Should players always choose Nash equilibria, or can they be better off switching strategies periodically?

Sebastian van Strien, Unversity of Warwick

April 10, 2010

Co-workers

- This work is in part with *Colin Sparrow* and in part with *Chris Harris* (from the Economics Department at Cambridge).
- References:

  - Sparrow, van Strien and Harris, *Fictitious play in* 3×3 *games: the transition between periodic and chaotic behaviour*, Games Econom. Behav., 2008, **63**, 259–291.
  - van Strien and Sparrow, *Fictitious Play in* 3×3 *Games: chaos and dithering behaviour*, submitted for publication.
  - van Strien, *A new class of Hamiltonian flows with random-walk behavior originating from zero-sum games and Fictitious Play*, submitted for publication.

Equilibria versus dynamics

- In game theory (and economics) there is a huge emphasis on the notion of Nash equilibria.
- In this talk I would like to argue that players can be better off playing non-equilibrium strategies.
- Some of the questions in my talk were triggered by an email from David Levine claiming:

  ‘Cycling of a learning procedure is a bad behavior’.

  Perhaps he was wrong?
- I will also compare our findings with recent work of Hart & Mas-Colell and of Foster & Young.

Best response

- Consider games with two players A and B who for each time $s \in [1, \infty)$ play n (possibly mixed) **actions** $m^A(s), m^B(s)$.
- At time s, players A and B obtain **utilities**

$$m^A(s)\, A\, m^B(s) \text{ resp. } m^A(s)\, B\, m^B(s).$$

- Let $p^A(t), p^B(t)$ be the **average past action**:

$$p^A(t) = \frac{1}{t}\int_1^t m^A(s) \text{ and } p^B(t) = \frac{1}{t}\int_1^t m^B(s).$$

- At time s, player A chooses the action $m^A(s) := \mathcal{BR}_A(p^B)$ which corresponds to the largest component of $Ap^B(s)$ (i.e. based on the average past actions of the other play).

Nash equilibria

- Let $\mathcal{BR}_A(p^B)$ and $\mathcal{BR}_A(p^B)$ be the **best responses** of player A respectively B.
- A **Nash equilibrium** is a choice of strategies from which no unilateral deviation by an individual player is profitable for that player. That is, (p_*^A, p_*^B) is a Nash equilibrium if

$$p_*^A \in \mathcal{BR}_A(p_*^B) \text{ and } p_*^B \in \mathcal{BR}_B(p_*^A).$$

Fictitious Play

In the 1950s Brown proposed **fictitious play**: this is a way in which players are able to naturally find the Nash equilibrium by flowing according to the following differential equation:

$$\begin{aligned} dp^A/dt &= \mathcal{BR}_A(p^B) - p^A \\ dp^B/dt &= \mathcal{BR}_B(p^A) - p^B \end{aligned} \tag{1}$$

where $BR_A(p^B) \in \Sigma_A$ is the best response of player A to players B position, and similarly for $BR_B(p^A) \in \Sigma_B$.
So each player's tendency is to adjust his or her strategy in a straight line from his/her strategy towards their (current) best response.

Examples

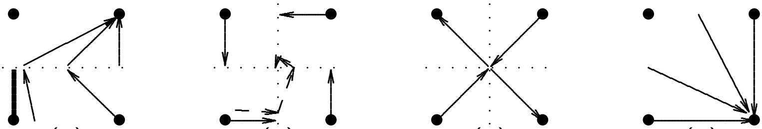


Figure: The possible motions in 2×2 games (up to relabeling, and shifting the indifference lines (drawn in dotted lines).

Case 1: $\begin{pmatrix} (0,-1) & (0,0) \\ (0,0) & (-1,-1) \end{pmatrix}$.

Case 2: $\begin{pmatrix} (-1,1) & (0,0) \\ (0,0) & (-1,1) \end{pmatrix}$, a zero sum game.

Case 3: $\begin{pmatrix} (1,1) & (0,0) \\ (0,0) & (1,1) \end{pmatrix}$, a coordination game.

