进化博弈论及模型应用

马欣悦　　1653515　　生命科学与技术学院

摘要

博弈论源于经济学家对经济行为的分析，但其精髓在于向着利益最大化做出最佳策略，这一点与达尔文的进化理论相仿，事实上，博弈论不但能很好的运用于进化生物学，解释生物行为，在应用过程中，生物进化也为传统博弈论提供了新思路。本文将介绍进化博弈论的基本概念，并分别介绍这一理论在生物进化及社会科学中的应用模型。

关键词: 进化, 博弈论, 进化博弈论, 进化稳态

Abstract

Game theory was used to explain ecological behavior, the key to which is choosing the best strategy for the greatest benefit. The main idea of game theory is surprisingly compatible with the Darwin's evolutionary theory. Virtually, game theory can explain the biological behavior during evolution as well. Moreover, the application on evolution also provides new-idea for classical game theory. This article will introduce fundamental concept of evolutionary game theory and generate two models based on evolutionary and social issue.

Keywords: evolution, game theory, evolutionary game theory, evolutionary stable state

1　前言

“进化博弈论“是将数学模型应用于生物学的典型例子之一，且由此启发的意义可以广泛延伸到生物进化学之外的学科，如经济学、社会学等，其原因在于 [van Benthem and Klein (2019)]：（1）“进化”这个词本身并非仅局限于生命群体，这个概念本身是可以向外引申的，到社会进化，如信仰、文化等；或者到经济学，如市场竞争、垄断等；（2）进化博弈论着眼于群体而非个体，策略方法为混合策略（mixed strategy），这补充了传统博弈论的“纯策略”（pure strategy）中认为“每一个个体都是理性的”的缺陷，因为现实生活中确实很难保证每一个个体的行为都是出于理智的；（3）进化博弈过程是一个动态变化过程，这继承于生物进化“传代”的特征，上一代的博弈结果（payoff）会作为下一代的博弈策略（strategy）。

2　发展历史

进化博弈论的概念由 R. A. Fisher[Fisher (1930)] 首次提出，在书中他尝试从博弈论的角度分析动物群体的性别比例，为什么大多数哺乳动物的性别比基本稳定在 1：1，在这些群体中，具有交配权的雄性只占全部雄性个体的一小部分，大部分不参与繁衍的雄性个体似乎对群体无甚贡献。Fisher 给出的解释是，任意一种性别的个体数量超过另一种性别的个体数时，其本身的适应性（fitness）会下降，而另一种的适应性会上升，这种适应性的不均等最终会推动性别比回到 1：1。这个解释虽然没有说明“适应性”指的是什么，而平衡比例又为什么是 1：1，但他指出了维稳的过程及原因。Fisher 的论述反映了进化博弈论的早期思想，但没有系统的将博弈论与进化生物学的概念相结合。1961 年 R. C. Lewontin 在著作“Evolution and the Theory of Games”中第一次提出博弈论在生物进化的应用 [Lewontin (1961)]；1972 年，Maynard Smith 在“Game Theory and the Evolution of Fighting.”提出了进化稳定策略（evolutionarily stable strategy，简称 ESS）的概念；1973 年 Maynard Smith 和 Price 出版“The Logic of Animal Conflict”一书 [Smith and Price (1973)]，其中包括经典的“Hawk-Dove”模型。随着该书的出版，ESS 的概念开始普及，博弈论开始广泛应用于描述及解释动物行为 [Dugatkin and Reeve (2000)]。1982 年 Maynard Smith 的“Evolution and the Theory of Games”[Smith (1982)] 及 1984 年 Robert Axelrod 的“The Evolution of Cooperation”[Suleiman and Fischer (1996)] 出版，进化博弈论这个概念开始受到经济学家及社会学家的关注，并渗透到其他学科当中。

