

基于仿真的博弈系统优化策略研究

田 伟, 刘懿芳

(上海理工大学 理学院, 上海 200093)

摘要: 以石头剪刀布博弈系统为例, 提出一种新的理论方法优化该系统, 目的是在不受其他因素影响下最大化玩家获得的收益, 这种新方法即凸优化. 引入非零和矩阵建立凸优化算法模型, 定量地创建了石头剪刀布博弈系统收益方程, 这种方法前人鲜有研究. 创新地提出了博弈系统最优值的临界方程即鞍点方程, 并用强对偶理论证明了该方程的正确性. 重点研究凸优化中的 Newton 算法对石头剪刀布博弈系统进行数据仿真和最大化玩家获得的收益. 仿真结果表明, 数值结果与理论假设相一致, 验证了该方法的可行性和正确性. 该研究对于理解博弈系统 and 应用凸优化具有十分重要的意义.

关键词: 仿真; 凸优化; 博弈系统; 鞍点

中图分类号: O 174.13 **文献标志码:** A

Optimizing Strategy for Game Theory System Based on the Convex Optimization Method

TIAN Wei, LIU Yifang

(College of Science, University of Shanghai for Science and Technology, Shanghai 200093, China)

Abstract: The convex optimization, which serves as a new optimized method, can be fully applied to the game theory system, aiming at maximizing the payoff of the game systems (for all players). The two-person non-zero-sum matrix was introduced to describe the equivalent of payoff in the hypothetical game theory system for making quantitative analyses and getting maximal payoff of the system. A saddle point equation was created to be the critical equation of the optimal value of the game theory system and Newton's algorithm was used to simulate the game model. The simulation results show that, the results are consistent with the theoretical hypothesis values. The method was tested and verified to be feasible and accurate. The work is helpful for deeply understanding the game theory system and reasonably applying the convex optimization.

Keywords: simulation; convex optimization; game theory system; saddle point

石头剪刀布 (rock-paper-scissors, RPS) 作为一个基本的非合作博弈系统^[1], 已广泛用于研究生物

学中的竞争现象, 如生态系统中物种的多样性^[2-7]. 纳什均衡理论^[8-10] 在研究经典博弈论和进化博弈

收稿日期: 2017-05-21

基金项目: 国家自然科学基金资助项目(10874118)

第一作者: 田 伟(1977-), 男, 讲师. 研究方向: 系统分析与集成. E-mail: tianwei@usst.edu.cn

论中起着重要的作用,该理论假设玩家足够理性,目的是为了确​​保玩家能够了解对手的战略,从而优化自身的策略^[1,11-13].很多研究致力于在博弈系统中建立模型获得系统最优,如石头剪刀布博弈系统动力学模型^[14-15]、无标度网络记忆模型^[16]和循环占优模型^[17-19].此外,由于凸优化具有较好的可扩展性和广泛的应用性,常用于解决电子技术^[20-21]和软件工程^[22-25]等方面的优化问题.

目前,对凸优化的研究越来越受到重视,但将其应用于最优化博弈系统的研究鲜有报道.本文提出用凸优化优化博弈系统中玩家的收益,通过在凸优化模型中设定不同的值进行仿真,使系统最大收益处鞍点方程成立,从而证明凸优化的可行性,这在理论研究和实际应用中具有一定的创新性和前瞻性.

1 模型

本文以2人非零和石头剪刀布为例研究非合作博弈系统,系统中每个玩家进行多次回合,每轮玩家可以选择石头(rock, R)、剪刀(scissors, S)、布(paper, P)中的任意一个动作,如图1所示.收益 g ($g>1$)被定义为决定胜负的唯一参数^[1],玩家根据 g 的大小决定选择的动作,如图2所示.当两个玩家(X和Y)选择相同的动作时,玩家X和玩家Y同时取得单位为1的收益;当玩家X击败玩家Y时,玩家X的收益为 g ,玩家Y的收益为0,反之亦然.

