

控制与决策

Control and Decision

鲁棒博弈动力学：分析与综合

王龙, 王国丞, 苏奇

引用本文:

王龙, 王国丞, 苏奇. 鲁棒博弈动力学：分析与综合[J]. *控制与决策*, 2025, 40(6): 1761-1773.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2024.1292>

您可能感兴趣的其他文章

Articles you may be interested in

考虑附加情绪的两阶段投资组合前景决策模型

Two-stage portfolio prospect decision model considering additional emotion

控制与决策. 2021, 36(3): 724-732 <https://doi.org/10.13195/j.kzyjc.2019.0565>

基于矩阵的双论域模糊概率粗糙集增量更新算法

Incremental updating of fuzzy probability rough sets over two universes based on matrix method

控制与决策. 2021, 36(3): 553-564 <https://doi.org/10.13195/j.kzyjc.2019.0692>

一种要素双模糊的限制交流结构合作博弈方法及应用

An allocation model of limited communication structure cooperative game with dual fuzzy elements

控制与决策. 2021, 36(2): 475-482 <https://doi.org/10.13195/j.kzyjc.2019.1048>

考虑供应商技术截断的“主-供”合作机制演化博弈分析

Evolutionary game analysis of “main manufacturer-supplier” collaboration mechanism considering supplier’s technology truncation

控制与决策. 2021, 36(10): 2547-2552 <https://doi.org/10.13195/j.kzyjc.2019.1678>

概率犹豫模糊决策理论与方法综述

An overview of probabilistic hesitant fuzzy decision-making theory and methods

控制与决策. 2021, 36(1): 42-51 <https://doi.org/10.13195/j.kzyjc.2020.0465>

鲁棒博弈动力学: 分析与综合

王龙^{1†}, 王国丞¹, 苏奇²

(1. 北京大学系统控制研究中心, 北京 100871; 2. 上海交通大学自动化系, 上海 200240)

摘要: 不确定性广泛存在于各种自然和工程系统中, 这对系统动力学的建模、分析、预测和控制提出巨大挑战。对于工程控制系统, 已经发展出鲁棒控制相关理论以研究不确定性对系统动力学的影响及其综合设计问题。然而, 对于大规模人在环路的集群自主系统, 不确定性下的系统动力学研究仍处于起步阶段。鉴于此, 将经典控制系统中鲁棒控制的概念和方法扩展到大规模集群自主系统, 提出鲁棒博弈动力学的分析和综合方法, 针对类典型的不确定性研究了系统的演化动力学, 并结合经济学、心理学、社会学中的期望效用理论探讨博弈群体对不确定性的主观响应。不确定性可以逆转经典博弈动力学的演化方向以及产生多稳态和周期振荡极限环等丰富的动力学行为, 所作研究揭示了不确定性对集群自主系统演化的重要影响以及群体博弈系统动力学的复杂性。

关键词: 不确定性; 鲁棒性; 演化博弈论; 风险偏好; 复杂系统

中图分类号: TP13 **文献标志码:** A

DOI: 10.13195/j.kzyjc.2024.1292

引用格式: 王龙, 王国丞, 苏奇. 鲁棒博弈动力学: 分析与综合 [J]. 控制与决策, 2025, 40(6): 1761-1773.

Robust game dynamics: Analysis and synthesis

WANG Long^{1†}, WANG Guo-cheng¹, SU Qi²

(1. Center for Systems and Control, Peking University, Beijing 100871, China; 2. Department of Automation, Shanghai Jiao Tong University, Shanghai 200240, China)

Abstract: Uncertainty exists in a wide range of natural and engineering systems, posing great challenges to the modeling, analysis, prediction and control of system dynamics. For engineering control systems, robust control theory has been developed to study the effect of uncertainty on system dynamics and its resistance to disturbance. However, the study of system dynamics under uncertainty in large-scale autonomous systems with humans in the loop is still in its infancy. In this paper, the concept of robust dynamics is extended to large-scale collective systems, and a robust game dynamics analysis method is proposed. Based on the proposed framework, three typical types of uncertainty are studied, and the subjective response of agents to uncertainty is also explored by combining the expected utility theory in economics. Unlike classical game dynamics, the study finds that uncertainty can reverse the direction of system evolution, and generate rich dynamical behaviors such as multistability and periodic oscillation limit rings. This study reveals the significant impact of uncertainty on the evolutionary dynamics and the complexity of collective behaviors.

Keywords: uncertainty; robustness; evolutionary game theory; risk preference; complex systems

0 引言

博弈论是用来描述理性个体在对抗冲突情况下决策方式的理论体系。von Neumann 等^[1]于1944年发表的巨著《博弈论与经济行为》标志着现代博弈论的诞生。经过以 Nash 为代表的众多学者的完善, 博弈论已经被广泛应用于经济学、物理学、控制科学、计算机科学、生物生态学等多个领域。1973年, Smith 等^[2]将群体动态演化的思想引入博弈论, 放弃了理性个体的假设, 用动态策略更新过程替代策略

选择, 创立了演化博弈论。该理论可以描述策略传播、社会学习、病毒传播等动态过程, 已被广泛应用于研究群体行为的涌现、多智能体的交互与学习、群体的智能决策、合作行为的演化、观点传播动力学等众多领域, 取得了巨大的成就^[3-9]。

在经典的演化博弈论中, 博弈的信息是完全的、确定的。每个个体获得确定的收益, 并且精确地知道其他个体的收益函数^[6,10-15]。个体可以通过修正自己的策略, 或者模仿其他个体的策略, 最终使得系统达

收稿日期: 2024-11-05; 录用日期: 2025-02-18.

基金项目: 国家自然科学基金项目 (62036002).

[†]通信作者. E-mail: longwang@pku.edu.cn.

到均衡状态. 然而, 在现实系统中不确定性广泛存在^[16-32]. 比如在自然界中, 温度、湿度等因素的变化都具有较大的不确定性. 在工程系统中, 物理参数的误差、未建模的动态等因素都可以看作真实系统相对于其数学模型的不确定性. 在控制理论中, 鲁棒性是指系统抵御各种不确定性 (如物理参数摄动、测量误差、未建模动态、环境变化、外部干扰等) 的能力. 研究控制系统的鲁棒性可以明确不确定性对系统动力学的影响, 以及帮助设计控制器使得系统在具有不确定性的情况下依然能够保证高效运行^[18-23,25]. 类似地, 在群体博弈系统中, 不确定性对系统的动力学行为的影响, 以及具有不确定性的博弈系统的综合设计是重要的研究课题, 但目前关于该问题的研究仍比较缺乏. 此外, 相对于传统的控制系统, 博弈系统中的个体往往具有一定的智能性. 所以在研究博弈系统的鲁棒性时, 还需要考虑个体对于不确定性的主观响应, 即不同个体对于不确定性具有不同的偏好程度, 进而导致不同个体在不确定性的环境中做出不同的决策. 这也意味着相对于传统控制系统而言研究博弈动力学中的不确定性以及鲁棒性更加复杂.

不确定性有许多种类型, 针对不确定性的分类尚无统一方式. 在控制理论中, 现有研究通常将不确定性划分为结构型和非结构型两类^[20, 22-23, 26]. 结构型不确定性一般指系统中的未建模动态, 非结构型不确定性一般指系统中的参数摄动. 在经济学中, 广义的不确定性包括风险与 Knightian 不确定性^[33-34]. 风险指能够被计算概率分布的不确定性, Knightian 不确定性指不能够获得概率分布 (即不可预测、未知) 的不确定性. 在机器学习领域, 不确定性可以分为偶然不确定性 (也称作随机不确定性) 与认知不确定性^[35-36]. 偶然不确定性是由于数据本身不可避免的误差导致的, 认知不确定性来源于知识的缺乏、模型本身不完备等因素. 在博弈动力学中, 已有文献研究了多种因素 (包括选择强度、环境容纳量、环境状态) 的不确定性对于系统动力学行为的影响^[16,37-41]. 这些研究发现, 不确定性对策略的演化具有一定程度的影响, 但只考虑了非常有限的确定性因素, 忽略了收益的随机性、博弈群体规模的不确定性等. 这些类型的不确定性对于博弈系统是更普遍、更广泛的. 更重要的是, 之前的研究都只是单纯探究了不确定性对于系统动力学的影响, 并未考虑个体对不确定性做出不同的主观响应. 对于由具有一定智能性个体组成的群体而言, 考虑个体对不确定性的响应是必要的, 这可能会本质地改变系统的动力学行

为.

受到鲁棒控制理论的启发, 本文综合博弈动力学中不确定性以及随机性的相关研究, 提出鲁棒博弈动力学的理论以及研究方法. 针对博弈系统中 3 种典型的不确定性, 即系统中观测不确定性^[42]、交互收益的不确定性^[43] 以及群体规模的不确定性^[44], 介绍群体博弈系统在不同类型不确定性下的动力学行为以及鲁棒性. 其中观测不确定性可以归为认知不确定性, 后两种不确定性可以归为随机不确定性. 具体而言, 针对观测不确定性, 提出博弈动力学下个体对于不确定性的风险响应模型, 并研究不同的风险偏好以及适应性风险偏好对于动力学的影响. 结果表明, 即使对于同样的博弈收益结构, 不同的风险偏好也会使得系统涌现出完全不同的动力学现象. 在适应性风险偏好下, 系统甚至会出现持续振荡的极限环现象. 此外, 对于收益不确定性以及群体规模不确定性, 本文结果表明这两种不确定性都能显著改变系统的动力学. 例如在背叛策略占优的囚徒困境博弈中, 不确定性可以逆转演化动力学的方向, 使得合作行为占优. 这些结果系统性地展示了博弈动力学的鲁棒特征, 揭示了不确定性对于博弈动力学的重要影响以及群体行为的复杂性.

