Time average replicator and best reply dynamics

Sylvain Sorin

Equipe Combinatoire et Optimisation, Faculté de Mathématiques, UPMC-Paris 6
Laboratoire d'Econométrie, Ecole Polytechnique
****
sorin@math.jussieu.fr
****

March 2008

Joint work with Josef Hofbauer (Wien) and Yannick Viossat (Dauphine).

Main examples of deterministic evolutionary game dynamics:
replicator dynamics (RD)
best reply dynamics (BRD).
A. replicator dynamics)
symmetric 2 person game with $K \times K$ payoff matrix A
single population, *replicator equation* on the simplex $\Delta(K)$ of $\mathbb{R}^K$

$$\dot{x}_t^k = x_t^k \left(e^k A x_t - x_t A x_t \right), \quad k \in K \qquad (RD) \qquad (1)$$

x_t^k : frequency of strategy k at time t.

Joint work with Josef Hofbauer (Wien) and Yannick Viossat (Dauphine).

Main examples of deterministic evolutionary game dynamics:
replicator dynamics (RD)
best reply dynamics (BRD).

A. *replicator dynamics*)
symmetric 2 person game with $K \times K$ payoff matrix A
single population, *replicator equation* on the simplex $\Delta(K)$ of $\mathbf{R}^K$

$$\dot{x}_t^k = x_t^k \left(e^k A x_t - x_t A x_t \right), \quad k \in K \qquad (RD) \qquad (1)$$

x_t^k : frequency of strategy k at time t.

Joint work with Josef Hofbauer (Wien) and Yannick Viossat (Dauphine).

Main examples of deterministic evolutionary game dynamics:
replicator dynamics (RD)
best reply dynamics (BRD).
A. replicator dynamics)
symmetric 2 person game with $K \times K$ payoff matrix A
single population, *replicator equation* on the simplex $\Delta(K)$ of $\boldsymbol{R}^K$

$$\dot{x}_t^k = x_t^k \left(e^k A x_t - x_t A x_t \right), \quad k \in K \qquad (RD) \tag{1}$$

x_t^k : frequency of strategy k at time t.

More generally fitness function $F : K \times \Delta(K) \to \boldsymbol{R}$

$$\dot{x}_t^k = x_t^k \left(F(k, x_t) - \sum_k x_t^k F(k, x_t) \right)$$

Taylor and Jonker (1978) : basic selection dynamics for the evolutionary games of Maynard Smith (1982).

Interpretation: infinite population of replicating players per capita growth rate of the frequencies of pure strategies linearly related to their payoffs.

More generally fitness function $F : K \times \Delta(K) \to \boldsymbol{R}$

$$\dot{x}_t^k = x_t^k \left(F(k, x_t) - \sum_k x_t^k F(k, x_t) \right)$$

Taylor and Jonker (1978) : basic selection dynamics for the evolutionary games of Maynard Smith (1982).
Interpretation: infinite population of replicating players per capita growth rate of the frequencies of pure strategies linearly related to their payoffs.

B. best reply dynamics
differential inclusion on $\Delta(K)$

$$\dot{z}_t \in BR(z_t) - z_t, \qquad t \geq 0 \qquad (BRD) \tag{2}$$

introduced by Gilboa and Matsui (1991).
$BR(z)$: the set of best replies to the strategy profile z

$$BR(z) = \{\, y \in \Delta(K); yAz \geq y'Az, \forall y' \in \Delta(K)\}.$$

interpretation: infinite population, in each small time interval, a small fraction
a) dies and is replaced by individuals best "adapted" to the current landscape
b) revises their strategies and changes to a best reply against the present population distribution.
It is the prototype of a population model of rational (but myopic)

B. best reply dynamics
differential inclusion on $\Delta(K)$

$$\dot{z}_t \in BR(z_t) - z_t, \qquad t \geq 0 \qquad (BRD) \tag{2}$$

introduced by Gilboa and Matsui (1991).
$BR(z)$: the set of best replies to the strategy profile z

$$BR(z) = \{\, y \in \Delta(K); yAz \geq y'Az, \forall y' \in \Delta(K)\}.$$

interpretation: infinite population, in each small time interval, a small fraction
a) dies and is replaced by individuals best "adapted" to the current landscape
b) revises their strategies and changes to a best reply against the present population distribution.
It is the prototype of a population model of rational (but myopic)

B. best reply dynamics

differential inclusion on $\Delta(K)$

$$\dot{z}_t \in BR(z_t) - z_t, \qquad t \geq 0 \qquad (BRD) \tag{2}$$

introduced by Gilboa and Matsui (1991).

