6.254 : Game Theory with Engineering Applications
Lecture 16: Repeated Games – II

Asu Ozdaglar
MIT

April 13, 2010

Outline

- Repeated Games – perfect monitoring
- Folk Theorems
- Repeated Games – imperfect monitoring
    - Price-trigger strategies

- Reference:
- Fudenberg and Tirole, Section 5.1 and 5.5.

Repeated Games

- By repeated games, we refer to a situation in which the same stage game (strategic form game) is played at each date for some duration of T periods.
- More formally, imagine that I players playing a strategic form game $G = \langle \mathcal{I}, (A_i)_{i \in \mathcal{I}}, (g_i)_{i \in \mathcal{I}} \rangle$ for T periods (infinitely-repeated game if $T = \infty$).
- At each period, *the outcomes of all past periods are observed by all players* $\Rightarrow$ perfect monitoring
- The notation $\mathbf{a} = \{a^t\}_{t=0}^{\infty}$ denotes the (infinite) sequence of action profiles.
- The payoff to player i for the entire repeated game is then

$$u_i(\mathbf{a}) = (1 - \delta) \sum_{t=0}^{\infty} \delta^t g_i(a_i^t, a_{-i}^t)$$

  where $\delta \in [0, 1)$.
- We have seen that grim trigger strategies can sustain "cooperation" in infinitely repeated games.
- This motivated the question what payoffs are achievable in equilibrium when players are sufficiently patient (i.e., when $\delta \approx 1$).

Folk Theorems for Infinitely Repeated Games

- We started last time talking about folk theorems which study equilibrium payoffs that can be obtained in infinitely repeated games.
- Recall that we use $\underline{v}_i$ to denote the minmax payoff of player i, i.e.,

$$\underline{v}_i = \min_{\alpha_{-i}} \max_{\alpha_i} g_i(\alpha_i, \alpha_{-i}).$$

- The strategy profile m^i_{-i} denotes the minmax strategy of opponents of player i and m^i_i denotes the best response of player i to m^i_{-i}, i.e., $g_i(m^i_i, m^i_{-i}) = \underline{v}_i$.
- The set $V^* \subset \mathbb{R}^I$ denotes the set of feasible and strictly individually rational payoffs.
- We have seen that equilibrium payoffs of repeated games are bounded from below by minmax payoffs.

Folk Theorems

Our first theorem is the Nash folk theorem which shows that any feasible and strictly individually rational payoff vector can be achieved if the player is sufficiently patient.

Theorem (Nash Folk Theorem)

If $(v_1, \ldots, v_I)$ is feasible and strictly individually rational, then there exists some $\underline{\delta} < 1$ such that for all $\delta > \underline{\delta}$, there is a Nash equilibrium of $G^\infty(\delta)$ with payoffs $(v_1, \cdots, v_I)$.

- The Nash folk theorem states that essentially any payoff can be obtained as a Nash Equilibrium when players are patient enough.
- However, the corresponding strategies involve this non-forgiving punishments, which may be very costly for the punisher to carry out (i.e., they represent non-credible threats).
- This implies that the strategies used may not be subgame perfect.
- Our next step is to get the set of feasible and strictly individually rational payoffs as the subgame perfect equilibria payoffs of the repeated game.

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 e^{NE}. For any feasible payoff v with $v_i > e_i^{NE}$ for all $i \in \mathcal{I}$, there exists some $\underline{\delta} < 1$ such that for all $\delta > \underline{\delta}$, there exists a subgame perfect equilibrium of $G^{\infty}(\delta)$ with payoffs v.

Proof: Simply construct the non-forgiving trigger strategies with punishment by the static Nash Equilibrium. Punishments are therefore subgame perfect. For δ sufficiently close to 1, it is better for each player i to obtain v_i rather than deviate and get a high deviation payoff for one period, and then obtain e_i^{NE} forever thereafter.

Subgame Perfect Folk Theorem (continued)

Theorem (Fudenberg and Maskin)

Assume that the dimension of the set V of feasible payoffs is equal to the number of players I. Then, for any $v \in V$ with $v_i > \underline{v}_i$ for all i, there exists a discount factor $\underline{\delta} < 1$ such that for all $\delta \geq \underline{\delta}$, there is a subgame perfect equilibrium of $G^\infty(\delta)$ with payoffs v.

- The proof of this theorem is more difficult, but the idea is to use the assumption on the dimension of V to ensure that each player i can be *singled out* for punishment in the event of a deviation, and then use rewards and punishments for other players to ensure that the deviator can be held down to her minmax payoff.

Cooperation in Finitely-Repeated Games

- We saw that finitely-repeated games with unique stage equilibrium do not allow corporation or any other outcome than the repetition of this unique equilibrium.
- But this is no longer the case when there are multiple equilibria in the stage game.
- Consider the following example

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

- The stage game has two pure Nash equilibria (B, B) and (C, C). The most cooperative outcome, (A, A), is not an equilibrium.
- **Main result in example:** in the twice repeated version of this game, we can support (A, A) in the first period.

Cooperation in Finitely-Repeated Games (continued)

- Idea: use the threat of switching to (C, C) in order to support (A, A) in the first period and (B, B) in the second.
- Suppose, for simplicity, no discounting.
- If we can support (A, A) in the first period and (B, B) in the second, then each player will receive a payoff of 4.
- If a player deviates and plays B in the first period, then in the second period the opponent will play C, and thus her best response will be C as well, giving her -1. Thus total payoff will be 3. Therefore, deviation is not profitable.

