Lecture XIII: Repeated Games

Markus M. Möbius

April 23, 2003

Readings for this class: Fudenberg and Tirole, page 146-154; Dutta, chapter 15

1 Introduction

So far one might get a somewhat misleading impression about SPE. When we first introduced dynamic games we noted that they often have a large number of (unreasonable) Nash equilibria. In the models we've looked at so far SPE has 'solved' this problem and given us a unique NE. In fact, this is not really the norm. We'll see today that many dynamic games still have a very large number of SPE.

2 Credible Threats

We introduced SPE to rule out non-credible threats. In many finite horizon games though credible threats are common and cause a multiplicity of SPE.

Consider the following game:

	L	C	R
T	3,1	0,0	5,0
M	2,1	1,2	3,1
B	1,2	0,1	4,4

The game has three NE: (T,L), (M,C) and $\left(\frac{1}{2}T + \frac{1}{2}M, \frac{1}{2}L + \frac{1}{2}C\right)$

Suppose that the players play the game twice and observe first period actions before choosing the second period actions. Now one way to get a SPE is to play any of the three profiles above followed by another of them (or same one). We can also, however, use credible threats to get other actions played in period 1, such as:

- Play (B,R) in period 1.
- If player 1 plays B in period 1 play (T,L) in period 2 - otherwise play (M,C) in period 2.

It is easy to see that no single period deviation helps here. In period 2 a NE is played so obviously doesn't help.

- In period 1 player 1 gets $4 + 3$ if he follows strategy and at most $5 + 1$ if he doesn't.
- Player 2 gets $4 + 1$ if he follows and at most 2+1 if he doesn't.

Therefore switching to the (M,C) equilibrium in period 2 is a *credible threat*.

3 Repeated Prisoner's Dilemma

Note, that the PD doesn't have multiple NE so in a finite horizon we don't have the same easy threats to use. Therefore, the finitely repeated PD has a unique SPE in which every player defects in each period.

	C	D
C	1,1	−1,2
D	2,−1	0,0

In infinite horizon, however, we do get many SPE because other types of threats are credible.

Proposition 1 *In the infinitely repeated PD with $\delta \geq \frac{1}{2}$ there exists a SPE in which the outcome is that both players cooperate in every period.*

Proof: Consider the following symmetric profile:

$$s_i(h_t) = \begin{cases} C & \text{If both players have played C in every} \\ & \text{period along the path leading to } h_t. \\ D & \text{If either player has ever played D.} \end{cases}$$

To see that there is no profitable single deviation note that at any h_t such that $s_i(h_t) = D$ player i gets:

$$0 + \delta 0 + \delta^2 0 + ..$$

if he follows his strategy and

$$-1 + \delta 0 + \delta^2 0 + ..$$

if he plays C instead and then follows s_i.

At any h_t such that $s_i(h_t) = C$ player i gets:

$$1 + \delta 1 + \delta^2 1 + .. = \frac{1}{1 - \delta}$$

if he follows his strategy and

$$2 + \delta 0 + \delta^2 0 + .. = 2$$

if he plays D instead and then follows s_i.

Neither of these deviations is worth while if $\delta \geq \frac{1}{2}$. QED

Remark 1 *While people sometimes tend to think of this as showing that people will cooperate in they repeatedly interact if does not show this. All it shows is that there is one SPE in which they do. The correct moral to draw is that there many possible outcomes.*

3.1 Other SPE of repeated PD

1. For any δ it is a SPE to play D every period.
2. For $\delta \geq \frac{1}{2}$ there is a SPE where the players play D in the first period and then C in all future periods.
3. For $\delta > \frac{1}{\sqrt{2}}$ there is a SPE where the players play D in every even period and C in every odd period.
4. For $\delta \geq \frac{1}{2}$ there is a SPE where the players play (C,D) in every even period and (D,C) in every odd period.

3.2 Recipe for Checking for SPE

Whenever you are supposed to check that a strategy profile is an SPE you should first try to classify all histories (i.e. all information sets) on and off the equilibrium path. Then you have to apply the SPDP for each class separately.

- Assume you want to check that the cooperation with grim trigger punishment is SPE. There are two types of histories you have to check. Along the equilibrium path there is just one history: everybody cooperated so far. Off the equilibrium path, there is again only one class: one person has defected.

- Assume you want to check that cooperating in even periods and defecting in odd periods plus grim trigger punishment in case of deviation by any player from above pattern is SPE. There are three types of histories: even and odd periods along the equilibrium path, and off the equilibrium path histories.

- Assume you want to check that TFT ('Tit for Tat') is SPE (which it isn't - see next lecture). Then you have you check four histories: only the play of both players in the last period matters for future play - so there are four relevant histories such as player 1 and 2 both cooperated in the last period, player 1 defected and player 2 cooperated etc.1

Sometimes the following result comes in handy.

Lemma 1 *If players play Nash equilibria of the stage game in each period in such a way that the particular equilibrium being played in a period is a function of time only and does not depend on previous play, then this strategy is a Nash equilibrium.*

The proof is immediate: we check for the SPDP. Assume that there is a profitable deviation. Such a deviation will not affect future play by assumption: if the stage game has two NE, for example, and NE1 is played in even periods and NE2 in odd periods, then a deviation will not affect future play.[1] Therefore, the deviation has to be profitable in the current stage game - but since a NE is being played no such profitable deviation can exist.

Corollary 1 *A strategy which has players play the same NE in each period is always SPE.*

In particular, the grim trigger strategy is SPE if the punishment strategy in each stage game is a NE (as is the case in the PD).