1 基本概念

1.1 经典博弈范式

一个完整的博弈模型包含三要素, 分别是参与者 (players)、策略 (strategies) 和收益 (payoffs). 最经典的两人两策略博弈模型包含囚徒困境 (prisoner's dilemma)、雪堆博弈 (snowdrift game)、猎鹿博弈 (stag-hunt game) 等. 这几类博弈可以通过如下收益矩阵进行刻画:

$$\begin{array}{cc} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{bmatrix} R & S \\ T & P \end{bmatrix} \end{array} \quad (1)$$

每个参与者都有两种可供选择的策略: 合作与背叛. 当双方都选择合作时, 双方均获得 R 的收益; 当双方都选择背叛时, 双方均获得 P 的收益; 当一方选择合作, 另一方选择背叛时, 合作者获得 S 的收益, 背叛者获得 T 的收益. 当收益矩阵中 4 个参数满足关系 $S < P < R < T$ 时, 无论对手选择什么策略, 背叛永远会获得更高的收益, 此时该博弈为囚徒困境. 当 $T > R$ 且 $S > P$ 时, 该博弈为雪堆博弈, 个体的最优选择为与对手相反的策略. 当 $T < R$ 且 $S < P$ 时, 该博弈为猎鹿博弈, 个体的最优响应策略与对手的策略相同.

1.2 演化博弈动力学

演化博弈将群体的概念引入到博弈中, 拓展了个体完全理性的假设, 进一步考虑有限理性的场景, 研究的基本对象为策略在群体之中传播的动力学过程. 本文研究中, 除第 3 节考虑结构化的群体之外, 其他地方如无特殊说明均假设群体为均匀混合的. 对于一个包含 N 个个体的均匀混合群体, 假设合作者数量为 k , 背叛者数量为 $N - k$, 每个个体的收益为与群体中所有个体 (包含自己) 博弈获得收益的平均值, 那么合作者与背叛者的收益分别为

$$\pi_C = \frac{kR + (N - k)S}{N}, \quad (2a)$$

$$\pi_D = \frac{kT + (N - k)P}{N}. \quad (2b)$$

个体的收益决定了其适应度, 适应度为收益的递增函数, 衡量个体传播自己策略的能力 (繁殖率). 比较经典的适应度形式为 $f = 1 + s\pi$, 其中 s 为选择强度, s 越大代表收益对于适应度的影响程度越大. 在演化博弈领域, 大多数理论研究主要关注弱选择 ($s \ll 1$) 的情况^[45-46]. 在所有个体获得收益后, 群体内部适应度更高的个体有更大的概率产生子代. 具体而言, 以灭生过程为例^[47], 每个时间步内, 随机选择一个个体 A , 其以正比于适应度的概率随机模仿另一个个体 B 的策略. 对于均匀混合群体, 系统的状态可以用合作者数量 k 来表示, 系统的动力学可以用马尔可夫链刻画. 该马尔可夫链的转移概率, 即 k 增加 1 或者减少 1 的概率分别为

$$T^+(k) = \frac{N - k}{N} \frac{k f_C}{k f_C + (N - k - 1) f_D}, \quad (3b)$$

$$T^-(k) = \frac{k}{N} \frac{(N - k) f_D}{(k - 1) f_C + (N - k) f_D}. \quad (3b)$$

其他经典的更新过程还包括生灭过程、对比较过程等^[47]. 不同的更新过程对应着 $T^+(k)$ 与 $T^-(k)$ 形式上的细微区别, 但研究方法基本相同, 故此处不再赘述. 该系统具有两个吸收态: 全为合作者 ($k = N$) 或全为背叛者 ($k = 0$). 假设初始时刻系统中只有一个合作者, 其余全是背叛者, 将系统最终能够到达全是合作者状态的概率定义为合作者的固定概率, 记作 ρ_C , 其表达式^[48] 如下:

$$\rho_C = 1 / \left(1 + \sum_{i=1}^{N-1} \prod_{k=1}^i \frac{T^-(k)}{T^+(k)} \right). \quad (4)$$

类似地, 也可以定义并给出背叛者的固定概率 ρ_D . 若 $\rho_C > \rho_D$, 则称合作相比背叛在群体中是演化占优的. 此外, 在中性选择下 (选择强度为 0), 合作者的固

定概率等于初始合作者的比例 $1/N$, 故有时也称自然选择促进合作演化的条件为 $\rho_C > 1/N$.

当群体规模 N 充分大时, 可以使用扩散过程将系统近似为一个连续系统. 用 x 表示合作者在群体中所占的比例 ($x = k/N$), x 的变化可由如下 Langevin 方程刻画:

$$dx = [T^+(x) - T^-(x)]dt + \sqrt{[T^+(x) + T^-(x)]/N}dW_t. \quad (5)$$

W_t 为标准的 Wiener 过程. 当群体规模趋于无限大 ($N \rightarrow \infty$) 时, 方程 (5) 的随机项消失, 此时系统可用一个常微分方程来描述. 忽略时间尺度的差异, 方程 (5) 与经典的复制动力学方程一致^[49-51], 有

$$\dot{x} = x(1 - x)(\pi_C - \pi_D). \quad (6)$$

对于囚徒困境, $x = 0$ (全是背叛者) 是系统唯一的稳定平衡点. 对于雪堆博弈, 系统唯一的稳定平衡点位于内部: $x = \frac{P - S}{R - S - T + P}$. 边界的两个平衡点是不稳定的, 这表示合作者与背叛者可以稳定共存. 对于猎鹿博弈, 系统具有两个稳定的平衡点 $x = 0$ 和 $x = 1$, 以及一个不稳定的内部平衡点 $x = \frac{P - S}{R - S - T + P}$.

2 风险响应的不确定性

在一个群体博弈系统中, 考虑经典的两人两策略博弈 (如式 (1)), 每个个体的收益由其策略决定, 由式 (2) 给出. 假设个体在模仿其他个体策略时存在不确定性^[42]: 个体无法准确地观测到其余每个个体的精确收益, 只能得到对方收益的一个估计值. 更准确地, 个体 i 观察到的 j 的收益具有一定的随机性, 由如下高斯随机变量给出:

$$\pi_{ji} = \pi_j + \sigma_{ji}\xi. \quad (7)$$

其中: ξ 是均值为 0 方差为 1 的高斯随机变量; π_j 为个体 j 的真实收益 (等于 π_C 或 π_D); σ_{ji} 为 i 观测 j 收益时随机性的标准差, 其衡量了 i 对 j 收益观察的精确程度. 在具体研究中, 假设观测的标准差仅取决于观测者与被观测者的策略, 即 $\sigma_{ji} = \sigma_{s_j s_i}$ (其中 $s_i, s_j \in \{C, D\}$ 代表个体 i 和 j 的策略). 此时 σ_{ji} 可能的取值有 $\sigma_{CC}, \sigma_{CD}, \sigma_{DC}, \sigma_{DD}$. 类似于收益结构 (式 (1)), 将如下矩阵称为不确定性结构:

$$\begin{bmatrix} \sigma_{CC} & \sigma_{CD} \\ \sigma_{DC} & \sigma_{DD} \end{bmatrix}. \quad (8)$$

另外, 用随机变量 $\pi_{s_j s_i} = \pi_{s_j} + \sigma_{s_j s_i}\xi$ 表示策略是 s_i 的个体观测策略为 s_j 个体的收益时得到的观测值.

由于收益观察的不确定性, 当观察到某个个体

的收益从而模仿其策略时,该策略模仿具有一定的风险性,因为被观察的收益不一定是该个体的真实收益.面对该风险,个体会采取不同的决策偏好,从而指导他们具体的行为策略.在经济学中,个体的风险偏好通常由期望效用理论进行描述.该理论由 von Neumann 等^[1]提出,已被广泛应用于各领域对风险决策的研究.该理论假设个体具有效用函数 $U(\pi)$, 效用函数描述 π 的实际收益给个体 i 带来的满足程度,个体总是选择可以最大化自己期望效用的选项或行动.效用函数不一定是线性的,其凹凸性反映了个体对于风险的偏好.凹的效用函数 ($U'' < 0$) 代表个体是风险厌恶 (risk averse) 的.比如某个体的效用函数为 $U(\pi) = \sqrt{\pi}$, 该个体面临两种选择:一是确定性地获得 50 元,另一种是以等概率获得 100 元或者 0 元.对于该个体而言,选择前者产生的效用是 $\sqrt{50}$, 选择后者产生的期望效用是 $0.5\sqrt{100}$, 所以该个体会选择前者.尽管两种选项的期望收益是相同的,但该个体仍偏向于选择风险较小的选项,所以被称作风险厌恶的.类似地,凸的效用函数 ($U'' > 0$) 表示风险追逐 (risk seeking), 此时个体倾向于选择后者 (风险较大的选项).线性的效用函数 ($U'' = 0$) 对应着风险中性 (risk neutral), 此时个体对两种选择没有偏向.给定效用函数 $U(\pi)$, 其表示的风险厌恶程度可由绝对风险厌恶指标 (absolute risk aversion, ARA) 描述^[52-53]

$$\text{ARA}(\pi) = -\frac{U''}{U'}. \quad (9)$$

ARA > 0 代表风险厌恶,其值越大表示风险厌恶程度越高.反之,ARA < 0 代表风险追逐,其值越小表示风险追逐程度越高.

采用经典的指数型效用函数^[52,54-56], 个体 i 的效用函数为

$$U_i(\pi) = \begin{cases} 1 - \frac{s}{\delta_i} + \frac{s}{\delta_i} \exp(\delta_i \pi), & \delta_i \neq 0; \\ 1 + s\pi, & \delta_i = 0. \end{cases} \quad (10)$$

其中: s 为选择强度; δ_i 为个体 i 的风险偏好程度.该效用函数的 ARA 为常数 $-\delta_i$, $\delta_i < 0$ 表示风险厌恶, $\delta_i > 0$ 表示风险追逐.在演化动力学中,通常假设 $s \ll 1$, 表示收益对于个体的效用影响较小.

