6.254: Game Theory with Engineering Applications	April 10, 2008

Lecture 16: Repeated Games -I

Lecturer: Asu Ozdaglar

Agenda

- Repeated games (perfect monitoring)
- Folk Theorems

In many strategic situation, players interact repeatedly over time, therefore it is important to understand the effect of repetition on the play of the game. The repeated game model we will present in this section is a simple model to capture the ongoing interaction. In particular,

- The players face the same stage game at all periods.
- Overall payoff is the sum of discounted payoffs at each stage.

We will see in this lecture how repeated play of the same strategic game introduces new (desirable) equilibria by allowing players to condition their actions on the way their opponents played in the previous periods.

Some Examples:

Example 1: (Cooperation in Prisoners Dilemma) The best known repeated game argument is that ongoing interaction can explain why people might behave cooperatively when it is against their self-interest in the short run. The classical example is the repeated prisoner's dilemma.

	Confess	Defect
Confess	$1, 1$	$-1, 2$
Defect	$2, -1$	$0, 0$

For this strategic form game, the strategy profile (D, D) is the unique NE. Moreover D strictly dominates C.

Suppose now the players 1, 2 play the game repeatedly at $0, 1, 2, \cdots$ and the payoff for the entire repeated game is:

$$u_1(\{\mathbf{a}^0, \mathbf{a}^1\}, \cdots) = (1 - \delta) \sum_{t=0}^{\infty} \delta^t g_i(a_i^t, a_{-i}^t)$$

in which $\delta \in [0, 1) \rightarrow$, i.e., players discount the future.

Play once: unique equilibrium $\rightarrow (D, D)$.

Play it T times: Do backward induction. Again, there is a unique subgame perfect equilibrium (SPE) in which both players defect in each period.

Now assume that the game is played infinitely often. Is playing (D, D) in every period still an SPE outcome?

Proposition 1 *If $\delta \geq \frac{1}{2}$ the repeated PD game has an SPE in which (C, C) is played in every period.*

Proof: Suppose the players use the "grim trigger" strategy given by the following. For player i:

I. Play C in every period unless someone plays D, in which case go to stage II.

II. Play D forever.

We next show that the preceding strategy is an SPE if $\delta \geq \frac{1}{2}$ using the one-stage-deviation principle.

There are two kinds of sub-games:

(1) Subgame following a history in which no player has ever defected (i.e., D has never been played).

(2) Any other subgame (i.e., D has been played at some point in the past).

Consider first a subgame of the form (1), i.e., suppose up to time t, D has never been played. Player i's continuation payoffs when he stick with his strategy and when he deviates for one stage and then conforms to his strategy thereafter are given respectively by:

$$\left.\begin{array}{l}\text{Play } C : (1-\delta)[1+\delta+\delta^2+\cdots] = 1 \\ \text{Play } D : (1-\delta)[2+0+0+\cdots] = 2(1-\delta)\end{array}\right\} \Longrightarrow \text{This shows that for } \delta \geq \tfrac{1}{2} \text{ deviation is}$$

not profitable.

Consider next a subgame of form (2), i.e., action D is played at some point before t. Since (D, D) is the NE of the static game and choosing C does not have an effect on subsequent outcomes, it follows that, for any discount factor, no player can profitably deviate by choosing C in one period. □

Remarks:

1. Depending on size of the discount factor, there may be many more equilibria.

2. If a^* is the NE of the stage game, then the strategies "each player, plays a_i^*" form an SPE. (Note that with these strategies, future play of the opponent is independent of how I play today, therefore, the optimal play is to maximize the current payoff, i.e., play a static best response.)

3. Sets of equilibria for finite and infinite horizon versions of the "same game" can be quite different.

The following example shows that repeated play can lead to <u>worse</u> outcomes than in the one shot game:

Example 2:

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

For the game defined above, the action A strictly dominates B, C for both players, therefore the unique NE is (A, A).

