区間値型マルコフ決定過程の
メインテナンスモデルへの適用例

秋田県立大学　システム科学技術研究科 *
Graduate School of Systems Science and Technology, Akita Prefectural University
秋田県立大学　システム科学技術学部　経営システム工学科 † ‡
Faculty of Systems Science and Technology, Akita Prefectural University

松森 俊哉（MATSUMORI, Shunya）*　荒谷　洋輔 (ARAYA, Yousuke) †
星野　満博（HOSHINO, Mitsuhiro）‡

1　研究背景・目的

マルコフ決定過程 (Markov decision processes; MDPs) は, 各時刻の状態を観測しながら各状態で意思決定を行う問題の数理的なモデルの一つで, 生産設備の発注指示問題や物流の効率化など幅広い分野で適用されている.

MDP を実問題へ適用する際には, MDP に用いる状態の推移確率の決定や推定が困難な場合が多いことが課題である．そこで, 推移確率を一意に定められない場合を扱うために, Song, 蔵野ら [14, 17] は, 制御マルコフ集合連鎖と呼ばれる新しい決定モデルを導入し, ある集合順序関係の下で割引期待総報酬 (期待総報酬の割引問題) 最適化問題を議論した.

その一方で, 多目的最適化問題・区間解析の自然な拡張である集合最適化問題が 30 年弱で大きく進展した．集合順序関係の導入 [15], 集合最適化問題の提唱 [16](最適解の概念の考察) から研究は始まり, 集合のスカラー化手法の解明が進んだ (詳細は [2, 3] を参照のこと)．また, Ide et al.[10] の研究から, 多目的ロバスト最適化問題は集合最適化問題の応用として変換できることが判明した. 実社会では、旅行のスケジューリングなどの最短経路を探すことと経費削減を両方達成したいなどの複数目的を最適化する例は多い．さらに, 最短経路問題では, 交通渋滞・悪天候などの「不確実性」が存在する．それらを共に考慮するために多目的ロバスト最適化問題を考える必要があり, その理論解析に集合最適化問題は不可欠である.

* (E-mail:　**M25D007@akita-pu.ac.jp**)
† (E-mail:　**y-araya@akita-pu.ac.jp**)
‡ (E-mail:　**hoshino@akita-pu.ac.jp**)

本稿では, 制御マルコフ集合連鎖の適用例として, 機械メンテナンス問題について考える. 4 状態の推移確率行列を用いて区間値型問題として取り扱い, 平均最適政策を導出する. また, 最適長期期待平均報酬も計算する.

2 制御マルコフ集合連鎖の導入

有限マルコフ決定過程とは, 次の 4 つから構成されている.

$$\{S, A, Q, r\}$$

ここで, $S = \{1, 2, \cdots, n\}$ を状態集合, $A = \{1, 2, \cdots, k\}$ を決定集合, $q_{ij}(a)$ を状態 i において決定 a を用いたとき, 状態 j へ推移する確率として, $Q = (q_{ij}(a))$ を推移確率行列とする. $r = r(i, a)$ は直積集合 $S \times A$ 上の報酬関数とする. 標準的な MDPs は状態 $i \in S$ にあり, 決定 $a \in A$ を選択するとき, S と $q(\cdot \mid i, a)$ 上の確率分布に従って選択された新しい状態 $j \in S$ へ移動して, 報酬 $r(i, a)$ を受け取る.

標準的な MDPs の推移確率行列 $Q = (q_{ij}(a))$ は未知で観測できない場合, 区間 $[\underline{q_{i\cdot}}(a), \overline{q_{i\cdot}}(a)]$ から $q_{i\cdot}(a)$ を推定する. このような制御集合マルコフ連鎖と呼ばれる区間値を含む推移確率行列を用いた決定モデルは [7, 8, 14, 17] などによって導入・議論されてきている.

次節以降では, 主に [17] に従って制御マルコフ集合連鎖の記号・概念等を導入する. そこで, 集合順序関係における新しい解釈も示す.

2.1 集合順序関係の導入とその演算・極限

$\mathbb{R}$ を実数全体の集合, $\mathbb{R}^n$ を実数を成分とする n 次の列ベクトル全体の集合, $\mathbb{R}^{n\times m}$ を実数を成分とする $n \times m$ 行列とする. $\mathcal{V}$ を $\mathbb{R}^n$ の空でない部分集合の族とする. ベクトル d の転置を d^T と表記する. $n \times 1$ 行列はベクトル, 1×1 行列は実数とみなす. $\mathbb{R}^n_+$ を $\mathbb{R}^n$ の非負象限の集合, $\mathbb{R}^{n\times m}_+$ を $\mathbb{R}^{n\times m}$ の非負象限の集合とする. $\mathbb{R}_+$ 内の有界閉区間内の集合を $C(\mathbb{R}_+)$, $C(\mathbb{R}_+)$ の n 次元列ベクトルの集合を $C(\mathbb{R}_+)^n$ と表記する.

$$C(\mathbb{R}_+)^n = \{D = (D_1, D_2, \ldots, D_n)^T | D_i \in C(\mathbb{R}_+), (1 \le i \le n)\}$$

$D, E \in C(\mathbb{R}_+)^n$, $\lambda \ge 0$ に対して, 以下の代数演算を定義する.

