第三期矩阵半张量积暑期班

第3讲: 有限博弈的半张量积方法
3.1 有限博弈初步

李长喜
lichangxi@pku.edu.cn

北京大学, 工学院

2023年08月14日, 聊城大学

主要内容

1. 博弈的数学模型
2. 纳什均衡
3. 混合策略
4. 博弈与伪逻辑函数
5. 作业

1.1 有限博弈的数学模型

☞ 什么是博弈论?

- 博弈论也称对策论(game theory). 各种博弈游戏和朴素的对策思想自从人类社会的出现就产生了, 中国田忌赛马的故事(公元前300余年), 国外如犹太教法典《塔木德》(约公元500年) 中的遗产分配, 等, 都是有记载的博弈论故事. 而现代博弈论则起源于上世纪. 通常认为, 冯诺伊曼(Von Neumann) 等的论著《博弈论与经济行为》标志着现代博弈论的诞生.
- 通常将博弈论分为两部分: 非合作博弈与合作博弈. 纳什(Nash) 和他提出的纳什均衡是非合作博弈的基础, 沙普利(Shapley) 等人是合作博弈的代表人物.

图 **1:** John von Neumann

J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, New Jersey, 1944.

☞ 为什么要学习博弈论?

- 《博弈论精粹》: “有两类人可以不学博弈论: 一是漂流到类似于鲁宾逊所在的荒岛上; 二是到了什么事都由自己说了算, 不受其他人影响.”
- Paul A. Samuelson (1915-2009, Nobel Memorial Prize in Economic Sciences(1970))“ 要想在现代社会做一个成功的人, 必须对博弈论有一个大致的了解”.
- 博弈控制论(Game Theoretic Control) 是系统控制理论的一个新生长点.
- 2023张嗣瀛奖, 关肇直奖...

- 刘培杰, “博弈论精粹”, 哈尔滨工业大学出版社, 2008.
- 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.
- 国务院. 新一代人工智能发展规划. 国发(2017) 35号, 2017.07.08

☞ 有限博弈

Definition 1.1

一个(有限)正规博弈(normal game) 由三个部分组成: $G = (N,\ S, C)$, 这里

- $N = \{1, 2, \cdots, n\}$ 表示玩家, 这里有n 个玩家;
- $S = \prod_{i=1}^{n} S_i$ 称为局势(profile), 这里

  $$S_i = \{s_1^i, \cdots, s_{k_i}^i\}, \quad i = 1, \cdots, n,$$

  是第i 个玩家的策略(strategy, action) 集, 表示第i 个玩家有k_i 个策略可选择. 局势是所有玩家策略集的笛卡尔积.
- $C = (c_1, \cdots, c_n)$, 这里$c_i : S \to \mathbb{R}$ 是第i 个玩家的收益函数(payoff function).

为方便计, 策略集通常记为

$$S_i = \{1, 2, \cdots, k_i\}, \quad i = 1, \cdots, n.$$

这是“有限”博弈指: (i) 玩家数$n < \infty$, (ii) 策略数$|S_i| < \infty$, $i = 1, \cdots, n$. 本课程只讨论有限博弈.

☞ 几个例子

Example 1.2 [囚徒困境 (prisoner's dilemma)]

两个小偷被捕. 由于证据不足, 如果两小偷合作(不招供), 则每人将被判刑1 年. 如果一个背叛(即招供), 另一个合作, 则背叛者无罪释放, 合作者被判9 年. 如果两个均背叛, 则每人将被判刑6 年. 讨论每个囚徒将如何决策? 这是个典型的博弈问题. (i) $N=\{1,2\}$ 即玩家1 与即玩家2; (ii) 策略集$S_1=S_2=\{1,2\}$, 这里1 代表合作, 2 代表背叛; (iii) 支付函数c_i. 通常支付函数用一个表来表示. 见表1.

Example 1.2 (cont'd)

表 1: 囚徒困境(双矩阵)

甲\乙	1	2
1	(1, 1)	(9, 0)
2	(0, 9)	(6, 6)

表1 给出的称为一个双矩阵(bi-matrix), 每个位置有两个数, 前面是玩家1 所得, 后面是玩家2 所得. 例如, 如果玩家1 取策略1, 玩家2 取策略2, 查$(1, 2)$ 位置可知, 玩家1 判9 年, 玩家2 判0 年.

