有限博弈的矩阵半张量积方法

程代展[1†], 刘泽群[1,2]

(1. 中国科学院 数学与系统科学研究院 系统控制重点实验室, 北京 100190;

2. 中国科学院大学 数学科学学院, 北京 100049)

摘要: 矩阵半张量积被广泛地应用在有限博弈的研究中, 例如: 1) 演化博弈; 2) 势博弈; 3) 有限博弈的向量空间分解; 4) 基于势博弈的优化与控制; 5) 合作博弈等. 本文的目的, 就是对上述各种应用做一个全面的介绍, 包括其原理、主要成果、以及尚待解决的问题.

关键词: 矩阵半张量积; 有限博弈; 演化博弈; 势博弈; Shapley值

引用格式: 程代展, 刘泽群. 有限博弈的矩阵半张量积方法. 控制理论与应用, 2019, 36(11): 1812 – 1819

DOI: 10.7641/CTA.2019.90595

Application of semi-tensor product of matrices to finite games

CHENG Dai-zhan[1†], LIU Ze-qun[1,2]

(1. Key Laboratory of Systems and Control, Academy of Mathematics and Systems Science,
Chinese Academy of Science, Beijing 100190, China;
2. School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract: Semi-tensor product of matrices has various applications for finite games, including 1) evolutionary game; 2) potential game; 3) vector space structure and decomposition of finite games; 4) potential-based optimization and control; 5) cooperative game, etc. The purpose of this paper is to provide a comprehensive introduction for the applications of semi-tensor product to finite games, including its principle, main results, and some faced challenging problems.

Key words: semi-tensor product of matrices; finite game; evolutionary game; potential game; Shapley value

Citation: CHENG Daizhan, LIU Zequn. Application of semi-tensor product of matrices to finite games. *Control Theory & Applications*, 2019, 36(11): 1812 – 1819

1 引言

博弈论也称对策论, 是研究主体行为发生直接相互作用时的决策以及这种决策的均衡问题的理论. 博弈问题历史久远, 在我国历史上有著名的田忌赛马、华容道等, 国外有较晚出现的双寡头垄断竞争的古诺模型等, 都是博弈论的经典案例. 但一般以冯·诺依曼等在1944年发表的名著《博弈论与经济行为》[1]为现代博弈论的起点, 该书首次对博弈问题引进了系统的数学方法, 从而使博弈论真正成为一个应用数学分支, 特别指出, 针对合作博弈的分配规则, Shapley提出了"Shapley值"这一概念, 被应用到很多实际问题中. 对于非合作博弈, Nash提出了"纳什均衡"这一核心概念, 标志着博弈新时代的开始. 由于博弈论在经济、军事等许多领域得到卓有成效的应用, 在其诞生后的半个多世纪中, 博弈论作为一个新的学科分支快速发展. 本文所涉及的博弈概念可见文[2], 这是一本很好的入门书, 有中译本. 另外, 文[3] 选编了大量博弈论的实例.

博弈论和控制论关系密切, 它们都是通过操控某个对象, 使之达到自己预期的目标. 只不过控制论的操控对象是机械, 而博弈论的操控对象是有反操控能力的智能个体. 因此, 有学者认为"广义而言, 控制论可视为博弈论的一个特例[4]". 又如, 文[5]以博弈论为统一框架讨论线性与非线性H_∞理论. 笔者以为, 控制论与博弈论确实有许多共通之点, 但在方法上各有所长, 有许多可以互相借鉴的地方. 因此, 它们的交叉领域是一个充满希望的新方向. 郭雷院士在对控制论现代发展方向的思考中曾说到[6]: "当面对真正的智能机器和智能网络等智能对象时, 现在的控制理论就不能套用了, 因为被控者与控制者往往存在博弈关

收稿日期: 2019−07−21; 录用日期: 2019−09−23.
†通信作者. E-mail: dcheng@iss.ac.cn.
本文责任编委: 梅生伟.
国家自然科学基金项目(61773371, 61733018)资助.
Supported by the National Natural Science Foundation of China (61773371, 61733018).
仅以此文献给导师关肇直先生百年诞辰!

系$\cdots\cdots$我认为, 把博弈因素恰当地放到控制理论框架中, 是一个非常重要的研究问题, 也是社会经济等领域问题中不可回避的, 将会大大拓展控制理论的研究与应用范围.” 现实的发展印证了郭雷的上述论断. 近年, 博弈控制论作为控制中一个新的分支得到迅速发展[7], 博弈论在控制问题中得到越来越多的应用[4,8–11]. 文[12]对矩阵半张量积从概念到新近的其对博弈论的应用做了较系统的概括.

矩阵半张量积将普通矩阵乘法推广到任意两个矩阵, 即, 设$A \in \mathcal{M}_{m\times n}$, $B \in \mathcal{M}_{p\times q}$, 记$t = \mathrm{lcm}(n, p)$为$n$与$p$的最小公倍数, 则其半张量积定义为

$$A \ltimes B := (A \otimes I_{t/n})(B \otimes I_{t/p}).$$

它是由中国学者提出, 并主要由中国学者发展起来的一个新的数学分支. 目前, 它已在多个领域, 特别是在逻辑系统的控制中得到了广泛应用. 本文遇到的矩阵乘法均默认为矩阵半张量积, 并且将乘法符号($\ltimes$)略去. 本文假定读者已熟悉矩阵半张量积的基本概念, 对矩阵半张量积不熟悉的读者可参考一些国内、外学者关于矩阵半张量积的综述文章[13–15]. 更详细的内容可参见文[16–18].

矩阵半张量积在博弈论中的应用是矩阵半张量积最成功的应用之一. 其基本原理是: 对于有限博弈, 每个玩家的策略是有限的. 因此, 它变为有限集上的函数(支付函数)及演化(演化博弈). 于是它本质上与多值逻辑类似[17–18]. 本文的目的, 就是要对矩阵半张量积在博弈论中的应用从原理、方法到进展做一个全面的介绍.

