

# 基于间接互惠的三策略囚徒困境演化博弈分析

王新颖

辽宁师范大学数学学院, 辽宁 大连

收稿日期: 2024年1月26日; 录用日期: 2024年2月21日; 发布日期: 2024年2月28日

## 摘要

本文首先将间接互惠中的辨别者策略(DIS)加入到经典囚徒困境博弈中, 得到三策略的囚徒困境博弈模型。然后利用复制动态方程及雅可比矩阵的稳定性分析, 探究博弈模型各个均衡点的稳定性。最后得出结论: 均衡点(0,0,1)和均衡点(1,0,0)是渐进稳定的, 进而解决了囚徒博弈困境。

## 关键词

间接互惠, 辨别者策略, 囚徒博弈, 演化稳定性

# Evolutionary Game Analysis of Three Strategies Prisoner's Dilemma Based on Indirect Reciprocity

Xinying Wang

School of Mathematics, Liaoning Normal University, Dalian Liaoning

Received: Jan. 26<sup>th</sup>, 2024; accepted: Feb. 21<sup>st</sup>, 2024; published: Feb. 28<sup>th</sup>, 2024

## Abstract

This article first incorporates the discriminator strategy (DIS) in indirect reciprocity into the classic prisoner's dilemma game and obtains a prisoner's dilemma game model with three strategies; then, using the stability analysis of replicating dynamic equations and Jacobian matrices, explores the stability of various equilibrium points in the game model. Finally, the conclusion is drawn that the equilibrium point (0,0,1) and equilibrium point (1,0,0) are asymptotically stable, thereby solving the prisoner's dilemma.

## Keywords

Indirect Reciprocity, Discriminator Strategy, Prisoner Game, Evolutionary Stability

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

自 2005 年《科学》杂志提出了“合作行为如何演化是 21 世纪最关键的 25 个科学问题之一”以来[1]，研究合作行为的演化稳定性一直是国内外学者试图解决的重要问题。Nowak [2]提出的五大合作机制：亲缘选择、直接互惠、间接互惠、网络互惠和群体选择，其中间接互惠是应用最广泛的合作机制之一。1986 年 Alexander 第一次提出间接互惠这个在人类社会中普遍存在的合作机制[3]。此后国内外学者围绕间接互惠进行了大量研究，在理论与应用方面取得了巨大的发展。E. E. Seidy 和 A. MAlmuntaser 将间接互惠与囚徒困境相结合，给出各个变化后的收益矩阵，并计算出演化稳定策略[4]。E. E. Seidy 随后又将这五种合作形式与鹰鸽博弈相结合，给出各个变化后的收益矩阵，并计算出了 ESS、AD 和 RD 策略[5]。S. Tanabe, H. Suzuki 在间接互惠模型中引入三倍声誉评估体系，得出三倍声誉模型在中性条件下允许形象评分合作[6]。Nowak 和 Sigmund 研究了间接互惠的辨别者策略、总是合作和总是背叛三种策略的复制者动力学[7]。张艳玲[8]等人研究了基于声望的间接互惠是如何促进合作的。孙熠讓等人[9]对现有研究间接互惠的文献进行梳理，并讨论了间接互惠的影响因素。

在众多博弈中，囚徒博弈是最为经典的应用最多的博弈模型，是体现非合作思想的经典案例。J. M. McNamara, Z. Barta 等人探究了行为变化能否促进囚徒困境向合作转化，结果表明外在条件的变化决定着演化博弈的结果[10]。C. Biely, K. Dragosits 等人在网络中模拟了完全理性下的囚徒困境问题，得出幂律度指数分布与通信网络模拟得到的结果完全匹配[11]。于维生利用囚徒困境博弈研究了多人非合作有限纯策略纳什均衡，并将囚徒困境博弈推广至非合作对策模型[12]。

综上，现有关于间接互惠和囚徒困境博弈的研究大多基于双人两策略的博弈模型，对于双参与人三策略囚徒博弈模型的研究较少。然而在现实生活中参与人的策略往往不止一个，因此对于三策略博弈模型的研究是十分必要的。故本文将间接互惠的辨别者策略应用到经典囚徒困境中，构造新的  $3 \times 3$  的收益矩阵，采用演化稳定性分析的方法，通过复制动态方程讨论辨别度  $a$  对演化博弈的影响，得到不同与经典囚徒博弈的演化稳定策略。

