Repeated Games

Harry R. Lloyd, September 19, 2019

Contents

1 Preliminaries

1.1 Notation

a_i^t / α_i^t	the action / strategy for player i at the tth stage game
a^t / α^t	the action / strategy profile played in the tth stage game
$u_i(\alpha^t)$	player i's utility *in the tth stage game*
h^t	$(\alpha^1, \ldots, \alpha^{t-1})$, the *history* of the game at time t
δ	probability of repetition × the discount factor
T	the number of stage games played in the entire game
n	the number of players

1.2 Payoffs of the Repeated Game

The payoff of history $h^{T+1} = (\alpha^1, \ldots, \alpha^T)$ for player i is given by the PDV of her stage-game utilities:

$$PV_i := \sum_{t=1}^{T} \delta^{t-1} u_i(\alpha^t)$$

It will be convenient to rescale PV_i so as to give us a value that can be compared more easily with the player's stage-game payoffs, *viz.* the average discounted value of player i's utility under history h^{T+1}:

$$ADV_i^{[1,T]} := \frac{1-\delta}{1-\delta^T} \cdot PV_i$$

ADV_i can also be defined recursively:

$$ADV_i^{[1,T]} \equiv (1-\delta) \cdot u_i(\alpha^1) + \delta \cdot ADV_i^{[2,T]}$$

The ADV of 'v for k periods, then $\hat{v}$ forever' is given by:

$$(1-\delta^k)v + \delta^k \hat{v}$$

1.3 Feasibility

$v = (v_1, \ldots, v_n)$ is a feasible average discounted payoff vector of the (T-repeated) stage game iff:

$$v \in \mathrm{co}\{u(a) \,:\, a \in A\}$$

…where $\mathrm{co}\{\cdot\}$ denotes the *convex hull* of the arguments: the *smallest-possible convex containing set,* or equivalently: the set of all possible convex combinations of the arguments

1.4 Minmax

Player i's minmax payoff [in the stage game] is:

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

1.4.1 Procedure #1 for Finding $\underline{v}_i$

1. Divide the set of possible values of α_{-i} into cases, and use this to derive piecewise the function $\max_{\alpha_i}[u_i(\alpha_i, \alpha_{-i})]$.

   - Take for example a case where there are only two players, and where player j only has two actions:
     Let 'mix' abbreviate '$(p, 1-p)$ over A_j'.
     Calculate $u_i(a_i, \text{mix})$ as a function of p for each *action* $a_i \in A_i$.
     Use these functions to derive piecewise $\max_{\alpha_i}[u_i(a_i, \text{mix})]$ as a function of p.

2. Minimise this function to obtain $\underline{v}_1$.

1.4.2 Procedure #2 for Finding $\underline{v}_i$

In some cases, it will be apparent that there is some value $\tilde{u}_1$ of u_1 such that:

1. P1 can *guarantee* herself a payoff of *at least* $\tilde{u}_1$. This $\Rightarrow \underline{v}_1 \geq \tilde{u}_1$.

2. The other players can *hold down* P1 so that she achieves *at most* $\tilde{u}_1$. This $\Rightarrow \underline{v}_1 \leq \tilde{u}_1$

This will allow us to conclude that $\underline{v}_1 = \tilde{u}_1$.

Remark 1.4.2.i The *manner* of the 'holding down' will, of course, be useful for constructing equilibria later one.

Remark 1.4.2.ii In a two-by-two game, if *mixing* is required to minimax Row, then Row must, when being minmaxed, be *indifferent* between her two strategies.

1.5 'Automaton' Diagrams

Automata can be a good way of checking that a strategy profile has been *fully specified*.

2 Types of Punishment Strategy

2.1 'Grim Trigger'

'Grim trigger' in the Prisoners' Dilemma: play C at $t = 1$ and as long thereafter as both players have always played C; if *anyone* (including *oneself*) plays D, then play D forever from then on.

2.2 'Stick, Carrot'

Punish hard for a limited time, then return to cooperation.

2.3 'Stick, Nash'

Punish hard for a limited time, then play a particular Nash equilibrium of the stage game (ad infinitum).

2.4 'Nash Reversion'

Revert to some unsatisfactory Nash of the stage game (often a *mixed* Nash of the stage game) forever.

3 Useful Results for Repeated Games

3.1 [NE] Individual Rationality

Suppose $v = (v_1, \ldots, v_n)$ are the average discounted payoffs for some *Nash equilibrium* of a repeated game. Then $v_i \geq \underline{v}_i$ for $\forall i \in N$.

3.2 [SPE] The One-Shot Deviation Principle

In an **infinitely repeated game** with a discount factor $\delta < 1$, a proposed strategy profile is an *SPE* iff neither player can increase her payoff by changing her action at the start of any subgame in which she is the first-mover, *given* the other players' strategies *and* the rest of her *own* strategy.

