Repeated games: 14.126

Sergei Izmalkov, Muhamet Yildiz
MIT

1 Introduction

1.1 Ideas

- repeated interaction, cooperation, social norms
- infinite horizon, uncertain terminal date
- strategic effects of own actions, threats
- punishment, revenge
- detection, monitoring
- reputation, collusion
- *learning

1.2 One-stage Deviation Principle

Theorem: Finite multi-stage game, observed actions.

Profile s is SPE $\iff$ for each player i no profitable one-stage deviation exists.

Proof: Suppose one-stage deviations are not profitable. Suppose a profitable deviation exists for some player i at some subgame, that is

$$\exists i, t, h^t, \hat{s}_i \quad (\hat{s}_i, \mathrm{s}_{-i})\,|_{h^t} \succ_i (s_i, \mathrm{s}_{-i})\,|_{h^t}.$$

Then $t^* = \max t'$ that $\hat{s}_i(h^{t^*}) \neq s_i(h^{t^*})$. Obviously $t^* > t$. One-stage deviation implies $(\hat{s}_i, \mathrm{s}_{-i})\,|_{h^{t^*}} \approx_i (s_i, \mathrm{s}_{-i})\,|_{h^{t^*}}$.

Define $\tilde{s}_i(h^\tau) = \hat{s}_i(h^\tau)$ for all $\tau < t^*$, $\tilde{s}_i(h^\tau) = s_i(h^\tau)$ for all $\tau \geq t^*$.

Thus, $(\hat{s}_i, \mathrm{s}_{-i})\,|_{h^\tau} \approx_i (\tilde{s}_i, \mathrm{s}_{-i})\,|_{h^\tau}$ for all h^τ.

Repeat for $\tilde{s}_i$ $(t^* \downarrow)$.

Definition: Game is *continuous at infinity* if for any i,

$$\sup_{h, \tilde{h}, \text{ s.t. } h^t = \tilde{h}^t} \left| u_i(h) - u_i(\tilde{h}) \right| \to 0 \text{ as } t \to 0.$$

Example: Discounting + bounded stage payoffs.

Theorem: One-stage deviation principle for games with continuous payoffs.

Proof: Let ε be the size of improvement. Cut the end that matters less than $\varepsilon/2$. The rest is the finite game. No improvement possible.

Remark: If payoffs defined differently, e.g. the average over time, the principle need not to hold.

1.3 Examples

Prisoner's Dilemma

	C	D
C	$1,1$	$-1,2$
D	$2,-1$	$0,0$

	L	M	R
U	$0,0$	$3,4$	$6,0$
M	$4,3$	$0,0$	$0,0$
D	$0,6$	$0,0$	$5,5$

2 Repeated Games with Observable Actions

2.1 The Model

Stage game G: $\mathcal{I}$–players, A_i–actions, $g_i : A = \times_{i\in\mathcal{I}} A_i \to \mathbb{R}$–payoffs, $\mathcal{A}_i$–probability distributions over A_i.

Repeated game: $\mathbf{a}^t \equiv (a_i^t)_{i\in\mathcal{I}}$, h^0–null history, $h^t = (\mathbf{a}^0, \mathbf{a}^1, \ldots, \mathbf{a}^{t-1})$, $H^t = (A)^t$.

Strategy (pure) $s_i \equiv (s_i^t)$, $s_i^t : H^t \to A_i$; (mixed) $\sigma_i \equiv (\sigma_i^t)$, $\sigma_i^t : H^t \to \mathcal{A}_i$.

Payoffs:

- Discounting: $G(\delta)$,

$$u_i = E_{\boldsymbol{\sigma}}(1-\delta)\sum_{t=0}^{\infty}\delta^t g_i(\sigma^t(h^t)) \to \max\,.$$

- Time-averaging criterion

$$\liminf_{T\to\infty} E\frac{1}{T}\sum_{t=0}^{T} g_i(\sigma^t(h^t)) \to \max\,.$$

- Overtaking criterion

$$g = (g^0, g^1, \ldots) \precsim_i h = (h^0, h^1, \ldots)$$
$$\text{if } \exists T', \forall T > T', \quad \sum_{t=0}^{T} g^t \le \sum_{t=0}^{T} h^t.$$

Observation: If α^* is a Nash eqm in G, then "play α^* (α_i^* for each i) for all t" is SPE. If m static equilibria exist, any combination of them is SPE.

2.2 Folk Theorems

Feasible, individually rational payoffs.

Reservation utility (minimax):

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

Define $\mathbf{m}^i$ to be a minimax profile.

Observation: In any static Nash eqm or Nash eqm of repeated game, i's payoff is not lower than $\underline{v}_i$.