## 2. 经典囚徒困境博弈模型

在一般的经典囚徒困境博弈模型中，两个参与人分别有两个策略：合作策略(简记为 C)与背叛策略(简记为 D)，则囚徒博弈的收益矩阵可如下表示。

$$\begin{array}{cc} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{pmatrix} b-c & -c \\ b & 0 \end{pmatrix} \end{array} \quad (1)$$

矩阵中  $b$  代表受助者的收益增加， $c$  代表助人者的收益减少，且有  $b > c > 0$ 。如果合作没有达成，

那么二者收益均不发生变化。由支付矩阵我们可以看出双方合作所得到的收益  $b-c$  明显高于双方背叛时所得到的收益  $0$ ，而当一个参与者选择合作，另一个参与者选择背叛时，背叛者所得到的收益  $b$  最大，合作者所得到的收益为  $c$  最小。根据纳什均衡概念和演化稳定策略的定义可知，(D, D) 是唯一的纳什均衡，且也是演化稳定策略。虽然策略(C, C)得到的收益大于策略(D, D)得到的收益，但当其中一个参与者选择合作时，另一个参与者总趋向于选择背叛从而使自己的收益最大化，即策略(C, C)不满足个人理性要求，或者说是个人理性与集体理性发生冲突。所以策略(C, C)不是纳什均衡，更不是一个演化稳定策略。

为解决这个困境，使博弈双方逃离背叛，我们将间接互惠合作机制中的辨别者策略引入囚徒博弈中，建立三策略的囚徒困境博弈模型。

### 3. 三策略囚徒博弈模型

#### 3.1. 模型的建设

假设 1. 人群中只有使用永远合作(ALLC)、永远背叛(ALLD)和辨别者(DIS)三种策略的人，分别为  $x_1$ 、 $x_2$ 、 $x_3$ ，且  $x_1 + x_2 + x_3 = 1$ ， $x_1, x_2, x_3 \in (0, 1)$ ， $x_1, x_2, x_3$  均为时间  $t$  的函数且连续可导。

假设 2. 辨别者可以辨别对手的信誉，信誉好的选择合作，信誉差的选择背叛， $d$  为其辨别度。

假设 3. 种群之间进行随机博弈， $f_i$  代表参与者  $x_i$  的收益，即适应度。

#### 3.2. 模型的建立

为计算方便我们将经典的  $2 \times 2$  囚徒博弈的支付矩阵进行化简，得到如下的  $2 \times 2$  支付矩阵。

$$\begin{array}{cc} & \text{ALLC} & \text{ALLD} \\ \text{ALLC} & (1 & -1) \\ \text{ALLD} & (2 & 0) \end{array} \quad (2)$$

Nowak 和 Sigmund [13] 提出辨别者策略(DIS)，设  $a$  为其策略的辨别因子，分辨博弈对手的名誉好坏，从而选择合作或背叛。将辨别者策略(DIS)加入到囚徒博弈中建立三策略囚徒博弈模型，进而得到收益矩阵。

$$\begin{array}{ccc} & \text{DIS} & \text{ALLC} & \text{ALLD} \\ \text{DIS} & (1 & 2-a & a-1) \\ \text{ALLC} & (2a-1 & 1 & -1) \\ \text{ALLD} & (2(1-a) & 2 & 0) \end{array} \quad (3)$$

### 4. 模型的演化稳定性分析

在演化博弈中，可以使用收益即适应度来描述选择策略的频率  $x_i$  随时间的动态变化。

适应度分别为：

$$\begin{cases} f_1 = x_1 + x_2(2-a) + x_3(a-1) \\ f_2 = x_1(2a-1) + x_2 - x_3 \\ f_3 = x_1 2(1-a) + 2x_2 \end{cases} \quad (4)$$

