Rational Learning Leads to Nash Equilibrium

Ehud Kalai, Ehud Lehrer

Econometrica 61(1993)

December 4, 2007

Introduction

- If each player correctly predicts the opponent's strategies and if each chooses the optimal strategy given own beliefs then the strategies form a Nash equilibrium of a repeated game.

- But under what circumstances will rational players actually learn to predict the behavior of others starting from *out-of-equilibrium* conditions?

Outline

- The Game.
- Statement of the result.
- "Intuition".
- Applicability and discussion of assumptions.
- Caveat: Impossibility Theorem by Foster and Young (2001).
- Conclusions.

The Repeated Game I: General Structure and Notation

The stage game is described by the following components:

- n finite sets Σ_1, Σ_2,..., Σ_n of actions with $\Sigma = \times_{i=1}^{n} \Sigma_i$ denoting the set of action combinations;
- n payoff functions $u_i : \Sigma \to R$;
- H_t denotes the set of histories of length t, $t = 0, 1, 2, \ldots$and $\overline{H} = \cup_t H_t$ the set of all histories;
- A *strategy* of player i is a function $f_i : \overline{H} \to \Delta(\Sigma_i)$ with $\Delta(\Sigma_i)$ denoting the set of probability distributions on Σ_i;
- Let $f = (f_1, \ldots, f_n)$ be a vector of behavior strategies and $z^t = (z_1^t(f_1), \ldots, z_n^t(f_n))$ be a vector of realized action combinations. Then the infinite vector $(z^1, z^2, \ldots)$ is the realized play path;
- Finally, induced by a strategy vector f on the set of infinite play paths is the probability distribution μ_f.

The Repeated Game II: The Payoffs and Beliefs

- Let $x_i^t = u(z^t)$ denote player i's payoff at stage t. Then the payoff to player i in the repeated game is defined by

$$U_i(f) = (1-\lambda_i)\sum_{t=0}^{\infty}\lambda_i^t E_f\left(x_i^{t+1}\right) = (1-\lambda_i)\int\left[\sum x_i^{t+1}\lambda_i^t\right] d\mu_f$$

- Let f_j^i denote the behavior strategy that player i thinks player j will follow. Then a strategy f_i is an ϵ-*best response* to $f_{-i}^i = \left(f_1^i, ..., f_{i-1}^i, f_{i+1}^i, f_n^i\right)$ if $U_i\left(\overline{f}_i, f_{-i}^i\right) - U_i\left(f_i, f_{-i}^i\right) \leq \epsilon\ \forall\epsilon > 0$.

The Repeated Game III: Assumptions

1. Knowledge of own payoff matrices.

2. Perfect monitoring.

3. Independence of strategies and beliefs.

4. Learning takes place through Bayesian updating of the individual prior beliefs.

5. Compatibility of beliefs with the truth i.e. with the "truly chosen" strategies (*absolute continuity*).

The Repeated Game IV: Definitions

Definition 1. Let $\epsilon > 0$ and let μ and $\widetilde{\mu}$ be two probability measures defined on the same space. We say that μ **is ϵ-close to** $\widetilde{\mu}$ if there is a measurable set Q such that:
(i) $\mu(Q)$ and $\widetilde{\mu}(Q)$ are greater than $1-\epsilon$;
(ii) for every measurable set $A \subseteq Q$

$$(1-\epsilon)\,\widetilde{\mu}(A) \leq \mu(A) \leq (1+\epsilon)\,\widetilde{\mu}(A)$$

Remark: Compare with $|\mu(A) - \widetilde{\mu}(A)| \leq \epsilon$.

Definition 2. Let $\epsilon > 0$. We say that f **plays ϵ-like** g if μ_f is ϵ-close to μ_g.

Definition 3. An n vector of strategies, g, is **a subjective ϵ-equilibrium** if there is a matrix of strategies $\left(g_j^i\right)_{1\leq i,j\leq n}$ with $g_i^i = g^i$ such that
(i) g_i is a best response to g_{-i}^i and;
(ii) g plays ϵ-like g^i.

Statement of the Main Result

Theorem 1. Let f and f^i be two n-vectors of strategies, representing the ones actually chosen and the beliefs of player i, respectively. Assume that f is absolutely continuous w.r.t f^i. Then for every $\epsilon > 0$ and for almost every play path z there is a time $T(z, \epsilon)$ such that for all $t \geq T$, $f_{z(t)}$ plays ϵ-like $f^i_{z(t)}$.