2　从静态博弈到演化博弈

定义 1　一个有限博弈$G = (N, S, C)$, 这里:

i) $N = \{1, 2, \cdots, n\}$是玩家集合, 表示这个博弈有n个玩家.

ii) $S = \prod_{i=1}^{n} S_i$称为局势, 这里, $S_i = \{1, 2, \cdots, k_i\}$表示第$i$个玩家有$k_i$个策略.

iii) $C = (c_1, c_2, \cdots, c_n)$是$n$个支付函数, 其中$c_i : S \to \mathbb{R}$是第$i$个玩家的支付函数.

矩阵半张量方法的核心技巧是将每个玩家的策略集合$S_i = \{1, 2, \cdots, k_i\}$转化为向量形式:

$$S_i \sim \varDelta_{k_i} = \{\delta_{k_i}^j \mid j = 1, \cdots, k_i\},\ i = 1, \cdots, n.$$

于是, 第i个玩家的策略$x_i \in \varDelta_{k_i}$, 记$k := \prod_{i=1}^{n} k_i$, 那么, 支付函数就可以用矩阵半张量积表示为

$$c_i(x_1, \cdots, x_n) = V_c^i \ltimes_{j=1}^{n} x_j,\ i = 1, \cdots, n. \tag{1}$$

这里$V_c^i \in \mathbb{R}^k$为一行向量, 称为c_i的结构向量.

如果一个博弈不断地重复进行, 那么, 每个玩家都可以根据以往的历史来更新自己的策略. 于是, 就可以得到策略演化方程. 一种策略更新规则是只依赖于上一次信息的, 即马尔科夫型的. 它可以表达为

$$\begin{cases} x_1(t+1) = f_1(x_1(t), \cdots, x_n(t)), \\ x_2(t+1) = f_2(x_1(t), \cdots, x_n(t)), \\ \qquad\vdots \\ x_n(t+1) = f_n(x_1(t), \cdots, x_n(t)). \end{cases} \tag{2}$$

这里$x_i(t), i = 1, \cdots, n$为第i个玩家在时刻t的策略. 利用向量表示方法并应用矩阵半张量积, 就可以将上述逻辑型映射(2)转换为代数形式:

$$x(t+1) = Lx(t), \tag{3}$$

这里: $x(t) = \ltimes_{j=1}^{n} x_j(t)$, $L \in \mathcal{L}_{k\times k}$是一个逻辑矩阵.

如果网络中某些结点为控制, $u_s \in \varDelta_{h_s}$, h_s表示第s个控制的策略个数, $s = 1, \cdots, m$, 即这些结点可不按策略演化规则而自由选择策略, 则策略演化方程变为控制逻辑网络:

$$x(t+1) = Lu(t)x(t), \tag{4}$$

这里: $u(t) = \ltimes_{s=1}^{m} u_s(t)$, $L \in \mathcal{L}_{k\times kh}(h = \prod_{s=1}^{m} h_s)$ 是一个逻辑矩阵.

文[19]首先使用矩阵半张量积方法对短视最优响应(myopic best response arrangement)规则更新策略的网络演化博弈建模并讨论了策略优化. 文[20]讨论了一般更新规则下的网络演化博弈的建模与控制, 并给出策略收敛的充要条件. 文[21]讨论了网络演化博弈的演化稳定策略. 文[22]讨论了演化博弈的稳定性与镇定设计等. 文[23]探讨了演化博弈的优化问题, 文[24]给出了网络演化博弈实现策略一致所需的最小控制玩家数的求解算法. 文[25]分析了网络演化博弈的策略同步. 文[26]讨论了变拓扑的网络演化博弈的建模与演化. 文[27]将代数状态空间理论引入到竞争扩散博弈的建模与分析中.

非马尔科夫型的演化博弈也是常见的, 特别是学习博弈经常导致这种演化形式. 文[28]在半张量积框架下对这种模型进行了探讨.

混合策略的使用在演化博弈中几乎是不可避免的, 这就导致了随机演化博弈. 文[29–32]等论文以矩阵半张量积为工具, 系统研究了随机演化布尔博弈的稳定性、镇定控制、输出调节、最优控制等问题.

3　势博弈

势博弈是用博弈论方法解决网络化系统优化与控制的关键, 在优化意义上可以说, 势函数对演化博弈系统的作用与李雅普诺夫函数对动力学系统的作用类似.

定义 2　考察一个有限博弈$G = (N, S, C)$, 如果存在一个函数$P : S \to \mathbb{R}$, 使得对任意$i \in \mathbb{N}$均成

立:

$$c_i(x_i,s_{-i})-c_i(y_i,s_{-i})=P(x_i,s_{-i})-P(y_i,s_{-i}),\\ x_i,y_i\in S_i,\ s_{-i}\in S_{-i}, \tag{5}$$

其中$S_{-i}=\prod\limits_{j\neq i}S_j$表示其他人的策略组合, 则称$G$为势博弈, P为势函数.

势函数有许多优良性质:

定理 1[33]

i) 设G为势博弈, 则势函数在除一个常数外唯一. 即如果P_1和P_2为G的两个势函数, 则$P_1-P_2=c_0\in\mathbb{R}$.

ii) 每一个有限势博弈都存在纯纳什均衡点.

iii) 串联短视最优等策略更新规则可使演化势博弈收敛到纯纳什均衡点.

