非对称协调博弈中的对数追踪与均衡选择

Yali Dong[a], Boyu Zhang[b, *]

[a] 系统科学学院，北京师范大学，中国，100875；

[b] 数学与复杂系统实验室，教育部，数学科学学院，北京师范大学，中国北京，100875，

[*] 通讯作者：zhangby@bnu.edu.cn

摘要： 我们研究有限个参与者和行动的非对称纯协调博弈中的均衡选择。对数正则化定义了一条从均匀策略组合出发的内部均衡追踪型路径。我们将均衡条件约化为单峰标量函数，并证明当一个相关收益指标具有唯一最小值时，从均匀策略组合出发的连通均衡分支将收敛于相应的协调结果。对于两名玩家，该条件等价于使收益平方根之和最大，从而将已知的双行动结果推广到有限个行动。一个三人博弈示例表明，平方根规则通常不能推广到更多参与者的情形。

关键词： 协调博弈；均衡选择；对数正则化；追踪程序。

1. 引言

协调博弈通常具有多个纳什均衡。尽管每个均衡都具有自我执行性，但仅凭均衡条件无法确定参与者将协调于哪种结果。均衡选择程序通过规定如何选择某个特定均衡，来解决这种不确定性。Harsanyi追踪程序通过从初始先验出发，追踪辅助博弈的均衡，直至原博弈的某个均衡，从而处理这一问题（Harsanyi，1975年；Harsanyi和Selten，1988年）。相关的同伦方法研究通过改变参数所生成的均衡路径，其中包括基于量化反应均衡的路径（McKelvey and Palfrey，1995年；Turocy，2005年；Zhang，2016年；Zhang，Hofbauer，2016年）。

本文研究一种追踪型选择方法，适用于参与者和行动数量有限的非对称纯协调博弈。当所有参与者选择相同行动时，他们会获得正收益，且收益可能各不相同；否则收益为零。我们考虑一类

博弈，其中每位参与者的物质收益与对数正则化项相结合。在初始参数值下，唯一的均衡是均匀策略组合。随着物质收益的权重增加，我们考察均衡图中包含该初始均衡的连通分支。这一构造与对数追踪（Harsanyi，1973；Zhang、Hofbauer，2016）有关，但不同于 Harsanyi 最初提出的追踪程序（Harsanyi，1975）。

我们的分析将均衡条件归结为一族单峰标量函数，每个协调结果对应一个函数。每个函数都有一个最大值，该最大值定义了相应的收益指标。我们证明，如果某个结果是该指标的唯一最小值，那么从均匀策略组合出发的均衡连通分支将收敛于该结果。对于两名玩家，最小化该指标等价于最大化收益平方根之和。这将 Zhang and Hofbauer（2016） 的双行动结果推广到了任意有限数量的行动。对于三名或更多玩家，平方根规则通常不成立，三人博弈示例对此作了说明。

本文结构如下。第2节介绍对数正则化博弈，并定义用于选择的均衡连通分支。第3节确立选择结果，推导两名玩家的规则，并给出三人博弈示例。第4节讨论该判据的适用范围及可能的扩展。

2. 模型

考虑一个有 $n \geq 2$ 名参与者和 $m \geq 2$个行动的非对称纯协调博弈。每名参与者的行动集均为$\{1, \dots, m\}$。如果所有参与者都选择行动 j，参与者i获得收益 $a_{ij} > 0$；其他所有行动组合下，每名参与者的收益均为零。

令 x_{ij} 表示参与者 i 选择行动 j的概率，其中 $\sum_{j=1}^{m} x_{ij} = 1$。参与者彼此独立地随机选择，因此所有参与者都选择行动 j 的概率为 $P_j(x) = \prod_{i=1}^{n} x_{ij}$。对于 $\lambda \geq 0$，参与者 i 选择混合策略以最大化

$$\lambda \sum_{j=1}^{m} a_{ij} P_j(x) + \sum_{j=1}^{m} \log x_{ij} \qquad (1)$$

满足 $\sum_{j=1}^{m} x_{ij} = 1$ 和 $x_{ij} \geq 0$。对数项在边界处赋值为 $-\infty$ 。参数 λ 衡量物质收益相对于对数正则化所占的权重。[1]