在每个时间步,随机选择一个个体 i 进行策略更新 (i 个体以 $1/N$ 的概率被选择).由于观测的不确定性,个体 i 发现不同的合作者获得的收益都不相同,不同的背叛者获得的收益也不相同,个体 i 认为如果自己模仿其余个体的策略也会获得随机的收益.因此,个体 i 会计算自己视角下每个个体的期望效用,

然后依据正比于期望效用的概率随机挑选一个个体进行模仿.该过程相当于将期望效用理论中的最大化期望效用的过程用动态的学习过程刻画.具体而言, i 个体被选择并且模仿 j 个体策略的概率为

$$e_{i \rightarrow j} = \frac{1}{N} \frac{\mathbb{E}[U_i(\pi_{ji})]}{\sum_{k \neq i} \mathbb{E}[U_i(\pi_{ki})]}. \quad (11)$$

2.1 固定风险偏好

基于以上模型,假设群体中存在两类风险偏好: $\delta_1 > 0$ (风险追逐) 与 $\delta_2 < 0$ (风险厌恶).首先考虑个体风险偏好保持不变的情况.群体中共有 4 种类型的个体: 风险追逐合作者、风险追逐背叛者、风险厌恶合作者、风险厌恶背叛者.假设合作者在整个群体中的比例为 x , 风险追逐个体的比例为 p (p 此时为一常数), 风险追逐合作者的比例为 q , 那么这 4 类个体的比例分别为 $q, p - q, x - q, 1 - x - p + q$.该系统为一个二维系统,其状态可以用 (x, q) 表示.单个时间步内, x 和 q 的期望变化量为

$$\mathbb{E}(\Delta x) = \frac{1}{N} (T_1 + T_2 - T_3 - T_4), \quad (12a)$$

$$\mathbb{E}(\Delta q) = \frac{1}{N} (T_2 - T_4). \quad (12b)$$

其中: T_1, T_2, T_3, T_4 为系统状态在一个时间步内从 (x, q) 转移到 $(x + 1/N, q), (x + 1/N, q + 1/N), (x - 1/N, q), (x - 1/N, q - 1/N)$ 的概率,对应发生的更新事件分别为风险厌恶背叛者模仿合作者的策略、风险追逐背叛者模仿合作者的策略、风险厌恶合作者模仿背叛者的策略、风险追逐合作者模仿背叛者的策略.这些转移概率的表达式如下:

$$T_1 = \frac{(1 - x - p + q)x\mathbb{E}[U_{\delta_2}(\pi_{CD})]}{x\mathbb{E}[U_{\delta_2}(\pi_{CD})] + (1 - x)\mathbb{E}[U_{\delta_2}(\pi_{DD})]}, \quad (13a)$$

$$T_2 = \frac{(p - q)x\mathbb{E}[U_{\delta_1}(\pi_{CD})]}{x\mathbb{E}[U_{\delta_1}(\pi_{CD})] + (1 - x)\mathbb{E}[U_{\delta_1}(\pi_{DD})]}, \quad (13b)$$

$$T_3 = \frac{(x - q)(1 - x)\mathbb{E}[U_{\delta_2}(\pi_{DC})]}{x\mathbb{E}[U_{\delta_2}(\pi_{CC})] + (1 - x)\mathbb{E}[U_{\delta_2}(\pi_{DC})]}, \quad (13c)$$

$$T_4 = q \frac{(1 - x)\mathbb{E}[U_{\delta_1}(\pi_{DC})]}{x\mathbb{E}[U_{\delta_1}(\pi_{CC})] + (1 - x)\mathbb{E}[U_{\delta_1}(\pi_{DC})]}, \quad (13d)$$

其中: $U_{\delta_1}(\cdot)$ 表示 δ_i 取值为 δ_1 的效用函数 (式 (10)), U_{δ_2} 的定义同理.对于弱选择 ($s \ll 1$) 以及弱风险偏好 ($|\delta_i| \ll 1$), 将式 (13) 泰勒展开并保留 s 的一阶项, 然后将其代入式 (12).在时间尺度变换 $t = \tau/N$ (τ 为时间步) 下, 得到系统的动力学方程 (详细推导过程见文献 [42])

$$\dot{x} = \frac{\mathbb{E}(\Delta x)}{1/N} = sx(1-x)[(\pi_C - \pi_D) + \frac{1}{2}((p\delta_1 + (1-p)\delta_2)(\sigma_{CD}^2 - \sigma_{DD}^2) + (q\delta_1 + (x-q)\delta_2)(\sigma_{CC}^2 - \sigma_{CD}^2 - \sigma_{DC}^2 + \sigma_{DD}^2))], \quad (14a)$$

$$\dot{q} = \frac{\mathbb{E}(\Delta q)}{1/N} = px - q + O(s). \quad (14b)$$

因此, 对于弱选择, q 的变化率远大于 p . 该系统具有快慢动力学的特征, 所以可以对其进行尺度分离. 可以认为在 x 几乎还未发生变化之前 q 就已经达到了平衡态

$$q = px. \quad (15)$$

将式 (15) 代入 (14), 方程简化为一维系统

$$\dot{x} = sx(1-x)\left[\pi_C - \pi_D + \frac{\bar{\delta}}{2}(V_C - V_D)\right]. \quad (16)$$

其中

$$V_C = x\sigma_{CC}^2 + (1-x)\sigma_{CD}^2, \quad (17a)$$

$$V_D = x\sigma_{DC}^2 + (1-x)\sigma_{DD}^2, \quad (17b)$$

$$\bar{\delta} = p\delta_1 + (1-p)\delta_2. \quad (17c)$$

这里: V_C 和 V_D 分别为合作者或者背叛者被一个随机个体 (x 的概率为合作者, $1-x$ 的概率为背叛者) 观测到的收益值的方差; $\bar{\delta}$ 为群体的平均风险偏好程度. 方程 (16) 表明风险偏好和不确定性可以显著改变系统的动力学行为, 当群体整体呈风险厌恶 ($\bar{\delta} < 0$) 且合作者的收益更透明, 即具有更小的不确定性 ($V_C < V_D$, 充分条件为 $\sigma_{CC}^2 < \sigma_{DC}^2, \sigma_{CD}^2 < \sigma_{DD}^2$) 时, 合作者相对于背叛者会获得更大的演化优势. 类似地, 如果群体为风险追逐的且合作者的收益更加模糊 (即 $V_C > V_D$, 充分条件为 $\sigma_{CC}^2 > \sigma_{DC}^2, \sigma_{CD}^2 > \sigma_{DD}^2$), 则合作行为也会获得更高的等效收益. 此外, 由于 V_C 与 V_D 依赖于合作者的比例 (x), 当 $V_C - V_D$ 不是恒为正或者恒为负时, 不确定性是促进还是抑制合作演化取决于当前系统合作者的比例 x .

此外, 将方程 (16) 与经典的复制动力学方程 (式 (6)) 作比较, 可以得到如下定理^[42], 其简洁地描述了风险偏好如何影响系统的动力学行为.

定理 1 式 (16) 所描述的动力学系统等价于在经典的复制动力学方程下 (无不确定性) 对收益结构做如下变换:

$$\begin{bmatrix} R & S \\ T & P \end{bmatrix} \Rightarrow \begin{bmatrix} R & S \\ T & P \end{bmatrix} + \frac{\bar{\delta}}{2} \begin{bmatrix} \sigma_{CC}^2 & \sigma_{CD}^2 \\ \sigma_{DC}^2 & \sigma_{DD}^2 \end{bmatrix}. \quad (18)$$

即给定不确定性结构 (式 (8)) 以及群体的平均风险

偏好程度 $\bar{\delta}$, 式 (18) 左边收益矩阵的演化动力学等价于右边收益矩阵在经典复制动力学下 (无不确定性) 的演化动力学.

在图 1 中, 选取 $\delta_1 = 0.2, \delta_2 = -0.2, \sigma_{CC} = \sigma_{DD} = 0, \sigma_{CD} = \sigma_{DC} = \sqrt{20}$. 此时 $p = 1$ 对应所有个体全为风险追逐 (δ_1) 的情况, $p = 0$ 对应所有个体全为风险厌恶 (δ_2) 的情况. 对于囚徒困境博弈, 风险中性的群体并不改变系统的动力学. 风险厌恶的群体可以使得系统的动力学变得与经典演化博弈中猎鹿博弈的动力学 (内部具有一个不稳定的平衡点, 两端边界为稳定的平衡点) 类似. 而风险追逐的群体可以使

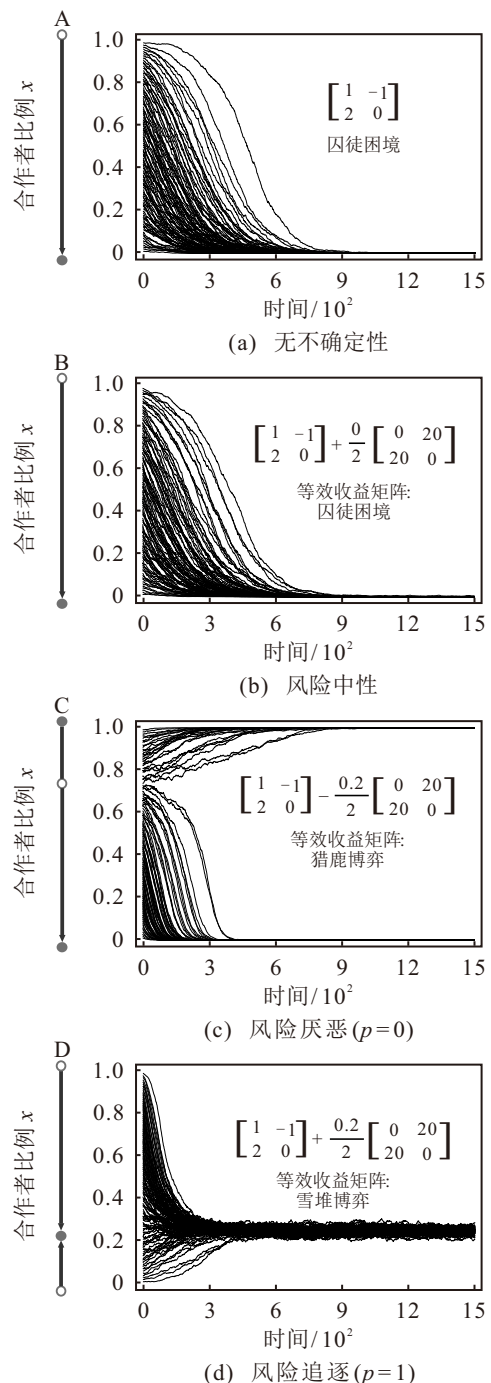


图1 风险偏好以及不确定性可以定性改变系统的动力学

得群体的动力学变得与雪堆博弈 (两端边界为不稳定的平衡点, 内部具有一个稳定的平衡点) 类似.