Example 1.2 (cont'd)

其实,一个有限博弈的全部信息都表现在双矩阵中了. 双矩阵虽然被广泛应用, 但是, 当玩家多于两人时它很不方便, 我们试图用单矩阵代替它. 表2 是表1 的另一种形式:

表 2: 囚徒困境(单矩阵)

$c_i \backslash s$	11	12	21	22
c_1	1	9	0	6
c_2	1	0	9	6

表2中第一行表示不同的局势, 第二行是相关局势下玩家1 所得, 第三行是相关局势下玩家2 所得. 下面的例子就看出它的好处.

Example 1.3

3 个人玩手心手背, 如果某人与其他两人不同, 则他输2 分, 其他二人各赢1 分. 记玩家为1, 2, 3, 则有如下收益单矩阵, 见表3.

表 3: 手心手背

$c_i \backslash s$	111	112	121	122	211	212	221	222
c_1	0	1	1	-2	-2	1	1	0
c_2	0	1	-2	1	1	-2	1	0
c_3	0	-2	1	1	1	1	-2	0

Remark 1.4

在支付表的第一行, 局势应严格按字典序排列. 即先让最后一个玩家的策略s_n 从1 变到k_n, 然后是倒数第二个玩家的策略s_{n-1} 从1 变到k_{n-1}, 等等, 直到第一个玩家的策略s_1 从1 变到k_1. 这样, 才能使支付矩阵有唯一性. 后面将看到, 这种排序在应用上十分有效.

如果在博弈中, 每种局势下所有收益和为零, 即

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

则称该博弈为零和博弈. 不难验证, 例1.3 为零和博弈.

1.2 纳什均衡

Definition 2.1

给定一个有限博弈$G = (N, S, C)$. 一个局势$s* = (s_1^*, s_2^*, \cdots, s_n^*)$ 称为一个纳什均衡, 如果

$$c_i(s_1^*, s_2^*, \cdots, s_n^*) \geq c_i(s_1^*, \cdots, s_{i-1}^*, s_i, s_{i+1}^*, \cdots, s_n^*), \quad s_i \in S_i, \ i = 1, \cdots, n. \tag{1}$$

纳什均衡表示竞争博弈中一个大家都能接受的策略集合, 因为任何人单独改变策略都会吃亏(至少不会占便宜).
对有限博弈, 很容易从支付矩阵找到纳什均衡点, 我们用几个例子来说明.

Example 2.2

(智猪博弈) 猪圈里有一大猪(B)、一小猪(S). 猪圈一边是猪食槽, 另一边是开关. 拱了开关, 猪食槽才能打开. 食槽一次装6斤食. 如果小猪去拱开关, 大猪将吃光饲料; 如果大猪去拱, 小猪将吃到5 斤, 大猪将吃到1 斤. 如果一起去拱, 小猪将吃到2 斤, 大猪将吃到4 斤. 如果拱一次开关消耗半斤能量,找出纳什均衡点.

图 2: 智猪博弈

表 4: 智猪博弈支付双矩阵

$B\backslash S$	拱	不拱
拱	3.5 1.5	$\underline{0.5}$ $\underline{5}$
不拱	$\underline{6}$ -0.5	0 $\underline{0}$

Nash Eqilibrium: B-拱, S-不拱.

Example 2.3

考虑一个有限博弈$G=(N,\ S,\ C)$, 这里$N=\{1,2,3\}$, $S_1=\{1,2,3\}$, $S_2=\{1,2\}$, $S_3=\{1,2,3\}$. 支付矩阵见表.

表 5: 例2.3 支付矩阵

$C\backslash P$	111	112	113	121	122	123	211	212	213
c_1	1	$\underline{2}$	-1	-2	0	1	-2	1	$\underline{1}$
c_2	2	$\underline{3}$	$\underline{4}$	$\underline{3}$	2	1	$\underline{3}$	2	$\underline{2}$
c_3	-2	-1	$\underline{0}$	-4	$\underline{-2}$	-3	-3	-2	$\underline{0}$
$C\backslash P$	221	222	223	311	312	313	321	322	323
c_1	$\underline{1}$	0	$\underline{2}$	$\underline{3}$	$\underline{2}$	$\underline{1}$	-1	$\underline{2}$	-2
c_2	2	$\underline{3}$	1	3	2	$\underline{4}$	$\underline{5}$	$\underline{3}$	1
c_3	-1	-1	$\underline{0}$	$\underline{0}$	-3	-3	-2	$\underline{-1}$	$\underline{-1}$

