Game Theory and Applications (博弈论及其应用)

Chapter 10: Repeated Games

南京大学

高 尉

Recap on Previous Chapter

- The extensive game is an alternative representation that makes the temporal structure explicit
- Subgame perfect equilibrium (SPE): an outcome is SPE if it is Nash Equilibrium in every subgame
- How to find SPE – back induction and one deviation

后向归纳使用很简单，证明难
数学的优美之处就在于：结论很易，但证明繁杂

Repeated Games

前进的，不间断的

- Many real interactions have an ongoing structure
  - Firms compete over and over time again
  - Chinese and American compete repeatedly for the future
- In such situation, players should consider their long-term and short-term payoff simultaneously
- This yields behaviors which is different from one interaction (extensive and strategy games)

A Simple Example

Firm 2

Firm 1	High	Low
High	2 2	0 3
Low	3 0	1 1

only NE (Low, Low)

- What happens if this game is played only once
- What do you think might happen if played repeatedly
  - Being caught cheating will yield punishment 产生
  - We give up payoffs now in the expectation that we will be paid back later
  - Is cooperation always good? 合作

暗示的，盲从的，含蓄的

Implicit Cooperation

- The firms cooperate with fixing prices (explicit cooper.)
- Could firms cooperate without explicitly fixing prices?
- Some reward/punishment mechanisms are used to keep the firms in line
- Repeated interaction provides the opportunity to implement such mechanisms

  权衡

  – A firm faces a trade-off between short- and long-term profits
- Repeated games is a model to study these questions

机制–保证博弈不断进行下去

Repeated Games

- Players plays a stage game repeated over time
  - Stage game includes strategic and extensive game
- If there is a finial period: finitely repeated game
- If there is no definite end period: infinite repeated game
  - We could think of firm having infinite lives
  - Players do not know when the game will end

Repeated Games (cont.)

类似通货膨胀? 货币必须贬值

- Denote the discount factor (贴现因子) by $\delta \in (0,1)$
  - Control the short-term and long-term profits
- Today's \$1 payoff is more valuable than tomorrow's \$1
  - Represents how patient the players are 有多少耐心
  - Think as probability with which the game will play next time 下一次发生的概率
  - Think as the factor to calculate the values for different period
  - Guarantee the convergence of payoff 收敛

Payoffs

δ 可以是变量（随时间变化）
这里简化了

- If starting now, a player receives an infinite sequence of payoffs $u_1, u_2,\ u_3, \ldots$ i.e., the payoff is u_t for each stage
- Discount factor $\delta \in (0,1)$. Payoff is defined by

$$u_1 + \delta\, u_2 + \delta^2 u_3 + \delta^3 u_4 + \cdots$$

Example: Period payoffs are all 2

Firm 2

Firm 1	High	Low
High	2 2	0 3
Low	3 0	1 1

$$2 + 2\delta\ + 2\delta^2 + 2\delta^3 + \cdots = \frac{2}{1-\delta}$$

Repeated Prisoners' Dilemma

		Prisoner 2	
		Confess(c)	Don't confess(d)
Prisoner 1	Confess(c)	0 0	2 -1
	Don't confess(d)	-1 2	1 1

- Suppose two players are going to play the prisoner's dilemma game for $t = 1, 2, \ldots, T + 1$
- The discount factor is $\delta \in (0,1)$
- What's the subgame perfect Nash equilibrium?
- Is cooperation is always good?

前T次都(D, D)
最后一次放手

Repeated Prisoners' Dilemma

Assume Prisoner's Dilemma game proceeds $T + 1$ periods

- If one player is a "nice" guy, who plays "d" as long as you play "d" in all previous periods, then selects "c" for all future periods once you choose "c"
  He plays "nice" until you cheat him.

欺骗前是合作状态
欺骗后变为竞争

- If you consider the payoff by selecting "d" for first T periods, then choosing "c" in the final period. Then the payoff from the strategy (d, d,..., d, c) is:

$$1 + \delta + \delta^2 + \cdots + \delta^{T-1} + 2\delta^T = \frac{1 - \delta^T}{1 - \delta} + 2\delta^T$$

Repeated Prisoners' Dilemma

Assume Prisoner's dilemma game proceeds $T + 1$ periods

- If one player is a "nice" guy, who plays "d" as long as you play "d" in all previous periods, then selects "c" for all future periods once you choose "c"
  He plays "nice" until you cheat him.

