第 22 卷第 12 期(总第 132 期) 系 统 工 程 Vol 22, No. 12
2004 年 12 月 Systems Engineering Dec, 2004

文章编号: 1001-4098(2004)12-0006-04

个体与群体之间的一类博弈问题分析*

刘德海[1,2], **徐寅峰**[1], **李纯青**[2]

(1. 西安交通大学 管理学院, 西安交通大学 机械制造系统工程国家重点实验室, 陕西 西安 710049;
2. 西安工业学院 经济管理学院, 陕西 西安 710049)

摘 要: 分析个体与群体之间的一类博弈问题, 构造一对多博弈模型。其同时具有 Stackelberg 模型、演化博弈模仿者动态模型的特点: 在博弈的第一阶段, 有限理性的群体参与者采取模仿者动态行为; 在博弈的第二阶段, 完全理性的个体参与者根据群体参与者的策略分布, 确定自己的反应函数。博弈达到稳定状态后, 群体参与者采取演化稳定策略, 而个体参与者的风险占优策略将收敛于 Nash 均衡策略。讨论模型的一些性质: 有限理性的群体在模仿过程中存在收益损失; 老练的个体参与者存在欺骗行为; 在个体参与者完全理性决策的带引下, 可以解决多重 Nash 均衡的选择问题。

关键词: 一对多博弈模型; Stackelberg 模型; 模仿者动态模型; 演化博弈

中图分类号: O225 **文献标识码**: A

传统博弈理论的理性要求是“完全理性的共同知识”[1], 均衡是在博弈规则、参与者的理性和收益函数等都是共同知识情况下, 通过参与者的分析推理过程得出的结果。其分析对象是相互作用的个体。如果参与者是某类群体, 那么应是同质的, 即群体中参与者的理性特征、采取的策略是完全相同的, 因此可以简化为个体进行分析。传统博弈论脱离现实的强理性假设正日益受到越来越多的博弈理论家和经济学家的批评[1]。20 世纪 90 年代兴起的博弈学习理论认为, “均衡是作为小于完全理性的参与者寻求最优化的长期过程的结果”[2]。其中, 演化博弈的分析对象是有限理性的群体行为, 根据不同的动态模仿过程, 参与者通过模仿具有更高收益的策略, 最终采取各种策略的群体比例达到均衡状态, 例如演化稳定策略(ESS, Evolutionary Stable Strategy)[2,3]等均衡概念。

社会经济中存在着大量的个体参与者与群体参与者之间的博弈问题, 例如我国目前的体制改革中, 政策制订者与广大公众之间的博弈; 农村劳动力转移过程中政府部门与农村劳动力群体之间的博弈; 企业主与员工之间的博弈等。这些博弈问题的特点是个体参与者往往具有较多的信息(包括对各种策略的收益、博弈的结构等), 从而能够做出理性的决策; 而在群体参与者内部由于信息拥有程度、分析计算能力不同, 因而其行为方式符合演化博弈理论中惯性、短视等有限理性的假设。

1 一对多博弈模型

一对多博弈模型中参与者的理性假设类似于 Stackelberg 模型的假设(双方分别为“老练的”和“幼稚的”)[4], 在博弈的第一阶段, 具有有限理性的、“幼稚的”群体选取不同的纯策略形成各种纯策略的比例分布; 在第二阶段, 具有完全理性的、“老练的”个体参与者根据对方(群体参与者)形成的策略分布确定自己的反应函数。不同于 Stackelberg 模型, 一对多博弈模型中参与者为个体与群体之间的博弈。在博弈的第一阶段中, “幼稚的”群体参与者的行为方式可以用演化博弈中模仿者动态模型加以描述。