2.2 自适应风险偏好

以上主要研究了个体风险偏好固定不变的情况, 在真实系统中, 收益更高的个体对于风险的容忍程度更高, 更乐意参与一些风险较高的活动^[52,55-57], 比如投资股票、赌博等. 所以, 假设收益高的个体有更高的概率变为风险追逐, 而收益低的个体有更高概率变为风险厌恶^[42]. 仍然考虑群体中存在两种风险偏好: 风险追逐 ($\delta_1 > 0$) 和风险厌恶 ($\delta_2 < 0$). 在每个时间步内, 随机选择一个个体 i 进行策略更新 (式 (11)), 然后随机选择另一个个体 j 更新其风险偏好. 个体 j 以 $1 - u$ 的概率保持原本的风险偏好, 以 u 的概率重新考虑其风险偏好. 当重新考虑风险偏好时, 个体 j 以概率

$$v_j^+ = \frac{1}{2} + \frac{\pi_j - \eta}{2D} \quad (19)$$

变为风险追逐, 以 $v_j^- = 1 - v_j^+$ 的概率变为风险厌恶. $D = |\pi - \eta|_{\max}$ 表示所有可能的收益与切换阈值 η 最大的差值. 在此模型下, 系统中风险追逐类型个体的比例 p 不再是常量. p 的动力学由如下方程给出:

$$\dot{p} = \frac{E(\Delta p)}{1/N} = u[(x - q)v_C^+ + (1 - x - p + q)v_D^+ - qv_C^-(p - q)v_D^-]. \quad (20)$$

$v_C^+, v_D^+, v_C^-, v_D^-$ 表示合作者或背叛者变为风险追逐或者风险厌恶的概率. 与式 (14) 相结合, 系统的动力学方程为一个三维系统. 类似于弱选择, 假设风险偏好的更新速率也足够慢, 即 $u \ll 1$, 此时 q 相对于 x 与 p 变化得都足够慢. 于是可以将式 (15) 代入 (20) 得到

$$\dot{p} = u \left[\frac{1}{2} + \frac{\bar{\pi} - \eta}{2D} - p \right], \quad (21)$$

其中 $\bar{\pi} = x\pi_C + (1 - x)\pi_D$ 为群体的平均收益. 式 (16) 与 (21) 构成一个二维系统. 该系统对于不同的收益矩阵、不同的不确定性结构、不同的风险偏好 (δ_1 与 δ_2) 都有不同的动力学性质. 为了减少参数的数目, 重点研究一类典型的动力学. 考虑收益结构为经典的单参数囚徒困境博弈: $R = 1, S = -r, T = 1 + r, P = 0$. 选取 $\eta = 1/2$, 此时 $D = 1/2 + r$. 即使对于如此简化的收益结构, 该系统仍然具有丰富的动力学现象^[42]. 此处重点关注极限环的情况, 其余动力学现象参见文献 [42].

定理 2 对于只有单参数 r 的囚徒困境博弈, 在自适应风险偏好系统中 (式 (16) 和 (21)) 存在全局渐

近稳定的极限环的充分必要条件为

$$\sigma_{CC}^2 - \sigma_{DC}^2 < \sigma_{CD}^2 - \sigma_{DD}^2 < -\frac{2r(2r + 1)}{\delta}, \quad (22a)$$

$$u < u^*, \quad (22b)$$

其中 u^* 为 u 的临界值, 由下文给出.

这里给出证明概要, 详细证明见文献 [42]. 用 (x, p) 表示系统的状态, 并令 $\alpha = \sigma_{CC}^2 - \sigma_{DC}^2, \beta = \sigma_{CD}^2 - \sigma_{DD}^2$. 该系统最多有 4 个平衡点, 其中两个平衡点位于系统的边界, 分别为

$$e_1 = (x_1, p_1) = \left(0, \frac{r}{2r + 1}\right), \quad (23a)$$

$$e_2 = (x_2, p_2) = \left(1, \frac{r + 1}{2r + 1}\right). \quad (23b)$$

这两个平衡点总是存在的. 此外, 系统还具有两个内部平衡点

$$e_+ = (x_+, p_+), e_- = (x_-, p_-). \quad (24)$$

其中

$$x_{+/-} = \frac{\alpha - 3\beta}{4(\alpha - \beta)} \pm \frac{\sqrt{(\alpha - 3\beta)^2 + 8(\alpha - \beta)(\beta + 2r(2r + 1)/\delta)}}{4(\alpha - \beta)}, \quad (25a)$$

$$p_{+/-} = \frac{1}{2} + \frac{x_{+/-} - 1/2}{r + 1/2}. \quad (25b)$$

这两个内部平衡点并不总是存在, 仅当 $(x_{+/-}, p_{+/-}) \in [0, 1]^2$ 时有意义. 将该二维系统在 e_1 和 e_2 处线性化, 得到对应的 Jacobi 矩阵为

$$J_1 = \begin{bmatrix} -s \left(r + \frac{\delta\beta}{4r + 2} \right) & 0 \\ \frac{u}{2r + 1} & -u \end{bmatrix}, \quad (26a)$$

$$J_2 = \begin{bmatrix} s \left(r - \frac{\delta\alpha}{4r + 2} \right) & 0 \\ \frac{u}{2r + 1} & -u \end{bmatrix}. \quad (26b)$$

若要得到全局渐近稳定的极限环, e_1 与 e_2 必须是不稳定的, 这要求 J_1 与 J_2 的特征值不能全为负, 据此可以得到

$$\alpha < \frac{2r(2r + 1)}{\delta}, \quad (27a)$$

$$\beta < -\frac{2r(2r + 1)}{\delta}. \quad (27b)$$

在此条件下, 通过计算可以得到 e_+ 并不存在 (位于区间 $[0, 1]$ 之外), 只有 e_- 是存在的. 用类似的方法, 将系统在 e_- 附近线性化, 求出其 Jacobi 矩阵 J_- ^[42]. 通过计算可以得到 $\det(J_-)$ 始终为正, 所以 e_- 不稳定当且

仅当 $\text{tr}(J_-)$ 为负. 通过计算可以得到 $\text{tr}(J_-)$ 为负的条件为

$$\alpha < \beta, \tag{28a}$$

$$u < u^*, \tag{28b}$$

其中

$$u^* = sx_-(1-x_-)(x_- - 1/2)(\alpha - \beta)/(2r + 1). \tag{29}$$

将以上条件结合起来, 利用 Poincaré-Bendixson 定理即可得到定理 2.

定理 2 的直观解释是式 (22a) 的必要条件为 $\alpha < 0$ 且 $\beta < 0$. 在该情况下, $V_C < V_D$ 总是成立, 这表明合作者的收益总是更加透明. 定理 1 已经表明在该情况下风险厌恶促进合作, 而风险追逐抑制合作的演化. 在囚徒困境中, 如果所有个体全都采取背叛, 则群体的平均收益最低, 个体会倾向于变为风险厌恶. 由于风险厌恶促进合作, 个体会逐渐转变为采取合作策略. 当足够多的个体采取合作策略后, 群体的平均收益上升, 个体倾向于变为风险追逐. 由于风险追逐抑制合作, 背叛行为再次占据群体. 在满足适当的条件后, 整个系统便会反复在合作与背叛、风险厌恶与风险追逐之间周期震荡. 在图 2 中, 图 A 直观上解释了这一循环机制, 图 B ~ E 中, 选取参数 $r = 0.5$, $\sigma_{CC} = \sigma_{CD} = 0$, $\sigma_{DC} = \sqrt{200}$, $\sigma_{DD} = \sqrt{100}$, $u = 0.000\ 3$ (B, C), $u = 0.000\ 9$ (D, E), $s = 0.01$, $N = 500\ 000$, $\eta = 0.5$, $\delta_1 = -\delta_2 = 0.03$, 对系统方程进行数值解计算以及 Monte Carlo 仿真, 结果证实了系统会出现振荡动力学. 当风险偏好更新率 u 比较小时, 振荡现象是持续的 (稳定的极限环). 当 u 比较大时, 系统会振荡收敛到内部稳定的平衡点.

这里只展示了最简单的囚徒困境博弈下的适应

性风险偏好动力学. 对于其他类型的博弈, 比如雪堆博弈和猎鹿博弈, 适应性风险偏好则会产生更加复杂丰富的动力学行为, 具体内容见文献 [42].

