Game-theoretic
Foundations of Multi-agent Systems

Lecture 6: Repeated Games

Seyed Majid Zahedi

UNIVERSITY OF WATERLOO

Outline

Repeated Games

- In a (typical) repeated game:
  - Agents play a given game (aka. stage game)
  - Then, they get their utilities
  - And, they play again ...
- Can be repeated finitely or infinitely many times
- Really, an extensive form game
  - Would like to find subgame-perfect equilibria

Repeated Games (cont.)

- One subgame-perfect equilibrium:
  - Keep repeating some Nash equilibrium of the stage game
  - Memoryless strategy, called a stationary strategy
- But are there other equilibria?
  - Strategy space of repeated game is much richer than that of stage game

Key Questions

- Do agents see what the other agents played earlier?
- Do they remember what they knew?
- Given utility of each stage game, what is the utility of the entire repeated game?

Finitely Repeated Games (with Perfect Monitoring)

- Agents play stage game G for R rounds
- At each round, outcomes of all past rounds are observed by all agents
- Agents' overall utility is sum of discounted utilities at each round
  - Discount factor is $0 \leq \delta \leq 1$
  - Game is denoted by $G^R(\delta)$
- Given sequence of utilities $u_i^{(1)}, \ldots, u_i^{(R)}$, $u_i = \sum_{r=1}^{R} \delta^{r-1} u_i^{(r)}$

Example: Finitely Repeated Prisoner's Dilemma

- Two agents play Prisoner's Dilemma for R rounds ($\delta = 1$)

	D	C
D	$-2, -2$	$-4, -1$
C	$-1, -4$	$-3, -3$

- Starting from last round, (C, C) is dominant strategy
- Hence, in second-to-last round, there is no way to influence what will happen
- So, (C, C) is dominant strategy at this round as well
- The unique SPE is (C, C) at each round

SPE in Finitely Repeated Games

[Theorem]

- If stage game G has unique strategy equilibrium s^*, then $G^R(\delta)$ has unique SPE in which $s^{(r)} = s^*$ for all $r = 1, ..., R$, regardless of history

[Proof]

- By backward induction, at round R, we have $s^{(R)} = s^*$
- Given this, then we have $s^{(R-1)} = s^*$, and continuing inductively, $s^{(r)} = s^*$ for all $r = 1, ..., R$, regardless of history

SPE: Example I

- Two agents play the following game for 2 rounds ($\delta = 1$)

	D1	D2	C
D1	$4, 4$	$1, 1$	$6, 0$
D2	$1, 1$	$2, 2$	$6, 0$
C	$0, 6$	$0, 6$	$5, 5$

- Consider the following strategy:
  - In round 1, cooperate;
  - In round 2, if someone defected in round 1, play D2; otherwise, play D1
- If both agents play this, is that SPE?

SPE: Example II

- Two agents play the following game for 2 rounds ($\delta = 1$)

	D	Crazy	C
D	$4, 4$	$1, 0$	$6, 0$
Crazy	$0, 1$	$0, 0$	$0, 1$
C	$0, 6$	$1, 0$	$5, 5$

- What are the subgame perfect equilibria?
- Consider the following strategy:
  - In round 1, cooperate;
  - In round 2, if someone played D or Crazy in round 1, play Crazy; otherwise, play D
- If both agents play this, is that NE (not SPE)?

TSPE: Example III

- If G has multiple equilibria, then $G^R(\delta)$ does not have unique SPE
- Consider following example

	x	y	z
x	$3, 3$	$0, 4$	$-2, 0$
y	$4, 0$	$1, 1$	$-2, 0$
z	$0, -2$	$0, -2$	$-1, -1$

- Stage game has two pure NE: (y, y) and (z, z)
- Consider the following policy:
  - Play x in first round
  - Play y in second round if opponent played x; otherwise, play z
- Is both agents playing this SPE?

Outline

1. Finitely Repeated Games
2. Infinitely Repeated Games
3. Folk Theorem
4. Repeated Games with Imperfect Monitoring

Utilities in Infinitely Repeated Games

- Average utility:

$$u_i = \lim_{R \to \infty} \frac{\sum_{r=1}^{R} u_i^{(r)}}{R}$$

- Discounted utility:

$$u_i = (1 - \delta) \sum_{r=1}^{\infty} \delta^{r-1} u_i^{(r)},$$

for some $0 \geq \delta < 1$

Subgame Perfection in Infinitely Repeated Games

- One-shot deviation from strategy s means deviating from s in single stage and conforming to it thereafter
- Strategy profile s^* is SPE if and only if there are no profitable one-shot deviation for each subgame and every agent
- This follows from principle of optimality of dynamic programming
- This applies to finitely repeated games as well

Trigger Strategies (TS)

- Agents get punished if they deviate from agreed profile
- In non-forgiving TS (or grim TS), punishment continues forever

$$s_i^{(t)} = \begin{cases} s_i^* & \text{if } s^{(r)} = s^* \;\; \forall r < t, \\ \underline{s}_i^j & \text{otherwise} \end{cases}$$

- Here, s^* is agreed profile, and $\underline{s}_i^j$ is punishment strategy of i against agent j
- Single deviation by j triggers agent i to switch to $\underline{s}_i^j$ forever

Example: Infinitely Repeated Prisoner's Dilemma

- Consider trigger strategy:
  - Cooperate as long as everyone cooperates
  - Once a player defects, defect forever
- Is both agents playing this SPE?
- Does it depend on δ?

