Repeated Game

Ichiro Obara

UCLA

March 1, 2012

Why Repeated Games?

- Prisoner's dilemma game:

	C	D
C	$1, 1$	$-1, 2$
D	$2, -1$	$0, 0$

- D is the strictly dominant action, hence (D, D) is the unique Nash equilibrium.
- However, people don't always play (D, D). Why? One reason would be that people (expect to) play this game repeatedly. Then what matters is the total payoff, not just current payoff.
- **Repeated game** is a model about such long-term relationships.

- A list of questions we are interested in:
    - When can people cooperate in a long-term relationship?
    - How do people cooperate?
    - What is the most efficient outcome that arise as an equilibrium?
    - What is the set of all outcomes that can be supported in equilibrium?
    - If there are many equilibria, which equilibrium would be selected?

Formal Model

Stage Game

- In repeated game, players play the same strategic game G repeatedly, which is called **stage game**.
- $G = (N, (A_i), (u_i))$ satisfies usual assumptions.
  - **Player:** $N = \{1, ..., n\}$
  - **Action:** $a_i \in A_i$ (finite or compact&convex in $\Re^K$).
  - **Payoff:** $u_i : A \to \Re$ (continuous).
- The set of **feasible payoffs** is $\mathcal{F} := co\,\{g(a) : a \in A\}$.

Now we define a repeated game based on G.

History

- A period t **history** $h_t = (a_1, ..., a_{t-1}) \in H_t = A^{t-1}$ is a sequence of the past action profiles at the beginning of period t.
- The initial history is $H_1 = \{\emptyset\}$ by convention.
- $H = \cup_{t=1}^{\infty} H_t$ is the set of all such histories.

Strategy and Payoff

- Player i's (pure) **strategy** $s_i \in S_i$ is a mapping from H to A_i.
    - Ex. Tit-for-Tat: "First play C, then play what your opponent played in the last period".
- A strategy profile $s \in S$ generates a sequence of action profiles $(a_1, a_2, ...) \in A^\infty$. Player i's **discounted average payoff** given s is

$$V_i(s) := (1-\delta)\sum_{t=1}^{\infty} \delta^{t-1} g_i(a_t)$$

where $\delta \in [0,1)$ is a discount factor.

Repeated Game

- This extensive game with simultaneous moves is called **repeated game** (sometimes called **supergame**).
- The repeated game derived from G and with discount factor δ is denoted by $G^{\infty}(\delta)$
- We use subgame perfect equilibrium.
- The set of all pure strategy SPE payoff profiles for $G^{\infty}(\delta)$ is denoted by $\mathcal{E}[\delta]$.

Public Randomization Device

- We may allow players to use a publicly observable random variable (say, throwing a die) in the beginning of each period.
- Formally we can incorporate a sequence of outcomes of such **public randomization device** as a part of history in an obvious way. To keep notations simple, we don't introduce additional notations for public randomization.

Minmax Payoff

- Let $\underline{v}_i$ be player i's **pure-action minmax payoff** defined as follows.

Pure-Action Minmax Payoff

$\underline{v}_i = \min_{a_{-i} \in A_{-i}} \max_{a_i \in A_i} g_i(a)$.

- Intuitively $\underline{v}_i$ is the payoff that player i can secure when player i knows the other players' actions.
- Ex. $\underline{v}_i = 0$ for $i = 1, 2$ in the previous PD.

Minmax Payoff

Minmax payoff serves as a lower bound on equilibrium payoffs in repeated games.

Lemma

Player i's payoff in any NE for $G^\infty(\delta)$ is at least as large as $\underline{v}_i$.

Proof

Since player i knows the other players' strategies, player i can deviate and play a "myopic best response" in every period. Then player i's stage game payoff would be at least as large as v_i in every period. Hence player i's discounted average payoff in equilibrium must be at least as large as v_i.

Trigger Strategy

- Consider the following PD ($g, \ell > 0$).

	C	D
C	$1, 1$	$-\ell, 1+g$
D	$1+g, -\ell$	$0, 0$

- When can (C, C) be played in every period in equilibrium?

Such an equilibrium exists if and only if the players are enough patient.