[1] 将（1）乘以正数 $\frac{1}{\lambda+1}$ 并记 $\theta = \frac{\lambda}{\lambda+1}$ 可得 $\theta \sum \quad a_{ij} P_j(\sum_{j=1}^{m} \quad x \ +)(1-\theta) \sum \log_{j=1}^{m} \quad x_{ij}$，这就是 Harsanyi（1973）提出的对数博弈，其形式为 Zhang and Hofbauer

对于每个 λ，目标函数关于参与者 i自身的策略严格凹：物质收益项关于 x_i呈线性，而对数项严格凹。目标函数在边界处趋于 $-\infty$ ，因此每位参与者都有唯一的内部最佳响应。最佳响应映射是连续的，因此布劳威尔不动点定理保证均衡存在。当 $\lambda = 0$时，每位参与者通过均匀选择使对数和达到唯一最大值：

$$x_{ij} = \frac{1}{m} \text{ for all } i, j. \tag{2}$$

这一参数化族定义了一种追踪型同伦：它追踪包含 $\lambda = 0$ 处均匀策略组合的均衡连通分支，并随着 λ 增大而跟踪均衡。我们将这种构造称为“竞速型”。

令 E 表示均衡图，

$$E = \{(\lambda, x) \in [0, \infty) \times \Delta_m^n : x \text{ is an equilibrium at } \lambda\},$$

并令 C_0 表示包含均匀均衡 $(0, \bar{x})$ 的 E 的连通分支，其中 $\bar{x}_{ij} = \frac{1}{m}$。将该连通分支的极限集定义为

$$L(C_0) = \{x^* : \exists\, (\lambda_r, x^r) \in C_0 \text{ with } \lambda_r \to \infty \text{ and } x^r \to x^*\}.$$

若 $L(C_0)$ 是单点集，则其元素就是追踪型程序选出的结果。此定义针对均衡图的一个连通分支，并不要求该分支上的均衡构成 λ (Zhang，2016年) 的单值函数。

为说明沿该连通分支取得的极限是基础博弈的纳什均衡，令$v_{ij}(x_{-i})$ 表示参与者 i 采用纯行动 j 时的期望物质收益，并令 $\bar{v}_i(x) = \sum_{j=1}^{m} x_{ij} v_{ij}\ x_{-i}$ ()表示参与者 i 的期望物质收益。(1) 的一阶条件为

$$\lambda v_{ij}(x_{-i}) + \frac{1}{x_{ij}} = \mu_i,\ j = 1, \ldots, m \tag{3}$$

其中， μ_i 是概率约束的乘子。将（3）乘以 x_{ij} 并对 j 求和，得到 $\mu_i = \lambda \bar{v}_i(x) + m$。因此，

$$\lambda \left(v_{ij}(x_{-i}) - \bar{v}_i(x) \right) = m - \frac{1}{x_{ij}}.$$

由于 $x_{ij} \leq 1$，单方面偏离所产生的遗憾满足

（2016年）。由于乘子为正，因此这两种设定在每个 λ下都有相同的均衡。

$$0 \le \max_j v_{ij}(x_{-i}) - \bar{v}_i(x) \le \frac{m-1}{\lambda}. \tag{4}$$

因此， $L(C_0)$ 的每个元素都是基础博弈的纳什均衡。问题在于， C_0 会达到哪个均衡。

3. 结果

对于每个行动 j，定义 $f_j(t) = \frac{t}{\prod_{i=1}^{n}(1+a_{ij}t)}$，$t > 0$，并令 $M_j = \max_{t>0} f_j(t)$。如下所示，该最大值在唯一一点处取得。我们的主要结果表明，如果某个行动是 M_j 的唯一最小值，那么从均匀策略组合出发的均衡连通分支将收敛至相应的纯协调均衡。

