

复制子动力学的反馈线性化：一个控制框架以促进演化博弈的收敛

Adil Faisal

2025 年 8 月 17 日

摘要

本文展示了反馈线性化首次应用于复制动态中的情况，推动了非收敛进化博弈向具有全局渐近稳定性的系统的演化。尽管复制动态是进化博弈理论的基础，但在纳什均衡点 [2] 具有中性稳定性，这导致进化过程在没有收敛到最优策略的情况下振荡。我们建立了一个控制论框架来消除复制动态中的非线性成分，然后应用一个线性反馈组件迫使在纳什均衡处发生策略变化。通过李雅普诺夫分析，我们展示了从概率单形中任何初始条件下的全局收敛。我们以点球大战游戏的数值示例说明了这种方法，其中我们的方法指导策略快速达到混合纳什均衡，而未控制的动力学则振荡。我们的工作作为非线性控制理论与进化博弈动力学之间已知的第一个联系之一，在多智能体系统、算法交易和战略优化等领域具有应用。

关键词： 博弈论，李雅普诺夫稳定性，复制动态，纳什均衡，反馈线性化，非线性控制，演化博弈

1 介绍

复制子动态是进化博弈理论中的一个基础模型，描述了在竞争环境中策略的频率如何根据其适应度发生变化 [5]。虽然复制子动态在建模生物和经济系统中的策略演化时通常非常成功，但复制子动态也有一个根本性的缺陷：纳什均衡通常是中性稳定的 [2]。因此，策略围绕最优策略振荡而不是收敛。更具体地说，这些振荡阻止了代理人向最佳玩法的收敛，并导致战略互动中的长期动态不可预测。

本文通过将完全状态反馈线性化应用于复制动态，解决了收敛问题，这标志着该非线性控制方法首次进入博弈演化。此方法将复制动态的不收敛动态转化为一个能够实现全局渐近稳定的系统，使得代理可以从复制动态中的任何初始条件可靠地收敛到纳什均衡策略。我们的方法消除了导致振荡动力学的非线性项，并应用线性反馈控制基本上引导策略向纳什均衡靠拢。通过使用李雅普诺夫理论，我们分析证明了所提出的控制律将使概率单形中的某一点实现全局渐近稳定。借助此控制律，我们可以确保所有代理在博弈设置中都将收敛到纳什均衡，而标准的复制动态则无法保证收敛。

我们通过前锋与守门员的点球大战来说明我们的方法，使用模拟展示这种方法与具有振荡的未控制复制动态在收敛到平衡状态方面的差异。

这项工作代表了控制理论和演化博弈论交叉领域的一个新方向，对多智能体学习、算法交易、网络中的资源分配以及任何试图利用演化型动力学实现可靠且可预测地收敛到平衡的战略环境都有更广泛的影响。

2 游戏设置和背景

玩家之间的游戏设定为零和博弈。这意味着游戏中总的利益之和为零，也就是说，如果一个玩家赢了，另一个玩家就必须输。在点球大战的背景下，这个假设是成立的，因为如果前锋进球，守门员就会失败，反之亦然。游戏也是同时进行的，因为守门员和球员必须同时做出决定 [1]。每个前锋和守门员都有三个方向可以选择射门或扑救：左、右和中路。因此，这个游戏是一个混合策略博弈。由于这是一个混合策略博弈，可以利用复制动态来分析随着时间演变的代理人的策略。

纳什均衡 (NE) 定义为一旦实现的结果，意味着没有一个玩家可以通过单方面改变他们的决策来增加自己的收益 [1]。在这个点球大战的背景下，由于比赛是非合作的，这意味着比赛中至少有一个点对前锋和守门员来说都是纳什均衡点 [2]。

3 游戏动力学

$$A = \begin{bmatrix} 0 & 0.8 & 0.7 \\ 0.9 & 0 & 0.2 \\ 0.75 & 0.45 & 0 \end{bmatrix} \quad (1)$$

A 是为进攻球员定义的收益矩阵。为了考虑点球的随机性，所有收益值均设定在区间 $[0, 1]$ 内，使得每个收益值代表给定射门方向和守门员扑救方向 [6] 的进球概率。如果球员向左侧射门而守门员留在中心位置，则进攻球员的收益为 0.8。

这允许在处罚过程中考虑射门的失误或偏转，因为矩阵 A 中的每个值都是给定球员做出任何决策组合下进球的概率。游戏中的一个关键假设是，如果前锋的射门方向和守门员的扑救方向相同，则收益为零，因此双方都没有输赢。