- If you consider the payoff by selecting "d" for first $T - 1$ periods, then choose "c" in the final two periods. Then the payoff from the strategy (d, ..., d, c, c) is

$$1 + \delta + \delta^2 + \cdots + \delta^{T-2} + 2\delta^{T-1} + 0 = \frac{1 - \delta^{T-1}}{1 - \delta} + 2\delta^{T-1}$$

Repeated Prisoners' Dilemma

- If your strategy is (d, ..., d, c, c), then

$$1 + \delta + \delta^2 + \cdots + \delta^{T-2} + 2\delta^{T-1} + 0 = \frac{1 - \delta^{T-1}}{1 - \delta} + 2\delta^{T-1}$$

- If your strategy is (d, ..., d, d, c), then

$$1 + \delta + \delta^2 + \cdots + \delta^{T-1} + 2\delta^{T} = \frac{1 - \delta^{T}}{1 - \delta} + 2\delta^{T}$$

- By comparison, we have, if $\delta \leq 1/2$,

$$\frac{1 - \delta^{T-1}}{1 - \delta} + 2\delta^{T-1} \geq \frac{1 - \delta^{T}}{1 - \delta} + 2\delta^{T}$$

- This looks like the noncooperation is going to occur, even if one player is willing to cooperate

Formal Definition of Repeated Game

Definition A repeated game $G^{\mathrm{T}}(\delta)$ is

- a stage game of finite length: $G = \{N, \{A_i\}, \{u_i\}\}$, which is usually independent of the calendar date.
- a terminal date $T = 1,2, \ldots,$ giving the number of interact times. The calendar date is given by $t = 1, 2, \ldots, T$.
- a discount factor, $0 \leq \delta \leq 1$, that represents both how patient the players are and how likely the game continues.

- If $a^t = (a_1^t, a_2^t, \ldots, a_N^t)$ is the strategy outcome that occurs in period t, the player i's payoff is

$$u_i(a^1) + \delta u_i(a^2) + \cdots + \delta^{T-1} u_i(a^T) = \sum_t \delta^{t-1} u_i(a^t)$$

History

- **Perfect information**: Players keep track of how players behave in previous periods; so as to choose strategies that reward or punish players for good or bad behavior.
- How to track what happens in repeated games?-**History**
- In prisoners' dilemma, all of the possible outcomes from the stage game are

$$\Sigma = \{(\mathrm{d}, \mathrm{d}), (\mathrm{d}, \mathrm{c}), (\mathrm{c}, \mathrm{d}), (\mathrm{c}, \mathrm{c})\}$$

- For the second period, there are 16 outcomes since we keep track of what happens in the first and second period

$$\Sigma^2 = \{[(\mathrm{d}, \mathrm{d}), (\mathrm{d}, \mathrm{d})], [(\mathrm{d}, \mathrm{d}), (\mathrm{d}, \mathrm{c})], [(\mathrm{d}, \mathrm{d}), (\mathrm{c}, \mathrm{d})], \ldots\}$$

History

- Let Σ be the set of all the strategy outcomes for the stage game. (for instance, $\Sigma = \{(\mathrm{d},\mathrm{d}),(\mathrm{d},\mathrm{c}),(\mathrm{c},\mathrm{d}),(\mathrm{c},\mathrm{c})\}$ for prisoners' dilemma)

- If we want to keep track of the outcome of a repeated game, we're interested in sequence observations from Σ.
  - For two periods, $\Sigma^2 = \Sigma \times \Sigma$ is the set of all possible outcomes of two repetitions of the game
  - For three periods, $\Sigma^3 = \Sigma \times \Sigma \times \Sigma$ is the set of all possible outcomes of three repetitions of the game
  - and so on

History and SPNE

Definition Let Σ be the set of all strategy outcomes for one stage game, and let $\Sigma^{\mathrm{t}} = \Sigma \times \Sigma \times \cdots \times \Sigma$ denote all possible outcomes.

A **history** at time t is an element $h_t \in H_t = \Sigma^t$.

