Consistency of Vanishingly Smooth Fictitious play

(In collaboration with Michel Benaïm, UNINE)

Mathieu Faure

Journées MAS, Toulouse, 2014.

Settings

- Player 1 (Decision Maker), Player 2 (Nature, environment),
- finite sets of actions I and L; sets of mixed actions : $X = \Delta(I)$, $Y = \Delta(L)$,

$$X := \{(x_i)_{i \in I},\ x_i \geq 0,\ \sum_i x_i = 1\}.$$

- payoff function of DM : π ($I \times L$ matrix)

Settings

- Player 1 (Decision Maker), Player 2 (Nature, environment),
- finite sets of actions I and L; sets of mixed actions : $X = \Delta(I)$, $Y = \Delta(L)$,

$$X := \{(x_i)_{i \in I},\ x_i \geq 0,\ \sum_i x_i = 1\}.$$

- payoff function of DM : π ($I \times L$ matrix)
- we assume that agents play **repeatedly**. Let $h_n = (i_1, l_1, ..., i_n, l_n)$ history at time n.

Settings

- Player 1 (Decision Maker), Player 2 (Nature, environment),
- finite sets of actions I and L; sets of mixed actions : $X = \Delta(I)$, $Y = \Delta(L)$,

$$X := \{(x_i)_{i \in I},\ x_i \geq 0,\ \sum_i x_i = 1\}.$$

- payoff function of DM : π ($I \times L$ matrix)
- we assume that agents play **repeatedly**. Let $h_n = (i_1, l_1, ..., i_n, l_n)$ history at time n.
- a **strategy** for DM is a map σ :

$$\cup_n \mathcal{H}_n \to \Delta(I),\ \ h_n \mapsto \sigma(h_n) \in \Delta(I).$$

Settings

- Player 1 (Decision Maker), Player 2 (Nature, environment),
- finite sets of actions I and L; sets of mixed actions : $X = \Delta(I)$, $Y = \Delta(L)$,

$$X := \{(x_i)_{i \in I},\ x_i \geq 0,\ \sum_i x_i = 1\}.$$

- payoff function of DM : π ($I \times L$ matrix)
- we assume that agents play **repeatedly**. Let $h_n = (i_1, l_1, ..., i_n, l_n)$ history at time n.
- a **strategy** for DM is a map σ :

$$\cup_n \mathcal{H}_n \to \Delta(I),\ \ h_n \mapsto \sigma(h_n) \in \Delta(I).$$

- A pair of strategies (σ, τ) induces a probability measure $\mathbb{P}$ on $(I \times L)^{\mathbb{N}}$; we assume that agents play **independently** :

$$\mathbb{P}\left(i_{n+1} = i, l_{n+1} = l \mid \mathcal{H}_n\right) = \mathbb{P}\left(i_{n+1} = i \mid \mathcal{H}_n\right) \mathbb{P}\left(l_{n+1} = l \mid \mathcal{H}_n\right).$$

Regret

Empirical distribution of moves and the average realized payoff up to time n :

$$x_n := \frac{1}{n}\sum_{k=1}^{n} \delta_{i_k} \in X, \;\; y_n = \frac{1}{n}\sum_{k=1}^{n} \delta_{l_k} \in Y, \;\; \pi_n := \frac{1}{n}\sum_{k=1}^{n} \pi(i_k, l_k).$$

Define

$$\Pi : Y \to \mathbb{R} : \;\; y \mapsto \max_i \pi(i, y).$$

Regret

Empirical distribution of moves and the average realized payoff up to time n :

$$x_n := \frac{1}{n}\sum_{k=1}^{n}\delta_{i_k} \in X, \;\; y_n = \frac{1}{n}\sum_{k=1}^{n}\delta_{l_k} \in Y, \;\; \pi_n := \frac{1}{n}\sum_{k=1}^{n}\pi(i_k, l_k).$$

Define

$$\Pi : Y \to \mathbb{R} : \;\; y \mapsto \max_i \pi(i, y).$$

Definition (regret at stage n)

$$e_n = \qquad\qquad .$$

Regret

Empirical distribution of moves and the average realized payoff up to time n :

$$x_n := \frac{1}{n}\sum_{k=1}^{n}\delta_{i_k} \in X, \;\; y_n = \frac{1}{n}\sum_{k=1}^{n}\delta_{l_k} \in Y, \;\; \pi_n := \frac{1}{n}\sum_{k=1}^{n}\pi(i_k, l_k).$$