Example 2.3 (cont'd)

考虑c_1, 比较它在111, 211 以及311 的值, 这里$s^{-1} := \prod_{i\neq 1} s_i$ 的策略是一样的. 准确地说, $s^{-1} = s_2 \times s_3 = 11$. 由于$c_1(111) = 1$, $c_1(211) = -2$, 及$c_1(311) = 3$, 故

$$c_1(311) = \max_{s_1 \in S_1} c_1(s_1, 1, 1).$$

那么, 我们在最大值3 下加一下划线. 类此, 我们对每个$s^{-1} \in S^{-1}$ 寻找最大值

$$\max_{s_1 \in S_1} c_1(s_1, s^{-1}), \quad s^{-1} \in S^{-1},$$

在最大值下加下划线.

Example 2.3 (cont'd)

对

$$\max_{s_2 \in S_2} c_2(s_2, s^{-2}), \quad s^{-2} \in S^{-2},$$

和

$$\max_{s_3 \in S_3} c_3(s_3, s^{-3}), \quad s^{-3} \in S^{-3},$$

做同样的事. 然后考察支付矩阵的每一列. 如果某一列所有元素都加了下划线, 这一列所对应的局势就是纳什均衡点. 最后, 不难看出, 本例有两个纳什均衡点, 即$(2, 1, 3)$及$(3, 2, 2)$.

1.3 混合策略

是否所有的有限博弈都有满足(1) 的纳什均衡呢? 我们考察下面这个例子.

Example 3.1

两 人 玩 石 头-剪 刀-布(Rock-Scissors-Paper) 游 戏. 记 策略$R = 1$, $S = 2$, $P = 3$. 则有如下支付矩阵(见表6

表 6: 石头-剪刀-布

$C\backslash P$	11	12	13	21	22	23	31	32	33
c_1	0	$\underline{1}$	-1	-1	0	$\underline{1}$	$\underline{1}$	-1	0
c_2	0	-1	$\underline{1}$	$\underline{1}$	0	-1	-1	$\underline{1}$	0

Example 3.1 (cont'd)

用前节办法我们将每种情况下的最佳反应划出. 从表中可以看出, 这个博弈没有定义2.1 所指的纳什均衡. 玩过石头-剪刀-布的人都知道, 应当随机地在“ 石头”, “剪刀”, “布”中任选一个. 这种策略称为混合策略(mixed strategy).

Definition 3.2

给定一个正规博弈.

- 记
$$\bar{S}_i = \left\{ \left(r_1^i, \cdots, r_{k_i}^i\right) | r_j^i \geq 0, \sum_{j=1}^{k_i} r_j^i = 1 \right\}.$$
$\bar{s}_i = \left(r_1^i, \cdots, r_{k_i}^i\right) \in \bar{S}_i$ 称为一个混合策略, 它表示第i 个玩家以概率r_j^i 取策略s_j^i (简称: 策略j), $j = 1, \cdots, k_i$.
相应的局势集合记为
$$\bar{S} = \prod_{i=1}^{n} \bar{S}_i.$$
- $\bar{s}^* = (\bar{s}_1^*, \cdots, \bar{s}_n^*)$ 称为一个纳什均衡, 如果支付函数的期望值满足
$$\begin{aligned} Ec_i(\bar{s}_1^*, \cdots, \bar{s}_n^*) \geq Ec_i(\bar{s}_1^*, \cdots, \bar{s}_{i-1}^*, \bar{s}_i, \bar{s}_{i+1}^*, \cdots, \bar{s}_n^*), \\ \bar{s}_i \in \bar{S}_i, \ i = 1, \cdots, n. \end{aligned} \quad (2)$$

由于引进了混合策略, 前面定义的$s_i \in S_i$ 就称为纯策略(pure strategy). 当然, 纯策略是一种特殊的混合策略.

