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

Chapter 12: Repeated Games II

南京大学

高 尉

Recap on Previous Chapter

- Repeated game: many real interactions have an ongoing structure; players consider short- and long-term payoffs.
- A repeated game $G^{\mathrm{T}}(\delta)$ consists of stage game G, terminal date T and discount factor δ
- Folk Theorem
  - An infinitely repeated game with a stage game equilibrium $a^* = (a_1^*, a_2^*, \ldots, a_N^*)$ with payoffs $u^* = (u_1^*, u_2^*, \ldots, u_N^*)$.
  - Suppose there is another $\hat{a} = (\hat{a}_1, \hat{a}_2, \ldots, \hat{a}_N)$ with payoffs $\hat{u} = (\hat{u}_1, \hat{u}_2, \ldots, \hat{u}_N)$, where, $\hat{u}_i \geq u_i^*$ for every player i
  - There is a Subgame Perfect Nash Equilibrium for some discount factor δ

Construct SPNE 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

Repeated Cournot Competition

- Two firms compete by choosing how much to produce

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

  - Price

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

  - Costs $(i = 1, 2)$

$$c_i(q_i) = 0$$

  - Payoffs $(i = 1, 2)$

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

  - Condition $a > 0, q_1 \geq 0, q_2 \geq 0$

Step 1: Nash Equilibrium for One Stage

Best Response Correspondence is given by

$$B_i(q_{-i}) = \max(0, (a - q_{-i})/2)$$

The Nash equilibria is give by

$$q^* = (q_1^*, q_2^*) = \left(\frac{a}{3}, \frac{a}{3}\right)$$

The payoff is

$$u^* = (u_1^*, u_2^*) = \left(\frac{a^2}{9}, \frac{a^2}{9}\right)$$

What happens if two firms cooperate for their profits?

Maximal Payoff for Cooperation

Summing the firms' profits, we get

$$u_1 + u_2 = (a - q_1 - q_2)q_1 + (a - q_1 - q_2)q_2$$
$$= (a - q_1 - q_2)(q_1 + q_2)$$

Maximizing the above gives

$$q_1 + q_2 = a/2$$

The total payoff for cooperation: $a^2/4 = 2a^2/8$

The total payoff for completion: $2a^2/9$

Cooperation is potentially profitable

Step 2: Cooperation

Suppose the two firms are playing the Cournot game an infinite number of times, and they share a discount factor δ.

Let

$$\hat{q} = (\hat{q}_1, \hat{q}_2) = \left(\frac{a}{4}, \frac{a}{4}\right)$$

$$\hat{u} = (\hat{u}_1, \hat{u}_2) = \left(\frac{a^2}{8}, \frac{a^2}{8}\right)$$

In competitive model,

$$\hat{q} = (\hat{q}_1, \hat{q}_2) = \left(\frac{a}{3}, \frac{a}{3}\right)$$

$$\hat{u} = (\hat{u}_1, \hat{u}_2) = \left(\frac{a^2}{9}, \frac{a^2}{9}\right)$$

Step 3: Trigger Strategy

Consider the strategy:

- If the two firms have both used $\hat{q} = (a/4, a/4)$ in all previous periods, use $\hat{q}_j = a/4$ this period

- If either firm ever did anything besides $\hat{q}$, play the stage Cournot quantity $q_j^* = a/3$

Is this a subgame perfect Nash equilibrium of the infinitely repeated game?

Check the NE of Cooperative Strategy

To check whether $\hat{q} = (\hat{q}_1, \hat{q}_2) = (a/4, a/4)$ is a NE?

By symmetry, it is sufficient to check player 1. We solve

$$\max_{q_2}(a - \hat{q}_1 - q_2)q_2 = \max_{q_2}(a - a/4 - q_2)q_2$$

Maximizing the above gives

$$q_2' = \frac{3a}{8}, u_2'\left(\frac{a}{4}, \frac{3a}{8}\right) = \left(\frac{3a}{8}\right)^2$$

$\hat{\boldsymbol{q}} = (\hat{\boldsymbol{q}}_1, \hat{\boldsymbol{q}}_2) = (\boldsymbol{a}/4, \boldsymbol{a}/4)$ is not a NE

