

第八章 博弈学习和进化博弈

湖南大学课程

本章内容

1. 有限理性及其博弈分析
2. 博弈学习模型
 - 最优反应动态
 - 虚拟行动 (略过)
3. 进化博弈
 - 进化稳定策略
 - 复制动态方程

有限理性

- 有限理性：博弈方存在**理性局限**。
- 存在有限理性博弈方的博弈称**有限理性博弈**。博弈方都有完全理性则称**完全理性博弈**。
- 我们之前的博弈分析主要是**完全理性博弈**：
 - 参与人以效用最大化为目标, 有完美的分析判断和决策执行能力, 不会犯错...
- 理性模型都很类似, 有限理性模型的内容相对要丰富得多。每种有限理性博弈模型有各自的适用情况。本章重点介绍两类有限理性博弈模型：
 1. **博弈学习模型** (最优反应动态: 通过不断试错来调整策略)
 2. **进化博弈** (复制动态, 进化稳定策略)

最优反应动态

1. 古诺调整动态
2. 西瓜甜瓜博弈
3. 多方协调博弈

古诺模型

- 两寡头厂商供应同样的产品, 同时定产: (q_A, q_B)
- 市场出清价格: $P = 8 - Q$
- 无固定成本, 边际成本均为 $c = 2$
- 双方最优反应函数:
- $q_A^*(q_B) = 3 - q_B/2, q_B^*(q_A) = 3 - q_A/2$
- 如果双方都是有预见能力的理性博弈方, 双方都将选择古诺均衡产量 $q^c = 2$.

古诺调整过程：最优反应动态

- 如果博弈方存在理性局限, 缺乏预测能力, 单次博弈中很难达成均衡产量.
- 用 q_i^t 表示厂商 $i \in \{A, B\}$ 在第 t 期的产量.
- 假设第一期中, 两家厂商盲目定产: $q_A^1 = 2.5, q_B^1 = 3$.
- 第二期中, 厂商分别针对上期对手的行动进行最优反应:
 - $q_A^2 = q_A^*(q_B^1) = 1.5, q_B^2 = q_B^*(q_A^1) = 1.75$
- 依此类推:
 - 第三期产量组合: $q_A^3 = 2.125, q_B^3 = 2.25$
 - 第四期产量组合: $q_A^4 = 1.875, q_B^4 = 1.9375$
- 可以证明, 上述最优反应动态中, 厂商 A 和 B 的产量都会趋于古诺均衡产量:
 - 随着 $t \rightarrow \infty, q_A^t \rightarrow 2$ 并且 $q_B^t \rightarrow 2$

古诺调整过程何时收敛?

- 对于一般情形, 分别用 $R_1(q_2)$ 和 $R_2(q_1)$ 表示最优反应函数
- 古诺调整过程收敛的充分条件:
$$|R_1'(q_2)R_2'(q_1)| < 1$$
 - 前面的例子中, $|R_1'(q_2)R_2'(q_1)| = 1/4 < 1$, 因此调整动态会逐渐收敛到对应的均衡产量.
- 最优反应动态的主要局限性: 参与人始终假设对手的行动不变.
 - 在下一个西瓜-甜瓜博弈的例子中, 如果参与人始终假设对方产量不变, 动态调整过程无法收敛到纳什均衡.

西瓜甜瓜博弈

		农场2	
		西瓜	甜瓜
农场1	西瓜	0, 0	8, 5
	甜瓜	5, 8	0, 0

协调博弈

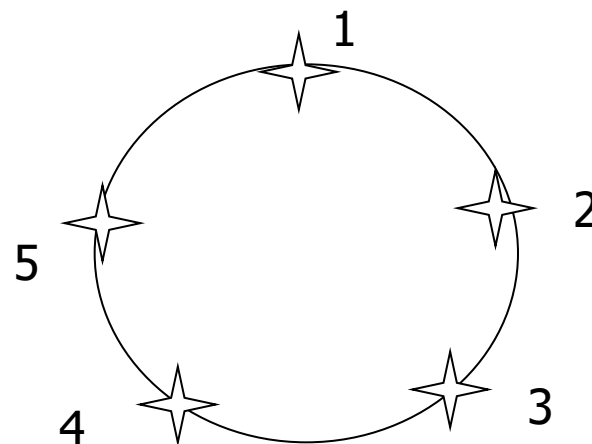
- 三个纳什均衡
 - (西瓜, 甜瓜)
 - (甜瓜, 西瓜)
 - 一个混合策略均衡
- 假设博弈重复无穷次, 双方具有有限理性, 并且第一年双方都选择利润较高的甜瓜. 则最优反应动态如下:
 1. (甜瓜, 甜瓜)
 2. (西瓜, 西瓜)
 3. (甜瓜, 甜瓜)
 4. ...
- 无法收敛到纳什均衡, 并且收益始终为 0.

复杂博弈关系的最优反应动态 (多方协调博弈)

模型:

		博弈方2	
		A	B
博弈方1	A	50, 50	49, 0
	B	0, 49	60, 60

协调博弈



- 假设有5个博弈方, 分别分布在右图圆周上的5个位置.
- 每个博弈方同时选择 A 和 B, 与相邻的两边进行博弈. 收益矩阵如左图

最优反应、策略调整规则推导

- 令 $x_i(t)$ 表示 t 时刻博弈方 i 的邻居中使用策略 **A** 的数量.

$$x_i(t) \in \{0, 1, 2\}$$

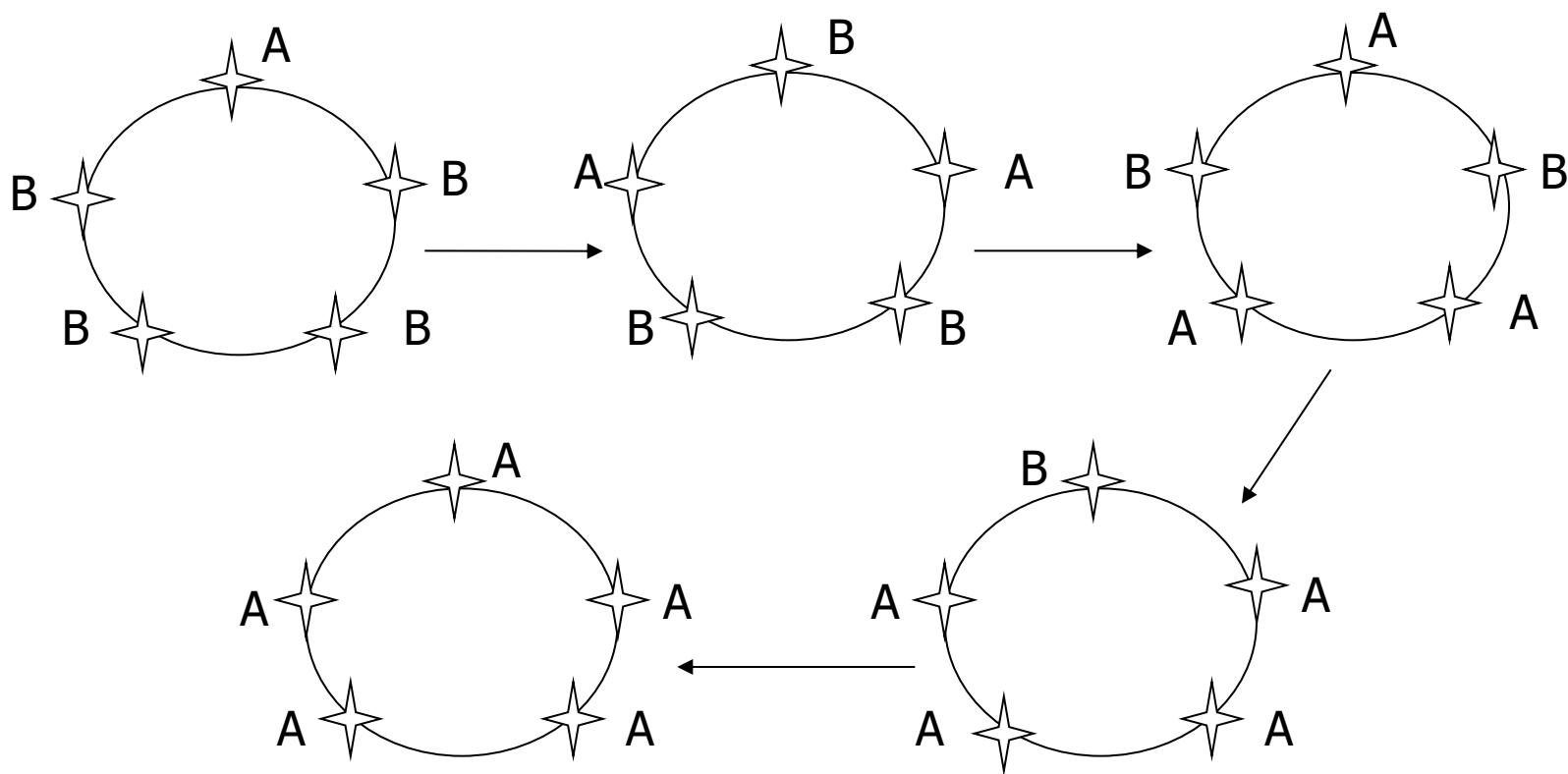
采用**A**的得益： $x_i(t) \times 50 + [2 - x_i(t)] \times 49$

