + \beta_n \end{pmatrix}$$

すなわち、第1成分については最大値をとり、第2成分については和をとるものである。このとき、落下回数が最小の落とし方の中でコストを最小にするものを求める問題は

$$Q(x_0) \quad \text{Minimize}^{\lesssim} \quad V(x_0; \sigma) \quad \text{subject to} \quad \sigma \in \Sigma$$

と表され、この問題 $Q(x_0)$ の最適値（ベクトル）を $v_0(x_0)$ とおく。ただし、 Σ は一般政策 ([2]) 全体をあらわし、

$$V(x_0; \sigma) = R_0 + \bigvee_{x_1 \in T_1} \left\{ R_1 + \bigvee_{x_2 \in T_2} \left\{ R_2 + \dots + \bigvee_{x_{N-1} \in T_{N-1}} \left\{ R_{N-1} + \bigvee_{x_N \in T_N} K \right\} \dots \right\} \right\},$$

$$x_0 \in X, \sigma = \{\sigma_0, \sigma_1, \dots, \sigma_{N-1}\} \in \Sigma$$

$$\left(R_n = R(x_n, \sigma_n(x_0, x_1, \dots, x_n)), \quad K = K(x_N), \quad T_{n+1} = T(x_n, \sigma_n(x_0, x_1, \dots, x_n)) \right)$$

である。なお、このベクトル値目的関数については

$$V(x_0; \sigma) = \begin{pmatrix} V_1(x_0; \sigma) \\ V_2(x_0; \sigma) \end{pmatrix}$$

であることを補足しておく。

この問題 $Q(x_0)$ に対して再帰式を導きたいのだが、前節までと同様の考え方では（類推では）再帰式を導出することができない。これは、順序 \lesssim がもつ次の性質に起因する。

$$a_1 \lesssim b_1, \quad a_2 \lesssim b_2 \quad \not\Rightarrow \quad \bigvee a_i \lesssim \bigvee b_i$$

このことより、最適政策がマルコフ政策の中には必ずしも存在しないのである。

そこで、第 n 期以降の部分問題として、パラメーター $\lambda \in \mathbf{R}$ を導入した次の問題群を考え、その最適値を w_n であらわす。

$$w_n(x_n, \lambda) = \min_{\sigma \in \Sigma(n)}^{(\lambda)} \left[R(x_n, u_n) + \sum_{x_{n+1} \in T_{n+1}} \left\{ R(x_{n+1}, u_{n+1}) + \cdots + \sum_{x_N \in T_N} K(x_N) \right\} \right]$$

$$x_n \in X, \lambda \in \mathbf{R}, n = 1, 2, \dots, N-1$$

$$\left(u_n = \sigma_n(x_n), u_{n+1} = \sigma_{n+1}(x_n, x_{n+1}), \dots, u_{N-1} = \sigma_{N-1}(x_n, x_{n+1}, \dots, x_{N-1}) \right)$$

ただし、 $\min^{(\lambda)}$ は第 1 成分が λ 以下のもののみを対象に、第 2 成分で最小化することを意味し、第 2 成分の最小値を与える対象が複数ある場合、第 1 成分が最小となるものを $\min^{(\lambda)}$ の値とする。第 1 成分が λ 以下となる対象がない場合は値なしとし、値なしを $-$ で表す。また、 $\Sigma(n)$ は n 期からはじまる $(N-n)$ 段階の一般政策全体をあらわすものとする。 $n = N$ に対しては

$$w_N(x_N, \lambda) = \begin{cases} K(x_N) & \langle K(x_N) \rangle_1 \leq \lambda \\ - & \text{その他} \end{cases}$$