4 Folk Theorem

The examples in 3.1 suggest that the repeated PD has a tremendous number of equilibria when δ is large. Essentially, this means that game theory tells us we can't really tell what is going to happen. This turns out to be an accurate description of most infinitely repeated games.

[1] If a deviation triggers a switch to only NE1 this statement would no longer be true.

Let G be a simultaneous move game with action sets A_1, A_2, .., A_I and mixed strategy spaces Σ_1, Σ_2,.., Σ_I and payoff function $\tilde{u}_i$.

Definition 1 *A payoff vector $v = (v_1, v_2, .., v_I) \subset \Re^I$ is feasible if there exists action profiles a^1, a^2, ..,$a^k \in A$ and non-negative weights ω_1,..,ω_I which sum up to 1 such that*

$$v_i = \omega_1 \tilde{u}_i \left(a^1\right) + \omega_2 \tilde{u}_i \left(a^2\right) + .. + \omega_k \tilde{u}_i \left(a^k\right) +$$

Definition 2 *A payoff vector v is strictly individually rational if*

$$v_i > \underline{v}_i = \min_{\sigma_{-i} \in \Sigma_{-i}} \max_{\sigma_i(\sigma_{-i}) \in \Sigma_i} \tilde{u}_i \left(\sigma_i \left(\sigma_{-i}\right), \sigma_{-i}\right) \tag{1}$$

We can think of this as the lowest payoff a rational player could ever get in equilibrium if he anticipates his opponents' (possibly non-rational) play.

Intuitively, the minmax payoff $\underline{v}_i$ is the payoff player 1 can guarantee to herself even if the other players try to punish her as badly as they can. The minmax payoff is a measure of the punishment other players can inflict.

Theorem 1 ***Folk Theorem.*** *Suppose that the set of feasible payoffs of G is I-dimensional. Then for any feasible and strictly individually rational payoff vector v there exists $\underline{\delta} < 1$ such that for all $\delta > \underline{\delta}$ there exists a SPE x^* of G^∞ such that the average payoff to s^* is v, i.e.*

$$u_i \left(s^*\right) = \frac{v_i}{1 - \delta}$$

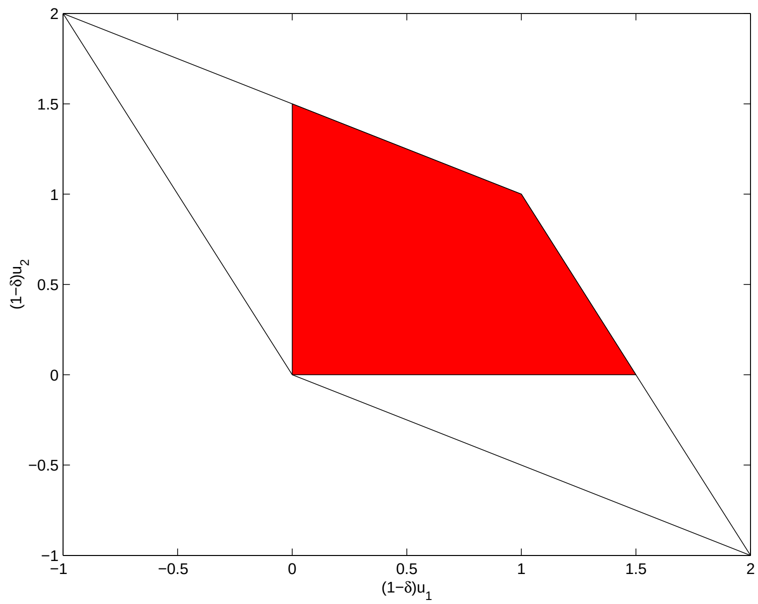
The normalized (or average) payoff is defined as $P = (1 - \delta)\, u_i \left(s^*\right)$. It is the payoff which a stage game would have to generate in each period such that we are indifferent between that payoff stream and the one generates by s^*:

$$P + \delta P + \delta^2 P + ... = u_i \left(s^*\right)$$

4.1 Example: Prisoner's Dilemma

- The feasible payoff set is the diamond bounded by (0,0), (2,-1), (-1,2) and (1,1). Every point inside can be generated as a convex combinations of these payoff vectors.

- The minmax payoff for each player is 0 as you can easily check. The other player can at most punish his rival by defecting, and each player can secure herself 0 in this case.



Hence the theorem says that anything in this trapezoid is possible. Note, that the equilibria I showed before generate payoffs inside this area.

4.2 Example: BOS

Consider the Battle of the Sexes game instead.

	F	O
F	2,1	0,0
O	0,0	1,2

Here each player can guarantee herself at least payoff $\frac{2}{3}$ which is the payoff from playing the mixed strategy Nash equilibrium. You can check that whenever player 2 mixes with different probabilities, player 1 can guarantee herself more than this payoff by playing either F or O all the time.

4.3 Idea behind the Proof

1. Have players on the equilibrium path play an action with payoff v (or alternate if necessary to generate this payoff).[2]

2. If some player deviates punish him by having the other players for T periods choose σ_{-i} such that player i gets $\underline{v}_i$.

3. After the end of the punishment phase reward all players (other than i) for having carried out the punishment by switching to an action profile where player i gets $v_i^P < v_i$ and all other players get $v_j^P + \epsilon$.

[2]For example, in the BoS it is not possible to generate $\left(\frac{3}{2}, \frac{3}{2}\right)$ in the stage game even with mixing. However, if players alternate and play (O,F) and then (F,O) the players can get arbitrarily close for large δ.