3　基本概念

3.1　定义

博弈有三要素：玩家（player），策略（strategy），效益（payoff）。在进化博弈中，生物个体即是玩家，个体的表型即是策略，个体存活率及繁衍能力（fitness）即是效益。博弈过程中，

个体的 fitness 由个体及对手的表型、及环境共同决定。类似传统博弈论，以上的规则可以以二联表的形式展示，种群状态及博弈过程可以用数学语言建立模型，这些概念模型可以辅助我们理解生物行为及进化过程，对选择结果及进化方向做出预测，比如花期、生产期、最适栖息地等 [Day and Taylor (2003)]。
在纯策略博弈中，如果 A 策略进化稳定，则群体中出现少量 B 时，一段时间后，B 会被逐渐淘汰，群体重新回到 A；在混合策略博弈中，如果 A 策略组合进化稳定，则当任意突变体导致策略组合偏离 A，一段时间后，策略组合会重新回到 A；我们把这个稳态称为进化稳定状态（evolutionary stable state, ESS）。假定生物体的策略为 z，个体的总效益表示为 $W(z,\overline{z})$。

$$W(z_i,\overline{z}) = p_1W(z_i,z_1) + p_2W(z_i,z_2) + ... + p_nW(z_i,z_n)$$

其中，$p_1 + p_2 + ... + p_n = 1$。在 ESS 状态下，有

$$W(z,\overline{z} < W(z^*,\overline{z})$$

称 z^* 为进化稳定策略（evolutionary stable state）。达到稳定状态后，倘若出现任意 z^*，z 的分布频率为 $\varepsilon > 0$，且满足 $\varepsilon \to 0$，这时 $\overline{z} = \varepsilon z + (1-\varepsilon)z^*$，满足

$$W(z,z^*) \leq W(z^*,z^*),\overline{z})$$

以上便是 ESS 的数学表示，其中 2 又被称为 evolutionary stability，3 被称为 convergence stability[Christiansen (1991)]。

3.2 应用及意义

纳什平衡是博弈论经典且重要的概念之一，当两个玩家的策略处于纳什平衡时，基于对方玩家的策略，没有其他策略可以让玩家得到更多的效益。例如经典的“囚徒困境”：

表 1: Prisoners’ Dilemma payoff matrix

AB	silent	betray
silent	(-1,-1)	(-3,0)
betray	(0,-3)	(-2,-2)

任意一方选择了“silent”，另一方的最佳选择（best-response）都会是“betray”，而最终（betray, betray）成为这个博弈的纳什平衡，尽管总体上来看，这似乎并不是一个“明智”的结果。
但是对于以下这种情况：
从纯策略的角度来说，四种策略都不是纳什平衡。例如在（Head, Head）时，玩家 2 知道玩家

表 2: Matching Pennies payoff matrix

A \ B	Head	Tail
Head	(0,1)	(1,0)
Tail	(1,0)	(0,1)

1 的选择是“Head“，为达到最大效益玩家 2 的策略会从“Head”变成“Tail”；当玩家 2 变成“Tail”后，玩家 1 为达到最大效益，策略会从“head”变成“Tail”；当玩家 1 变成“Tail”时，玩家 2 的最佳选择又从“Tail”变成了“Head”……如此循环往复。这个情况其实并不少见，比如“剪刀、石头、布”。

而通过进化博弈中的混合策略思想，我们就能提出达到纳什平衡的策略。因为混合策略模型着眼于群体或者是多次实验，给出的策略不是（A，B）形式，而是 p 频率采用 Head，1-p 频率采用 Tail，其中 $0 \leq q \leq 1$。由于玩家策略对称，AB 采取的混合策略相同，有

$$\begin{cases} W_A = p[p \times 0 + (1-p) \times 1] + (1-p) \times [p \times 1 + (1-p) \times 0] \\ W_B = p[p \times 1 + (1-p) \times 0] + (1-p) \times [p \times 0 + (1-p) \times 1] \end{cases}$$