采用**B**的得益： $x_i(t) \times 0 + [2 - x_i(t)] \times 60$

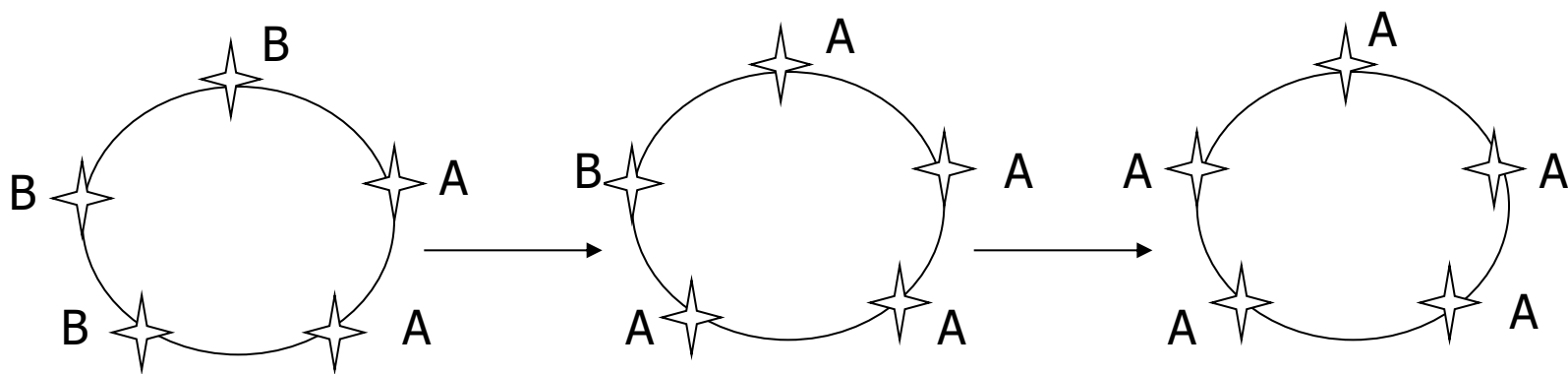
当 $x_i(t) > 22/61$ 时，采用**A**；当 $x_i(t) < 22/61$ 时，采用**B**

- 结论：尽管单期博弈有两个纯策略纳什均衡，但只要一开始有博弈方选择 **A**，最后所有参与人的策略在最优反应动态下都会收敛到 **A**.

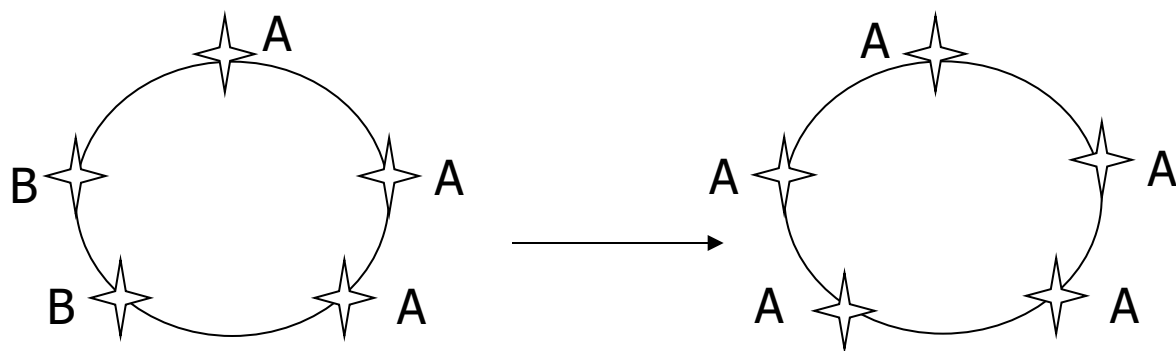
最优反应动态模拟：初次博弈1A



初次博弈相邻2个A



初次博弈相连3个A



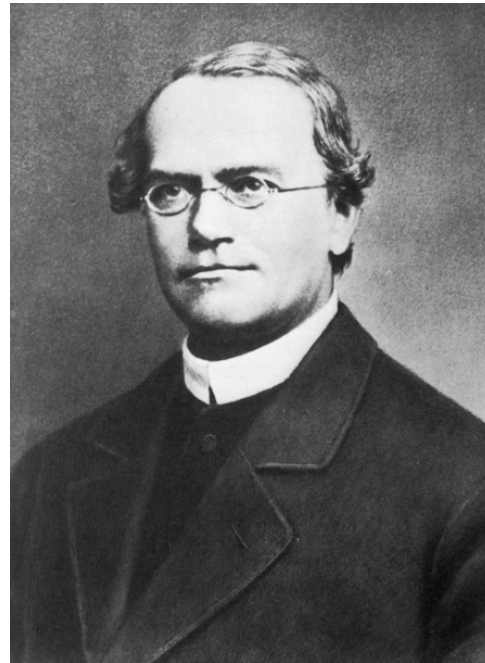
注：前面的三类分析，已经遍历了所有可能的动态.

图 1 包括了 “2A 不相邻”，“3A间隔” 和 “4A” 的情况.

8.3 进化博弈论 (Nash meets Darwin and Mendel)

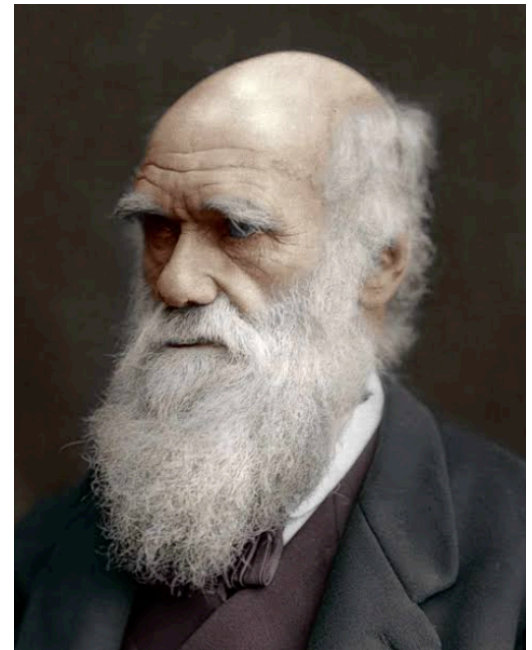
孟德尔 (1822–1884)

通过豌豆实验发现遗传定律,并猜测生命体存在基本的遗传物质 (即今天人们所说的基因).