(In a **finite horizon game with perfect information,** a proposed strategy profile is an *SPE* iff neither player can increase her payoff by changing her action at the start of any subgame in which she is the first-mover, *given* the other players' strategies *and* the rest of her *own* strategy.)

Take-home: there is no need to check *complicated deviations* when working out whether a strategy profile is an SPE of a repeated game.

3.3 [SPE] Generating 'Coordination with Randomness' in an SPE

One way that we can generate randomness in the selection of one of a set of coordination outcomes within an SPE (we may want to do this to *achieve certain expected payoffs*) is to get one player to *randomise over equivalent strategies* in an early-period stage game, and then to condition later actions on the randomised action played earlier. **Example:** 2011 finals paper, Q5(d) [repeated as problem 2 in the 2017 version of the Repeated Games problem sheet].

Another option is to use an observed random variable as public randomization device (as a 'sun spot'). **Example:** 2016 finals paper, Q7(f).

4 Renegotiation-Proof SPE

4.1 Weak Renegotiation-Proofness

- An SPE consists of a (subgame perfect) 'continuation' after every play – call the collection of these continuations the 'book of plays'.
- An SPE is *weakly renegotiation-proof* iff no continuation play in its book of plays is Pareto dominated by any other.

4.2 Gloss

Why is this only a 'weak' concept of renegotiation proofness? Because we're comparing *within* the playbook of a given SPE, rather than across different SPE.

5 Folk Theorems

	Repetitions $T = \infty$	Repetitions $T \in \mathbb{N}$
NE	Folk Theorem, §5.1	
SPE	Fudenberg & Maskin's Perfect Folk Theorem, §5.2	*Benoît & Krishna's Finite-Repetition Folk Theorem, §5.4*
Renegotiation-Proof SPE	*Folk Theorem for Renegotiation-Proof SPE §5.3*	

Key: theorems in *italics* hold only when $n = 2$; theorems not in italics hold for all $n \in \mathbb{N}$.

5.1 Folk Theorem

If v is feasible,
and $v_i \geq \underline{v}_i$ for $\forall i \in N$,
then $\exists \underline{\delta} < 1$ such that for $\forall \delta \geq \underline{\delta}$,
the ∞-repeated game
has an NE in which $ADV_i = v_i$ for $\forall i \in N$.

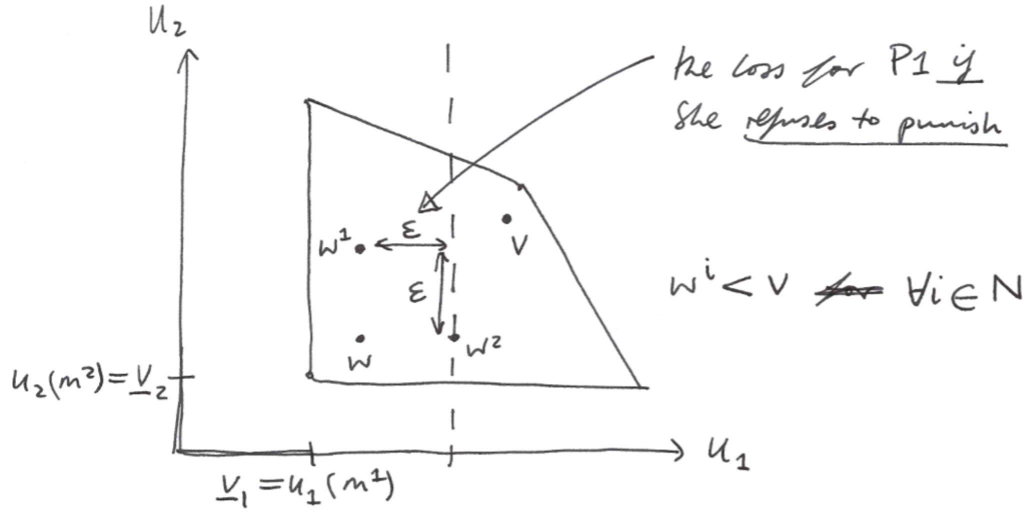
5.2 Fudenberg & Maskin's Perfect Folk Theorem

5.2.1 Theorem

[Assume that the set of feasible and strictly individually rational payoffs is n-dimensional.]

If v is feasible,
and $v_i > \underline{v}_i$ for $\forall i \in N$,
then $\exists \underline{\delta} < 1$ such that for $\forall \delta \geq \underline{\delta}$,
the ∞-repeated game
has an SPE in which $ADV_i = v_i$ for $\forall i \in N$.