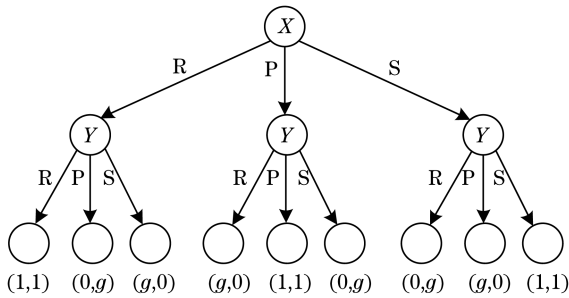


图1 收益树状图

Fig.1 Payoff trees

	R	P	S
R	1	0	g
P	g	1	0
S	0	g	1

图2 收益矩阵

Fig.2 Payoff matrix

2 RPS 博弈系统的凸优化

2.1 凸优化方程的引入

玩家X和玩家Y每轮收益的期望值分别为

$$W_X = \mathbf{X}^T \mathbf{A} \mathbf{Y} \quad (1)$$

$$W_Y = \mathbf{Y}^T \mathbf{A} \mathbf{X} \quad (2)$$

式中: $\mathbf{A} = \begin{bmatrix} 1 & 0 & g \\ g & 1 & 0 \\ 0 & g & 1 \end{bmatrix}$; $\mathbf{X}^T = (\mathbf{X}_R, \mathbf{X}_P, \mathbf{X}_S)$, $\mathbf{X}_R, \mathbf{X}_P, \mathbf{X}_S$ 为玩家X的收益概率; $\mathbf{Y}^T = (\mathbf{Y}_R, \mathbf{Y}_P, \mathbf{Y}_S)$, $\mathbf{Y}_R, \mathbf{Y}_P, \mathbf{Y}_S$ 为玩家Y的收益概率.

与传统的收益方程相比,本文的收益方程是严格凸的.以下为凸优化问题

$$\mathbf{X}_R + \mathbf{X}_P + \mathbf{X}_S = 1; \mathbf{Y}_R + \mathbf{Y}_P + \mathbf{Y}_S = 1 \quad (3)$$

$$\begin{cases} \mathbf{X}_R > 0, \mathbf{X}_P > 0, \mathbf{X}_S > 0 \\ \mathbf{Y}_R > 0, \mathbf{Y}_P > 0, \mathbf{Y}_S > 0 \end{cases} \quad (4)$$

令 $\mathbf{B}_1^T = (1, 0, 0)$, $\mathbf{B}_2^T = (0, 1, 0)$, $\mathbf{B}_3^T = (0, 0, 1)$, 根据事件矩阵将式(3)和式(4)表示为

$$\begin{cases} \sum_i \mathbf{B}_i^T \mathbf{X} = 1, i = 1, 2, 3 \\ \sum_j \mathbf{B}_j^T \mathbf{Y} = 1, j = 1, 2, 3 \end{cases} \quad (5)$$

$$\begin{cases} \mathbf{B}_i^T \mathbf{X} \geq 0, i = 1, 2, 3 \\ \mathbf{B}_j^T \mathbf{Y} \geq 0, j = 1, 2, 3 \end{cases} \quad (6)$$

式中: 玩家X的策略 $i \in \{1, 2, 3\}$; 玩家Y的策略 $j \in \{1, 2, 3\}$.

如上所述,式(1)~(4)具有凸性,因此所有最优解是全局最优解.Boyd等^[26]表示凸优化是一个基本属性,任何局部优化鞍点都是(全局)最优的.进一步假设,式(5)和式(6)严格可行.

2人非零和博弈系统的矩阵形式表示为 $\mathbf{Y}^T \mathbf{A} \mathbf{X} + \mathbf{X}^T \mathbf{A} \mathbf{Y} \neq 0$.为定量研究该系统,假设玩家X先作选择,然后玩家Y根据玩家X的选择再作决定.玩家X想最小化收益 W_Y ,而玩家Y想最大化收益 W_Y .同理,玩家X想最大化收益 W_X ,而玩家Y想最小化收益 W_X .