$BR(z)$: the set of best replies to the strategy profile z

$$BR(z) = \{\, y \in \Delta(K); yAz \geq y'Az, \forall y' \in \Delta(K)\}.$$

interpretation: infinite population, in each small time interval, a small fraction

a) dies and is replaced by individuals best “adapted” to the current landscape

b) revises their strategies and changes to a best reply against the present population distribution.

It is the prototype of a population model of rational (but myopic)

(BRD) closely related to fictitious play, process introduced by Brown (1949).

two person repeated game : choices I and J
dynamics :

$$x_{n+1} \in BR^1(Y_n), \quad y_{n+1} \in BR^2(X_n)$$

with $x_{n+1} \in \Delta(I)$, $y_{n+1} \in \Delta(J)$ and $X_n = \frac{1}{n}\sum_{m=1}^{n} x_m$.continuous time analog: $X_t = \frac{1}{t}\int_0^t x_s ds$, $Y_t = \frac{1}{t}\int_0^t y_s ds$ with $x_t \in BR^1(Y_t)$, $y_t \in BR^2(X_t)$.
This implies that $Z_t = (X_t, Y_t)$ satisfies the *continuous fictitious play* equation

$$\dot{Z}_t \in \frac{1}{t}(BR(Z_t) - Z_t), \quad t > 0 \qquad (CFP) \qquad (3)$$

which is equivalent to (BRD) via a change in time $Z_{e^s} = Z_s$.

(BRD) closely related to fictitious play, process introduced by Brown (1949).
two person repeated game : choices I and J
dynamics :

$$x_{n+1} \in BR^1(Y_n), \quad y_{n+1} \in BR^2(X_n)$$

with $x_{n+1} \in \Delta(I)$, $y_{n+1} \in \Delta(J)$ and $X_n = \frac{1}{n}\sum_{m=1}^{n} x_m$.continuous time analog: $X_t = \frac{1}{t}\int_0^t x_s ds$, $Y_t = \frac{1}{t}\int_0^t y_s ds$ with $x_t \in BR^1(Y_t)$, $y_t \in BR^2(X_t)$.
This implies that $Z_t = (X_t, Y_t)$ satisfies the *continuous fictitious play* equation

$$\dot{Z}_t \in \frac{1}{t}(BR(Z_t) - Z_t), \quad t > 0 \qquad (CFP) \tag{3}$$

which is equivalent to (BRD) via a change in time $Z_{e^s} = Z_s$

(BRD) closely related to fictitious play, process introduced by Brown (1949).
two person repeated game : choices I and J
dynamics :

$$x_{n+1} \in BR^1(Y_n), \quad y_{n+1} \in BR^2(X_n)$$

with $x_{n+1} \in \Delta(I)$, $y_{n+1} \in \Delta(J)$ and $X_n = \frac{1}{n}\sum_{m=1}^{n} x_m$.continuous time analog: $X_t = \frac{1}{t}\int_0^t x_s ds$, $Y_t = \frac{1}{t}\int_0^t y_s ds$ with $x_t \in BR^1(Y_t)$, $y_t \in BR^2(X_t)$.
This implies that $Z_t = (X_t, Y_t)$ satisfies the *continuous fictitious play* equation

$$\dot{Z}_t \in \frac{1}{t}\left(BR(Z_t) - Z_t\right), \quad t > 0 \qquad (CFP) \qquad (3)$$

which is equivalent to (BRD) via a change in time $Z_{e^s} = Z_s$

(BRD) closely related to fictitious play, process introduced by Brown (1949).
two person repeated game : choices I and J
dynamics :

$$x_{n+1} \in BR^1(Y_n), \quad y_{n+1} \in BR^2(X_n)$$

with $x_{n+1} \in \Delta(I)$, $y_{n+1} \in \Delta(J)$ and $X_n = \frac{1}{n}\sum_{m=1}^{n} x_m$.continuous time analog: $X_t = \frac{1}{t}\int_0^t x_s ds$, $Y_t = \frac{1}{t}\int_0^t y_s ds$ with $x_t \in BR^1(Y_t)$, $y_t \in BR^2(X_t)$.
This implies that $Z_t = (X_t, Y_t)$ satisfies the *continuous fictitious play* equation

$$\dot{Z}_t \in \frac{1}{t}(BR(Z_t) - Z_t), \quad t > 0 \qquad (CFP) \qquad (3)$$