Example 3.3

回忆石头、剪刀、布游戏，如果允许使用混合策略, 不妨设, 第一个玩家的混合策略为$(p_1, p_2, 1-p_1-p_2)$, 第二个玩家的混合策略为$(q_1, q_2, 1-q_1-q_2)$. 于是, 期望值为

$$\begin{aligned} Ec_1 &= p_1q_2 - p_1(1-q_1-q_2) - p_2q_1 + p_2(1-q_1-q_2) \\ &\quad +(1-p_1-p_2)q_1 - (1-p_1-p_2)q_2 \\ Ec_2 &= -p_1q_2 + p_1(1-q_1-q_2) + p_2q_1 - p_2(1-q_1-q_2) \\ &\quad -(1-p_1-p_2)q_1 + (1-p_1-p_2)q_2. \end{aligned}$$

Example 3.3(cont'd)

注意到纳什均衡是最佳响应函数的解, 计算

$$\frac{\partial Ec_1}{\partial p_1} = q_2 - (1 - q_1 - q_2) - q_1 + q_2 := 0,$$

可得$3q_2 = 1$, 于是

$$q_2^* = \frac{1}{3}.$$

同样, 由

$$\frac{\partial Ec_1}{\partial p_2} = 0; \quad \frac{\partial Ec_2}{\partial q_1} = 0; \quad \frac{\partial Ec_2}{\partial q_2} = 0$$

可分别得到

$$q_1^* = \frac{1}{3}; \quad p_2^* = \frac{1}{3}; \quad p_1^* = \frac{1}{3}.$$

Example 3.3(cont'd)

因此, $\sigma^* = (\sigma_1^*, \sigma_2^*)$ 是"剪刀石头布"博弈的混合策略纳什均衡, 其中 $\sigma_1^* = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$, $\sigma_2^* = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$.

为了区别, 由纯策略组成的纳什均衡称为纯纳什均衡.
以下定理是纳什的一个主要贡献.

Theorem 1

(Gibbons 92) 任何有限博弈总存在纳什均衡, 但它可能是由混合策略构成.

R. Gibbons, *A Primer in Game Theory*, Pearson Education Lim., Harlow, 1992.

☞ 纳什均衡的意义

John W. Milnor (1931-) Topologist, Fields Medal(1962), Wolf Prize(1989), Abel Prize(2011)：

“纯粹数学家对任何数学工作的评价往往基于它在数学上的深度和广度, 即或者引进了新的数学思想和方法, 或者解决了长期悬而未决的问题. 按照这种方式看, 纳什的获奖工作只是一个巧妙但并不出人意料的对熟知方法的应用. 但是, 当数学被应用到人类科学的其他分支时,我们必须认真地提出一个完全不同的问题: 这个新的工作能使我们对现实世界的理解增加到何等程度? 基于这个理由, 那么Nash 的论文完全不逊色于一场革命. ”

1.4 博弈与伪逻辑函数

☞ 伪逻辑函数

Definition 4.1

设$x_i \in \mathcal{D}$, $i = 1, \cdots, n$. 一个映射$f : \mathcal{D}^n \to \mathbb{R}$, 记作$f(x_1, \cdots, x_n)$, 称为一个伪布尔函数.

当x_i 用向量表示时, 即令$x_i \in \Delta$, 类似于布尔函数, 伪布尔函数也有如下代数表达式.

Proposition 4.2

考查一个伪布尔函数$f : \mathcal{D}^n \to \mathbb{R}$. 当$x_i$ 表示为向量形式时, 则存在唯一行向量$V_f \in \mathbb{R}^{2^n}$, 称为f 的结构向量, 使得

$$f(x_1, \cdots, x_n) = V_f \ltimes_{i=1}^{n} x_i. \tag{3}$$

伪布尔函数有很广泛的应用，特别是其优化问题曾被广泛讨论. 伪布尔函数的一个直接推广就是伪逻辑函数.

Definition 4.3

设$x_i \in \mathcal{D}_{k_i}$, $i = 1, \cdots, n$. 一个映射$f : \prod_{i=1}^{n} \mathcal{D}_{k_i} \to \mathbb{R}$, 记作$f(x_1, \cdots, x_n)$, 称为一个伪逻辑函数. 当$k_1 = k_2 = \cdots = k_n := k_0$ 时, f 称为k_0-值伪逻辑函数, 否则称为混合值伪逻辑函数.

同样, 在向量形式下, 令$x_i \in \Delta_{k_i}$, 伪逻辑函数也有如下代数表达式.