达尔文 (1809–1882)

提出进化论: 用进废退, 适者生存



经济博弈 v.s. 进化博弈

- 生物学家也使用博弈论的工具 (**进化博弈**) 来研究动植物的行为和种群竞争等生物学现象.
- **进化博弈**和**经济博弈**的核心区别在于, 进化博弈中的策略就是生命体的基因.
 - 由于进化博弈中的策略就是**基因 (或某类基因)**, 进化博弈中参与方的策略不是它基于个体理性选择的, 而是它从出生开始就与生俱来的.
- 进化博弈对经济学的影响:
 - 虽然进化博弈的相关分析工具主要是博弈论学者在研究生物学现象时提出的, 但经济学家也会使用进化博弈的工具来分析经济学现象, 如商业竞争.
 - 这类模型也可以归类为**有限理性模型**. 例: 模型中企业的策略主要由它的“基因” (如企业文化) 决定, 而不是由企业基于利润最大化的考量所决定的.

从纳什均衡到进化稳定策略

- 进化博弈中, 策略(or, 基因)不是博弈方基于个人理性选择的. 因此, 纳什均衡这个概念就不再适用, 因为纳什均衡要求所有博弈方的策略是最优反应.
- 进化博弈论中使用的主要分析工具是**进化稳定策略**.
 - 进化稳定策略这个想法由史密斯和普莱斯于 1973 年在论文 “The logic of animal conflict” 中提出, 这个分析工具的核心思想就是达尔文的进化论 (“用进废退, 适者生存”).
- **进化稳定策略**的非严格定义:
 - 假设种群中所有个体都使用纯策略 s (或携带基因 s). 如果任何可能的突变基因 s' 最终都会在博弈中被淘汰, 则称 s 是进化稳定的.
- 下面我们用“囚徒困境”的例子来解释进化稳定策略.

囚徒困境的进化博弈分析

- 初始时, 假设种群中所有个体携带的基因均为 s_1 .
- 假设有 ε 比例的个体发生突变. 现在种群中有 $1 - \varepsilon$ 比例的个体携带基因 s_1 , ε 比例的个体携带基因 s_2 .
- 博弈进行无穷期, 每一期里种群中的个体随机配对进行囚徒困境博弈.
 - 每个个体有概率 $1 - \varepsilon$ 和 s_1 配对, 概率 ε 和 s_2 配对.
 - 携带 s_1 基因个体的期望收益: $2(1 - \varepsilon) + 0 \cdot \varepsilon = 2(1 - \varepsilon)$
 - 携带 s_2 基因个体的期望收益: $3(1 - \varepsilon) + 1 \cdot \varepsilon = 3 - 2\varepsilon > 2(1 - \varepsilon)$

博弈方2

		s_1	s_2
博弈方1	s_1	2,2	0,3
	s_2	3,0	1,1

囚徒困境的进化博弈分析

- 演化博弈中, 博弈方的收益代表其与环境的适应度 (**fitness**). 由于突变基因 s_2 的适应度比基因 s_1 更高, 因此基因 s_2 不会灭绝.
 - 结论1: 策略 s_1 不是进化稳定的.
- 另一方面, 如果初始时所有个体均携带基因 s_2 . 当有 ε 比例的个体基因突变为 s_1 时, 这些突变基因 s_1 对环境的适应度不如 s_2 , 携带基因 s_1 的个体会灭绝.
 - 结论2: 策略 s_2 是进化稳定的.
- 囚徒困境这个进化博弈的启示:
 - 自然选择并非总能最大化种群整体的适应度. 尽管所有个体都携带基因 s_1 时种群整体的适应度最大, 但基因 s_1 不是进化稳定的: 当出现突变基因 s_2 时, 携带基因 s_1 的个体会灭绝.

进化稳定策略和纳什均衡

- 命题1: 如果策略 s_1 是严格下策, s_1 一定不是进化稳定策略.
- 命题2: 如果策略 s_1 是进化稳定策略, (s_1, s_1) 是纳什均衡.
- 命题3: 如果 (s_1, s_1) 是纳什均衡, 策略 s_1 不一定是进化稳定策略.
 - 下例中, (s_1, s_1) 和 (s_2, s_2) 都是纳什均衡, 但 s_1 不是进化稳定策略.

	s_1	s_2
s_1	0,0	0,0
s_2	0,0	2,2

从纯策略到混合策略

- 前面给出的“进化稳定策略”判定方法只对于**纯策略**适用.
- “进化稳定策略”这个概念也可以用于**混合策略**. 下面是进化博弈中混合策略的一个例子.
 - 令博弈方的策略集为 $\{s_1, s_2\}$, 考虑混合策略 σ :
$$\sigma(s_1) = 0.4, \sigma(s_2) = 0.6$$
 - σ 表示种群中 **40%** 的个体携带基因 s_1 , 的个体携带基因 s_2 .
- 对于两方进化博弈, 如果要找出所有可能的进化稳定状态, 我们需要描述所有个体在博弈过程中的繁殖过程.
 - 这个繁殖过程可以用关于时间 t 的微分方程来描述, 这个微分方程一般称为**复制动态方程**.

复制动态方程

- 对两方博弈, 假设每个博弈方的策略集均为 $\{s_1, s_2\}$.
- 记 t 时刻使用策略 s_1 个体的比例为 $x_1(t)$, 则使用策略 s_2 个体的比例为 $x_2(t) = 1 - x_1(t)$.
 - 后文中, 我们经常会把 $x_1(t)$ 简写为 x_1 .
- 使用策略 s_1 个体的繁殖速度 (或基因 s_1 的繁殖速度) 取决于 (i) 基因 s_1 在博弈结果中和环境的适应度以及(ii) 基因 s_2 在博弈结果中和环境的适应度.
 - 令 u_1 和 u_2 分别为 t 时刻使用策略 s_1 和 s_2 的适应度
 - 令 $\bar{u} = x_1 u_1 + (1 - x_1) u_2$ 为 t 时刻种群的平均适应度.
 - 则基因 s_1 在 t 时刻的繁殖速度为 $u_1 - \bar{u}$, 对应的复制动态方程为:

$$x_1'(t) = (u_1 - \bar{u})x_1(t)$$

稳定点

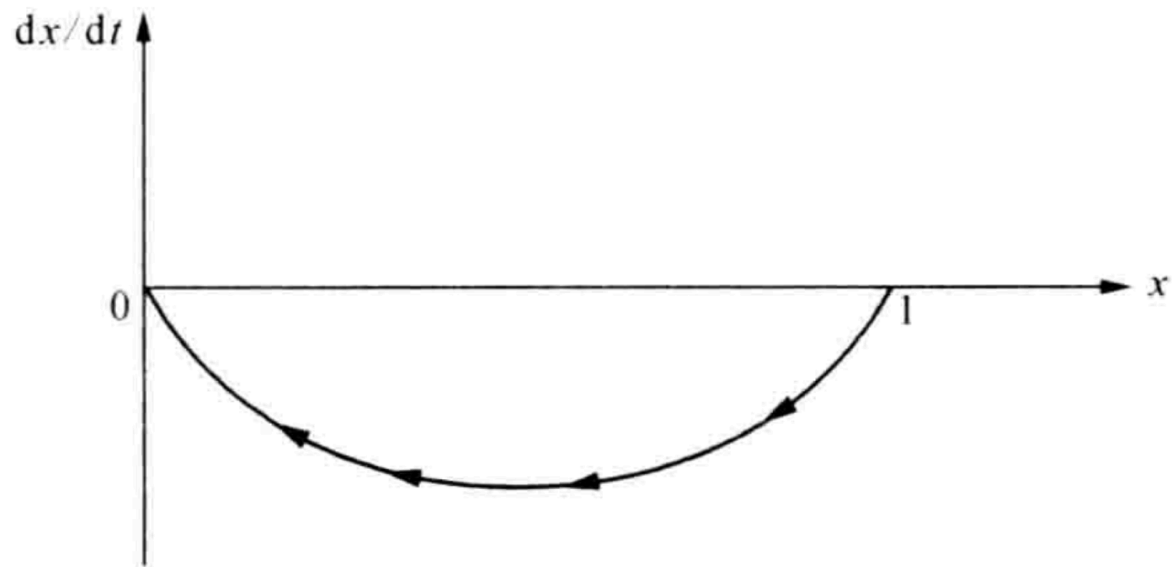
- 给定某个复制动态方程 $x_1'(t) = (u_1 - \bar{u})x_1(t)$, 我们可以把 x_1 作为横轴, x_1' 作为纵轴, 绘制出对应的复制动态曲线.
- 对曲线上的每个点,
 - 若 $x_1' > 0$, 则基因 s_1 的比例 x_1 会上升.
 - 若 $x_1' < 0$, 则基因 s_1 的比例 x_1 会下降.
 - 若 $x_1' = 0$, 则基因 s_1 的比例 x_1 在不发生突变的情况下是稳定的. 我们称此时 x_1 对应的取值 x_1^* 为复制动态方程的稳定点.