ただし、ベクトル $a = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ に対して、その第 i 成分をあらわす記号を

$$\langle a \rangle_i = a_i, \quad i = 1, 2$$

と定義する。

このとき、元の問題の最適値との間に次の関係が成り立つ。

定理 4.1 $\lambda = 1, 2, \dots$ に対し、 $w_0(x_0, \lambda)$ が値なしにならない、最小の λ を $\hat{\lambda}$ とするとき

$$v_0(x_0) = w_0(x_0, \hat{\lambda})$$

である。

証明 $m = \min_{\sigma \in \Sigma} \langle V(x_0; \sigma) \rangle_1$ とおくと

$$\begin{aligned} v_0(x_0) &= \min_{\sigma \in \Sigma} \tilde{V}(x_0; \sigma) \\ &= \left(\begin{array}{c} m \\ \min_{\sigma \in \Sigma; \langle V(x_0; \sigma) \rangle_1 = m} \langle V(x_0; \sigma) \rangle_2 \end{array} \right) = \left(\begin{array}{c} m \\ \min_{\sigma \in \Sigma; \langle V(x_0; \sigma) \rangle_1 \leq m} \langle V(x_0; \sigma) \rangle_2 \end{array} \right) \\ &= \min_{\sigma \in \Sigma}^{(m)} V(x_0; \sigma) = w_0(x_0, m) \end{aligned}$$

また、 $\hat{\lambda}$ の定め方により

$$\hat{\lambda} = \min_{\sigma \in \Sigma} \langle V(x_0; \sigma) \rangle_1 = m$$

である。 □

さらに、次の再帰式が成り立つ。

定理 4.2

$$w_N(x, \lambda) = \begin{cases} K(x) & \langle K(x) \rangle_1 \leq \lambda \\ - & \text{その他} \end{cases}$$

$$w_n(x, \lambda) = \min_{u \in U(x)}^{[2]} \left[R(x, u) + \sum_{y \in T_{n+1}} w_{n+1}(y, \lambda - \langle R(x, u) \rangle_1) \right] \quad (1)$$

$$x \in X, 0 \leq n \leq N-1.$$

ここで、 $\min^{[2]}$ は第 2 成分のみで最小化することを意味し、第 2 成分の最小値を与える対象が複数ある場合、第 1 成分が最小となるものを $\min^{[2]}$ の値とする。なお、 $\min^{[2]}$ の対象に $-$ がある場合、 $-$ は無視され、全ての対象が $-$ の場合は $\min^{[2]}$ の値も $-$ とする。また $a \in \mathbf{R}^2$ に対し

$$a + - = -$$

とし、 \sum_y についてはその対象に 1 つでも値なしがあれば、値なしとする。

証明 まず、 $w_n(x_n, \lambda) = -$ であるときは、(1) の右辺において、すべての $u_n \in U(x_n)$ に対し

$$w_{n+1}(x_{n+1}, \lambda - \langle R(x_n, u_n) \rangle_1) = -$$

を満たす $x_{n+1} \in T_{n+1}$ が存在する。もし、存在しないとすると、任意の $x_{n+1} \in T_{n+1}$ に対して $\{\bar{\sigma}_{n+1}, \bar{\sigma}_{n+2}, \dots, \bar{\sigma}_{N-1}\} \in \Sigma(n+1)$ が存在し、次をみताす。

$$w_{n+1}(x_{n+1}, \lambda - \langle R(x_n, u_n) \rangle_1) = R(x_{n+1}, \bar{u}_{n+1}) + \sum_{x_{n+2} \in T_{n+2}} \left\{ R(x_{n+2}, \bar{u}_{n+2}) + \dots + \sum_{x_{N-1} \in T_{N-1}} \left\{ R(x_{N-1}, \bar{u}_{N-1}) + \sum_{x_N \in T_N} K(x_N) \right\} \dots \right\}_1 \neq - \quad (2)$$

ただし

$$\bar{u}_{n+1} = \bar{\sigma}_{n+1}(x_{n+1}), \bar{u}_{n+2} = \bar{\sigma}_{n+2}(x_{n+1}, x_{n+2}), \dots, \bar{u}_{N-1} = \bar{\sigma}_{N-1}(x_{n+1}, x_{n+2}, \dots, x_{N-1})$$

