Handout: Repeated Games with Observable Actions[1]

link to latest version

No Profitable One-Shot Deviation Property

Dynamic programming gives us two results about one-stage deviations. In a finite-horizon game, we have the following result.

Theorem 1 (Finite-Game One-Stage-Deviation Principle)**.** *Let Γ be a finite-horizon multi-stage game with perfect information. s is a SPNE of Γ if, and only if: there is no player i and strategy s'_i that agrees with s_i except at a single history h, such that: $s'_{i|h}$ is a strictly better response to $s_{-i|h}$ than $s_{i|h}$.*

Proof sketch: the nontrivial direction is to construct a one-shot deviation from a deviation in multiple stages. We take a deviation $\widetilde{s}_i$ and consider the longest history h^t at which the two strategies $\widetilde{s}_i$, s_i differ for the first stage, to construct a one-shot deviation from the equilibrium strategy s_i.

Note that the No Profitable One-Shot Deviation Property is necessary for SPNE, but not for NE. In an infinite-horizon game, we have a similar result, assuming a continuity property of payoffs.

Definition 1. In an infinite-horizon multi-stage game, we say that *payoffs are continuous at infinity* if:

$$\sup_{h,h' \in H^T \text{ s.t. } h^k = h'^k} |u_i(h) - u_i(h')| \to 0 \text{ as } k \to \infty.$$

Payoffs are continuous at infinity whenever: (i) every player i's payoff is the discounted sum of per-period payoffs, and (ii) per-period payoffs are uniformly bounded.

Theorem 2 (Infinite-Horizon-Game One-Stage-Deviation Principle)**.** *Let Γ be an infinite-horizon multi-stage game with perfect information with payoffs continuous at infinity. s is a SPNE if, and only if: there is no player i and strategy s'_i that agrees with s_i except at a single history h, such that: $s'_{i|h}$ is a strictly better response to $s_{-i|h}$ than $s_{i|h}$.*

Proof sketch: the nontrivial direction is to construct a one-shot deviation from a deviation in multiple stages, possibly infinitely many. First, we prove that if $\widetilde{s}_i$ is a

[1] Pietro Dall'Ara, Boston College. Corrections and typos: `pietro.dallara@bc.edu`.

deviation in infinitely many stages, there must be a finite deviation by continuity at infinity. Second, we observe that the finite-game one-stage-deviation principle establishes that: if a strategy profile has the no profitable one-shot deviation property, it is not improvable by finitely many deviations.

The results collected in this handout apply to repeated games in which: (i) the stage game is a simultaneous-move game, and (ii) all past actions are observable. These results are collected in the slide deck on repeated games, with slightly different presentation.

The next result shows that the set of equilibrium payoff vectors weakly expands when a game is repeated.

Theorem 3. *If a is a NE of the stage game of a repeated game Γ, the strategy profile s, where s_i equals a_i at every history, for every player i, is a SPNE of the repeated game Γ.*

Proof Sketch. Note that the future actions of $-i$ is independent of player i's current action.

Let's define minmax values and individual rationality. For a given stage game — of a repeated game —, we use $\mathcal{A}_i$ for the set of mixed actions of player i, and A for the set of pure action profiles, and $g_i : \times_{j\in\{1,\ldots,N\}}\mathcal{A}_j \to \mathbb{R}$ as the per-period expected payoff to player i.

Definition 2. For a given repeated game with observable actions:
(A) Player i's *minmax value* $\underline{v}_i$ is:

$$\underline{v}_i = \min_{\alpha_{-i}\in\times_{j\in -i}\mathcal{A}_j} \max_{\alpha_i\in\mathcal{A}_i} g_i(\alpha_i, \alpha_{-i}).$$

(B) A payoff profile v is *Individually Rational* (IR) if: $v_i \geq \underline{v}_i$, for all i; A payoff vector v is *strictly Individually Rational* (strictly IR) if: $v_i > \underline{v}_i$, for all i.
(C) The set of *feasible average payoff profiles* is:

$$V = \left\{ v \in \mathbb{R}^N : \text{there exists } (\alpha_a)_{a\in A} \in \mathbb{Q}_+^{|A|} \text{ s.t. } v_i = \sum_{a\in A} \alpha_a g_i(a) \text{ for all } i, \text{ and } \sum_{a\in A} \alpha_a = 1 \right\}.$$

Infinitely Repeated Games

The next results are valid for an infinitely repeated game with discounted payoffs: $G^\infty(\delta)$, $\delta \in [0, 1]$. In particular, players maximize the discounted average stage-game payoff. First, NE payoffs are individually rational (IR).

Theorem 4 (Individual Rationality of Nash Equilibria)**.** *If v is the payoff profile of a NE of $G^{\infty}(\delta)$, then v is IR.*

Proof sketch: Player i can use the myopic strategy that prescribes: for each period t, to play a best response to the mixed action of the opponents at t.