$$D + E := \{d + e \mid d \in D, e \in E\}, \qquad \lambda D := \{\lambda d \mid d \in D\}.$$

加法の特別な場合として, $D \in C(\mathbb{R}_+)^n$, $h \in \mathbb{R}^n_+$ について, $h + D := \{h + d \mid d \in D\}$ とする. ここで $\underline{d} := (\underline{d}_1, \ldots, \underline{d}_n)^T, \overline{d} := (\overline{d}_1, \ldots, \overline{d}_n)^T$, $[\underline{d}, \overline{d}] := \{d \mid d \in \mathbb{R}^n_+, \underline{d} \le d \le \overline{d}\}$ と定義すると, $D = ([\underline{d}_1, \overline{d}_1], [\underline{d}_2, \overline{d}_2] \ldots [\underline{d}_n, \overline{d}_n])^T$ は $D = [\underline{d}, \overline{d}]$ と表される.

任意の部分集合 $D_1 \subset \mathbb{R}^{1\times n}_+$ と $D_2 \in C(\mathbb{R}_+)^n$ に対して, 積 $D_1 D_2$ は以下で定義される.

$$D_1 D_2 := \{d_1 d_2 \mid d_1 \in D_1, d_2 \in D_2\}$$

任意の $\underline{A}=(\underline{a}_{ij}),\ \overline{A}=(\overline{a}_{ij})\in\mathbb{R}_+^{n\times m},\ \underline{A}\le\overline{A}$ に対して, $\langle\underline{A},\overline{A}\rangle$ を以下のように定義する.

$$\langle\underline{A},\overline{A}\rangle=\{Q=(q_{ij})\in\mathbb{R}_+^{n\times m}|\underline{a}_{ij}\le q_{ij}\le\overline{a}_{ij},\ q_{ij}\ge 0,\ \sum_{j=1}^{m}q_{ij}=1\ (1\le i\le n, 1\le j\le m)\}$$

$n\times n$ 個の要素を持つすべての区間値型行列の集合を M_n とする.

$$M_n=\{\langle\underline{Q},\overline{Q}\rangle\mid\langle\underline{Q},\overline{Q}\rangle\ne\emptyset,\quad\underline{Q}\preceq\overline{Q},\quad\underline{Q},\overline{Q}\in\mathbb{R}_+^{n\times n}\}$$

$\mathcal{Q}_1,\mathcal{Q}_2\in M_n$ に対して, $\mathcal{Q}_1,\mathcal{Q}_2$ の積を以下で定義する.

$$\mathcal{Q}_1\mathcal{Q}_2:=\{Q_1Q_2\mid Q_1\in\mathcal{Q}_1,Q_2\in\mathcal{Q}_2\}$$

$\mathcal{Q}\in M_n$ の多重積は, 帰納的な $\mathcal{Q}^k$ で定義できる.

$$\mathcal{Q}^k:=\mathcal{Q}^{k-1}\mathcal{Q}\quad(k\ge 2).$$

補題 2.1 ([14])**.** 次が成り立つ.

(1) 任意の $\mathcal{Q}\in M_n$ はベクトル空間 $\mathbb{R}^{n\times n}$ 内の凸多面体である.
(2) コンパクトな部分集合 $D_1\subset\mathbb{R}_+^{1\times n}$ と $D_2\in C(\mathbb{R}_+)^n$ に対して, $D_1D_2\in C(\mathbb{R}_+)$ である.

定義 2.2 (極大要素)**.** $(\mathcal{Q},\le)$ を前順序集合 (反射律・推移律を満たす集合) とする. $\mathcal{A}$ を $\mathcal{Q}$ の空でない集合, $\bar{A}\in\mathcal{A}$ とする. $\mathcal{A}$ の極大要素を以下で定義する.

($\star$) $\bar{A}\in\mathcal{A}$ が $\mathcal{A}$ の極大要素であるとは, ある $A\in\mathcal{A}$ に対して $\bar{A}\le A\Rightarrow A\le\bar{A}$ であること.

2 項関係 $\le$ が半順序であるとき, $\bar{A}$ が $\mathcal{A}$ の極大要素であるとは, 任意の $A\in\mathcal{A},\ A\ne\bar{A}$ に対して $\bar{A}\not\le A$ と書き直すことができる.

$\boldsymbol{a},\boldsymbol{b}\in\mathbb{R}^n$ に対して, ベクトル順序と呼ばれるものを以下で定義する.

$$\boldsymbol{a}\le_{\mathbb{R}_+^n}\boldsymbol{b}\iff\boldsymbol{b}-\boldsymbol{a}\in\mathbb{R}_+^n$$

順序錐 $\mathbb{R}_+^n$ は $\mathbb{R}_+^n\cap(-\mathbb{R}_+^n)=\{0\}$ が成り立つため, $\le_{\mathbb{R}_+^n}$ は反射性, 推移性, 反対称性を満たす.

定義 2.3 (極大元)**.** 点 $\bar{z}\in A\subseteq\mathbb{R}^n$ が $\le_{\mathbb{R}_+^n}$ に関して**極大**であるとは, 以下を満たすときである.

$$z\in A,\quad\bar{z}\le_{\mathbb{R}_+^n}z\quad\Longrightarrow\quad z=\bar{z}$$

A の極大元の集合を $\mathrm{Max}A$ とかく. $\mathrm{Min}A$ も同様に定義される.