であり

$$\left\langle R(x_{n+1}, \bar{u}_{n+1}) + \sum_{x_{n+2} \in T_{n+2}} \left\{ R(x_{n+2}, \bar{u}_{n+2}) + \dots + \sum_{x_{N-1} \in T_{N-1}} \left\{ R(x_{N-1}, \bar{u}_{N-1}) + \sum_{x_N \in T_N} K(x_N) \right\} \dots \right\}_1 \right\rangle_1 \leq \lambda - \langle R(x_n, u_n) \rangle_1$$

を満たす。 $x_{n+1} \in T_{n+1}$ は任意であったので

$$\bigvee_{x_{n+1} \in T_{n+1}} \left\langle R(x_{n+1}, u_{n+1}) + \sum_{x_{n+2} \in T_{n+2}} \left\{ R(x_{n+2}, u_{n+2}) + \dots + \sum_{x_{N-1} \in T_{N-1}} \left\{ R(x_{N-1}, u_{N-1}) + \sum_{x_N \in T_N} K(x_N) \right\} \dots \right\}_1 \right\rangle_1 \leq \lambda - \langle R(x_n, u_n) \rangle_1$$

を満たす決定列（政策）が構成できる。このとき

$$\langle R(x_n, u_n) \rangle_1 + \bigvee_{x_{n+1} \in T_{n+1}} \left\langle R(x_{n+1}, u_{n+1}) + \bigvee_{x_{n+2} \in T_{n+2}} \left\{ R(x_{n+2}, u_{n+2}) + \dots + \bigvee_{x_{N-1} \in T_{N-1}} \left\{ R(x_{N-1}, u_{N-1}) + \bigvee_{x_N \in T_N} K(x_N) \right\} \dots \right\} \right\rangle_1 \leq \lambda$$

$$\langle R(x_n, u_n) \rangle_1 + \left\langle \bigvee_{x_{n+1} \in T_{n+1}} R(x_{n+1}, u_{n+1}) + \bigvee_{x_{n+2} \in T_{n+2}} \left\{ R(x_{n+2}, u_{n+2}) + \dots + \bigvee_{x_{N-1} \in T_{N-1}} \left\{ R(x_{N-1}, u_{N-1}) + \bigvee_{x_N \in T_N} K(x_N) \right\} \dots \right\} \right\rangle_1 \leq \lambda$$

$$\left\langle R(x_n, u_n) + \bigvee_{x_{n+1} \in T_{n+1}} \left\{ R(x_{n+1}, u_{n+1}) + \bigvee_{x_{n+2} \in T_{n+2}} \left\{ R(x_{n+2}, u_{n+2}) + \dots + \bigvee_{x_{N-1} \in T_{N-1}} \left\{ R(x_{N-1}, u_{N-1}) + \bigvee_{x_N \in T_N} K(x_N) \right\} \dots \right\} \right\} \right\rangle_1 \leq \lambda$$

よって、左辺 $\langle \cdot \rangle_1$ 内は $w_n(x_n, \lambda)$ において $\min^{(\lambda)}$ の対象となりうる。すなわち

$$w_n(x_n, \lambda) \neq -$$

これは仮定に矛盾する。ゆえに各 $u_n \in U(x_n)$ に対して

$$\exists x_{n+1} \in T_{n+1}; w_{n+1}(x_{n+1}, \lambda - \langle R(x_n, u_n) \rangle_1) = -$$

である。したがって

$$\bigvee_{x_{n+1} \in T_{n+1}} w_{n+1}(x_{n+1}, \lambda - \langle R(x_n, u_n) \rangle_1) = -$$

となり

$$R(x_n, u_n) + \bigvee_{x_{n+1} \in T_{n+1}} w_{n+1}(x_{n+1}, \lambda - \langle R(x_n, u_n) \rangle_1) = -$$

を得る。これがすべての $u_n \in U(x_n)$ についていえるので

$$\min_{u_n \in U(x_n)}^{[2]} \left[R(x_n, u_n) + \bigvee_{x_{n+1} \in T_{n+1}} w_{n+1}(x_{n+1}, \lambda - \langle R(x_n, u_n) \rangle_1) \right] = -$$