Case 4: $\begin{pmatrix} (3,3) & (0,5) \\ (5,0) & (1,1) \end{pmatrix}$, a prisoner dilemma game.

Fictitious Play: learning

There is an interpretation of this game as a mechanism by which the players **learn** from the other players previous actions and in that case one often writes

$$\begin{aligned} dp^A/dt &= \left(\mathcal{BR}_A(p^B) - p^A\right)/t \\ dp^B/dt &= \left(\mathcal{BR}_B(p^A) - p^B\right)/t \end{aligned} \tag{2}$$

The dynamics of this system and the previous are the same up to time-parametrisation.

There are many papers on Fictitious Play as a model for learning, see for example Fudenberg and Levine's monograph.

Convergence and non-convergence of fictitious play

- If $A + B = 0$ then we have a so-called **zero-sum** game. It was shown in the 1950s by Robinson that then the game converges (albeit slowly) to the set of Nash equilibria.
- There are a few other non-zero sum examples for which it is shown that the game converges (for example, $2 \times n$ games), but in those cases the Nash equilibrium is not unique and the flow does not have unique attractor.
- There is a famous example due to **Shapley** from the 1960s which shows that in general the evolution does NOT converge to an equilibrium of the game, but can have periodic behavior.

Aim of this talk

In this talk:

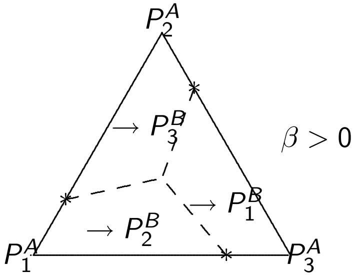
1. I will describe a family of examples depending on a parameter β where both players have three stategies
   - For $\beta = 0$ this corresponds to **Shapley's example**.
   - For $\beta = \sigma$ where $\sigma := (\sqrt{5} - 1)/2 \approx 0.618$ is the golden mean, the game is equivalent to a **zero-sum game**, so then play always converges to the interior equilibrium E^A, E^B.
   - For $\beta \in (\sigma, \tau)$ where $\tau \approx 0.915$ the game has infinitely many periodic orbits and there is **chaos**.
2. It turns out that the mean pay-off of the players *Pareto dominates* the Nash equilibrium, and even correlated equilibrium.

A specific family

- We consider a family of games with 3 strategies, so that Σ_A, Σ_B are the simplex in $\mathbb{R}^3$, i.e. just triangles. So the flow takes place in the four-dimensional ball $\Sigma_A \times \Sigma_B$.
- In normal form the games we consider are:

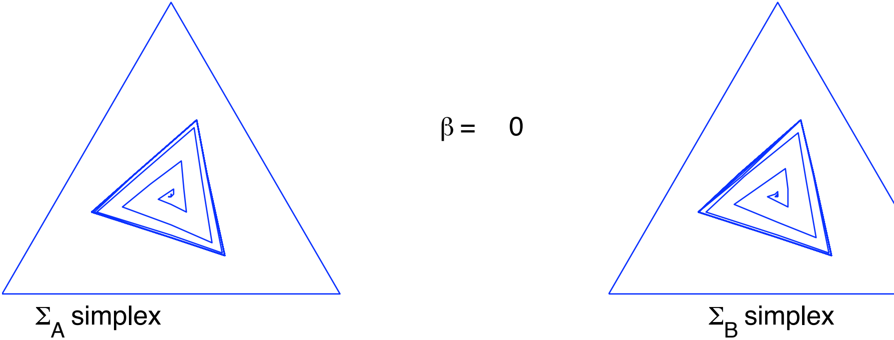
$$A_\beta = \begin{pmatrix} 1 & 0 & \beta \\ \beta & 1 & 0 \\ 0 & \beta & 1 \end{pmatrix} \quad B_\beta = \begin{pmatrix} -\beta & 1 & 0 \\ 0 & -\beta & 1 \\ 1 & 0 & -\beta \end{pmatrix}, \qquad (3)$$

- It is not hard to see that when $\sigma := (\sqrt{5} - 1)/2 \approx 0.618$, the game is equivalent to a zero-sum game.
    - rescaling B to $\tilde{B} = \sigma(B - 1)$ gives $A + \tilde{B} = 0$, so then play always converges to the interior equilibrium E^A, E^B.