守门员的季后赛矩阵 B 由方程 (2) 定义，因为比赛是不对称的，这意味着角色不能互换，因为前锋的任务是将球踢入网中，而守门员的策略则包括向一侧扑救。因此，这个点球场景被定义为一个“非对称零和”游戏。

$$B = -A^T \quad (2)$$

为了分析游戏动态以及玩家策略如何演变，使用了复制动态来分析该系统。复制动态涉及分析随连续时间演化且随着时间推移代理之间学习的混合策略游戏。由于该游戏是零和博弈，总收益为零 [4]，这意味着：

$$x^T A y + y^T B x = 0 \quad (3)$$

由于我们的游戏是两人游戏，复制动态可以定义为 [4]：

$$\dot{x}_i = x_i (\{A y\}_i - x^T A y) \quad (4)$$

$$\dot{y}_j = y_j (\{B x\}_j - y^T B x) \quad (5)$$

方程 (4) 涵盖了前锋的复制动态及其策略如何演化。 $\{A y\}_i$ 是前锋采用策略 i 的收益，给定守门员使用混合策略 y 和复制动态比较纯策略的收益与使用混合策略 x 所获得的平均收益，后者用 $x^T A y$ 表示。方程 (5) 表示相同的内容，但以守门员的角度来看， $\{B x\}_j$ 表示守门员当前策略相对于前锋策略组合 x 的收益和 $y^T B x$ 是守门员的平均收益。因此，这些动态使得超出平均收益的策略更加盛行而表现不佳的策略逐渐消失。这使玩家能够根据对方调整自己的策略以优化他们的收益。对于这个游戏， x 和 y 是 3×1 列向量，分别表示前锋和守门员策略的概率分布，定义如下：

$$i \in \{\text{kickleft}, \text{kickcenter}, \text{kickright}\} = \{1, 2, 3\}$$

$$j \in \{\text{diveleft}, \text{divecenter}, \text{diveright}\} = \{1, 2, 3\}$$

由于该游戏是非合作的，因此至少存在一个纳什均衡，这对我们的系统是成立的。系统的纳什均衡被定义为单方面改变策略不会使 [2] 受益的点。对于复制动态来说，如果代理认为它们各自的收益已被优化，则混合策略 x^* 和 y^* 是纳什均衡点。在零和博弈中，纳什策略也被称为最大化最小化策略。以下表达式描述了纳什均衡 [4] 的标准：

$$\forall i \in [n]: (Ay^*)_i = x^{*T} Ay^* \quad (6)$$

$$\forall j \in [n]: (Bx^*)_j = y^{*T} Bx^* \quad (7)$$

通过将所有可能策略的方程 (4) 和 (5) 设为 0, 使得 $\frac{dx_i}{dt} = 0$ 和 $\frac{dy_j}{dt} = 0$, 可以计算并比较均衡策略混合, 并确定所获得的策略是否是纳什均衡。纯策略也可以是方程 (6) 和 (7) 中的不动点; 然而, 这些很容易被其他玩家利用, 因此不被视为博弈中的纳什均衡。

4 李雅普诺夫稳定性分析

由于我们的点球游戏是动态和非线性的, 可以使用合适的李雅普诺夫候选函数来分析我们的双智能体系统。从现有文献来看, KL 散度可以用作定义如下所示的李雅普诺夫函数 [5]:

$$V(x, y) = \sum_{i=1}^3 x_i^* \ln \left(\frac{x_i^*}{x_i} \right) + \sum_{j=1}^3 y_j^* \ln \left(\frac{y_j^*}{y_j} \right) \quad (8)$$

使用李雅普诺夫函数的稳定性准则 [7]:

- $\frac{dV}{dt} \leq 0$: 系统是稳定的
- $\frac{dV}{dt} < 0$: 系统是渐近稳定的
- $\frac{dV}{dt} > 0$: 系统不稳定

KL 散度衡量了当前策略组合与纳什均衡之间的差异。它满足以下条件:

- $V(x^*, y^*) = 0$ 是该函数的唯一零点
- $V(x, y) > 0$ 对所有 $x \neq x^*, y \neq y^*$ 成立

纳什均衡时, 时间导数方程 (8) 的结果为 0, 这意味着纳什均衡是中性稳定的, 即如果系统从该平衡点开始, 则会保持在该处; 但如果从平衡点附近开始, 策略的轨迹将在平衡点周围振荡。证明如下:

证明 1: 中性稳定

$$\begin{aligned}
\dot{V} &= -\sum_{i=1}^3 \frac{x_i^* (\dot{x}_i)}{x_i} - \sum_{j=1}^3 \frac{y_j^* (\dot{y}_j)}{y_j} \\
\dot{x}_i &= x_i (\{Ay\}_i - x^T Ay) \text{ and } \dot{y}_j = y_j (\{Bx\}_j - y^T Bx) \\
\dot{V} &= -\left(\sum_{i=1}^3 x_i^* \frac{(x_i (\{Ay\}_i - x^T Ay))}{x_i} + \sum_{j=1}^3 y_j^* \frac{(y_j (\{Bx\}_j - y^T Bx))}{y_j} \right) \\
\dot{V} &= -\left(\sum_{i=1}^3 x_i^* (\{Ay\}_i - x^T Ay) + \sum_{j=1}^3 y_j^* (\{Bx\}_j - y^T Bx) \right) x^T Ay + y^T Bx = 0 \\
\dot{V} &= -\left(\sum_{i=1}^3 x_i^* (\{Ay\}_i) + \sum_{j=1}^3 y_j^* (\{Bx\}_j) \right) B = -A^T
\end{aligned}$$

$\dot{V} = -(x^* - x)^T A(y - y^*)$ For Nash equilibrium $(x, y) = (x^*, y^*)$, $\dot{V} = 0$ neutrally stable

接近平衡: $\dot{V} = \delta x^T A \delta y \rightarrow \delta x \approx 0$ 和 $\delta y \approx 0$, 然后是 $\dot{V} \approx 0$