Proof: Play $a_i(h^t)$ to maximize $Eg_i(a_i, \sigma_{-i}(h^t))$, where σ is Nash strategies.

Feasible payoffs (with randomization):

$$V = \text{c.h.}\left\{\mathbf{v} = g(\mathbf{a}),\ \text{for } \mathbf{a} \in A\right\}.$$

Theorem: (folk theorem) For any $\mathbf{v} \in V$, with $v_i > \underline{v}_i$ for all i, there exists a $\delta^* < 1$, such that for all $\delta \in (\delta^*, 1)$ there exist Nash eqm with payoffs $\mathbf{v}$.

Proof: Punish by minimax.

Theorem: (Friedman, Nash-threats) α^* is a static Nash with payoffs $\mathbf{e}$. Then for any $\mathbf{v} \in V$, with $v_i > e_i$ for all i , there exists a $\delta^* < 1$, such that for all $\delta \in (\delta^*, 1)$ there exist SPE of $G(\delta)$ with payoffs $\mathbf{v}$.

Proof: Punish by Nash. SPE follows from above observation.

Theorem: (Aumann, Shapley) Time-average criterion, then or any $\mathbf{v} \in V$, with $v_i > \underline{v}_i$ for all i , there exists a SPE of $G(\delta)$ with payoffs $\mathbf{v}$.

Proof: Punish by minimax for a limited time. Long-Run effects are zero.

Theorem: (Fudenberg, Maskin) Suppose $\dim V = \#\mathcal{I}$. Then for any $\mathbf{v} \in V$, with $v_i > \underline{v}_i$ for all i , there exists a $\delta^* < 1$, such that for all $\delta \in (\delta^*, 1)$ there exist SPE of $G(\delta)$ with payoffs $\mathbf{v}$.

Proof: Idea is to reward punishers. Suppose for all considered $\mathbf{v}$, there exists $\mathbf{a}$, $g(\mathbf{a}) = \mathbf{v}$. Since $\dim V = \#\mathcal{I}$, $\exists \mathbf{v}' \in V$, $\underline{v}_i < v'_i < v_i$ for all i, and $\mathbf{v}'(i) \in V$, that

$$\mathbf{v}'(i) = \left(v'_1 + \varepsilon, \ldots, v'_{i-1} + \varepsilon, v'_i, v'_{i+1} + \varepsilon, \ldots, v'_I + \varepsilon\right).$$

Suppose $\mathbf{a}'(i)$ exist that $g(\mathbf{a}'(i)) = \mathbf{v}'(i)$.

Phase 1. Play $\mathbf{a}$ until realized action is $\mathbf{a}$ or differs from $\mathbf{a}$ in ≥ 2 components. If $a'_j \neq a_j$, switch to Phase 2_j.

Phase 2_j. Play $\mathbf{m}^j$ for N periods as long as realized action is $\mathbf{m}^j$ or differs from $\mathbf{m}^j$ in ≥ 2 components. Switch to Phase 3_j. If some k deviates switch to Phase 2_k.

Phase 3_j. Play $\mathbf{v}'(j)$ forever as long as realized action is $\mathbf{a}'(j)$ or differs from $\mathbf{a}'(j)$ in ≥ 2 components. If bidder k deviates switch to Phase 3_k.

Use one-time deviation principle.

Problem: If $\mathbf{a}'(i)$ is mixed, the same continuation payoff has to be guaranteed for all actions in support.

Theorem: (Abreu, Dutta, Smith) NEU condition instead of $\dim V = \#\mathcal{I}$.

Definition: *NEU* (non-equivalent utilities) is satisfied if for any (i,j), $\nexists c, d \in \mathbb{R}_+$ that $g_i(\mathbf{a}) = c + dg_j(\mathbf{a})$ for all $\mathbf{a} \in A$.

Proof: NEU $\Longrightarrow \exists \left[\mathbf{v}^1, \ldots, \mathbf{v}^I\right]$, such that $\forall i, j$, $v_i^i < v_i^j$.

Roughly: Substitute $\mathbf{v}^i$ in place of $\mathbf{v}'(i)$.

2.3 Finite Games

Theorem: (Benoit, Krishna) Time-averaging criterion. Suppose $\forall i$ exists static Nash $\alpha^*(i)$ with $g_i(\alpha^*(i)) > \underline{v}_i$. Then the set of Nash eqm payoffs of the G^T converges as $T \to \infty$ to the set of feasible, IR payoffs of G^∞.

Proof: Terminal reward phase. $R \times I$ cycle:

$([\alpha^*(1), \ldots, \alpha^*(I)])^R$—Nash-eqm path.

Gives strictly more than $\underline{v}_i$ to each i. For large R the threat of minimaxing over RI periods prevents all deviations.

Fix $\varepsilon > 0$. Exists T, such that payoff over $T - RI$ periods approximates v_i for all i within ε. ...