定義 2.4 (集合順序関係：黒岩-田中-Ha[15, 16])**.** $A,\ B\in\mathcal{V}$ に対して, 以下の集合関係を定義する.

$$[\textbf{lower}]\quad A\preceq_L B\quad\text{by}\quad B\subset A+\mathbb{R}_+^n,\qquad[\textbf{upper}]\quad A\preceq_U B\quad\text{by}\quad A\subset B-\mathbb{R}_+^n,$$

$$A\preceq B\quad\text{by}\quad B\subset A+\mathbb{R}_+^n\quad\text{and}\quad A\subset B-\mathbb{R}_+^n.$$

[13] では, $\preceq$ を **KNY 集合順序関係**と呼んでいる. KNY は, それを提案した研究者の頭文字に由来している.

命題 2.5 ([3]). $A, B, D \in \mathcal{V}$, $\alpha \geq 0$ に対して, 次が成り立つ.

(i) $A \preceq_L [\preceq_U] B \implies A + D \preceq_L [\preceq_U] B + D \quad \textit{and} \quad A \preceq_L [\preceq_U] B \implies \alpha A \preceq_L [\preceq_U] \alpha B.$
(ii) $\preceq_L$ と $\preceq_U$ は, 反射律と推移律が成り立つ.

定義 2.6. $V_1, V_2 \in \mathcal{V}$ とする. $\mathcal{V}$ に次のような同値関係を導入する.

$$V_1 \sim_L V_2 \Longleftrightarrow V_1 \preceq_L V_2 \quad \text{and} \quad V_2 \preceq_L V_1, \qquad V_1 \sim_U V_2 \Longleftrightarrow V_1 \preceq_U V_2 \quad \text{and} \quad V_2 \preceq_U V_1,$$

$$V_1 \sim V_2 \Longleftrightarrow V_1 \preceq V_2 \quad \text{and} \quad V_2 \preceq V_1.$$

同値類の集合をそれぞれ $[\cdot]^{\preceq_L}$, $[\cdot]^{\preceq_U}$, $[\cdot]^{\preceq}$ とかく. 同値関係の定義より次が分かる.

$$A \in [B]^{\preceq_L} \Leftrightarrow A + \mathbb{R}^n_+ = B + \mathbb{R}^n_+, \qquad A \in [B]^{\preceq_U} \Leftrightarrow A - \mathbb{R}^n_+ = B - \mathbb{R}^n_+,$$

$$A \in [B]^{\preceq} \Leftrightarrow A + \mathbb{R}^n_+ = B + \mathbb{R}^n_+ \quad \text{and} \quad A - \mathbb{R}^n_+ = B - \mathbb{R}^n_+.$$

命題 2.7. $A, B \in C(\mathbb{R}_+)^n$ とする. すると, 次が成り立つ.

$$A \in [B]^{\preceq} \iff A = B.$$

Proof. $A, B \in C(\mathbb{R}_+)^n$ とする. すると, $a_1 \leq_{\mathbb{R}^n_+} a_2$, $b_1 \leq_{\mathbb{R}^n_+} b_2$ を満たす $a_1, a_2, b_1, b_2 \in \mathbb{R}^n$ が存在して, A, B は次のように表現される.

$$A = [a_1, a_2] := \left\{ y \in Y \,\middle|\, a_1 \leq_{\mathbb{R}^n_+} y \leq_{\mathbb{R}^n_+} a_2 \right\}, \quad B = [b_1, b_2] := \left\{ y \in Y \,\middle|\, b_1 \leq_{\mathbb{R}^n_+} y \leq_{\mathbb{R}^n_+} b_2 \right\}.$$

[12] より, 次が言える.

$$[a_1, a_2] \preceq_L [b_1, b_2] \iff a_1 \leq_{\mathbb{R}^n_+} b_1, \qquad [a_1, a_2] \preceq_U [b_1, b_2] \iff a_2 \leq_{\mathbb{R}^n_+} b_2.$$

したがって, 次の関係が言える.

$$A \in [B]^{\preceq_L} \iff [a_1, a_2] + \mathbb{R}^n_+ = [b_1, b_2] + \mathbb{R}^n_+$$

$$\iff A_L := \{y \in \mathbb{R}^n | a_1 \leq_{\mathbb{R}^n_+} y\} = \{z \in \mathbb{R}^n | b_1 \leq_{\mathbb{R}^n_+} z\} =: B_L.$$

$a_1 \neq b_1$ と仮定する. もし $a_1 \leq_{\mathbb{R}^n_+} b_1$ ならば, $A_L \supset B_L$ となり矛盾する. もし $b_1 \leq_{\mathbb{R}^n_+} a_1$ ならば, $A_L \subset B_L$ となり矛盾する. 他の場合は明白に $A_L = B_L$ に反する. よって, $a_1 = b_1$ を得る. 同様にして, 以下も得ることが出来る.

$$A \in [B]^{\preceq_U} \iff A_U := \{y \in \mathbb{R}^n | y \leq_{\mathbb{R}^n_+} a_2\} = \{z \in \mathbb{R}^n | z \leq_{\mathbb{R}^n_+} b_2\} =: B_U$$

そして, $a_2 = b_2$ から $A = B$ を得る. 逆は自明である. □

定義 2.4 の特別な場合として, 以下のものを考えることが出来る.