Theorem

There exists a subgame perfect equilibrium in which (C, C) is played in every period if and only if $\delta \geq \frac{g}{1+g}$.

Proof.

- Consider the following **trigger strategy**:
  - Play C in the first period and after any cooperative history $(C, C), ..., (C, C)$.
  - Otherwise play D.
- This is a SPE if the following one-shot deviation constraint is satisfied

$$1 \geq (1 - \delta)(1 + g)$$

, which is equivalent to $\delta \geq \frac{g}{1+g}$.
- By our previous observation, each player's continuation payoff cannot be lower than 0 after a deviation to D. Hence $\delta \geq \frac{g}{1+g}$ is also necessary for supporting (C, C) in every period.

Stick and Carrot

- Consider the following modified PD.

	C	D	E
C	$1,1$	$-1,2$	$-4,-4$
D	$2,-1$	$0,0$	$-4,-4$
E	$-4,-4$	$-4,-4$	$-5,-5$

- The standard trigger strategy supports (C, C) in every period if and only if $\delta \geq 1/2$.

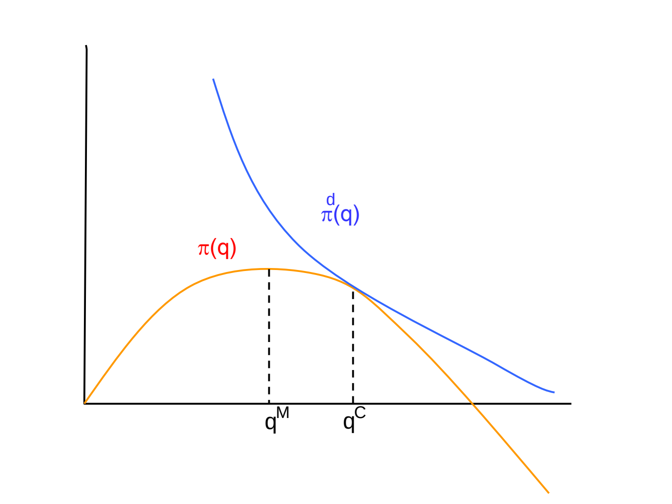
- The following strategy ("Stick and Carrot" strategy) support cooperation for even lower δ.
  - Cooperative Phase: Play C. Stay in Cooperative Phase if there is no deviation. Otherwise move to Punishment Phase.
  - Punishment Phase: Play E. Move back to Cooperative Phase if there is no deviation. Otherwise stay at Punishment Phase.
- There are two one-shot deviation constraints.
  - $1 \geq (1-\delta)2 + \delta\left[(1-\delta)(-5)+\delta\right]$
  - $(1-\delta)(-5)+\delta \geq (1-\delta)(-4) + \delta\left[(1-\delta)(-5)+\delta\right]$.
- They are satisfied if and only if $\delta \geq \frac{1}{6}$.

Optimal Collusion

- We study this type of equilibrium in the context of dynamic Cournot duopoly model.
- Consider a repeated game where the stage game is given by the following Cournot duopoly game.
  - $A_i = \Re_+$
  - Inverse demand: $p(q) = \max\{A - (q_1 + q_2), 0\}$
  - $\pi_i(q) = p(q)q_i - cq_i$
- Discount factor $\delta \in (0, 1)$.

Cournot-Nash equilibrium and Monopoly

- In Cournot-Nash equilibrium, each firm produces $q^C = \frac{A-c}{3}$ and gains $\pi^C = \frac{(A-c)^2}{9}$.
- The total output that would maximize the joint profit is $\frac{A-c}{2}$. Let $q^M = \frac{A-c}{4}$ be the monopoly production level per firm.
- Let $\pi(q) = \pi_i(q, q)$ be each firm's profit when both firms produce q.
- Let $\pi^d(q) = \max_{q_i \in \Re_+} \pi_i(q_i, q)$ be the maximum profit each firm can gain by deviating from q when the other firm produces q.