这意味着如果系统从平衡点开始, 它将保持在那里; 否则, 它会在平衡点附近振荡, 但永远不会收敛到该平衡点, 前提是初始条件足够接近平衡点 [7]。由于收益矩阵 A 是非对称的 (即 $A \neq A^T$), 但它是实矩阵, 则需要对该矩阵进行平均处理以使其对称, 从而使用 Johnson 的方法分析实矩阵 A 的定性, 如下所示 [3]:

$$A_s = \frac{1}{2}(A + A^T)$$

这导致了以下更新后的矩阵 A_s :

$$A_s = \begin{bmatrix} 0 & 0.85 & 0.725 \\ 0.85 & 0 & 0.325 \\ 0.725 & 0.325 & 0 \end{bmatrix}$$

该矩阵是对称的, 因此通过寻找特征值可以确定矩阵的正定性。特征值是 $\lambda_1 = 1.29, \lambda_2 = -0.970, \lambda_3 = -0.320$ 。由于特征值既有正数也有负数, 所以矩阵被判定为不定, 并且不能推断出系统在平衡点附近轨迹的行为。如果初始轨迹保持接近平衡状态, 即 $\delta x \approx 0$ 和 $\delta y \approx 0$, 那么 $\dot{V} \approx 0$, 导致系统呈现出中性稳定。这意味着如果系统从平衡点开始, 策略的混合不会收敛到纳什均衡; 相反, 游戏会演变成一种不断智胜对手和被对手赶上循环的状态, 从而使系统的动态呈现周期性。

5 模拟与结果

使用背景和游戏设置，该问题在 Python 中进行了模拟（代码附在这个项目报告中）。系统的纳什均衡如下：

$$x^* = [0.503, 0.422, 0.075], y^* = [0.416, 0.276, 0.308]$$

这表明射手更喜欢向左和中路射门，而不是向右射门，而守门员则采取了更为平衡的扑救方向策略，由于射手倾向于更多地朝左边射门，因此左边是较为优选的方向。通过方程（6）和（7）来分析这一点可以证明该系统是一个纳什均衡。表 1 和表 2 分别显示了对射手和守门员的这一分析结果。

Strategy	$\{Ay\}_i$	Optimal payoff for the striker $x^{*T}Ay^*$	$\Delta = \{Ay\}_i - x^{*T}Ay^*$
Kick left ($i = 1$)	0.436	0.436	0
Kick center ($i = 2$)	0.436	0.436	0
Kick right ($i = 3$)	0.436	0.436	0

表 1: 得分者收益分析

Strategy	$\{Bx\}_j$	Optimal payoff for the goalkeeper $y^{*T}Bx^*$	$\Delta = \{Bx\}_j - y^{*T}Bx^*$
Dive left ($j = 1$)	-0.436	-0.436	0
Dive center ($j = 2$)	-0.436	-0.436	0
Dive right ($j = 3$)	-0.436	-0.436	0

表 2: 守门员收益分析

从平衡点开始轨迹，如图 1 所示，它们将保持在那里：

Lyapunov 稳定性理论表明系统是中性稳定的，这意味着如果我们从平衡点的邻域开始，轨迹会在平衡点附近振荡。图 2 显示了这一现象对于守门员和射手随机选择的初始概率分布的发生。

进一步在不同的扰动下检查系统，在已知的平衡点上添加了均值为 0 且标准差不同的随机生成高斯噪声，从而允许在远离平衡点的邻域内随机生成初始条件。图 3-4 分别显示了使用 $\sigma = 0.02$ 和 $\sigma = 0.2$ 生成的轨迹。