进化稳定点

- 如果稳定点 x_1^* 周围存在某个小区间 D 使得:
 1. 对任意 $x_1 \in D$ 且 $x_1 > x_1^*$, 有 $x_1' < 0$ 成立
 2. 对任意 $x_1 \in D$ 且 $x_1 < x_1^*$, 有 $x_1' > 0$ 成立则称 x_1^* 是进化稳定点.
- 若 x_1^* 是进化稳定点, 此时即使发生基因突变使得基因 s_1 的比例暂时偏离 x_1^* , 复制动态方程也会使得 x_1 最终收敛到 x_1^* . (想想为什么?)

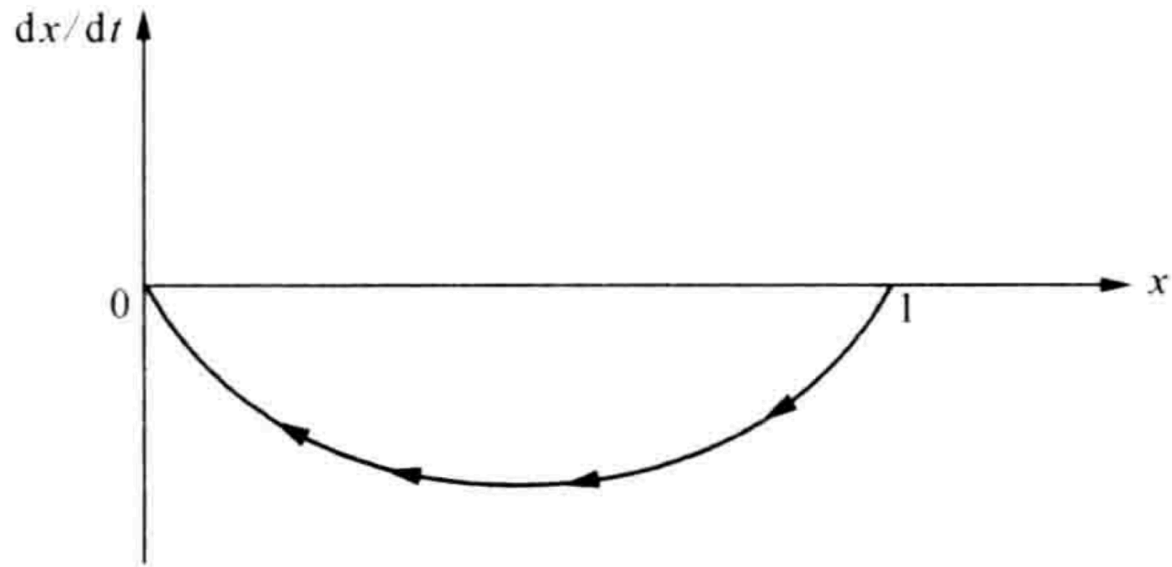
练习

下图中, 哪些是稳定点, 哪些是进化稳定点?



练习

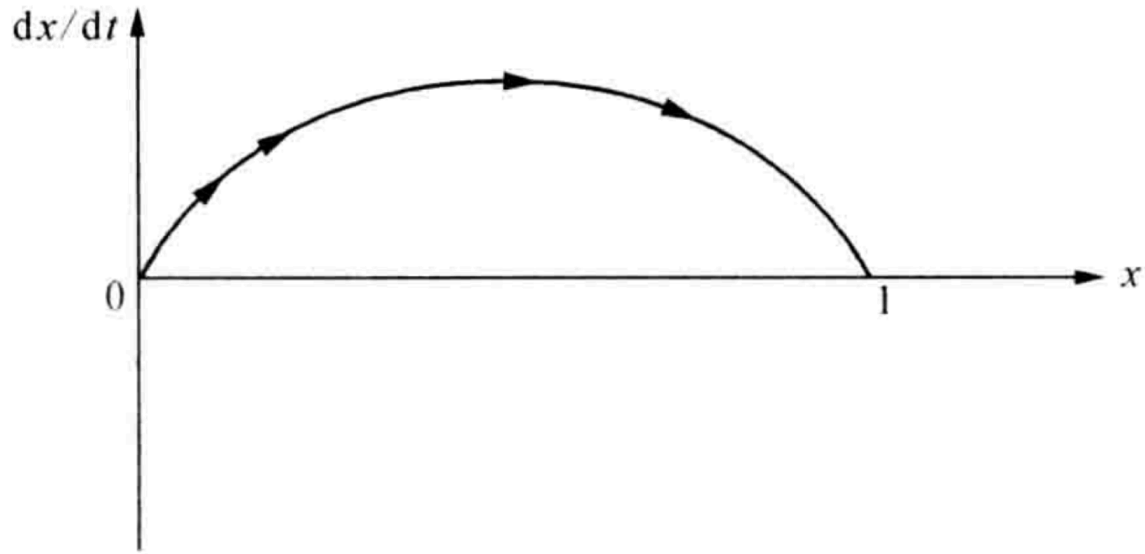
下图中, 哪些是稳定点, 哪些是进化稳定点?



$x=0$ 和 $x=1$ 是稳定点, 其中 $x=0$ 是进化稳定点.

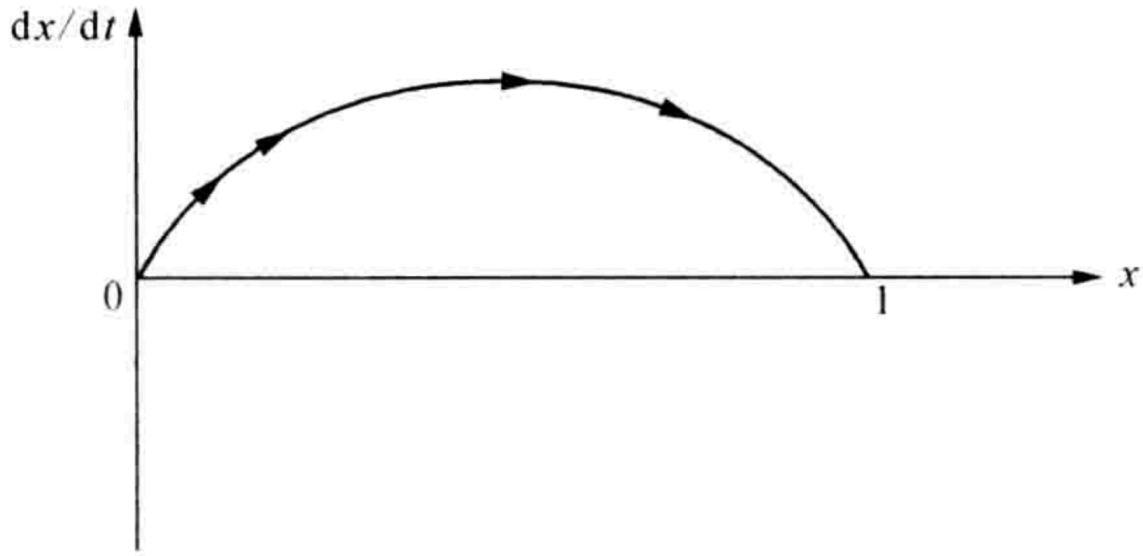
练习

下图中, 哪些点是稳定点, 哪些点是进化稳定点?