π(q)
d
π(q)
q^M
q^C

- We look for a SPE to maximize the joint profit.
- The firms like to collude to produce less than the Cournot-Nash equilibrium to keep the price high.
- We focus on **strongly symmetric SPE**. When the stage game is symmetric, an SPE is strongly symmetric if every player plays the same action after any history.

Structure of Optimal Equilibrium

- We show that the best SSSPE and the worst SSSPE has a very simple structure.
- Consider the following strategy:
  - **Phase 1:** Play q^*. Stay in Phase 1 if there is no deviation. Otherwise move to Phase 2.
  - **Phase 2:** Play q_*. Move to Phase 1 if there is no deviation. Otherwise stay in Phase 2.
- The best SSSPE is achieved by a strategy that starts in Phase 1 (denoted by $s(q^{*\infty})$) and the worst SSSPE is achieved by a strategy that starts in Phase 2 (denoted by $s(q_*, q^{*\infty})$) for some q^*, q_*.

- Let $\overline{V}$ and $\underline{V}$ be the best SSSPE payoff and the worst SSSPE payoff respectively (**Note:** this needs to be proved).
- First note that the equilibrium action must be constant for $\overline{V}$.
  - Let q^* be the infimum of the set of all actions above q^M that can be supported by some SSSPE. Let $q^k, k = 1, 2, ..$ be a sequence within this set that converges to q^*.
  - One-shot deviation constraint implies

$$(1-\delta)\pi(q^k) + \delta\overline{V} \geq (1-\delta)\pi^d(q^k) + \delta\underline{V}$$

Taking the limit and using $\pi(q^*) \geq \overline{V}$, we have

$$\pi(q^*) \geq (1-\delta)\pi^d(q^*) + \delta\underline{V},$$

which means that it is possible to support q^* in every period.

- Secondly, we can show that the worst SSSPE can be achieved by $s(q_*, q^{*\infty})$ ("**stick and carrot**") for some $q_* \geq q^C$.
  - Take any path $Q' = (q'_1, q'_2, ...,)$ to archive the worst SSSPE.
  - Since $\pi(q'_t) \leq \pi(q^*)$ for all t and $\pi(q)$ is not bounded below, we can find $q_* \geq q'_1$ such that $Q_* = (q_*, q^*, ...)$ generates the same discounted average payoff as Q'.
  - Then $c(q_*, q^{*\infty})$ is a SPE that archives the worst SSSPE payoff because

$$V(Q') \geq (1-\delta)\pi^d(q'_1) + \delta V(Q')$$

$$\Downarrow$$

$$V(Q_*) \geq (1-\delta)\pi^d(q_*) + \delta V(Q_*).$$

To summarize, we have the following theorem.

Theorem (Abreu 1986)

There exists $q^* \in [q^M, q^C]$ and $q_* \geq q^C$ such that $s(q^{*\infty})$ achieves the best SSSPE and $s(q_*, q^{*\infty})$ achieves the worst SSSPE.

- **Note:** This can be generalized to the case with nonlinear demand function and many firms.

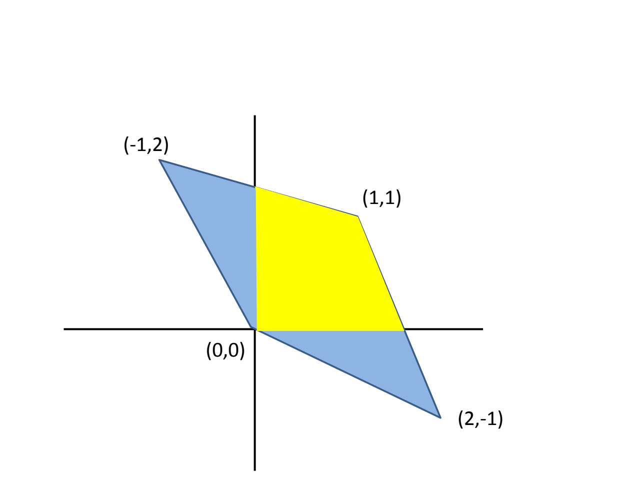
- **Q:** How many SPE? Which payoff can be supported by SPE?
- **A:** Almost all "reasonable" payoffs if δ is large.