which is equivalent to (BRD) via a change in time $Z_{e^s} = Z_s$

(BRD) closely related to fictitious play, process introduced by Brown (1949).
two person repeated game : choices I and J
dynamics :

$$x_{n+1} \in BR^1(Y_n), \quad y_{n+1} \in BR^2(X_n)$$

with $x_{n+1} \in \Delta(I)$, $y_{n+1} \in \Delta(J)$ and $X_n = \frac{1}{n}\sum_{m=1}^{n} x_m$.continuous time analog: $X_t = \frac{1}{t}\int_0^t x_s ds$, $Y_t = \frac{1}{t}\int_0^t y_s ds$ with $x_t \in BR^1(Y_t)$, $y_t \in BR^2(X_t)$.
This implies that $Z_t = (X_t, Y_t)$ satisfies the *continuous fictitious play* equation

$$\dot{Z}_t \in \frac{1}{t}\left(BR(Z_t) - Z_t\right), \quad t > 0 \qquad (CFP) \tag{3}$$

which is equivalent to (BRD) via a change in time $Z_{e^s} = z_s$.

Despite the different interpretation and the different dynamic character there are amazing similarities in the long run behaviour of these two dynamics, that have been summarized in the following heuristic principle,

For many games, the long run behaviour ($t \to \infty$) of the time averages $X_t = \frac{1}{t}\int_0^t x_s ds$ of the trajectories x_t of the replicator equation is the same as for the BR trajectories.

We provide here a rigorous statement that largely explains this heuristics.

Despite the different interpretation and the different dynamic character there are amazing similarities in the long run behaviour of these two dynamics, that have been summarized in the following heuristic principle,
For many games, the long run behaviour ($t \to \infty$) of the time averages $X_t = \frac{1}{t}\int_0^t x_s ds$ of the trajectories x_t of the replicator equation is the same as for the BR trajectories.
We provide here a rigorous statement that largely explains this heuristics.

Despite the different interpretation and the different dynamic character there are amazing similarities in the long run behaviour of these two dynamics, that have been summarized in the following heuristic principle,

For many games, the long run behaviour ($t \to \infty$) of the time averages $X_t = \frac{1}{t}\int_0^t x_s ds$ of the trajectories x_t of the replicator equation is the same as for the BR trajectories.

We provide here a rigorous statement that largely explains this heuristics.

We show that for any interior solution of (RD), for every $t \geq 0$, x_t is an approximate best reply against X_t and the approximation gets better as $t \to \infty$.

This implies that X_t is an asymptotic pseudo trajectory of (BRD) and hence the limit set of X_t has the same properties as a limit set of a true orbit of (BRD), i.e. it is invariant and internally chain transitive under (BRD).

The main tool to prove this is via the logit map which is a canonical smoothing of the best response correspondence. We show that x_t equals the logit approximation at X_t with error rate $\frac{1}{t}$.

We show that for any interior solution of (RD), for every $t \geq 0$, x_t is an approximate best reply against X_t and the approximation gets better as $t \rightarrow \infty$.

This implies that X_t is an asymptotic pseudo trajectory of (BRD) and hence the limit set of X_t has the same properties as a limit set of a true orbit of (BRD), i.e. it is invariant and internally chain transitive under (BRD).

The main tool to prove this is via the logit map which is a canonical smoothing of the best response correspondence. We show that x_t equals the logit approximation at X_t with error rate $\frac{1}{t}$.

We show that for any interior solution of (RD), for every $t \geq 0$, x_t is an approximate best reply against X_t and the approximation gets better as $t \to \infty$.
This implies that X_t is an asymptotic pseudo trajectory of (BRD) and hence the limit set of X_t has the same properties as a limit set of a true orbit of (BRD), i.e. it is invariant and internally chain transitive under (BRD).
The main tool to prove this is via the logit map which is a canonical smoothing of the best response correspondence. We show that x_t equals the logit approximation at X_t with error rate $\frac{1}{t}$.

The model will be in the framework of an N-person game but we consider the dynamics for one player, without hypotheses on the behavior of the others.

