Mixed-Strategy Subgame-Perfect Equilibria in Repeated Games

Kimmo Berg
Department of Mathematics and Systems Analysis
Aalto University, Finland
(joint with Gijs Schoenmakers)

July 8, 2014

Outline of the presentation

- Illustrative example
  - Shows how players may randomize in repeated games
  - Convert into various normal-form games by using different continuation payoffs
- Abreu-Pierce-Stacchetti fixed-point characterization
  - Extension to behavior strategies
- Self-supporting sets to find equilibria in behavior strategies
- Comparison between pure, behavior and correlated strategies

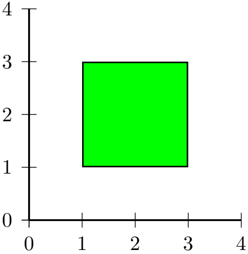
The model

- Infinitely repeated game
- Stage game with finitely many actions
- Discounting (possibly unequal discount factors)
- Behavior strategies (randomization and history-dependent)
- Players observe realized pure actions (not randomizations)

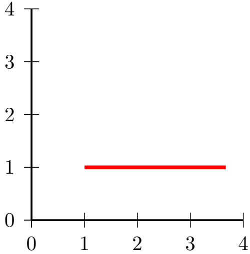
The model (2)

- Finite set of players $N = \{1, \dots, n\}$
- Finite set of pure actions A_i, $i \in N$, $A = \times_{i \in N} A_i$
- Mixed action $q_i(a_i) \geq 0$, profile $q = (q_1, \dots, q_N)$
- Probability of pure action profile $a \in A$: $\pi_q(a) = \prod_{j \in N} q_j(a_j)$
- Stage game payoff $u_i(q) = \sum_{a \in A} u_i(a) \pi_q(a)$
- Histories $H^k = A^k$ for stage $k \geq 0$, $H^0 = \varnothing$
- Behavior strategy $\sigma_i : H \mapsto Q_i$
- Discounted payoff $U_i(\sigma) = \mathbb{E}\left[(1 - \delta_i) \sum_{k=0}^{\infty} \delta_i^k u_i^k(\sigma)\right]$

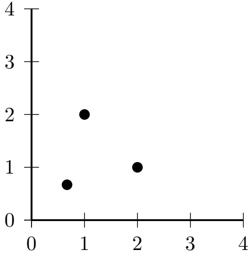
Payoffs from stage games



3,3	1,3
3,1	1,1



7/3,7/3	1/3,3
11/3,1	1,1



0,0	2,1
1,2	0,0

Prisoner's Dilemma

3,3 (a)	0,4 (b)
4,0 (c)	1,1 (d)

- What are equilibria in pure, behavior and correlated strategies?
- Common discount factor $\delta = 1/3$
- The pure action profiles are called a, b, c and d

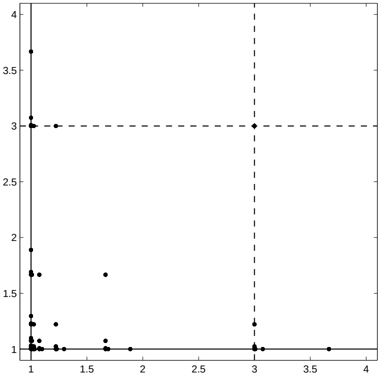
Prisoner's Dilemma (2)

3,3	1/3,3
3,1/3	5/3,5/3

7/3,7/3	1/3,3
3,1/3	1,1

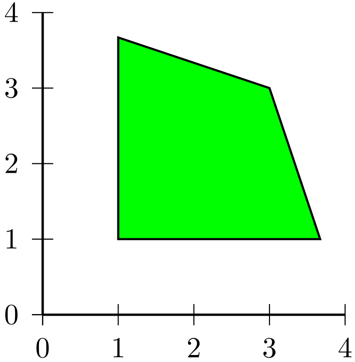
- Left: No unilateral deviation, a and d followed by cooperation, b and c by punishment
- Right: d^{∞} after all pure action profiles

Prisoner's Dilemma: Pure strategies



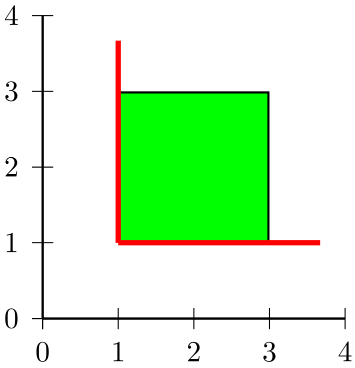
- Berg and Kitti (2010): elementary subpaths d,aa,ba,bc,ca,cb
- Equilibrium paths are compositions of the elementary subpaths, e.g., $d^7(bc)^3a^\infty$

Prisoner's Dilemma: Correlated strategies



- All reasonable (feasible and individually rational) payoffs

Prisoner's Dilemma: Behavior strategies



- Union of rectangle $(1,3) \times (1,3)$ and two lines
- How do we get these payoffs?