化简求得 $p_{W_{max}} = 0.5$,A,B 的最大效益分别为

$$\begin{cases} W_{A(max)} = 0.5 \\ W_{B(max)} = 0.5 \end{cases}$$

混合策略除了解决这种纯策略无法分析的博弈情况，还补充了另一种情况，及并非所有玩家都是理智的。这在生物进化过程中液由很好的体现，例如相对劣势的性状面临的并不是在第二代就完全消失，而是逐渐小减少，甚至能够与相对优势性状维持一定比例的稳定分布。“相对”这一概念指出，在现实中影响博弈玩家的策略选择不不单单是游戏规则中明确指出的效益，还有很多其他因素。同时进化博弈的动态模型，例如“replicator dynamics”[Taylor and Jonker (1978)]，上一代的博弈结果导致下一代的策略分布。这些都为传统博弈论做了补充。

得益于进化博弈论的补充，大量社会行为的分析都采用了这样一种博弈思想，例如利他性 [Fletcher and Zwick (2007)]，共情能力 [Page and Nowak (2002)]，社会学习 [Kameda and Nakanishi (2003)]，文化的起源与传播 [Enquist et al. (2007)]。接下来将各介绍一种模型，利用进化博弈论解决生物进化及社会行为问题。

4 基本模型

4.1 Hawk-Dove 模型

Hawk-Dove 模型是 Maynard Smith 和 Price 所著 The Logic of Animal Conflict[Smith and Price (1973)] 的经典例子。下面介绍此类模型的分析方法。

4.1.1 假设及定义

1. 策略（strategy）
Hawk：侵略型玩家，直到受伤或者打败对手才停止战斗；
Dove：保守型玩家，遇到危险时立刻撤退；

2. 效益（payoff）

表 3: Hawk and Dove payoff matrix

A \ B	Hawk	Dove
Hawk	$\frac{V-C}{2}$	V
Dove	0	$\frac{V}{2}$

上表中，V 为两个玩家相遇时争夺的资源，C 为两个玩家势均力敌的斗争造成的资源损失。如上表，一只 Dove 另一只 Dove，二者和平均分 V 的资源；而当 Dove 面临 Hawk，Hawk 的侵略性会让 Dove 选择逃跑，因此鸽子效益为 0 而 Hawk 为 V；当 Hawk 与 Hawk 针锋相对，战斗最终会使双方都损失 $\frac{1}{2}C$ 的效益。

4.1.2 纯策略 ESS 分析

$s \in (hawk, dove)$。$w(s_1, s_2)$ 为 s_1 个体遇到 s_2 时的效益，W(s) 为 s 策略的玩家的总效益，W_0 为所有玩家初始效益。δ 为 ESS 下的策略，μ 为 ESS 后出现的突变。δ 的分布频率为 p, 且满足 $(1-p) \to 0$。有如下关系式

$$\begin{cases} W(\delta) = W_0 + (1-p) \triangle W(\delta,\delta) + p \triangle W(\delta,\mu) \\ W(\mu) = W_0 + (1-p) \triangle F(\mu,\delta) + p \triangle W(\mu,\mu) \end{cases}$$

当 δ 为 ESS 的策略时，满足 $W(\delta) > W(\mu)$，由于 $(1-p) \to 0$，因此只有在

$$\triangle W(\delta,\delta) > \triangle W(\mu,\delta)$$

或者

$$\begin{cases} \triangle W(\delta,\delta) = \triangle W(\mu,\delta) \\ \triangle W(\delta,\mu) = \triangle W(\mu,\mu) \end{cases}$$

时成立。

有以上结果可知，(Hawk,Dove) 的效益关系符合（δ,μ）的效益关系，因此从纯策略的角度分析，Hawk 策略严格由于 Dove。

4.1.3 混合策略 ESS 分析

群体中（Hawk，Dove）的分布比为 (p,1-p)。根据效益表分别计算 Hawk 和 Dove 的效益

$$\begin{cases} W_{dove} = (1-p)[p \times 0 + (1-p) \times \dfrac{V}{2}] \\ W_{hawk} = p[p \times \dfrac{V-C}{2} + (1-p) \times V] \end{cases}$$