Step 4: Select Discounting Factor

For cooperating case, all players keep the cooperation model, and the payoff for player 2 is

$$\hat{u}_2(1 + \delta + \delta^2 + \cdots) = \frac{a^2}{8}\frac{1}{1-\delta}$$

For competitive case, deviating optimally in some period t after a history, and all players cooperated switches the game to competition. The pay off for player 2 is

$$u_2' + u_2^*(\delta + \delta^2 + \cdots) = \left(\frac{3a}{8}\right)^2 + \left(\frac{a}{3}\right)^2 \frac{\delta}{1-\delta}$$

Step 5: SPNE

- The cooperating is better than deviating if

$$\frac{a^2}{8}\frac{1}{1-\delta} \geq \left(\frac{3a}{8}\right)^2 + \left(\frac{a}{3}\right)^2 \frac{\delta}{1-\delta}$$

This implies $\delta \geq 9/17$.

If $\delta \geq 9/17$, then the strategy:

- If the two firms have both used $\hat{q}$ in all previous periods, use $\hat{q}_j = a/4$ this period

- If either firm ever did anything besides $\hat{q}$, play the stage Cournot quantity $q_j^* = a/3$

is a SPNE of the infinitely repeated game?

Convex Hull

- A set is said to be convex if it contains the line segments connecting each pair of its points

- The convex hull of set $S = \{x_1, \dots, x_n\}$ is defined as

$$\mathrm{Conv}(S) = \left\{ \sum_i a_i x_i \mid a_i \in [0,1], \sum_i a_i = 1 \right\}$$

  - The (unique) minimal convex set containing *S*
  - The intersection of all convex sets containing *S*
  - The set of all convex combinations of points in *S*

Feasible Payoffs

- Consider stage game $G = \{N, \{A_i\}, \{u_i\}\}$ and infinitely repeated game $G^{\infty}(\delta)$.

- Let us introduce the set of feasible payoffs:

$$U = \text{conv}\left\{\begin{matrix} \text{u} \in R^N\text{:} & \text{there exists } a = (a_1, \dots, a_N) \\ & \text{s.t.} \quad u = (u_1(a), \dots, u_N(a)) \end{matrix}\right\}$$

That is, U is the convex hull of all N-dimensional vectors that can be obtained by some (possibly mixed) strategy outcome.

Minmax Payoffs

- Minmax payoff of player i: the lowest payoff that player i's opponent can hold him to:

$$\underline{u}_i = \min_{a_{-i}} \left[\max_{a_i} u_i(a_i, a_{-i}) \right]$$

- The player can never receive less than this amount.
- Minmax strategy outcome against to i

$$a_{-i}^{i} = \arg\min_{a_{-i}} \left[\max_{\mathrm{a}} u_i(a_i, a_{-i}) \right]$$

- Let a_i^i denote the strategy of player i such that

$$u_i\left(a_i^i, a_{-i}^i\right) = \underline{u}_i$$

Notice that a_i may be a mixed strategy for each player i

Example

		Player 2 L	Player 2 R
Player 1	U	-2, 2	1, -2
	M	1, -2	-2, 2
	D	0, 1	0, 1

How to find the minmax payoff for player 1

The payoffs of player 1 for different strategies are

‘U’ : $1 - 3q$

‘M’: $-2 + 3q$

‘D’: 0

Example (cont.)

We have

$$\underline{u}_1 = \min_{q\in[0,1]}[\max\{1-3q, -2+3q, 0\}]$$

Then, $\underline{u}_1 = 0$, and $a^1_{-1} = a^1_2$ is the mixed strategy with probability $q \in [\frac{1}{3}, \frac{2}{3}]$ over strategy ‘L’

For player 2, we have

$$\underline{u}_2 = \min_{\substack{q_1,q_2\in[0,1] \\ q_1+q_2\leq 1}}[\max\{1+q_1-3q_2, 1+q_2-3q_1\}]$$

Then, $\underline{u}_2 = 0$, and a^2_{-1} is the mixed strategy with probability $q_1 = q_2 = 1/2$ over strategy ‘U’ and ‘M’

Minmax Payoff Lower Bounds

Theorem

- Let $a' = (a'_1, a'_2, \dots, a'_n)$ be a (possibly mixed) Nash Equilibrium of game G and $u_i(a')$ be its payoff. Then

$$u_i(a') \geq \underline{u}_i$$

Proof. See board.

