

MA300.2 Game Theory II, LSE

Summary of Lecture 2

Repeated Games

1. THE GENERAL SETUP

Let $G = (A_1, \dots, A_n; f_1, \dots, f_n)$ be an n -person game in normal form. We are interested in *repetitions* of G , perhaps infinitely many times. This gives rise to a new game which we shall call G^T , where T is number of repetitions (possibly infinite). The strategies in the “stage game” will now be referred to as actions, and we shall reserve the term strategy for behavior in the repeated game. Hence the change in notation from S to A .

The important assumption that we make here is that *past actions are perfectly observable*. (Much) later we shall relax this assumption. We shall also suppose that a pure action is taken at every stage of the repeated game.

A *path* is given by a sequence $\mathbf{a} \equiv \{a(t)\}_{t=0}^T$, where $a(t) \in A$ (set of action *profiles*) for all t . Along a path \mathbf{a} , the (normalized) payoff to player i is given by

$$F_i(\mathbf{a}) = \sum_{t=0}^T \beta^t f_i(a(t)).$$

where $\beta \in (0, 1)$ is a discount factor common to all players.

Game starts at date 0. *history* at date $t \geq 1$ is a full description of all that has transpired in the past, up to and including date $t - 1$. Obviously the space of histories at date t is just the time product of action-profile sets A^t .

A *strategy* for player i is a specification of an action at date 0, and thereafter an action conditional on every history. Formally, a strategy σ_i spells out an action $\sigma_i(h(t)) \in A_i$ for every conceivable history $h(t)$ at date t .

A strategy profile generates a path $\mathbf{a}(\sigma)$, as well as a path $\mathbf{a}(\sigma, h(t))$ conditional on every history at date t . In the latter case, the path obviously starts with the given history in the first t dates.

A strategy profile σ is a *Nash equilibrium* if for every player i and every strategy σ'_i ,

$$F_i(\mathbf{a}(\sigma)) \geq F_i(\mathbf{a}(\sigma_{-i}, \sigma'_i)).$$

And σ is a *subgame perfect Nash equilibrium* (SGPE) if it is a Nash equilibrium, and for every history $h(t)$, every player i , and every alternative strategy σ'_i ,

$$F_i(\mathbf{a}(\sigma, h(t))) \geq F_i(\mathbf{a}(\sigma_{-i}, \sigma'_i, h(t))).$$

2. A FIRST LOOK AT REPETITIONS

Example. Prisoner's Dilemma (PD)

| | | |
|-------|-------|-------|
| | C_2 | D_2 |
| C_1 | 2, 2 | 0, 3 |
| D_1 | 3, 0 | 1, 1 |

Repeat this. Various strategy profiles possible.

Cooperate forever: start with C_i , and set $\sigma_i(h(t)) = C_i$ for all conceivable t -histories.

Defect forever: start with D_i , and set $\sigma_i(h(t)) = D_i$ for all conceivable t -histories.

Grim trigger: Start with C_i , thereafter $\sigma_i(h(t)) = C_i$ if and only if every *opponent* entry in $h(t)$ is a " C_j ".

Modified grim trigger: Start with C_i , thereafter $\sigma_i(h(t)) = C_i$ if and only if *every* entry (mine or the opponent's) in $h(t)$ is a " C ".

Tit for tat: Start with C_i , thereafter $\sigma_i(h(t)) = x_i$, where x has the same letter as the opponent's last action under $h(t)$.

THEOREM 1. *The only SGPE in a finitely repeated PD is "defect forever".*

How general is this observation? Well, it works for any game in which the stage Nash equilibrium is unique, such as the Cournot oligopoly.

What if there is more than one equilibrium in the stage game?

Example. Variant of the PD

| | | | |
|-------|-------|-------|-------|
| | C_2 | D_2 | E_2 |
| C_1 | 5, 5 | 0, 0 | 0, 6 |
| D_1 | 0, 0 | 4, 4 | 0, 1 |
| E_3 | 6, 0 | 1, 0 | 1, 1 |

Let $T = 1$; i.e., game is played at periods 0 and 1. Consider the strategy for player i :

Start with C_i .

Next, if both players have played C in period 0, play D_i in period 1.

Otherwise, play E_i in period 1.

Claim: if discount factor is no less than $1/3$, these strategies form a SGPE.