当且仅当 $p = \frac{1}{2} = \frac{V}{V+C}$ 时，Dove 和 Hawk 的效益均达到最大。

4.2 Divide-the-cake 模型

“Divide-the-cake”模型用于解释社会行为中的“公平”, 这个模型 1996 年由 Skyrms 在 Evolution of the Social Contract[Skyrms (1996)] 中提出。书中描述的游戏规则为：

“现在有一个意外得到的蛋糕，要求是将蛋糕分给所有玩家，已知所有玩家的需求完全相同。如果所有人都同意分蛋糕的方法，每个人都能得到划分到的份额；但只要有人不满划分的方法，没有人能得到蛋糕。”

现在将这个模型简化。由 A，B 两位玩家分享体积为 C 的蛋糕。D_A,D_B 为分配方法，W_A, W_B 表示两个玩家分到的蛋糕量。已知 $0 \leq W_A, W_B \leq C$。游戏规则为表示如下：

$$\begin{cases} W_A = D_A; W_B = D_B & 0 \leq (W_A + W_B) \leq C \\ W_A = W_B = 0 & (W_A + W_B) > C \end{cases}$$

凭直觉，我们会觉得 $(\frac{C}{2}, \frac{C}{2})$ 就是最佳策略，但是如何解释这种“公平”分配呢? 当玩家是依次做选择时，已知玩家 1 选择 p($0 \leq p \leq C$)，玩家 2 的最佳选择是 C-p。该博弈的纳什平衡策略有无穷种，为 $(p, C-p)(p \in [0, C])$。

而更多时候，玩家不清楚自己是那个先做选择的还是后做选择的，这时倘若玩家凭经验推测自己有 q 的可能先做选择，（1-q）的可能后做选择，玩家的期望效益为：

$$p \times q + (C - p) \times (1 - q) \qquad p \in [0, C], q \in [0, 1]$$

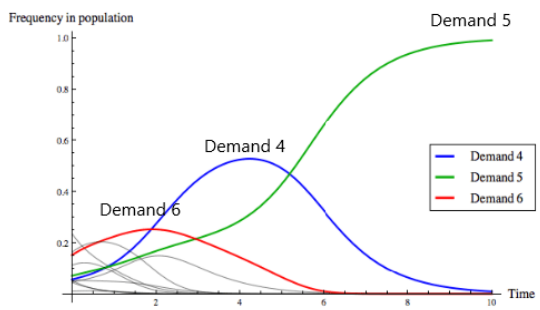
回到进化博弈的思想。我们假设“Divide-the-cake”实验进行不止一次，且在群体中两两配对进行，下一次的策略取决于上一次的结果（即上面谈到的 replicator dynamics）。为方便分析，我们将蛋糕等分为 10 份，第 n 代的策略表示为 $< p_0, p_1, ..p_{10} >_n$，其中 p_i 表示这一代中选择 i 份蛋糕的频率。一代博弈后，每一种选择的效益表示为 $W(p_i)$，值为：

$$W(p_i) = \sum_{j=0}^{10-i} p_j \times i$$

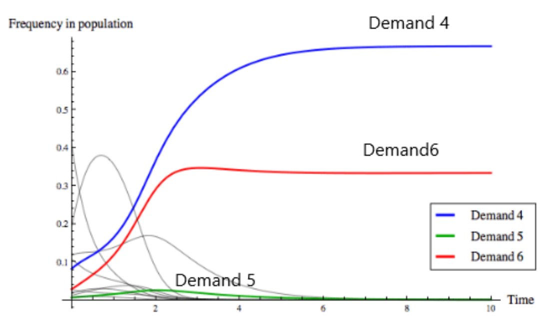
下一代策略则为：

$$p_i^{'} = \frac{W(p_i)}{\sum_{j=0}^{10} W(p_j)}$$

因此，只要给出第一代的策略 $< p_0, p_1, ..p_{10} >_{n=1}$，我们就可以通过计算，并画图看到策略随代数的变化。如下图：



(a) The evolution of fair division.



