

# Research on Semi-Tensor Product Solution Method for Optimal Decision of a Class of Multi-Valued Control Networks

Fanqiang Fu<sup>1</sup>, Wei Wei<sup>1</sup>, Ying Zhou<sup>2</sup>, Liu Qian<sup>1</sup>

<sup>1</sup>School of Data Science and Information Engineering, Guizhou Minzu University, Guiyang Guizhou

<sup>2</sup>School of Mathematics and Statistics, Guizhou University, Guiyang Guizhou

Email: 1113200865@qq.com

Received: Oct. 1<sup>st</sup>, 2018; accepted: Oct. 17<sup>th</sup>, 2018; published: Oct. 24<sup>th</sup>, 2018

---

## Abstract

The finite dynamic game process can be transformed into the optimal control problem of multi-valued logic control networks given the strategy of some players. In this paper, we study the finite dynamic game problem of single input and output. The paper uses the method of matrix half tensor product to derive the algebraic expression of multi-valued logic dynamic system and the expression of the semi-tensor product of the income objective functional, and proves the equivalence between the new expression of this income maximization problem and the problem. Furthermore, the dynamic programming method and algorithm for solving the optimal control problem of semi-tensor product system are given. Finally, an algorithm is used to solve an example.

## Keywords

Semi-Tensor Product, Multi-Valued Logic Control Network, Optimal Control, Discrete Dynamic Programming

---

# 一类多值控制网络最优决策的半张量积求解法研究

符繁强<sup>1</sup>, 韦维<sup>1</sup>, 周莹<sup>2</sup>, 钱柳<sup>1</sup>

<sup>1</sup>贵州民族大学, 数据科学与信息工程学院, 贵州 贵阳

<sup>2</sup>贵州大学, 数学与统计学院, 贵州 贵阳

Email: 1113200865@qq.com

收稿日期: 2018年10月1日; 录用日期: 2018年10月17日; 发布日期: 2018年10月24日

## 摘要

有限动态博弈过程在部分玩家的策略给定情形下,可以转化为多值逻辑控制网络的最优控制问题。本文研究这一类单输入输出的有限动态博弈问题,论文利用矩阵半张量积的方法,推导出多值逻辑动态系统的代数表达式和收益目标泛函的半张量积的表达形式,并证明此收益最大化问题新的表达方式与原问题的等价性;进而,给出了求解半张量积系统最优控制问题的动态规划法和算法;最后,应用此算法求解一个实例。

## 关键词

矩阵半张量积, 多值逻辑控制网络, 最优控制, 离散型动态规划方法

Copyright © 2018 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

布尔网络是描述人工智能系统、神经网络系统、基因调控网络等的一种简单可行工具[1] [2] [3] [4],在 1969 年, Kauffman 首次提出用布尔网络来描述、模拟基因网络[5],但是由于布尔网络在描述一些实际问题时难免会有一些的局限性,例如:温度高低,天气阴晴,人的年龄老少等,因此,这时,多值逻辑网络和混合值逻辑网络应运而生。在一定程度上,多值逻辑网络是布尔网络的自然推广。同时逻辑网络在博弈论中也有应用。在一个网络演化博弈中,当局中人的对策是多个或者不同多个时,我们可以用多值逻辑网络来建模,详细结果参见[6] [7] [8]。2011 年程代展及其团队提出了矩阵半张量积的概念,将传统的普通矩阵乘法推广到一般的矩阵上,即,不再需要前一个矩阵的列数与后一个矩阵的行数相等,然而它却保持着矩阵的普通乘积的基本性质。这种工具有效地将逻辑表达融入到基本的数学运算框架中,在专著[9] [10] [11]中详细介绍了矩阵的半张量积的定义与性质,并讨论了其在非线性系统、布尔网络、模糊逻辑等领域中的应用。