	D	C
D	$-2, -2$	$-4, -1$
C	$-1, -4$	$-3, -3$

Trigger Strategy for Infinitely Repeated Prisoners' Dilemma

- We can use one-stage deviation principle
- There are two types of subgames:
  - Type 1: Both agents cooperated so far
  - Type 2: At least one agent defeated in the past
- Type-1 subgames: (D is best response to D)
  - Utility from no deviation: $(1-\delta)(-2-2\delta-2\delta^2+\dots)=-2$
  - Utility from on-shot deviation: $(1-\delta)(-1+(-3\delta-3\delta^2+\dots)=-(1-\delta)-3\delta$
  - Deviation is not beneficial if $\delta \geq 1/2$
- Type-2 subgames: (C is best response to C)
  - Other agents will always play C, thus C is best response

Tit-for-tat Strategy

- Consider tit-for-tat strategy:
    - Cooperate in 1st round
    - Then, do whatever other agent did in previous round
- Is both agents playing this NE?
- Is both agents playing this SPE?
- What about one playing TFT and other trigger?

Remarks

- If s^* is NE of G, then "each agent plays s_i^*" is SPE of $G^R(\delta)$
  - Future play of other agents is independent of how each agent plays
  - Optimal play is to maximize current utility, i.e., play static best response
- Sets of equilibria for finite and infinite horizon versions can be different
  - Multiplicity of equilibria in repeated prisoner's dilemma only occurs at $R = \infty$
  - For any finite R (thus for $R \to \infty$), repeated prisoners' dilemma has unique SPE

Repetition Could Lead to Bad Outcomes

- Consider the following game

	x	y	z
x	$2, 2$	$2, 1$	$0, 0$
y	$1, 2$	$1, 1$	$-1, 0$
z	$0, 0$	$0, -1$	$-1, -1$

- Strategy x strictly dominates y and z for both agents
- Unique NE of stage game is (x, x)
- If $\delta \geq 1/2$, this game has SPE in which (y, y) is played in every round
- It is supported by slightly more complicated strategy than grim trigger
  - I. Play y in every round unless someone deviates, then go to II
  - II. Play z. If no one deviates go to I. If someone deviates stay in II

Outline

Characterizing NE of Infinitely Repeated Games

- Characterizing all equilibrium strategy profiles might be challenging
- Instead, we can characterize utilities obtained in them
- Such utilities must be feasible
  - There must be outcomes of game such that agents, on average, get these utilities
- They must also be enforceable
  - Deviation should lead to punishment that outweighs benefits of deviation
- Folk theorem states that utility vector can be realized by some NE iff it is both feasible and enforceable

Feasibility

	Left	Right
Left	$2, 2$	$0, 3$
Right	$3, 0$	$1, 1$

- Utility vector (2, 2) is feasible as it is one of outcomes of game
- Utility vector (1, 2.5) is feasible as agents can alternate between (2, 2) and (0, 3)
- What about (0.5, 2.75)?
- What about (3, 0.1)?
- In general, convex hall of outcomes of game are feasible
  - $p_1 x_1 + \cdots + p_n x_n$ is convex hall of x_i if p_i sum to 1 and are non negative

Feasible and Individually Rational Utilities

- Feasible utilities:

$$V = \mathrm{Conv}\{v \in \mathbb{R}^{|N|} \mid \text{ there exists } a \in A \text{ such that } u(a) = v\}$$

- Note that $V \neq \{v \in \mathbb{R}^{|N|} \mid \text{ there exists } s \in S \text{ such that } u(s) = v\}$

- Recall minmax value of agent i:

$$\underline{v}_i = \min_{s_{-i}} \max_{s_i} u_i(s_i, s_{-i})$$

- Utility vector $v \in \mathbb{R}^{|N|}$ is strictly individually rational if $v_i > \underline{v}_i$ for all i

Nash Folk Theorem

- Consider infinitely repeated game G played by agents with average utilities
- If u is utility profile for any NE of repeated G, then u_i is enforceable for all i
- If u is both feasible and enforceable, then u is utility profile for some NE of G
- Folk theorem can be stated for agents with discounted utilities as well