势博弈虽然重要, 但检验它却是一个长期未解决的问题. 文[34]指出: "检验一个博弈是否为一个势博弈不是一件容易的事(It is not easy, however, to verify whether a given game is a potential game)." 利用矩阵半张量积方法, 本文彻底解决了这个问题[35]. 构造势方程

$$E\xi=b, \tag{6}$$

这里:

$$E=\begin{bmatrix}-E_1 & E_2 & 0 & \cdots & 0\\ -E_1 & 0 & E_3 & \cdots & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ -E_1 & 0 & 0 & \cdots & E_n\end{bmatrix},\ \xi=\begin{bmatrix}\xi_1\\ \xi_2\\ \vdots\\ \xi_n\end{bmatrix},\ b=\begin{bmatrix}b_2\\ b_3\\ \vdots\\ b_n\end{bmatrix},$$

其中:

$$E_i=I_{k^{i-1}}\otimes\mathbf{1}_k\otimes I_{k^{n-i}}\in\mathcal{M}_{k^n\times k^{n-1}},\\ i=1,\cdots,n,\\ \xi_i:=(V_i^{\rm d})^{\rm T}\in\mathbb{R}^{k^{n-1}},\ i=1,\cdots,n,\\ b_i:=(V_i^{\rm c}-V_1^{\rm c})^{\rm T}\in\mathbb{R}^{k^n},\ i=2,\cdots,n,$$

则有以下结论:

定理 2 一个有限博弈G是势博弈, 当且仅当, 其势方程(6)有解. 并且, 如果解存在, 则其势函数的结构向量为

$$V_{\rm P}=V_1^{\rm c}-V_1^{\rm d}E_1^{\rm T}=V_1^{\rm c}-\xi_1^{\rm T}(\mathbf{1}_k^{\rm T}\otimes I_{k^{n-1}}). \tag{7}$$

利用势方程, 文[36]设计了有效的数值算法. 文[37]讨论了策略集受限情况下的势博弈, 给出了相应的势方程. 文[38]将这种方法推广到连续策略的情况.

关于势博弈还有许多需要继续深入讨论的问题. 例如, 关于分群势博弈的问题[39], 基于超图的网络演化势博弈[40]等. 文[41]给出一个工程应用的设计.

4 有限博弈的向量空间结构

记$\mathcal{G}_{[n;k_1,\cdots,k_n]}$为一类有限博弈, 其玩家个数为$|N|=n$, 每个玩家的策略数为$|S_i|=k_i, i=1,\cdots,n$. 那么, 这些博弈的区别仅在于它们不同的支付函数. 注意到

$$c_i(x_1,\cdots,x_n)=V_i^{\rm c}\ltimes_{j=1}^{n}x_j,\ i=1,\cdots,n,$$

这里$V_i^{\rm c}\in\mathbb{R}^k(k=\prod\limits_{i=1}^{n}k_i)$是$c_i$的结构向量, 它完全决定了支付函数. 令

$$V_G:=[V_1^{\rm c}\ V_2^{\rm c}\ \cdots\ V_n^{\rm c}]\in\mathbb{R}^{nk},$$

则每个V_G一一对应于$G\in\mathcal{G}_{[n;k_1,\cdots,k_n]}$. 因此, $\mathcal{G}_{[n;k_1,\cdots,k_n]}$有一个天然的向量空间结构, 即

$$\mathcal{G}_{[n;k_1,\cdots,k_n]}\sim\mathbb{R}^{nk}.$$

下面讨论它的一些重要子空间. 首先, 不难看出, 势博弈集合是其子空间. 以下介绍它的一些其他重要子空间.

定义 3 设$G\in\mathcal{G}_{[n;k_1,\cdots,k_n]}$. G称为一个非策略博弈, 如果

$$c_i(x_i,s_{-i})=c_i(y_i,s_{-i}),\ \forall x_i,y_i\in S_i,\\ i=1,\cdots,n. \tag{8}$$

不难看出, 一个非策略博弈是一个势博弈, 其势函数为任意常数.

定义 4 设$G\in\mathcal{G}_{[n;k_1,\cdots,k_n]}$.

1) G是一个调和博弈, 如果

$$\sum_{i=1}^{n}(c_i(s)-\frac{1}{k_i}\sum_{x_i\in S_i}c_i(x_i,s_{-i}))=0. \tag{9}$$

2) G是一个纯调和博弈, 如果

$$\sum_{i=1}^{n}c_i(s)=0, s\in S, \tag{10}$$

且

$$\sum_{x\in S_i}c_i(x,y)=0,\ \forall y\in S_{-i}:=\prod_{j\neq i}S_j,\\ i=1,\cdots,n. \tag{11}$$

石头–剪刀–布就是一个纯调和博弈.

下面这个分解定理在有限博弈理论中是很重要的.

定理 3[42]

$$\mathcal{G}_{[n;k_1,\cdots,k_n]}=\underbrace{\mathcal{P}\oplus\overbrace{\mathcal{N}}^{\text{Harmonic games}}}_{\text{Potential games}}\oplus\mathcal{H}. \tag{12}$$

这里$\mathcal{P},\mathcal{N},\mathcal{H}$分别为: 纯势博弈、非策略博弈和纯调和博弈.

文[42]利用代数拓扑及图论中的分解定理才得到式(12)中的正交分解, 并且其内积是加权的. 本文利用

矩阵半张量积也得到同样的正交分解, 并且内积是普通欧氏空间的内积. 本文还给出了各子空间的基底与维数[43]. 加权后的势博弈和调和博弈有类似的向量空间结构, 但在应用上有更大自由度[44].

对称博弈的空间结构是一个有趣的问题[45]. 利用矩阵半张量积[46]讨论了对称博弈、反对称博弈, 给出有限博弈关于对称博弈、反对称博弈, 及非对称博弈的正交分解式(13), 并给出对称博弈和反对称博弈的许多性质. 文[47]给出了对称博弈及其正交补空间的基底. 文[48]在布尔博弈框架下, 深入讨论了对称博弈及势博弈的关系.

$$\mathcal{G}_{[n;k_1,\cdots,k_n]}=\mathcal{S}\oplus\mathcal{A}\oplus\mathcal{K}, \tag{13}$$

这里$\mathcal{S}, \mathcal{A}, \mathcal{K}$分别为对称博弈、非对称博弈和反对称博弈.

5 博弈控制论

文[7]用一个顺序结构(图1)当作博弈控制理论的示意图.

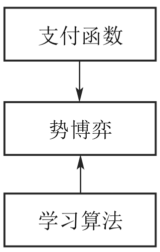

图 1 博弈控制理论的示意图

Fig. 1 Sketch of game control theory

从图中可以看出, 博弈控制论的中心思想是: 对于给定优化目标的系统:

第1步 设计每个个体的支付函数, 使系统变为以整体目标函数为势函数的势博弈.

第2步 设计每个个体的控制策略(或者优化算法), 使系统收敛于势函数的纳什均衡点.