Define

$$\Pi : Y \to \mathbb{R} : \;\; y \mapsto \max_i \pi(i, y).$$

Definition (regret at stage n)

$$e_n = \underbrace{\max_{i \in I} \pi(i, y_n)}_{\Pi(y_n)} \qquad .$$

Regret

Empirical distribution of moves and the average realized payoff up to time n :

$$x_n := \frac{1}{n}\sum_{k=1}^{n}\delta_{i_k} \in X, \;\; y_n = \frac{1}{n}\sum_{k=1}^{n}\delta_{l_k} \in Y, \;\; \pi_n := \frac{1}{n}\sum_{k=1}^{n}\pi(i_k, l_k).$$

Define

$$\Pi : Y \to \mathbb{R} : \;\; y \mapsto \max_i \pi(i, y).$$

Definition (regret at stage n)

$$e_n = \underbrace{\max_{i \in I} \pi(i, y_n)}_{\Pi(y_n)} - \underbrace{\frac{1}{n}\sum_{k=1}^{n}\pi(i_k, l_k)}_{\pi_n}.$$

Consistency

Definition

Player 1's strategy is *consistent* if, regardless of the strategy τ of nature,

$$\limsup_{n \to +\infty} e_n \leq 0, \text{ almost surely.}$$

It is η-consistent provided

$$\limsup_{n \to +\infty} e_n \leq \eta, \text{ almost surely.}$$

Fictitious play

Let **br** be the *best response map* :

$$\mathbf{br} : Y \rightrightarrows X, \;\; y \mapsto Argmax_{x \in X} \pi(x, y)$$

Fictitious play

Let **br** be the *best response map* :

$$\mathbf{br} : Y \rightrightarrows X, \;\; y \mapsto Argmax_{x \in X} \pi(x, y)$$

Definition

The strategy σ of DM is a *fictitious play* strategy if, $\forall n \in \mathbb{N}$,

$$\sigma(h_n) \in \mathbf{br}(\tilde{y}_n),$$

Fictitious play

Let **br** be the *best response map* :

$$\mathbf{br} : Y \rightrightarrows X, \;\; y \mapsto Argmax_{x \in X} \pi(x, y)$$

Definition

The strategy σ of DM is a *fictitious play* strategy if, $\forall n \in \mathbb{N}$,

$$\sigma(h_n) \in \mathbf{br}(\tilde{y}_n),$$

with $\tilde{y}_n = \frac{1}{n+1} \underbrace{y_0}_{\text{prior}} + \frac{n}{n+1} y_n$.

Remark

FP is not consistent

Perturbed payoff function and Smooth best response

Let ε be a small positive parameter

Definition (Perturbed payoff)

The (ρ, ε)-perturbed payoff relative to the original payoff function π is defined by

$$\pi^{\varepsilon}(x, y) = \pi(x, y) + \varepsilon\rho(x),$$

where ρ is **concave** and its gradient explodes at the boundary.

We have :

Perturbed payoff function and Smooth best response

Let ε be a small positive parameter

Definition (Perturbed payoff)

The (ρ, ε)-perturbed payoff relative to the original payoff function π is defined by

$$\pi^{\varepsilon}(x, y) = \pi(x, y) + \varepsilon\rho(x),$$

where ρ is **concave** and its gradient explodes at the boundary.

We have :

- hence for all $y \in Y$, $Argmax_{x\in X}\pi^{\varepsilon}(\cdot, y)$ reduces to one point,
- Thus we can define the smooth best response map $\mathbf{br}^{\varepsilon} : Y \to Int(X)$:

$$\mathbf{br}^{\varepsilon}(y) := Argmax_{x\in X}\pi^{\varepsilon}(x, y).$$

Particular case

Example

$$\rho(x) = -\sum_i x_i \log x_i.$$

Particular case

Example

$$\rho(x) = -\sum_i x_i \log x_i.$$

In that case, we can give an explicit formula for the perturbed best response map :

$$(\mathbf{br}^\varepsilon(y))_i = \frac{\exp\left(\frac{1}{\varepsilon}\pi(i,y)\right)}{\sum_k \exp\left(\frac{1}{\varepsilon}\pi(k,y)\right)}$$

Particular case

Example

$$\rho(x) = -\sum_i x_i \log x_i.$$

In that case, we can give an explicit formula for the perturbed best response map :

$$(\mathbf{br}^\varepsilon(y))_i = \frac{\exp\left(\frac{1}{\varepsilon}\pi(i,y)\right)}{\sum_k \exp\left(\frac{1}{\varepsilon}\pi(k,y)\right)}$$

Definition (Smooth fictitious play)

σ is a smooth fictitious play (SFP(ε)) strategy for player 1 if, for all $n \in \mathbb{N}$

$$\sigma(h_n) = \mathbf{br}^\varepsilon(y_n),$$

Interpretations

- One way to interpret SFP$((\varepsilon))$ strategies is that the agent chooses to randomize his moves, playing the best response to the average moves of the opponent, with respect to a slightly perturbed version of his payoff function ;
- Another possible interpretation of SFP(ε) strategies is that his payoffs are actually perturbed by i.i.d. random shocks (usually called *stochastic fictitious play*).