多值逻辑网络比布尔网络的应用更为广泛,所以在研究多值逻辑网络时同样可应用矩阵的半张量积。如文献[12]讨论了如何从演化动态博弈导出多值逻辑系统的优化控制问题,并在多值逻辑系统下讨论平均收益和带贴因子的总收益两种性能指标下的最优控制的基本结论和算法,同时从逻辑动态系统的最优控制中导出演化博弈的纳什均衡。文献[13]在博弈理论框架下对无穷时域上逻辑控制网络的最优控制问题进行求解。讨论逻辑控制网络的代数形式下对周期的计算。文献[14]研究了人与机器之间重复无穷的囚徒困境模型,从逻辑控制网络的代数形式出发,讨论逻辑控制网络的拓扑结构,最后将这一结果推广到带有记忆的博弈中。文献[15]研究了兼具逻辑表达式和代数表达式的布尔网络的最优控制问题。经查阅相关文献,对有限博弈的最优收益问题研究比较少,因其系统是逻辑形式描述,而目标泛函须用代数式表达,将逻辑表达式和代数表达式在统一的框架下处理有一定的难度。本文利用矩阵半张量积的方法研究了多值逻辑网络的最优收益问题,将多值逻辑动态表达式转化为对应的半张量形式,在此基础上给出目标泛函的半张量形式,得到等价的半张量形式的最优控制问题,进而利用动态规划法求解半张量系统的最优控制问题,再根据等价公式得到系统的最优解。最后给出一个算例。

## 2. 准备知识

在这一节里，我们将本文的符号，涉及到的半张量积的定义和性质总结归纳如下。

### 2.1. 符号的说明

为叙述方便，做以下的符号说明：

- $\mathcal{M}_{m \times n}$  :  $m \times n$  维实矩阵集合；
- $D_k = \left\{ 0, \frac{1}{k-1}, \dots, \frac{k-2}{k-1}, 1 \right\}$ ；

注：在博弈中也用

$$D_k = \{1, 2, \dots, k\}$$

这时，将  $i \sim \frac{i-1}{k-1}, i=1, 2, \dots, k$ ，即可

- $\delta_n^k$  为单位阵  $I_n$  的第  $k$  列；
- $Col_i(A)$ ：为矩阵  $A$  的第  $i$  列，矩阵  $A$  的列集合记作  $Col(A)$ ；
- $\Delta_k = \left\{ \delta_k^i \mid i=1, 2, \dots, k \right\}$ ；这里  $\delta_k^i = Col_i(I_k)$ ；
- 一个矩阵  $L \in \mathcal{L}_{n \times m}$  称为逻辑矩阵，如果它的列满足  $Col(L) \subset \Delta_n$ ；
- 矩阵  $L = \left[ \delta_n^1, \delta_n^2, \dots, \delta_n^m \right] \in \mathcal{L}_{n \times m}$ ，记为  $v$ 。

### 2.2. 矩阵的半张量积的定义及性质

定义 2.2.1 [9]：设  $A \in \mathcal{M}_{m \times n}$  和  $B \in \mathcal{M}_{p \times q}$ 。记  $n$  与  $p$  的最小公倍数为

$$a = lcm\{n, p\}$$

定义  $A$  与  $B$  的半张量为

$$A \times B := (A \otimes I_{l/n})(B \otimes I_{l/p}),$$

这里  $\otimes$  是矩阵的 Kronecker 积(亦称张量积)。

注：当  $n=p$  时，矩阵  $A, B$  满足等维条件，定义化为普通矩阵乘法，因此，半张量积是普通乘法的推广。本文所说的半张量积均为左半张量积。

命题 2.2.1 [9]：若  $F, G, H \in \mathcal{M}_{m \times n}$ ，则矩阵半张量积满足

1) 结合律

$$(F \times G) \times H = F \times (G \times H).$$

2) 分配律

$$\begin{cases} F \times (aG \pm bH) = aF \times G \pm bF \times H, \\ (aF \pm bG) \times H = aF \times H \pm bG \times H, \end{cases} \quad a, b \in R$$

命题 2.2.2 [9]：设  $A \in \mathcal{M}_{m \times n}$  和  $B \in \mathcal{M}_{p \times q}$ ，关于半张量积的置换，有：

$$(A \times B)^T = B^T \times A^T.$$

为了进一步实现交换运算，为此定义换位矩阵。

定义 2.2.2 [9]：换位矩阵  $W_{[m,n]}$  是一个  $mn \times mn$  矩阵，定义如下：它的行和列都是由双指标  $(i, j)$  标注，列是按照索引  $Id(i, j; m, n)$  排列，行是由双指标  $(I, J)$  标注按照索引  $Id(J, I; n, m)$  排列，并且位于

$[(I, J), (i, j)]$  上的元素的值为

$$w_{(I, J)(i, j)} = \delta_{(i, j)}^{(I, J)} = \begin{cases} 1, & I = i \text{ 且 } J = j; \\ 0, & \text{其他.} \end{cases}$$