Definition A set of strategies is a **Subgame Perfect Nash Equilibrium (SPNE)** of a repeated game if, for any t-period history h_t, there is no subgame in which any player has a profitable deviation. 越轨，偏离

- No player can have a profitable deviation for any history, even if only one history actually occurs
- The players know the consequences of their actions

SPEN in Repeated Games

Proposition If one stage game has an Nash equilibrium $a^* = (a_1^*, \dots, a_N^*)$, then the strategy

$$(a^*, a^*, \dots, a^*)$$

is a subgame perfect Nash equilibrium (SPNE) of the repeated game, i.e., each player i plays a_i^* for every history.

Is this the only equilibrium of a repeated game?

SPNE of Finite Repeated Game

Theorem Consider a repeated game $G^T(\delta)$ with $T < +\infty$. Suppose that the stage game G has an unique pure strategy NE a^*. G^T has a unique SPNE with $a^t = a^*$ for each t.

Proof We use the backward induction.

For period T, we will have $a^T = a^*$ regards of history.

For period T-1, we also have $a^{T-1} = a^*$

…

By Induction, we have $a^t = a^*$ for $1 \leq t \leq T$

Infinite Repeated Prisoners' Dilemma

		Prisoner 2	
		Confess(c)	Don't confess(d)
Prisoner 1	Confess(c)	0 0	2 -1
	Don't confess(d)	-1 2	1 1

改变只有一次机会
对手会接着改变

- Consider the following strategies
  - If the history at time t is {(d, d), ... , (d, d)}, play d
  - Else play c

之前是合作，则继续
否则竞争

Is this a SPNE of the infinite repeated game?

Infinite Repeated Prisoners' Dilemma

Check all the possible histories for profitable deviations.

There's really just two cases:

$$\{(d, d), (d, d), \dots, (d, d)\}$$

and anything else.

- Suppose the history is not $\{(d, d), (d, d), \dots, (d, d)\}$. In now and future periods, the opponent will choose c. You should choose c. So there are no profitable deviations from these histories.
- Suppose the history is $\{(d, d), (d, d), \dots, (d, d)\}$. Is playing d an optimal strategy?

Infinite Repeated Prisoners' Dilemma

- Suppose the history is {(d, d), (d, d), ... }, the payoff is

$$1 + \delta + \delta^2 + \cdots + \delta^t + \cdots$$

- Suppose the history is not {(d, d), (d, d), ... }

$$2 + 0\delta + 0\delta^2 + \cdots$$

- So it all comes down to whether it's better to cooperate than cheat in any periods,

$$\frac{1}{1-\delta} > = < 2$$

δ小于1/2（即1/(1-x)<2）时早点竞争好

Infinite Repeated Prisoners' Dilemma

If $\delta > 1/2$, both players using the strategy

- if the history at time t is $\{(\mathrm{d}, \mathrm{d}), \ldots, (\mathrm{d}, \mathrm{d})\}$, play d;
- for any other history at time t, play c.

is a Subgame Perfect Nash Equilibrium of the infinitely repeated prisoners' dilemma.

Bertrand Model(伯特兰德模型)

边际成本

- Two firms 1 and 2 have the same marginal costs c and compete in prices: $p_1, p_2 \in \{0, 1, \dots, c, \dots, 9, 10\}$
- Fixed demand = 1
- If $p_1 < p_2$, all the consumers go to firm 1
- If $p_1 = p_2$, the firms split the market equally
- If $p_1 > p_2$, all the consumers go to firm 2
- The payoffs for firm 1 are:

$$u_1(p_1, p_2) = \begin{cases} p_1 - c & \text{if } p_1 < p_2 \\ (p_1 - c)/2 & \text{if } p_1 = p_2 \\ 0 & \text{if } p_1 > p_2 \end{cases}$$

and similarly for firm 2.