Problems with Nash Folk Theorem

- Any feasible and enforceable utility can be achieved (for patient enough agents)
- Enforcement is often done by grim trigger strategy
  - Play certain strategy as long as no one deviates
  - If some agent j deviates, then play minmax strategy against that agent thereafter
- NE involves non-forgiving TS which may be costly for punishers
- NE may include non-credible threats
- NE may not be subgame perfect

Example

	L	R
U	$6, 6$	$0, -100$
D	$7, 1$	$0, -100$

- Unique NE in this game is (D, L)
- Minmax values are given by $\underline{v}_1 = 0$ and $\underline{v}_2 = 1$
- Minmax strategy against agent 1 requires agent 2 to play R
- R is strictly dominated by L for agent 2

Outline

Motivation

- So far, we assumed that agents observe actions of others at each round of game
- Next, we consider games where agents' actions may not be directly observable
- We assume that agents observe only an imperfect signal of stage game actions

Example: Cournot Competition with Noisy Demand

[Green and Porter, Non-cooperative Collusion under Imperfect Price Information, 1984]

- Firms set production levels $q_1^{(r)}, \ldots, q_n^{(r)}$ privately at round r
- Firms do not observe each others' output levels
- Market demand is stochastic
- Market price depends on total production and market demand
- Low price could be due to high production or low demand
- Firms utility depends on their own production and market price

Model

- We focus on game with public information
- At each round, all agents observe some public outcome
- Let $y^{(r)} \in Y$ denote publicly observed outcome at round r
- Each action profile a induces probability distribution over y
- Let $\pi(y, a)$ denote probability distribution of y under action profile a
- Public information at round r is $h^{(r)} = (y^{(1)}, \ldots, y^{(r-1)})$
- Strategy of agent i is sequence of maps $s_i^{(r)} : h^{(r)} \to S_i$

Model (cont.)

- Agents utility depends only on their own action and public outcome
- Dependence on actions of others is through their effect on distribution of y
- Agent i's realized utility at round r is $u_i(a_i^{(r)}, y^{(r)})$
- Agent i's expected stage utility is

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

- Agent i's average discounted utility when sequence $\{a^{(t)}\}$ is played is

$$(1 - \delta) \sum_{r=1}^{\infty} \delta^{r-1} u_i(a^{(r)})$$

Simpler Example: Noisy Prisoner's Dilemma

- Prisoners do not observe each others actions, instead, they observe signal y
  - $u_1(D, y) = 1 + y$ $\qquad u_1(C, y) = 4 + y$
  - $u_2(D, y) = 1 + y$ $\qquad u_2(C, y) = 4 + y$
- Signal y is defined by cont. random variable X with CDF $F(x)$ and $\mathbb{E}[X] = 0$
  - If $a = (D, D)$, then $y = X$
  - If $a = (D, C)$ or (C, D), then $y = X - 2$
  - If $a = (D, D)$, then $y = X - 4$
- Normal-form stage game is

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

Trigger-price Strategy

- Consider following trigger strategy
  - (I) - Play (D, D) until $y \leq y^*$, then go to (II)
  - (II) - Play(C, C) for R rounds, then go back to (I)
- Notice that strategy is stationary and symmetric
- Also notice that punishment uses NE of stage game
- We can choose y^* and R such that this strategy profile is SPE

Trigger-price Strategy (cont.)

- We use one-shot deviation principle
- Deviation in (II) is obviously not beneficial
- In (I), if agents do not deviate, their expected utility is

$$v = (1-\delta)\left((1+0) + \delta\left(F(y^*)\delta^R v + (1-F(y^*))v\right)\right)$$

- From this, we obtain

$$v = \frac{1-\delta}{1-\delta(1-\delta)\left(1-F(y^*)(1-\delta^R)\right)}$$

Trigger-price Strategy (cont.)

- If some agent deviates in (1), then her expected utility is

$$v_d = (1-\delta)\left((2+0) + \delta\left(F(y^*+2)\delta^R v + (1-F(y^*+2))v\right)\right)$$

- Deviation provides immediate utility, but increases probability of entering (II)
- To have SPE, we mush have $v \geq v_d$ which means

$$v \geq \frac{2(1-\delta)}{1-\delta(1-\delta)\left(1-F(y^*+2)(1-\delta^R)\right)}$$

$$\Rightarrow F(y^*+2) - 2F(y^*) \geq \frac{1-\delta(1-\delta)}{\delta(1-\delta)(1-\delta^R)}$$

- Any R and y^* that satisfy this constraint construct SPE
- Best trigger-price strategy can be found by maximizing v s.t. this constraint

Acknowledgment

- This lecture is a slightly modified version of ones prepared by
  - Asu Ozdaglar [MIT 6.254]
  - Vincent Conitzer [Duke CPS 590.4]
- Elly Khodaie helped with importing slides from PowerPoint to LaTeX