定理1。假设存在唯一行动 k，使得对于每个 $j \neq k$ 都有 $M_k < M_j$。那么 $\mathrm{L}(C_0) = \{(e_k, \ldots, e_k)\}$，在此均衡中，每位参与者都选择行动 k。

证明概要：我们首先将均衡条件表示为每个行动对应的一个标量方程，这些方程具有共同水平 h。每个标量函数都是单峰的。从均匀均衡出发，所有行动都位于各自函数的递增分支上。由于 M_k 是唯一最小的，行动 k 最先达到其峰值。随后，中心分支沿着 k 的递减分支继续，而所有其他行动仍处于各自的递增分支上。当辅助变量 U_k 趋于无穷大时，策略组合收敛至围绕 k 的协调。

证明：我们首先刻画 $\lambda > 0$ 的均衡。令 $U_j = \lambda P_j(x) = \lambda \prod_{i=1}^{n} x_{ij}$。将一阶条件 (3) 乘以 x_{ij}，得 $\mu_i x_{ij} = 1 + a_{ij}U_j$。对 j 求和可得

$$\mu_i = m + \sum_{l=1}^{m} a_{il}\, U_l. \tag{5}$$

利用 $U_j = \lambda \prod_{i=1}^{n} x_{ij}$ 和 $x_{ij} = \frac{1+a_{ij}U_j}{\mu_i}$，可得

$$f_j(U_j) = h \text{ for every } j,\ h = \frac{\lambda}{\prod_{i=1}^{n}\mu_i} \tag{6}$$

反过来，假设正数 U_1、…、U_m 和 h 满足 (6)。定义

$$\mu_i(U) = m + \sum_{l=1}^{m} a_{il}\, U_l,\ x_{ij}(U) = \frac{1+a_{ij}U_j}{\mu_i(U)}, \text{and } \lambda(U) = h \prod_{i=1}^{n} \mu_i\,(U). \tag{7}$$

这些定义意味着 $\sum_{j=1}^{m} x_{ij}(U) = 1$。于是，方程（6）给出 $\lambda(U)\prod_{i=1}^{n} x_{ij}(U) = U_j$，因此一阶条件成立。严格凹性使其成为均衡的充分条件。因此，（6）结合重构公式（7），刻画了 $\lambda > 0$ 的均衡图。

接下来我们证明每个 f_j 都是严格单峰的。对于每个 j，

$$\frac{f_j'(t)}{f_j(t)} = \frac{1}{t}\left(1 - \sum_{i=1}^{n} \frac{a_{ij}t}{1+a_{ij}t}\right). \tag{8}$$

(8) 中的总和严格递增，并从 0 增至 n。因此，存在唯一的 $\tau_j > 0$ 满足

$$\sum_{i=1}^{n} \frac{a_{ij}\tau_j}{1+a_{ij}\tau_j} = 1. \tag{9}$$

由此可知，f_j 在 $(0,\tau_j)$ 上严格递增，在 (τ_j,∞) 上严格递减，并在两个端点处趋于零。特别地，$M_j = f_j(\tau_j)$。对于每个 $\mathrm{h} \in (0,M_j)$，方程 $f_j(t) = \mathrm{h}$ 因而有两个根。将递增分支上的根记为 $t_j^-(\mathrm{h})$，将递减分支上的根记为 $t_j^+(\mathrm{h})$。

接下来，我们识别中心分支。由于 $0 < U_j \le \lambda$，每个充分小的正 λ 对应的均衡都位于各个 f_j 的递增分支上。假设 $M_k < M_j$ 对每个 $j \neq k$ 都成立，这意味着随着 h 增大，行动 k 最先达到其峰值，此时 $\mathrm{h} = M_k$。对于 $j \neq k$，$\mathrm{h} \le M_k < M_j$，因此这些坐标无法达到其峰值。于是，经过均匀均衡的分支可描述为

$$U_j = t_j^-\left(f_k(U_k)\right),\ j \neq k,\ U_k > 0. \tag{10}$$

对于 $U_k < \tau_k$，所有坐标都位于各自的递增分支上；在 $U_k = \tau_k$ 处，k的两个根相遇；而对于 $U_k > \tau_k$，U_k 位于其递减分支上，其余坐标仍位于各自的递增分支上。由于对于每个 $j \neq k$，$f_k(U_k) \le M_k < M_j$，因此式 (10) 中的所有根都有定义。它们随 U_k 连续变化，且没有其他坐标会切换分支。标量表示及其逆映射 $U_j = \lambda P_j(x)$ 在正的 λ 均衡图上连续。因此，式 (10) 及其在 $\lambda = 0$ 处的极限描述了中心分支。