Smooth fictitious play

Theorem (Fudenberg-Levine)

Given $\eta > 0$, SFP(ε) is η-consistent, provided ε is small enough.

Smooth fictitious play

Theorem (Fudenberg-Levine)

Given $\eta > 0$, SFP(ε) is η-consistent, provided ε is small enough.

Benaïm-Hofbauer-Sorin (2006) gave an alternative proof using stochastic approximations technics : we have $x_{n+1} - x_n = \frac{1}{n+1}\left(\delta_{i_{n+1}} - x_n\right)$. Hence

$$x_{n+1} - x_n = \frac{1}{n+1}\left(\underbrace{\mathbb{E}\left(\delta_{i_{n+1}} \mid \mathcal{H}_n\right)}_{\mathbf{br}^{\varepsilon}(y_n)} - x_n + \underbrace{\left(\delta_{i_{n+1}} - \mathbb{E}\left(\delta_{i_{n+1}} \mid \mathcal{H}_n\right)\right)}_{\text{martingale difference}}\right)$$

$$y_{n+1} - y_n = \frac{1}{n+1}\left(\underbrace{\mathbb{E}\left(\delta_{l_{n+1}} \mid \mathcal{H}_n\right)}_{\tau(h_n)} - y_n + \underbrace{\left(\delta_{l_{n+1}} - \mathbb{E}\left(\delta_{l_{n+1}} \mid \mathcal{H}_n\right)\right)}_{\text{martingale difference}}\right)$$

Stochastic Approximation Algorithms, the ODE method

M compact subset of $\mathbb{R}^d$. Let $(v_n)_n$ be a M-valued stochastic process governed by the recursive formula

$$v_{n+1} - v_n = \frac{1}{n+1}(f(v_n) + U_{n+1}),$$

where

Stochastic Approximation Algorithms, the ODE method

M compact subset of $\mathbb{R}^d$. Let $(v_n)_n$ be a M-valued stochastic process governed by the recursive formula

$$v_{n+1} - v_n = \frac{1}{n+1}(f(v_n) + U_{n+1}),$$

where

- f is a *Lipschitz* vector field, inducing a flow Φ on M,
- $(U_n)_n$ is a *bounded* sequence of random variables.

Stochastic Approximation Algorithms, the ODE method

M compact subset of $\mathbb{R}^d$. Let $(v_n)_n$ be a M-valued stochastic process governed by the recursive formula

$$v_{n+1} - v_n = \frac{1}{n+1}(f(v_n) + U_{n+1}),$$

where

- f is a *Lipschitz* vector field, inducing a flow Φ on M,
- $(U_n)_n$ is a *bounded* sequence of random variables.

Question : can we say anything about the *qualitative* asymptotic behavior of $(v_n)_n$?

Mean ODE

$$v_{n+1} = v_n + \frac{1}{n+1}(f(v_n) + U_{n+1}), \tag{1}$$

Consider the mean ODE :

$$\dot{v} = f(v). \tag{2}$$

Mean ODE

$$v_{n+1} = v_n + \frac{1}{n+1}(f(v_n) + U_{n+1}), \tag{1}$$

Consider the mean ODE :

$$\dot{v} = f(v). \tag{2}$$

Link between (1) and (2) : if $(U_n)_n$ is a martingale difference : $\mathbb{E}(U_{n+1} \mid \mathcal{F}_n) = 0$, the asymptotic behavior of the paths $(v_n(\omega))_n$ should be related to the solution curves of (2) (*ODE method*)

Convergence of Stochastic Approximation Algorithms

Theorem (Limit set theorem, Benaïm, 1996)

$a)$ *The limit set of* $(v_n)_n$ *is almost surely compact convex, invariant and attractor-free,*

$b)$ *if* A *is a global attractor,* $\mathcal{L}((v_n)_n) \subset A$ *almost surely.*