5.2.2 Procedure for Constructing the SPE



(I) Collaboration: Play the stage game actions that generate v. Repeat unless player i deviates, in which case go to phase II^i.

(II^i) Punishment: Play m^i (i.e.: minmax i) for exactly τ periods [for particular τ]. If no-one deviates, then go to phase III^i. If player k deviates, then go to the start of phase II^k.

(III^i) Reconciliation: Play the stage-game actions that generate w^i. (In phase III^i, those who *did the punishing* get *rewarded*; those who *deviated* do *not*.) Repeat unless player k deviates: in that case, go to phase II^k.

Question: Is this structure really plausible? If player k *fails to punish* player i, then player i will respond by *punishing player k*! ...this is certainly not *renegotiation proof*.

5.3 Folk Theorem for Renegotiation-Proof SPE

5.3.1 Thereom

Assume $n = 2$,
and let v be a feasible payoff pair
such that $v_i > \underline{v}_i$ for $\forall i \in \{1, 2\}$.
If $\exists a', a'' \in A$ such that:
(1) $v_1 > u_1(BR_1(a'_2), a'_2)$ and $v_2 < u_2(a'_1, a'_2)$, and

(2) $v_1 < u_1(a''_1, a''_2)$ and $v_2 > u_2(a''_1, BR_2(a''_1))$,
then $\exists \underline{\delta} < 1$ such that for $\forall \delta \geq \underline{\delta}$,
the ∞-repeated game
has a renegotiation-proof SPE in which $ADV_i = v_i$ for $\forall i \in \{1, 2\}$.

5.3.2 Gloss

(?) a' can be used to punish P1 whilst rewarding P2 for doing so; a'' can be used to punish P2 whilst rewarding P1 for doing so; $u(a')$, $u(a'')$ and v are Pareto unranked.

5.3.3 Converse Theorem

If an ∞-repeated game has a renegotiation-proof SPE in which $ADV_i = v_i$ for $\forall i \in \{1, 2\}$, then $\exists a', a'' \in A$ satisfying (1) and (2) *with weak inequalities*.

5.3.4 Procedure for Constructing a Renegotiation-Proof SPE

- Punishment must be followed by forgiveness (i.e. a return to cooperation).
- But, the punishment phase must *reward the punisher* before cooperation can resume.

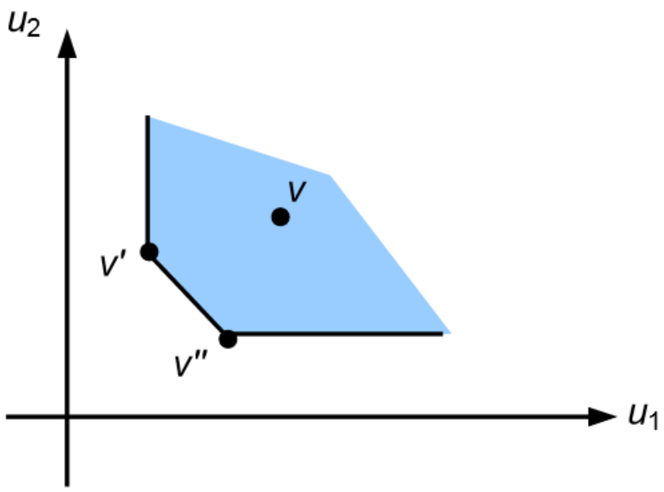
5.4 Benoît & Krishna's Finite-Repetition Folk Theorem

5.4.1 Theorem

Assume $n = 2$,
and $\delta = 1$.
Let v' and v'' be *Pareto non-ranked, NE payoff pairs* in the stage game.
If v is a feasible payoff vector
such that $v_i \geq v'_i, v''_i$ for $\forall i \in \{1, 2\}$,
and $v \geq \lambda v' + (1 - \lambda)v''$ for $\forall \lambda \in [0, 1]$,
then for some sufficiently large T,
v can be approximated arbitrarily closely by the average payoff of an SPE of the T-repeated game.

Summary of the Lesson: Finite repetition only expands the set of ADVs sustainable in an equilibrium if there are *multiple Nash equilibria* of the stage game, that the players *rank differently*; *Nash reversion* is the punishment style used.

5.4.2 Diagram



5.4.3 A Tip for Constructing the SPE

Keep things simple! Use *cycling* over pure NE, rather than mixed-strategy NE, wherever possible.

6 Tips for the Examination

In constructing punishment phases, be guided by hints in earlier parts of the question. It's sensible to try Nash reversion first; stick-Nash is also often useful.

Note that if in the cooperation phase of our proposed SPE *player i is best-responding to player j, even if only weakly,* then we *don't need to construct any punishment phase for player i*.