最后，我们研究中心分支的极限。当 $U_k \to \infty$ 时，$f_k(U_k) \to 0$，且式 (10) 意味着对于每个 $j \neq k$，$U_j \to 0$。因此，方程 (5) 给出 $x_{ik} \to 1$ 和 $x_{ij} \to 0$，其中 $j \neq k$。由此，$P_k(x) \to 1$，且 $\lambda = \frac{U_k}{P_k\ x}( ) \to \infty$。反过来，重构公式 (7) 可连续延拓至 $U_k = 0$，其中 $\lambda(0) = 0$，并且在 U_k 的每个有界区间上有界。因此，C_0 上的任意序列若满足 $\lambda \to \infty$，则必有 $U_k \to \infty$。由此可得 $\mathrm{L}(C_0) = \{(e_k, \ldots, e_k)\}$。定理得证。□

当收益指数的最小值唯一时，该定理为追踪型连通分支给出了一个选择准则。如果多个行动共享最小值，则该论证无法确定该连通分支是否具有唯一极限。

对于两名玩家，该指数的形式更为简单。对于每个行动 j，

$$\frac{1}{M_j} = \min_{t>0} \frac{(1+a_{1j}t)(1+a_{2j}t)}{t} = (\sqrt{a_{1j}} + \sqrt{a_{2j}})^2. \tag{11}$$

最后一个等式由对 $\frac{1}{t} + a_{1j}a_{2j}t$ 进行最小化得出。因此，对 M_j 进行最小化等价于最大化收益平方根之和。

推论1. 对于两名参与者和任意有限数量的行动，若某一行动唯一地最大化 $\sum a_{1j} + \sum a_{2j}$，则从均匀策略组合出发的分支会收敛到该行动上的纯协调均衡。

这将 Zhang and Hofbauer（2016年）的双行动平方根选择规则推广到了更多行动。值得注意的是，该规则通常不能推广到三名或更多玩家。例如，设行动 A 和 B 处的收益向量分别为 $a = (16, 1, 1)$，$b = \left(\frac{441}{100}, \frac{441}{100}, \frac{441}{100}\right)$。收益平方根之和更倾向于 B，因为 $\sum_{i=1}^{3} \overline{b_i} = 6.3 > 6 = \sum_{i=1}^{3} \overline{a_i}$。然而，指标 M_A 和 M_B 则更倾向于 A。对于 A，$\frac{1}{M_A} = \min_{t>0} \left(\frac{1}{t} + 18 + 33t + 16t^2\right)$，其中唯一的最小值点是 $\tau_A = \frac{\sqrt{129}-1}{64}$。代入后得到 $M_A \approx 0.03340$。对于 B，三个收益相等，直接最大化得到 $M_B \approx 0.03359$。由于 $M_A < M_B$，定理1选择 A。这个例子表明，双人平方根规则通常不能刻画三名参与者情形下的选择结果。

4. 讨论

本文为对数正则化下的非对称纯协调博弈提出了一项选择准则。均衡图从均匀策略组合出发进行追踪，此时物质收益的权重为零。当某一行动唯一地使指标 M_j 最小化时，连通分支 C_0 收敛于对该行动的协调。对于两名玩家，使 M_j 最小化等价于使收益平方根之和最大化。对于三名或更多玩家，这种简化通常不成立；在唯一最小化条件下，决定选择的是 M_j，而非收益平方根之和。