(b) The evolution of an unequal division rule.

其中，$< p_0, p_1, ..p_{10} >_{n=1,a} = < 0.0544685, 0.236312, 0.0560727, 0.0469244, 0.0562243, 0.0703294, 0.151136, 0.162231, 0.0098273, 0.111366, 0.0451093 >$,

$< p_0, p_1, ..p_{10} >_{n=1,b} = < 0.410376, 0.107375, 0.0253916, 0.116684, 0.0813494, 0.00573677,$
$0.0277155, 0.0112791, 0.0163166, 0.191699, 0.00607705 >$

有意思的是，并非所有结果都如预期想的 p_5 的频率会最大甚至逼近于 0，由进化博弈论思想分析得到的结果可知，最优策略并不唯一，会因为初代的分布不同，使得群体分布最终向不同的方向发展。

5 总结

以上是对进化博弈论及模型构建的介绍。通过数学公式解释进化过程，并将数据可视化使得进化过程被清晰展示。进化博弈轮的发展，从引入博弈论的思想辅助进化生物学的研究，到生物进化的规律补充博弈论的不足，两者相得益彰。甚至 Maynard Smith 在所著的 Evolution and the Theory of Games 前言中这样写道 [Smith (1982)]:

“[p]aradoxically, it has turned out that game theory is more readily applied to biology than to the field of economic behaviour for which it was originally designed。”

参考文献

Christiansen, F. B. (1991). On conditions for evolutionary stability for a continuously varying character. *The American Naturalist*, 138(1):37–50.

Day, T. and Taylor, P. D. (2003). Evolutionary dynamics and stability in discrete and continuous games. *Evolutionary Ecology Research*, 5(4):605–613.

Dugatkin, L. A. and Reeve, H. K. (2000). *Game theory and animal behavior*. Oxford University Press on Demand.

Enquist, M., Eriksson, K., and Ghirlanda, S. (2007). Critical social learning: a solution to rogers's paradox of nonadaptive culture. *American Anthropologist*, 109(4):727–734.

Fisher, R. (1930). The genetical theory of natural selection. oxford: Clarendon press. 272 p.

Fletcher, J. A. and Zwick, M. (2007). The evolution of altruism: game theory in multilevel selection and inclusive fitness. *Journal of theoretical biology*, 245(1):26–36.

Kameda, T. and Nakanishi, D. (2003). Does social/cultural learning increase human adaptability?: Rogers's question revisited. *Evolution and Human Behavior*, 24(4):242–260.

Lewontin, R. C. (1961). Evolution and the theory of games. *Journal of theoretical biology*, 1(3):382–403.

Page, K. M. and Nowak, M. A. (2002). Empathy leads to fairness. *Bulletin of mathematical biology*, 64(6):1101–1116.

Skyrms, B. (1996). *Evolution of the social contract*. Cambridge University Press.

Smith, J. M. (1982). *Evolution and the Theory of Games*. Cambridge university press.

Smith, J. M. and Price, G. R. (1973). The logic of animal conflict. *Nature*, 246(5427):15.

Suleiman, R. and Fischer, I. (1996). The evolution of cooperation in a simulated inter-group conflict. In *Frontiers in Social Dilemmas Research*, pages 419–438. Springer.

Taylor, P. D. and Jonker, L. B. (1978). Evolutionary stable strategies and game dynamics. *Mathematical biosciences*, 40(1-2):145–156.

van Benthem, J. and Klein, D. (2019). Logics for analyzing games.