Hence, from the point of view of this player, he is facing a (measurable) vector outcome process $\mathcal{U} = \{ U_t, t \geq 0\}$, with values in the cube $C = [-c, c]^K$ where K is his action's set and c is some positive constant. U_t^k is the payoff at time t if k is the action at that time.The cumulative vector outcome up to stage t is $S_t = \int_0^t U_s ds$ and its time average is denoted $\bar{U}_t = \frac{1}{t} S_t$.

br denotes the (payoff based) best reply correspondence from C to Δ defined by

$$\mathbf{br}(U) = \{ x \in \Delta; \langle x, U \rangle = \max_{y \in \Delta} \langle y, U \rangle \}$$

The model will be in the framework of an N-person game but we consider the dynamics for one player, without hypotheses on the behavior of the others.
Hence, from the point of view of this player, he is facing a (measurable) vector outcome process $\mathcal{U} = \{ U_t, t \geq 0\}$, with values in the cube $C = [-c, c]^K$ where K is his action's set and c is some positive constant. U_t^k is the payoff at time t if k is the action at that time. The cumulative vector outcome up to stage t is $S_t = \int_0^t U_s ds$ and its time average is denoted $\bar{U}_t = \frac{1}{t} S_t$.
br denotes the (payoff based) best reply correspondence from C to Δ defined by

$$\mathbf{br}(U) = \{ x \in \Delta; \langle x, U \rangle = \max_{y \in \Delta} \langle y, U \rangle \}$$

The model will be in the framework of an N-person game but we consider the dynamics for one player, without hypotheses on the behavior of the others.
Hence, from the point of view of this player, he is facing a (measurable) vector outcome process $\mathcal{U} = \{ U_t, t \geq 0\}$, with values in the cube $C = [-c, c]^K$ where K is his action's set and c is some positive constant. U_t^k is the payoff at time t if k is the action at that time.The cumulative vector outcome up to stage t is $S_t = \int_0^t U_s ds$ and its time average is denoted $\bar{U}_t = \frac{1}{t} S_t$.
br denotes the (payoff based) best reply correspondence from C to Δ defined by

$$\mathbf{br}(U) = \{ x \in \Delta; \langle x, U \rangle = \max_{y \in \Delta} \langle y, U \rangle \}$$

The $\mathcal{U}$-fictitious play process (*FPP*) is defined on Δ by

$$\dot{X}_t \in \frac{1}{t}[\mathbf{br}(\bar{U}_t) - X_t] \tag{4}$$

The $\mathcal{U}$-replicator process (*RP*) is specified by the following equation on Δ:

$$\dot{x}_t^k = x_t^k[U_t^k - \langle x_t, U_t \rangle], \qquad k \in K. \tag{5}$$

The $\mathcal{U}$-fictitious play process (*FPP*) is defined on Δ by

$$\dot{X}_t \in \frac{1}{t}[\mathbf{br}(\bar{U}_t) - X_t] \tag{4}$$

The $\mathcal{U}$-replicator process (*RP*) is specified by the following equation on Δ:

$$\dot{x}_t^k = x_t^k[U_t^k - \langle x_t, U_t\rangle], \qquad k \in K. \tag{5}$$

Explicitely, in the framework of a N-player game with payoff for player 1 defined by a function G from $\prod_{i\in N} K^i$ to $\boldsymbol{R}$, with $X^i = \Delta(K^i)$, U is the vector payoff i.e. $U_t = G(, x_t^{-1})$.
If all the players follow a (payoff based) fictitious play dynamics, each time average strategy satisfies (4). For $N = 2$ this is (CFP).
If all the players follow the replicator dynamics then (5) is the replicator dynamics equation.

Explicitely, in the framework of a N-player game with payoff for player 1 defined by a function G from $\prod_{i \in N} K^i$ to $\boldsymbol{R}$, with $X^i = \Delta(K^i)$, U is the vector payoff i.e. $U_t = G(, x_t^{-1})$.
If all the players follow a (payoff based) fictitious play dynamics, each time average strategy satisfies (4). For $N = 2$ this is (CFP).
If all the players follow the replicator dynamics then (5) is the replicator dynamics equation.

Define a map L from $\boldsymbol{R}^K$ to Δ by

$$L^k(V) = \frac{\exp V^k}{\sum_j \exp V^j}. \tag{6}$$

Given $\eta > 0$, let $[\mathbf{br}]^\eta$ be the correspondence from C to Δ with graph being the η-neighborhood for the uniform norm of the graph of **br**.
The L map and the **br** correspondence are related as follows:

Define a map L from $\boldsymbol{R}^K$ to Δ by

$$L^k(V) = \frac{\exp V^k}{\sum_j \exp V^j}. \tag{6}$$

Given $\eta > 0$, let $[\mathbf{br}]^\eta$ be the correspondence from C to Δ with graph being the η-neighborhood for the uniform norm of the graph of **br**.
The L map and the **br** correspondence are related as follows:

Proposition

For any $U \in C$ *and* $\varepsilon > 0$

$$L(U/\varepsilon) \in [\mathbf{br}]^{\eta(\varepsilon)}(U)$$

with $\eta(\varepsilon) \to 0$ *as* $\varepsilon \to 0$.