Prisoner's Dilemma: Behavior strategies (2)

3,3	0,4
4,0	1,1

$\Rightarrow$

7/3,7/3	1/3,3
11/3,1	1,1

- Find follow-up strategies and continuation payoffs so that payoffs correspond to the game on right
- Action profiles a, b and d are followed by d^∞ (SPEP) and c is followed by a^∞ (SPEP)
- ad^∞: $(1-\delta)(3,3)+\delta(1,1)=(7/3,7/3)$
- ca^∞: $(1-\delta)(4,0)+\delta(3,3)=(11/3,1)$
- Produces the red lines of payoffs

Prisoner's Dilemma: Behavior strategies (3)

3,3	0,4
4,0	1,1

$\Rightarrow$

3,3	1,3
3,1	1,1

- Find continuation payoffs: a $(3,3)$, b $(3,1)$, c $(1,3)$, d $(1,1)$
- $(1-\delta)(0,4)+\delta(3,1)=(1,3)$
- a is followed by a^∞, d is followed by d^∞
- b is followed by $(cb)^\infty$:
  $(1-\delta)(1-\delta^2)^{-1}\left[(4,0)+\delta(0,4)\right]=(3,1)$
- No randomization needed (not as easy in general!)
- Produces the green rectangle of payoffs

Characterization of Equilibria à la APS

- Carrier of mixed action $Car(q_i) = \{a_i \in A_i | q_i(a_i) > 0\}$
- Most profitable deviation $d_i(q) = \max_{a_i' \in A_i \setminus Car(q_i)} u_i(a_i', q_{-i})$.
- Smallest payoff from a set $p_i(W) = \min\{w_i,\ w \in W\}$
- A pair (q, w) is admissible with respect to $(w \in)W$ if

$$(1-\delta)u_i(q) + \delta w_i \geq (1-\delta)d_i(q) + \delta p_i(W)$$

- Each $a \in Car(q)$ may follow by different continuation play
- Continuation payoff $w = \sum_{a \in Car(q)} x(a)\pi_q(a)$, $x(a) \in W$

Characterization (2)

- Stage game payoffs $\tilde{u}_\delta(a) \doteq (1-\delta)u(a) + \delta x(a)$
- Set of all equilibrium payoffs $M(x)$ of stage game with $\tilde{u}$
- V is the set of subgame-perfect equilibrium payoffs

Theorem

V is the largest fixed point of B:

$$W = B(W) = \bigcup_{x \in W^{|A|}} M(x),$$

where (q, w) admissible, w formed by x, and q equilibrium of stage game with payoffs x.

Comparison to Pure Strategies

- V^P is the set of pure-strategy subgame-perfect equilibrium payoffs

Theorem (Abreu-Pearce-Stacchetti 1986/1990)

V^P is the largest fixed point of B^P:

$$W = B^P(W) = \bigcup_{a \in A} \bigcup_{w \in C_a(W)} (1-\delta)u(a) + \delta w,$$

where $C_a(W) = \{w \in W \ s.t. \ (a, w) \ admissible\}$.

Comparison to Pure Strategies (2)

- Complexity of fixed-point is higher
- Structure of equilibria different
- In pure strategies, enough to have high enough continuation payoff
- Randomization requires exact continuation payoffs

Self-supporting sets

Definition

S is self-supporting set if $S \subseteq M(x)$ for $x \in \mathbb{R}^{|A|}$ and

- $x(a) \in S$ for $a \in Car(q(s))$,
- if player i plays an action $\tilde{a}_i$ outside $Car(q(s)_i)$ (an observable deviation), while $a_{-i} \in Car(q(s)_{-i})$, then $x_i(\tilde{a}_i, a_{-i})$ is player i's punishment payoff.
- if at least two players make an observable deviation, then the continuation payoff is a predetermined equilibrium payoff.

- Strongly self-supporting if $x(a) \in S$ for all $a \in A$

Self-supporting sets (2)

- Required continuation payoffs are within the set itself
- Easy way to produce (subsets of) equilibrium payoffs

Theorem (Monotonicity in δ)

If S is self-supporting set for δ,

- *S is convex,*
- *$\tilde{u}_\delta(a) = (1-\delta)u(a) + \delta x(a) \in S$ for all $a \in Car(q(s))$, and*
- *$p_i(V(\delta))$ is not increasing in δ for all $i \in N$.*

Then there exists a self-supporting set $S' \supseteq S$ for $\delta' > \delta$.