Theorem 2. Let f and f^1, f^2, f^3, ..., f^n be the strategy vectors representing respectively the one actually played and the beliefs of the players. Suppose that for every player i:
(i) f_i is a best response to f^i_{-i}; and
(ii) f is absolutely continuous w.r.t. f^i.
Then for every $\epsilon > 0$ and for almost every play path z there is a time $T(z, \epsilon)$ such that for all $t \geq T$, there exists an ϵ-equilibrium $\overline{f}$ of the repeated game satisfying $f_{z(t)}$ plays ϵ-like $\overline{f}$ and $\overline{f}$ is an ϵ-like Nash equilibrium.

Intuition Behind the Result in Three Steps

Step 1. Establish the self-correcting property of Bayesian updating.

Step 2. Conjecture that this self-correcting property implies that the probability distributions describing the players' beliefs about the future play converge. Call this situation a subjective equilibrium.

Step 3. Show that the behavior induced by a subjective equilibrium approximates the behavior of ϵ-Nash equilibrium.

Does it really work?

Does it really work?

Consider the following simple coordination game:

	L	R
L	$1, 1$	$0, 0$
R	$0, 0$	$1, 1$

Does it really work?

Consider the following simple coordination game:

	L	R
L	$1, 1$	$0, 0$
R	$0, 0$	$1, 1$

In one-shot play this game has three equilibria: both play L, both play R, and both randomize fifty-fifty.

Does it really work?

Consider the following simple coordination game:

	L	R
L	$1,1$	$0,0$
R	$0,0$	$1,1$

In one-shot play this game has three equilibria: both play L, both play R, and both randomize fifty-fifty.

In the repeated game the players believe that their opponents consider the following strategies: always play R ($\overline{R}$), always play L ($\overline{L}$), or always play fifty-fifty ($\overline{F}$).

Does it really work?

Consider the following simple coordination game:

	L	R
L	$1,1$	$0,0$
R	$0,0$	$1,1$

In one-shot play this game has three equilibria: both play L, both play R, and both randomize fifty-fifty.

In the repeated game the players believe that their opponents consider the following strategies: always play R ($\overline{R}$), always play L ($\overline{L}$), or always play fifty-fifty ($\overline{F}$).

Arrange the beliefs so that (R, L) is being played in the first period.

Does it really work?

Consider the following simple coordination game:

	L	R
L	$1,1$	$0,0$
R	$0,0$	$1,1$

In one-shot play this game has three equilibria: both play L, both play R, and both randomize fifty-fifty.

In the repeated game the players believe that their opponents consider the following strategies: always play R ($\overline{R}$), always play L ($\overline{L}$), or always play fifty-fifty ($\overline{F}$).

Arrange the beliefs so that (R, L) is being played in the first period.

In period 2, both players eliminate $\overline{R}$ and $\overline{L}$ as possible strategies of the opponent and play (L, R).

Does it really work?

Consider the following simple coordination game:

	L	R
L	$1,1$	$0,0$
R	$0,0$	$1,1$

In one-shot play this game has three equilibria: both play L, both play R, and both randomize fifty-fifty.

In the repeated game the players believe that their opponents consider the following strategies: always play R ($\overline{R}$), always play L ($\overline{L}$), or always play fifty-fifty ($\overline{F}$).

Arrange the beliefs so that (R, L) is being played in the first period.

In period 2, both players eliminate $\overline{R}$ and $\overline{L}$ as possible strategies of the opponent and play (L, R).

At the end of period 2 each concludes that the opponent is playing $\overline{F}$.

Does it really work?

Consider the following simple coordination game:

	L	R
L	$1, 1$	$0, 0$
R	$0, 0$	$1, 1$

In one-shot play this game has three equilibria: both play L, both play R, and both randomize fifty-fifty.

In the repeated game the players believe that their opponents consider the following strategies: always play R ($\overline{R}$), always play L ($\overline{L}$), or always play fifty-fifty ($\overline{F}$).

Arrange the beliefs so that (R, L) is being played in the first period.

In period 2, both players eliminate $\overline{R}$ and $\overline{L}$ as possible strategies of the opponent and play (L, R).

At the end of period 2 each concludes that the opponent is playing $\overline{F}$.

But neither player is actually playing $\overline{F}$.....

Caveat: Impossibility Theorem by Foster and Young (2001)

There exists a broad class of games that are unlearnable in principle by rational agents. That is, under any learning rule the players' strategies fail to come close to Nash equilibrium with probability one.

Conclusions

- The paper argues in favor of Rational Expectations Hypothesis.
- (Un)fortunately the main result of the paper does not seem to hold for a broad class of games.