Remarks

L is also given by

$$L(V) = \text{argmax}_{x \in \Delta}\{ \langle x, V \rangle - \sum_k x^k \log x^k \}.$$

Hence introducing the (payoff based) perturbed best reply $\mathbf{br}^\varepsilon$ from C to Δ defined by

$$\mathbf{br}^\varepsilon(U) = \text{argmax}_{x \in \Delta}\{ \langle x, U \rangle - \varepsilon \sum_k x^k \log x^k \}$$

one has

$$L(U/\varepsilon) = \mathbf{br}^\varepsilon(U)$$

The map $\mathbf{br}^\varepsilon$ is the logit approximation.

Remarks

L is also given by

$$L(V) = \text{argmax}_{x \in \Delta}\{\, \langle x, V \rangle - \sum_k x^k \log x^k \}.$$

Hence introducing the (payoff based) perturbed best reply $\mathbf{br}^{\varepsilon}$ from C to Δ defined by

$$\mathbf{br}^{\varepsilon}(U) = \text{argmax}_{x \in \Delta}\{\, \langle x, U \rangle - \varepsilon \sum_k x^k \log x^k \}$$

one has

$$L(U/\varepsilon) = \mathbf{br}^{\varepsilon}(U)$$

The map $\mathbf{br}^{\varepsilon}$ is the logit approximation.

Remarks

L is also given by

$$L(V) = \text{argmax}_{x \in \Delta}\{\langle x, V\rangle - \sum_k x^k \log x^k\}.$$

Hence introducing the (payoff based) perturbed best reply $\mathbf{br}^\varepsilon$ from C to Δ defined by

$$\mathbf{br}^\varepsilon(U) = \text{argmax}_{x \in \Delta}\{\langle x, U\rangle - \varepsilon \sum_k x^k \log x^k\}$$

one has

$$L(U/\varepsilon) = \mathbf{br}^\varepsilon(U)$$

The map $\mathbf{br}^\varepsilon$ is the logit approximation.

The following procedure has been introduced in discrete time in the framework of on-line algorithms under the name "multiplicative weight algorithm". We use here the name (CEW) (continuous exponential weight) for the process defined, given $\mathcal{U}$, by

$$x_t = L(\int_0^t U_s ds).$$

The main property of (CEW) that will be used is that it provides an explicit solution of (RD).

The following procedure has been introduced in discrete time in the framework of on-line algorithms under the name "multiplicative weight algorithm". We use here the name (CEW) (continuous exponential weight) for the process defined, given $\mathcal{U}$, by

$$x_t = L(\int_0^t U_s ds).$$

The main property of (CEW) that will be used is that it provides an explicit solution of (RD).

The following procedure has been introduced in discrete time in the framework of on-line algorithms under the name "multiplicative weight algorithm". We use here the name (CEW) (continuous exponential weight) for the process defined, given $\mathcal{U}$, by

$$x_t = L(\int_0^t U_s ds).$$

The main property of (CEW) that will be used is that it provides an explicit solution of (RD).

Proposition

(CEW) satisfies (RP).

Proof

Straightforward computations lead to

$$\dot{x}_t^k = x_t^k U_t^k - x_t^k \sum_j \frac{U_t^j \exp \int_0^t U_v^j dv}{\sum_j \exp \int_0^t U_v^j dv}$$

which is

$$\dot{x}_t^k = x_t^k [U_t^k - \langle x_t, U_t \rangle]$$

hence gives the previous (RP) equation (5). ∎

Proposition

(CEW) satisfies (RP).

Proof

Straightforward computations lead to

$$\dot{x}_t^k = x_t^k U_t^k - x_t^k \sum_j \frac{U_t^j \exp \int_0^t U_v^j dv}{\sum_j \exp \int_0^t U_v^j dv}$$

which is

$$\dot{x}_t^k = x_t^k [U_t^k - \langle x_t, U_t \rangle]$$

hence gives the previous (RP) equation (5). ∎

The link with the best reply correspondence is the following.

Proposition

CEW satisfies

$$x_t \in [\mathbf{br}]^{\delta(t)}(\bar{U}_t)$$

with $\delta(t) \to 0$ *as* $t \to \infty$.

Proof

Write

$$x_t = L(\int_0^t U_s ds) = L(t\,\bar{U}_t)$$

$$= L(U/\varepsilon) \in [\mathbf{br}]^{\eta(\varepsilon)}(U)$$

with $U = \bar{U}_t$ and $\varepsilon = 1/t$, by Proposition 3.1. Let $\delta(t) = \eta(1/t)$. ∎

The link with the best reply correspondence is the following.