观察到对于这些初始条件，轨迹在平衡点附近持续振荡，这意味着两个代理都没有学会在这种点球大战中采取最优策略，要在射手和守门员之间实现战略平衡仍有待

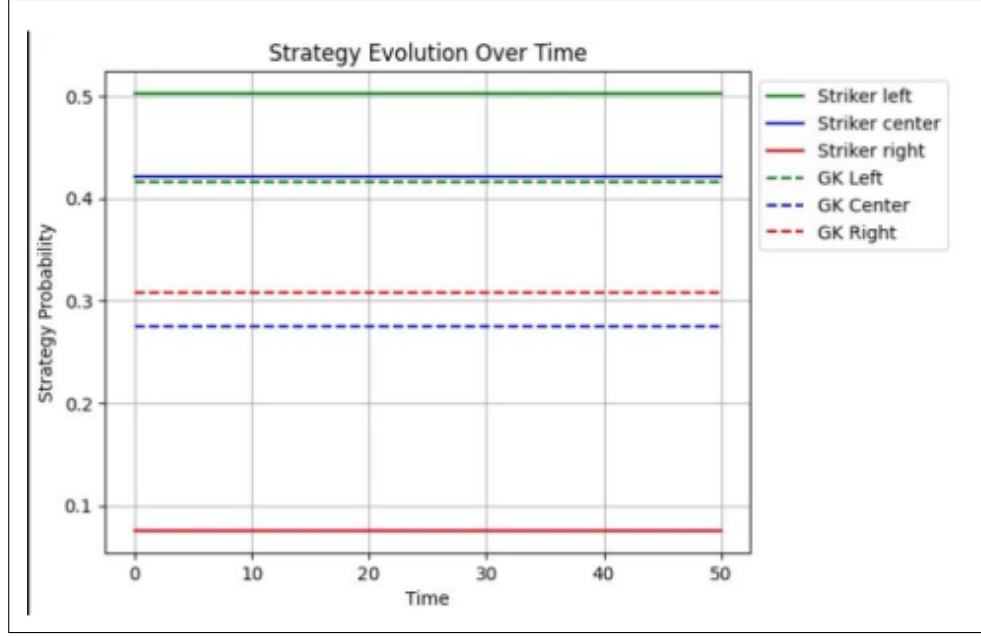


图 1: 初始条件与纳什均衡相同, 使轨迹保持在起始位置

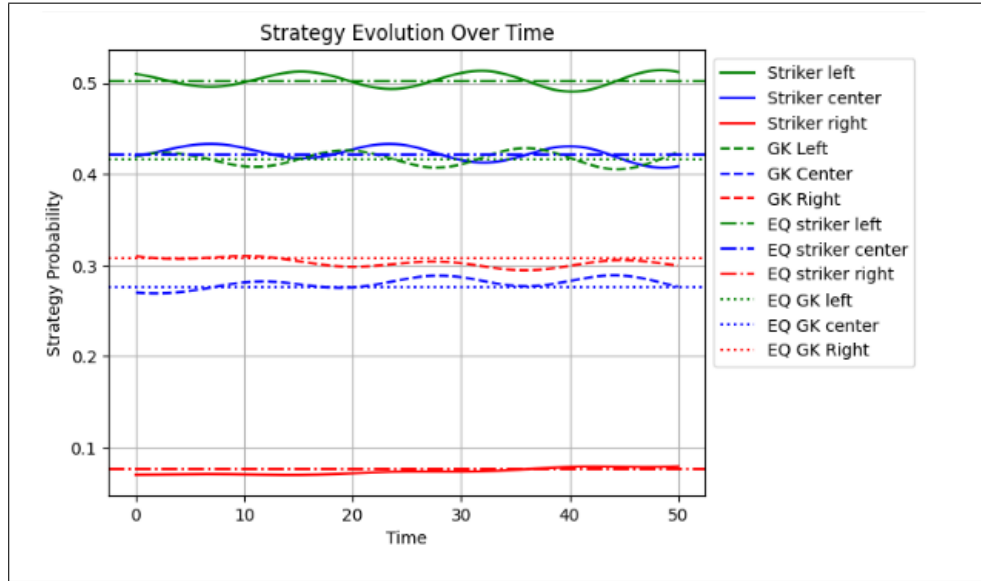


图 2: 初始概率分布随时间的策略演化, $x = [0.51, 0.42, 0.07]$ 和 $y = [0.42, 0.27, 0.31]$

提高。从图 3 和图 4 可以看出, 增加噪声分布的宽度会增大初始条件与系统平衡点之间的距离。这导致围绕平衡点的振荡增加, 并反映出系统如 Lyapunov 理论所强调的那样具有中性稳定性。为了进一步研究这一点, 分析了每个玩家随时间的收益情况。修改方程 (6) 和 (7), 使我们可以考察对于射手和守门员而言, 玩家策略相对于每个玩家的平均收益是如何产生回报的。

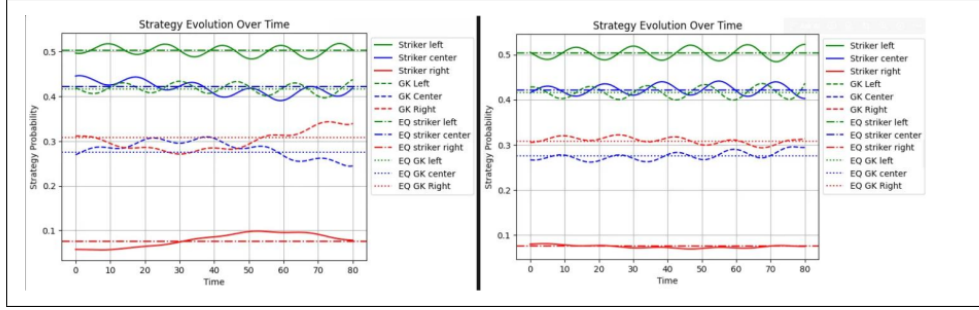


图 3: (a) (b): 初始概率分布随时间的战略演化, $\sigma = 0.02$

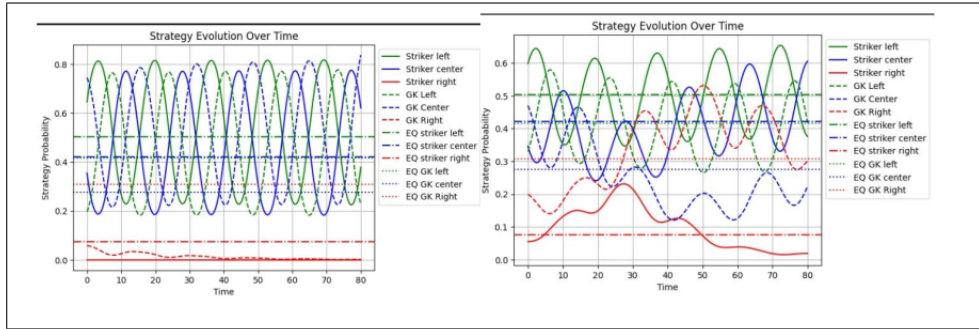


图 4: (a) (b): 初始概率分布随时间的战略演化, $\sigma = 0.2$

$$\{Ay\}_i - x^T Ay \neq 0 \text{ and } \{Bx\}_j - y^T Bx \neq 0$$

通过绘制图 3(a) 和 4(a) 中这些差异, 我们将两个玩家的收益与平均值进行比较, 得到如图 5-6 所示的图表:

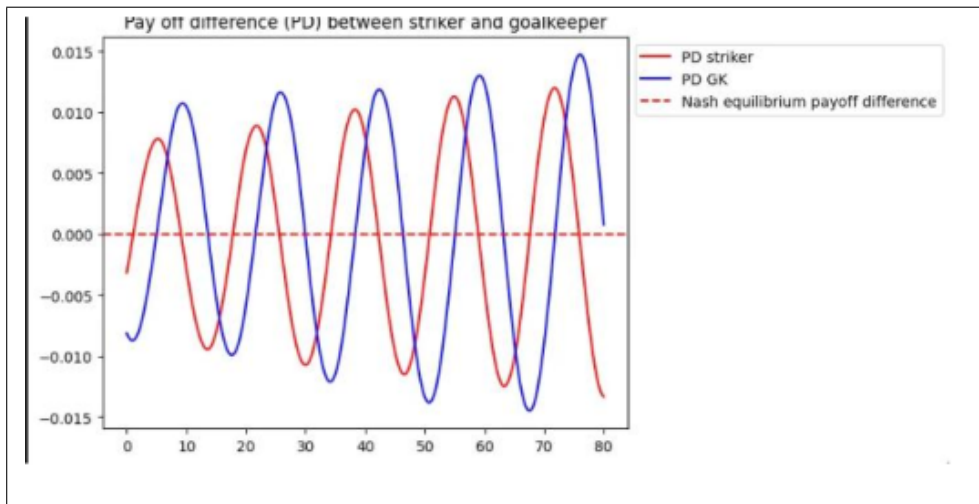


图 5: 3(a) 的收益差异随时间 (秒) 变化

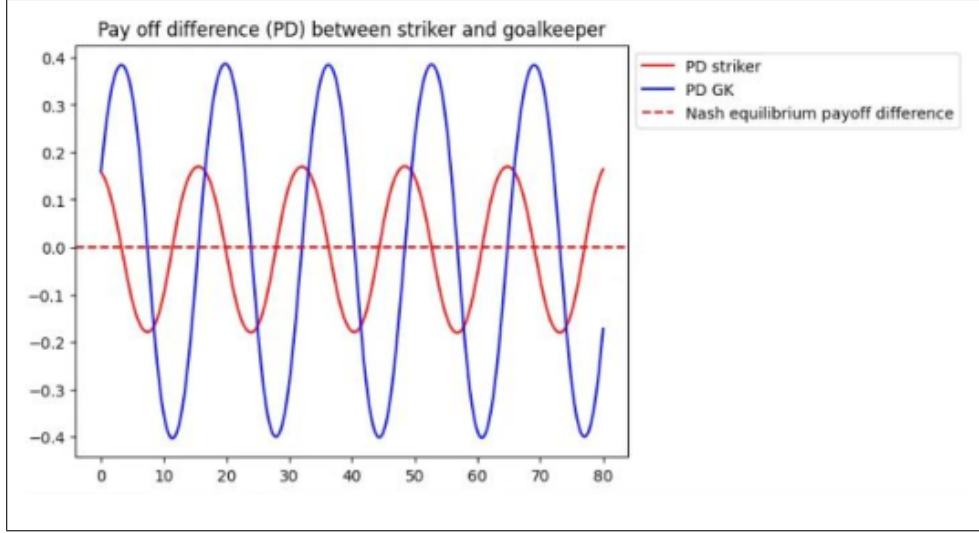


图 6: 收益差异对比于时间（秒）对于 4(a)

射门者和守门员的收益差异在 0 附近波动，这意味着双方都无法达到纳什均衡，并且球员们一直在不断适应对方的战略。高斯分布宽度的增加使得这些差异更加明显，意味着收益轨迹偏离平衡点更远。正的差异表明当前策略为玩家带来的收益高于平均水平，反之亦然，这意味着表现更好的策略正在被选择，而表现不佳的策略则在减少。射门者和守门员的收益差异也朝着相同的方向移动，这表明他们正在适应对方的战略。结果导致游戏发展成为一种相互适应战略的循环，并且由于这些战略无法收敛到纳什均衡，因此玩家有动机通过单方面改变自己的战略来获益。由于复制动态本身无法实现战略平衡，需要通过反馈线性化进行更多的工作以达到战略平衡，从而使比赛更具可预测性。

6 反馈线性化

由于复制子动态本身证明在教会代理最优策略方面无效，因此引入了全状态反馈线性化到系统中。反馈线性化使系统全局渐近稳定 [7]，允许无论初始起始条件如何都收敛到纳什均衡。这使得两个代理都能适应优化后的策略，从而使他们相对于另一个玩家最大化各自的收益。方程 (9) 和 (10) 显示了为反馈线性化添加到系统中的更新复制子动态。

$$\frac{dx_i}{dt} = x_i (\{Ay\}_i - x^T Ay) + u_i, i \in \{1, 2, 3\} \quad (9)$$

$$\frac{dy_j}{dt} = y_j (\{Bx\}_j - y^T Bx) + v_j, j \in \{1, 2, 3\} \quad (10)$$