となり再帰式は両辺値なしとして成り立つ。

次に、 $w_n(x_n, \lambda) \neq -$ であるときは、最小値を与える $\{\bar{\sigma}_n, \bar{\sigma}_{n+1}, \dots, \bar{\sigma}_{N-1}\} \in \Sigma(n)$ が存在し

$$w_n(x_n, \lambda) = R(x_n, \bar{u}_n) + \bigvee_{x_{n+1} \in T_{n+1}} \left\{ R(x_{n+1}, \bar{u}_{n+1}) + \dots + \bigvee_{x_{N-1} \in T_{N-1}} \left\{ R(x_{N-1}, \bar{u}_{N-1}) + \bigvee_{x_N \in T_N} K(x_N) \right\} \dots \right\}$$

を満たす。ただし

$$\bar{u}_n = \bar{\sigma}_n(x_n), \quad \bar{u}_{n+1} = \bar{\sigma}_{n+1}(x_n, x_{n+1}), \dots, \quad \bar{u}_{N-1} = \bar{\sigma}_{N-1}(x_n, x_{n+1}, \dots, x_{N-1})$$

であり、

$$\left\langle R(x_n, \bar{u}_n) + \sum_{x_{n+1} \in T_{n+1}} \left\{ R(x_{n+1}, \bar{u}_{n+1}) + \dots + \sum_{x_{N-1} \in T_{N-1}} \left\{ R(x_{N-1}, \bar{u}_{N-1}) + \sum_{x_N \in T_N} K(x_N) \right\} \dots \right\} \right\rangle_1 \leq \lambda \quad (3)$$

も満たす。ここで

$$\lambda(x_{n+1}) := \left\langle R(x_{n+1}, \bar{u}_{n+1}) + \sum_{x_{n+2} \in T_{n+2}} \left\{ R(x_{n+2}, \bar{u}_{n+2}) + \dots + \sum_{x_{N-1} \in T_{N-1}} \left\{ R(x_{N-1}, \bar{u}_{N-1}) + \sum_{x_N \in T_N} K(x_N) \right\} \dots \right\} \right\rangle_1, \quad x_{n+1} \in T_{n+1} \quad (4)$$

とおくとき、 $w_{n+1}(x_{n+1}, \lambda(x_{n+1}))$ は値なしにならず、部分問題の定義により

$$\left\langle w_{n+1}(x_{n+1}, \lambda(x_{n+1})) \right\rangle_i = \left\langle R(x_{n+1}, \bar{u}_{n+1}) + \sum_{x_{n+2} \in T_{n+2}} \left\{ R(x_{n+2}, \bar{u}_{n+2}) + \dots + \sum_{x_{N-1} \in T_{N-1}} \left\{ R(x_{N-1}, \bar{u}_{N-1}) + \sum_{x_N \in T_N} K(x_N) \right\} \dots \right\} \right\rangle_i, \quad i = 1, 2 \quad (5)$$

が成り立つ。

また、式(3)より $x_{n+1} \in T_{n+1}$ に対し

$$\lambda(x_{n+1}) \leq \lambda - \langle R(x_n, \bar{u}_n) \rangle_1$$

これは $w_{n+1}(x_{n+1}, \lambda - \langle R(x_n, \bar{u}_n) \rangle_1)$ における最小化の対象となるベクトル全体は、 $w_{n+1}(x_{n+1}, \lambda(x_{n+1}))$ におけるそれを含むことを意味する。よって

$$\left\langle w_{n+1}(x_{n+1}, \lambda(x_{n+1})) \right\rangle_1 \leq \left\langle w_{n+1}(x_{n+1}, \lambda - \langle R(x_n, \bar{u}_n) \rangle_1) \right\rangle_1 \quad (6)$$

$$\left\langle w_{n+1}(x_{n+1}, \lambda(x_{n+1})) \right\rangle_2 \geq \left\langle w_{n+1}(x_{n+1}, \lambda - \langle R(x_n, \bar{u}_n) \rangle_1) \right\rangle_2 \quad (7)$$