玩家X最佳策略是利用变量 $\mathbf{X}$ 最小化收益 $\mathbf{Y}^T \mathbf{A} \mathbf{X}$

$$\min_{\mathbf{X}} \mathbf{Y}^T \mathbf{A} \mathbf{X} \quad (7)$$

而玩家Y利用变量 $\mathbf{Y}$ 最大化收益

$$\max_{\mathbf{Y}} (\min_{\mathbf{X}} \mathbf{Y}^T \mathbf{A} \mathbf{X}) \quad (8)$$

$$\begin{aligned} Y^T A X &= (Y_R, Y_P, Y_S) \begin{pmatrix} 1 & 0 & g \\ g & 1 & 0 \\ 0 & g & 1 \end{pmatrix} \begin{pmatrix} X_R \\ X_P \\ X_S \end{pmatrix} = \\ &= (Y_R + gY_P, Y_P + gY_S, gY_R + Y_S) \begin{pmatrix} X_R \\ X_P \\ X_S \end{pmatrix} = \\ &= Y_R X_R + Y_P X_P + Y_S X_S + g(Y_R X_S + Y_P X_R + Y_S X_P) = \\ &= (Y_R + gY_P)X_R + (Y_P + gY_S)X_P + (Y_S + gY_R)X_S \quad (9) \end{aligned}$$

假设 $\chi = Y_R + gY_P$, $\xi = Y_P + gY_S$, $\mu = Y_S + gY_R$. X_R, X_P, X_S 表示收益概率, χ, ξ, μ 表示收益值.

当 $\chi < \xi, \chi < \mu$ 时, 玩家 X 通过最大化变量 X_R (χ 的系数) 来最小化收益 $Y^T A X$. 因此, 当 $X_R = 1, X_P = 0, X_S = 0$ 时, $Y^T A X$ 收益最小, 根据式(8)得 $e_1^T = (X_R, X_P, X_S) = (1, 0, 0)$. 同理, 当 $\xi < \chi, \xi < \mu$ 时, 玩家 X 令 $X_P = 1, X_R = 0, X_S = 0$ 来最小化收益 $Y^T A X$, 得到 $e_2^T = (X_R, X_P, X_S) = (0, 1, 0)$; 如果 $\mu < \chi, \mu < \xi$, 玩家 X 最优的方式是令 $X_S = 1, X_P = 0, X_R = 0$ 来最小化收益 $Y^T A X$, 得到 $e_3^T = (X_R, X_P, X_S) = (0, 0, 1)$.

对于 $\max_Y (\min_X Y^T A X)$, 内部优化为

$$\min_X Y^T A X = Y^T A e_i, i = 1, 2, 3 \quad (10)$$

式中: i 为确定性策略; e_i 为第 i 个位置为 1, 其他位置全为零的矩阵.

凸优化函数标准形式为

$$\begin{cases} \min & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, i = 1, \dots, m \\ & h_i(x) = 0, i = 1, \dots, p \end{cases} \quad (11)$$

引入变量 α 表示内部最小化的值

$$\begin{cases} \max & \alpha \\ \text{subject to} & \alpha \leq Y^T A e_i, i = 1, 2, 3 \\ & \sum_j B_j^T Y = 1, j = 1, 2, 3 \\ & B_j^T Y \geq 0, j = 1, 2, 3 \end{cases} \quad (12)$$

矩阵向量表示为

$$\begin{cases} \max & \alpha \\ \text{subject to} & \alpha \mathbf{1} - A Y \leq 0 \\ & \mathbf{1}^T Y = 1 \\ & B^T Y \geq 0 \end{cases} \quad (13)$$

式中: $\alpha = \min_X Y^T A X$; $\mathbf{1}^T = (1, 1, 1)$; $\mathbf{0}^T = (0, 0, 0)$;

$$B = (B_1, B_2, B_3) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}; \geq \text{和} \leq \text{为广义不等式}; \equiv \text{为广义等式}.$$

2.2 鞍点方程

假设玩家顺序颠倒: 玩家 Y 先选择 Y_R, Y_P, Y_S ; 玩家 X 再选择 X_R, X_P, X_S . 如果双方玩家遵循最佳策略, 玩家 X 应该利用变量 X 最小化收益 $\max_Y Y^T A X$, 从而系统收益表示为

$$\min_X \max_Y Y^T A X$$

相当于

$$\begin{cases} \min & \beta \\ \text{subject to} & \beta \mathbf{1} - A X \geq 0 \\ & \mathbf{1}^T X = 1 \\ & B^T X \geq 0 \end{cases} \quad (14)$$

式中: $\beta = \max_Y Y^T A X$; $\mathbf{1}^T = (1, 1, 1)$; $\mathbf{0}^T = (0, 0, 0)$;

$$B = (B_1, B_2, B_3) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

式(13)和式(14)分别为玩家 X 和 Y 的收益, 两式为对偶关系(在博弈系统中玩家 Y 和玩家 X 是对偶关系).