Second, the Folk Theorem for NE.

Theorem 5 (Nash Folk Theorem)**.** *Let v be a strictly IR feasible average payoff profile of $G^{\infty}(\delta)$. There exists a discount factor $\underline{\delta} \in [0,1)$ such that: for all $\delta \in (\underline{\delta}, 1)$, there exists a NE of $G^{\infty}(\delta)$ with payoffs v.*

Proof sketch: for simplicity, assume that a pure action profile of the stage game exists such that v is the payoff profile. Use grim-trigger strategies: A *grim-trigger* strategy profile s is a profile of strategies such that any deviation triggers reversion to NE play (hence *trigger*) and the NE minmaxes the players' payoffs (hence *grim*).

Folk Theorem for SPNE:

Theorem 6 (Fudenberg-Maskin (1986) Perfect Folk Theorem)**.** *Let v be a strictly IR feasible average payoff profile of $G^{\infty}(\delta)$, and suppose that the dimension of V is equal to the number of players. There exists a discount factor $\underline{\delta} \in [0,1)$ such that: for all $\delta \in (\underline{\delta}, 1)$, there exists a SPNE of $G^{\infty}(\delta)$ with payoffs v.*

Proof sketch: for simplicity, assume that a pure action profile of the stage game exists such that v is the payoff profile. Use reconciliation phase after a minimax punishment phase, in order to reward punishers. The full-dimension condition is needed to construct strategies that reward players in $-i$, without rewarding the deviator player i.

Finitely Repeated Games

The next results are valid for a finitely repeated game with discounted payoffs: $G^T(\delta)$, $T \in \mathbb{N}$, $\delta \in [0,1]$. In particular, players maximize the discounted average stage-game payoff. The first two results are observations about a game $G^T(\delta)$ with a unique stage-game NE. In particular, SPNE offers a sharp prediction — unique SPNE —, while NE does so only in particular cases — such as in the Prisoner's Dilemma (PD).

Theorem 7 (NE when the stage game has a unique NE)**.**

1. *If s is a NE of the finitely repeated PD, (d,d) is the only action profile that is played on the path of the play according to s, for every $\delta \in [0,1]$.*

2. *There exists a finitely repeated game $G^T(\delta)$ with a unique stage-game NE $a^\star$ such that: $a \neq a^\star$ is played on the path of a NE of $G^T(\delta)$.*

Proof Sketch: induction on the length of terminal histories with positive probability for the first part. Intuition for the second part: PD has minimax payoffs in the unique NE, so one needs to allow the NE to have larger payoffs (v) than minimax ($\underline{v}$), in order to use minimax as off-path "grim" punishment. In this way, a deviator i's continuation play is some average of $\underline{v}_i$'s, while the equilibrium play implies some average of v_i, for periods close to T (by backward induction). See the Extra section for my complete proof.

The first part holds more generally. Suppose players have a unique dominant action in the stage game...

Theorem 8 (SPNE when the stage game has a unique NE)**.** *If $a^\star$ is the unique NE of the stage game of $G^T(\delta)$, and $s^\star$ is a SPNE of $G^T(\delta)$, then: $s_i^\star$ is equal to $a_i^\star$ at every history, for every player i.*

Proof sketch: backward induction.

The next result is related to two-player games when the stage game has multiple NE's, ordered differently between players.

Theorem 9 (Benoit-Krishna, 1985)**.** *Suppose the game $G^T(\delta)$ has two players. Let v' and v'' be stage-game NE payoffs with $v_i' < v_i''$ and $v_j'' < v_j'$. For all v in the feasible average payoff set and greater than or equal to any convex combination of v' and v'', and sufficiently small $\varepsilon > 0$, there exists a T such that: $G^{T'}(1)$ has a SPNE with (average) payoffs within ε of v, for all finite horizons $T' > T$.*

Proof Sketch. For simplicity, assume that a pure action profile a of the stage game exists such that v is the payoff profile. Play a until a sufficiently large cutoff period, after which players alternate between the two NE's. Use the deviator-worst NE as punishment. A sufficiently large cutoff period ensures that the SPNE payoffs are within ε of v, if the horizon T' is large.

The Benoit-Krishna result holds with more players, assuming a full-dimension condition on V. Note that the Benoit-Krishna result extends Nash and perfect folk theorems to a class of finitely repeated games.

Extra

The next result shows that the set of NE continuation payoff vectors is the same in every proper subgame.

Theorem 10. *Let's fix two proper subgames of a repeated game with observable actions, $\Gamma|_{h_1}$ and $\Gamma|_{h_1}$. There exists a one-to-one correspondence between the spaces of strategy profiles, in $\Gamma|_{h_1}$ and in $\Gamma|_{h_2}$, that preserves the payoff to every player.*

Proof: See Exercise 5.3 in Fudenberg and Tirole for a hint of the isomorphism.