Nash Equilibrium for Bertrand Model

- $$B_1(p_2) = \begin{cases} \{p_1 : p_1 < p_2\} & \text{if } p_2 < c \\ \{p_1 : p_1 \leq p_2\} & \text{if } p_2 = c \\ \{p_1 : p_1 > p_2\} & \text{if } p_2 > c \end{cases}$$

- $$B_2(p_1) = \begin{cases} \{p_2 : p_2 < p_1\} & \text{if } p_1 < c \\ \{p_2 : p_2 \leq p_1\} & \text{if } p_1 = c \\ \{p_2 : p_2 > p_1\} & \text{if } p_1 > c \end{cases}$$

- The Nash Equilibrium is (c,c)

Repeated Bertrand Model

Consider the repeated game:

- $T = \infty$
- Discount factor $0 < \delta < 1$

（0～10，0～10）共121种

- Stage game: Bertrand Competition

0～10 增长

Notice that there are 11 price increments, and 121^t possible outcomes time t. If t = 5, there are 45,444,082,772 histories. On a very good computer, computing the extensive form and payoffs would take a lot of time.

Repeated Bertrand Model

- We know $p_1^* = p_2^* = c$ is a Nash equilibrium of the stage game. Let's use this as the "punishment" for a breakdown in cooperation.

- The maximum payoff is $p_1 = p_2 = 10$.

- Consider the strategies:

  – If the history is $\{(10, 10), (10, 10), (10, 10), \ldots, (10, 10)\}$, then play 10 this period.

  – For any other history, play c this period.

- Is it an subgame perfect Nash equilibrium?

Analysis of Repeated Bertrand Model

Suppose the history is $\{(10, 10), (10, 10), \ldots, (10, 10)\}$, then the payoff of cooperation is

$$\mathrm{I} = \frac{10 - c}{2}(1 + \delta + \cdots + \delta^{t-1} + \cdots)$$

Suppose a deviation occurs at time t, and the opponent uses the best strategies 9, in all future periods. The payoff is

$$\mathrm{II} = \frac{10 - c}{2}(1 + \delta + \cdots + \delta^{t-1}) + \delta^{t}(9 - c) + 0 + \cdots$$

Analysis of Repeated Bertrand Model

The cooperation is better than deviation if I $\geq$ II, that is

$$\frac{10 - c}{2(1 - \delta)} \geq 9 - c$$

or

$$\delta \geq \frac{8 - c}{18 - 2c}$$

For example, we have

$\delta \geq 3/7$ for $c = 2$

$\delta \geq 1/3$ for $c = 6$

SPNE for Repeated Bertrand Model

As long as

$$\delta \geq \frac{8-c}{18-2c}$$

The strategy

- If the history is $\{(10, 10), (10, 10), (10, 10), \ldots, (10, 10)\}$, then play 10 this period.

- For any other history, play c this period.

is a Subgame Perfect Nash Equilibrium of the infinitely repeated Bertrand game. So cooperation is possible in the infinite-horizon version of the repeated game.

The Steps for Repeated Games

- Solve for all of the equilibria of the stage game (Competitive Play)
- Find a strategy profile that gives all the players a higher payoff (Cooperative Play)
- Enforce cooperation: If all players have previously cooperated, continue cooperating. If any player has previously defected, play competitively
- For sufficiently large discount factor δ, this will be an equilibrium of the repeated game

Examples of Equilibria in Repeated Games

- For any game, playing the equilibrium of the stage game forever is a Subgame Perfect Nash Equilibrium.

- In prisoners' dilemma, as long as the discount factor is not smaller too much ($\delta \geq 0.5$), one SPNE of the game was in all periods unless your opponent had previously confessed at some point, and then to confess forever.

We will show "As long as players are patient, they can cooperate in infinitely repeated games in ways that aren't possible in finitely repeated games"

Folk Theorem

- Consider a N-player infinitely repeated game with a stage game equilibrium $a^* = (a_1^*, a_2^*, \dots, a_N^*)$ with payoffs $u^* = (u_1^*, u_2^*, \dots, u_N^*)$.

- Suppose there is another $\hat{a} = (\hat{a}_1, \hat{a}_2, \dots, \hat{a}_N)$ with payoffs $\hat{u} = (\hat{u}_1, \hat{u}_2, \dots, \hat{u}_N)$, where, for every player i,

$$\hat{u}_i \geq u_i^*$$

For some discount factor δ, there is a Subgame Perfect Nash Equilibrium in which the players use $\hat{a}$ in every period of the infinitely repeated game.

The Folk Theorem for Repeated Prisoner Dilemma

Consider prisoners' dilemma with a stage game equilibrium $a^* = (c, c)$ with payoffs $u^* = (0,0)$.