练习

下图中, 哪些点是稳定点, 哪些点是进化稳定点?



$x=0$ 和 $x=1$ 是稳定点, 其中 $x=1$ 是进化稳定点.

协调博弈的进化博弈分析

	s_1	s_2
s_1	50,50	49,0
s_2	0,49	60,60

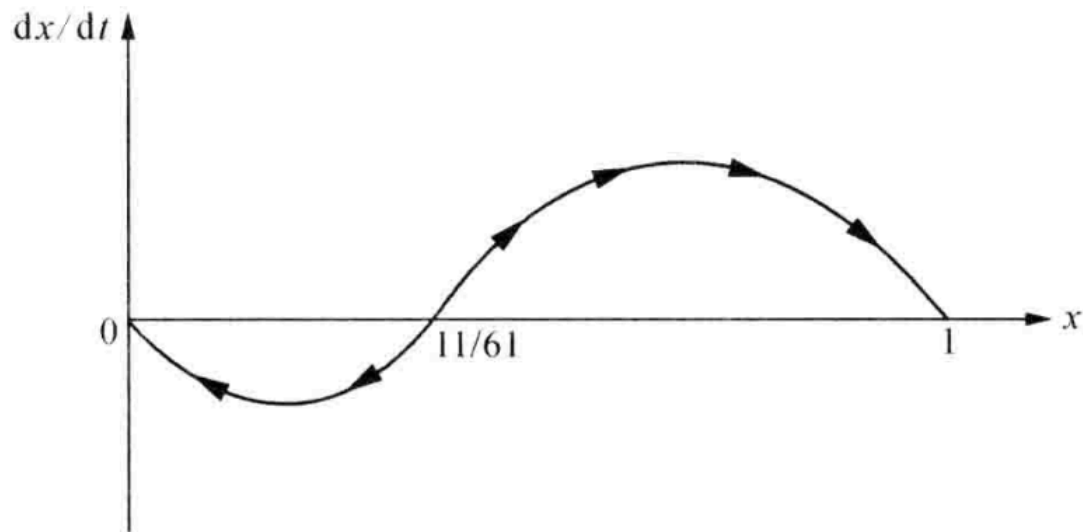
- 这个博弈有两个纯策略纳什均衡： (s_1, s_1) 和 (s_2, s_2) ，以及一个混合策略纳什均衡 (σ, σ) .
 - 混合策略 σ 表示：11/61 的概率使用 s_1 ，50/61 的概率使用 s_2 .
 - 两个纯策略均衡中， (s_1, s_1) 是风险上策均衡， (s_2, s_2) 是帕累托上策均衡.
- 下面我们用复制动态方程来检验 s_1 ， s_2 和 σ 这三个策略哪些是进化稳定的.

协调博弈：复制动态方程

- 记 t 时刻使用策略 s_1 个体的比例为 $x_1(t)$, 则使用策略 s_2 个体的比例为 $x_2(t) = 1 - x_1(t)$.
- 基因 s_1 和基因 s_2 的期望收益分别为:
- $u_1 = 50x_1 + 49(1 - x_1) = x_1 + 49, u_2 = 60(1 - x_1)$
- 种群的平均期望收益: $\bar{u} = 60 - 71x_1 + 61x_1^2$
- 复制动态方程:
- $x_1'(t) = (u_1 - \bar{u})x_1(t) = (61x_1 - 11)(1 - x_1)x_1$

协调博弈：相位图

- 3个稳定点 $x_1 = 0$, $x_1 = 1$ 和 $x_1 = 11/61$
- 其中前两个是进化稳定点, 最后的 $x_1 = 11/61$ 不是进化稳定点.



协调博弈：相位图

- 3个稳定点 $x_1 = 0$, $x_1 = 1$ 和 $x_1 = 11/61$
- 其中前两个是进化稳定点, 最后的 $x_1 = 11/61$ 不是进化稳定点.

相位图还说明了:

- 当初始状态为 $x_1 < 11/61$ 时, 种群会收敛到进化稳定策略 s_2 (即进化稳定点 $x_1 = 0$);
- 当初始状态为 $x_1 > 11/61$ 时, 种群会收敛到进化稳定策略 s_1 (即进化稳定点 $x_1 = 1$).
- 复制动态趋向于效率较高均衡的 "机会" 是 $11/61$

鹰鸽博弈的进化博弈分析

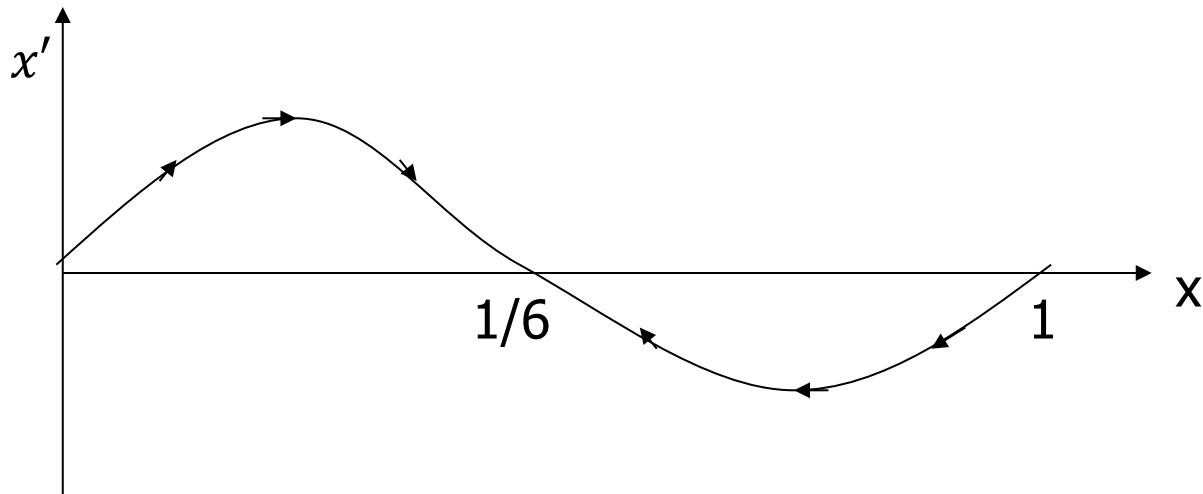
鹰鸽博弈

- 鹰鸽博弈研究的是种群内部竞争和冲突中的策略和均衡问题. 其中:
 - 策略 s_1 表示“鹰”, 即攻击型策略
 - 策略 s_2 表示“鸽”, 即和平型策略
- 鹰鸽博弈的进化博弈分析可以揭示人类社会或动物世界发生战争或激烈冲突的可能性及其频率, 国际关系中霸道和软弱、侵略和反抗等.

	s_1	s_2
s_1	$(v - c)/2, (v - c)/2$	$v, 0$
s_2	$0, v$	$v/2, v/2$

鹰鸽博弈：复制动态方程和相位图

- 记 t 时刻使用策略 s_1 个体的比例为 $x(t)$. 则复制动态方程如下:
- $x'(t) = x(1-x)[x(v-c) - (1-x)v]/2$
- 代入数值 $v = 2, c = 12$, 可以绘制出如下相位图:
只有 $x = 1/6$ 是进化稳定状态.



鹰鸽博弈启示

- 同协调博弈不同, 无论种群中初始时基因 s_1 的比例 x 是多少, 最后种群都会收敛到状态 $x = 1/6$.
- 进化博弈分析结论的现实意义:
 - 当战争和冲突的结果不确定, 失败方损失大大超过所争夺利益时, 长期进化会使大多数国家或其他冲突主体放弃攻击型策略, 和平主义会成为主流。和平共处将成为常态, 但仍然存在发生战争、冲突的少量风险。和平主义者会经常受霸权主义者欺负。这与当前国际政治、军事、经济、外交的现实格局很相似。

小结

- 本章介绍的博弈学习理论和进化博弈分析方法, 反映了有限理性博弈与完全理性博弈的差异, 对于更好地理解博弈关系, 特别是现实条件下的各种长期稳定关系、群体行为特征等, 有非常重要的意义.
- 核心概念: 最优反应动态, 进化稳定策略, 复制动态方程.
- 鹰鸽博弈的进化博弈分析结果说明, 进化稳定这个概念也可以用来对纳什均衡进行筛选. 尽管鹰鸽博弈本身有三个纳什均衡, 但只有一个进化稳定状态. 鹰鸽博弈中, 当博弈方理性程度较低或彼此之间缺乏信任时, 双方均使用进化稳定策略是更合理的预测.