The subgame strategies generate Nash equilibria. So all we have to do is check deviations in period 0. Given the opponent's strategy as described, if player i plays according to her prescribed strategy she gets $5 + \beta 4$. If she misbehaves, the best she can get is $6 + \beta$. The former is better than the latter if $\beta \geq 1/3$.

Calculus of cooperation and punishment, constrained by the requirement that the punishment has to be credible.

Finitely repeated games can therefore take us in interesting directions provided we are willing to rely on multiple stage equilibria.

Example. Prisoner's Dilemma again, this time infinitely repeated.

| | | |
|-------|-------|-------|
| | C_2 | D_2 |
| C_1 | 2, 2 | 0, 3 |
| D_1 | 3, 0 | 1, 1 |

Consider the modified grim trigger strategy: Start with C_i , thereafter $\sigma_i(h(t)) = C_i$ if and only if every entry in $h(t)$ is a "C".

Claim: if discount factor is no less than $1/2$, the modified grim trigger profile is a SGPE.

Two kinds of histories to consider:

1. Histories in which nothing but C has been played in the past (this includes the starting point).
2. Everything else.

In the second case, we induce the infinite repetition of the one-shot equilibrium which is subgame perfect. [This will need more discussion; see Theorem 2 below.]

So all that's left is the first case. In that case, by complying you get $2/(1-\beta)$. By disobeying, the best you can get is $3 + \beta/(1-\beta)$. If $\beta \geq 1/2$, the former is no smaller than the latter.

Examine perfection of the grim trigger, the modified trigger with limited punishments, tit-for-tat.

3. THE ONE-SHOT DEVIATION PRINCIPLE

In the last example, we only checked *one* deviation from an ongoing strategy. How do we know that this is enough to deter *all* deviations, however complicated? In the PD, perhaps, the answer is clear because of the simplicity of the game. But what about more complicated games? In this section, we show that a remarkable reduction in complexity can be achieved by the use of the *one-shot deviation principle*.

The first cut on this problem is to think about finite game trees that terminate after a certain number of repetitions.

STEP 1. Pick a node x and suppose that from that point on the strategy of a particular agent — call her Sally — is not a best response. Then Sally has another strategy (compared to the original one) that does better. Apply Sally's new strategy (along with the given strategies of the other agents) and generate a path along the tree starting from x . Look at every node on the tree at which Sally needs to move (starting from x), and look at the *last node* along this path at which Sally plays *differently* from her original strategy. Call this node y .

Suppose at y we simply switched Sally's action back to her original action. Would this destroy the gain from this entire chain of deviations? There are only two answers:

Yes. In that case, it is this last deviation at y that is responsible for Sally's gain. We have thus found a one-shot deviation at y !

No. In that case, simply consider the "reduced" chain of deviations by replacing Sally's move at y with her original move. Now move up to the last node for *this* reduced chain for which Sally's action is different from the original prescription. Call this y' .

Ask the same question again: if we switch Sally's action back to her original action at y' , does this destroy the gain from this entire chain of deviations? Again, if the answer is "yes", we are done, but if the answer is "no" we replace the deviation at y' with Sally's original action, leading to a still smaller chain of deviations. In this way, we must either get a "yes" to our question at some stage, or we will have reduced the chain to a single profitable deviation. Either way, we will have found a one-shot profitable deviation.

Now, this argument does *not* work in the infinite case free of charge.

Example. Consider the following (rather silly, but perfectly legitimate) one-person game, with an infinite number of nodes. At each node the person can choose a or b . If she chooses a infinitely many times, she gets a payoff of 1, otherwise she gets a payoff of 0. Consider the strategy in which the person chooses b each time she moves. This strategy is sub-optimal, but there is no one-shot profitable deviation from it.

However, in repeated games with discounting, or indeed in all games with discounting (and satisfying some natural consistency conditions), this sort of counterexample cannot be provided. Here is the argument for repeated games. [If you are interested, you can look at a general statement and proof of the one-shot deviation principle, available on the course webpage. But the material here is only for those of you who are really getting into game theory, and is not required reading for this course.]