There is another $\hat{a} = (d, d)$ with payoffs $\hat{u} = (1,1)$, and we have $\hat{u} \geq u^*$.

For some discount factor δ, there is a Subgame Perfect Nash Equilibrium in which the players use (d, d) in every period of the infinitely repeated game.

How to prove the Folk Theorem?

The Folk Theorem: Trigger Strategies (触发策略)

Consider the following trigger strategy for player i:

- If the history at t is $h_t = (\hat{a}, \hat{a}, \ldots, \hat{a})$, play $\hat{a}_i$ in period t
- For any other history at time t, play a_i^* in period t

This is called a "trigger strategy" because it starts in "cooperative" mode, but after any defection by any player, it switches to "punishment" or "competitive" mode, and they play the stage game strategies forever.

The Folk Theorem: Optimal Deviations

Since $\hat{u}$ is presumably not a Nash equilibrium of the stage game, there are at least some players (player j) for whom

$$u_j^d > \hat{u}_j \geq u_j^*.$$

While they prefer cooperating to the equilibrium of the stage game, they prefer defection to cooperation.

The above inequality implies

$$u_j^d - u_j^* \geq u_j^d - \hat{u}_j$$

The Folk Theorem: Cooperating and Deviating

The payoff to cooperating to player j is

$$\hat{u}_j + \hat{u}_j\delta_j + \cdots + \hat{u}_j\delta_j^{t-1} + \cdots = \frac{\hat{u}_j}{1-\delta_j}$$

The payoff to deviating to player j is

$$u_j^d + u_j^*\delta_j + u_j^*\delta_j^2 + \cdots = u_j^d + \frac{u_j^*\delta_j}{1-\delta_j}$$

Then the cooperating is better than deviating for player j if

$$\frac{\hat{u}_j}{1-\delta_j} \geq u_j^d + \frac{u_j^*\delta_j}{1-\delta_j} \text{ implying } \delta_j \geq \frac{u_j^d-\hat{u}_j}{u_j^d-u_j^*}$$

But $\delta_j < 1$ from the previous slides.

The Folk Theorem: Equilibrium

We set

$$\delta^* = \max\{\delta_1, \delta_2, \dots, \delta_N\},$$

i.e., we select the highest discount factor for which cooperating is better than deviating for all the players.

If all players are sufficiently patient, i.e., each of their discount factors are greater than δ^*, then the trigger strategies are a subgame perfect Nash equilibrium

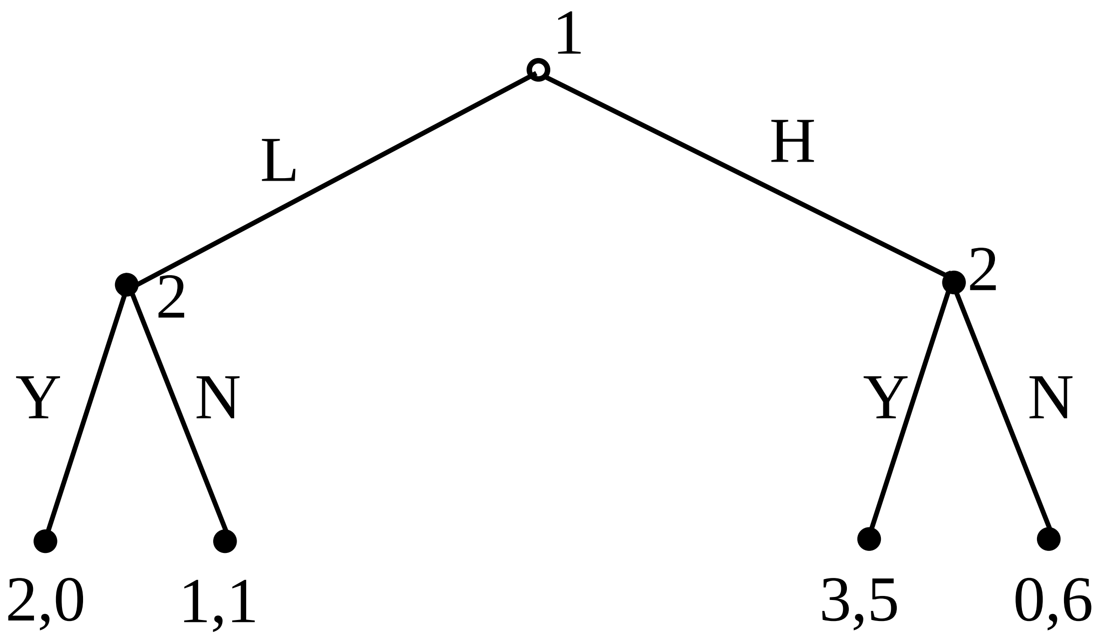
Players select $\hat{a}$ in every period to get highest profits.