博弈的第一阶段: 根据演化博弈理论的基本假设, 由于转换新策略存在一定的成本, 群体中大部分参与者保持某种纯策略不变, 少部分参与者根据模仿行为采取具有更高收益的策略。在群体参与者的纯策略空间 $s(s_1, \cdots, s_i, \cdots, s_n)$ 中, 设 $\theta_t(s_i)$ 为在 t 阶段采用纯策略 s_i 的参与者群体比例向量, 采用某种纯策略 $s_i(i=1, 2, \cdots, n)$ 的群体比例增长率 $\dot{\theta}_t(s_i)$ 是此纯策略效用 $u_t(s_i)$ 与群体平均期望效用 $\overline{u_t}$

---

* **收稿日期**: 2004-07-31
基金项目: 国家自然科学基金委员会优秀创新群体项目(70121001); 国家自然科学基金资助项目(10371094)
作者简介: 刘德海(1974-)男, 辽宁辽阳人, 西安交通大学管理学院博士研究生, 研究方向: 农村劳动力转移问题, 演化博弈; 徐寅峰(1962-), 男, 吉林东丰人, 西安交通大学管理学院教授, 博士生导师, 研究方向: 组合优化, 博弈论; 李纯青(1970-)女, 河南社旗人, 西安工业学院经济管理学院副教授, 研究方向: 组合优化, 博弈论。

差的严格增函数[5]:

$$\dot{\theta}_t(s_i) = \theta_t(s_i)[u_t(s_i) - \overline{u_t}] \tag{1}$$

博弈的第二阶段: 在个体参与者的纯策略空间 $S'(s_1', \cdots, s_i', \cdots, s_n')$ 中,“老练的”个体参与者根据对方采取不同策略 $s_i(i=1, 2, \cdots, n)$ 的比例分布 $\theta_t(s)$, 确定自己的反应函数 $u_t(s') = \sum_s \theta_t(s) \cdot u(s', s)$。

在具有多重 Nash 均衡的重复博弈中, 个体参与者的收益如果满足:

$$u_t(s_k', s) \geq u_t(s', s), \quad k \in (1, 2, \cdots, m) \tag{2}$$

那么对于追求当期收益最大化的“短视的”个体参与者, 其将采取风险占优策略 s_k'.

对于具有“远见的”、追求长期利益最大化的个体参与者, 贴现因子记为 δ, 如果其收益满足:

$$\sum_t \delta \cdot u_t(s_l') > \sum_t \delta \cdot u_t(s_k'), \quad l \in (1, 2, \cdots, m) \tag{3}$$

那么个体参与者采取具有 Pareto 最优的 Nash 均衡策略实现的收益 s_l', 大于采取风险占优策略 s_k' 实现的收益。

在博弈的收益结构 $u(s', s)$、重复博弈的次数 t 和群体参与者的策略比例分布 $\theta_t(s)$ 既定的条件下, (3) 式主要取决于贴现因子 δ 的大小, 其意义为个体参与者是否具有“远见的”, 比较注重未来阶段的收益。举例说明如下。

图 1 中的博弈具有 2 个 Nash 均衡 (1, 1)、(2, 2), 其中 (2, 2) 是 Pareto 最优的。当群体参与者在不同策略之间具有一定的比例分布 $(\theta(s_A), 1-\theta(s_A))$ 时, 个体参与者采取“N”策略的收益为 $u(s_N) = 2 - \theta(s_A)$, 采取“M”策略的收益为 $u(s_M) = 2 - 2\theta(s_A)$, 根据 (2) 式可知, 个体参与者采取“N”策略是风险占优策略, 在重复博弈中双方最终形成的稳定 Nash 均衡为处于 Pareto 劣位的 (1, 1) 均衡, 其中 $\theta(s_A) \to 1$。

在无限次重复博弈中, 个体参与者采取风险占优策略“N”最终实现的收益为 $u(s_N) \to 1$, 而具有“远见的”个体参与者采取具有 Pareto 最优的 Nash 均衡策略“M”最终实现的收益为 $u(s_M) \to 2$, 当 $\theta(s_A) \to 0$。