由于每个系统具有三次度，由三个状态方程决定，因此每个代理必须有 3 个控制输出，这意味着整个系统共有 6 个控制输出。完全状态线性化可以通过以下方程定义 [7]：

$$\begin{aligned}\frac{dx}{dt} &= f(x) + g(x)u \\ y &= h(x)\end{aligned}\quad (11)$$

对于如方程 (9) 和 (10) 所示的更新动力学，可以定义击球者和守门员的方程 (11) 参数为：

$$\frac{dx}{dt} = \begin{bmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \\ \frac{dx_3}{dt} \end{bmatrix}, f(x) = \begin{bmatrix} x_1\{Ay\}_1 - x^T Ay \\ x_2\{Ay\}_2 - x^T Ay \\ x_3\{Ay\}_3 - x^T Ay \end{bmatrix}, g(x) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, u = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \quad (12)$$

$$\frac{dy}{dt} = \begin{bmatrix} \frac{dy_1}{dt} \\ \frac{dy_2}{dt} \\ \frac{dy_3}{dt} \end{bmatrix}, f(y) = \begin{bmatrix} y_1\{Bx\}_1 - y^T Bx \\ y_2\{Bx\}_2 - y^T Bx \\ y_3\{Bx\}_3 - y^T Bx \end{bmatrix}, g(y) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, v = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \quad (13)$$

由于在这个实例中系统的动态特性是已知的，因此方便取消系统中的非线性项，并添加一个涉及每种策略达到纳什均衡的概率收敛性的线性项。控制输入 u 和 v 如下：

$$\begin{aligned}u &= \begin{bmatrix} -x_1\{Ay\}_1 + x^T Ay - k_1(x_1 - x_1^*) \\ -x_2\{Ay\}_2 + x^T Ay - k_2(x_2 - x_2^*) \\ -x_3\{Ay\}_3 + x^T Ay - k_3(x_3 - x_3^*) \end{bmatrix} \\ v &= \begin{bmatrix} -y_1\{Bx\}_1 + y^T Bx - c_1(y_1 - y_1^*) \\ -y_2\{Bx\}_2 + y^T Bx - c_2(y_2 - y_2^*) \\ -y_3\{Bx\}_3 + y^T Bx - c_3(y_3 - y_3^*) \end{bmatrix}\end{aligned} \quad (14)$$

将控制输入 u 和 v 代入方程 (14) 中的方程 (12) 和 (13)，复制器动力学的非线性被抵消，结果得到以下动态系统方程：

$$\begin{bmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \\ \frac{dx_3}{dt} \end{bmatrix} = \begin{bmatrix} -k_1(x_1 - x_1^*) \\ -k_2(x_2 - x_2^*) \\ -k_3(x_3 - x_3^*) \end{bmatrix}, \begin{bmatrix} \frac{dy_1}{dt} \\ \frac{dy_2}{dt} \\ \frac{dy_3}{dt} \end{bmatrix} = \begin{bmatrix} -c_1(y_1 - y_1^*) \\ -c_2(y_2 - y_2^*) \\ -c_3(y_3 - y_3^*) \end{bmatrix}, \quad k_i, c_i > 0 \quad (15)$$

使用方程 (15)，系统可以实现渐近稳定，对于大于零的任何值都能达到全局渐近稳定。为了证明这一概念，重新计算了 $\frac{dx}{dt}$ ，结果小于零，证明了如证明 2 所示的渐近稳定性。这些常数值决定了收敛到纳什均衡的速度，较高的值允许更快地收敛反之

亦然，并且如果它们都大于零，则可以任意选择。

证明 2：渐近稳定性

$$\begin{aligned}\dot{V} &= -\sum_{i=1}^3 \frac{x_i^* (\dot{x}_i)}{x_i} - \sum_{j=1}^3 \frac{y_j^* (\dot{y}_j)}{y_j} \\ \dot{x}_i &= -k_i (x_i - x_i^*), \dot{y}_j = -c_j (y_j - y_j^*) \\ \dot{V} &= -\left(\sum_{i=1}^3 \frac{x_i^* (-k_i (x_i - x_i^*))}{x_i} + \sum_{j=1}^3 \frac{y_j^* (-c_j (y_j - y_j^*))}{y_j}\right) \\ \dot{V} &= -\left(\sum_{i=1}^3 x_i^* \left(k_i \left(-\frac{x_i}{x_i} + \frac{x_i^*}{x_i}\right)\right) + \sum_{j=1}^3 y_j^* \left(c_j \left(-\frac{y_j}{y_j} + \frac{y_j^*}{y_j}\right)\right)\right)\end{aligned}$$

$$\dot{V} = \sum_{i=1}^3 x_i^* \left(k_i \left(1 - \frac{x_i^*}{x_i}\right)\right) + \sum_{j=1}^3 y_j^* \left(c_j \left(1 - \frac{y_j^*}{y_j}\right)\right), \text{ 对于 } \frac{y_j^*}{y_j}, \frac{x_i^*}{x_i} > 1 \dot{V} < 0$$