当  $n = m$  时, 我们简记  $W_{[n]} := W_{[n, n]}$ 。

命题 2.2.3 [9]: 设  $x \in \Delta_k$ 。则

$$x^2 = R_k^p x.$$

其中  $R_k^p := \text{diag}(\delta_k^1, \delta_k^2, \dots, \delta_k^k) = \delta_{k^2} [1, k+2, 2k+3, \dots, k^2]$  称为  $k$  维降价矩阵。

利用降价矩阵与换位矩阵, 我们可以降低计算逻辑函数的结构矩阵的计算量。

通过换位矩阵, 引入半张量积的重要性质——伪交换性。

命题 2.2.4 [9]: 给定矩阵  $A \in \mathcal{M}_{m \times n}$ ,

设  $Y \in R^l$  为行向量, 则

$$A \times Y = Y \times (I_l \otimes A).$$

设  $X \in R^l$  为列向量, 则

$$X \times A = (I_l \otimes A) \times X.$$

设两个行向量  $X \in R^n$ ,  $Y \in R^m$ , 则

$$X \times Y \times W_{[n, m]} = Y \times X.$$

设两个列向量  $X \in R^n$ ,  $Y \in R^m$ , 则

$$W_{[n, m]} \times X \times Y = Y \times X.$$

### 2.3. 博弈的定义

本文只关心有限博弈。

定义 2.3.1 [16]: 一个(有限)正规博弈, 有三个部分组成:  $G = (N, S, C)$ , 其中

1)  $N = \{1, 2, \dots, n\}$  表示这里有  $n$  玩家(局中人);

2)  $S = \prod_{i=1}^n S_i$  称为局势, 其中

$$S_i = \{S_1^i, \dots, S_k^i\}, \quad i = 1, \dots, n$$

是第  $i$  个玩家的策略集, 它表示第  $i$  个玩家有  $k_i$  个策略可供选择, 局势是所有玩家策略的笛卡尔积。

3)  $C = (c_1, \dots, c_n)$ , 其中  $c_i: S \rightarrow R$  是第  $i$  个玩家的收益函数为方便计, 策略集通常记

$$S_i = \{1, \dots, k_i\}, \quad i = 1, \dots, n.$$

这里“有限”博弈指:

a) 玩家数  $n < \infty$ ,

b) 策略数  $|S_i| < \infty, i = 1, \dots, n.$

### 3. 多值逻辑网络最大收益问题

本文我们考虑单输入输出有限博弈(参与者与行动集有限)的动态过程, 并假定部分玩家被视为“机器”, 其策略固定已知, 另一部分玩家作为“人”, 他的策略选择的准则是使其收益最大化。本文讨论所有玩家的策略都只有 1 步记忆的情形。

根据如上假设, 此类博弈问题可以转化为下面的多值逻辑控制网络的优化问题。

控制系统为:

$$x(k+1) = f(x(k), u(k)), \quad k = 1, 2, \dots \quad (1)$$

其中  $x(k)$  为状态表示  $k$  时刻“机器”采取的行动;  $u(k)$  为控制输入表示  $k$  时刻“人”采取的行动, “机器”的策略  $f$  为从  $D_k^2$  到  $D_k$  的多值逻辑函数, 即:

$$x(k) \in D_k, u(k) \in D_k, f: D_k^2 \rightarrow D_k$$

容许控制集为:

$$U_{ad} = \{u: N^+ \rightarrow D_k, u(k) \in D_k, k = 1, 2, \dots\}$$

最优控制问题(P): 假设初始时刻  $t_1$ 、初端  $x(t_1) = x_1$  和终端时刻  $t_N$  已定, 终端满足约束

$$J_N^*[x(N)] = 0$$

机器策略  $x(t)$  按(1)更新的情况下, 寻找人的最佳策略  $u^* \in U_{ad}$ , 使得系统状态从给定的初始状态  $x(1) = x$  到  $N$  时刻达到给定的终端状态  $x(N)$ , 并使得目标泛函最大化

$$J(u) = \sum_{t=1}^{N-1} \lambda^{t-1} c(x(t), u(t)) \rightarrow \max \quad (2)$$