一般两人对称博弈复制动态和进化稳定策略

一般模型

	策略1	策略2
策略1	a, a	b, c
策略2	c, b	d, d

一般 2x2 对称博弈

- 进化博弈设定是在一个大群体的成员中进行随机配对的反复博弈。
- 基本模型是两个博弈方之间的对称博弈。含义是两个博弈位置是无差异的。
- 其中 a, b, c, d 可以是任何得益，根据问题设定。

复制动态分析

复制动态的进化规则是生物学中生物特征进化规则

设 x 为采用策略**1**的比例

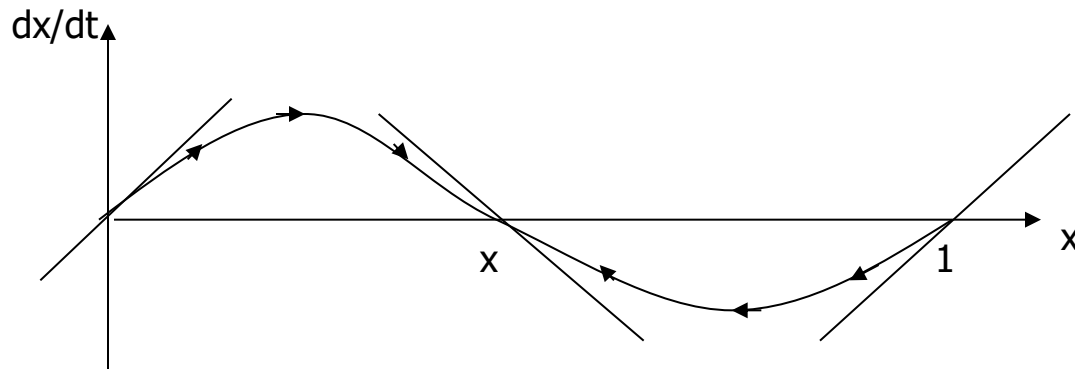
$$u_1 = x \times a + (1 - x) \times b$$

$$u_2 = x \times c + (1 - x) \times d$$

$$\bar{u} = x \times u_1 + (1 - x) \times u_2$$

$$\begin{aligned} \frac{dx}{dt} &= x(u_1 - \bar{u}) = x[u_1 - xu_1 - (1-x)u_2] \\ &= x(1-x)(u_1 - u_2) \\ &= x(1-x)[x(a-c) + (1-x)(b-d)] \end{aligned}$$

复制动态
相位图

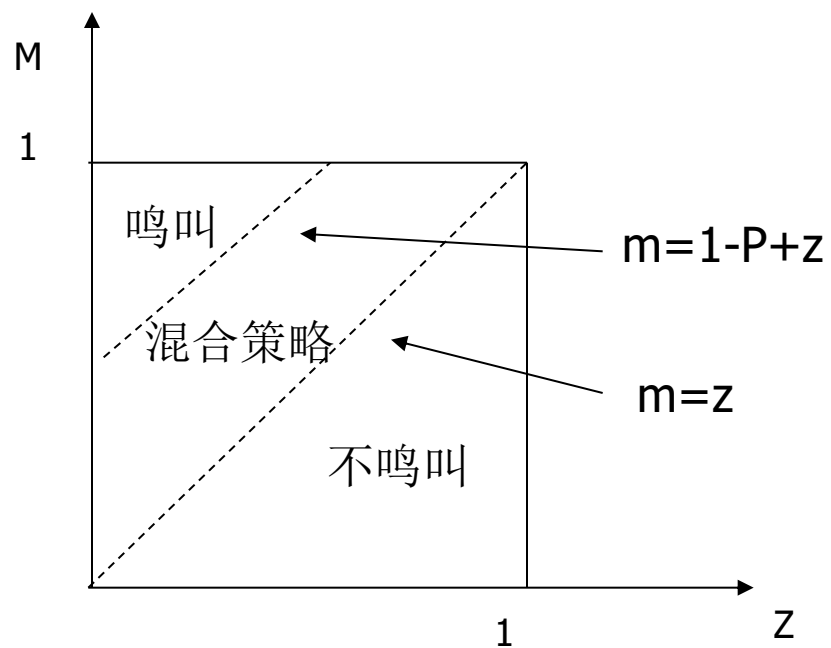


蛙鸣博弈的复制动态和进化稳定策略

蛙鸣博弈：

		雄蛙 2	
		鸣叫	不鸣
雄蛙 1	鸣叫	$P-z, P-z$	$m-z, 1-m$
	不鸣	$1-m, m-z$	$0, 0$

蛙鸣博弈不同均衡的条件



蛙鸣博弈复制动态方程和不动点

x ——鸣叫雄蛙比例

复制动态方程

$$\frac{dx}{dt} = x(1-x)[x(P-z-1+m) + (1-x)(m-z)]$$

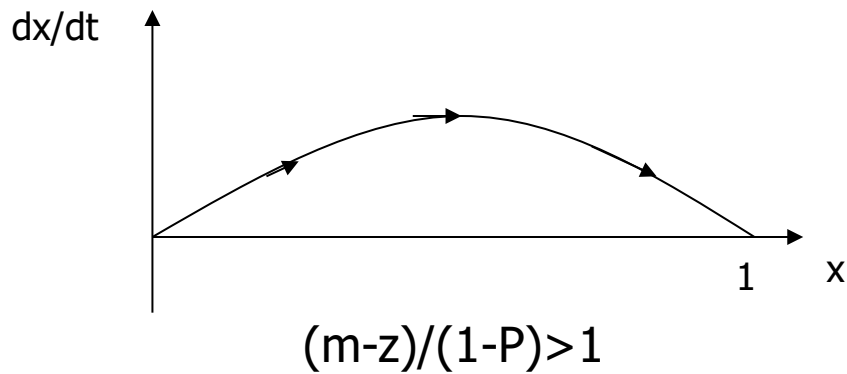
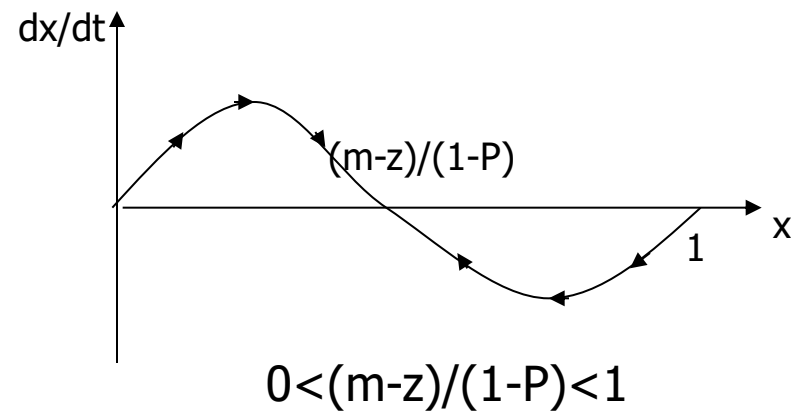
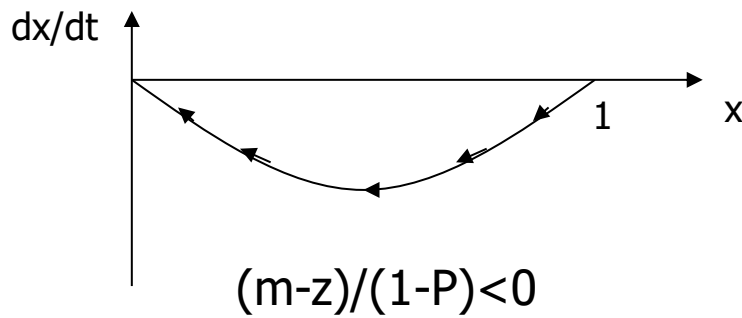
可能的稳定点：