Proposition

CEW satisfies

$$x_t \in [\mathbf{br}]^{\delta(t)}(\bar{U}_t)$$

with $\delta(t) \to 0$ *as* $t \to \infty$.

Proof

Write

$$x_t = L(\int_0^t U_s ds) = L(t\ \bar{U}_t)$$

$$= L(U/\varepsilon) \in [\mathbf{br}]^{\eta(\varepsilon)}(U)$$

with $U = \bar{U}_t$ and $\varepsilon = 1/t$, by Proposition 3.1. Let $\delta(t) = \eta(1/t)$. ∎

We describe here the consequences for the time average process.
Define

$$X_t = \frac{1}{t}\int_0^t x_s ds$$

Proposition

If x_t follows (CEW) then X_t satisfies

$$\dot{X}_t \in \frac{1}{t}([\mathbf{br}]^{\delta(t)}(\bar{U}_t) - X_t) \qquad (*)$$

with $\delta(t) \to 0$ as $t \to \infty$.

We describe here the consequences for the time average process.
Define

$$X_t = \frac{1}{t}\int_0^t x_s ds$$

Proposition

If x_t follows (CEW) then X_t satisfies

$$\dot{X}_t \in \frac{1}{t}([\mathbf{br}]^{\delta(t)}(\bar{U}_t) - X_t) \qquad (*)$$

with $\delta(t) \to 0$ as $t \to \infty$.

Consider a 2 person (bimatrix) game (A, B).
If the game is symmetric this gives rise to the single population replicator dynamics (RD) and best reply dynamics (BRD) as defined in section 1.
Otherwise, we consider the two population replicator dynamics

$$\dot{x}_t^k = x_t^k \left(e^k A y_t - x_t A y_t \right), \quad k \in K_1 \tag{7}$$

$$\dot{y}_t^k = y_t^k \left(x_t B e^k - x_t B y_t \right), \quad k \in K_2$$

and the corresponding BR dynamics as in (3).
Let M be the state space (a simplex Δ or a product of simplices $\Delta_1 \times \Delta_2$).

We now use the previous results with the $\mathcal{U}$ process being defined by $U_t = Ay_t$ for player 1, hence $\bar{U}_t = AY_t$. Note that $\mathbf{br}(AY) = BR_1(Y)$.

Proposition

The limit set of every replicator time average process X_t starting from an initial point $x_0 \in M$ is a closed subset of M which is invariant and internally chain transitive (ICT) under (BRD).

Proof

Equation (*) implies that X_t satisfies a perturbed version of (CFP) hence X_{e^t} is a perturbed solution to the differential inclusion (BRD), according to Definition II in Benaim Hofbauer Sorin. Now apply Theorem 3.6 of that paper. ∎

We now use the previous results with the $\mathcal{U}$ process being defined by $U_t = Ay_t$ for player 1, hence $\bar{U}_t = AY_t$. Note that $\mathbf{br}(AY) = BR_1(Y)$.

Proposition

The limit set of every replicator time average process X_t starting from an initial point $x_0 \in M$ is a closed subset of M which is invariant and internally chain transitive (ICT) under (BRD).

Proof

Equation (*) implies that X_t satisfies a perturbed version of (CFP) hence X_{e^t} is a perturbed solution to the differential inclusion (BRD), according to Definition II in Benaim Hofbauer Sorin. Now apply Theorem 3.6 of that paper.

We now use the previous results with the $\mathcal{U}$ process being defined by $U_t = Ay_t$ for player 1, hence $\bar{U}_t = AY_t$. Note that $\mathbf{br}(AY) = BR_1(Y)$.

Proposition

The limit set of every replicator time average process X_t starting from an initial point $x_0 \in M$ is a closed subset of M which is invariant and internally chain transitive (ICT) under (BRD).

Proof

Equation (*) implies that X_t satisfies a perturbed version of (CFP) hence X_{e^t} is a perturbed solution to the differential inclusion (BRD), according to Definition II in Benaim Hofbauer Sorin. Now apply Theorem 3.6 of that paper. ∎

In particular this implies:

Proposition

Let $\mathcal{A}$ be the global attractor (i.e., the maximal invariant set) of (BRD). Then the limit set of every replicator time average process X_t is a subset of $\mathcal{A}$.