3 交互收益的不确定性

在文献 [43] 中, 考虑了博弈交互的收益结构具有不确定性的情况. 此时, 个体通过与其他个体交互得到的收益也是随机的. 假设收益结构如下:

$$\begin{bmatrix} R & S \\ T & P \end{bmatrix} + \begin{bmatrix} \theta_R \xi & \theta_S \xi \\ \theta_T \xi & \theta_P \xi \end{bmatrix}, \tag{30}$$

其中 ξ 为标准的高斯随机变量, 均值为 0, 方差为 1. 不同的两个个体交互, 以及相同个体在不同时间的交互所对应的高斯随机变量都是互相独立的. 例如两个合作者 i 与 j 交互, 那么他们会分别获得收益 $R + \theta_R \xi_{ij}$ 和 $R + \theta_R \xi_{ji}$. 这里 ξ_{ij} 与 ξ_{ji} 是互相独立的标准高斯变量. 类似地, 当一个合作者与一个背叛者交互, 或者两个背叛者交互, 也会分别得到叠加随机扰动的收益.

这里考虑结构化的群体, 用网络 $(\mathcal{N}, \mathcal{E})$ 来表示. \mathcal{N} 代表网络中节点的集合, $\mathcal{E} \in \mathcal{N} \times \mathcal{N}$ 代表网络中边的集合. 每个节点代表一个个体, 边代表个体之间的交互关系. 个体的最终收益为其与所有邻居博弈的收益之和. 个体 i 的策略记为 s_i , $s_i = 1$ 表示个体 i 采取合作, $s_i = 0$ 表示个体 i 采取背叛. i 的总收益为

$$\begin{aligned} \pi_i = & \sum_{j \in \mathcal{N}_i} s_i [s_j (R + \theta_R \xi_{ij}) + (1 - s_j) (S + \theta_S \xi_{ij})] + \\ & (1 - s_i) [s_j (T + \theta_T \xi_{ij}) + (1 - s_j) (P + \theta_P \xi_{ij})], \end{aligned} \tag{31}$$

其中 \mathcal{N}_i 代表个体 i 的邻居节点集合. 博弈所产生的收益通过映射 $f_i = \exp(s\pi_i)$ 转化为适应度. 然后通过

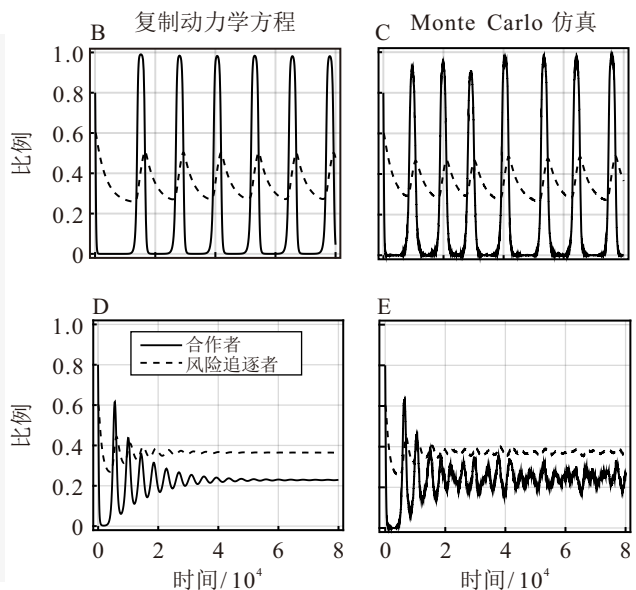
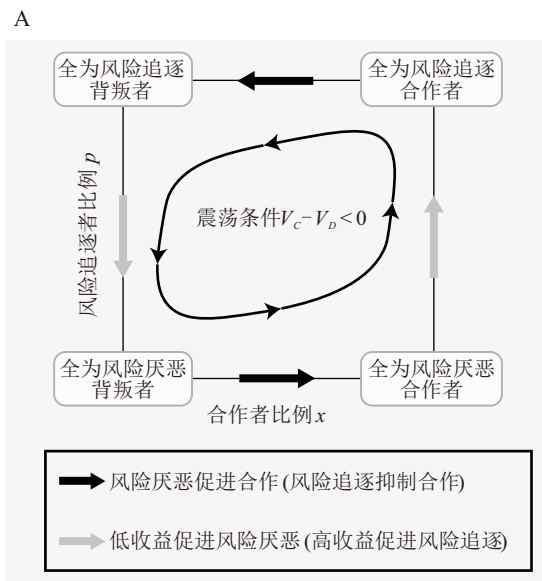


图2 适应性风险偏好导致周期震荡极限环

天生更新过程进行策略更新. 即每次随机选择一个个体, 该个体以正比于适应度的概率模仿某个邻居的策略. 具体而言, 若 j 是 i 的邻居 ($j \in \mathcal{N}_i$), 则 i 模仿 j 的概率为

$$e_{i \rightarrow j} = \frac{f_j}{\sum_{k \in \mathcal{N}_i} f_k}. \quad (32)$$

本文主要考虑的网络为随机规则网络, 假设每个节点的度都为 k , 即每个个体都有 k 个邻居.

定理 3 在度为 k 的随机规则网络中, 合作者的固定概率为

$$\rho_C = \frac{1}{N} + \frac{s}{6k}(\alpha + 3\beta) + o(s). \quad (33)$$

其中

$$\begin{aligned} \alpha &= (k+1)(k-2) \left[(R-S-T+P) + s \left(\frac{1}{2} - \frac{1}{k} \right) (\theta_R^2 - \theta_S^2 - \theta_T^2 + \theta_P^2) \right], \\ \beta &= (k+1) \left[R + s \left(\frac{1}{2} - \frac{1}{k} \right) \theta_R^2 \right] + \\ &\quad (k^2 - k - 1) \left[S + s \left(\frac{1}{2} - \frac{1}{k} \right) \theta_S^2 \right] - \\ &\quad \left[T + s \left(\frac{1}{2} - \frac{1}{k} \right) \theta_T^2 \right] - \\ &\quad (k^2 - 1) \left[P + s \left(\frac{1}{2} - \frac{1}{k} \right) \theta_P^2 \right]. \end{aligned} \quad (34)$$

定理 3 证明见文献 [43].

通过类似的方式也可以得到背叛者的固定概率 ρ_D 的表达式^[43]为

$$\rho_D = \frac{1}{N} - \frac{s}{6k}(2\alpha + 3\beta) + o(s). \quad (36)$$

如果合作者的固定概率大于背叛者的固定概率, 则称合作相对于背叛是占优的. 结合 ρ_C 与 ρ_D 的表达式, 有如下定理.

定理 4 在度为 k 的随机规则网络中, 假设选择强度非常弱 ($s \ll 1$), 合作行为相对于背叛行为占优 ($\rho_C > \rho_D$) 当且仅当

$$\begin{aligned} &(k+1) \left[R + s \left(\frac{1}{2} - \frac{1}{k} \right) \theta_R^2 \right] + \\ &(k-1) \left[S + s \left(\frac{1}{2} - \frac{1}{k} \right) \theta_S^2 \right] - \\ &(k-1) \left[T + s \left(\frac{1}{2} - \frac{1}{k} \right) \theta_T^2 \right] - \\ &(k+1) \left[P + s \left(\frac{1}{2} - \frac{1}{k} \right) \theta_P^2 \right] > 0. \end{aligned} \quad (37)$$

将该理论应用到捐赠博弈中 (一类简化的囚徒困境博弈), 即 $R = b - c$, $S = -c$, $T = b$, $P = 0$. 该博弈中, 合作意味着自己付出 c 的代价为对方提供 b 的收益, 而背叛不付出成本也不为对方提供收益.

假设这两个参数受到独立的随机扰动, 即 b 叠加强度 (标准差) 为 θ_b 的高斯分布, c 叠加强度 (标准差) 为 θ_c 的高斯分布, 那么 $b - c$ 的扰动强度为 $\sqrt{\theta_b^2 + \theta_c^2}$. 将这些值代入式 (37), 得到合作行为占优的条件为

$$\frac{b + s \left(\frac{1}{2} - \frac{1}{k} \right) \theta_b^2}{c - s \left(\frac{1}{2} - \frac{1}{k} \right) \theta_c^2} > k. \quad (38)$$

相较于经典的结果 bck -rule^[14], 合作占优的条件为

$$\frac{b}{c} > k, \quad (39)$$

存在收益不确定性的合作占优条件更加宽松, 所以相较于没有不确定性的情形, 不确定性收益可以促进合作的演化.

在图 3 中, 固定 $c = 1$ 和 $k = 4$, 采用 Monte Carlo 仿真验证合作相对于背叛占优的条件. 结果表明相比于确定性收益, 无论是 b 具有不确定性还是 c 具有不确定性, 都可以显著放松合作演化的条件.

