信息经济学

第八课：重复博弈

彭世喆

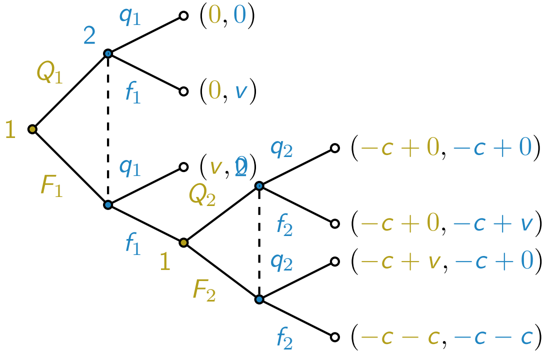
数字经济系
长沙理工大学经济与管理学院

消耗战游戏

- 两位玩家每阶段**同时**选择进攻还是撤退
- 当有人撤退时，游戏就结束了
- 好消息：如果对手先撤退，你将得到奖励 $v = 1$
- 坏消息：如果两人都选择进攻，则每人都需要付出成本 $c = 0.75$
- 如果两人同时撤退，则都得到 0
- 玩游戏（允许中途交流一次；一直进攻，越亏越多，远超奖励；成本已沉没，每一次都跟第一次玩一样；好斗形象）
- 例子：长期战争消耗，长期商业竞争消耗，行贿竞赛（钱不能要回）
- 理性人面对很小的奖励时可能会陷入消耗战

两阶段游戏

- 假设如果两人在第二轮都进攻，则都得不到奖励，游戏结束
- 效用的第一个 $-c$ 是沉没成本，与决策无关
- 假设 $v > c$



子博弈完美纳什均衡（第二个子博弈）

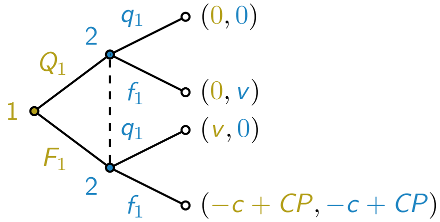
- 两个纯策略纳什均衡 (F_2, q_2) 和 (Q_2, f_2)

		玩家 2	
		f_2	q_2
玩家 1	F_2	$-c, -c$	$\underline{v}, \underline{0}$
	Q_2	$\underline{0}, \underline{v}$	$0, 0$

沉没成本 $-c+$

子博弈完美纳什均衡（第一个子博弈）

- CP =Continuation payoffs，即阶段二的纳什均衡效用
- 当玩家们知道 (F_2, q_2) 是阶段二的纳什均衡时，阶段一的 $NE = (F_1, q_1)$
- 如果预判自己未来会获胜，那么自己现在进攻就可以取得胜利



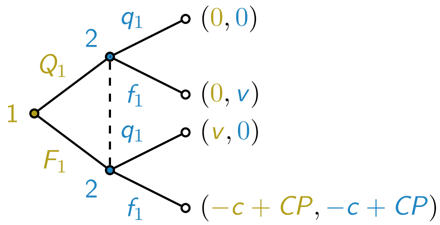
玩家 2

玩家 1

	f_1	q_1
F_1	$\underline{-c+v}, -c+0$	$\underline{v}, \underline{0}$
Q_1	$0, \underline{v}$	$0, 0$

子博弈完美纳什均衡（第一个子博弈）

- 当玩家们知道 (Q_2, f_2) 是阶段二的纳什均衡时，阶段一的 $NE = (Q_1, f_1)$



玩家 2

玩家 1		f_1	q_1
	F_1	$-c+0$, $\underline{-c+v}$	$\underline{v}$, 0
	Q_1	$\underline{0}$, $\underline{v}$	$0, 0$

子博弈完美纳什均衡

- 纯策略子博弈完美纳什均衡：$[(F_1, F_2), (q_1, q_2)]$ 和 $[(Q_1, Q_2), (f_1, f_2)]$
- 两个子博弈完美纳什均衡都有一个勇者和懦夫（如果知道对手是个懦夫，我就当一个勇者）
- 没有真的消耗战出现，与实际不符
- 寻找包含消耗战的子博弈完美纳什均衡

子博弈完美纳什均衡（第二个子博弈）

- 考虑混合策略
- $-c * p + v * (1 - p) = 0 * p + 0 * (1 - p) \rightarrow p = \frac{v}{v+c}$
- 由对称性可得：混合策略纳什均衡 $= ((\frac{v}{v+c}, \frac{c}{v+c}), (\frac{v}{v+c}, \frac{c}{v+c}))$，期望效用为 0

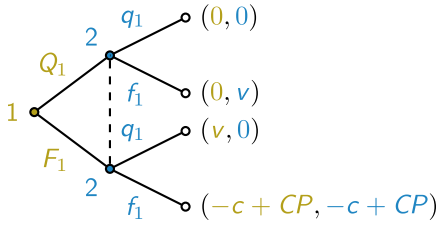
玩家 2

玩家 1	f_2	q_2
F_2	$-c, -c$	$\underline{v}, \underline{0}$
Q_2	$\underline{0}, \underline{v}$	$0, 0$

