1997年度日本オペレーションズ・リサーチ学会
春季研究発表会

結合型動的計画問題のベクトル・マトリクス計算による解法

02501655　九州大学　藤田敏治　FUJITA Toshiharu

マルコフ決定過程上における結合型評価の最大化問題に対し、効率的な計算手法を紹介する。マルコフ決定過程では、通常各ステージにおける利得の和を評価値として考えるが、ここでは、加法だけでなくより一般の結合法則を満たす演算による評価を考える。なお、結合型動的計画に関しては[2]で詳しく議論されている。

1　結合型動的計画問題

次のデータが与えられているとする：

$N \geq 2$	**終端時刻**
$X = \{s_1, \cdots, s_m\}$	**状態集合**
$U = \{a_1, \cdots, a_l\}$	**決定集合**
$x_n \in X$	時刻 n における**状態**
$u_n \in U$	時刻 n における**決定**
$r_n : X \times U \to \mathbf{R}$	時刻 n における**利得**
$r_G : X \to \mathbf{R}$	**終端利得**
p	**マルコフ推移法則** $p(y\|x,u) \geq 0$ $\sum_{y \in X} p(y\|x,u) = 1$
$\circ : \mathbf{R} \times \mathbf{R} \to \mathbf{R}$	*associative* binary relation $(x \circ y) \circ z = x \circ (y \circ z)$

$y \sim p(\cdot|x,u)$ は現時刻の状態が x, 決定が u であるとき，次の時刻で状態 y へ確率 $p(y|x,u)$ で推移することをあらわす。このとき、次の問題を考える。

$$\begin{aligned} &\mathrm{Max} E[\, r_1(x_1,u_1) \circ \cdots \circ r_N(x_N,u_N) \circ r_G(x_{N+1})] \\ &\quad \text{s.t.}\ \ (\mathrm{i})_\mathrm{n}\ \ x_{n+1} \sim p(\cdot|\, x_n, u_n) \quad 1 \leq n \leq N \\ &\qquad\quad (\mathrm{ii})_\mathrm{n}\ \ u_n \in U \quad 1 \leq n \leq N \end{aligned}$$

ここで E は条件付き確率 $p(x_{n+1}|\, x_n, u_n)$、政策 $\pi = \{\pi_1, \pi_2, \ldots, \pi_N\}$ 及び初期状態 $x_1 \in X$ に依存して定まる $X \times U \times X \times U \times \cdots \times U \times X$ 上の期待値を表す。

この問題に対し次のような部分問題を考える。

$$\begin{aligned} &v^{N+1}(x_{N+1};\lambda) := \lambda \circ r_G(x_{N+1}) \quad x \in X \quad \lambda \in \mathbf{R} \\ &v^\nu(x_\nu;\lambda) \\ &:= \mathrm{Max}\ E[\, \lambda \circ r_\nu(x_\nu,u_\nu) \circ \cdots \circ r_G(x_{N+1}) \\ &\qquad\qquad |\, (\mathrm{i})_\mathrm{n},\ (\mathrm{ii})_\mathrm{n} \quad \nu \leq n \leq N] \\ &\qquad\qquad x \in X \quad \lambda \in \mathbf{R} \quad \nu = 1,2,\cdots,N \end{aligned}$$

このとき，次の再帰式が成り立つ。

定理 1.1

$$\begin{aligned} &v^{N+1}(x;\lambda) = \lambda \circ r_G(x) \qquad x \in X \quad \lambda \in \mathbf{R} \\ &v^\nu(x;\lambda) = \mathop{\mathrm{Max}}_{u \in U} \sum_{y \in X} v^{\nu+1}(y; \lambda \circ r_\nu(x,u)) p(y|x,u) \\ &\qquad\qquad x \in X \quad \lambda \in \mathbf{R} \quad \nu = 1,2,\cdots,N \end{aligned}$$

例 1.1　最小型評価系の場合[1, §5] ($\circ := \wedge$, $a \wedge b := \min(a,b)$)：

$$\begin{aligned} &\mathrm{Max} E[\, r_1(x_1,u_1) \wedge \cdots \wedge r_N(x_N,u_N) \wedge r_G(x_{N+1})] \\ &\quad \text{s.t.}\ \ (\mathrm{i})_\mathrm{n},\ (\mathrm{ii})_\mathrm{n} \quad 1 \leq n \leq N \end{aligned}$$