2.4 Varying opponents

2.4.1 Short-Run vs Long-Run players

If Short-Run players move first, "cooperation" is attainable.

Principle: S-R player(s) plays C, L-R player(s) responds C as long as (C, C) was played in the past. Otherwise D.

Simultaneous moves: S-R player always plays BR.

$1, \ldots, l$ – L-R players,

$l+1, \ldots, I$ – S-R players,

$B : \times_{i=1}^{l} \mathcal{A}_i \to \times_{j=l+1}^{I} \mathcal{A}_j$ – BR correspondence.

$$\underline{\mathsf{v}}_i = \min_{\alpha \in \mathrm{graph}(B)} \left[\max_{a_i} g_i(a_i, \alpha_{-i}) \right],$$

$$V = \text{c.h.} \left\{ \mathbf{v} = (g_i(\mathbf{a}))_{i=1}^{l} \in \mathbb{R}^l, \text{ for } \alpha \in \mathrm{graph}(B) \right\}.$$

Observability of mixed actions is important. Long-Run players have to be indifferent between the pure actions they assign positive probabilities.

$$\bar{\mathsf{v}}_i = \max_{\alpha \in \mathrm{graph}(B)} \left[\min_{a_i \in \mathrm{supp}(\alpha_i)} g_i(a_i, \alpha_{-i}) \right].$$

Theorem: (Fudenberg, Kreps, Maskin).

Suppose $\dim V = l$.

For any $\mathbf{v} \in V$, with $\bar{\mathsf{v}}_i > v_i > \underline{\mathsf{v}}_i$ for all i , there exists a $\delta^* < 1$, such that for all $\delta \in (\delta^*, 1)$ there exist SPE of $G(\delta)$ with payoffs $\mathbf{v}$.

2.4.2 Overlapping generations

Players live for T periods. Every generation has the same mass.

Actions are observable: $work$ or $shirk$ (IR,static NE). All $work$ is efficient.

Payoffs are averages over lifetime.

Result (Crémer): Nash eqm exists where all except the oldest $work$.

Folk theorems: Candori, Smith.

2.4.3 Random matching

What is observable? What is remembered? Public vs Private information.

Prisoner's dilemma: Play C as long as (C, C) was played. D otherwise.

Supportable as long as δ is high enough and some info about opponent is known.

If only past private outcomes are observable, with N high enough, "contagion" strategies may not be an equilibrium.

Reason: Responding C on D slows contagion.

2.5 Pareto-Perfection

$$Eff(C) = \{x \in C, \nexists y \in C, y \geqq x, y \neq x\}.$$

Definition: (Bernheim, Peleg, Whinston) Consider G^T, P^T is the set of pure-strategy SPE payoffs of G^T. $Q^1 = P^1$, $R^1 = Eff(P^1)$.

For $T > 1$, $Q^T \subseteq P^T$—the set of pure-strategy perfect equilibrium payoffs enforced by R^{T-1}. Let $R^T = Eff(Q^T)$.

SPE σ is *Pareto-Perfect* if, $\forall t$ and $\forall h^t$, continuation payoffs implied by σ are in R^{T-t}.

Example: (Benoit, Krishna) $\delta = 1$.

	b_1	b_2	b_3	b_4
a_1	0, 0	2, 4	0, 0	5.5, 0
a_2	4, 2	0, 0	0, 0	0, 0
a_3	0, 0	0, 0	3, 3	0, 0
a_4	0, 5.5	0, 0	0, 0	5, 5

3 Repeated Games with Imperfect Public Information

3.1 The Model

$\mathbf{a} \in A \to \Delta(y)$, $y \in Y$—publicly observable.

$\pi_y(a) \in \Delta(y)$; $\pi(a)$

$r_i(a_i, y)$—payoff to i, (!) independent of a_{-i}.

$g_i(a) = \sum_y \pi_y(a) r_i(a_i, y)$.

$h^t = (y^0, y^1, \ldots, y^{t-1})$—public history.

z_i^t—private history (past actions).

Strategy (mixed) $\sigma_i \equiv (\sigma_i^t)$, $\sigma_i^t : H^t \times Z_i^t \to \mathcal{A}_i$.

Definition: σ_i is a *public strategy* if $\sigma_i(h^t, z_i^t) = \sigma_i(h^t, \tilde{z}_i^t)\ \forall t, h^t, z_i^t, \tilde{z}_i^t$.

Observation: Pure-Strategy eqm payoff can be supported as a payoff of an equilibrium in Public strategies.

Definition: σ is a *perfect public equilibrium* if for all i, σ_i is a public strategy, and $\forall t, h^t$, strategies $\sigma|_{h^t}$ form Nash eqm.