注意: 1) 对于多自主体系统, 控制算法应基于局部信息, 方可实现. 2) 对多个纳什均衡的情况, 收敛到纳什均衡只是局部最优. 要达到全局最优, 需要更精细的学习算法.

5.1 基于资源的拥塞博弈

拥塞博弈与势博弈是等价的, 最典型的拥塞博弈是道路交通系统[33]. 文[49]利用矩阵半张量积方法给出资源型系统可设计为拥塞博弈的充要条件.

定义 5 一个基于资源的系统记作$\Sigma=(M,N,(\mathcal{A}^i)_{i\in\mathbb{N}},P)$, 这里$M=\{1,2,\cdots,m\}$是可使用的资源, $N=\{1,2,\cdots,n\}$为用户, $\mathcal{A}^i\subset 2^M$为第i个用户的策略. $\mathcal{A}:=\prod_{i=1}^{n}\mathcal{A}^i$称为局势. $P:\mathcal{A}\to\mathbb{R}$是总花费. 资源优化的目的是找到最优策略$a^*\in\mathcal{A}$, 使总花费最小, 即

$$P(a^*)=\min_{a\in\mathcal{A}}P(a). \tag{14}$$

设$|\mathcal{A}|=\ell$, 记作$\mathcal{A}=\{A_1,A_2,\cdots,A_\ell\}$. 构造

$$B\varXi^{\rm T}=P, \tag{15}$$

这里: $\varXi$是依赖于使用次数的资源费用, $B:\mathcal{A}\to\mathbb{R}^n$为玩家费用向量, P为总花费函数.

定理 4[49] 给定一个资源型系统$\varSigma=(M,N,(\mathcal{A}^i)_{i\in\mathbb{N}},P)$, 存在资源费用使系统成为以总花费$P(a)$为势函数的拥塞博弈, 当且仅当, 方程(15)有解.

5.2 具有用户特定代价的拥塞博弈

在一般拥塞博弈的基础上, 文献[50]考虑用户具有自己的特定代价的情形, 将用户的成本函数扩展成用户特定的代价和用户使用公共设备支付的和, 提出一种新的拥塞博弈. 结合文[33]的定理2.8, 证明了新提出的具有用户特定代价的拥塞博弈仍然是势博弈, 利用矩阵半张量积方法给出具有用户特定代价的拥塞博弈的势函数表达式, 并且运用此博弈对基于网络功能虚拟化(network function virtualization, NFV)的网络中虚拟服务功能链的资源配置建模, 有效解决其资源配置的问题.

定义 6[50] 一个具有用户特定代价的拥塞博弈, 记作$G=(M,N,(\mathcal{A}^i)_{i\in\mathbb{N}},(u_i)_{i\in\mathbb{N}})$, 这里$M=\{1,2,\cdots,m\}$是可使用的资源, $N=\{1,2,\cdots,n\}$为用户, $\mathcal{A}^i\subset 2^M$, 为第i个用户的策略. $\mathcal{A}:=\prod_{i=1}^{n}\mathcal{A}^i$称为局势. $u_i:\mathcal{A}\to\mathbb{R}$是第$i$个局中人的成本函数, 对于局势空间$\mathcal{A}$里的任一局势$a=(a_1,\cdots,a_n)\in\mathcal{A}$, 使用$n_l(a)$表示使用第$l$个设备的用户人数, 成本函数$u_i$表达式如下:

$$u_i(a)=r_i(a_i)+w_i\sum_{\{l\in a_i\}}c_l(n_l(a)),\ i=1,\cdots,n, \tag{16}$$

这里: $r_i:\mathcal{A}\mapsto\mathbb{R}$是第$i$个用户特定的代价函数, $c_l:\{0,1,2,\cdots,n\}\mapsto\mathbb{R}$是第$i$个用户使用第$l$个设备的支付函数, w_i是刻画第i个用户的支付受设备拥塞影响程度的权重参数.

定理 5[50] 一个用户有特定代价的拥塞博弈$G=(M,N,(\mathcal{A}^i)_{i\in\mathbb{N}},(u_i)_{i\in\mathbb{N}})$, 它一定是个加权势博弈且势函数由下式给出:

$$P(a)=\sum_{i\in\mathbb{N}}\frac{R_i}{w_i}a_i+\sum_{l=1}^{m}{\rm Row}_l C[\mathbf{1}^{\rm T}_{n_l(a)}\ltimes N(a)], \tag{17}$$

这里R_i是博弈G的第i个用户特定代价函数的结构向量, C是博弈G的设备支付矩阵,

$$\mathbf{1}_{n_l(a)}:=\underbrace{(1\ 1\ \cdots\ 1)}_{n_l(a)}{}^{\rm T},$$

$$N(a):=((\delta_n^1)^{\rm T}\ (\delta_n^2)^{\rm T}\ \cdots\ (\delta_n^{n_l(a)})^{\rm T})^{\rm T}.$$

5.3 控制设计

控制设计就是要设计一个依赖于(局部)信息的算法, 使系统收敛到最优点. 常见的算法, 如:

・带有惯性的联合策略虚拟行为 (joint strategy fictitious play with inertia)[51]

$$p_i^{a_i(t-1)}(t) = \epsilon,$$
$$p_i^{a_i^*(t)}(t) = 1 - \epsilon,$$

这里: $p_i^{a_i}$是玩家i取策略a_i的概率, $0 < \epsilon < 1$是惯性系数, $a_i^*(t) \in \arg\max\limits_{a_i \in A_i} u_i^{a_i}(t)$为玩家虚拟收益的最优响应, $u_i^{a_i}(t) = \dfrac{1}{t}\sum\limits_{\tau=0}^{t-1} c_i(x_i, x_{-i}(\tau))$.

・对数线性学习(log-linear learning)[52]

$$p_i^{a_i}(t) = \frac{\mathrm{e}^{\frac{1}{T}u_i(a_i, a_{-i}(t-1))}}{\sum\limits_{a_i' \in \mathcal{A}_i} \mathrm{e}^{\frac{1}{T}u_i(a_i', a_{-i}'(t-1))}},$$

这里$T \geqslant 0$称为温度系数.