最初假设玩家 X 和 Y 有先后顺序, 实际上是玩家 X 和 Y 同时作决定. 在此基础上, 结合式(13)和式(14), 有

$$\max_Y \min_X Y^T A X = \min_X \max_Y Y^T A X \quad (15)$$

式(15)称为鞍点方程, 用于描述玩家获得最大收益时的鞍点. 式(13)和式(14)分别为式(15)的两边.

同理, 玩家 X 的收益用鞍点方程表示为

$$\max_X \min_Y X^T A Y = \min_Y \max_X X^T A Y \quad (16)$$

用于描述系统收益最大值. 其中式(16)的两边分别用式(17)和式(18)表示.

$$\begin{cases} \max & \delta \\ \text{subject to} & \delta \mathbf{1} - A X \leq 0 \\ & \mathbf{1}^T X = 1 \\ & B^T X \geq 0 \end{cases} \quad (17)$$

$$\begin{cases} \min & \gamma \\ \text{subject to} & \gamma \mathbf{1} - A Y \geq 0 \\ & \mathbf{1}^T Y = 1 \\ & B^T Y \geq 0 \end{cases} \quad (18)$$

3 算法

3.1 Newton 算法

本文提出的鞍点方程是用于表示系统收益最优的临界方程, 为验证其正确性, 利用 Newton 算法证明.

对于 $z \in \text{dom } f$, 称向量

$$\Delta \mathbf{z}_{nt} = -\nabla^2 f(\mathbf{z})^{-1} \nabla f(\mathbf{z}) \quad (19)$$

为(f 在 $\mathbf{z}$ 处的)Newton步径^[26].由 $\nabla^2 f(\mathbf{z})$ 的正定性可知,除非 $\nabla f(\mathbf{z})=0$,否则就有

$$\begin{aligned} \nabla f(\mathbf{z})^T \Delta \mathbf{z}_{nt} = \\ -\nabla f(\mathbf{z})^T \nabla^2 f(\mathbf{z})^{-1} \nabla f(\mathbf{z}) < 0 \end{aligned} \quad (20)$$

因此 Newton 步径为下降方向.

函数 f 在 $\mathbf{z}$ 处的二阶 Taylor 近似(或模型) $\hat{f}$ 为

$$\begin{aligned} \hat{f}(\mathbf{z} + \mathbf{v}) = f(\mathbf{z}) + \nabla f(\mathbf{z})^T \mathbf{v} + \\ \frac{1}{2} \mathbf{v}^T \nabla^2 f(\mathbf{z}) \mathbf{v} \end{aligned} \quad (21)$$

这是 $\mathbf{v}$ 的二次凸函数,在 $\mathbf{v} = \Delta \mathbf{z}_{nt}$ 处达到最小值.因此,将 $\mathbf{z}$ 加上 Newton 步径 $\Delta \mathbf{z}_{nt}$ 能够极小化 f 在 $\mathbf{z}$ 处的二阶近似^[26].

凸优化问题

$$\begin{cases} \min \mathbf{Y}^T \mathbf{A} \mathbf{X} \\ \text{subject to } \mathbf{C} \mathbf{X} = \mathbf{D} \end{cases} \quad (22)$$

$$\begin{cases} \min f(\mathbf{z} + \mathbf{v}) = f(\mathbf{z}) + \\ \nabla f(\mathbf{z})^T \mathbf{v} + \frac{1}{2} \mathbf{v}^T \nabla^2 f(\mathbf{z}) \mathbf{v} \\ \text{subject to } \mathbf{C}(\mathbf{z} + \mathbf{v}) = \mathbf{D} \end{cases} \quad (23)$$