$$x^*=0$$

$$x^*=1$$

$$x^*=(m-z)/(1-p)$$

蛙鸣博弈复制动态相位图



复制动态和进化稳定性：两人非对称 博弈

- 市场准入博弈
- 非对称鹰鸽博弈

市场准入博弈的复制动态和进化稳定策略

$$u_{1e} = y \times 0 + (1 - y) \times 2 = 2(1 - y)$$

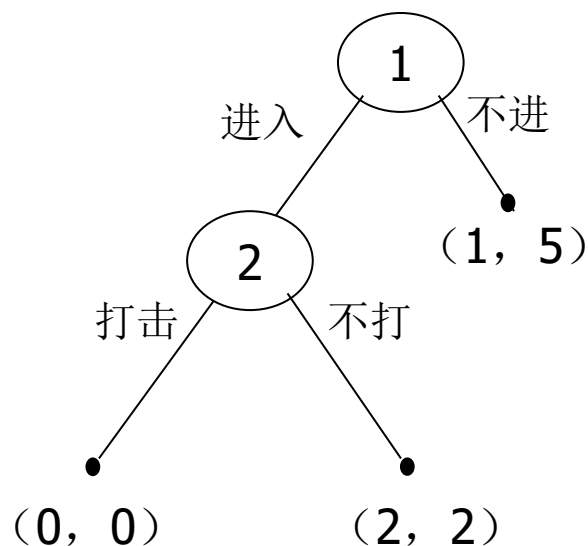
$$u_{1n} = y \times 1 + (1 - y) \times 1 = 1$$

$$\bar{u}_1 = x \times u_{1e} + (1 - x)u_{1n} = 2x(1 - y) + (1 - x)$$

$$u_{2s} = x \times 0 + (1 - x) \times 5 = 5 - 5x$$

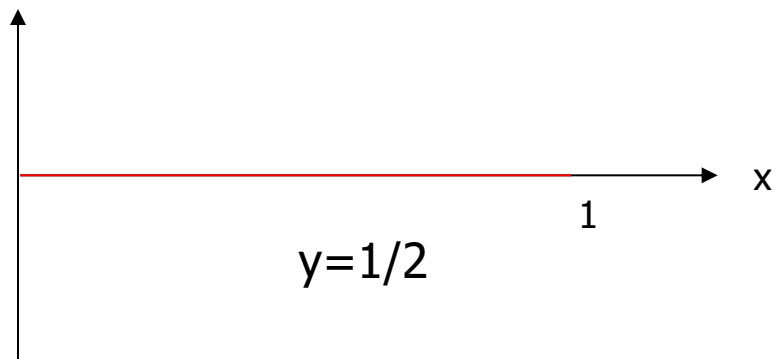
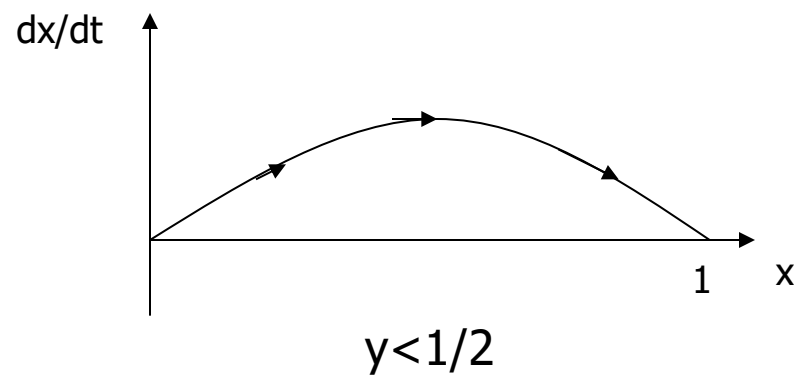
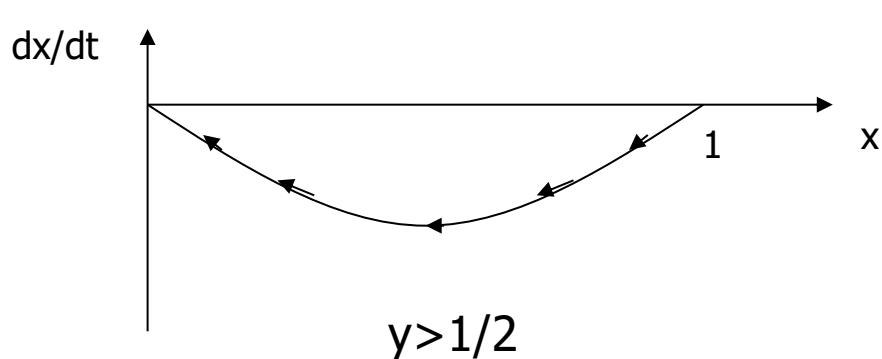
$$u_{2n} = x \times 2 + (1 - x) \times 5 = 5 - 3x$$

$$\bar{u}_2 = y \times u_{2s} + (1 - y)u_{2n} = 5 - 2xy - 3x$$



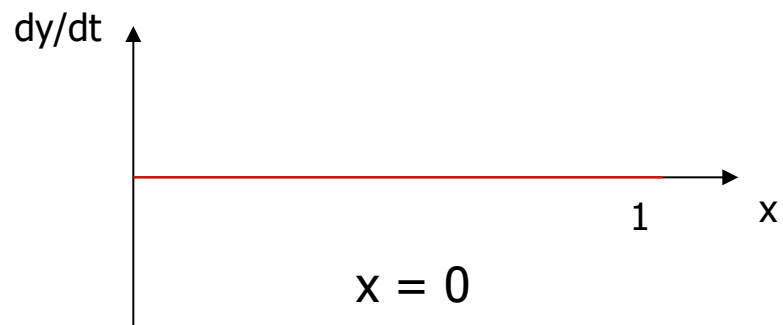
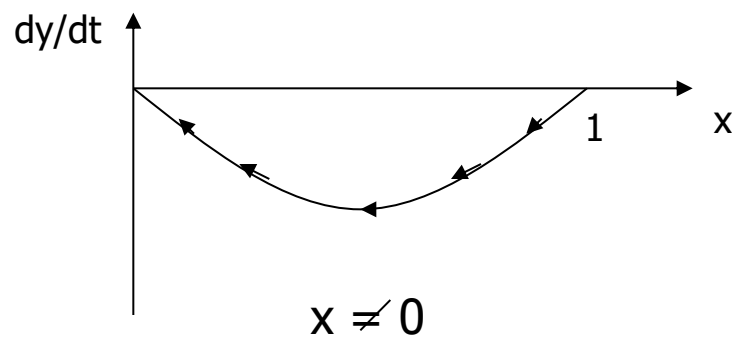
博弈方1位置博弈群体复制动态相位图