沉没成本 $-c+$

子博弈完美纳什均衡（第一个子博弈）

- 当玩家们知道阶段二是混合策略纳什均衡时，阶段一的效用矩阵与阶段二的相同
- 混合策略下的子博弈完美纳什均衡 $= ((p,p),(p,p))$，期望效用是 0，游戏以概率 p^2 继续下去
- 进攻概率 p 随着 v 增加，随着 c 减少



		玩家 2	
		f_1	q_1
玩家 1	F_1	$\underline{-c+0}, -c+0$	$\underline{v}, \underline{0}$
	Q_1	$0, \underline{v}$	$0, 0$

无穷阶段游戏

- 博弈树怎么画？
- 对于任一阶段，仍是沉没成本 + 当期效用 $+CP$
- 假设玩家们在未来采用混合策略，则 $CP = 0$
- 双方都以概率 p 进攻
- 消耗战持续多轮的概率随着阶段数递减。长期消耗战不太可能发生

设定

- 现实许多关系不是用契约等外在强制力去保证的，而是重复的
- 重复交互（Repeated Interaction）能够促进合作，让我们走出囚徒困境吗？
- 玩多次囚徒困境游戏

经验 18

在持续的关系中，未来可能获得的奖励以及未来可能面临的惩罚，有时能够激励人们在当下选择合作。但若要让此结论生效，必须要有一个明确的未来。

分析

- 最后一阶段是一次性游戏
- 倒数第二阶段的决策不会影响下一阶段的行为，依此类推，最终所有阶段都不合作
- 末期效应（End effect）：任期将要结束会损害某人为他人提供合作激励的能力（连任失败、退休、短时间关系）

		B	
		coop	*defect*
A	*coop*	$2, 2$	$-1, 3$
	defect	$3, -1$	$0, 0$

B

coop *defect*

设定

- 即使有一个已知的终点，仍然有合作的可能
- 如下游戏玩两轮
- 纯策略纳什均衡为 (B, B) 和 (C, C)
- 想维持 (A, A) 作为合作结果，但是 (A, A) 不是一次性游戏的纳什均衡

		玩家 2		
		A	B	C
	A	$4, 4$	$0, \underline{5}$	$0, 0$
玩家 1	B	$\underline{5}, 0$	$\underline{1}, \underline{1}$	$0, 0$
	C	$0, 0$	$0, 0$	$\underline{3}, \underline{3}$

分析

- 虽然在阶段二不能维持 (A, A)，但在阶段一可能
- 考虑策略：先选 A。然后如果上阶段结果是 (A, A)，则选 C，否则选 B
- 验证这是一个策略（$1 + 9 = 10$ 个信息集）

玩家 2

玩家 1		A	B	C
	A	$4, 4$	$0, \underline{5}$	$0, 0$
	B	$\underline{5}, 0$	$\underline{1}, \underline{1}$	$0, 0$
	C	$0, 0$	$0, 0$	$\underline{3}, \underline{3}$

子博弈完美纳什均衡

- 在阶段二
  - 在 (A, A) 之后，这个策略会选择 (C, C)
  - 在其他结果之后，这个策略会选择 (B, B)
- 在阶段一
  - 选择 A 的效用 $=4+3=7>$ 背叛选 B 的效用 $=5+1=6$
  - 或者，今天背叛的诱惑 $=5-4=1<$ 明天效用减少的惩罚 $3-1=2$

玩家 2

玩家 1	A	B	C
A	4, 4	0, <u>5</u>	0, 0
B	<u>5</u>, 0	<u>1</u>, <u>1</u>	0, 0
C	0, 0	0, 0	<u>3</u>, <u>3</u>

启示

经验 19

在有限游戏中，如果一个一次性游戏有多个纳什均衡，那么可以通过在未来选择不同均衡这一方式来为今天的合作提供激励。（一个作为奖励，一个作为惩罚。）

- 重新谈判可能会阻碍上述动机：假设阶段一有一方选择了 B，之后他找到对方，说服其在阶段二选择 C 而非原定的 B（让过去成为过去、说服对方这个结果更好）
- 在 19 世纪的英国，借钱不还的人要被抓去坐牢，但是坐牢的费用需要债权人支付（别抓我，让我出去工作还能偿还一部分债务）
- 但人们也会有愤怒和惩罚背叛的情况

概率型囚徒困境

- 之前的游戏都有明确的最后阶段
- 在概率型囚徒困境中，每次博弈结束后，以概率 $\delta < 1$ 继续（如抛一枚硬币决定）
- 游戏有一个随机的持续周期

		B: *coop*	B: *defect*
A	*coop*	$2, 2$	$-1, 3$
A	*defect*	$3, -1$	$0, 0$

分析

- 冷酷触发策略（Grim trigger strategy）：选择 *C*，然后如果没人背板，则选择 *C*；否则选择 *D*
- 验证这是一个 SPE：今天背叛的诱惑 $= 3 - 2 = 1 \leq$ 折现乘以在合作后明天的奖励减去在背叛后明天的惩罚 $\delta(2 + 2\delta + 2\delta^2 + \cdots - 0) = 2$
- 明天的奖励是指到游戏结束一直合作的效用，明天的惩罚是指到游戏结束结果一直是背叛的效用