We prove Theorem 7 (I came up with the following counterexample for part 2., while most textbooks contain a version of the proof of part 1.). First, we prove 1. We use *d*efect and *c*ooperate as the per-period actions available to each player in the Prisoner's Dilemma.

Proof. Let's show that, if terminal history $h \in H^T$ has positive probability under the NE σ, then $\sigma(h) = (d, d)$. If σ_i puts positive probability on c, then player i has a strictly profitable deviation: mimick σ_i except at h, where i plays d. Note that the deviation is profitable because it puts positive probability on h, by the assumption on σ.

Let's show that, if history $h \in H$ is such that $(h, a) \in H^T$, $a \in A$, and has positive probability under the NE σ, then $\sigma(h) = (d, d)$. Suppose $\sigma_i(h)(c) > 0$, i.e. σ_i puts positive probability on c, and let's construct a strictly profitable deviation to σ_i, given σ_{-i}, which we call σ'_i. i mimicks σ_i, except that he plays d at h. Player i's $T-1$ payoff is strictly larger under σ'_i, given σ_{-i}, than uder σ_i. At period T, player i's payoff depends on her opponent's off-path play, and there are only two possibilities. (1) If $\sigma_{-i}(h, (c, \sigma_{-i}(h))) = d$, player i gets the same payoff at period T under σ'_i, given σ_{-i}, than uder σ_i. (2) If $\sigma_{-i}(h, (c, \sigma_{-i}(h))) = c$, player i gets a strictly larger payoff at period T under σ'_i, given σ_{-i}, than uder σ_i. In each case, the deviation is strictly profitable. Thus, σ is not a NE.

An induction argument completes the proof. □

The next example is used to prove part 2. of the result about NE when the stage game has a unique NE. Let $b > a > 1$. G is the following game.

	C_2	D_2
C_1	$-a, 1$	$-1, 2$
D_1	$-b, -1$	$0, 0$

Intuition. Prisoner's dilemma has minimax payoffs as the unique NE payoff profile. We modify prisoner's dilemma to make a player's minimax payoff lower than NE payoff. Say for player i, $\underline{v}_i$ is less than NE payoff in the unique NE. Then we can construct strategies sustaining a non-NE action profiles where: *off-path* player $-i$ punishes player i by forcing i's minimax payoff.

We prove part 2 using the next claim.

Claim A NE of $G^2(1)$ exists such that: the non-NE action profile (C_1, D_2) is played on-path.

Proof. Consider the following strategy profile s. C_1, D_1 for player 1, and: D_2 if $h = \phi$, D_2 if $h = C_1, D_2$, C_2 otherwise. Note that s_1 is not history contingent at period 2, and note that we can't use one-shot deviation arguments. Let's consider deviations by player 2. She does not have period-2 deviations. If player 2 plays C_2 in period 1, she gets weakly less than 1 in the overall game, while her equilibrium payoff is 2.

Let's consider deviations by player 1. She does not have period-2 deviations. If she plays D_1 in period 1, she gets weakly less than $-a$ in the overall game, while her equilibrium payoff is -1. Therefore, s is a NE of $G^2(1)$. □

[Why is s not a SPNE? Let's observe that the "grim" punishment is not a credible threat to deter 1 from playing D_1. Under s: If player 1 deviates to D_1 (for a period-1 gain of a to 1), player 2 should punish her by playing C_2 and force the period-2 play to be C_1, C_2, which minimaxes player 1's payoff (for a period-2 loss of a to 1). Thus, using "grim" punishment by 2, we sustain the non-NE profile in period 1. However, after 1's deviation — i.e., at history D_1, D_2 —, 2 does not find it profitable to play C_2 — actually, D_2 is strictly dominant for 2 in the stage game! s is not an equilibrium in the subgame that starts at D_1, D_2.]

References

For an in-depth treatment of repeated games, I refer to the relevant chapters in the following texts, in ascending order of difficulty.

- Maschler, Solan and Zamir, Game Theory, 2013. Chapter 13. Chapter 14 includes a simple treatment of Blackwell's approachability theorem for repeated games, which is rarely found in undergradute-level textbooks.

- Fudenberg and Tirole, Game Theory, 1992. Chapters: 4, 5.

- Mailath and Samuelson, Repeated Games and Reputations: Long-Run Relationships, 2006.

- Sorin, Laraki and Renault, Mathematical Foundations of Game Theory, 2019. Chapter 8.

- Mertens, Sorin and Zamir, Repeated Games, 2015.

For a gentle introduction: Gibbons, Game Theory for Applied Economists, 1992, Chapter 2. And, if you speak a little french: Laraki, Renault and Tomala, Théorie des Jeux, 2006.