最近, 文[53]给出基于状态的势博弈的一种两步学习算法, 并证明了在一定条件下这种算法在这种复杂演化博弈中可以收敛到纳什均衡.

5.4 基于状态的演化博弈

基于状态的演化博弈最先是Marden等在文[54]中提出的. 它同时考虑两个演化过程: 状态空间的演化过程和策略的演化过程, 是博弈控制系统的一个合理模型. 在这个框架下, 文[55]探讨了如下系统:

定义 7[55]

i) 一个基于有限状态的演化博弈可由$G = \{N, S, c, X, P\}$来刻画, 这里$N = \{1, 2, \cdots, n\}$是玩家集合; $S = \prod\limits_{i=1}^{n} S_i$ 是局势; $c = \{c_1, \cdots, c_n\}$是支付函数集, $c_i : X \times S \to \mathbb{R}$; X 是状态空间; $P : X \times S \to \Delta(X)$是依赖于局势的状态转移函数, 而$\Delta(X)$是有限状态空间$X$上的一个概率分布;

ii) $X = \{x_1, \cdots, x_r\}$, 这里$r = |X|$. 类似于$a_i \in \varDelta_{k_i}$, 也可以将状态表达为向量形式, 即

$$x_i \sim \delta_r^i,\ i = 1, \cdots, r.$$

iii) 状态$x(t)$满足

$$x(t+1) = P(x(t), a(t)), \tag{18}$$

用代数状态空间表示, 可得

$$x(t+1) = M_{\mathrm{P}} x(t) a(t), \tag{19}$$

这里: $a(t) = \ltimes_{i=1}^{n} a_i(t)$, $M_{\mathrm{P}} \in \varUpsilon_{r \times rk}$ $(k = \prod\limits_{i=1}^{n} k_i)$ 是 P 的结构矩阵.

iv) 演化博弈的策略演化方程为

$$\begin{cases} a_1(t+1) = f_1(x(t+1), a(t), c(t)), \\ a_2(t+1) = f_2(x(t+1), a(t), c(t)), \\ \qquad\vdots \\ a_n(t+1) = f_n(x(t+1), a(t), c(t)), \end{cases} \tag{20}$$

这里$c(t) = (c_1(t), c_2(t), \cdots, c_n(t))$. 策略的演化决定于策略更新规则, 本文假定$f_i$与$c$无关. 因此, 它可以用代数状态空间方法来描述策略演化过程, 得到状态方程如下:

$$a(t+1) = M_{\mathrm{F}} x(t+1) a(t), \tag{21}$$

这里 $M_{\mathrm{F}} \in \varUpsilon_{k \times rk}$ 是 $F = (f_1, f_2, \cdots, f_n)$ 的结构矩阵.

文[55]给出了上述系统是依赖状态的演化博弈的充分条件, 并给出优化控制设计方法.

5.5 考虑有限时域收益的线性动态博弈

考虑如下的具有n–决策者(decision makers)的线性系统:

$$\begin{cases} x(k+1) = Ax(k) + \sum\limits_{i=1}^{n} B_i u_i(k), \\ x(0) = x_0, \end{cases} \tag{22}$$

其中: $x \in \mathbb{R}^{m_0}$, $u_i \in \mathbb{R}^{m_i}, i = 1, \cdots, n$分别为系统状态以及第$i$个决策者的控制输入, 并且$A \in \mathbb{R}^{m_0 \times m_0}$, $B_i \in \mathbb{R}^{m_0 \times m_i}$, $i = 1, \cdots, n$. 假设$U_i \subset \mathbb{R}^{m_i}$为第$i \in N$个决策者的决策集, 并把联合决策集记为

$$\mathcal{U} = \prod_{i \in \mathbb{N}} U_i = U_1 \times U_2 \times \cdots \times U_n.$$

每一个决策者$i (i \in \mathbb{N})$都想最大化如下的有限时域收益函数:

$$G_i(x(0), u_1(\cdot), \cdots, u_n(\cdot)) = \sum_{k=0}^{K-1} g_i(x(k), u_1(k), \cdots, u_n(k)), \tag{23}$$

其中g_i: $\mathbb{R}^{m_0} \times U_1 \times U_2 \cdots \times U_n \to \mathbb{R}$ $(i = 1, \cdots, n)$为第i个决策者的单步收益函数.

定理 6[56] 假设单步收益函数g_i: $\mathbb{R}^{m_0} \times U_1 \times U_2 \cdots \times U_n \to \mathbb{R}$ $(i = 1, \cdots, n)$线性的, 即

$$g_i(x, u_1, \cdots, u_n) = c_i^{\mathrm{T}} x + \sum_{j=1}^{n} d_{ij}^{\mathrm{T}} u_j, \tag{24}$$

其中: $c_i \in \mathbb{R}^{m_0}$, $d_{ij} \in \mathbb{R}^{m_j}$, $i, j = 1, \cdots, n$. 那么线性动态博弈(22)–(23)是势博弈, 并且如果每一个决策集$U_i \subset \mathbb{R}^{m_i} (i \in \mathbb{N})$是有界的, 那么线性动态博弈(22)–(23)有纯纳什平衡点.

根据以上的结果, 文[56]进一步提供了, 当决策集$U_i \subset \mathbb{R}^{m_i} (i \in \mathbb{N})$为多面体时, 寻找线性动态博弈(22)–(23)纳什均衡点的有效算法.

6 其他博弈问题

其他许多与博弈相关的问题也可以通过基于矩阵半张量积的代数状态空间表示来进行建模和分析. 例如电子游戏中的探路[57], 传教士与野人过河及狼羊白菜过河[25]等. 这些应用无法一一介绍, 下面介绍一些较系统的结果.