		个体参与者	
		N	M
群体参与者	A	1, 1	0, 0
	B	0, 2	2, 2

图 1

2　博弈的均衡分析

命题　个体与群体之间重复博弈模型的稳定状态为 Nash 均衡。

证明　根据模型的假设, 群体参与者的行为方式符合演化博弈的模仿者动态模型。根据 Fudenberg (1998) 提出的命题:“模仿者动态的稳定均衡是 Nash 均衡”[2]。当模仿者动态允许群体选择混合策略时, Bomze 进一步证明模仿者动态的一个稳定均衡是演化稳定均衡 ESS[6]。下面分析个体参与者的最终稳定状态:

第一种情况: 如果个体参与者是具有“远见的”, 即满足条件 (3) 式, 其采取的策略 s_l' 为具有 Pareto 最优的 Nash 均衡策略。

第二种情况: 如果个体参与者是“短视的”, 其将采取风险占优策略 s_k'. 当 $t \to \infty$ 时, 群体参与者的模仿者动态趋于稳定的 Nash 均衡 s_i^*, 其策略比例分布 $\theta_t(s) = \begin{cases} 0, s \neq s_i^* \\ 1, s = s_i^* \end{cases}$, 则个体参与者的反应函数 $u_t(s') = \sum_s \theta_t(s) \cdot u(s', s) = u_t(s', s_i^*)$, 将反应函数代入 (2) 式, 此时风险占优策略满足:

$$u_t(s_k', s_i^*) > u_t(s_j', s_i^*) \tag{4}$$

由于 (4) 式中风险占优策略 s_k' 满足严格均衡[7]的定义, 而严格均衡必定是纯策略 Nash 均衡, 因此当群体参与者的模仿者动态行为收敛到 Nash 均衡 s_i^* 时, 个体参与者的风险占优策略也将收敛到纯策略 Nash 均衡。

3　一对多重复博弈模型的性质

(1) 群体参与者在模仿的动态过程中存在着收益损失, 由此将对个体参与者产生额外的收益影响。由于模型假设群体参与者为“幼稚的”, 因此在纯策略空间 $S(s_1, \cdots, s_i, \cdots, s_n)$ 上存在一个比例分布 $\theta_t(s_1), \cdots, \theta_t(s_i), \cdots, \theta_t(s_n)$, 群体参与者在模仿具有更高收益的策略过程中存在着损失。设群体参与者有 p 人, 最终采取的稳定 Nash 均衡策略为 s_i^*, 则无限次重复博弈中群体参与者收益总损失的贴现值 $C = \sum_{t=0}^{\infty} \sum_{i=1}^{n} \delta \cdot [u_t(s_i, s_k') - u_t(s_i^*, s_k')] \cdot \theta_t(s_i) \cdot p$, 其中 s_k' 是“短视的”个体参与者采取的风险占优策略; 如果对方是具有“远见的”, 则上式中风险占优策略 s_k' 替换为 Pareto 最优的 Nash 均衡策略 s_l'.

如果个体参与者是“短视的”, 采取的风险占优策略 (这里假设个体具有贝叶斯理性 Bayesian rationality, 即参与者根据事后信息和贝叶斯法则对主观事前概率进行更新) 可以实现其各阶段的当前收益最大化。但是在具有多重 Nash 均衡的重复博弈中, 如 (3) 式所示, 该策略并不等于个体参与者总收益的最大化。