定義 2.8 ([14, 17])**.** $[c_1, c_2], [d_1, d_2] \in C(\mathbb{R}_+)$ に対して, 集合順序関係 $\preceq$, $\prec$ を以下で定義する.

$$[c_1, c_2] \preceq [d_1, d_2] \iff c_1 \le d_1, c_2 \le d_2$$
$$[c_1, c_2] \prec [d_1, d_2] \iff [c_1, c_2] \preceq [d_1, d_2] \quad \text{and} \quad [c_1, c_2] \neq [d_1, d_2]$$

$v, w \in C(\mathbb{R}_+)^n$ に対して、集合順序関係 $\preceq$, $\prec$ を以下で定義する.

$$v \preceq w \iff v_i \preceq w_i \quad (v_i, w_i \in C(\mathbb{R}_+),\ 1 \le i \le n)$$
$$v \prec w \iff v \prec w,\ v \neq w$$

命題 2.5 (ii), 2.7 から, 定義 2.8 にある $\preceq$ は半順序であることが分かる. 次に, 集合順序関係における極大元を導入する. これは, 多目的最適化における極大元（定義 2.3）の自然な拡張である.

定義 2.9. $\mathcal{A} \subset \mathcal{V}$ とし, $A \in \mathcal{A}$ とする. A が集合順序関係 $\preceq$ に関する $\mathcal{A}$ の**極大元 (maximal element)** であるとは, 以下を満たすものである.

$$A' \in \mathcal{A}, A \preceq A' \Rightarrow A' \preceq A.$$

極小元も同様に定義される.

$[a, b], [c, d] \in C(\mathbb{R}_+)$ に対して, $C(\mathbb{R}_+)$ 上のハウスドルフ距離 δ を以下で定義する.

$$\delta([a, b], [c, d]) := \max\Big(|a - c|, |b - d|\Big)$$

$v, w \in C(\mathbb{R}_+)^n$ に対して, $C(\mathbb{R}_+)^n$ 上の距離を以下で定義する.

$$\Delta(v, w) := \max_{i \in S} \delta(v_i, w_i)$$

$(C(\mathbb{R}_+)^n, \Delta)$ は, 完備距離空間であることが分かっている ([1] の 3.16, 3.17 節を参照). $x \in \mathbb{R}^n$, $D \subset C(\mathbb{R}_+)^n$, $\mathbb{R}^n$ のハウスドルフ距離 δ_2 に対して, $\delta_1(x, D) := \inf_{y \in D} \delta_2(x, y)$ と定義する. 数列 $\{D_k\} \subset C(\mathbb{R}_+)^n$ に対して, 集合の下極限を以下で定義する.

$$\liminf_{k \to \infty} D_k := \{x \in \mathbb{R}^n |\ \limsup_{k \to \infty} \delta_1(x, D_k) = 0\}$$

2.2 ロバスト多目的最適化問題

多目的最適化問題への頑健性 (robustness) の導入を検討する. 不確実な多目的最適化問題を定義するため, 私たちは [4, 5] の考え方を採用する. 問題定式化における不確実性は, 既知の不確実性集合 $\mathcal{U} \subseteq \mathbb{R}^m$ からのシナリオとして与えられると仮定する. また, $g : \mathcal{X} \times \mathcal{U} \to \mathbb{R}^k$, すなわち $\mathcal{U}$ 内のシナリオが g の値に影響を与えることを仮定する. さらに, 実現可能集合 $\mathcal{X}$ は不確実性によるものではなく, 異なるシナリオにおいても不変であると仮定する.

定義 2.10 (ロバスト多目的最適化問題 [6, 10, 11])**.** $\mathcal{X}$ を決定空間, $\mathcal{U}$ を状況集合（不確実集合とも呼ばれる）, $g : \mathcal{X} \times \mathcal{U} \to \mathbb{R}^n$ をパラメータ付きベクトル値の目的関数とする．不確実性を考慮したロバスト多目的最適化を以下のように定式化する．

$$\mathrm{P}(\mathcal{U}) \begin{cases} \text{minimize} \quad g(x, \xi) \\ \text{subject to} \quad x \in \mathcal{X}, \xi \in \mathcal{U} \end{cases}$$

ここで, minimize とは定義 2.3 の意味である．$g_{\mathcal{U}}(x) := \{g(x, \xi) | \xi \in \mathcal{U}\}$ を全てのシナリオの下での全ての目的値の集合とする．

定義 2.11 ([11])**.** 不確実な多目的最適化問題 $\mathcal{P}(\mathcal{U})$ に対し, 実行可能な解 $\bar{x} \in \mathcal{X}$ は, 以下を満たす $x' \in \mathcal{X} \setminus \{\bar{x}\}$ が存在しない場合, alternative set less ordered efficient と呼ばれる．

$$g_{\mathcal{U}}(x') \subseteq g_{\mathcal{U}}(\bar{x}) - \mathbb{R}^k_+ \qquad \text{or} \qquad g_{\mathcal{U}}(\bar{x}) \subseteq g_{\mathcal{U}}(x') + \mathbb{R}^k_+.$$

上記の解は集合順序関係 $\preceq$ の極小元の解と解釈できる．また, ロバスト多目的最適化の文脈では, 与えられた解の最良ケースと最悪ケースを最適化すると [11] で説明されている．