然而，由于约束条件的存在，这无法证明反馈线性化定理所提出的全局渐近稳定性，也就是说，如果给定方向的概率大于 1，则全局渐近稳定性的证明将失效。因此，必须使用不同的李雅普诺夫函数来证明全局渐近稳定性。考虑：

$$V(x, y) = \sum_{i=1}^3 \frac{1}{2} (x_i - x_i^*)^2 + \sum_{j=1}^3 \frac{1}{2} (y_j - y_j^*)^2$$

这个候选函数也满足以下条件：

- $V(x^*, y^*) = 0$ 是该函数的唯一零点
- $V(x, y) > 0$ 对所有 $x \neq x^*, y \neq y^*$ 成立

证明 2 使用新的李亚普诺夫候选函数：

$$\begin{aligned}\dot{V}(x, y) &= \sum_{i=1}^3 \dot{x}_i (x_i - x_i^*) + \sum_{j=1}^3 \dot{y}_j (y_j - y_j^*) \\ \dot{x}_i &= -k_i (x_i - x_i^*), \dot{y}_j = -c_j (y_j - y_j^*) \\ \dot{V}(x, y) &= \sum_{i=1}^3 -k_i (x_i - x_i^*) (x_i - x_i^*) - \sum_{j=1}^3 c_j (y_j - y_j^*) (y_j - y_j^*) \\ \dot{V}(x, y) &= \sum_{i=1}^3 -k_i (x_i - x_i^*)^2 + \sum_{j=1}^3 -c_j (y_j - y_j^*)^2, \forall k, c > 0 \dot{V} < 0 \rightarrow \text{asymptotic stability}\end{aligned}$$

这个第二个候选李亚普诺夫函数能够证明动态系统全局渐近稳定。这意味着无论从哪个起点开始，动态系统都将保证收敛到纳什均衡，从而两个玩家都能学会点球大战中的最优策略。

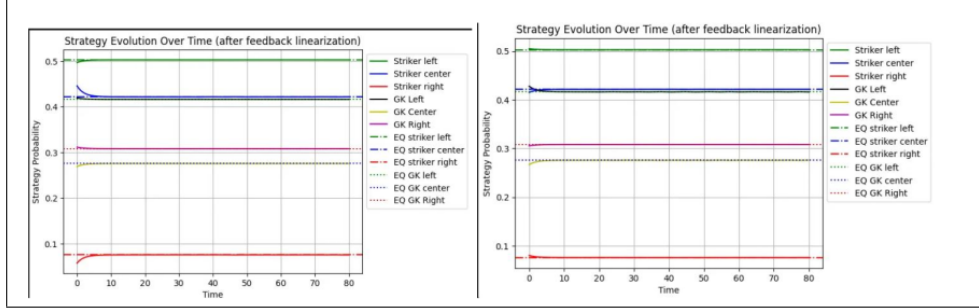


图 7: (a) (b): 反馈线性化后的时间演化策略, $\sigma = 0.02, k = c = 0.5$

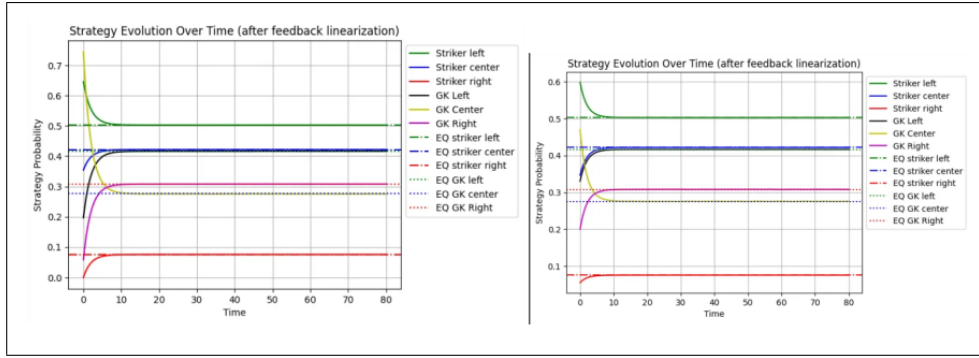


图 8: (a) (b): 反馈线性化后的时间演化策略, $\sigma = 0.2, k = c = 0.5$

图 7 和图 8 显示，无论初始条件中加入的扰动大小如何，方程 (15) 所示的反馈线性化控制都能使玩家策略收敛到系统的期望纳什均衡点，从而使系统渐近稳定。

为了进一步可视化系统的渐近稳定性，可以分析收益与平均值的偏差，因为收敛于纳什条件表明收益差异 (PD) 应收敛至 0，这标志着向纳什均衡的收敛。

为说明这一概念，下方显示了图 9 和图 10，并给出了标准差的数值。

与复制子动力学本身的中性稳定性相比，反馈线性化保证了收敛到纳什均衡，这意味着玩家选择的策略表现得如同最优混合策略一样好。因此，通过收敛到纳什均衡，这种反馈线性化使双方都能学习到各自的最优策略，并在比赛中实现平衡，防止比赛无限循环，并最小化游戏中被利用或出现错误的风险，从而在玩家和守门员之间创造出战略上的平衡。在现实生活中，这可以想象为基于侦察报告和之前的点球大战数据的教练指导，这些都允许球员优化他们在点球大战中的各自策略。系统稳定下来——不是因为每个人都在赢，而是因为没有任何一方能够通过单方面改变自己的策略来获胜。