由  $\dot{x}_i = f_i - \bar{f}$ ，其中  $\bar{f} = x_1 f_1 + x_2 f_2 + x_3 f_3$ ， $\sum_{i=1}^3 x_i = 1$ ， $x_3 = 1 - x_1 - x_2$ ，可得复制动态方程为：

$$\begin{cases} \frac{dx_1}{dt} = x_1(f_1 - \bar{f}) = x_1(x_1 + 2x_2 - \partial x_2 + \partial x_3 - x_3 - x_2x_3 - x_1x_3 + \partial x_1x_3 - x_1^2 - x_2^2 - x_1x_2 - \partial x_1x_2) \\ \frac{dx_2}{dt} = x_2(f_2 - \bar{f}) = x_2(2\partial x_2 - x_1 + x_2 - x_3 - x_2x_3 - x_1x_3 + \partial x_1x_3 - x_1^2 - x_2^2 - x_1x_2 - \partial x_1x_2) \\ \frac{dx_3}{dt} = x_3(f_3 - \bar{f}) = x_3(2x_1 - 2\partial x_1 + 2x_2 - x_2x_3 - x_1x_3 + \partial x_1x_3 - x_1^2 - x_2^2 - x_1x_2 - \partial x_1x_2) \end{cases} \quad (5)$$

化简降维得:

$$\begin{cases} \dot{x}_1 = x_1(2x_2 - 2ax_2 + a - 1 + x_1x_2 - ax_1^2 + x_1 - 2ax_1x_2) \\ \dot{x}_2 = x_2(-x_1 + 3ax_1 - 1 + x_1x_2 - 2ax_1x_2 + x_2 - x_1^2a) \end{cases} \quad (6)$$

通过求解可得满足要求的均衡点:  $(\frac{1-a}{a}, 0, 2 - \frac{1}{a})$ 、 $(0, 0, 1)$ 、 $(1, 0, 0)$ 、 $(0, 1, 0)$ 。

根据系统的稳定性得到定理 1:

定理 1:  $(0, 0, 1)$  和  $(1, 0, 0)$  为渐进稳定点, 其余均衡点均不稳定。

证明: 非线性系统的雅可比矩阵为:

$$J = \begin{bmatrix} \frac{\partial \dot{x}_1}{\partial x_1} & \frac{\partial \dot{x}_1}{\partial x_2} \\ \frac{\partial \dot{x}_2}{\partial x_1} & \frac{\partial \dot{x}_2}{\partial x_2} \end{bmatrix}$$

其中  $\frac{\partial \dot{x}_1}{\partial x_1} = 2x_2 - 2ax_2 + a - 1 + 2x_1x_2 - 3ax_1^2 + 2x_1 - 4ax_1x_2$ ;

$$\frac{\partial \dot{x}_1}{\partial x_2} = 2x_1 - 2ax_1 + x_1^2 - 2ax_1^2;$$

$$\frac{\partial \dot{x}_2}{\partial x_1} = -x_2 + 3ax_2 + x_2^2 - 2ax_2^2 - 2ax_1x_2;$$

$$\frac{\partial \dot{x}_2}{\partial x_2} = -x_1 + 3ax_1 - 1 + 2x_1x_2 - 4ax_1x_2 + 2x_2 - ax_1^2.$$

计算各均衡点对应的雅可比矩阵的特征值, 当特征值都小于 0 时, 在均衡点处才是渐进稳定。

平衡点  $(\frac{1-a}{a}, 0, 2 - \frac{1}{a})$  对应雅可比矩阵的特征值为  $\lambda_1 = -\frac{1}{a^2} - 4a + 5$ 、 $\lambda_2 = \frac{3}{a^2} + 5a - 2$ 。当

$0 < a < 0.6404$  时,  $\lambda_1 < 0$ , 当  $a > 0.6404$  时,  $\lambda_1 > 0$ , 而  $\lambda_2 > 0$ 。因此平衡点  $(\frac{1-a}{a}, 0, 2 - \frac{1}{a})$  为鞍点或源点。