Nash Folk Theorem

Definition A payoff vector $(u_1, u_2, \dots, u_N) \in R^N$ is strictly individually rational if $u_i > \underline{u}_i$ for all i

Nash Folk Theorem If $(u_1, u_2, \dots, u_N) \in U$ is strictly individually rational, then there exists some $\delta_0 < 1$ such that for all $\delta \geq \delta_0$, there is Nash equilibrium of $G^\infty(\delta)$ with payoff $(u_1, u_2, \dots, u_N)$

Any strictly individually rational payoff can be obtained as a Nash Equilibrium when players are patient enough

Proof. See board

Problem with Nash Folk Theorem

- Nash Folk Theorem may be not a subgame perfect

		Player 2	
		L	R
Player 1	U	6, 6	0, -20
	D	7, 1	0, -20

- The unique NE in this game is (D,L).
- The minmax payoff are given by

$$\underline{u}_1 = 0 \quad \text{and} \quad \underline{u}_2 = 1$$

and $a^2_{-1} = R$

Problem with Nash Folk Theorem

		Player 2	
		L	R
Player 1	U	6, 6	0, -20
	D	7, 1	0, -20

- Nash Folk Theorem: the strategy
  - Play (U,L) as long as no one deviates
  - If Player 1 deviates, then player 2 select R
- While this will hurt player 1, it will hurt player 2 a lot.
- It is an threat, and it is not a SPNE

Subgame Perfect Folk Theorem

The first subgame perfect folk theorem shows that any payoff above the static Nash payoffs can be sustained as a subgame perfect equilibrium of the repeated game

Theorem (Friedman)

- Let a^{NE} be a static equilibrium of the stage game with payoffs u^{NE};
- For any feasible payoff u with $u_i > u_i^{\mathrm{NE}}$ for all i;

There exists some $\delta_0 < 1$ s.t. for all $\delta \geq \delta_0$, there is subgame perfect Nash equilibrium of $G^{\infty}(\delta)$ with payoff u.

Cooperation in Finitely-Repeated Games

- There are unique stage equilibrium in previous example
- There are multiple Nash equilibria in the stage game.

		Player 2		
		A	B	C
Player 1	A	3, 3	0, 4	-2, 0
	B	4, 0	1, 1	-2, 0
	C	0, -2	0, -2	-1, -1

The Nash equilibria are (B,B) and (C,C)

For cooperation, the best strategy is (A,A)

Cooperation in Finitely-Repeated Games

For $T = 2$, the strategy is

- Each plays (A, A) in the first period, and plays (B, B) in the second period
- If some player plays B in the first period, then the other plays C in the second period

If each player agrees the strategy, then the payoff is 4 for each player

If some one deviates, then the other will play C, and the payoff is 3

Deviation is not profitable

Repeated Games with Imperfect Information

- So far, we assume that in the repeated game, each player observes the actions of others at the end of each stage
  - I could observe if you stick or deviated from the agreement
- In several cases, player's actions may not be directly observable, e.g.,
  - Firm productions in Cartel
  - Antiballistic Missile treaty between the US and USSR in 1972 (ABM treaty).
  - Every country can imperfectly observe each other's compliance (despite spies, satellites, etc.)

ABMs treaty with Imperfect Information

We introduce ABMs treaty as follows:

Number of ABMs	Probability of detection ABMs
None	0
Low	10%
High	50%

- If a country has no ABMs, then the probability that satellite detects ABMs is zero
- If a country has a low level of ABMs, then the probability that my satellite detects ABMs is 10%
- If a country has a high level of ABMs, then the probability that my satellite detects ABMs is 50%

ABMs treaty with Imperfect Information

USSR

USA		No		Low		High	
	No	10	10	6	12	0	18
	Low	12	6	8	8	2	14
	High	18	0	14	2	3	3

- The unique NE is (High,High)
- (Low,Low) is more efficient, and (No,No) is the most efficient
- Can we cooperate playing (No,No) in the SPNE of the infinitely repeated game

ABMs treaty with Imperfect Information

Strategy

- In period $t = 1$, choose No ABMs (cooperation)
- For $t \geq 1$, the strategy is:
  - No ABMs if neither country has observed ABMs in other countries during the previous period, or
  - High ABMs if either country has observed ABMs in other countries during previous periods
- At any time t, if no country has detected ABMs, the payoff from sticking to the agreement is:

$$10 + 10\delta + \cdots + 10\delta^{t-1} + \cdots = \frac{10}{1-\delta}$$

ABMs treaty with imperfect monitoring

- In contract, the payoff from deviating to low ABMs during one period

$$12 + \delta\left(0.1 \times \frac{3}{1-\delta} + 0.9 \times \frac{10}{1-\delta}\right)$$