Results: Prisoner's Dilemma

$\mathfrak{a}, \mathfrak{a}$	$\mathfrak{b}, \mathfrak{c}$
$\mathfrak{c}, \mathfrak{b}$	$\mathfrak{d}, \mathfrak{d}$

with $\mathfrak{c} > \mathfrak{a} > \mathfrak{d} > \mathfrak{b}$

Theorem

The rectangle $[\mathfrak{d}, \mathfrak{a}] \times [\mathfrak{d}, \mathfrak{a}]$ *is a subset of the subgame-perfect equilibrium payoffs for*

$$\delta \geq \max \left[\frac{\mathfrak{c} - \mathfrak{a}}{\mathfrak{c} - \mathfrak{d}}, \frac{\mathfrak{d} - \mathfrak{b}}{\mathfrak{a} - \mathfrak{b}} \right].$$

Results: Nonmonotonicity

Theorem (Nonmonotonicity of payoffs)

The set of subgame-perfect equilibrium payoffs are not monotone in the discount factor in the following symmetric game:

$3, 3$	$-\frac{1}{10}, 4$	$-10, -10$	$1, -10$
$4, -\frac{1}{10}$	$1, 1$	$-10, -10$	$-10, -10$
$-10, 1$	$-10, -10$	$\frac{43}{10}, -\frac{1}{10}$	$-10, -10$
$-10, -10$	$-10, -10$	$-10, -10$	$-\frac{1}{10}, \frac{43}{10}$

- $[1, 3] \times [1, 3]$ is a subset of the subgame-perfect equilibrium payoffs when $\delta = 1/3$ but not for a higher discount factor
- Rectangle gets contracted and relies on outside payoff

Results: Comparison of pure, mixed and correlated

- Feasible payoffs $V^{\dagger} = co\left(v \in \mathbb{R}^n : \exists q \in A \text{ s.t. } v = u(q)\right)$
- Reasonable payoffs $V^*(\delta) = \left\{v \in V^{\dagger},\ v_i \geq p_i(V(\delta)),\ i \in N\right\}$
- Critical discount factor

$$\delta^M = \inf\left\{\delta : V(\delta') = V^*(\delta'), \forall \delta' \geq \delta\right\}$$

Theorem

For all δ, $V^P(\delta) \subseteq V^M(\delta) \subseteq V^C(\delta)$.

Theorem

If $p^P(V^P(\delta')) = p(V(\delta')) = p^C(V^C(\delta'))$ for all $\delta' \geq \min\left[\delta^P, \delta^M, \delta^C\right]$, then it holds that $\delta^P \geq \delta^M \geq \delta^C$.

Results: Comparison in Prisoner's Dilemma

Theorem

In symmetric Prisoner's Dilemma, it holds that

$$\delta^P = \delta^M = \frac{\mathfrak{c} - \mathfrak{b}}{\mathfrak{a} + \mathfrak{c} - \mathfrak{b} - \mathfrak{d}} > \max\left[\frac{\mathfrak{c} - \mathfrak{a}}{\mathfrak{c} - \mathfrak{d}}, \frac{\mathfrak{d} - \mathfrak{b}}{\mathfrak{a} - \mathfrak{b}}\right] = \delta^C,$$

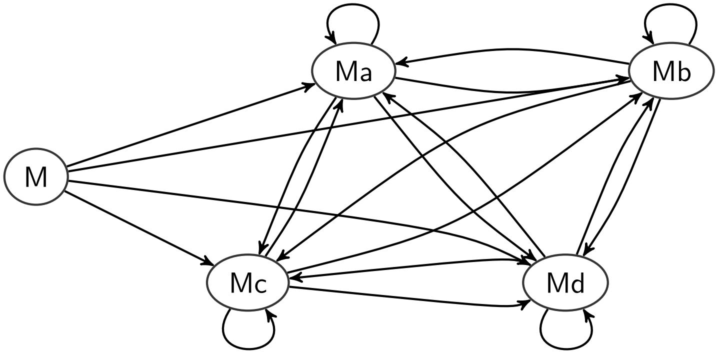
when $\mathfrak{b} + \mathfrak{c} < 2\mathfrak{a}$, *and otherwise*

$$\delta^P = \frac{2(\mathfrak{c} - \mathfrak{d})}{\mathfrak{b} + 3\mathfrak{c} - 4\mathfrak{d}} > \delta^M = \frac{\mathfrak{c} - \mathfrak{b}}{2(\mathfrak{c} - \mathfrak{d})} > \frac{\mathfrak{d} - \mathfrak{b}}{\mathfrak{c} - \mathfrak{d}} = \delta^C,$$

Conclusion

- Characterization of equilibria in behavior strategies
- Self-supporting sets offer easy way to find behavior strategies
- It is possible to compare equilibria under different assumptions
- Open problem: punishment strategies in pure and behavior strategies

That's all folks...



Thank you! Any questions?