Proposition 4.4

考查一个伪逻辑函数$f : \prod_{i=1}^{n} \mathcal{D}_{k_i} \to \mathbb{R}$. 当$x_i$ 表示为向量形式时, 则存在唯一行向量$V_f \in \mathbb{R}^k$ (这里$k = \prod_{i=1}^{n} k_i$), 称为f 的结构向量, 使得

$$f(x_1, \cdots, x_n) = V_f \ltimes_{i=1}^{n} x_i. \tag{4}$$

回到有限博弈, 考虑策略集$S_i = \{1, 2, \cdots, k_i\}$. 只要将$j$ 与$\delta_{k_i}^j$ 等同, 则$S_i \sim \Delta_{k_i}$, $i = 1, \cdots, n$. 则支付函数$c_i : \prod_{i=1}^n \Delta_{k_i} \to \mathbb{R}$, $i = 1, \cdots, n$, 就成为伪逻辑函数. 记c_i 的结构向量为V_i, 则$(V_1, V_2, \cdots, V_n)$ 唯一确定了一个有限博弈.

Proposition 4.5

$|N| = n$, $|S_i| = k_i$, $i = 1, \cdots, n$ 的有限博弈全体构成一个nk 维线性空间$\mathbb{R}^{nk}$, 这里$k = \prod_{i=1}^n k_i$. 记

$$V := (V_1, V_2, \cdots, V_n) \in \mathbb{R}^{nk}, \tag{5}$$

这里V_i 是玩家i 的支付函数的结构向量.

以后会看到, 有限博弈的线性空间结构极具重要性. 这也是为什么STP可以处理有限博弈.

Example 4.6

考虑例2.3, 支付矩阵为:

表 7: 例2.3 支付矩阵

$C \backslash P$	111	112	113	121	122	123	211	212	213
c_1	1	$\underline{2}$	-1	-2	0	1	-2	1	$\underline{1}$
c_2	2	$\underline{3}$	$\underline{4}$	$\underline{3}$	2	1	$\underline{3}$	2	$\underline{2}$
c_3	-2	-1	$\underline{0}$	-4	$\underline{-2}$	-3	-3	-2	$\underline{0}$

$C \backslash P$	221	222	223	311	312	313	321	322	323
c_1	$\underline{1}$	0	$\underline{2}$	$\underline{3}$	$\underline{2}$	$\underline{1}$	-1	$\underline{2}$	-2
c_2	2	$\underline{3}$	1	3	2	$\underline{4}$	$\underline{5}$	$\underline{3}$	1
c_3	-1	-1	$\underline{0}$	$\underline{0}$	-3	-3	-2	$\underline{-1}$	$\underline{-1}$

Example 4.6(cont'd)

$$V = [V_1, V_2, V_3]$$
$$= [\underbrace{1, 2, -1, -2, 0, 1, -2, 1, 1, 1, 0, 2, 3, 2, 1, -1, 2, -2}_{V_1},$$
$$\underbrace{2, 3, 4, 3, 2, 1, 3, 2, 2, 2, 3, 1, 3, 2, 4, 5, 3, 1}_{V_2}$$
$$\underbrace{-2, -1, 0, -4, -2, -3, -3, -2, 0, -1, -1, 0, 0, -3, -3, -2, -1, -1}_{V_3}$$

基于STP的两个研究方向

- 静态分析: 向量空间结构, 类型验证
- 动态分析: 演化博弈/网络演化博弈分析

作业

1. Benoit & Krishna 博弈的支付双矩阵见表8，其中策略1: 抵赖, 其中策略2: 胡扯, 其中策略3: 坦白.

表 8: Benoit & Krishna 博弈

$B\backslash K$	1	2	3
1	$10,\ 10$	$-1,\ -12$	$-1,\ 15$
2	$-12,\ -1$	$8,\ 8$	$-1,\ -1$
3	$15,\ -1$	$-1,\ -1$	$0,\ 0$

- 将支付表表示为单矩阵形式.
- 它是否有纯纳什均衡?

2. 两个人猜硬币, 每人各出一枚. 揭开后, 若均为正面朝上或正面朝下, 则甲赢. 若一枚正面朝上而另一枚正面朝下, 则乙赢. 找出其纳什均衡点.

3. 将Benoit & Krishna 博弈(见题1) 中两个玩家的支付函数写成伪逻辑函数的形式.

谢　谢　！

Q&A