が成り立つ。ここで、

$$\left\langle w_{n+1}(\tilde{x}_{n+1}, \lambda(\tilde{x}_{n+1})) \right\rangle_1 < \left\langle w_{n+1}(\tilde{x}_{n+1}, \lambda - \langle R(x_n, \bar{u}_n) \rangle_1) \right\rangle_1$$

を満たす $\tilde{x}_{n+1} \in T_{n+1}$ は存在しないことが示されるので、(6) は等号で成り立ち

$$\left\langle \sum_{x_{n+1} \in T_{n+1}} w_{n+1}(x_{n+1}, \lambda - \langle R(x_n, \bar{u}_n) \rangle_1) \right\rangle_1 = \left\langle \sum_{x_{n+1} \in T_{n+1}} w_{n+1}(x_{n+1}, \lambda(x_{n+1})) \right\rangle_1$$

よって、 $i = 1$ に対し (5) を用いると

$$\begin{aligned} & \left\langle \sum_{x_{n+1} \in T_{n+1}} w_{n+1}(x_{n+1}, \lambda - \langle R(x_n, \bar{u}_n) \rangle_1) \right\rangle_1 \\ & \leq \left\langle \sum_{x_{n+1} \in T_{n+1}} \left\{ R(x_{n+1}, \bar{u}_{n+1}) + \cdots + \sum_{x_{N-1} \in T_{N-1}} \left\{ R(x_{N-1}, \bar{u}_{N-1}) + \sum_{x_N \in T_N} K(x_N) \right\} \cdots \right\} \right\rangle_1 \quad (8) \end{aligned}$$

である。

次に、(7) および $i = 2$ に対する (5) を用いると、

$$\begin{aligned} & \left\langle \sum_{x_{n+1} \in T_{n+1}} w_{n+1}(x_{n+1}, \lambda - \langle R(x_n, \bar{u}_n) \rangle_1) \right\rangle_2 \leq \left\langle \sum_{x_{n+1} \in T_{n+1}} w_{n+1}(x_{n+1}, \lambda(x_{n+1})) \right\rangle_2 \\ & \leq \left\langle \sum_{x_{n+1} \in T_{n+1}} \left\{ R(x_{n+1}, \bar{u}_{n+1}) + \cdots + \sum_{x_{N-1} \in T_{N-1}} \left\{ R(x_{N-1}, \bar{u}_{N-1}) + \sum_{x_N \in T_N} K(x_N) \right\} \cdots \right\} \right\rangle_2 \quad (9) \end{aligned}$$

が成り立つ。よって式 (8) および式 (9) より

$$\begin{aligned} & \left\langle R_n(x_n, \bar{u}_n) + \sum_{x_{n+1} \in T_{n+1}} w_{n+1}(x_{n+1}, \lambda - \langle R(x_n, \bar{u}_n) \rangle_1) \right\rangle_i \\ & \leq \left\langle R_n(x_n, \bar{u}_n) + \sum_{x_{n+1} \in T_{n+1}} \left\{ R(x_{n+1}, \bar{u}_{n+1}) + \right. \right. \\ & \quad \left. \left. \cdots + \sum_{x_{N-1} \in T_{N-1}} \left\{ R(x_{N-1}, \bar{u}_{N-1}) + \sum_{x_N \in T_N} K(x_N) \right\} \cdots \right\} \right\rangle_i \\ & = \left\langle w_n(x_n, \lambda) \right\rangle_i, \quad i = 1, 2 \quad (10) \end{aligned}$$

が成り立つ。したがって

$$\left\langle \min_{u_n \in U(x_n)}^{[2]} \left[R_n(x_n, u_n) + \sum_{x_{n+1} \in T_{n+1}} w_{n+1}(x_{n+1}, \lambda - \langle R(x_n, u_n) \rangle_1) \right] \right\rangle_2 \leq \left\langle w_n(x_n, \lambda) \right\rangle_2 \quad (11)$$