可将 Newton 方向 $\Delta \mathbf{z}_{nt}$ 及其相关向量 $\mathbf{w}$ 解释为最优性条件

$$\begin{aligned} \mathbf{C} \mathbf{z}^* = \mathbf{D}, \nabla f(\mathbf{z}^*) + \mathbf{C}^T \mathbf{z}^* = \\ \mathbf{A}^T \mathbf{y} + \mathbf{C}^T \mathbf{v}^* = 0 \end{aligned} \quad (24)$$

的线性近似方程组的解.用 $\mathbf{z} + \Delta \mathbf{z}_{nt}$ 代替 $\mathbf{z}^*$,用 $\mathbf{w}$ 代替 $\mathbf{v}^*$,并将第二个方程中的梯度项换成其在 $\mathbf{z}$ 附近

的线性近似,从而得到

$$\begin{aligned} \mathbf{C}(\mathbf{z} + \Delta \mathbf{z}_{nt}) = \mathbf{D}, \nabla f(\mathbf{z} + \Delta \mathbf{z}_{nt}) + \\ \mathbf{C}^T \mathbf{w} \approx \mathbf{A}^T \mathbf{y} + 0 + \mathbf{C}^T \mathbf{w} = 0 \end{aligned} \quad (25)$$

利用 $\mathbf{C} \mathbf{X} = \mathbf{D}$,以上方程变为

$$\mathbf{C} \Delta \mathbf{z}_{nt} = 0, \mathbf{C}^T \mathbf{w} \approx -\mathbf{A}^T \mathbf{y} \quad (26)$$

Newton 方向 $\Delta \mathbf{z}_{nt}$ 由以下方程确定

$$\begin{bmatrix} 0 & \mathbf{C}^T \\ \mathbf{C} & 0 \end{bmatrix} \begin{bmatrix} \Delta \mathbf{z}_{nt} \\ \mathbf{w} \end{bmatrix} = \begin{bmatrix} -\mathbf{A}^T \mathbf{y} \\ 0 \end{bmatrix} \quad (27)$$

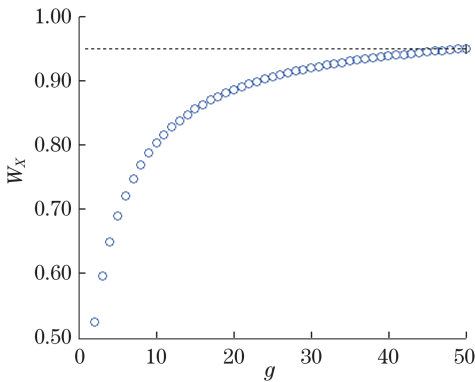
式中, $\mathbf{w}$ 是该二次问题的最优对偶向量.

Newton 方向 $\Delta \mathbf{z}_{nt}$ 和 Newton 减量 $\lambda(\mathbf{z})$ 分别为

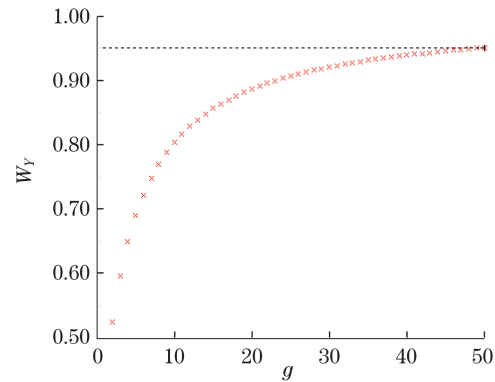
$$\begin{cases} \Delta \mathbf{z}_{nt} = -\nabla^2 f(\mathbf{z})^{-1} \nabla f(\mathbf{z}) \\ \lambda(\mathbf{z}) = (\Delta \mathbf{z}_{nt}^T \nabla^2 f(\mathbf{z}) \Delta \mathbf{z}_{nt})^{1/2} \end{cases} \quad (28)$$