2.3 定式化

$S = \{1, 2, \cdots, n\}$ を状態集合, $A = \{1, 2, \cdots, k\}$ を決定集合とする．それぞれの $(i, a) \in S \times A$ に対して, $\underline{q} = \underline{q}(\cdot \mid i, a) \in \mathbb{R}^{n\times 1}_+$, $\overline{q} = \overline{q}(\cdot \mid i, a) \in \mathbb{R}^{n\times 1}_+$, $\underline{q} \preceq \overline{q}$, $\langle \underline{q}, \overline{q} \rangle \neq \emptyset$ として, $r(\cdot, \cdot) : S \times A \to C(\mathbb{R}_+)^n$ を $0 \preceq r$ を満たす利得写像とする．

F を S から A への関数全体の集合とする．戦略 π は $f_t \in F (t \geq 1)$ となる関数の列 $(f_1, f_2, \cdots)$ である．政策のクラスを Π と表記する．任意の $t \geq 1$ とある $f \in F$ について $f_t = f$ となる政策 $(f_1, f_2, \cdots)$ を f^∞ と表記する．そのような政策は定常政策と呼ばれ, 単に f と表記される．定常政策の集合を Π_F と表記する．

それぞれの $f \in F$ に対して, i 番目の要素が $r(i, f(i))$ となる n 次元列ベクトル $r(f) = (\cdots, r(i, f(i)), \cdots) \in \mathbb{R}^n_+$ とし, $Q(f) := \langle \underline{Q}(f), \overline{Q}(f) \rangle \in M_n$ とする．ただし, $\underline{Q}(f), \overline{Q}(f)$ の (i, j) の成分は $\underline{q}(j \mid i, f(i)) \in \mathbb{R}^{n\times 1}_+$, $\overline{q} = \overline{q}(j \mid i, (f(i)) \in \mathbb{R}^{n\times 1}_+$ とする．

任意の $\pi = (f_1, f_2, \cdots) \in \Pi$ について, $v_1(\pi) = r(f_1)$ と定義して, $v_T(\pi)$ を以下で定義する．$v_T(\pi)$ は区間値の長期期待総報酬と解釈することが出来る．

$$v_T(\pi) := \{r(f_1) + Q_1 r(f_2) + \cdots + Q_1 Q_2 \cdots Q_{T-1}(f_T) | Q_i \in Q(f_i),\ i = 1, \cdots, T-1\}\ (T \geq 2)$$

補題 2.1 より, すべての $T \geq 1$ に対して $v_T(\pi) \in C(\mathbb{R}_+)^n$ が成り立つ．

$$v(\pi) := \liminf_{T\to\infty} \frac{1}{T} v_T(\pi)$$

$v(\pi) \in C(\mathbb{R}_+)^n$ なので, $v(\pi) = [\underline{v}(\pi), \overline{v}(\pi)]$ と表される．定義 2.2 から次の概念を導入する．

定義 2.12 (平均最適政策 [17]). $v(f^*) \prec v(f)$ を満たすような $f \in \Pi_F$ が存在しないとき, 政策 $f^* \in \Pi_F$ は, 平均最適と呼ばれる.

2.2 節から, 上記の定義は最良の場合と最悪の場合を考慮したものであることが分かる.

(仮定 1) 任意の $n \times n$ 確率行列 $Q = (q_{ij})$ と $Q' = (q'_{ij})$ について,

$$\kappa_1 := \max_{f,g\in F} \sup_{Q\in Q(f),Q'\in Q(g)} \left(\frac{1}{2} \max_{i,j} \sum_{l=1}^{n} |q_{il} - q'_{jl}| \right) < 1$$

すべての $i, j \in S$ と $a \in A$ に対して、$\underline{q}(j \mid i, a) > 0$ であれば, 仮定 1 が成り立つ.

(仮定 2) 任意の $f \in F$ に対して, 各 $Q \in Q(f)$ は既約, すなわち, ある $t \geq 1$ に対して $Q^t > 0$.

上記を仮定すると, 縮小写像の不動点定理から次の $(*)$, $(**)$ を満たす $\underline{v}^* \in \mathbb{R}_+$, $\overline{v}^* \in \mathbb{R}_+$, $\underline{h} = (\underline{h_1}, \cdots, \underline{h_n})^T \in \mathbb{R}^n$, $\overline{h} = (\overline{h_1}, \cdots, \overline{h_n})^T \in \mathbb{R}^n$ が存在する.

$$\underline{v}^* + \underline{h_i} = \max_{a\in A} \left(r(i,a) + \min_{q\in q(i,a)} \sum_{j=1}^{n} q(j \mid i,a)\underline{h_j} \right) \qquad (1 \leq i \leq n) \qquad (*)$$

$$A_i := \arg\max_{a\in A} \left(r(i,a) + \min_{q\in q(i,a)} \sum_{j=1}^{n} q(j \mid i,a)\underline{h_j} \right)$$

$$\overline{v}^* + \overline{h_i} = \max_{a\in A_i} \left(r(i,a) + \max_{q\in q(i,a)} \sum_{j=1}^{n} q(j \mid i,a)\overline{h_j} \right) \qquad (1 \leq i \leq n) \qquad (**)$$

$$A_i^* := \arg\max_{a\in A_i} \left(r(i,a) + \max_{q\in q(i,a)} \sum_{j=1}^{n} q(j \mid i,a)\overline{h_j} \right)$$

定理 2.13 ([17]). $e = (1, 1, \cdots, 1)^T$ とする. すべての $i \in S$ に対して $f^*(i) \in A_i^*$ である政策を f^* とすると, f^* は平均最適であり, $v(f^*) = [\underline{v}^* e, \overline{v}^* e]$ である.