さらに、上式において逆向きの不等号も成り立つことを示す。左辺に対し、 $u_n^* \in U(x_n)$ が存在し

$$\begin{aligned} & \min_{u_n \in U(x_n)}^{[2]} \left[R_n(x_n, u_n) + \sum_{x_{n+1} \in T_{n+1}} w_{n+1}(x_{n+1}, \lambda - \langle R(x_n, u_n) \rangle_1) \right] \\ & = R_n(x_n, u_n^*) + \sum_{x_{n+1} \in T_{n+1}} w_{n+1}(x_{n+1}, \lambda - \langle R(x_n, u_n^*) \rangle_1) \end{aligned}$$

を満たす。さらに、各 $x_{n+1} \in T_{n+1}$ に対して $\{\hat{\sigma}_{n+1}, \hat{\sigma}_{n+2}, \dots, \hat{\sigma}_{N-1}\} \in \Sigma(n+1)$ が存在して

$$\begin{aligned} & w_{n+1}(x_{n+1}, \lambda - \langle R(x_n, u_n^*) \rangle_1) \\ & = R(x_{n+1}, \hat{u}_{n+1}) + \sum_{x_{n+2} \in T_{n+2}} \left\{ R(x_{n+2}, \hat{u}_{n+2}) + \cdots + \sum_{x_{N-1} \in T_{N-1}} \left\{ R(x_{N-1}, \hat{u}_{N-1}) + \sum_{x_N \in T_N} K(x_N) \right\} \cdots \right\} \end{aligned}$$

$$\hat{u}_{n+1} = \hat{\sigma}_{n+1}(x_{n+1}), \quad \hat{u}_{n+2} = \hat{\sigma}_{n+2}(x_{n+1}, x_{n+2}), \dots, \quad \hat{u}_{N-1} = \hat{\sigma}_{N-1}(x_{n+1}, \dots, x_{N-1})$$

を満たす。このとき

$$\left\langle R(x_{n+1}, \hat{u}_{n+1}) + \sum_{x_{n+2} \in T_{n+2}} \left\{ R(x_{n+2}, \hat{u}_{n+2}) + \dots + \sum_{x_{N-1} \in T_{N-1}} \left\{ R(x_{N-1}, \hat{u}_{N-1}) + \sum_{x_N \in T_N} K(x_N) \right\} \dots \right\} \right\rangle_1 \leq \lambda - \langle R(x_n, u_n^*) \rangle_1$$

なので

$$\left\langle R(x_n, u_n^*) + \sum_{x_{n+1} \in T_{n+1}} \left\{ R(x_{n+1}, \hat{u}_{n+1}) + \sum_{x_{n+2} \in T_{n+2}} \left\{ R(x_{n+2}, \hat{u}_{n+2}) + \dots + \sum_{x_{N-1} \in T_{N-1}} \left\{ R(x_{N-1}, \hat{u}_{N-1}) + \sum_{x_N \in T_N} K(x_N) \right\} \dots \right\} \right\} \right\rangle_1 \leq \lambda$$

したがって、左辺 $\langle \cdot \rangle_1$ 内は $w_n(x_n, \lambda)$ において $\min^{(\lambda)}$ の対象となりうるので

$$\left\langle w_n(x_n, \lambda) \right\rangle_2 \leq \left\langle R(x_n, u_n^*) + \sum_{x_{n+1} \in T_{n+1}} \left\{ R(x_{n+1}, \hat{u}_{n+1}) + \sum_{x_{n+2} \in T_{n+2}} \left\{ R(x_{n+2}, \hat{u}_{n+2}) + \dots + \sum_{x_{N-1} \in T_{N-1}} \left\{ R(x_{N-1}, \hat{u}_{N-1}) + \sum_{x_N \in T_N} K(x_N) \right\} \dots \right\} \right\} \right\rangle_2$$