- The payoff from deviating to high ABMs during one period

$$18 + \delta\left(0.5 \times \frac{3}{1-\delta} + 0.5 \times \frac{10}{1-\delta}\right)$$

We need Coop ≥ Low and Coop ≥ High by

$$\delta \geq 0.74 \quad \text{and} \quad \delta \geq 0.7$$

Cournot Competition with Noisy Demand

- Firms set output levels $q_1^t, \ldots, q_N^t$ privately at time t
- The level of demand is stochastic
- Each firm's payoff depends on his own output and on the publicly observed market price
- Firms do not observe each other's outputs
- The market price depends on uncertain demand and the total outputs (low market price is due to high outputs and low demands)

Games with public information: At each period, all players observe a public outcome, is correlated with stage actions

Formulations

We formulize a game with actions and a public outcome.

- Let $A_1, \ldots, A_N$ be finite action sets
- Let y denote the publicly observed outcome (stochastic) from a (finite) set Y
- Let $\pi(y, a)$ denote the probability distribution of y under action outcome a, i.e., each outcome induces a probability distribution over $y \in$ Y

Formulations

- Player i payoff: $r_i(a_i, y)$, which depends on the actions of the others and y
- Player i's expected stage payoff is given by

$$u_i(a) = \sum_{y \in Y} \pi(y, a) r_i(a_i, y)$$

- The public information at the start of period t is

$$h_t = (y_0, \ldots, y_{t-1})$$

- We consider public strategies for player i, which is a sequence of maps

$$s_i^t : h_t \to A_i$$

Example: Noisy Prisoner's Dilemma

- Public signal: p
- Actions: (a_1, a_2), $a_i \in \{\mathrm{C}, \mathrm{D}\}$
- Payoffs:

$$r_1(C,p) = 1 + p, r_1(D,p) = 4 + p,$$
$$r_2(C,p) = 1 + p, r_2(D,p) = 4 + p.$$

- Probability distribution for public signal p:
  - $a_1 = a_2 = \mathrm{C} \;\; \rightarrow \; p = X$
  - $a_1 = a_2 \;\; \rightarrow \; p = X - 2$
  - $a_1 = a_2 = \mathrm{D} \;\; \rightarrow \; p = X - 4$

  where X is a continuous random variable with cumulative distribution function $F(x)$ and $E[X] = 0$

Trigger-Price Strategy

The payoff matrix is

	C	D
C	$(1+X, 1+X)$	$(-1+X, 2+X)$
D	$(2+X, -1+X)$	(X, X)

Consider the strategy

- I: Play (C, C) until $p \leq p^*$, then go to II.
- II: Play (D, D) for T periods, then go back to I.

- The strategy is symmetric, and the punishment phase uses a static NE.
- We next show that we can choose p^* and T such that the proposed strategy profile is an SPNE.

Trigger-Price Strategy

In phrase I, if players do not deviate, the expected payoff is

$$v = 1 + \delta\Big(F(p^*)\delta^T v + (1 - F(p^*)v)\Big)$$

$$v = \frac{1}{1 - F(p^*)\delta^T - \delta(1 - F(p^*))}$$

If the player deviates, the expected payoff is

$$v_d = 2 + \delta\Big(F(p^* + 2)\delta^T v + (1 - F(p^* + 2)v)\Big)$$

The deviating provides immediate payoff, but increases the probability of entering Phase II

Trigger-Price Strategy

- Incentive Compatibility Constraint: $v \geq v_d$, that is
$$v \geq 2 + \delta\left(F(p^* + 2)\delta^T v + (1 - F(p^* + 2)v)\right)$$
- Substituting v, we have
$$\frac{1}{1-F(p^*)\delta^T-\delta(1-F(p^*))} \geq \frac{2}{1-F(p^*+2)\delta^{T+1}-\delta(1-F(p^*+2))}$$
- Any T and p^* satisfying the constraint would construct an SPNE
- The best possible trigger-price equilibrium strategy could be found if we could maximize v subject to the incentive compatibility constraint

Summaries

- Repeated Game
- Minmax strategy
- Nash Folk Theorem (NE, Not SPNE)
- Folk Theorem (previous chapter)
- Repeated game with imperfect information