How Do People Play Repeated Games?

- In lab experiments, there is more cooperation in prisoners' dilemma games than predicted by theory.
- More interestingly, cooperation increases as the game is repeated, even if there is only finite rounds of repetition.
- Why?
- Most likely, in labs, people are confronted with a payoff matrix of the form:

	Cooperate	Defect
Cooperate	$1, 1$	$-1, 2$
Defect	$2, -1$	$0, 0$

- Entries are monetary payoffs. But we should really have people's **full payoffs**.
- These may differ because of social preferences.

Social Preferences

- Types of social preferences:
  1. **Altruism**: people receive utility from being nice to others.
  2. **Fairness:** people receive utility from being fair to others.
  3. **Vindictiveness:** people like to punish those deviating from "fairness" or other accepted norms of behavior.
- All of these types of social preferences seem to play some role in experimental results.

Repeated Games with Imperfect Monitoring

- So far, we have assumed that in the repeated play of the game, each player observes the actions of others at the end of each stage.
- Here, we consider the problem of repeated games where player's actions may not be directly observable.
- We assume that players observe only an imperfect signal of the stage game actions.

Motivational Example

Cournot competition with noisy demand. [Green and Porter 84]

- Firms set output levels $q_1^t, \ldots, q_I^t$ 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 output levels.
- The market price depends on uncertain demand and the total output, i.e., a low market price could be due to high produced output or low demand.

We will focus on games with public information: At the end of each period, all players observe a public outcome, which is correlated with the vector of stage game actions. Each player's realized payoff depends only on his own action and the public outcome.

Model

We state the components of the model for a stage game with finite actions and a public outcome that can take finitely many values. The model extends to the general case with straightforward modifications.

- Let $A_1, \ldots, A_I$ be finite action sets.
- Let y denote the publicly observed outcome, which belong to a (finite) set Y. Each action profile a induces a probability distribution over $y \in Y$. Let $\pi(y, a)$ denote the probability distribution of y under action profile a.
- Player i realized payoff is given by $r_i(a_i, y)$. Note that this payoff depends on the actions of the other players through their effect on the distribution of y.
- Player i's expected stage payoff is given by

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

Model (Continued)

- A mixed strategy is $\alpha_i \in \Delta(A_i)$. Payoffs are defined in the obvious way.
- 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 $s_i^t : h^t \longrightarrow A_i$.
- Player i's average discounted payoff when the action sequence $\{a^t\}$ is played is given by

$$(1-\delta)\sum_{t=0}^{\infty}\delta^t g_i(a^t)$$

Example: Noisy Prisoner's Dilemma

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

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

- Probability distribution for public signal p:

  $a_1 = a_2 = C \rightarrow p = X,$
  $a_1 \neq a_2 \quad\quad \rightarrow p = X - 2,$
  $a_1 = a_2 = D \rightarrow p = X - 4,$
  where X is a continuous random variable with cumulative distribution function $F(x)$ and $E[X] = 0$.

- The payoff matrix for this game takes the following form:

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

Trigger-price Strategy

- Consider the following trigger type strategy for the noisy prisoner's dilemma:
  - **(I)** - Play (C, C) until $p \leq p^*$, then go to Phase II.
  - **(II)** - Play (D, D) for T periods, then go back to Phase I.
- Notice the strategy is stationary and symmetric. Also notice 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 SPE.
- Find the continuation payoffs:
- In Phase I, if players do not deviate, their expected payoff will be

$$v = (1-\delta)(1+0) + \delta\Big[F(p^*)\delta^T v + (1-F(p^*))v\Big].$$

From this equation, we obtain

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

Trigger-price Strategy

- If the player deviates, his payoff will be

$$v_d = (1-\delta)\{(2+0) + \delta[F(p^*+2)\delta^T v + (1-F(p^*+2))v]\}$$

- Note that deviating provides immediate payoff, but increases the probability of entering Phase II. In order for the strategy to be an SPE, the expected difference in payoff from the deviation must not be positive.
- Incentive Compatibility Constraint: $v \geq v_d$

$$v \geq (1-\delta)\{2 + F(2+p^*)\delta^{T+1}v + [1-F(2+p^*)]\delta v\}$$

Substituting v, we have

$$\frac{1}{1-\delta^{T+1}F(p^*) - \delta(1-F(p^*))} \geq \frac{2}{1-F(2+p^*)\delta^{T+1} - \delta[1-F(2+p^*)]}$$

- Any T and p^* that satisfy this constraint would construct an SPE. The best possible trigger-price equilibrium strategy could be found if we could maximize v subject to the incentive compatibility constraint.

Other Equilibria

- Question: Are there other equilibria with higher payoffs ?
- There's an explicit characterization of the set of equilibrium payoffs using a multi-agent dynamic programming-type operator, by Abreu, Pearce and Stachetti (90). The set of equilibrium payoffs is a fixed point of this operator and the value iteration converges to the set of equilibrium payoffs.
- For more on this topic, see Fudenberg and Tirole section 5.5 and Abreu, Pearce and Stachetti (90).

6.254 Game Theory with Engineering Applications
Spring 2010