1998年度日本オペレーションズ・リサーチ学会
秋季研究発表会

あるクラスのマルコフ決定過程問題へのLemke法の適用

01605850 日本電信電話(株) *松林 伸生 MATSUBAYASHI Nobuo
01400760 慶応義塾大学 西野 寿一 NISHINO hisakazu

1 はじめに

線形相補性問題を解くことに帰着できる問題としては、従来より、凸２次計画の最適点を求める問題や、双行列ゲームにおけるNash均衡点を求める問題などがよく知られており、また前者については、Lemkeの相補掃き出しアルゴリズム（Lemke法）によって解けることが知られている。

本稿では、これらとは異なり、

（１）最適停止問題

（２）動的最短経路問題 [3]

等に現れる, 従来はマルコフ決定過程の枠組で取扱われていた最適意思決定問題を扱う。具体的には、

有限個（n個）の状態が定常マルコフ過程に従って推移している時、あるアクションを起こして所定の報酬（または損失）を得て決定値とするか、または所定の（定数項を加えるか、報酬・損失のα倍で表現される）ペナルティーを払ってアクションを一期間保留した時の期待報酬（または損失）を求め、どちらか有利な方を選択するという二者択一型の意思決定問題

を考える。

本稿では、この問題が線形相補性問題として定式化でき、Lemke法を用いて、状態の数（n）以下のステップ数で必ず解けることを示す。

2 モデル

問題MAX：

$$\boldsymbol{x} = \max(\boldsymbol{f}, \alpha \boldsymbol{P}\boldsymbol{x} - \boldsymbol{c}),\ \boldsymbol{c} \geq \boldsymbol{o}$$

ただし、

$$0 < \alpha \leq 1,\ \text{もし、}\boldsymbol{c} = \boldsymbol{o}\ \text{ならば、}\alpha < 1$$

問題MIN：

$$\boldsymbol{x} = \min(\boldsymbol{f}, \alpha \boldsymbol{P}\boldsymbol{x} + \boldsymbol{c}),\ \boldsymbol{c} \geq \boldsymbol{o}$$

ただし、

$$\alpha \geq 1,\ \text{もし、}\boldsymbol{c} = \boldsymbol{o}\ \text{ならば、}\alpha > 1$$

ここで、遷移確率行列$\boldsymbol{P}$はエルゴード行列と仮定する。

3 本モデルに対する従来の解法

本稿で取り扱う問題は、従来より、以下のようなマルコフ決定過程の典型的な解法を適用することによって解けることが知られている。

（１）Successive Approximation 法

（２）Policy Iteration 法

（３）線形計画法

この中では、（２）のPolicy Iteration法がnステップ以下の反復回数で解けることなどから、多くの場合において最良解法とされてきたが、今回のLemke法を用いた解法は、それとComparableであると言えるものである。

4 線形相補性問題としての定式化

Reduced MAX:

$$\begin{aligned} \boldsymbol{z} &\equiv \boldsymbol{x} - \boldsymbol{f} \\ \boldsymbol{w} &\equiv (\boldsymbol{I} - \alpha \boldsymbol{P})\boldsymbol{z} + \{\boldsymbol{c} + (\boldsymbol{I} - \alpha \boldsymbol{P})\boldsymbol{f}\} \end{aligned}$$

とおくことにより、MAX は $\boldsymbol{z} \geq \boldsymbol{o}$，$\boldsymbol{w} \geq \boldsymbol{o}$ $\boldsymbol{z}^T\boldsymbol{w} = 0$ を満たす n-次元ベクトル $\boldsymbol{z}, \boldsymbol{w}$ を見つけよ、という線形相補性問題に帰着される。(Reduced MINも同様にして定式化される)

5　Lemke 法の適用

与えられた線形相補性問題が Lemke 法によって解けるかどうかは、一般に問題が含む行列の構造に依存する。(解けるための必要条件)