指标 M_j 与若干已有的均衡选择准则相关，但又有所不同。在非对称纯协调博弈中，帕累托支配其他所有协调结果的结果，其指标必然严格小于 M_j。对于Zhang（2016）中的多数占优结果，同样有这一推论；该判据适用于一大类量化响应函数。在非对称纯协调博弈中，多数支配等价于严格帕累托支配，而这又意味着 M_j 在 j 处唯一取得最小值。然而，与此相反，该判据仍不同于

总收益最大化和风险占优。在双人双行动情形中，平方根排序可能不同于按总收益进行的排序，也可能不同于按偏离损失乘积进行的排序；后者在这一情形下刻画了风险占优（Zhang andHofbauer 2016）。

我们的结果依赖于博弈的结构和对数正则化。具体而言，证明中使用的标量表示利用了这样一个事实：只有协调的行动组合才能产生正收益。因此，本分析无法为一般协调博弈或其他量化响应函数（例如 Logit响应Zhang and Hofbauer 2016）得出类似的规则。主定理还要求 M_j存在唯一的极小值点。当多个行动都达到最小值时，我们的论证无法说明 C_0 是否存在唯一极限。

有若干值得进一步研究的拓展方向。其一是考察平方根排序能否为具有两个以上行动、且采用更一般量化响应函数的双人博弈提供有用的一阶近似，并确定局部近似能够决定全局均衡分支的条件（Zhang and Hofbauer, 2016）。其二是刻画多行动协调博弈中 Logit 及其他量化响应对应的极限均衡（McKelvey and Palfrey, 1995; Turocy,2005; Zhang, 2016）。最后，分析收益变换或加入劣策略如何改变均衡选择（Zhang, 2016），并将对数追踪与其他选择方法进行比较，也会很有意义，包括线性追踪程序（Harsanyi, 1975； Cao and Dang, 2022）、同伦方法（Turocy, 2005)以及演化动力学（Zhang and Hofbauer, 2015）。

致谢：Yali Dong 感谢国家自然科学基金委员会（项目编号：72573024）的资助。Boyu Zhang 感谢国家自然科学基金委员会（项目编号：72131003 和 72671038）以及北京市自然科学基金（项目编号：Z220001）的资助。作者感谢使用 ChatGPT 协助润色语言。所有内容均已在作者的监督和控制下完成审阅与编辑。未使用人工智能工具进行思想生成或研究成果开发。

利益冲突：作者声明不存在任何竞争性利益。

作者贡献：概念化，B.Z.；方法论，B.Z.；形式分析，Y.D.、B.Z.；撰写初稿，Y.D.、B.Z.；审阅与编辑，Y.D.、B.Z.；经费获取，Y.D.、B.Z。

参考文献：

[1] Cao, Y., Dang, C, 2022. 运用 Harsanyi追踪程序在标准式博弈中选择完美均衡的一种变体。《博弈与经济行为》134, 127-150.

[2] Harsanyi, J. C., 1973. 均衡点数量的奇数性：一种新的证明。《国际博弈论杂志》2, 235-250.

[3] Harsanyi, J.C., 1975. 追踪程序：一种为 N人非合作博弈定义解的贝叶斯方法。《国际博弈论杂志》4, 61-94.

[4] Harsanyi, J.C., Selten, R., 1988. 《博弈中的均衡选择一般理论》。麻省理工学院出版社，剑桥。

[5] McKelvey, R.D., Palfrey, T.A., 1995. 标准式博弈的量化反应均衡。《博弈与经济行为》10, 6-38.

[6] Turocy, T.L., 2005年. 逻辑斯蒂量化反应均衡对应的动态同伦解释. 《博弈与经济行为》51, 243-263.

[7] Zhang, B., 2016年. 标准式博弈中均衡选择的量化反应方法. 《数学经济学杂志》64, 113-123.

[8] Zhang, B., Hofbauer, J., 2015. 通过复制者动态在2×2协调博弈中进行均衡选择. 《国际博弈论杂志》44, 433-448.

[9] Zhang, B., Hofbauer, J., 2016. 2×2协调博弈中均衡选择的量化反应方法. 《博弈与经济行为》97, 19-31.