STEP 2. Suppose that a strategy for Sally — call it σ — is not a best response to the strategy profile her opponents are using, at some subgame. Then there is a history h_t following which Sally makes a sequence of moves different from those specified by σ . If the number of different moves or deviations is finite, use Step 1. If it is infinite, proceed as follows. Let g be the total discounted value of the gain experienced by Sally, starting from date t when her deviations first begin. Choose a date $s > t$ such that $\beta^{s-t}M < g/2$, where M is the best conceivable one-period payoff to Sally in the stage game.¹

If all deviations by Sally after period s are simply restored to their original moves (indeed, even if we ignore all of Sally's payoffs after date s), the effect that this will have on date t payoffs will be no more than $g/2$, by construction. This means that the deviations between periods t and s create a gain of at least $g/2$. We have therefore found a finite set of profitable deviations, and now Step 1 can be applied again.

4. COLLUSION AND PUNISHMENTS I: REVERSION TO NASH EQUILIBRIUM

The one-shot deviation principle significantly simplifies our study of repeated games.

¹I am assuming that this best payoff is well-defined and finite, as it will be, for instance, when the stage game has a finite number of actions for everybody.

For instance, the following proposition is true:

THEOREM 2. *The repetition of any one-shot equilibrium is subgame perfect, and therefore any outcome which is preferred (by each player) to a one-time deviation followed by a Nash equilibrium forever is supportable as a subgame perfect equilibrium outcome.*

Proof. Consider any strategy profile which prescribes the play of a one-shot Nash equilibrium action profile forever, regardless of history. By the OSD principle, it suffices to check whether a single deviation is profitable. But a single deviation cannot be profitable *in that period*, because the profile is one-shot Nash, and thereafter the same Nash profile will be repeated. So there is no gain (present or future) from a one-shot deviation, and therefore the strategy profile in question is subgame perfect.

Now fix some Nash equilibrium with payoffs (v_1^*, \dots, v_n^*) . Consider some action profile \mathbf{a} . For player i , define $d_i(\mathbf{a})$ to be the best one-shot payoff that i can generate under the assumption that others hold their actions at \mathbf{a}_{-i} ; i.e.,

$$d_i(\mathbf{a}) \equiv \max_{a_i'} f_i(a_i', \mathbf{a}_{-i}).$$

Now suppose that for every player i ,

$$(1) \quad f_i(\mathbf{a}) \geq (1 - \beta)d_i(\mathbf{a}) + \beta v_i^*.$$

Then \mathbf{a} must be supportable as a SGP equilibrium outcome, by the OSD principle. \square

The modified grim strategy is the kind of strategy profile that supports the outcomes described in Theorem 2.

This sort of argument rapidly leads to the first of some famous theorems in game theory, known as the *folk theorems*. The idea is simple: if players are sufficiently patient, lots of different outcomes can be supported. The first theorem of this kind that we look at is called the Nash reversion theorem, because it features strategies that retreat to one-shot Nash equilibria in the event of a deviation.

THEOREM 3. [Friedman's Nash Reversion Theorem.] *Suppose that w_i^* is the worst payoff to player i generated by the entire set of one-shot Nash equilibria. Consider any action profile \mathbf{a} such that $f_i(\mathbf{a}) > w_i^*$ for every i . Then, if players are sufficiently patient, there is a subgame perfect equilibrium with outcome \mathbf{a} .*

Consider a strategy profile σ that dictates the play of \mathbf{a} initially, and in all histories where \mathbf{a} has been faithfully played in the past. For histories in which this is not true, look at the first date of departure from \mathbf{a} at which only a single player deviated.

If there is such a date, with deviating player i , play the Nash equilibrium that yields i 's worst payoff.

If all deviations from \mathbf{a} involve multi-player deviations, play some preassigned Nash equilibrium (it doesn't matter which one).

Notice that the above covers all possible types of histories.

To check whether σ is SGP, it suffices (by Theorem 2 and the OSD principle) to only look at those histories in which \mathbf{a} has been faithfully played throughout. For such histories, it suffices to check (1) — replacing v by w — for every player i ; that is, check whether

$$(2) \quad f_i(\mathbf{a}) \geq (1 - \beta)d_i(\mathbf{a}) + \beta w_i^*.$$

But $f_i(\mathbf{a}) > w_i^*$ by assumption, so that there exists β_i^* close enough to 1 such that (2) holds for all β between β_i^* and 1. This completes the proof. \square