其中  $c$  为每一步的收益函数,  $0 < \lambda < 1$  为折扣因子。

针对这种多值逻辑控制的最优控制问题不但包含状态系统的逻辑运算, 而且又包含了性能指标的代数运算, 如果仅仅只有单一的逻辑系统或一般最优控制方法是很难解决的, 为解决此类问题。我们引入矩阵半张量工具将逻辑动态网络转化为对应的代数状态方程。

## 4. 多值控制网络最大收益问题的半张量方法及其问题的求解

### 4.1. 多值控制网络最大收益问题的半张量表示

定理 4.1.1 [9]: 给定一个  $n$  元多值逻辑函数  $f(x_1, x_2, \dots, x_n)$ , 则存在一个唯一的逻辑矩阵  $M_f \in \mathcal{L}_{k \times k^n}$ , 使得  $f$  在半张量形式下可表示为

$$f(x_1, x_2, \dots, x_n) = M_f \bowtie x \quad (3)$$

这里,  $x = \times_{i=1}^n x_i$ ,  $M_f$  称为  $f$  的结构矩阵。称上式为多值逻辑函数  $f$  的代数形式。

根据定理以及矩阵半张量积的性质, 问题(P)可以等价表示为如下问题:

状态方程

$$\hat{x}(k+1) = L \bowtie \hat{x}(k) \bowtie \hat{u}(k)$$

其中  $\hat{x}(k) \in \Delta_k, \hat{u}(k) \in \Delta_k, L \in \mathcal{L}_{k \times k^2}$ .

初始状态  $\hat{x}(1) = \hat{x}_1$ , 终端  $\hat{x}(N) = \hat{x}_d$  容许控制集  $\hat{U}_{ad} = \{\hat{u}(k) \in \Delta_k, k = 1, 2, \dots\}$

收益函数所对应的半张量形式为:  $c(x(t), u(t)) = \hat{u}^T(t) \bowtie P \bowtie \hat{x}(t)$

目标泛函为

$$\hat{J}(\hat{u}) = \sum_{t=1}^{N-1} \lambda^{t-1} \hat{u}^T(t) \bowtie P \bowtie \hat{x}(t) \quad (4)$$

( $P$  为  $\hat{u}(t)$  对应的收益矩阵,  $P \in \mathcal{L}_{k \times k}$ ; 即: “人”相对于“机器”的收益值)。

最优控制问题  $\hat{P}$ : 寻找  $\hat{u}^* \in \hat{U}_{ad}$

$$st. \quad \hat{J}(\hat{u}^*) \geq \hat{J}(\hat{u}), \quad \forall \hat{u} \in \hat{U}_{ad}$$

## 4.2. 最优控制问题的求解

变分法、庞特里亚金最大值原理、动态规划法是求解最优控制问题经典的三种方法, 本文采用动态规划法求解离散系统的最优控制问题。针对半张量积表示的最优控制问题( $\hat{P}$ ), 有如下定理:

**定理 4.2.1:** 若  $\hat{P}_{1N}^* = \{\hat{u}^*(1), \hat{u}^*(2), \dots, \hat{u}^*(N-1)\}$  是最优策略, 则对于任意  $2 \leq k < N$ 。它的子策略  $\hat{P}_{kN}^* = \{\hat{u}^*(k), \hat{u}^*(k+1), \dots, \hat{u}^*(N-1)\}$  对于  $\hat{x}(1)$ ,  $\hat{P}_{1k-1}^*$  确定的  $\hat{x}^*(k)$  为起点的第  $k$  到  $N$  后部子过程而言, 也是最优策略。

证明: 以  $\max$  为例证明此定理, 反证法:

假设子策略  $\hat{P}_{kN}^* = \{\hat{u}^*(k), \hat{u}^*(k+1), \dots, \hat{u}^*(N-1)\}$  对于  $\hat{x}(1)$ ,  $\hat{P}_{1k-1}^*$  确定的  $\hat{x}^*(k)$  为起点的第  $k$  到  $N$  后部子过程而言, 不是最优策略。则必定存在  $2 \leq k < N$  和策略  $\tilde{P}_{kN} = \{\tilde{u}(k), \tilde{u}(k+1), \dots, \tilde{u}(N-1)\}$ , 使得

$$\hat{J}_k[\hat{x}^*(k), \tilde{P}_{kN}] > \hat{J}_k[\hat{x}^*(k), \hat{P}_{kN}^*]$$