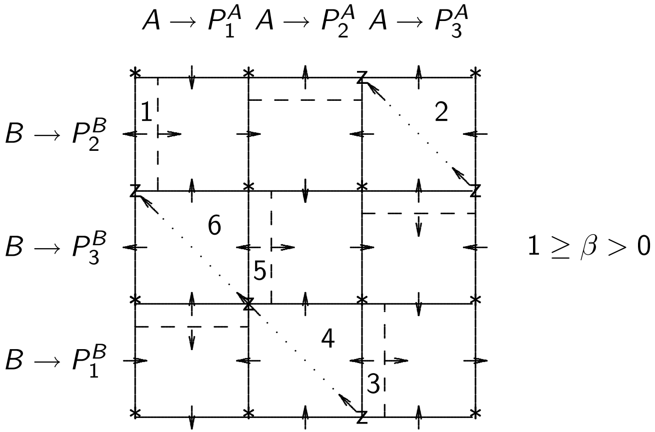


Note that $\mathcal{BR}_A$ and $\mathcal{BR}_B$ are multivalued when you are on the dashed lines, and so orbits are not necessarily unique. Also note that the vector field is not continuous and so the flow is not necessarily continuous:

For $\beta = 0$ all orbits spiral clockwise towards the Shapley orbit.



One can visualize an orbit as a path in the following diagram, indicating actions the players choose. Only certain transitions are allowed.



Various periodic orbits

Theorem (Simple periodic orbits)

- *For $\beta \in (-1, \sigma)$, where σ is the golden mean ($\sigma = 0.618..$), there exists a periodic orbit with play* $(1,2),(2,2),(2,3),(3,3),(3,1),(1,1)$ *(the* **Shapley orbit***) which attracts an open set of initial conditions.*
- *For $\beta \in (\sigma, 1)$ there exists another periodic orbit with play* $(1,2),(3,2),(3,1),(2,1),(2,3),(1,3)$ *(the* **anti-Shapley orbit***) which becomes attracting when $\sigma \in (\tau, 1)$ where $\tau \approx 0.915$.*
- *For $\beta \in (\sigma, 1)$, there is also a periodic orbit of mixed strategies*

$$(\bar{1},\bar{1}),(\bar{1},\bar{2}),(\bar{2},\bar{2}),(\bar{2},\bar{3}),(\bar{3},\bar{3}),(\bar{3},\bar{1}).$$

Here $(\bar{i},\bar{j})$ means that player A is **not** *playing strategy i, but is indifferent between the other two strategies, and similarly for B.*

Basin of attraction of Nash equilibrium

One can ask, do *all* orbits converge to the attracting periodic Shapley orbit when $\beta \in (0, \sigma)$. NO:

Theorem

Define the basin of attraction of the Nash equilibrium E as

$$W^s(E) = \{(p^A, p^B);\ \textit{orbits starting in } (p^A, p^B) \textit{ converges to } E\}.$$

Then for $\beta \in (0, 1)$, $\beta \neq \sigma$, $W^s(E)$ contains a countable union of codimension-one sets (each of them a cone with apex E and as base some polygonal set).

So the stable set of E is extremely complicated: note that the differential equation is *not* smooth and not even continuous. However, or $\beta \in (0, 1)$ the flow *is continuous*.

Chaos

Theorem (An abundance of periodic play and chaos)

For each $\beta \in (0, 1)$ and each $n \geq 1$ there are infinitely many different orbits γ_s, $s = 1, 2, \ldots$ with cyclic play of period $N_s \to \infty$ as $s \to \infty$ but with essential period equal to $6n$. Moreover,

- *for $\beta \in (0, \sigma)$, these orbits with cyclic play reach the interior equilibrium E in finite time;*
- *for $\beta \in (\sigma, 1)$ these orbits with cyclic play are genuine periodic orbits.*
- *for $\beta \in (\sigma, \tau)$, where $\tau \approx 0.915$, the dynamical system is chaotic.*

Chaos occurs in a rather extraordinary way, but I will not discuss this here.

