Introduction to Game Theory
Lecture Note 5: Repeated Games

Haifeng Huang

University of California, Merced

Spring 2020

Repeated games

- **Repeated games**: given a simultaneous-move game G, a repeated game of G is an extensive game with perfect information and simultaneous moves in which a history is a sequence of action profiles in G.
- Denote the repeated game, if repeated T times, as G^T. G is often called a **stage game**, and G^T a **supergame**.
- Why repeated games matter? E.g., in Prisoner's Dilemma, if players interact repeatedly, maybe they can be induced to cooperate rather than defect, by the threat of punishment, if they defect, by other players in later rounds of the interactions.

Discounting

- Discounting: people value \$100 tomorrow less than \$100 today. (Why? You can earn interest, for example.)
- The **discount factor** δ $(0 \leq \delta \leq 1)$ denotes how much a future payoff is valued at the current period, or how patient a player is.
  - If a player has a δ of 0.8, then \$100 tomorrow is equivalent to \$80 today for her.
- Relationship between discount factor (δ) and discount rate/interest rate (r): for a fixed discount rate r, discretely compounded over T periods, the discount factor $\delta = \frac{1}{(1+r)^T}$; if the discount rate is continuously compounded, $\delta = e^{-rT}$ $((\frac{1}{1+\frac{r}{n}})^{nT} = (1+\frac{r}{n})^{-(\frac{n}{r})rT} = e^{-rT}$ when $n \to \infty$).

Payoffs from a repeated game

- A player gets a payoff from each stage game, so her total payoff from the supergame is the discounted sum of the payoffs from each stage game.
- Let's call a sequence (with T periods) of action profiles as (a_1, a_2,...a_T), then a player i's total payoff from this sequence, when her discount factor is δ, is

$$u_i(a_1) + \delta u_i(a_2) + \delta^2 u_i(a_3) + ... + \delta^{T-1} u(a_T) = \sum_{t=1}^{T} \delta^{t-1} u_i(a_t).$$

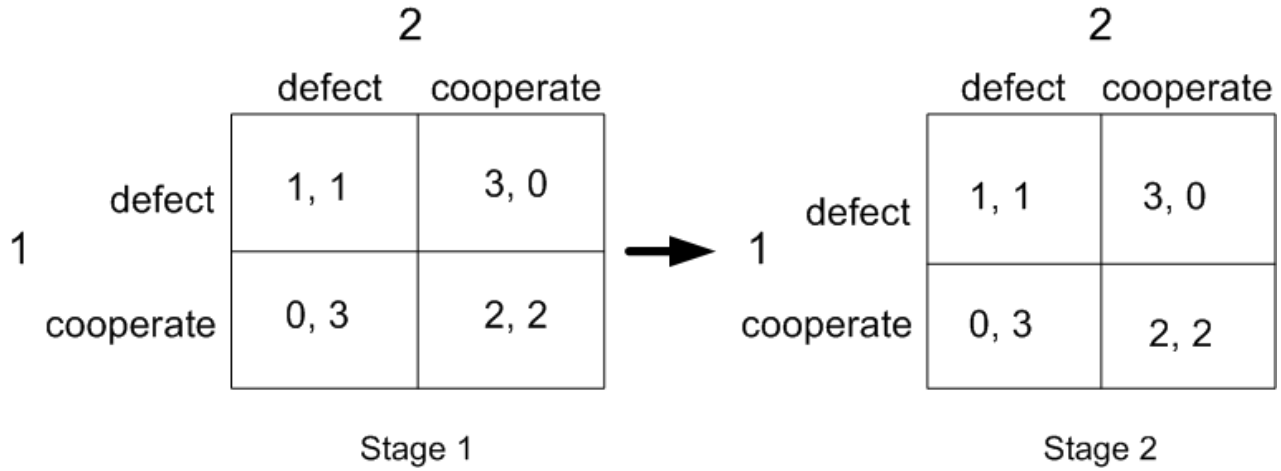
- If the sequence is infinite, then the discounted sum is $\sum_{t=1}^{\infty} \delta^{t-1} u_i(a_t)$.

Geometric series

- The discounted sum of a player's payoffs from a repeated game is a geometric series.
- **Geometric sequence**: b, bk, bk^2,...,bk^T; **geometric series**: $b + bk + bk^2 + ... + bk^T = \sum_{t=1}^{T+1} bk^{t-1}$.
- If T is finite, $S_T = b + bk + bk^2 + ... + bk^T = \frac{b(1-k^{T+1})}{1-k}$.
- If T is infinite, $b + bk + bk^2 + ... + bk^\infty = \frac{b}{1-k}$, if $-1 < k < 1$.

PD finitely repeated

- If a Prisoner's Dilemma game is repeated twice, what is the SPNE? Does the discount factor matter here?



- What if the PD game is repeated T times, $T < \infty$?

Unraveling in finitely repeated games

- Proposition (**unraveling**): Suppose the simultaneous-move game G has a unique Nash equilibrium, σ^*. If $T < \infty$, then the repeated game G^T has a unique SPNE, in which each player plays her strategy in σ^* in each of the stage games.