令  $\tilde{P}_{1N} = \{\hat{u}^*(1), \hat{u}^*(2), \dots, \hat{u}^*(k-1), \tilde{u}(k), \tilde{u}(k+1), \dots, \tilde{u}(N-1)\}$ , 则

$$\begin{aligned} \hat{J}_{1N}[\hat{x}(1), \tilde{P}_{kN}] &= \hat{J}_{1k-1}[\hat{x}(1), \hat{P}_{1k-1}^*] + \hat{J}_k[\hat{x}^*(k), \tilde{P}_{kN}] \\ &> \hat{J}_{1k-1}[\hat{x}(1), \hat{P}_{1k-1}^*] + \hat{J}_k[\hat{x}^*(k), \hat{P}_{kN}^*] \\ &= \hat{J}_{1N}[\hat{x}(1), \hat{P}_{1N}^*] \end{aligned}$$

这与  $\hat{P}_{1N}^* = \{\hat{u}^*(1), \hat{u}^*(2), \dots, \hat{u}^*(N-1)\}$  是最优策略矛盾。证毕。

下面我们将定理 4.2.1 的思想, 转化为可计算的递归算法。对于任意给定的  $t \in \{1, \dots, N-1\}$  和  $\hat{x} \in \Delta_k$ , 考虑如下的受控系统

$$\begin{cases} \hat{x}(k+1) = L \times \hat{x}(k) \times \hat{u}(k) \\ \hat{x}(t) = \hat{x} \end{cases} \quad k \in \{t, t+1, \dots, N-1\} \quad (5)$$

设其性能指标为  $\hat{J}(\hat{u}(\bullet), t) = \sum_{k=t}^{N-1} \lambda^{t-k} \hat{u}^T(k) \times P \times \hat{x}(k)$ 。

则最优控制问题  $\hat{P}_{kN}^*$  为: 对于给定  $t \in \{1, 2, \dots, N-1\}$ ,  $\hat{x}(t) \in \Delta_k$  对于系统 (5) 找到策略  $\{\hat{u}^*(t), \hat{u}^*(t+1), \dots, \hat{u}^*(N-1)\}$ , 使得  $\hat{x}^*(N) = \hat{x}_d$ , 并  $\hat{J}(\hat{u}(\bullet), t) = \max$ 。

当  $t=1$  且  $\hat{x}(1) = \hat{x}_1$  时, 问题  $\hat{P}_{kN}^*$  即为我们需求的最优控制问题  $\hat{P}$ , 定义如下值函数:

$$\begin{cases} V(t, \hat{x}) = \max_{\hat{u}^*(k), \hat{u}^*(k+1), \dots, \hat{u}^*(N-1)} J(\hat{u}(\bullet), t) \\ V(N, \hat{x}) = 0 \end{cases} \quad \forall t \in \{1, 2, \dots, N-1\}, \hat{x} \in \Delta_k \quad (6)$$

则有如下定理:

**定理 4.2.2:** 对  $\forall t \in \{1, 2, \dots, N-1\}$ ,  $\hat{x} \in \Delta_k$ , 有

$$\begin{cases} V(t, \hat{x}) = \max_{\hat{u}(t)} \left\{ \lambda^{t-1} \hat{u}^T(t) P \hat{x}(t) + V(t+1, \hat{x}((t+1), \hat{u}(\bullet))) \right\} \\ V(N, \hat{x}) = 0 \end{cases} \quad \forall t \in \{N-1, \dots, 2, 1\}, \hat{x} \in \Delta_k$$

证明: 由值函数  $V(t, \hat{x})$  的定义可知:

$$\begin{aligned}
V(t, \hat{x}) &\geq \hat{J}(\hat{u}(\cdot), t) = \sum_{t=1}^{N-1} \lambda^{t-1} \hat{u}^T(t) P \hat{x}(t) \\
&\geq \lambda^{t-1} \hat{u}^T(t) P \hat{x}(t) + \sum_{k=t+1}^{N-1} \lambda^{t-1} \hat{u}^T(t) P \hat{x}(t), \quad \forall \{\hat{u}(k), \hat{u}(k+1), \dots, \hat{u}(N-1)\}
\end{aligned}$$