1. 任意の定数項（$\boldsymbol{q}$とする）に対し、Lemke 法が*unbounded ray*に抜けて終了することが問題に解がないことを意味するクラス

   ・*Copositive Plus*-matrix（凸２次計画問題)、*Positive Semi Definite*-matrix、Z-matrix（全ての非対角成分が$non-positive$）など。

2. 任意の$\boldsymbol{q}$に対して、必ず問題に解があって、Lemke 法で解けるクラス

   ・P-matrix（全ての *Positive Principal Minors* が正）など。

しかし、本問題における行列$\boldsymbol{I}-\alpha\boldsymbol{P}$は、$Z$-matrix 以外のクラスについては、多くの場合条件を満たしておらず、この事実だけでは、本問題に解が必ずあることや、nステップ以内で解けることなどは証明できない。そこで今回は、小島、西野、関根[2]による、「Piecewise Linear Complementarity Problem に対する拡張された Lemke 法（問題が線形の時は通常の Lemke 法）が解を求めて終了する必要十分条件」を用いて、これらを証明した。

Lemma 5.1 *The extended Lemke's method solves the piecewise linear complementarity problem if and only if there exists a set $B\subset R^n_+$ such that*

1. *every unbounded closed path connected set in R^n_+ which contains $\mathbf{o}\in R^n_+$ intersects with B, and*
2. *for every $\boldsymbol{z}\in B$, we have $E(\boldsymbol{z})\cap F(\boldsymbol{z})\neq\emptyset$ where*

$$
\begin{aligned}
E(\boldsymbol{z}) &\equiv \{\boldsymbol{y}\in R^n_+ \mid \sum_{i=1}^{n} y_i \le \sum_{i=1}^{n} z_i\} \\
F(\boldsymbol{z}) &\equiv \{\boldsymbol{y}\in R^n_+ \mid \boldsymbol{y}^T\boldsymbol{w} < \boldsymbol{z}^T\boldsymbol{w}\}.
\end{aligned}
$$

Theorem 5.1 *Reduced MAX is solvable by Lemke's method*

Proof *let C be any positive number, and define B as*

$$
\begin{aligned}
B &\equiv \{\boldsymbol{z}\in R^n_+ \mid \boldsymbol{z}\le f^*\mathbf{1},\ {}^\exists k\ s.t.\ z_k=f^*\},\\
f^* &\equiv |f_{\max}|+|f_{\min}|+C,
\end{aligned}
$$

以下省略（この場合のBは、一辺がf^*の立方体の、第一象限を２つに分割するへりの部分に当たる）

Theorem 5.2 *If $\boldsymbol{f}>\mathbf{o}$, Reduced MIN is solvable by Lemke's method.*

証明省略。

6　Lemke 法を適用した場合のステップ数

Theorem 6.1 *The Lemke's method applied to Reduced MAX terminates within n iterations.*

Theorem 6.2 *The Lemke's method applied to Reduced MIN with $\boldsymbol{f}>\mathbf{o}$ terminates within n iterations.*

証明は、基底変数中の$\boldsymbol{z}$の成分の数に関する帰納法を用いて行う。すなわち、各ステップにおいて、新しく基底変数になるのは常に$\boldsymbol{z}$の成分であることを示すことによって、最悪のケースでも、全ての$\boldsymbol{z}$の成分が基底に入るnステップ目までで解を求められることが証明される。

参考文献

[1] N.Matsubayashi and H.Nishino, "An Application of Lemke's method to a Class of Markov Decision Problems ", to appear in *European Journal of Operational Research*

[2] M.Kojima, H.Nishino and T.Sekine, "An extension of Lemke's Method to the Piecewise Linear Complementarity Problem", *SIAM Journal on Applied Mathematics* 37 (1976), 513-538.

[3] H.N.Psaraftis and J.N.Tsitsiklis, "Dynamic Shortest Paths in Acyclic Networks with Markovian Arc Costs", *Operations Research* 41-1 (1993), 91-101.