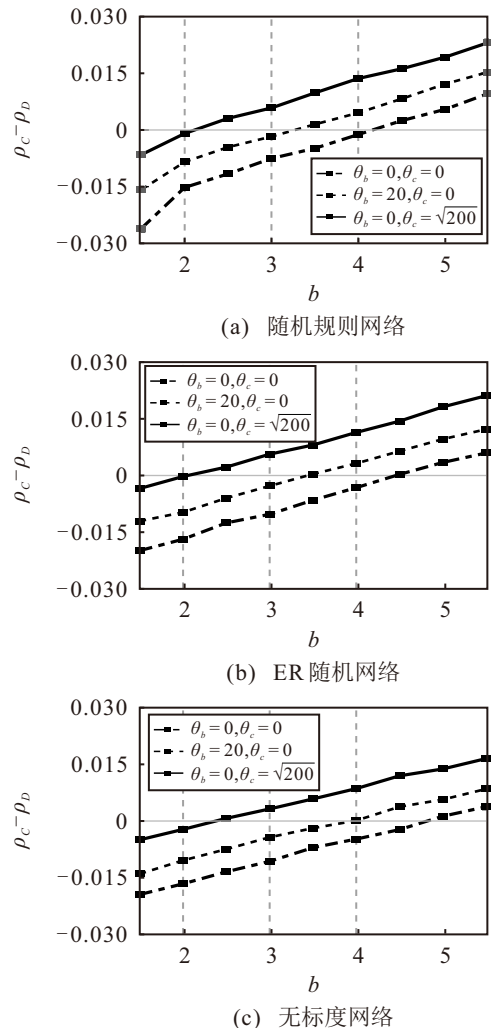


图3 收益结构不确定性可以促进合作的演化

4 群体规模的不确定性

经典的演化博弈模型往往假设无限大种群或者

有限固定规模的种群. 无限大群体通常采用复制动力学方程描述, 有限固定规模的群体通常采用生灭、灭生等更新过程刻画. 在传统的生灭过程 (birth-death process) 下, 每个时间步内, 一个以正比于适应度的概率被随机选择的个体会复制产生一个相同的个体替代随机选择的另一个个体, 这会导致每个时间步内群体的个体数目保持不变. 但是在自然以及工程等各领域, 群体的规模往往是不确定的、随时间随机波动的. 为了刻画这一现象, 解耦生灭过程中的出生和死亡过程, 使得复制产生新个体和个体的死亡是互相独立的两个过程, 即单位时间内产生的新个体和死亡的个体数目不一定相等^[44].

具体而言, 群体中合作者与背叛者的收益由式 (2) 给出. 用 x 表示群体中合作者的数量 (不是比例), y 表示背叛者的数量. 在生过程 (birth process) 中, 个体 i 单位时间内复制产生的后代数用随机变量 ξ_i 表示, 也称作个体 i 的出生率. ξ_i 的期望和方差满足

$$\mathbb{E}(\xi_i) = B + s\pi_i, \tag{40a}$$

$$\text{Var}(\xi_i) = \delta_1 B + \delta_2 s\pi_i. \tag{40b}$$

其中: B 为出生率期望的基准值; s 为选择强度, 表示收益对于出生率的影响程度; π_i 为个体 i 通过群体内部博弈交互得到的收益; δ_1 和 δ_2 为控制方差与收益关系的两个参数, δ_1 描述了基准方差的大小, δ_2 描述了不同收益个体出生率方差的差异. 针对演化博弈中主要关注的弱选择的情况 ($s \ll 1$), 出生率受到收益的影响非常微弱, 可以近似认为期望和方差都是收益的线性函数. 因此式 (40) 是一个比较一般化的假设, 不同的出生模型对应于不同的参数 δ_1 与 δ_2 . 例如经典的 Poisson 过程对应着 $\delta_1 = \delta_2 = 1$ (期望等于方差). 对于死过程 (death process), 假设其为一个经典的 Poisson 过程, 单个个体的死亡速率为 $D + \lambda(x + y)$. 其中第 1 项 D 代表个体的自然死亡率; 第 2 项 $\lambda(x + y)$ 代表群体中竞争所导致的死亡率. 这里的竞争表示由于资源和空间有限, 群体中个体数越多会导致单个个体的死亡率越大. 更具体的模型假设参见文献 [44].

基于该模型, 可以推导得到系统的动力学方程

$$dx = x[\alpha + s\pi_C - \lambda(x + y)]dt + \sqrt{x[\delta_1 B + \delta_2 s\pi_C + D + \lambda(x + y)]}dW_t^{(1)}, \tag{41a}$$

$$dy = y[\alpha + s\pi_D - \lambda(x + y)]dt + \sqrt{y[\delta_1 B + \delta_2 s\pi_D + D + \lambda(x + y)]}dW_t^{(2)}. \tag{41b}$$

其中: $\alpha = B - D$ 为基准净增长率, $W_t^{(1)}$ 和 $W_t^{(2)}$ 为两个独立的 Wiener 过程. 由于主要关注的是合作者与背叛者的相对数量, 可以引入两个新的变量: 合作者比例 $p = x/(x + y)$ 和群体规模 $n = x + y$. 在该坐标变换下, 使用 Itô 公式, 系统方程变为

$$dp = sp(1 - p)(\pi_C - \pi_D)\left(1 - \frac{\delta_2}{n}\right)dt + \frac{y}{n^2}\sqrt{x(\delta_1 B + D + \lambda n)}dW_t^{(1)} - \frac{x}{n^2}\sqrt{y(\delta_1 B + D + \lambda n)}dW_t^{(2)}, \tag{42a}$$

$$dn = [n\alpha + s(b - c)pn - \lambda n^2]dt + \sqrt{x(\delta_1 B + D + \lambda n)}dW_t^{(1)} + \sqrt{y(\delta_1 B + D + \lambda n)}dW_t^{(2)}. \tag{42b}$$

假设基准净增长率是显著量, 即 $\alpha \sim O(1)$, 那么可以得到 $s \ll \alpha$, 此时 p 相对于 n 变化非常快. 系统具有快慢动力学两个尺度, 即 p 还没有变化之前, n 已经几乎收敛到如下平衡点:

$$n = \frac{\alpha + s(b - c)p}{\lambda} \approx \frac{\alpha}{\lambda} := M. \tag{43}$$

M 为稳态时群体的大致规模, 称为环境容纳量 (carrying capacity). 此时可以将 $n = M$ 代入式 (42c), 得到一个一维系统

$$dp = sp(1 - p)\left[\left(1 - \frac{\delta_2}{M}\right)(\pi_C - \pi_D)\right]dt + \sqrt{\frac{(\delta_1 + 1)Bp(1 - p)}{M}}(\sqrt{1 - p}dW_t^{(1)} - \sqrt{p}dW_t^{(2)}). \tag{44}$$

针对该系统, 首先考虑经典的捐赠博弈: $R = b - c$, $S = -c$, $T = b$, $P = 0$. 给定合作者的初始比例为 p_0 , 系统最终总会到达两个吸收态之一: 全是合作者 ($p = 1$) 或全是背叛者 ($p = 0$). 系统最终演化到全是合作者的概率被称为合作者的固定概率, 其与初始比例 p_0 、选择强度、收益结构等均有关. 使用 Fokker-Planck 后向方程, 可以得到合作者的固定概率为 (详细推导见文献 [44])

$$\rho(p_0, n_0) \approx p_0 + \frac{sc}{(\delta_1 + 1)B}(\delta_2 - M)p_0(1 - p_0). \tag{45}$$

在中性选择下 ($s = 0$, 合作者与背叛者的出生率无差别), 合作者的固定概率等于初始合作者比例 p_0 . 对于弱选择, 如果 $\delta_2 < M$, 则与经典演化博弈动力学相同, 合作者固定概率小于中性选择下的固定概率. 此时背叛者更具有演化优势. 但是若 $\delta_2 > M$, 则合作者的固定概率大于中性选择下的固定概率. 合

作相对于背叛具有更大的演化优势(见图4). 该结论表明, 即使背叛者的期望增长率更高, 足够大的方差异性也可以逆转系统的动力学方向, 使得合作者在演化中占优. 在图4中, 考虑环境容纳量为200的群体, 对比了通过对式(41)进行数值计算得到的固定概率和固定概率的近似理论解(式(45)), 分别展示了合作可以演化($\rho > p_0$)的临界条件. 结果表明, 理论近似得到的合作占优条件 $\delta_2 > M$ 与数值仿真吻合较好.