Solving for Equilibria in Repeated Games

1. Solve all equilibria of the stage game (Competition)
2. Find an outcome (equilibrium/not) where all the players do at least as well as in the stage game (Cooperation)
3. Design trigger strategies that support cooperation and punish with competition
4. Compute the maximum discount factor so that cooperation is an equilibrium
5. The trigger strategies are an SPEN of the infinitely repeated game for some larger discount factor

Exercise

What happens for the following repeated extensive game



（1）竞争时，后向归纳法，1>0, 6>5 so NN, 1>0 so L
得(L, NN) (1, 1)
（2）有更好的吗？（保证每个玩家都有在增加）
(3, 5)这个 (H, ?Y)
（3）都是3,5 3,5 3,5
否则1,1
求δ

P1不会想变得，P2可能想变为6
不妨令在第一步就越轨
$5+5\cdot\delta+5\cdot\delta$^2 ...
$6+1\cdot\delta+1\cdot\delta$^2 ...

$5\cdot(1/(1-\delta))>5+1/(1-\delta)$
$\delta > 1/5$ 就能保持一直合作

Equilibria with Forgiveness

The trigger strategies are pretty harsh: Mess up once, and cooperation is cut off forever

How about if punish by playing the stage game equilibrium K rounds and then return to cooperative mode?

For repeated prisoners' dilemma, cooperating is better than deviating if

$$1 + \delta + \delta^2 + \cdots \geq 2 + 0 + \cdots + 0 + \delta^{K+1} + \delta^{K+2} + \cdots$$

$$\frac{1}{1-\delta} \geq 2 + \frac{\delta^{K+1}}{1-\delta}$$

For larger K and δ, the equality holds.

Equilibria with Forgiveness (cont.)

For repeated prisoners' dilemma, cooperating is better than deviating if

$$2\delta \geq 1 + \delta^{K+1} \quad \text{or} \quad K \leq \frac{\log(2\delta \quad 1)}{\log \delta} - 1$$

If we take limit as $\delta \to 1$ (use L'Hopital's rule twice), we get that the minimal punishment period is $K \to 0$. So players that are sufficiently patient will never cheat on each other.

Home work 1: Analysis of repeated Cournot Model

- Two firms compete by choosing how much to produce

$$G = \{\{1,2\},\{q_1,q_2\},\{u_1,u_2\}\}$$

  - Price

$$p(q_1+q_2) = \max(0, a - b(q_1+q_2))$$

  - Costs $(i = 1,2)$

$$c_i(q_i) = cq_i$$

  - Payoffs $(i = 1,2)$

$$u_i(q_1,q_2) = (\max\big(0, a - b(q_1+q_2)\big) - c)q_i$$

  - Condition $a > \mathrm{b}, c > 0, q_1 \geq 0, q_2 \geq 0$

Find SPNE and discount factor

下一章的开场白就是了！
我自己也算出了9/17

Home work 2: Analysis of Repeated Bertrand Model

- $N = \{1,2\}$; Price $\{q_1, q_2\}$; Market price q $= min\{q_1, q_2\}$;
- Demand $d(q) = a - q$; Cost of firm i is $c_i(x) = cx$
- Payoff $\{u_1, u_2\}$

$$u_1(q_1, q_2) = \begin{cases} q_1(a - q_1) - c(a - q_1) & \text{if } q_1 < q_2 \\ q_1(a - q_1)/2 - c(a - q_1)/2 & \text{if } q_1 = q_2 \\ 0 & \text{if } q_1 > q_2 \end{cases}$$

Here a $> c$.

Find SPNE and discount factor