2　ベクトル・マトリクス表記

定義 2.1　区間 $I \subset \mathbf{R}$ をとり、区間族 $\{I_i|\ i = 1,2,\ldots,m\}$ が次を満たすとする

$$\begin{aligned} &I_i \subset I\ (i = 1,2,\ldots,m)\ ,\ \bigcup_{i=1}^m I_i = I\ , \\ &I_i \cap I_j = \emptyset\ (i \neq j)\ , \\ &\sup(I_i) = \inf(I_{i+1})\ (i = 1,2,\ldots,m-1). \end{aligned}$$

このとき記号 $[\![\cdots]\!](\cdot)$ を次のように定義する

$$\left[\!\!\left[\frac{I_1 \ \cdots \ I_m}{f_1 \ \cdots \ f_m} \right]\!\!\right](x) := f_1(x)\,\chi_{I_1}(x) + \cdots + f_m(x)\,\chi_{I_m}(x)$$

$$\left[\!\!\left[\begin{array}{ccc} I_1 & \cdots & I_m \\ \hline f_{11} & \cdots & f_{1m} \\ \vdots & \ddots & \vdots \\ f_{n1} & \cdots & f_{nm} \end{array} \right]\!\!\right](x) := \begin{pmatrix} \left[\!\!\left[\dfrac{I_1 \ \cdots \ I_m}{f_{11} \ \cdots \ f_{1m}} \right]\!\!\right](x) \\ \vdots \\ \left[\!\!\left[\dfrac{I_1 \ \cdots \ I_m}{f_{n1} \ \cdots \ f_{nm}} \right]\!\!\right](x) \end{pmatrix}$$

ただし、χ は特性関数を表す。

例 2.1

$$\begin{aligned} &I = [0,1], \quad m = 3, \\ &f_1(x) = x,\ f_2(x) = 0.1x + 0.18,\ f_3(x) = 0.24 \\ &I_1 = [0, 0.2], \quad I_2 = (0.2, 0.6], \quad I_3 = (0.6, 1] \end{aligned}$$

このとき

$$\left[\!\!\left[\frac{I_1 \quad I_2 \quad I_3}{f_1 \quad f_2 \quad f_3} \right]\!\!\right](x)$$

$$= \left[\!\left[\frac{[0,0.2] \quad (0.2,0.6] \quad (0.6,1]}{x \quad 0.1x+0.18 \quad 0.24}\right]\!\right](x)$$
$$= x \times \chi_{[0,0.2]}(x) + (0.1x+0.18) \times \chi_{(0.2,0.4]}(x) + 0.24 \times \chi_{(0.4,1]}(x)$$

よって

$$\left[\!\left[\frac{I_1 \quad I_2 \quad I_3}{f_1 \quad f_2 \quad f_3}\right]\!\right](0.1) = 0.1,$$
$$\left[\!\left[\frac{I_1 \quad I_2 \quad I_3}{f_1 \quad f_2 \quad f_3}\right]\!\right](0.5) = 0.23$$

補題 2. 1

(i)

$$\left[\!\left[\frac{I_1 \ \cdots \ I_n}{f_1 \ \cdots \ f_n}\right]\!\right](x) + \left[\!\left[\frac{I_1 \ \cdots \ I_n}{g_1 \ \cdots \ g_n}\right]\!\right](x) = \left[\!\left[\frac{I_1 \quad \cdots \quad I_n}{f_1+g_1 \quad \cdots \quad f_n+g_n}\right]\!\right](x)$$

(ii)

$$r\left[\!\left[\frac{I_1 \ \cdots \ I_n}{f_1 \ \cdots \ f_n}\right]\!\right](x) = \left[\!\left[\frac{I_1 \quad \cdots \quad I_n}{rf_1 \quad \cdots \quad rf_n}\right]\!\right](x) \qquad r \in \mathbf{R}$$

(iii)

$$\begin{pmatrix} p_{11} & \cdots & p_{1m} \\ \vdots & \ddots & \vdots \\ p_{l1} & \cdots & p_{lm} \end{pmatrix} \left[\!\left[\begin{array}{ccc} I_1 & \cdots & I_n \\ \hline f_{11} & \cdots & f_{1n} \\ \vdots & \ddots & \vdots \\ f_{m1} & \cdots & f_{mn} \end{array}\right]\!\right](x) = \left[\!\left[\begin{array}{ccc} I_1 & \cdots & I_n \\ \hline \sum_{j=1}^m p_{1j}f_{j1} & \cdots & \sum_{j=1}^m p_{1j}f_{jn} \\ \vdots & \ddots & \vdots \\ \sum_{j=1}^m p_{lj}f_{j1} & \cdots & \sum_{j=1}^m p_{lj}f_{jn} \end{array}\right]\!\right](x)$$