平衡点  $(0, 0, 1)$  对应雅可比矩阵的特征值为  $\lambda_1 = -1 < 0$ 、 $\lambda_2 = a - 1 < 0$ , 故平衡点  $(0, 0, 1)$  是渐进稳定的。

平衡点  $(1, 0, 0)$  对应雅可比矩阵特征值为  $\lambda_1 = 1 - \frac{2}{a} < 0$ 、 $\lambda_2 = 2a - 2 < 0$ , 故平衡点  $(1, 0, 0)$  是渐进稳定的。

平衡点  $(0, 1, 0)$  对应雅可比矩阵特征值为  $\lambda_1 = 1 > 0$ ,  $\lambda_2 = 1 - a > 0$ , 因此平衡点  $(0, 1, 0)$  为源点, 不是渐进稳定的。

## 5. 结论

通过对三策略囚徒博弈的演化稳定性分析, 得到均衡点  $(0, 0, 1)$  和  $(1, 0, 0)$  为渐进稳定点, 解决了经典囚徒博弈中背叛策略为纳什均衡和演化稳定策略的困境。大多数文章讨论的是两策略的囚徒困境, 本文

的创新点在于将间接互惠这种合作策略衍生出的辨别者策略与囚徒博弈结合，三策略囚徒博弈模型，并将合作机制与演化博弈结合。本文只结合了一种合作机制，而其他四种合作机制能否解决囚徒困境是今后可以研究的方向。

## 参考文献

- [1] Pennisi, E. (2005) How Did Cooperative Behavior Evolve? *Science*, **309**, 93. <https://doi.org/10.1126/science.309.5731.93>
- [2] Nowak, M.A. (2006) Five rules for the Evolution of Cooperation. *Science*, **314**, 1560-1563. <https://doi.org/10.1126/science.1133755>
- [3] Alexander, R.D. (1986) Biology and Law. *Ethology & Sociobiology*, **7**, 167-173. [https://doi.org/10.1016/0162-3095\(86\)90045-2](https://doi.org/10.1016/0162-3095(86)90045-2)
- [4] Seidy, E.E. and Almunaser, A.M. (2015) On the Evolution of Cooperative Behavior in Prisoner's Dilemma. *Journal of Game Theory*, **4**, 1-5. <https://doi.org/10.1155/2015/647246>
- [5] Seidy, E.E. (2016) On the Behavior of Strategies in Hawk-Dove. *Journal of Game Theory*, **211**, 1390-1396.
- [6] Tanabe, S., Suzuki, H. and Masuda, N. (2013) Indirect Reciprocity with Trinary Reputations. *Journal of Theoretical Biology*, **317**, 338-347. <https://doi.org/10.1016/j.jtbi.2012.10.031>
- [7] Nowak, M. and Sigmund, K. (2005) Evolution of Indirect Reciprocity. *Nature*, **437**, 1291-1298. <https://doi.org/10.1038/nature04131>
- [8] 张艳玲, 刘爱志, 孙长银. 间接互惠与合作演化的若干问题研究进展[J]. 自动化学报, 2018, 44(1): 1-12.
- [9] 孙熠譞, 张建华, 李菁萍. 间接互惠理论研究进展[J]. 经济学动态, 2022(1): 146-160.
- [10] Mcnamara, J.M., Barta, Z. and Houston, A.I. (2004) Variation in Behaviour Promotes Cooperation in the Prisoner's Dilemma Game. *Nature*, **428**, 745-748. <https://doi.org/10.1038/nature02432>
- [11] Biely, C., Dragosits, K. and Thurner, S. (2007) The Prisoner's Dilemma on Co-Evolving Networks under Perfect Rationality. *Physica D Nonlinear Phenomena*, **228**, 40-48. <https://doi.org/10.1016/j.physd.2007.02.004>
- [12] 于维生. 非合作对策的纳什均衡求解及囚徒困境问题的推广[C]//中国现场统计研究会. 中国现场统计研究会学术年会, 北京, 1999.
- [13] Nowak, M.A. and Sigmund, K. (1998) The Dynamics of Indirect Reciprocity. *Journal of Theoretical Biology*, **194**, 561-574. <https://doi.org/10.1006/jtbi.1998.0775>