$$\frac{dx}{dt} = x[u_{1e} - \bar{u}_1] = x(1-x)(1-2y)$$

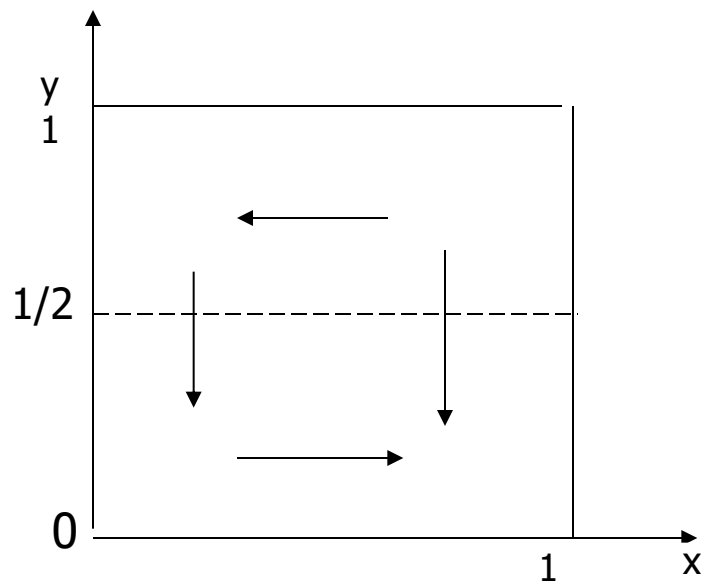


博弈方2位置博弈群体复制动态相位图

$$\frac{dy}{dt} = y[u_{2s} - \bar{u}_2] = y(1-y)(-2x)$$



两群体复制动态的关系和稳定性



非对称鹰鸽博弈的进化分析

$$u_{1e} = y \times (-1) + (1 - y) \times 10 = 10 - 11y$$

$$u_{1d} = y \times 0 + (1 - y) \times 5 = 5 - 5y$$

$$\overline{u_1} = x \times u_{1e} + (1 - x) \times u_{1d} = 5 + 5x - 5y - 6xy$$

$$u_{2e} = x \times (-5) + (1 - x) \times 2 = 2 - 7x$$

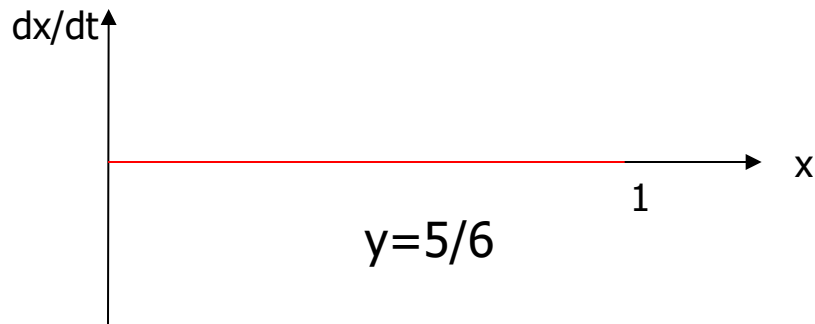
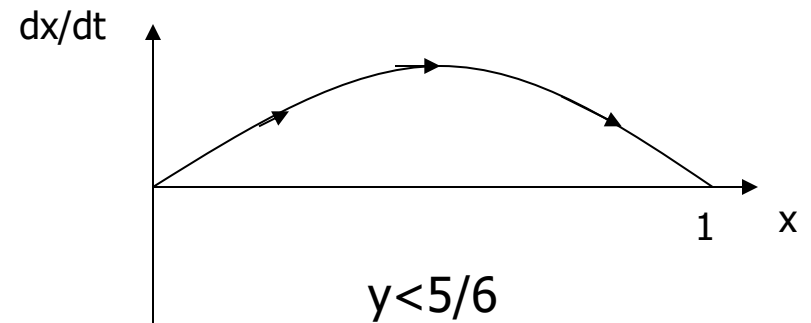
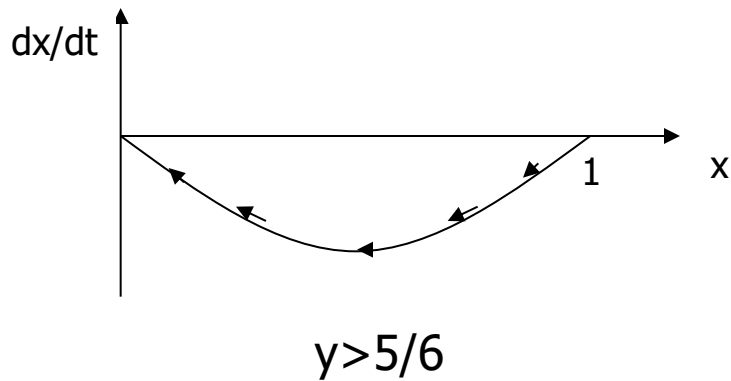
$$u_{2d} = x \times 0 + (1 - x) \times 5 = 1 - x$$

$$\overline{u_2} = y \times u_{2e} + (1 - y) \times u_{2d} = 1 - x + y - 6xy$$

博弈方2		鹰	鸽
博弈方1	鹰	$\frac{v_1 - c}{2}, \frac{v_2 - c}{2}$	$v_1, 0$
	鸽	$0, v_2$	$\frac{v_1}{2}, \frac{v_2}{2}$

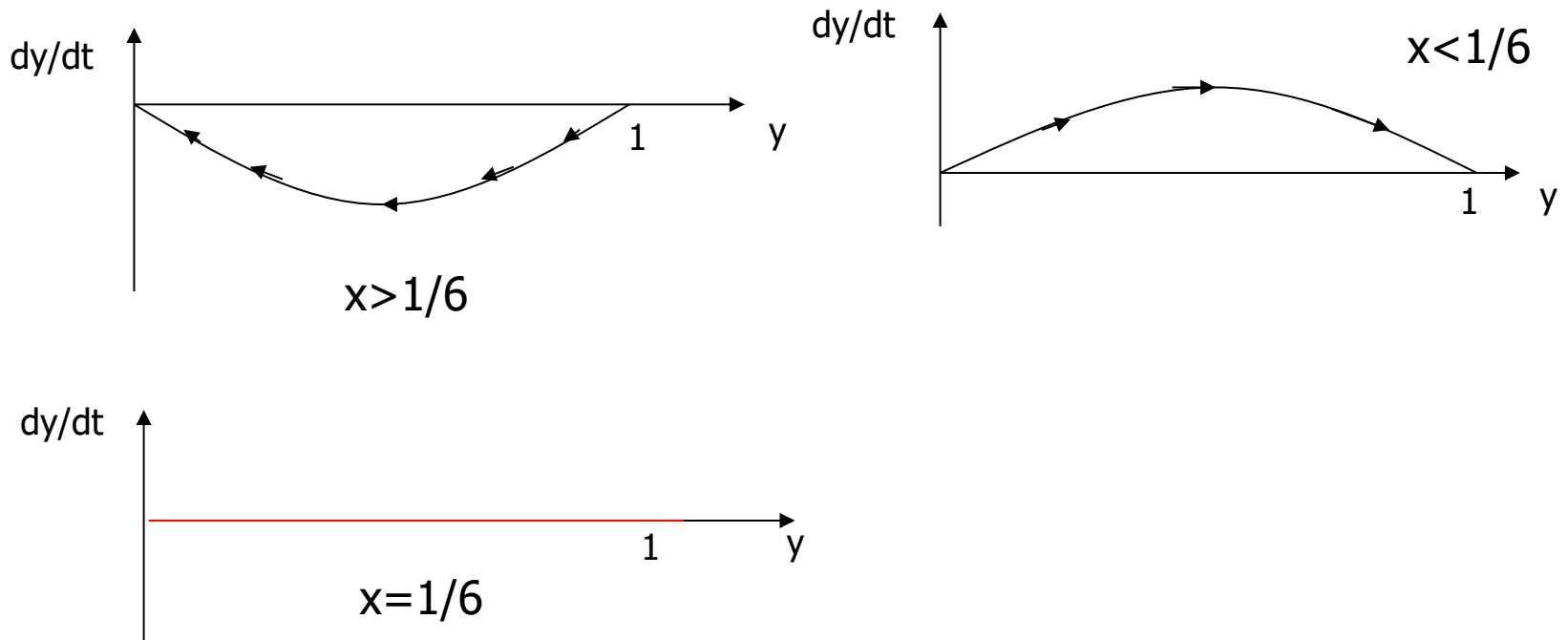
非对称鹰鸽博弈博弈方1群体复制动态相位图

$$\frac{dx}{dt} = x[u_{1e} - \bar{u}_1] = x(1-x)(5-6y)$$



非对称鹰鸽博弈博弈方2群体复制动态相位图

$$\frac{dy}{dt} = y[u_{2e} - \bar{u}_2] = y(1-y)(1-6x)$$



两群体复制动态关系和稳定性