が成り立つ。よって、式(11)より、

$$\left\langle w_n(x_n, \lambda) \right\rangle_2 = \left\langle \min_{u_n \in U(x_n)}^{[2]} \left[R_n(x_n, u_n) + \sum_{x_{n+1} \in T_{n+1}} w_{n+1}(x_{n+1}, \lambda - \langle R(x_n, u_n) \rangle_1) \right] \right\rangle_2$$

を得る。このとき、 $\min^{(\lambda)}$ の定義より、

$$\left\langle w_n(x_n, \lambda) \right\rangle_1 \leq \left\langle \min_{u_n \in U(x_n)}^{[2]} \left[R_n(x_n, u_n) + \sum_{x_{n+1} \in T_{n+1}} w_{n+1}(x_{n+1}, \lambda - \langle R(x_n, u_n) \rangle_1) \right] \right\rangle_1$$

が成り立ち、さらに、式(10)においても $i = 2$ に対して等号が成り立つことから、 $\min^{[2]}$ の定義より、第1成分について

$$\left\langle \min_{u_n \in U(x_n)}^{[2]} \left[R_n(x_n, u_n) + \sum_{x_{n+1} \in T_{n+1}} w_{n+1}(x_{n+1}, \lambda - \langle R(x_n, u_n) \rangle_1) \right] \right\rangle_1 \leq \left\langle w_n(x_n, \lambda) \right\rangle_1$$

も成り立つ。したがって

$$\left\langle w_n(x_n, \lambda) \right\rangle_1 = \left\langle \min_{u_n \in U(x_n)}^{[2]} \left[R_n(x_n, u_n) + \sum_{x_{n+1} \in T_{n+1}} w_{n+1}(x_{n+1}, \lambda - \langle R(x_n, u_n) \rangle_1) \right] \right\rangle_1$$

が成り立ち、再帰式

$$w_n(x_n, \lambda) = \min_{u_n \in U(x_n)}^{[2]} \left[R_n(x_n, u_n) + \sum_{x_{n+1} \in T_{n+1}} w_{n+1}(x_{n+1}, \lambda - \langle R(x_n, u_n) \rangle_1) \right]$$

を得る。 □

最後に最適政策の構成について述べておく。再帰式において、最小ベクトル $w_n(x, \lambda)$ を与える決定 (の一つ) をパラメータ付マルコフ決定関数 $\pi_n^*(x, \lambda)$ であらわすとき ($n = 0, 1, \dots, N-1$)、問題 $Q(x_0)$ に対する最適一般政策 $\sigma^* = \{\sigma_0^*, \sigma_1^*, \dots, \sigma_{N-1}^*\} \in \Sigma$ は次のように構成される：

$$\begin{aligned} \sigma_0^*(x_0) &= \pi_0^*(x_0, \lambda_0), \quad \lambda_0 = \hat{\lambda} \\ \sigma_n^*(x_0, x_1, \dots, x_n) &= \pi_n^*(x_n, \lambda_n), \quad \lambda_n = \lambda_{n-1} - r(x_{n-1}, \sigma_{n-1}^*(x_0, x_1, \dots, x_{n-1})) \\ & \qquad \qquad \qquad n = 1, 2, \dots, N-1 \end{aligned}$$

ただし、 $\hat{\lambda}$ は定理 4.1 で定めたものとする。

References

- [1] T. Fujita, T. Ueno and S. Iwamoto A Nondeterministic Dynamic Programming Model Proceedings of KES 2004 Part 2, Lecture Notes in Artificial Intelligence, Vol. 3214, Springer-Verlag, Berlin, 2004, pp. 1208-1214
- [2] S. Iwamoto, T. Ueno and T. Fujita, Controlled Markov Chains with Utility Functions, Ed. H. Zhenting, J. A. Filar and A. Chen, Markov Processes and Controlled Markov Chains, Chap.8, 135-148, Kluwer, 2002
- [3] M. Sniedovich, OR/MS Games: 4. The Joy of Egg-Dropping in Braunschweig and Hong Kong, INFORMS Transactions on Education, Vol. 4(1)(2003), pp.48-64