- What does "almost all" mean?
- We know that player i's (pure strategy) SPE payoff is never strictly below $\underline{v}_i$. We show that every feasible v strictly above $\underline{v}$ can be supported by SPE. This is so called **Folk Theorem** in the theory of repeated games.

Definitions

- $v \in \mathcal{F}$ is **strictly individually rational** if v_i is strictly larger than $\underline{v}_i$ for all $i \in I$. Let $\mathcal{F}^* \subset \mathcal{F}$ be the set of feasible and strictly individually rational payoff profiles.
- Normalize $\underline{v}_i$ to 0 for every i without loss of generality.
- Let $\overline{g} := \max_i \max_{a,a' \in A} \left| g_i(a) - g_i(a') \right|$.

For the repeated PD, the yellow area is the set of strictly individually rational and feasible payoffs.



Folk Theorem

There are many folk theorems. This is one of the most famous ones.

Theorem (Fudenberg and Maskin 1986)

Suppose that $\mathcal{F}^*$ is full-dimensional (has an interior point in $\Re^N$). For any $v^* \in \mathcal{F}^*$, there exists a strategy profile $s^* \in S$ and $\underline{\delta} \in (0,1)$ such that s^* is a SPE and achieves v for any $\delta \in (\underline{\delta}, 1)$.

Folk Theorem

Idea of Proof

- Players play $a^* \in A$ such that $v^* = g(a^*)$ every period in equilibrium.
- Any player who deviates unilaterally is punished by being minmaxed for a finite number of periods.
- The only complication is that minmaxing someone may be very costly, even worse than being minmaxed.
- In order to keep incentive of the players to punish the deviator, every player other than the deviator is "rewarded" after minmaxing the deviator.

Proof

- **Step 1.** Pick $v^j \in \mathcal{F}^*$ for each $j \in N$ so that $v_j^* > v_j^j$ for all j and $v_j^i > v_j^j$ for all $i \neq j$. We assume that there exists $a^*, a^j \in A, j = 1, ..., N$ such that $v^* = g(a^*)$ and $v^j = g(a^j)$ for simplicity (use a public randomization device otherwise).
- **Step 2.** Pick an integer T to satisfy $\overline{g} < T \min_{i,j} v_i^j$.

Proof

- **Step 3.** Define the following strategy.
  - Phase I. Play $a^* \in A$. Stay in Phase I if there is no unilateral deviation from a^*. Go to Phase II(i) if player i unilaterally deviates from a^*.
  - Phase II(i). Play $\underline{a}(i) \in A$ (the action minmaxing player i) for T periods and go to Phase III(i) if there is no unilateral deviation. Go to Phase II(j) if player j unilaterally deviates from $\underline{a}(i)$.
  - Phase III(i). Play $a^i \in A$. Stay in Phase III(i) if there is no unilateral deviation from a^i. Go to Phase II(j) if player j unilaterally deviates from a^i.

Proof

- **Step 4.** Check all one shot deviation constraints.
  - **Phase I**

$$(1-\delta)\,\overline{g} \leq (1-\delta)\,(\delta+,...,+\delta^{T})v_j^{*} + \delta^{T+1}\left(v_j^{*} - v_j^{j}\right) \text{ for all } j \in N$$

  - **Phase II(i)**(the first period): IC is clearly satisfied for i. For $j \neq i$,

$$\left(1-\delta^{T+1}\right)\overline{g} \leq \delta^{T+1}\left(v_j^{i} - v_j^{j}\right)$$

  - **Phase III(i)**

$$(1-\delta)\,\overline{g} \leq (1-\delta)\,(\delta+,...,+\delta^{T})v_j^{i} + \delta^{T+1}\left(v_j^{i} - v_j^{j}\right) \text{ for all } j \in N$$

These constrains are satisfied for some large enough $\underline{\delta}$ and any $\delta \in (\underline{\delta}, 1)$.

References

- Abreu, "On the theory of infinitely repeated games with discounting," *Econometrica* 1988.
- Fudenberg and Maskin, "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information," *Econometrica* 1986.
- Mailath and Samuelson, *Repeated Games and Reputations*, Oxford Press 2006.