6.1 具有破产风险的网络演化博弈

具有破产风险的网络演化博弈在经济学问题中具有特殊的重要性. 利用矩阵半张量积方法研究这类网络演化博弈的策略调控与优化问题是一个富有挑战性的问题. 这方面的工作包括: 文[58]研究了具有破产风险和一步记忆的网络演化博弈策略调控问题, 并设计了自由控制序列使得从一初始局势出发, 所有玩家在博弈演化过程中均不破产. 文[59]针对这类博弈设计了一类状态反馈控制以避免博弈中玩家破产. 此外, 对于具有破产风险和多步记忆的网络演化. 文[60]设计了自由控制序列使得博弈整体收益达到最优. 最后, 文[61]对一类时变拓扑结构下具有多步记忆的网络演化博弈进行了稳定性分析, 并给出了博弈全局收敛到纯策略纳什均衡的充分条件.

6.2 从合作博弈到Shapley值

前面讨论的都是非合作博弈. 博弈论的另一个分支是合作博弈.

定义 8[62]

i) 一个合作博弈可表示为$G=(N,v)$, 其中$N=\{1,2,\cdots,n\}$为玩家集合, $v:2^N\to\mathbb{R}v(\varnothing)=0$称为特征函数. 特征函数$v(s)$, $s\subset N$ 表示当s 中的玩家结为联盟时这个联盟的收益.

ii) 合作博弈的基本问题是找到一个合理的分配. $x=(x_1,x_2,\cdots,x_n)\in\mathbb{R}^n$ 称为一个分配, 如果它满足: 1) "个体合理性", 即$x_i\geqslant v(\{x_i\})$; 2) "全局合理性", 即$\sum\limits_{i=1}^n x_i=v(N)$.

合作博弈的目的就是根据特征函数找到最合理的分配. 目前较为有效的分配有核、核心、核仁、Shapley值等.

矩阵半张量积在合作博弈中的应用包括: 特征函数基底(无异议特征函数)的矩阵表示[63], Shapley值的计算[64]等. 近期的一个工作是用半张量积方法给出模糊联盟下的分配, 称为路解[65].

下面以Shapley值的计算为例, 讨论如何用半张量积方法设计分配.

定义 9[66] 考察一个合作博弈$G=(N,v)$, 其中$N=\{1,2,\cdots,n\}$, $v\in\varGamma$ 为特征函数($\varGamma$为特征函数集合). Shapley值$\varphi:\varGamma\to\mathbb{R}^n$定义如下:

$$\varphi_i(v)=\frac{1}{n!}\sum_{\sigma\in\pi(N)}[v(S_\sigma^i\cup\{i\})-v(S_\sigma^i)]=\sum_{S\subseteq N\backslash\{i\}}\frac{|S|!(n-1-|S|)!}{n!}[v(S\cup\{i\})-v(S)],\quad i=1,2,\cdots,n,\tag{25}$$

这里$\pi(N)$指$N=\{1,2,\cdots,n\}$的置换集. S_σ^i指在置换σ中在i之前的玩家, 即

$$S_\sigma^i=\{j|\sigma^{-1}j\leqslant\sigma^{-1}i\}.$$

Shapley值作为分配有许多基本的的优良性质. 例如: 有效性、对称性、可加性, 因此得到广泛应用. 但是, 直接从定义(25)出发, 其计算极其复杂. 利用矩阵半张量积, 本文给出了简单的递推计算公式.

定理 7 考察一个合作博弈$G=(N,v)$, 设v给定, 其结构向量为C_v, 则相应的 Shapley 值 $\varphi(v)=(\varphi_1(v),\cdots,\varphi_n(v))$可计算如下:

$$\varphi(v)=C_v\varPhi_n,\tag{26}$$

这里

$$\varPhi_n=\frac{1}{n!}\left[\begin{pmatrix}\varepsilon_1^1\\-\varepsilon_1^1\end{pmatrix}\begin{pmatrix}\varepsilon_2^1\\-\varepsilon_2^1\\\varepsilon_2^2\\-\varepsilon_2^2\end{pmatrix}\cdots\begin{pmatrix}\varepsilon_{2^{n-1}}^1\\-\varepsilon_{2^{n-1}}^1\\\varepsilon_{2^{n-1}}^2\\-\varepsilon_{2^{n-1}}^2\\\vdots\\\varepsilon_{2^{n-1}}^{2^{n-1}}\\-\varepsilon_{2^{n-1}}^{2^{n-1}}\end{pmatrix}\right]\in\mathcal{M}_{2^n\times n},\tag{27}$$

参数$\varepsilon_{2^s}^i$可递归计算.

文[64]给出其计算细节及部分应用.

7 结束语

矩阵半张量积作为一个新的数学工具, 在有限博弈的建模与策略优化中起着全方位的作用. 它解决了一些用经典方法难以解决的问题, 显示了这种方法的有效性. 这篇综述不可能介绍半张量积在博弈中的全部应用, 但它对半张量积方法在博弈中的应用原理和一些主要结果做了尽可能详细的论述. 应该说半张量积方法在博弈中的应用尚处于尝试阶段, 还有许多问题值得探讨.

下面举出一些博弈中应用矩阵半张量积方法值得探索的方向: 1) 不完全信息(即贝叶斯)博弈的建模与优化策略; 2) 多目标优化的博弈方法, 特别是向量势博弈方法; 3) 近似势博弈, 特别是加权势博弈的优化和应用; 4) 依赖于状态的网络演化博弈的建模与学习算法.

矩阵半张量积对博弈问题的解决提供了一个新的武器. 对于熟悉或正在学习矩阵半张量积的学者, 博弈是一个宽阔的天地, 在那里是可以大有作为的.

感谢: 本文在写作过程中得到许多同行学者的帮助,

包括: 冯俊娥、吴玉虎、付世华、王金环、孟敏、王元华、李长喜、郝亚琦、张潇等, 特此致谢.

参考文献:

[1] VON NEUMANN J, MORGENSTERN O. *Theory of Game and Economic Behavior.* Princeton: Princeton University Press, 1944.

[2] GIBBONS R A. *Primer in Game Theory*. London: Prentice Hall, 1992.

[3] 王国成. 博弈论精粹. 哈尔滨: 哈尔滨工业大学出版社, 2008.

[4] 梅生伟, 刘锋, 魏韡. 工程博弈论基础及电力系统应用. 北京: 科学出版社, 2016.