- If there are more than one Nash equilibria in the stage game, however, players' SPNE strategies in any given period (except the last period) need not coincide with any one-shot equilibrium.

Finitely repeated PD with punishment

- How many Nash equilibria in the following stage game (a variant of PD with punishment)?

		Player 2		
		Cooperate	Defect	Punish
Player 1	Cooperate	$-1, -1$	$-10, 0$	$-15, -10$
	Defect	$0, -10$	$-9, -9$	$-15, -10$
	Punish	$-10, -15$	$-10, -15$	$-12, -12$

- (Cooperate, Cooperate) is not a NE in the stage game, but can be supported as part of a SPNE if the stage game is played twice and the players are patient enough.

SPNE in finitely repeated PD with punishment (1)

Let the discount factor $\delta \geq \frac{1}{3}$. For $i = 1, 2$, let s_{i0} be player i's strategy after history $\emptyset$, s_{i1} be player i's strategy after the first stage game, and a_1 be the action profile in the first stage game. The following strategy profile that induces the players to cooperate in the first stage is a SPNE:

$$s_{i0} = C, \text{ and } s_{i1}(a_1) = \begin{cases} D, & \text{if } a_1 = (C, C); \\ P, & \text{otherwise.} \end{cases}$$

SPNE in finitely repeated PD with punishment (2)

- The stage game again:

		Player 2		
		Cooperate	Defect	Punish
Player 1	Cooperate	$-1, -1$	$-10, 0$	$-15, -10$
	Defect	$0, -10$	$-9, -9$	$-15, -10$
	Punish	$-10, -15$	$-10, -15$	$-12, -12$

- Proof: First, the strategies induce a NE in the 2nd stage game regardless of the history. Now, if the other player is sticking to the above strategy, player i's discounted payoff from adhering to the strategy is $-1 - 9\delta$. Her best deviation, given the other player sticks to the above strategy, is to play D in the 1st round and P in the 2nd round, yielding a discounted payoff of $0 - 12\delta$. Therefore the proposed strategies constitute a SPNE if $-1 - 9\delta \geq 0 - 12\delta$, or if $\delta \geq \frac{1}{3}$.

PD infinitely repeated

- An infinitely repeated game usually has more SPNEs than a finitely repeated game, and it may have multiple SPNEs even if the stage game has a unique NE.
- But playing the NE strategies in each stage game, regardless of history, is still a SPNE in the infinitely repeated game.
  - So each player choosing defection is a SPNE in infinitely repeated PD as in finitely repeated PD.
  - But there are other SPNEs in infinitely repeated PD.

One-deviation property

- **One-deviation property**: no player can increase her payoff by changing her action at the start of any subgame in which she is the first-mover, given the other players' strategies *and* the rest of her own strategy.
- One-deviation property of SPNE of finite horizon games: A strategy profile in an extensive game with perfect information and a finite horizon is a SPNE if and only if it satisfies the one-deviation property.
- A strategy profile in an infinitely repeated game with a discount factor less than 1 is a SPNE if and only if it satisfies the one-deviation property.

Grim trigger strategies

- Consider an infinitely repeated standard PD game, where (D, D) is the unique stage game NE.

		Player 2: Defect	Player 2: Cooperate
Player 1	Defect	1, 1	3, 0
	Cooperate	0, 3	2, 2

- Let s_{it} be i's strategy after period t (history x_t). The following strategy profile will be a SPNE if $\delta \geq \frac{1}{2}$. For $i = 1, 2$,

$$s_{i0} = C, \text{ and } s_{it}(x_t) = \begin{cases} C, & \text{if } x_t = ((C,C),(C,C),...,(C,C)); \\ D, & \text{otherwise.} \end{cases}$$

- You start by playing C, but if any player (including yourself) ever deviates to D, then both players play D forever.

Grim trigger strategies: proof of SPNE (1)

- When both players follow the grim trigger strategy, the outcome path of a subgame will either be (C, C) in every period or (D, D) in every period.
- When the outcome is (C, C) in every period,
  - A player's payoff (in that subgame) from following the strategy is $2 + 2\delta + 2\delta^2 + ... = \frac{2}{1-\delta}$.
  - If she deviates in one period, she gets 3 in that period, but 1 in every period thereafter (since both players will then play D), so her payoff is $3 + \delta\frac{1}{1-\delta}$.
  - She has no incentive to make that one deviation if $\frac{2}{1-\delta} \geq 3 + \delta\frac{1}{1-\delta}$, or $\delta \geq \frac{1}{2}$.

Grim trigger strategies: proof of SPNE (2)

- When the outcome is (D, D) in every period, clearly no player wants to deviate to C in any one period (getting 0 rather than 1), while what follows after that period is still (D, D) forever.
- By the one-deviation property, the proposed grim trigger strategies constitute a SPNE for $\delta \geq 1/2$.
- Cooperation can also be induced by punishment for a finite number of periods rather than punishment forever, as long as δ sufficiently high. But still it has to be the case that D is triggered for both players whenever one player deviates.
  - Triggering D for a player only if the other player deviates will not sustain cooperation in SPNE.