定義 2.12 において, [17] は $\prec$ を利用して定義しているが, 著者は $\preceq$ に緩めることが出来るのではないかと予想している. $\preceq$ に緩めた場合でも定理 2.13 が成り立つかは今後の研究課題である.

3 機械メンテナンス問題への適用

機械メンテナンス問題について, 状態数が 4 の推移確率行列の数値例が [9] の 19 章にある。これを基に区間値を持つ推移確率行列を設定することで, 長期期待平均報酬 $v(f^*)$ を（集合順序関係の意味で）最大化する最適化問題を考える. そして, 平均最適政策 f^*, 最適下限値 $\underline{v}^*$, 最適上限値 $\overline{v}^*$ を計算する. 状態空間 S と決定空間 A を以下のように設定した.

$$S = \{1, 2, 3, 4\}, \qquad A = \{1, 2, 3\},$$

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

状態 i	機械の状態
1	新品同様
2	軽度の劣化
3	重度の劣化
4	許容できない

決定 a	状態 i	報酬 $r(i,a)$
1. 何もしない	1	6000
	2	5000
	3	3000
2. 徹底調査	3	2000
3. 交換する	2,3,4	0

長期期待平均報酬 $v(f^*)$ を求める際に用いる政策

政策	政策の内容	$f(1)$	$f(2)$	$f(3)$	$f(4)$
R_a	状態 4 で交換する	1	1	1	3
R_b	状態 4 で交換して, 状態 3 で徹底調査する	1	1	2	3
R_c	状態 3 と状態 4 で交換する	1	1	3	3
R_d	状態 2, 状態 3 と状態 4 で交換する	1	3	3	3
R_e	状態 2 と状態 4 で交換する	1	3	1	3
R_f	状態 3 で徹底調査をして, 状態 2 と状態 4 で交換する	1	3	2	3

各政策ごとに推移確率行列 (q_{ij}), (q_{ij}) の第 2 列と第 4 列の成分を $\frac{1}{16}$ 増やした行列 $(\overline{q_{ij}})$, (q_{ij}) の第 2 列と第 4 列の成分を $\frac{1}{16}$ 減らした行列 $(\underline{q_{ij}})$ を考える.

- 政策 R_a の場合

$$(q_{ij}) = \begin{pmatrix} 0 & \frac{7}{8} & \frac{1}{16} & \frac{1}{16} \\ 0 & \frac{3}{4} & \frac{1}{8} & \frac{1}{8} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad (\underline{q_{ij}}) = \begin{pmatrix} 0 & \frac{13}{16} & \frac{1}{16} & 0 \\ 0 & \frac{11}{16} & \frac{2}{16} & \frac{1}{16} \\ 0 & 0 & \frac{8}{16} & \frac{7}{16} \\ 1 & 0 & 0 & 0 \end{pmatrix}, (\overline{q_{ij}}) = \begin{pmatrix} 0 & \frac{15}{16} & \frac{1}{16} & \frac{2}{16} \\ 0 & \frac{13}{16} & \frac{2}{16} & \frac{3}{16} \\ 0 & 0 & \frac{8}{16} & \frac{9}{16} \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

- 政策 R_e の場合

$$(q_{ij}) = \begin{pmatrix} 0 & \frac{7}{8} & \frac{1}{16} & \frac{1}{16} \\ 1 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad (\underline{q_{ij}}) = \begin{pmatrix} 0 & \frac{13}{16} & \frac{1}{16} & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & \frac{8}{16} & \frac{7}{16} \\ 1 & 0 & 0 & 0 \end{pmatrix}, (\overline{q_{ij}}) = \begin{pmatrix} 0 & \frac{15}{16} & \frac{1}{16} & \frac{2}{16} \\ 1 & 0 & 0 & 0 \\ 0 & 0 & \frac{8}{16} & \frac{9}{16} \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

- 政策 R_b の場合

$$(q_{ij}) = \begin{pmatrix} 0 & \frac{7}{8} & \frac{1}{16} & \frac{1}{16} \\ 0 & \frac{3}{4} & \frac{1}{8} & \frac{1}{8} \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad (\underline{q_{ij}}) = \begin{pmatrix} 0 & \frac{13}{16} & \frac{1}{16} & 0 \\ 0 & \frac{11}{16} & \frac{2}{16} & \frac{1}{16} \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, (\overline{q_{ij}}) = \begin{pmatrix} 0 & \frac{15}{16} & \frac{1}{16} & \frac{2}{16} \\ 0 & \frac{13}{16} & \frac{2}{16} & \frac{3}{16} \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