两端关于序列  $\{\hat{u}(k), \hat{u}(k+1), \dots, \hat{u}(N-1)\}$  取最大泛函

$$\begin{aligned}
V(t, \hat{x}) &\geq \max_{\hat{u}(k), \dots, \hat{u}(s-1)} \left\{ \lambda^{t-1} \hat{u}^T(t) P \hat{x}(t) + \sum_{k=t+1}^{N-1} \lambda^{t-1} \hat{u}^T(t) P \hat{x}(t) \right\} \\
&\geq \max_{\hat{u}(t)} \left\{ \lambda^{t-1} \hat{u}^T(t) P \hat{x}(t) + V(t+1, \hat{x}((t+1), \hat{u}(\cdot))) \right\}
\end{aligned}$$

对任何序列  $\{\hat{u}(t), \hat{u}(t+1), \dots, \hat{u}(N-1)\}$ , 有

$$\begin{aligned}
\hat{J}(\hat{u}(\cdot), t) &= \sum_{k=t}^{N-1} \lambda^{t-1} \hat{u}^T(t) P \hat{x}(t) \\
&\geq \lambda^{t-1} \hat{u}^T(t) P \hat{x}(t) + \max_{\hat{u}(t+1), \dots, \hat{u}(s-1)} \sum_{k=t+1}^{N-1} \lambda^{t-1} \hat{u}^T(t) P \hat{x}(t) \\
&= \max_{\hat{u}(t)} \left\{ \lambda^{t-1} \hat{u}^T(t) P \hat{x}(t) + V(t+1, \hat{x}((t+1), \hat{u}(\cdot))) \right\}
\end{aligned}$$

证毕。

### 4.3. 动态规划算法

根据定理 4.2.2, 我们容易得到如下算法。

逆推算法 4.3.1:

Step 1: 利用终端条件  $\hat{x}(N) = \hat{x}_d$  开始由后向前逆推基本方程, 计算出最优收益函数  $\hat{J}^*(\hat{u}(\cdot), k), (k = N, \dots, 3, 2)$ , 最后通过计算  $\hat{J}_1^*[\hat{x}(1, \hat{u}(\cdot)), 1]$  从而得到最优决策序列

$$\{\hat{u}^*(1), \hat{u}^*(2, \hat{x}(2)), \dots, \hat{u}^*(N-1, \hat{x}(N-1))\}$$

Step 2: 根据状态方程  $\hat{x}(k+1) = L \times \hat{u}(k) \times \hat{x}(k)$  和初始条件  $\hat{x}(1) = \hat{x}$ , 从  $k=1$  开始由前向后确定  $\hat{x}^*(k)$ 。应用此算法通过反复的迭代和回代从而得到

最优轨线  $\{\hat{x}^*(k), k = 1, 2, \dots, N\}$ 。

最优策略  $\hat{p}_{1N}^* = \{\hat{u}^*(1), \hat{u}^*(2, \hat{x}^*(2)), \dots, \hat{u}^*(N-1, \hat{x}^*(N-1))\}$ 。

### 4.4. 应用实例

考虑有限的人机博弈, 双方均有三个可选择的策略  $\{1, 2, 3\}$ 。支付双矩阵见表 1。

这里机器采取的策略为: 在本次博弈中如果赢或平, 下次不改变策略, 如果本次博弈输, 则下次采用对手的策略。

Table 1. Payoff Bi-Matrix

表 1. 支付双矩阵

人 \ 机	机		
	1	2	3
1	3, 3	0, 4	9, 2
2	4, 0	4, 4	5, 3
3	2, 9	3, 5	6, 6

给定初始状态  $x(1)=1$ ，终端状态  $x(4)=3$ ；性能指标为

$$J(u) = \sum_{t=1}^4 \lambda^{t-1} c(x(t), u(t))$$

其中， $x(t) \in D_k$  为逻辑变量表示机器的策略， $u(t) \in D_k$  为控制变量表示人的策略， $c(x(t), u(t))$  为人的收益函数，求使性能指标最大的控制序列。

解：令  $1 \sim \delta_3^1$ ， $2 \sim \delta_3^2$ ， $3 \sim \delta_3^3$ ，则上述博弈可以写成

$$\hat{x}(k+1) = L \times \hat{x}(k) \times \hat{u}(k)$$

其中

$$L = \delta_3 [1, 1, 3, 1, 2, 2, 3, 2, 3]$$

$\hat{x}(t)$  为状态变量，表示机器  $t$  时刻的策略。 $\hat{x}(t) \in \Delta_3$ ；

$\hat{u}(t)$  为状态变量，表示人  $t$  时刻的策略。 $\hat{u}(t) \in \Delta_3$ 。

初始条件  $\hat{x}(1) = \delta_3^1$ ，终端  $\hat{x}(4) = \delta_3^3$  性能指标

$$\hat{J}(\hat{u}) = \sum_{t=1}^4 \lambda^{t-1} \hat{u}^T(t) \times P \times \hat{x}(t)$$