Tit-for-tat: another SPNE in infinitely repeated PD

- Tit-for-tat: a player starts by playing C, and then do whatever the other player did in the previous period.
- When players follow tit-for-tat, behavior in a subgame hinges on the last outcome in the history that preceded the subgame: (C, C), (C, D), (D, C), or (D, D). Suppose player 2 sticks to t-f-t, and consider player 1's choices.

Tit-for-tat (2)

- After (C, C): if player 1 sticks to t-f-t, she gets $2/(1-\delta)$; if she deviates in and only in the period following (C, C) (choosing D in that period), the subgame outcome path will be (D, C), (C, D), (D, C), (C, D)..., and her payoff

$$\begin{aligned} &= 3 + 0 + 3\delta^2 + 0 + 3\delta^4 + ... \\ &= 3 + 3(\delta^2) + 3(\delta^2)^2 + 3(\delta^2)^3 + ... \\ &= 3/(1-\delta^2). \end{aligned}$$

- She has no incentive to deviate if $2/(1-\delta) \geq 3/(1-\delta^2)$, or $\delta \geq 1/2$.

Tit-for-tat (3)

- After (C, D): if player 1 sticks to t-f-t, the subgame outcome path will again be $(D, C), (C, D), (D, C), (C, D)...$, and her payoff is $3/(1-\delta^2)$. If she deviates in and only in the period following (C, D) (choosing C), the outcome is (C, C) in every subsequent period. So her payoff is $2/(1-\delta)$. For her not to deviate, we must have $3/(1-\delta^2) \geq 2/(1-\delta)$, or $\delta \leq 1/2$.
- After (D, C): if player 1 sticks to t-f-t, the subgame outcome path will again be $(C, D), (D, C), (C, D), (D, C)...$, and her payoff is $3\delta/(1-\delta^2)$. If she deviates in and only in the period following (D, C) (choosing D), the outcome is (D, D) in every subsequent period. So her payoff is $1/(1-\delta)$. For her not to deviate, we must have $3\delta/(1-\delta^2) \geq 1/(1-\delta)$, or $\delta \geq 1/2$.

Tit-for-tat (4)

- Finally, after a history ending in (D, D): if player 1 sticks to t-f-t, the subgame outcome path will again be (D, D) in every subsequent period, and her payoff is $1/(1-\delta)$. If she deviates in and only in the period following the history ending in (D, D) (choosing C), the outcome path in subsequent periods will be $(C, D), (D, C), (C, D), (D, C)$.... So her payoff is $3\delta/(1-\delta^2)$. For her not to deviate, we must have $\delta \leq 1/2$.
- By the one-deviation property, the strategy profile (tit-for-tat, tit-for-tat) constitutes a SPNE if and only if $\delta = 1/2$ in this PD.

The (Perfect) Folk Theorem

- So a cooperative outcome (C, C) can be supported in an infinitely repeated PD game by grim trigger strategies if the discount factor is sufficiently high. More generally,
- **Proposition**: Suppose that σ^* is a Nash equilibrium of the simultaneous-move game G, and consider the supergame $G^\infty(\delta)$ with $\delta < 1$. Let σ be any one-shot strategy profile that is strictly preferred to σ^* by all players: $u_i(\sigma) > u_i(\sigma^*)$, $\forall i$. There exists a $\underline{\delta}$ such that if $\delta \in [\underline{\delta}, 1)$, the following strategies, $\forall i = 1, 2, ..., N$, constitute a SPNE of $G^\infty(\delta)$.

$$\sigma_{i0} = \sigma_i, \text{ and } \sigma_{it}(x_t) = \begin{cases} \sigma_i, & \text{if } x_t = (\sigma, \sigma, ..., \sigma) \\ \sigma_i^*, & \text{otherwise,} \end{cases}$$

where σ_{it} refers to player i's strategy after history x_t.

Proof of the (Perfect) Folk Theorem

- As in the grim trigger strategy SPNE for PD, the above strategies constitute a SPNE following a deviation by any player.
- To show that no one can do better than unilaterally deviating from σ_i, we just need to show no one can benefit from a one-shot deviation σ_i':

$$u_i(\sigma_i', \sigma_{-i}) + \frac{\delta u_i(\sigma^*)}{1-\delta} \leq \frac{u_i(\sigma)}{1-\delta}.$$

  the inequality is satisfied when

$$\delta \geq \frac{u_i(\sigma_i', \sigma_{-i}) - u_i(\sigma)}{u_i(\sigma_i', \sigma_{-i}) - u_i(\sigma^*)}.$$

- If σ is a mixed strategy, the randomization of each player needs to be verifiable ex post.

Significance of the Folk Theorem

- Any strategy profile that is Pareto superior to some Nash equilibrium of the stage game can be supported as infinite play on the equilibrium path of a SPNE if the players are patient enough.
- There are other versions of the Folk Theorem, in which any feasible payoff profile of the stage game that exceeds each player's min-max payoff can be sustained in a SPNE of the infinitely repeated game, as long as the discount factor is high enough.