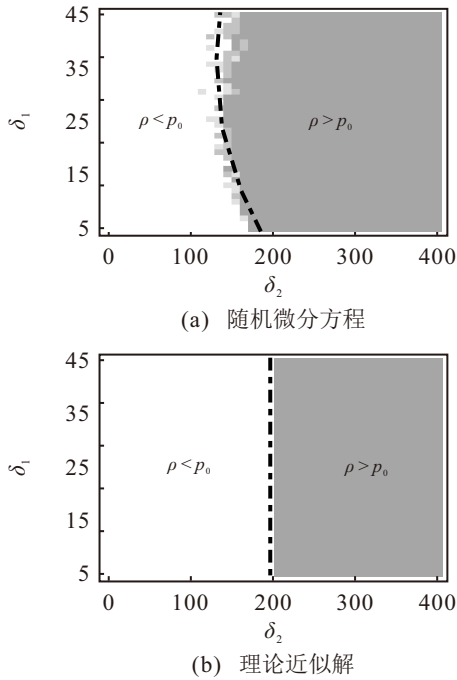


图4 不确定群体规模可以促进合作演化

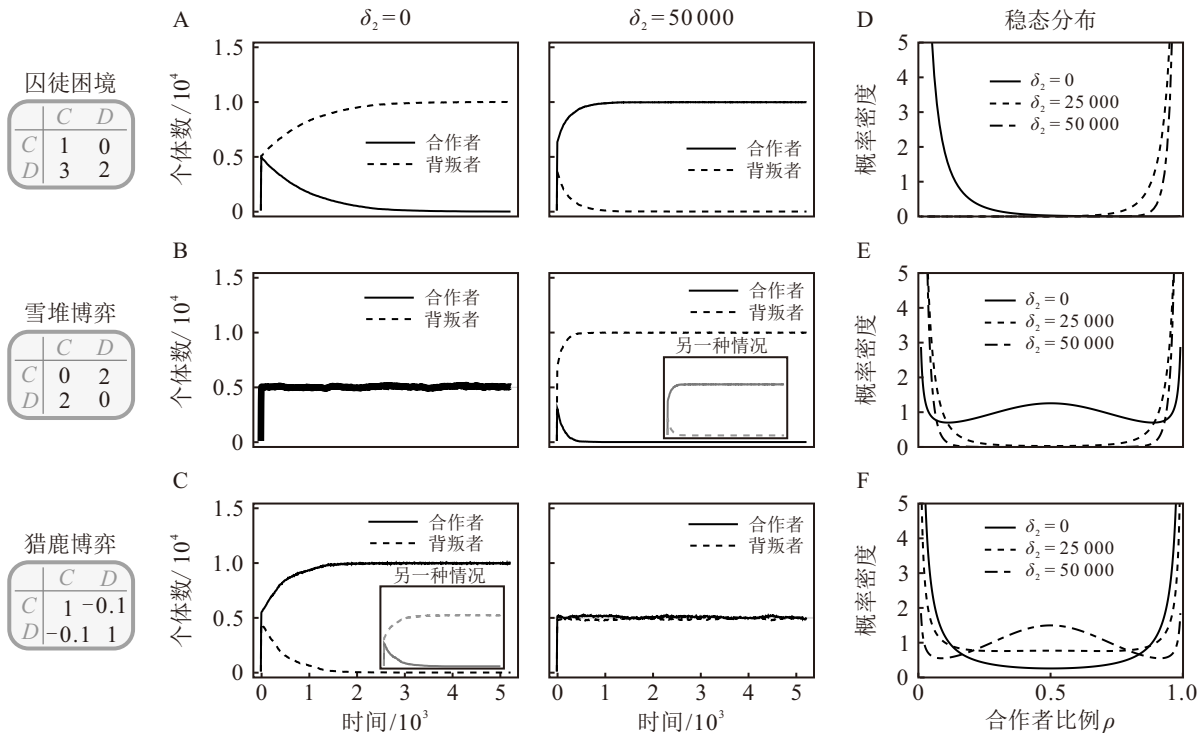


图5 不确定群体规模可以逆转系统演化的方向

对于其他类型的博弈, 不确定的群体规模依然可以逆转演化动力学方向. 在图5中, 考虑3种经典的博弈类型: 囚徒困境、雪堆博弈和猎鹿博弈. 在囚徒困境博弈中, 当条件 $\delta_2 > M$ 不成立或群体规模固定时, 背叛者会在演化中占主导地位. 然而, 当 $\delta_2 > M$ 时, 演化方向会发生逆转, 合作变为占优策略. 在雪堆博弈以及猎鹿博弈中, 当 $\delta_2 > M$ 成立时, 系统的演化方向也发生逆转. 原本稳定的平衡点变为不稳定的平衡点, 而原本不稳定的平衡点变为稳定的平衡点. 例如在雪堆博弈中, 经典的博弈动力学表明系统唯一稳定的平衡点位于内部, 而边界上两点 $p = 1$ 与 $p = 0$ 为不稳定的平衡点. 但是当 $\delta_2 > M$ 时, 演化方向发生了逆转, 这使得内部的平衡点不再稳定, 而边界上两平衡点变得稳定. 该动力学与经典博弈动力学中猎鹿博弈是相同的. 图5中, 子图(A ~ C)为多次数值仿真得到的平均演化样本轨道, 而(D ~ F)为假设了反射性边界所得到的合作者比例的稳态分布. 以上结果可以归纳为当 $\delta_2 > M$ 时, 系统的动力学改变相当于对收益矩阵做了如下变换:

$$\begin{bmatrix} R & S \\ T & P \end{bmatrix} \Rightarrow \begin{bmatrix} -R & -S \\ -T & -P \end{bmatrix}. \quad (46)$$

即当 $\delta_2 > M$ 时, 在左边收益矩阵下的群体动力学等价于右边收益矩阵在经典复制动力学方程下的动力学.

5 总结与展望

在控制理论中, 鲁棒性刻画了系统抵御各种不

确定性的能力,包括物理参数摄动、测量误差、未建模动态等.在博弈动力学系统中,不确定性也广泛存在,不确定性对于博弈系统的动力学的影响以及群体对于不确定性的响应是非常重要的研究课题.

本文刻画了博弈动力学系统中几类典型的不确定性,包括观测不确定性、收益不确定性、群体规模不确定性,这些不确定性都可以显著地改变系统的动力学行为.这表明不确定性对于系统的动力学的影响是不可忽略的,也间接地揭示了群体行为的复杂性.此外,本文还刻画了群体中个体对不确定性的响应,发现了极限环等复杂的动力学现象,这为研究由多个智能体构成的群体系统的鲁棒博弈动力学提供了坚实的理论框架.

值得注意的是,本文所研究的博弈动力学的不确定性和鲁棒性与经典博弈论中不完全信息博弈并不是同一个概念.不完全信息博弈是指参与者不清楚博弈的收益函数、决策方式、决策顺序等信息^[58-59].本文所描述的不确定性是指一个动力学系统中某些参数或者特征是不确定的,重点研究系统的动力学行为.相比于不完全信息博弈,本文的重点在于如何在博弈动力学的框架下建模不确定性,以及如何刻画不确定性对于系统动力学的影响.此外,本文所描述的不确定性与演化博弈动力学中的环境状态变化或者环境切换并不相同.这些研究考虑了环境状态(即收益结构)与群体行为构成的反馈系统^[60-64].尽管在这些模型中环境状态也是变化的,但与本文所考虑的不确定性仍有本质的区别.这些模型中环境状态的变化是由群体行为驱动的,而本文中的不确定性是由系统外部扰动所导致的.

本文对鲁棒博弈动力学进行了初步探索,仍然有许多问题有待进一步深入研究.比如,本文主要探索了均匀混合群体下风险偏好对于系统动力学的影响,结构化的群体已经被广泛研究并证明可以促进合作行为的演化.在结构化群体中考虑不确定性以及个体的风险偏好是否能产生协同效应或者抵消效应仍有待进一步探索.此外,本文主要研究两人博弈,对于多人博弈比如公共品博弈,不确定性对系统的影响也是十分重要的研究课题.最后,描述个体对于不确定性的响应方式并不是唯一的.本文采用了最经典的期望效用理论描述个体对于不确定性的响应,其适用于描述完全理性个体的决策方式,而现实中的个体往往是非完全理性的.对此,Kahneman等^[65]提出了新的理论框架——前景理论描述非理性个体面对风险时的决策方式,并因此成果获得诺贝尔经济学奖.使用该理论描述个体的风险响应并探究

其对博弈动力学的影响也具有重要的意义.

参考文献 (References)

- [1] von Neumann J, Morgenstern O. Theory of games and economic behavior[M]. Princeton: Princeton University Press, 1944.
- [2] Smith J M, Price G R. The logic of animal conflict[J]. *Nature*, 1973, 246(5427): 15-18.
- [3] 王龙, 伏锋, 陈小杰, 等. 演化博弈与自组织合作[J]. *系统科学与数学*, 2007, 27(3): 330-343. (Wang L, Fu F, Chen X J, et al. Evolutionary games and self-organizing cooperation[J]. *Journal of Systems Science and Mathematical Sciences*, 2007, 27(3): 330-343.)
- [4] Wang L, Fu F, Chen X R. Mathematics of multi-agent learning systems at the interface of game theory and artificial intelligence[J]. *Science China Information Sciences*, 2024, 67(6): 166201.
- [5] Weibull J W. Evolutionary game theory[M]. Cambridge: MIT Press, 1997.
- [6] Nowak M A. Evolutionary dynamics: Exploring the equations of life[M]. Cambridge: Harvard University Press, 2006.
- [7] Tuyls K, Parsons S. What evolutionary game theory tells us about multiagent learning[J]. *Artificial Intelligence*, 2007, 171(7): 406-416.
- [8] Mann R P, Helbing D. Optimal incentives for collective intelligence[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2017, 114(20): 5077-5082.
- [9] Wang G C, Su Q, Wang L. Evolution of cooperation with joint liability[J]. *Journal of the Royal Society, Interface*, 2022, 19(188): 20220082.
- [10] Nowak M A. Five rules for the evolution of cooperation[J]. *Science*, 2006, 314(5805): 1560-1563.
- [11] Nowak M A, Sigmund K. A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game[J]. *Nature*, 1993, 364(6432): 56-58.
- [12] Su Q, McAvoy A, Mori Y, et al. Evolution of prosocial behaviours in multilayer populations[J]. *Nature Human Behaviour*, 2022, 6(3): 338-348.
- [13] Allen B, Lippner G, Chen Y T, et al. Evolutionary dynamics on any population structure[J]. *Nature*, 2017, 544(7649): 227-230.
- [14] Ohtsuki H, Hauert C, Lieberman E, et al. A simple rule for the evolution of cooperation on graphs and social networks[J]. *Nature*, 2006, 441(7092): 502-505.
- [15] McAvoy A, Allen B, Nowak M A. Social goods dilemmas in heterogeneous societies[J]. *Nature Human Behaviour*, 2020, 4(8): 819-831.
- [16] Shirado H, Christakis N A. Locally noisy autonomous agents improve global human coordination in network experiments[J]. *Nature*, 2017, 545(7654): 370-374.
- [17] 王龙, 黄锋. 多智能体博弈、学习与控制[J]. *自动化学报*, 2023, 49(3): 580-613. (Wang L, Huang F. An interdisciplinary survey of multi-