Results are robust

The above results do not require the matrices to be of precisely this form; they hold for games corresponding to an open set of matrices:

Theorem (Robustness)

For each $\beta \in (0,1)$ with $\beta \neq \sigma$, there exists $\epsilon > 0$ so that for each 3×3 matrices A and B with

$$||A - A_\beta||, ||B - B_\beta|| < \epsilon$$

the previous theorems also hold.

Is Chaos interesting from an economic point of view?

- When I emailed David Levine the papers with these results, he emailed me back:

  'This is an interesting mathematical fact; ... but from an economic point of view not that important. Cycling of a learning procedure is a bad behavior; chaos isn't appreciably worse.'

- His motivation for saying this is that if players notice they are cycling, they surely notice this and change their behavior accordingly.

- But perhaps the players DO notice this, but realize that they are doing very well, as we will see.

Non-convergence

Indeed, during the last 10 years there were a number of papers giving examples where fictitious play does not lead to convergence to Nash equilibrium:

- Shapley (1964)
- Jordan (1993)
- Foster and Young (1998): non-convergence in some coordination games
- Hart and Mas-Colell (2003): convergence is not possible in open classes of games

Since one does not have convergence, fictitious play is ‘useless’. For this reason fictitious play has received somewhat less attention recently.

Alternative learning processes

Instead, recently various other processes were proposed to get convergence to Nash equilibria, or to the somewhat more general 'set of correlated equilibria'. Indeed,

- Kalai & Lehrer (1993)....
- In a seminal paper, Hart and Mas-Colell (2000) show that, if the players of a repeated game choose their actions according to a simple regret-based rule of thumb, then
  - the empirical distribution of the action profile may not itself converge,
  - but the distance between this distribution and the set of correlated equilibria converges to zero with probability one.
- Foster and Young (2003) have a modified version of this.

What are players aiming for?

But why is going to a Nash equilibrium so important? Surely players are just trying to maximize (average) pay-off. It may be natural to consider

$$\pi^A(T) = \frac{1}{T-1}\int_1^T \mathcal{BR}_A(p^B(s))\, A\, \mathcal{BR}_B(p^A(s))\, ds$$

$$\pi^B(T) = \frac{1}{T-1}\int_1^T \mathcal{BR}_A(p^B(s))\, B\, \mathcal{BR}_B(p^A(s))\, ds$$

(Note that fictitious play

$$\begin{aligned} dp^A/ds &= \left(\mathcal{BR}_A(p^B) - p^A\right)/s \\ dp^B/ds &= \left(\mathcal{BR}_B(p^A) - p^B\right)/s \end{aligned}$$

starts at time $s = 1$).

Theorem

In our game mean pay-off for **both** *players (when starting in the basin of the (attracting) Shapley orbit), Pareto dominates the Nash equilibrium.*

Nash equilibria and correlated equilibria

$\triangleleft$ Of course one might say: the learning models by Hart & Mass-Colell, Foster & Young and others converge (in some sense) to correlated equilibria, and these are known to (Pareto) better than Nash. So you have not shown anything....

$\triangleleft$ But it turns out that in our family of games, the set of correlated equilibria is unique (and is equal to the Nash equilibria). So pay-off beats that of the other learning models mentioned before.

Conclusion

- We saw that fictitious play has associated to it complicated dynamical behaviour
- Of course an economist can say: people behave rationally and if they do not converge then they will notice this. So chaos is a mathematical curiosity. I'd like to argue this is not so:
- As we saw, stationary solutions can be Pareto-suboptimal.
- Perhaps, without necessarily trusting each other, there are ways in which players can dynamically develop an interaction which beats Nash equilibria and which - in the non-zero sum case - allows both players to benefit from this.