		B: *coop*	B: *defect*
A	*coop*	$2, 2$	$-1, 3$
A	*defect*	$3, -1$	$0, 0$

分析

- $1 \leq \delta(2/(1-\delta)-0) \rightarrow \delta \geq 1/3$
- 未来的奖励和惩罚需要可信度（Credibility），是真实存在的
- 比如在有限阶段的囚徒困境下一直合作就不可信

A \ B	*coop*	*defect*
coop	$2, 2$	$-1, 3$
defect	$3, -1$	$0, 0$

其他偏离方式

- 为什么是永远的 2 或 0 呢？
- 一人采取冷酷触发策略
- 另一人良心发现 $D, C, D, D, D, \ldots$
  - $3 + \delta(-1) + 0 + 0 + \ldots$ 比一直选 D 更差
- 或另一人在后面阶段才背叛 $C, D, D, D, D, \ldots$
  - 当 $\delta \geq 1/3$ 时不值得这样做，因为到了后面阶段往后看，它就是第一阶段，与之前的分析一样

B

A		coop	defect
	coop	$2, 2$	$-1, 3$
	defect	$3, -1$	$0, 0$

启示

- 在囚徒困境中，如果 $\delta \geq 1/3$，则可使用冷酷触发策略作为子博弈完美纳什均衡，以实现合作
- δ 代表未来的权重

经验 20

在持续的关系中，为了为当下合作行为提供激励，让关系以高概率继续延续是有帮助的。

- 例子：大四情侣因毕业异地而分手，模型预测情侣会背叛对方；在小区开小卖部时采用诚信策略，而在旅游景点开小卖部时就采用欺骗策略

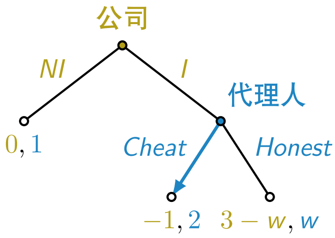
只惩罚一次的策略

- 冷酷触发策略太过严苛和极端，小的或无意的背叛也毫无挽回的余地（比如今天去楼下粉店吃早餐，老板一时疏忽给我放了太多的盐，于是我就一辈子都不去这家粉店了 ?）
- 考虑只惩罚一次的策略：先选择 C。然后如果上阶段结果是 (C, C) 或 (D, D)，则选择 C；否则选择 D
- 验证这是一个 SPE：今天背叛的诱惑 $= 3 - 2 = 1 \leq$ 折现乘以在合作后明天的奖励减去在背叛后明天的惩罚 $\delta(2/(1-\delta) - 0 - \delta \times 2/(1-\delta)) \to \delta \geq 1/2$（需要更大的未来权重）

		B	
		coop	defect
A	coop	$2, 2$	$-1, 3$
	defect	$3, -1$	$0, 0$

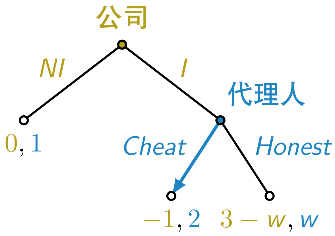
设定

- 假设一家公司正计划投资一个海外新兴市场，成本为 1，投资成功的收益为 4
- 这个市场的人力成本很低，令为 1
- 但司法不健全，合同难以得到保障
- 拟以工资 w 雇佣一个代理人，代理人可能偷走投资 1，并可以重新找到工资为 1 的工作



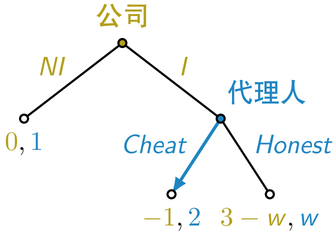
一阶段游戏

- 重复道德风险
- 机制设计（Incentive design）
- 要想代理人采取诚信策略，则要求工资 $w \geq 2$
- 在均衡中，$w^* = 2 + \epsilon$，代理人认真工作，但工资溢价为 100%



重复游戏

- 长期投资：投资继续的概率为 δ（贸易政策、战争）
- 在均衡中，今天欺骗的诱惑 $= 2 - w = 1 \leq$ 折现乘以在诚实后明天的奖励减去在欺骗后明天的惩罚（被解雇）$\delta(w/(1-\delta) - 1/(1-\delta)) \rightarrow w^{**} \geq 2 - \delta$（在 0 和 2 之间）
- 当 $\delta = 0$ 时，$w^{**} = 2$（一阶段游戏的结果）；当 $\delta = 1$ 时，$w^{**} = 1$（市场价）；当 $\delta = 0.5$ 时，$w^{**} = 1.5$（溢价只有 50%，显著降低）；



启示

经验 19

想要达成合作，必须要有奖励。未来的概率越小，奖励需要越大。

Thanks!