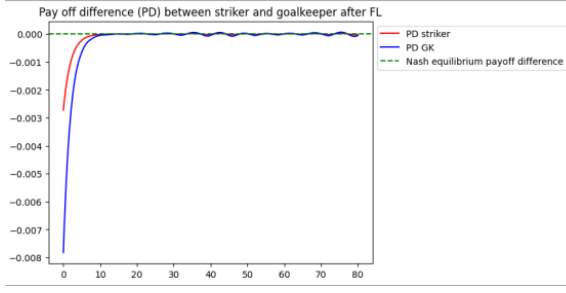


图 9: FL 实施后 7(a) 的 PD

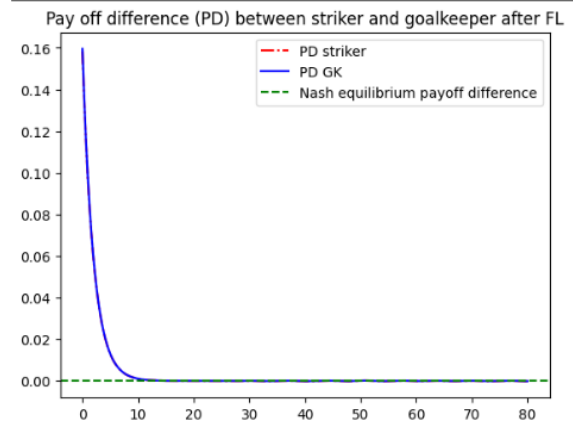


图 10: FL 实施后 8(a) 的 PD

7 限制和未来工作

该框架依赖于一系列大假设，这些假设限制了其立即实际应用的范围。首先，该框架需要确切地知道收益矩阵 A ，而在现实世界的战略互动中通常无法获得这样的信息，因为收益必须来自于噪声观测或一个变动的目标，这会改变游戏条件。其次，实施过程假设有集中控制能力，即外部协调者可以修改每个群体内玩家策略演化率——这是一个强烈的假设，并不一定反映经历策略演化的决策者的自主倾向。第三，分析集中在概率单纯形的内部，需要更多的考虑来捕捉通过投影方法或屏障函数接近策略边界 ($x_i = 0$) 的行为。在数值实现中，使用了裁剪和归一化技术以保持总和不超过 1 且概率值始终为正的约束。最后，该方法假设玩家具有连续的策略调整，这些调整遵循演化微分方程；然而，正如我们所知的战略行为通常由固有认知或物理限制的离散决策组成。

这些约束暗示了几种有用的扩展。首先，这将限制模型不确定性下的鲁棒性收敛到纳什均衡，滑模控制的动机是它会限制实际系统与估计系统之间的误差。其次，可以放宽固定收益矩阵的假设，以允许时变博弈和特定玩家属性，从而实现更现实的战略建模。第三，分布式实现方法可以缓解对集中式控制的需求，使得每个自主玩家可以根据局部交互在有限知识的基础上收敛到纳什均衡。最后，可以构造离散时间公式来展示理论上连续动力学与离散决策过程之间更现实的转换。

8 结论：

在这篇论文中，我们展示了首次将反馈线性化用于简化复制动态框架的方法，这为从控制理论的角度实现进化博弈中的保证收敛提供了一个框架。严格的李雅普诺夫分析能够证明该控制律产生全局渐近稳定性，将复制动态的振荡自然行为转化为对纳

什均衡的可预测收敛。

我们展示了这种方法论如何说明控制理论可以从根本上改变进化博弈的收敛性质，而传统学习方法对此没有提供理论保证。虽然我们通过点球分析的例子来阐述这些性质，但该理论适用于任何由复制动态描述的战略互动类型。

未来的研究将利用滑模控制展开这项基础工作，以适应未知的收益结构、作为自主代理的分布式实现以及现实世界中的离散状态形式。建立这一新的方法论途径在控制理论和演化博弈的交叉领域创造了一种新的研究流派，现在可以在多代理系统、算法交易和战略优化等多个领域加以利用。

参考文献

- [1] Hayes A., *Game Theory: A Comprehensive Guide*, Investopedia, June 27, 2024.
- [2] Fryer D., *The Kullback-Leibler Divergence as a Lyapunov Function for Incentive-Based Game Dynamics*, 2018.
- [3] Weisstein E.W., “Positive Definite Matrix,” *MathWorld—A Wolfram Web Resource*, [Online]. Available: <https://mathworld.wolfram.com/PositiveDefiniteMatrix.html>. [Accessed: Apr. 29, 2025].
- [4] Engel G., *A Stochastic Variant of Replicator Dynamics in Zero-Sum Games and Its Invariant Measures*, Physica D, vol. 456, pp. 1 – 15, 2023.
- [5] Allesina S., *Game Theory and Replicator Dynamics*, Chicago, 2025.
- [6] Monstad V., *The Game Theory of Penalty Kicks: A Framework for Approximating Nash Equilibria*, NHH Norwegian School of Economics, 2023.
- [7] Khalil H.K., *Nonlinear Systems*, Prentice Hall, 2002.

附录：

该项目的 GitHub 仓库：

<https://github.com/adilfaisal01/SE762--Game-theory-and-Lyapunov-calculations>

欢迎留下评论和问题，我很乐意回答它们。