- 政策 R_c の場合

$$(q_{ij}) = \begin{pmatrix} 0 & \frac{7}{8} & \frac{1}{16} & \frac{1}{16} \\ 0 & \frac{3}{4} & \frac{1}{8} & \frac{1}{8} \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad (\underline{q_{ij}}) = \begin{pmatrix} 0 & \frac{13}{16} & \frac{1}{16} & 0 \\ 0 & \frac{11}{16} & \frac{2}{16} & \frac{1}{16} \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, (\overline{q_{ij}}) = \begin{pmatrix} 0 & \frac{15}{16} & \frac{1}{16} & \frac{2}{16} \\ 0 & \frac{13}{16} & \frac{2}{16} & \frac{3}{16} \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

$$(1)\quad \underline{v}^* + \underline{h_1} = r(1,1) + \min_{q\in\langle \underline{q}(\cdot|1,1),\overline{q}(\cdot|1,1)\rangle} \sum_{j=1}^{4} q(j \mid 1,1)\underline{h_j}$$

$$= 6000 + \min_{\frac{13}{16}\le x\le \frac{15}{16}} \left\{ x\underline{h_2} + \frac{1}{16}\underline{h_3} + \left(\frac{15}{16} - x\right)\underline{h_4} \right\}$$

$$(2)\quad \underline{v}^* + \underline{h_2} = \max\Bigg(r(2,1) + \min_{q\in\langle \underline{q}(\cdot|2,1),\overline{q}(\cdot|2,1)\rangle} \sum_{j=1}^{4} q(j \mid 2,1)\underline{h_j},$$

$$r(2,3) + \min_{q\in\langle \underline{q}(\cdot|2,3),\overline{q}(\cdot|2,3)\rangle} \sum_{j=1}^{4} q(j \mid 2,3)\underline{h_j} \Bigg)$$

$$= \max\left\{ 5000 + \min_{\frac{11}{16}\le x\le \frac{13}{16}} \left\{ x\underline{h_2} + \frac{2}{16}\underline{h_3} + \left(\frac{14}{16} - x\right)\underline{h_4} \right\}, \underline{h_1} \right\}$$

$$(3)\quad \underline{v}^* + \underline{h_3} = \max\Bigg(r(3,1) + \min_{q\in\langle \underline{q}(\cdot|3,1),\overline{q}(\cdot|3,1)\rangle} \sum_{j=1}^{4} q(j \mid 3,1)\underline{h_j},$$

$$r(3,2) + \min_{q\in\langle \underline{q}(\cdot|3,2),\overline{q}(\cdot|3,2)\rangle} \sum_{j=1}^{4} q(j \mid 3,2)\underline{h_j},$$

$$r(3,3) + \min_{q\in\langle \underline{q}(\cdot|3,3),\overline{q}(\cdot|3,3)\rangle} \sum_{j=1}^{4} q(j \mid 3,3)\underline{h_j} \Bigg)$$

$$= \max\left\{ 3000 + \min_{\frac{7}{16}\le x\le \frac{9}{16}} \left\{ x\underline{h_3} + (1-x)\underline{h_4} \right\}, 2000 + \underline{h_2}, \underline{h_1} \right\}$$

$$(4)\quad \underline{v}^* + \underline{h_4} = r(4,3) + \min_{q\in\langle \underline{q}(\cdot|4,3),\overline{q}(\cdot|4,3)\rangle} \sum_{j=1}^{4} q(j \mid 4,3)\underline{h_j} = \underline{h_1}$$

$\underline{h_4} = 0$ と置いて, 連立方程式 (1), (2), (3), (4) を解くと,

$$\underline{h} = (\underline{h_1}, \underline{h_2}, \underline{h_3}, \underline{h_4})^T = \left(\frac{490000}{117}, \frac{112000}{39}, \frac{8000}{117}, 0\right)^T, \qquad \underline{v}^* = 4188.03.$$

このとき, $f^*(1) = 1$, $f^*(2) = 1$, $f^*(3) = 2$, $f^*(4) = 3$ が得られる.

$$(5)\quad \overline{v}^* + \overline{h_1} = r(1,1) + \max_{q\in\langle \underline{q}(\cdot|1,1),\overline{q}(\cdot|1,1)\rangle} \sum_{j=1}^{4} q(j \mid 1,1)\overline{h_j}$$

$$= 6000 + \max_{\frac{13}{16}\le x\le \frac{15}{16}} \left\{ x\overline{h_2} + \frac{1}{16}\overline{h_3} + \left(\frac{15}{16} - x\right)\overline{h_4} \right\}$$

$$(6)\quad \overline{v}^* + \overline{h_2} = \max\Bigg(r(2,1) + \max_{q\in\langle \underline{q}(\cdot|2,1),\overline{q}(\cdot|2,1)\rangle} \sum_{j=1}^{4} q(j \mid 2,1)\overline{h_j},$$

$$r(2,3) + \max_{q\in\langle \underline{q}(\cdot|2,3),\overline{q}(\cdot|2,3)\rangle} \sum_{j=1}^{4} q(j \mid 2,3)\overline{h_j}\Bigg)$$

$$= \max\left\{5000 + \max_{\frac{11}{16}\le x\le \frac{13}{16}} \left\{x\overline{h_2} + \frac{2}{16}\overline{h_3} + \left(\frac{14}{16} - x\right)\overline{h_4}\right\}, \overline{h_1}\right\}$$

$$(7)\quad \overline{v}^* + \overline{h_3} = \max\Bigg(r(3,1) + \max_{q\in\langle \underline{q}(\cdot|3,1),\overline{q}(\cdot|3,1)\rangle} \sum_{j=1}^{4} q(j \mid 3,1)\overline{h_j},$$