求使性能指标最大的控制序列  $\hat{u}^*(1), \hat{u}^*(2), \hat{u}^*(3), \hat{u}^*(4)$ 。

应用动态规划法求解

$$\hat{x}(4) = L \times \hat{x}(3) \times \hat{u}(3)$$

$$\text{有, } \hat{J}^*(\hat{u}(4)) = \max_{\hat{u}(4) \in \Delta_3} \lambda^3 \hat{u}^T(4) \times \begin{bmatrix} 3 & 0 & 9 \\ 4 & 4 & 5 \\ 2 & 3 & 6 \end{bmatrix} \times \hat{x}(4)。$$

当  $\hat{u}(4) = \delta_3^1, \hat{x}(4) = \delta_3^3$  时， $\hat{J}^*(\hat{u}(4))$  取最大值为，其值为： $\lambda^3 9$

$$\hat{x}(4) = L \times \hat{x}(3) \times \hat{u}(3) = \delta_3^3$$

$$\hat{J}^*(\hat{u}(3)) = \max_{\hat{u}(3) \in \Delta_3} \lambda^2 \hat{u}^T(3) \times \begin{bmatrix} 3 & 0 & 9 \\ 4 & 4 & 5 \\ 2 & 3 & 6 \end{bmatrix} \times \hat{x}(3)$$

有，

$$= \max_{\hat{u}(3) \in \Delta_3} \begin{cases} \lambda^2 6 + \hat{J}^*(\hat{u}(4)) & \hat{u}(3) = \delta_3^2, \hat{x}(3) = \delta_3^3 \\ \lambda^2 9 + \hat{J}^*(\hat{u}(4)) & \hat{u}(3) = \delta_3^1, \hat{x}(3) = \delta_3^3 \\ \lambda^2 3 + \hat{J}^*(\hat{u}(4)) & \hat{u}(3) = \delta_3^3, \hat{x}(3) = \delta_3^2 \end{cases}$$

则当取  $\hat{u}(3) = \delta_3^1, \hat{x}(3) = \delta_3^3$  时， $\hat{J}^*(\hat{u}(3))$  取得最大值，其值为： $\lambda^2 9 + \lambda^3 9$

$$\hat{x}(3) = L \times \hat{x}(2) \times \hat{u}(2) = \delta_3^3$$

$$\text{有, } \hat{J}^*(\hat{u}(1)) = \max_{\hat{u}(1) \in \Delta_3} \lambda^0 \hat{u}^T(1) \times \begin{bmatrix} 3 & 0 & 9 \\ 4 & 4 & 5 \\ 2 & 3 & 6 \end{bmatrix} \times \hat{x}(1)$$

$$= \max_{\hat{u}(1) \in \Delta_3} \begin{cases} 2 + \hat{J}^*(\hat{u}(2)) & \hat{u}(1) = \delta_3^3, \hat{x}(1) = \delta_3^1 \end{cases}$$

则当取  $\hat{u}(1) = \delta_3^3, \hat{x}(1) = \delta_3^1$  时， $\hat{J}^*(\hat{u}(1))$  取得最大值，其值为： $2 + \lambda 9 + \lambda^2 9 + \lambda^3 9$ 。



因此所求的最优控制序列为:  $\hat{u}^*(1) = \delta_3^3, \hat{u}^*(2) = \delta_3^1, \hat{u}^*(3) = \delta_3^1, \hat{u}^*(4) = \delta_3^1$ , 最优轨线为:  
 $\hat{x}(1) = \delta_3^1, \hat{x}(2) = \delta_3^3, \hat{x}(3) = \delta_3^3, \hat{x}(4) = \delta_3^3$ , 性能指标为  $J^*(\hat{u}(1)) = 2 + \lambda 9 + \lambda^2 9 + \lambda^3 9$ 。

将向量转化为标量则最优控制序列为:  $u(1) = 3, u(2) = 1, u(3) = 1, u(4) = 1$ , 最优轨线为:

$x(1) = 1, x(2) = 2, x(3) = 3, x(4) = 3$ , 性能指标  $J^*(u(1))$  为:  $2 + \lambda 9 + \lambda^2 9 + \lambda^3 9$ 。则人的最大收益为:  
 $2 + \lambda 9 + \lambda^2 9 + \lambda^3 9$ 。