- agent games, learning, and control[J]. *Acta Automatica Sinica*, 2023, 49(3): 580-613.)
- [18] Doyle J C, Glover K, Khargonekar P P, et al. State-space solutions to standard H_2 and H_∞ control problems[J]. *IEEE Transactions on Automatic Control*, 1989, 34(8): 831-847.
- [19] Başar T, Bernhard P. H_∞ infinity optimal control and related minimax design problems: A dynamic game approach[M]. Berlin: Springer Science & Business Media, 2008.
- [20] Doyle J C, Francis B A, Tannenbaum A R. Feedback control theory[M]. North Chelmsford: Courier Corporation, 2013.
- [21] Wang L, Huang L. Robustness analysis of discrete systems — A geometric approach[J]. *Chinese Science Bulletin*, 1992, 37(20): 1747-1752.
- [22] Wang L. On strict positive realness of multilinearly parametrized interval systems[J]. *Science in China Series E: Technological Sciences*, 1998, 41(5): 552-560.
- [23] Kharitonov V L. Asymptotic stability of an equilibrium position of a family of systems of linear differential equations[J]. *Differential Equations*, 1977, 14: 1483-1485.
- [24] 王龙, 伏锋, 陈小杰, 等. 复杂网络上的演化博弈[J]. *智能系统学报*, 2007, 2(2): 1-10.
(Wang L, Fu F, Chen X J, et al. Evolutionary games on complex networks[J]. *CAAI Transactions on Intelligent Systems*, 2007, 2(2): 1-10.)
- [25] Minnichelli R J, Anagnost J J, Desoer C A. An elementary proof of Kharitonov's stability theorem with extensions[J]. *IEEE Transactions on Automatic Control*, 1989, 34(9): 995-998.
- [26] Barmish B R. New tools for robustness analysis[C]. *Proceedings of the 27th IEEE Conference on Decision and Control*. Austin, 1988: 1-6.
- [27] Wang L. Composite interval control systems: Some strong kharitonov-like properties[J]. *Reliable Computing*, 2000, 6(3): 231-246.
- [28] Wang L. Kharitonov-like theorems for robust performance of interval systems[J]. *Journal of Mathematical Analysis and Applications*, 2003, 279(2): 430-441.
- [29] Wang L, Wang Z Z, Yu W S, et al. Edge theorem for MIMO systems[J]. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 2003, 50(12): 1577-1580.
- [30] Wang L, Jiang F C, Xie G M, et al. Controllability of multi-agent systems based on agreement protocols[J]. *Science in China Series F: Information Sciences*, 2009, 52(11): 2074-2088.
- [31] Wang L, Xiao F. Finite-time consensus problems for networks of dynamic agents[J]. *IEEE Transactions on Automatic Control*, 2010, 55(4): 950-955.
- [32] 王龙, 丛睿, 李昆. 合作演化中的反馈机制[J]. *中国科学: 信息科学*, 2014, 44(12): 1495-1514.
(Wang L, Cong R, Li K. Feedback mechanism in cooperation evolving[J]. *Scientia Sinica Informationis*, 2014, 44(12): 1495-1514.)
- [33] Knight F H. Risk, uncertainty and profit[M]. Hart: Schaffner and Marx, 1921.
- [34] LeRoy S F, Singell L D. Knight on risk and uncertainty[J]. *Journal of Political Economy*, 1987, 95(2): 394-406.
- [35] Hüllermeier E, Waegeman W. Aleatoric and epistemic uncertainty in machine learning: An introduction to concepts and methods[J]. *Machine Learning*, 2021, 110(3): 457-506.
- [36] der Kiureghian A, Ditlevsen O. Aleatory or epistemic? Does it matter?[J]. *Structural Safety*, 2009, 31(2): 105-112.
- [37] Wienand K, Frey E, Mobilia M. Evolution of a fluctuating population in a randomly switching environment[J]. *Physical Review Letters*, 2017, 119(15): 158301.
- [38] Taitelbaum A, West R, Assaf M, et al. Population dynamics in a changing environment: Random versus periodic switching[J]. *Physical Review Letters*, 2020, 125(4): 048105.
- [39] Ashcroft P, Altrock P M, Galla T. Fixation in finite populations evolving in fluctuating environments[J]. *Journal of the Royal Society*, 2014, 11(100): 20140663.
- [40] Assaf M, Mobilia M, Roberts E. Cooperation dilemma in finite populations under fluctuating environments[J]. *Physical Review Letters*, 2013, 111(23): 238101.
- [41] Constable G W A, Rogers T, McKane A J, et al. Demographic noise can reverse the direction of deterministic selection[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2016, 113(32): E4745-E4754.
- [42] Wang G C, Su Q, Wang L, et al. The evolution of social behaviors and risk preferences in settings with uncertainty[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2024, 121(30): e2406993121.
- [43] Wang G C, Su Q, Wang L. Evolutionary games with payoff perturbations[C]. *The 42nd Chinese Control Conference*. Tianjin, 2023: 8082-8087.
- [44] Wang G C, Su Q, Wang L, et al. Reproductive variance can drive behavioral dynamics[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2023, 120(12): e2216218120.
- [45] Wu B, Altrock P M, Wang L, et al. Universality of weak selection[J]. *Physical Review E, Statistical, Nonlinear, and Soft Matter Physics*, 2010, 82(4): 046106.
- [46] Wu B, García J, Hauert C, et al. Extrapolating weak selection in evolutionary games[J]. *PLoS Computational Biology*, 2013, 9(12): e1003381.
- [47] Ohtsuki H, Nowak M A. The replicator equation on graphs[J]. *Journal of Theoretical Biology*, 2006, 243(1): 86-97.
- [48] Hauert C, Doebeli M. Spatial structure often inhibits the evolution of cooperation in the snowdrift game[J].

- Nature*, 2004, 428(6983): 643-646.
- [49] Taylor P D, Jonker L B. Evolutionary stable strategies and game dynamics[J]. *Mathematical Biosciences*, 1978, 40(1/2): 145-156.
- [50] Hofbauer J, Sigmund K. Evolutionary games and population dynamics[M]. Cambridge, UK: Cambridge University Press, 1998.
- [51] Hofbauer J, Sigmund K. Evolutionary game dynamics[J]. *Bulletin of the American Mathematical Society*, 2003, 40(4): 479-519.
- [52] Arrow K J. Essays in the theory of risk-bearing[M]. Chicago: Markham Publish Co, 1971.
- [53] Pratt J W. Risk aversion in the small and in the large[J]. *Uncertainty in Economics*, 1978: 59-79.
- [54] Nicholson W, Snyder C. Microeconomic theory: basic principles and extensions[M]. Mason: Cengage Learning, 2012.
- [55] Levy H. Absolute and relative risk aversion: An experimental study[J]. *Journal of Risk and Uncertainty*, 1994, 8(3): 289-307.
- [56] Kimball M S. Standard risk aversion[J]. *Econometrica*, 1993, 61(3): 589.
- [57] Wik M, Aragie Kebede T, Bergland O, et al. On the measurement of risk aversion from experimental data[J]. *Applied Economics*, 2004, 36(21): 2443-2451.
- [58] Kajii A, Morris S. The robustness of equilibria to incomplete information[J]. *Econometrica*, 1997, 65(6): 1283.
- [59] Harsanyi J C. Games with incomplete information[J]. *American Economic Review*, 1955, 44(3): 291-303.
- [60] Hilbe C, Šimsa Š, Chatterjee K, et al. Evolution of cooperation in stochastic games[J]. *Nature*, 2018, 559(7713): 246-249.
- [61] Su Q, McAvoy A, Wang L, et al. Evolutionary dynamics with game transitions[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2019, 116(51): 25398-25404.
- [62] Chen X J, Szolnoki A. Punishment and inspection for governing the commons in a feedback-evolving game[J]. *PLoS Computational Biology*, 2018, 14(7): e1006347.
- [63] Wang G C, Su Q, Wang L. Evolution of state-dependent strategies in stochastic games[J]. *Journal of Theoretical Biology*, 2021, 527: 110818.
- [64] Tilman A R, Plotkin J B, Akçay E. Evolutionary games with environmental feedbacks[J]. *Nature Communications*, 2020, 11: 915.
- [65] Kahneman D, Tversky A. Prospect theory: An analysis of decision under risk[J]. *Econometrica*, 1979, 47(2): 263.

作者简介

王龙(1964-),男,教授,博士,博士生导师,主要研究方向为人工智能理论、智能控制理论、演化博弈动力学, E-mail: longwang@pku.edu.cn;

王国丞(1999-),男,博士生,主要研究方向为演化博弈动力学、群体智能与决策, E-mail: gcwang@pku.edu.cn;

苏奇(1991-),男,副教授,博士,主要研究方向为演化博弈理论、网络化系统、群体智能, E-mail: qisu@sjtu.edu.cn.

科研团队简介

王龙教授科研团队立足于北京大学系统与与控制研究中心,长期专注于人工智能理论、博弈控制系统、群体智能和自主无人系统的研究,倡导将前沿性基础研究成果与国民经济发展和国家重大需求紧密结合.团队近年来在控制理论与人工智能交叉科学领域做出了一系列系统性的创新工作,在国内外著名学术期刊(如 *Science*、*PNAS*、*Nature Computational Science*、*Nature Communications*、*IEEE Trans. on Automatic Control* 等)发表论文 100 余篇.先后获得国家自然科学奖、关肇直奖、《中国科学:信息科学》10 年经典论文奖、《控制理论与应用》创刊 30 周年最具影响力论文奖、教育部自然科学一等奖等多项奖励,相关成果被国内外著名学者大量引用和高度评价,在学术界产生了重要影响.

团队带头人王龙教授连续多年入选 Clarivate、Elsevier 高被引学者,曾先后在加拿大多伦多大学电子与计算机工程系、德国宇航院(DLR)机器人与系统动力学研究所进行合作研究,曾获得国家自然科学奖二等奖(2017 年第 1 获奖人)、三等奖(1999 年第 2 获奖人)以及关肇直奖(1994 年)、张嗣瀛奖(2019 年)、杰出青年科学基金、长江学者特聘教授、德国 Alexander von Humboldt 基金、《中国科学:信息科学》10 年经典论文奖(2017 年)、控制理论与应用》创刊 30 周年最具影响力论文奖(2014)、教育部自然科学一等奖(2015 年第 1 获奖人)、北京市优秀研究生指导教师(2022 年)、北京大学优秀博士生导师等多项奖励.王龙教授现任北京人工智能学会副理事长、中国系统仿真学会常务理事、智能物联系统专业委员会主任,《控制与决策》、《智能系统学报》、《控制理论与应用》等杂志编委,国家应用数学中心(四川)学术委员会委员.