$$r(3,2) + \max_{q\in\langle \underline{q}(\cdot|3,2),\overline{q}(\cdot|3,2)\rangle} \sum_{j=1}^{4} q(j \mid 3,2)\overline{h_j},$$

$$r(3,3) + \max_{q\in\langle \underline{q}(\cdot|3,3),\overline{q}(\cdot|3,3)\rangle} \sum_{j=1}^{4} q(j \mid 3,3)\overline{h_j}\Bigg)$$

$$= \max\left\{3000 + \max_{\frac{7}{16}\le x\le \frac{9}{16}} \left\{x\overline{h_3} + (1-x)\overline{h_4}\right\}, 2000 + \overline{h_2}, \overline{h_1}\right\}$$

$$(8)\quad \overline{v}^* + \overline{h_4} = r(4,3) + \max_{q\in\langle \underline{q}(\cdot|4,3),\overline{q}(\cdot|4,3)\rangle} \sum_{j=1}^{4} q(j \mid 4,3)\overline{h_j} = \overline{h_1}$$

$\overline{h_4} = 0$ と置いて, 連立方程式 (5), (6), (7), (8) を解くと,

$$\overline{h} = (\overline{h_1}, \overline{h_2}, \overline{h_3}, \overline{h_4})^T = \left(\frac{1442000}{321}, \frac{336000}{107}, \frac{20800}{321}, 0\right)^T, \qquad \overline{v}^* = 4492.21.$$

このとき, $f^*(1) = 1,\ f^*(2) = 1,\ f^*(3) = 2,\ f^*(4) = 3$ が得られる.

＜結論＞

区間値を設けていない推移確率 (q_{ij}) の下での長期期待平均報酬 V は,

$$V = 4333.3$$

と求められる. このときの最適政策は

$$f^*(1) = 1, f^*(2) = 1, f^*(3) = 2, f^*(4) = 3$$

となる. それに対して, 区間値を設けた平均最適政策は

$$f^*(1) = 1, f^*(2) = 1, f^*(3) = 2, f^*(4) = 3$$

であり, 両方の最適政策は一致することが確認できる. 最適区間は,

$$v(f^*) = [\underline{v}^* e, \overline{v}^* e] = [4188.03e, 4492.21e]$$

となり, 求められた区間の中に 4333.3 が含まれていることが分かった. 尚, この例において仮定 1 は満たされていない.

参考文献

[1] C. D Aliprantis, K. C. Border, *Infinite dimensional analysis*, Springer, Berlin, 2006.

[2] 荒谷 洋輔, 集合最適化問題における非線形スカラー化手法：*20* 年余の進展, 京都大学数理解析研究所講究録, 2242, 26–38, 2023 年 1 月.

[3] Y. Araya, *Conjugate duality in set optimization via nonlinear scalarization*, J. Optim. Theory Appl. 199 (2023), no.2, 466–498.

[4] A. Ben-Tal, A. Nemirovski, *Robust convex optimization*, Math. Oper. Res. 23 (1998), no. 4, 769–805.

[5] A. Ben-Tal, A. Nemirovski, *Robust optimization –methodology and applications*, Math. Program. 92 (2002), no. 3, Ser. B, 453–480.

[6] M. Ehrgott, J. Ide, A. Schöbel, *Minmax robustness for multi-objective optimization problems*, European J. Oper. Res. 239 (2014), 17–31.

[7] D. J. Hartfiel, *Markov set-chains*, Lecture Notes in Math., 1695, Springer-Verlag, Berlin, 1998.

[8] D. J. Hartfiel, *Component bounds for Markov set-chain limiting sets*, J. Statist. Comput. Simul. 38, (1991), 15–24.

[9] F. Hillier, G. J. Lieberman, *Introduction to Operations Research*, 8th edition, McGraw-Hill Higher Education.

[10] J. Ide, E. Köbis, D. Kuroiwa, A. Schöbel, C. Tammer, *The relationship between multi-objective robustness concepts and set-valued optimization*, Fixed Point Theory Appl. (2014), 2014:83, 20.

[11] J. Ide, A. Schöbel, *Robustness for uncertain multi-objective optimization: a survey and analysis of different concepts*, OR Spectrum **38** (2016), 235–271.

[12] J. Jahn, T.X.D. Ha, *New order relations in set optimization*, J. Optim. Theory Appl. 148(2), (2011), 209–236.

[13] J. Jahn, *Vector optimization*, Springer-Verlag, Berlin, 2004.

[14] M. Kurano, J. Song, M. Hosaka, Y. Huang, *Controlled Markov Set-Chains with Discounting*, J. Appl. Prob. 35, (1998), 293–302.

[15] D. Kuroiwa, T. Tanaka and T. X. D. Ha, On cone convexity of set-valued maps, Nonlinear Analysis 30 (1997), 1487–1496.

[16] D. Kuroiwa, *On set-valued optimization*, Nonlinear Anal. 47(2), (2001), 1395–1400.

[17] J. Song, M. Hosaka, M. Kurano, *A Span Seminorm Approach to Controlled Markov Set-Chains*, 千葉大学教育学部研究紀要. III, 自然科学編.46, (1998), 13–23.

[18] D. J. White, *Markov decision processes*, John Wiley & Sons, Ltd., Chichester, 1993.