A procedure satisfies external consistency if for each process $\mathcal{U} \in \boldsymbol{R}^K$, it produces a process $x_t \in \Delta$, such that for all k

$$\int_0^t [U_s^k - \langle x_s, U_s \rangle] ds \leq C_t = o(t)$$

This property says that the (expected) average payoff induced by x_t along the play is asymptotically not less than the payoff obtained by any fixed choice $k \in K$.

A procedure satisfies external consistency if for each process $\mathcal{U} \in \boldsymbol{R}^K$, it produces a process $x_t \in \Delta$, such that for all k

$$\int_0^t [U_s^k - \langle x_s, U_s \rangle] ds \leq C_t = o(t)$$

This property says that the (expected) average payoff induced by x_t along the play is asymptotically not less than the payoff obtained by any fixed choice $k \in K$.

We recall this result from Sorin (2007), where the aim was to compare discrete and continuous time procedures.

Proposition

(CEW) satifies external consistency.

We recall this result from Sorin (2007), where the aim was to compare discrete and continuous time procedures.

Proposition

(CEW) satifies external consistency.

In fact a direct and more simple proof is available, see Hofbauer (2004):

Proposition

(RP) satifies external consistency.

Proof

By integrating equation (5), one obtains, on the support of x_0:

$$\int_0^t [U_s^k - \langle x_s, U_s \rangle] ds = \int_0^t \frac{\dot{x}_s^k}{x_s^k} ds = \log(\frac{x_t^k}{x_0^k}) \leq -\log x_0^k.$$

In fact a direct and more simple proof is available, see Hofbauer (2004):

Proposition

(RP) satifies external consistency.

Proof

By integrating equation (5), one obtains, on the support of x_0:

$$\int_0^t [U_s^k - \langle x_s, U_s \rangle] ds = \int_0^t \frac{\dot{x}_s^k}{x_s^k} ds = \log(\frac{x_t^k}{x_0^k}) \leq -\log x_0^k.$$

■

Back to a game framework this implies that if player 1 follows (*RP*) the set of accumulation points of the empirical correlated distribution process will belong to her reduced Hannan set:

$$\bar{H}^1 = \{ \theta \in \Delta(S); G^1(k, \theta^{-1}) \leq G^1(\theta), \forall k \in S^1 \}$$

with equality for one component.

The example due to Viossat (2007) of a game where the limit set for the replicator dynamics is disjoint from the unique correlated equilibrium shows that (*RP*) does not satisfy internal consistency.

Back to a game framework this implies that if player 1 follows (*RP*) the set of accumulation points of the empirical correlated distribution process will belong to her reduced Hannan set:

$$\bar{H}^1 = \{ \theta \in \Delta(S); G^1(k, \theta^{-1}) \leq G^1(\theta), \forall k \in S^1 \}$$

with equality for one component.
The example due to Viossat (2007) of a game where the limit set for the replicator dynamics is disjoint from the unique correlated equilibrium shows that (*RP*) does not satisfy internal consistency.

We can now compare several processes in the spirit of (payoff based) fictitious play.

The original fictitious play process (I) is defined by

$$x_t \in \mathbf{br}(\bar{U}_t)$$

The corresponding time average satisfies (CFP).

With a smooth best reply process one has (II)

$$x_t = \mathbf{br}^{\varepsilon}(\bar{U}_t)$$

and the corresponding time average satisfies a smooth fictitious play process.

We can now compare several processes in the spirit of (payoff based) fictitious play.
The original fictitious play process (I) is defined by

$$x_t \in \mathbf{br}(\bar{U}_t)$$

The corresponding time average satisfies (CFP).
With a smooth best reply process one has (II)

$$x_t = \mathbf{br}^{\varepsilon}(\bar{U}_t)$$

and the corresponding time average satisfies a smooth fictitious play process.

We can now compare several processes in the spirit of (payoff based) fictitious play.
The original fictitious play process (I) is defined by

$$x_t \in \mathbf{br}(\bar{U}_t)$$

The corresponding time average satisfies (CFP).
With a smooth best reply process one has (II)

$$x_t = \mathbf{br}^{\varepsilon}(\bar{U}_t)$$

and the corresponding time average satisfies a smooth fictitious play process.

Finally the replicator process (III) satisfies

$$x_t = \mathbf{br}^{1/t}(\bar{U}_t)$$

and the time average follows a time dependent perturbation of the fictitious play process.

While in (*I*), the process x_t follows exactly the best reply correspondence, the induced average X_t does not have good unilateral properties.

One the other hand for (*II*), X_t satisfies a weak form of external consistency, with an error term $\alpha(\varepsilon)$ vanishing with ε.