Proposition 2 *If $\delta \geq \frac{1}{2}$ this game has an SPE in which (B, B) is played in every period.*

Proof Sketch: Here, we construct a slightly more complicated strategy than grim trigger:

I. Play B in every period unless someone deviates, then go to II.

II. Play C. If no one deviates go to I. If someone deviates stay in II.

□

Exercise: Show that the preceding strategy profile is an SPE of the repeated game for $\delta \geq \frac{1}{2}$.

General Model

- Let G be a strategic form game with action spaces $A_1, \cdots, A_I$ and (stage) payoff functions $g_i : A \to \mathbb{R}$, where $A = A_1 \times \cdots \times A_I$.
- Let $G^\infty(\delta)$ be the infinitely repeated version of G played at $t = 0, 1, 2, \ldots$, where players discount payoffs with the factor δ and observe all previous actions.
- Payoffs for player i:

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

We now investigate what average payoffs could result from different equilibria when δ is close to 1. That is, what can happen in equilibrium when players are very patient?

Some terminology related to payoffs:

1. *Set of feasible payoffs*:

$$V = \text{Conv}\{v \mid \text{there exists } a \in A \text{ such that } g(a) = v\}.$$

Example 1: Consider the following game:

	L (q)	R $(1-q)$
U	$-2, -2$	$1, -2$
M	$1, -1$	$-2, 2$
D	$0, 1$	$0, 1$

Sketch set V for this example.

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

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

Let

$$m^i_{-i} = \arg\min_{\alpha_{-i}} [\max_{\alpha_i} \; g_i(\alpha_i, \alpha_{-i})],$$

i.e. minimax strategy profile against player i.

Example 2: We compute the minimax payoffs for Example 1. To compute $\underline{v}_1$, let q denote the probability that player 2 chooses action L. Then player 1's payoffs for playing different actions are given by:

$$\left.\begin{array}{l} U \rightarrow 1 - 3q \\ M \rightarrow -2 + 3q \\ D \rightarrow 0 \end{array}\right\}$$

Therefore, we have

$$\underline{v}_1 = \min_{0 \leq q \leq 1} \quad [\max\{1 - 3q, -2 + 3q, 0\}] = 0,$$

and $m^1_2 \in [\frac{1}{3}, \frac{2}{3}]$.

Similarly, one can show that: $\underline{v}_2 = 0$, and $m^2_1 = (1/2, 1/2, 0)$ is the unique minimax profile.

Verify that payoff of player 1 at any NE is 0.

Remark:

1. Player i's payoff is at least $\underline{v}_i$ in any static NE, and in any NE of the repeated game.

   - Static equilibrium: $\hat{\alpha} \Longrightarrow \underline{v}_i = \min_{\alpha_i} [\max_{\alpha_i} \; g_i(\alpha_i, \alpha_{-i})] \leq \max_{\alpha_i} \; g_i(\alpha_i, \alpha_{-i})$
   - Extend the above idea for the repeated case.

   Check the following: In the Prisoner's Dilemma example, NE payoff is $(0, 0)$ and minmax payoff is $(0, 0)$. In the second example introduced in the beginning of the lecture, NE payoff is $(2, 2)$ and minmax payoff is $(0, 0)$.

   We say that a payoff vector $\mathbf{v} \in \Re^I$ is strictly individually rational if $v_i > \underline{v}_i$ for all i.

Folk Theorems: We have observed above that no equilibrium of the repeated game can go below $\underline{v}_i$. We next investigate the payoffs that can be obtained at any equilibrium of the repeated game.

Theorem 1 *(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 NE of $G^\infty(\delta)$ with payoffs $(v_1, \cdots, v_I)$.*

Proof: Assume that there exists an action profile $a = (a_1, \cdots, a_I)$ s.t. $g_i(a) = v$. Recall:

m^i_{-i}: minimax strategy of opponent of i

m^i_i: i's best response to m^i_{-i}.

Now consider the following grim trigger strategy. For player i:

I. Play $(a_1, \cdots, a_I)$ as long as no one deviates (or more than one player deviates). If one player deviates go to II.

II. If some player j deviates, then play m^j_i thereafter.

We will complete the proof of this theorem in the next lecture.

□