(2) 在群体参与者模仿最优策略的过程中, 老练的个体参与者通过误导对方可以实现更大的收益, 因此存在着欺骗行为。

图 2 所示的博弈具有唯一的 Nash 均衡 (1, 1), 个体参与者采取风险占优策略“N”后, 博弈收敛于 Nash 均衡。此

方案中, 个体参与者仅能获得部分群体参与者的“B ”策略分布造成的额外收益 $P_k = \sum_t 8\cdot\delta\cdot\theta_t(B)$, 其中 $\theta_t(B)\in(0, 1)$。如果个体参与者先选择“M ”策略, 当群体参与者逐渐收敛于“B ”策略后, 再转换为“N ”策略, 则此方案中, 个体参与者先是通过发布误导的信息, 使博弈暂时收敛于具有 Pareto 最优的收益(4, 7); 然后在转换策略过程中, 最初阶段获得了对方几乎全部参与者停留在“B ”策略造成的额外收益 $P_c = \sum_t 8\cdot\delta\cdot\theta_t(B)$, 其中 $\theta_t(B)\approx 1$。显然, 此方案个体参与者获得的收益大于(或者等于)风险占优策略的收益。

个体参与者

群体参与者	N	M
A	1, 1	0, 0
B	0, 8	4, 7

图 2

(3) 为了避免有限理性的群体参与者在模仿过程中分散行为造成的收益损失, 以及博弈对方老练的个体参与者故意采取的欺骗行为造成更大损失, 分散的群体参与者可以形成组织中的合谋(collusion), 增加自身讨价还价的能力, 从而将一对多博弈模型转化为传统的非合作博弈模型。Downs 和 Olson 的利益集团理论指出, 一个利益集团的力量取决于它的酬金(stake)和组织成本[8, 9]。对于本模型的群体参与者, 形成合谋组织后群体可以获得的最大酬金即为各个阶段群体参与者收益总损失的贴现值 C. 但是, 群体参与者形成合谋组织后, 完全理性的两个博弈方又出现了传统博弈理论中经典的多重 Nash 均衡选择问题。

(4) 在“有远见的”个体参与者锁定 Pareto 最优的 Nash 均衡策略下, 博弈模型可以解决传统博弈的多重 Nash 均衡选择问题。

非合作博弈论的基础问题之一是多重 Nash 均衡的选择问题[1]。为此, Schelling(1960) 提出“聚点”理论[10]; Aumann(1974) 提出了“相关均衡”概念[11]。Binmore (1990) 指出, 改进博弈论必须模拟参与者的思考方式, 更多地注意均衡过程而不是均衡点[1]。博弈学习理论将侧重点转向“过程理性”(procedural rationality), 但是演化博弈的 ESS 均衡仅是对 Nash 均衡的精练[12], 其并没有解决多重 Nash 均衡的选择问题。在传统博弈“目的理性”的理论框架内, 通过参与者自身的内省式推理演绎过程是无法解决多重 Nash 均衡选择问题的。从博弈进行的历史过程上看, 这一问题是通过参与者的“过程理性”导致特定的路径依赖, 产生“锁定”现象而自然解决的。演化博弈理论虽然提供了历史发展产生路径依赖的演化观念, 但是其所以没有解决多重 Nash 均衡的选择问题, 原因在于博弈双方对称的有限理性假设, 与传统博弈理论的完全理性假设一样, 都使得参与者在面临多重 Nash 均衡时无法做出协调一致的反应(除非设计某种信号装置 Signaling device)。只有打破博弈双方理性水平的对称性假设, 在完全理性的博弈方通过锁定具有 Pareto 最优的 Nash 均衡策略带引下, 有限理性的对方在重复博弈不断试错的过程中最终收敛于具有 Pareto 最优的 Nash 均衡, 双方才有可能在历史的发展过程中走出多重 Nash 均衡选择的困境。

从更深层次上看, 解决传统博弈理论多重 Nash 均衡选择问题的思路, 完全理性的“相关均衡”概念是沿着人为设计的机制设计理论选择方法; 而有限理性的 ESS 均衡则沿着自发演化的演化博弈理论的选择(精练)方法。在制度发展的历史进程中, 人为设计和自发演化两者不是相互割裂的, 而应该是相互作用的。本文分析的群体与个体之间一对多重复博弈模型正是揭示了制度演化的历史特点, 从而走出多重 Nash 均衡选择的困境。