[5] 杨宪东, 叶芳柏. 线性与非线性控制理论. 台北: 全华科技图书股份有限公司, 2007.

[6] 郭雷. 关于控制理论发展的某些思考. 系统科学与数学, 2012, 31(9): 1014 – 1018.

[7] GOPALAKRISHNAN R, MARDEN J R, WIERMAN A. An architectural view of game theoretic control. *ACM Sigmetrics Performance Evaluation Review*, 2011, 38(3): 31 – 36.

[8] MEI S, WANG Y, LIU F, et al. Game approaches for hybrid power system planning. *IEEE Transactions Sustainable Energy*, 2012, 3(3): 506 – 517.

[9] TEMBINE H, ALTMAN E, EL-AZOUZI R, et al. Evolutionary games in wireless networks. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 2010, 40(3): 634 – 646.

[10] TOMLIN C J, LYGEROS J, SASTRY S S. A game theoretic approach to controller design for hybrid systems. *Proceedings of the IEEE*, 2000, 88(7): 949 – 970.

[11] YAZICIOGLU A Y, EGERSTEDT M, SHAMMA J S. A game theoretic approach to distributed coverage of graphs by heterogeneous mobile agents. *Est Control Network System*, 2013, 4: 309 – 315.

[12] CHENG D, QI H, LIU Z. From STP to game-based control. *Science China Information Science* 2018, 61(1): 1 – 19.

[13] FORNASINI E, VALCHER M E. Recent developments in Boolean networks control. *Journal of Control and Decision*, 2016, 3(1): 1 – 18.

[14] LU J, LI H, LIU Y, et al. Survey on semi-tensor product method with its applications in logical networks and other finite-valued systems. *IET Control Theory and Applications*, 2017, 11(13): 2040 – 2047.

[15] MUHAMMAD A, RUSHDI A, GHALEB F A M. A tutorial exposition of semi-tensor products of matrices with a stress on their representation of Boolean function. *Journal of King Abdulaziz University*, 2016, 5: 3 – 30.

[16] 程代展, 齐洪胜. 矩阵的半张量积—理论与应用. 北京: 科学出版社, 2007.

[17] CHENG D, QI H, LI Z. *Analysis and Control of Boolean Networks: A Semi-tensor Product Approach*. London: Springer, 2011.

[18] CHENG D, QI H, ZHAO Y. *An Introduction to Semi-tensor Product of Matrices and Its Applications*. Singapoer: World Scientific, 2012.

[19] GUO P, WANG Y, LI H. Algebraic formulation and strategy optimization for a class of evolutionary networked games via semi-tensor product method. *Automatica*, 2013, 49(11): 3384 – 3389.

[20] CHENG D, HE F, QI H, et al. Model, analysis and control of networked evolutionary games. *IEEE Transaction Automation Control*, 2015, 61(9): 2402 – 2415.

[21] CHENG D, XU T, QI H. Evolutionarily stable strategy of networked evolutionary games. *IEEE Transactions on Neural Networks and Learning Systems*, 2014, 25(7): 1335 – 1345.

[22] WANG Y, CHENG D. Stability and stabilization of a class of finite evolutionary games. *Journal of Franklin Institute*, 2017, 354(3): 1603 – 1617.

[23] ZHAO G, WANG Y, LI H. A matrix approach to modeling and optimization for dynamic games with random entrance. *Applied Mathematics and Computation*, 2016, 290: 9 – 20.

[24] LI Y, LI H, XU X, et al. Semi-tensor product approach to minimal-agent consensus control of networked evolutionary games. *IET Control Theory and Applications*, 2018, 12(16): 2269 – 2275.

[25] ZHAO G, LI H, SUN W, et al. Modelling and strategy consensus for a class of networked evolutionary games. *International Journal System Science*, 2018, 49(12): 2548 – 2557.

[26] ZHAO G, WANG Y. Formulation and optimization control of a class of networked evolutionary games with switched topologies. *Nonlinear Analysis-Hybrid Systems*, 2016, 22: 98 – 107.

[27] LI H, DING X, YANG Q, et al. Algebraic formulation and Nash equilibrium of competitive diffusion games. *Dynamic Games and Applications*, 2018, 8: 423 – 433.

[28] ZHANG X, CHENG D. Profile-dynamic based fictitious play. *Science China Information Science*, DOI: 10.1007/s11432-019-9926-2.

[29] LI H, DING X, ALSAEDI A, et al. Stochastic set stabilization of n-person random evolutionary Boolean games and its applications. *IET Control Theory and Applications*, 2017, 11(13): 2152 – 2160,.

[30] DING X, LI H, YANG Q, et al. Stochastic stability and stabilization of n-person random evolutionary Boolean games. *Mathematics and Computation*, 2017, 306: 1 – 12.

[31] DING X, LI H, ALSAADI F E. Regulation of game result for n-person random evolutionary Boolean games. *Asian Journal of Control*, 2019, DOI: 10.1002/asjc.2119.

[32] DING X, LI H. Optimal control of random evolutionary Boolean games. *International Journal of Control*, 2019, DOI: 10.1080/00207179.2019.1585957.

[33] MONDERER D, SHAPLEY L S. Potential games. *Games and Economic Behavior*, 1996, 14: 124 – 143.

[34] HINO Y. An improved algorithm for detecting potential games. *International Journal of Game Theory*, 2011, 40: 199 – 205.

[35] CHENG D. On finite potential games. *Automatica*, 2014, 50(7): 1793 – 1801.

[36] LIU X, ZHU J. On potential equations of finite games. *Automatica*, 2016, 68: 245 – 253.

[37] ZHANG X, HAO Y, CHENG D. Incomplete-profile potential games. *Journal of the Franklin Institute*, 2018, 355: 862 – 877.

[38] HAO Y, CHENG D. Finite element approach to continuous potential games. *Science China Information Science*, DOI: 10.1007/s11432-018-9763-7.

[39] LI C, HE F, LIU T, et al. Verification and dynamics of group-based potential games. *IEEE Transaction Control Netwk System*, 2018, 6(1): 215 – 224.