注 4.4.1: 本文只研究了单输入输出的决策最优控制问题, 本文所应用的定理和算法对于多输入多输出系统的决策最优控制问题仍然适用。

## 5. 总结

本文研究了有限动态博弈过程中多值逻辑控制网络的收益优化问题, 利用矩阵半张量积的方法将逻辑系统转化为对应的代数系统, 并给出了半张量形式下目标泛函的表达式。在引入值函数的情况下对其最优性原理进行了证明并给出了算法, 同时举例说明算法的有效性。

## 基金项目

国家自然科学基金(No. 11761021), 贵州民族大学科研基金资助项目(No. 2017ZD017)。

## 参考文献

- [1] Kauffman, S.A. (1992) The Origins of Order: Self-Organization and Selection in Evolution. *Journal of Evolutionary Biology*, **13**, 133-144.
- [2] Kurten, K.E. (1999) Correspondence between Neural Threshold Networks and Kauffman Boolean Cellular Automata. *Journal of Physics A General Physics*, **21**, L615.
- [3] Davidson, E.H., Rast, J.P., Oliveri, P., et al. (2002) A Genomic Regulatory Network for Development. *Science*, **295**, 1669. <https://doi.org/10.1126/science.1069883>
- [4] Harris, S.E., Sawhill, B.K., Wuensche, A., et al. (2002) A Model of Transcriptional Regulatory Networks Based on Biases in the Observed Regulation Rules. *Complexity*, **7**, 23-40. <https://doi.org/10.1002/cplx.10022>
- [5] Kauffman, S.A. (1969) Metabolic Stability and Epigenesis in Randomly Constructed Genetic Nets. *Journal of Theoretical Biology*, **22**, 437-467. [https://doi.org/10.1016/0022-5193\(69\)90015-0](https://doi.org/10.1016/0022-5193(69)90015-0)
- [6] Smith, J.M. and Price, G.R. (1973) The Logic of Animal Conflict. *Nature*, **246**, 15-18. <https://doi.org/10.1038/246015a0>
- [7] Smith, J.M. (1982) Evolution and the Theory of Games. Evolution and the Theory of Games. Cambridge University Press, Cambridge, 41-45. <https://doi.org/10.1017/CBO9780511806292>
- [8] Nowak, M.A. and May, R.M. (1992) Evolutionary Games and Spatial Chaos. *Nature*, **359**, 826-829. <https://doi.org/10.1038/359826a0>
- [9] 程代展, 齐洪胜. 矩阵的半张量积理论与应用[M]. 北京: 科学出版社, 2007.
- [10] Cheng, D., Qi, H. and Zhao, Y. (2012) An Introduction to Semi-Tensor Product of Matrices and Its Applications. World Scientific, Singapore. <https://doi.org/10.1142/8323>
- [11] Cheng, D., Qi, H. and Li, Z. (2010) Analysis and Control of Boolean Networks. In: Communications & Control Engineering, Vol. 13, Springer, Berlin.
- [12] 程代展, 赵寅, 徐听听. 演化博弈与逻辑动态系统的优化控制[J]. 系统科学与数学, 2012, 32(10): 1226-1238.
- [13] Zhao, Y., Li, Z. and Cheng, D. (2011) Optimal Control of Logical Control Networks. *IEEE Transactions on Automatic Control*, **56**, 1766-1776. <https://doi.org/10.1109/TAC.2010.2092290>
- [14] Mu, Y. and Guo, L. (2009) Optimization and Identification in a Non-Equilibrium Dynamic Game. *Proceedings of the 48th IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference, Shanghai, 15-18 December 2009*, 5750-5755. <https://doi.org/10.1109/CDC.2009.5400029>
- [15] 钱柳, 韦维, 符繁强, 周荧. 布尔控制网络最小能耗问题的半张量方法[J]. 应用数学进展, 2018, 7(1): 95-103.
- [16] 程代展, 夏元清, 马宏宾, 等. 矩阵代数、控制与博弈[M]. 北京: 北京理工大学出版社, 2016.

**知网检索的两种方式:**

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>  
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2324-7991, 即可查询
2. 打开知网首页 <http://cnki.net/>  
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: [aam@hanspub.org](mailto:aam@hanspub.org)