In contrast, (*III*) satisfies exact external consistency due to a both smooth and time dependent approximation of **br**.

While in (*I*), the process x_t follows exactly the best reply correspondence, the induced average X_t does not have good unilateral properties.
One the other hand for (*II*), X_t satisfies a weak form of external consistency, with an error term $\alpha(\varepsilon)$ vanishing with ε.
In contrast, (*III*) satisfies exact external consistency due to a both smooth and time dependent approximation of **br**.

While in (*I*), the process x_t follows exactly the best reply correspondence, the induced average X_t does not have good unilateral properties.
One the other hand for (*II*), X_t satisfies a weak form of external consistency, with an error term $\alpha(\varepsilon)$ vanishing with ε.
In contrast, (*III*) satisfies exact external consistency due to a both smooth and time dependent approximation of **br**.

Bibliography I

- Benaïm M., J. Hofbauer and S. Sorin (2005) Stochastic approximations and differential inclusions, *SIAM J. Opt. and Control*, **44**, 328-348.
- Benaïm M., J. Hofbauer and S. Sorin (2006) Stochastic approximations and differential inclusions. Part II: applications, *Mathematics of Operations Research*, **31**, 673-695.
- Brown G. W.(1951) Iterative solution of games by fictitious play, in Koopmans T.C. (ed.) *Activity Analysis of Production and Allocation* , Wiley, 374-376.

Bibliography II

- CRESSMAN, R. (2003) *Evolutionary Dynamics and Extensive Form Games*, M.I.T. Press.
- FREUND Y. AND R.E. SCHAPIRE (1999) Adaptive game playing using multiplicative weights, *Games and Economic Behavior*, **29**, 79-103.
- FUDENBERG D. AND D. K. LEVINE (1995) Consistency and cautious fictitious play, *Journal of Economic Dynamics and Control*, **19**, 1065-1089.

Bibliography III

- GAUNERSDORFER A. AND J. HOFBAUER (1995) Fictitious play, Shapley polygons and the replicator equation, *Games and Economic Behavior*, **11**, 279-303.
- GILBOA I. AND A. MATSUI (1991) Social stability and equilibrium, *Econometrica*, **59**, 859-867.
- HANNAN J. (1957) Approximation to Bayes risk in repeated plays, *Contributions to the Theory of Games, III*, Drescher M., A.W. Tucker and P. Wolfe eds., Princeton University Press, 97-139.

Bibliography IV

- HART S. (2005) Adaptive heuristics, *Econometrica*, **73**, 1401-1430.
- HART S. AND A. MAS-COLELL (2000) A simple adaptive procedure leading to correlated equilibria, *Econometrica*, **68**, 1127-1150.
- HOFBAUER J. (1995) Stability for the best response dynamics, mimeo.
- HOFBAUER J. (2004) Time averages of the replicator dynamics and correlated equilibria, preprint.

Bibliography V

- HOFBAUER J. AND W. H. SANDHOLM (2002) On the global convergence of stochastic fictitious play, *Econometrica*, **70**, 2265-2294.
- HOFBAUER J. AND K. SIGMUND (1998) *Evolutionary Games and Population Dynamics*, Cambridge U.P.
- HOFBAUER J. AND K. SIGMUND (2003) Evolutionary Game Dynamics, *Bulletin of the A.M.S.*, **40**, 479-519.
- HOFBAUER J. AND S. SORIN (2006) Best response dynamics for continuous zero-sum games, *Discrete and Continuous Dynamical Systems-series B*, **6**, 215-224.

Bibliography VI

- HOPKINS E. (2002) Two competing models of how people learn in games, *Econometrica*, **70**, 2141-2166.
- LITTLESTONE N. AND M.K. WARMUTH (1994) The weighted majority algorithm, *Information and Computation*, **108**, 212-261.
- MAYNARD SMITH J. (1982) *Evolution and the Theory of Games*, Cambridge U.P.
- ROBINSON J. (1951) An iterative method of solving a game, *Annals of Mathematics*, **54**, 296-301.

Bibliography VII

- SORIN S. (2007) Exponential weight algorithm in continuous time, *Mathematical Programming*, Ser. B , DOI: 10.1007/s10107-007-0111-y .
- TAYLOR P. AND L. JONKER (1978) Evolutionary stable strategies and game dynamics, *Mathematical Biosciences*, **40**, 145-156.
- VIOSSAT Y. (2007) The replicator dynamics does not lead to correlated equilibria, *Games and Economic Behavior*, **59**, 397-407