Convergence of Stochastic Approximation Algorithms

Theorem (Limit set theorem, Benaïm, 1996)

$a)$ *The limit set of* $(v_n)_n$ *is almost surely compact convex, invariant and attractor-free,*

$b)$ *if* A *is a global attractor,* $\mathcal{L}((v_n)_n) \subset A$ *almost surely.*

Theorem (Benaïm, Hofbauer and Sorin, 2005)

It also holds when f *is a (reasonably regular) set-valued map.*

Back to SFP(ε)

State variable : $v_n = (x_n, y_n, \pi_n)$.
We have

$$v_{n+1} - v_n \in \frac{1}{n+1}(F^\varepsilon(v_n) + U_{n+1}),$$

where

$$F^\varepsilon(x,y,\pi) = \{(\mathbf{br}^\varepsilon(y_n), \tau, \pi(\mathbf{br}^\varepsilon(y_n), \tau), \ \tau \in Y\} - (x,y,\pi)$$

- The set-valued map F^ε is very regular,
- As a consequence, BHS results apply and, if the differential inclusion $\dot{v}(t) \in F^\varepsilon(v(t))$ admits a global attractor A then $\mathcal{L}((v_n)_n) \subset A$.

Proof of η-consistency via stochastic approximations

Theorem (Benaïm-Hofbauer-Sorin, 2006)

Given $\eta > 0$, for ε small enough we have

- *the set $A := \{v = (x, y, \pi) : \ \Pi(y) - \pi \leq \eta\}$ contains a global attractor for the differential inclusion $\dot{v}(t) \in F^{\varepsilon}(v(t))$,*
- *consequently*

$$\limsup_n \Pi(y_n) - \pi_n \leq \eta \text{ almost surely.}$$

A natural question

What happens when the parameter ε is replaced by a vanishing sequence $\varepsilon_n \downarrow 0$?

A natural question

What happens when the parameter ε is replaced by a vanishing sequence $\varepsilon_n \downarrow 0$?

Definition (Vanishingly Smooth fictitious play)

Given a sequence $\varepsilon_n \downarrow 0$, we say that DM plays accordingly to a *vanishingly smooth fictitious play strategy* (VSFP(ε_n)) if, for all $n \in \mathbb{N}$,

$$\sigma(h_n) = \mathbf{br}^{\varepsilon_n}(y_n).$$

A natural question

What happens when the parameter ε is replaced by a vanishing sequence $\varepsilon_n \downarrow 0$?

Definition (Vanishingly Smooth fictitious play)

Given a sequence $\varepsilon_n \downarrow 0$, we say that DM plays accordingly to a *vanishingly smooth fictitious play strategy* (VSFP(ε_n)) if, for all $n \in \mathbb{N}$,

$$\sigma(h_n) = \mathbf{br}^{\varepsilon_n}(y_n).$$

Remark

VSFP is not consistent, if $\varepsilon_n = \frac{1}{n}$

A counter-Example

Example

2-player matching pennies and that nature uses $(l, r, l, r...)$. $\varepsilon_n = 1/n$, prior $y_0 = (1/3, 2/3)$, then

$$\tilde{y}_{2n} = \frac{1}{2n+1} y_0 + \frac{n}{n+1} y_{2n} = \left(\frac{1}{2} - \frac{1}{6(2n+1)}, \frac{1}{2} + \frac{1}{6(2n+1)} \right).$$

After a few lines of calculus one gets :

$$\mathbf{br}^{\varepsilon_\mathbf{n}}(\tilde{y}_{2n}) \longrightarrow_{n \to +\infty} \left(\frac{1}{2} - c, \frac{1}{2} + c \right).$$

Statement of the main result

Theorem (Benaïm, F.)

If, for some $\alpha < 1$, $\varepsilon_n \geq \frac{1}{n^\alpha}$ then VSFP(ε_n) is consistent.

Statement of the main result

Theorem (Benaïm, F.)

If, for some $\alpha < 1$, $\varepsilon_n \geq \frac{1}{n^\alpha}$ then VSFP(ε_n) is consistent.

- The proof relies on set-valued dynamical systems approach, similarily to BHS,

Statement of the main result

Theorem (Benaïm, F.)

If, for some $\alpha < 1$, $\varepsilon_n \geq \frac{1}{n^\alpha}$ then VSFP(ε_n) is consistent.

- The proof relies on set-valued dynamical systems approach, similarily to BHS,
- unfortunately, we now need to deal with *nonautonomous* differential inclusions,