4 结论

本文介绍了一类介于演化博弈和传统博弈之间的博弈问题, 即个体与群体之间的一对多博弈模型, 其同时具有了 Stackelberg 模型、演化博弈模仿者动态模型的特点。文章证明了, 当博弈达到稳定状态后, 个体参与者的风险占优策略将收敛于 Nash 均衡。该模型具有如下一些性质: 有限理性的群体在模仿过程中存在收益损失; 老练的个体参与者存在欺骗行为; 在个体参与者完全理性决策的带引下, 可以实现多重 Nash 均衡的精炼。这些性质具有很强的现实意义。在我国当前的体制改革和社会转型的历史时期, 各级政府部门拥有各种政治、经济资源, 处于政策制订者的有利位置; 而广大社会公众, 尤其是农民群体, 由于缺少合适的组织表达途径, 在维护自身利益时处于分散的群体状态。如何避免农民群体的盲动行为、以及地方政府实施“政绩工程”等官员自利行为带来的损失? 在现代化多种途径的历史选择中, 政府部门如何制定政策引导社会公众的行为, 使得社会整体尽快步入现代化的帕累托最优路径? 这些问题需要从理论探讨和实践总结上进一步加以深入研究。

参考文献:

[1] Binmore K. Foundation of game theory[A]. Laffont J. Advances in Economic Theory, Sixth World Congress(Vol 1)[C]. Cambridge University Press, 1992: 1～31.

[2] Fudenberg D, Levine D K. Theory of Learning in Games[Z]. 1998.

[3] Smith J M. Evolution and the theory of games[M]. Cambridge University Press, 1982.

[4] Basu K. Stackelberg equilibrium in oligopoly: an explanation based on managerial incentives[J]. Economics Letters, 1996, 49: 459～464.

[5] Taylor P D, Jonker L. Evolutionarily stable strategies and game dynamics[J]. Math. Biosci, 1978, 40: 145～156.

[6] Bomze I. Non-cooperative two-person games in biology: a classification[J]. International Journal of Game Theory, 1986, 15: 31～57.

[7] Harsanyi J. Oddness of the number of equilibrium points: a new proof[J]. International Journal of Game Theory, 1973, 2: 235～250.

[8] Downs A. Inside bureaucracy[M]. Boston: Little, Brown and Co., 1964.

[9] Olson M. The logic of collective action: public goods and the theory of collective action[M]. MA: Harvard University Press, 1965.

[10] Schelling T. The strategy of conflicts[M]. Harvard University Press, 1960.

[11] Aumann R J. Subjectivity and correlation in randomized strategies[J]. Journal of Mathematical Economics, 1974, 1: 67～96.

[12] Van Damme E. Refinement of Nash equilibrium [A]. Laffont J. Advances in Economic Theory, Sixth World Congress (Vol. 1) [C]. Cambridge University Press, 1992: 32～75.

An Analysis of a Game Problem Analyze between Individual and Group

LIU De-HAI[1,2], XU Yin-feng[1], LI Chun-qing[2]

(1. School of Management, Xi'an Jiaotong University, Xi'an 710049, China;

2. School of Economics and Management, Xi'an College of Industry, Xi'an 710049, China)

Abstract: This paper analyses the game model between individual and colony that has the characteristics of Stackelberg model and replicator dynamic model in evolutionary game. In the game first phase, the bounded rationality colony players adopt the replicator dynamic behavior. In the second phase, the full rationality individual player decides the own response function by the strategies distribution of colony players. This model has some characteristics as follows. The bounded rationality colony players have the payoff losing in the imitating course. The individual player could take the cheat behavior. The colony players can build the form of collusion avoiding the payoff losing. If the game has multiple Nash equilibriums, the foresighted individual player will take the Nash equilibrium strategy that is the Pareto-optimality. In this case the game could reach the multiple Nash equilibriums select.

Key words: The Game Model between Individual and Group; Stackelberg Model; Replicator Dynamic Model; Evolutionary Game