[40] LIU T, WANG Y, CHENG D. Dynamics and stability of potential hyper-networked evolutionary games. *International Journal of Automation and Computing*, 2017, 14(2): 229 – 238.

[41] YU Yongyuan, FENG Jun'e, PAN Jinfeng. Ordinal potential game and its application in agent wireless networks. *Control and Decision*, 2017, 32(3): 393 – 402.
(于永渊, 冯俊娥, 潘金凤. 有序势博弈及其在智能体无线网络中的应用. 控制与决策, 2017, 32(3): 393 – 402.)

[42] CANDOGAN O, MENACHE I, OZDAGLAR A, et al. Flows and decompositions of games: Harmonic and potential games. *Mathematcs of Operations Research*, 2011, 36(3): 474 – 503.

[43] CHENG D, LIU T, ZHANG K. On decomposed subspaces of finite games. *IEEE Transaction Automation Control*, 2016, 61(11): 3651 – 3656.

[44] WANG Y, LIU T, CHENG D. From weighted potential game to weighted harmonic game. *IET Control Theory and Applications*, 2017, 11(13): 2161 – 2169.

[45] CHENG D, LIU T. Linear representation of symmetric games. *IET Control Theory and Applications*, 2017, 11(18): 3278 – 3287.

[46] HAO Y, CHENG D. On skew-symmetric games. *Journal of the Franklin Institute*, 2018, 355: 3196 – 3220.

[47] LI C, HE F, LIU T, et alx. Symmetry-based decomposition of finite games. *Science China Information Science*, 2019, 62(1): 164 – 176.

[48] CHENG D, LIU T. From Boolean game to potential game. *Automatica*, 2018, 96: 51 – 60.

[49] HAO Y, PAN S, QIAO Y, et al. Cooperative control via congestion game approach. *IEEE Transaction Automation Control*, 2018, 63(12): 4361 – 4366.

[50] LE S, WU Y, SUN X. Congestion games with player-specific utility functions and its application to NFV networks. *IEEE Transactions on Automation Science and Engineering*, 2019, DOI: 10.1109/TASE. 2019.2899504.

[51] MARDEN J R, ARSLAN G, SHAMMA J S. Joint strategy fictitious play with inertia for potential games. *IEEE Transactions on Automatic Control*, 2009, 54(2): 208 – 220.

[52] ALÓS-FERRER C, NETZER N. The logit-response dynamics. *Games and Economic Behavior*, 2010, 68(2): 413 – 427.

[53] LI C, XING Y, HE F, et al. A strategic learning algorithm for state-based games. *Automatica*, 2019, 55(12).

[54] MARDEN J R. State based potential games. *Automatica*, 2012, 48(12): 3075 – 3088.

[55] LIU T, WANG J, ZHANG X, et al. Game theoretic control of multi-agent systems. *SIAM Journal on Control and Optimization*, 2019, 57(3): 1691 – 1709.

[56] WU Y, TOYODA M, SHEN T. Linear dynamic games with polytope strategy sets. *IET Control Theory and Applications*, 2017, 11(13): 2146 – 2151.

[57] ZHANG L, FENG J. Mix-valued logic-based formation control. *International Journal of Control*, 2013, 86(6): 1191 – 1199.

[58] FU S, WANG Y, ZHAO G. A matrix approach to the analysis and control of networked evolutionary games with bankruptcy mechanism. *Asian Journal of Control*, 2017, 19(2): 717 – 727.

[59] DENG L, FU S, ZHU Y. State feedback control design to avoid players going bankrupt. *Asian Journal of Control*, 2019, DOI: 10.1002/ asjc.2126.

[60] FU S, LI H, ZHAO G. Modeling and strategy optimization for a kind of networked evolutionary games with memories under the bankruptcy mechanism. *International Journal of Control*, 2018, 91(5): 1104 – 1117.

[61] FU S, ZHAO G, LI H, et al. Model and control for a class of networked evolutionary games with finite memories and time-varying networks. *Circuits, Systems, and Signal Processing*, 2017, 37(7): 3093 – 3114.

[62] XIE Zheng. *Game Theory: An Introduction*. Beijing: Science Press, 2010.
(谢政. 对策论导引. 北京: 科学出版社, 2010.)

[63] CHENG D, XU T. Application of STP to cooperative games. *2013 10th IEEE International Conference on Control and Automation (ICCA)*, Hangzhou: IEEE, 2013: 1680 – 1685.

[64] WANG Y, CHENG D, LIU X. Matrix expression of Shapley values and its application to distributed resource allocation. *Science China Information Science*, 2019, 62(2): 46 – 56.

[65] GE Meixia, ZHAO Jianli, LI Ying, et al. Application of STP to path solution in cooperative games. *Journal of Mathematics in Practice and Theory*, 2018, 48(14): 237 – 242.
(葛美侠, 赵建立, 李莹, 等. 半张量积在合作博弈路解中的应用. 数学的实践与认识, 2018, 48(14): 237 – 242.)

[66] SHAPLEY L S. A value for n-person games. *Contributions to the Theory of Games*, 1953, 2(28): 307 – 317.

作者简介:

程代展 研究员, 本科毕业于清华大学, 在中国科学院数学与系统科学研究院取得硕士学位, 在美国(St.Louis)华盛顿大学取得博士学位. 发表论著14本, 期刊论文250 余篇, 会议论文150余篇. 曾任中国自动化学会理事, 控制理论专业委员会主任, 国际自动化学会(IFAC)Fellow, 理事(Council Member), IEEE Fellow, IEEE CSS执 委 会(Board of Governors)成员. 两次(第一完成人)获国家自然科学二等奖; 中国科学院杰出科技成就奖(个人)金质奖章. 并曾获由IFAC颁发的其旗舰杂志Automatica三年一篇的最佳理论/方法论文奖(2008–2010), 为迄今唯一由华人学者完成的获奖论文. 目前研究方向为非线性控制、矩阵半张量积的理论与应用、逻辑动态系统、博弈论等, E-mail: dcheng@iss.ac.cn;

刘泽群 目前研究方向为矩阵半张量积的理论与应用、逻辑动态系统、博弈论等, E-mail: liuzequn@amss.ac.cn.