補題 2. 2

任意の $a \in I$ に対し、整数 $n_0 \leq n$ 及び Definition 2. 1 の条件を満たす区間族 $\{I_1, I_2, \ldots, I_{n_0-1}, I'_{n_0}, I'_{n_0+1}\}$ が存在し次を満たす（c は定数）

$$\left[\!\left[\frac{I_1 \quad I_2 \quad \cdots \quad I_n}{f_1 \quad f_2 \quad \cdots \quad f_n}\right]\!\right](x \wedge a) = \left[\!\left[\frac{I_1 \quad I_2 \quad \cdots \quad I_{n_0-1} \quad I'_{n_0} \quad I'_{n_0+1}}{f_1 \quad f_2 \quad \cdots \quad f_{n_0-1} \quad f_{n_0} \quad c}\right]\!\right](x)$$

補題 2. 3

$$\max_j \left[\!\left[\frac{I_1 \ \cdots \ I_n}{f_1^j \ \cdots \ f_n^j}\right]\!\right](x) = \left[\!\left[\frac{I_1 \quad \cdots \quad I_n}{\max_j f_1^j \quad \cdots \quad \max_j f_n^j}\right]\!\right](x)$$

定理 2. 1 二項演算 $\circ$ は $\wedge$ とする。このとき、定理1. 1 の再帰式は次のように書き表される。

$$\begin{pmatrix} v^{N+1}(s_1;\lambda) \\ \vdots \\ v^{N+1}(s_m;\lambda) \end{pmatrix} = \left[\!\left[\begin{array}{ccc} I_1^{N+1} & \cdots & I_n^{N+1} \\ \hline f_{11}^{N+1} & \cdots & f_{1n}^{N+1} \\ \vdots & \ddots & \vdots \\ f_{m1}^{N+1} & \cdots & f_{mn}^{N+1} \end{array}\right]\!\right](\lambda)$$

$$v^\nu(s_j;\lambda) = \max_{k=1,2,\ldots,l} \left(p_{1jk} \cdots p_{mjk}\right) \left[\!\left[\begin{array}{ccc} I_1^{\nu+1} & \cdots & I_{n_\nu}^{\nu+1} \\ \hline f_{11}^{\nu+1} & \cdots & f_{1n_\nu}^{\nu+1} \\ \vdots & \ddots & \vdots \\ f_{m1}^{\nu+1} & \cdots & f_{mn_\nu}^{\nu+1} \end{array}\right]\!\right](\lambda \circ r_{\nu jk})$$
$$j = 1,2,\ldots,m, \ \nu = 1,2,\ldots,N$$

よって、

$$\begin{pmatrix} v^{\nu}(s_1;\lambda) \\ \vdots \\ v^{\nu}(s_m;\lambda) \end{pmatrix} = \left[\!\left[\begin{array}{ccc} I_1^{\nu} & \cdots & I_{n^\nu}^{\nu} \\ \hline f_{11}^{\nu} & \cdots & f_{1n^\nu}^{\nu} \\ \vdots & \ddots & \vdots \\ f_{m1}^{\nu} & \cdots & f_{mn^\nu}^{\nu} \end{array}\right]\!\right](\lambda)$$

(ただし、$p_{ijk} := p(s_i|s_j, a_k), r_{\nu jk} := r_\nu(s_j, a_k)$)

3 Example

講演当日、次の例題に関し実際の計算の流れを示す。

【問題】

$\mathrm{Max}E[r_1(u_1) \wedge r_2(u_2) \wedge r_G(x_3)]$

s.t. (i) $x_{n+1} \sim p(\cdot|x_n, u_n) \quad n = 1,2$

(ii) $u_1 \in U, \ u_2 \in U$

【利得】

$r_G(s_1) = 0.3 \quad r_G(s_2) = 1.0 \quad r_G(s_3) = 0.8$

$r_2(a_1) = 1.0 \quad r_2(a_2) = 0.6$

$r_1(a_1) = 0.7 \quad r_1(a_2) = 1.0$

【推移確率】

$u_t = a_1$

$x_t \backslash x_{t+1}$	s_1	s_2	s_3
s_1	0.8	0.1	0.1
s_2	0.0	0.1	0.9
s_3	0.8	0.1	0.1

$u_t = a_2$

$x_t \backslash x_{t+1}$	s_1	s_2	s_3
s_1	0.1	0.9	0.0
s_2	0.8	0.1	0.1
s_3	0.1	0.0	0.9

References

[1] Bellman, R.E. and Zadeh, L.A.: *Decision-making in a fuzzy environment*, *Management Science*, **17**, 1970, pp. B141-B164.

[2] Iwamoto, S.: *Associative dynamic programs*, *J. Math. Anal. Appl.*, **201**, 1996, pp. 195-211.