3.2 验证理论

对于鞍点方程 $\max_Y \min_X \mathbf{Y}^T \mathbf{A} \mathbf{X} = \min_X \max_Y \mathbf{Y}^T \mathbf{A} \mathbf{X}$, 参数 $\mathbf{A}$ 因 g 的变化而变化,但它不受其他因素影响,如图 3 所示,其中 EPR(the expected payoff per round)指每轮游戏的预期回报, $g_1 = (g + 3.7)/3$. 赋予 $\mathbf{A}$ 两个不同值,当 $\mathbf{A}_1 = [1, 0, 2.7; 2.7, 1, 0; 0, 2.7, 1]$, 利用 CVX (convex optimization) 软件仿真得到鞍点方程左边为 $\max_Y \min_X \mathbf{Y}^T \mathbf{A} \mathbf{X} = +1.233\ 33$, 鞍点方程右边为 $\min_X \max_Y \mathbf{Y}^T \mathbf{A} \mathbf{X} = +1.233\ 33$; 当 $\mathbf{A}_2 = [1, 0, 4.1; 4.1, 1, 0; 0, 4.1, 1]$, 仿真得到鞍点方程左边为 $\max_Y \min_X \mathbf{Y}^T \mathbf{A} \mathbf{X} = +1.7$, 鞍点方程右边为 $\min_X \max_Y \mathbf{Y}^T \mathbf{A} \mathbf{X} = +1.7$. 因此,本文的仿真结果与理论假设是相契和的.



(a) 玩家X每轮预期收益的最优值



(b) 玩家Y每轮预期收益的最优值

图3 W_X 为玩家X获得的最大回报值, W_Y 为玩家Y获得的最大回报值(单位为 g_1).

Fig.3 Biggest gain of player X is W_X , the biggest gain of player Y is W_Y (in units of g_1).

4 结 论

本文旨在利用凸优化方法得到博弈系统的最大

收益, $\mathbf{A} = [1, 0, g; g, 1, 0; 0, g, 1]$ 作为非零和收益矩阵, 是玩家量化收益的主要参数; $\mathbf{X}^T \mathbf{A} \mathbf{Y}$ 和 $\mathbf{Y}^T \mathbf{A} \mathbf{X}$ 作为描述玩家 X 和 Y 的收益矩阵, 是系统获得最大收益时鞍点方程 $\max_Y \min_X \mathbf{Y}^T \mathbf{A} \mathbf{X} = \min_X \max_Y \mathbf{Y}^T \mathbf{A} \mathbf{X}$ 的重要组

成部分. Newton 算法模拟博弈系统的重点在于设 α 为目标函数, $\alpha \mathbf{1} - \mathbf{A}\mathbf{Y} \leq 0, \mathbf{1}^T \mathbf{Y} = 1, \mathbf{B}^T \mathbf{Y} \geq 0$ 为约束函数. 仿真结果表明 W_X, W_Y 随着收益参数 g 的增加而增加. 此外, 为证明鞍点方程的有效性, 在 2 人非零和 RPS 博弈系统中给 $\mathbf{A}$ 设定两个不同值. 当 $\mathbf{A}_1 = [1, 0, 2.7; 2.7, 1, 0; 0, 2.7, 1]$, 仿真结果为 $\max_Y \min_X \mathbf{Y}^T \mathbf{A} \mathbf{X} = \min_X \max_Y \mathbf{Y}^T \mathbf{A} \mathbf{X} = +1.233\ 33$; 当 $\mathbf{A}_2 = [1, 0, 4.1; 4.1, 1, 0; 0, 4.1, 1]$, 仿真结果为 $\max_Y \min_X \mathbf{Y}^T \mathbf{A} \mathbf{X} = \min_X \max_Y \mathbf{Y}^T \mathbf{A} \mathbf{X} = +1.7$. 仿真结果证明, 在博弈系统中采用凸优化来最大化收益的方法是正确的.

参考文献:

- [1] WANG Z J, XU B, ZHOU H J. Social cycling and conditional responses in the rock-paper-scissors game [J]. *Scientific Reports*, 2014, 4: 5830.
- [2] BIERNASKIE J M, GARDNER A, WEST S A. Multicoloured greenbeards, bacteriocin diversity and the rock-paper-scissors game [J]. *Journal of Evolutionary Biology*, 2013, 26(10): 2081–2094.
- [3] LAIRD R A. Population interaction structure and the coexistence of bacterial strains playing ‘rock-paper-scissors’ [J]. *Oikos*, 2014, 123(4): 472–480.
- [4] LOERTSCHER S. Rock-scissors-paper and evolutionarily stable strategies [J]. *Economics Letters*, 2013, 118(3): 473–474.
- [5] HE Q, MOBILIA M, TÄUBER U C. Spatial rock-paper-scissors models with in homogeneous reaction rates [J]. *Physical Review E*, 2010, 82(5): 051909.
- [6] SINERVO B, HEULIN B, SURGET-GROBA Y, et al. Models of density-dependent genic selection and a new rock-paper-scissors social system [J]. *The American Naturalist*, 2007, 170(5): 663–680.
- [7] KERR B, RILEY M A, FELDMAN M W, et al. Local dispersal promotes biodiversity in a real-life game of rock-paper-scissors [J]. *Nature*, 2002, 418(6894): 171–174.
- [8] BI Z D, ZHOU H J. Optimal cooperation-trap strategies for the iterated rock-paper-scissors game [J]. *PLoS One*, 2014, 9(10): e111278.
- [9] ZHOU H J. The rock-paper-scissors game [J]. *Contemporary Physics*, 2016, 57(2): 151–163.
- [10] AUMANN R J, BRANDENBURGER A. Epistemic conditions for Nash equilibrium [J]. *Econometrica*, 1995, 63(5): 1161–1180.
- [11] 汪筱阳, 吴德伟, 戴传金. 基于纳什博弈论的功率控制策略及其牛顿迭代算法 [J]. *电子设计工程*, 2013, 21(1): 74–76.
- [12] BAHREL E, HALLER H. Cycles with undistinguished actions and extended rock-paper-scissors games [J]. *Economics Letters*, 2013, 120(3): 588–591.
- [13] BAHREL E. Rock-paper-scissors and cycle-based games [J]. *Economics Letters*, 2012, 115(3): 401–403.
- [14] WESSON E, RAND R. Hopf bifurcations in delayed rock-paper-scissors replicator dynamics [J]. *Dynamic Games and Applications*, 2016, 6(1): 139–156.
- [15] SEMMANN D, KRAMBECK H J, MILINSKI M. Volunteering leads to rock-paper-scissors dynamics in a public goods game [J]. *Nature*, 2003, 425(6956): 390–393.
- [16] LUBASHEVSKY I, KANEMOTO S. Scale-free memory model for multiagent reinforcement learning. Mean field approximation and rock-paper-scissors dynamics [J]. *The European Physical Journal B*, 2010, 76(1): 69–85.
- [17] MOBILIA M. Oscillatory dynamics in rock-paper-scissors games with mutations [J]. *Journal of Theoretical Biology*, 2010, 264(1): 1–10.
- [18] VERMA G, CHAN K, SWAMI A. Zealotry promotes coexistence in the rock-paper-scissors model of cyclic dominance [J]. *Physical Review E*, 2015, 92(5): 052807.
- [19] SZOLNOKI A, PERC M. Zealots tame oscillations in the spatial rock-paper-scissors game [J]. *Physical Review E*, 2016, 93(6): 062307.
- [20] 曾勇. 基于博弈论的无线通信抗干扰关键技术研究 [D]. 成都: 电子科技大学, 2014.
- [21] JOSHI S, BOYD S. Sensor selection via convex optimization [J]. *IEEE Transactions on Signal Processing*, 2009, 57(2): 451–462.
- [22] 唐璜, 毛璐. 基于博弈论的软件动态调控策略的研究与实现 [J]. *信息与电脑*, 2012(2): 50–51.
- [23] MATTINGLEY J, BOYD S. Cvxgen: a code generator for embedded convex optimization [J]. *Optimization and Engineering*, 2012, 13(1): 1–27.
- [24] WANG T, JOBREDEAUX R, PANTEL M, et al. Credible autocoding of convex optimization algorithms [J]. *Optimization and Engineering*, 2016, 17(4): 781–812.
- [25] 刘海蛟, 赵书田. 一种关联博弈的软件调度线性规划控制算法 [J]. *微电子学与计算机*, 2016, 33(7): 121–124.
- [26] BOYD S, VANDENBERGHE L. Convex optimization [J]. *IEEE Transactions on Automatic Control*, 2006, 51(11): 1859–1859.

(编辑: 丁红艺)