Fictitious Play and Replicator Dynamic

Duarte Gonçalves

University College London

Topics in Economic Theory

Overview

1. Learning in Games
2. Fictitious Play
3. Stability and Stochastic Approximation
4. Evolutionary Game Theory and Replicator Dynamic

Overview

Learning in Games

How do people get to play equilibrium?

Main question of interest in 'learning in games' ($\neq$ games with learning)

Goals

- Provide foundations for existing equilibrium concepts.
- Capture lab behaviour.
- Predict adjustment dynamics transitioning to new equilibrium. (akin to 'impulse response' in macro; uncommon but definitely worth investigating)
- Select equilibria.
- Algorithm to solve for equilibria.
- Explain persistence of heuristics/nonequilibrium behaviour.

Overview

Fictitious Play

		Col Player	
		L	R
Row Player	L	-1, -1	3, 0
	R	0, 3	0, 0

Suppose play 1st time

Period 0: Say you choose best-response against uniform $a_1 = L$.

Period 1: best-respond against empirical frequency of past play $\bar{\sigma}^1_{-i}(L) = 1$, $\sigma_i^*(\bar{\sigma}^1_{-i}(L)) = R$.

Period 2: best-respond against empirical frequency of past play $\bar{\sigma}^2_{-i}(L) = 1/2$, $\sigma_i^*(\bar{\sigma}^2_{-i}(L)) = L$. etc.

Fictitious Play

(Stage) Game

Player $i \in I$; $-i = I \setminus \{i\}$.

Actions $A = \times_i A_i$. Mixed $\boldsymbol{\sigma}_i \in \boldsymbol{\Sigma}_i = \Delta(A_i)$. Strategies (whenever different) $S = \times_i S_i$.
Abuse notation: $\boldsymbol{\sigma}_{-i} \in \Delta(A_{-i})$ and $\boldsymbol{\sigma}_{-i} \in \times_{j \neq i} \boldsymbol{\Sigma}_j$.

Payoffs $u = (u_i)$, $u_i(a) = u_i(a_i, a_{-i})$ also $u_i(s)$.

Prior $\boldsymbol{\mu}_i \in \times_{j \neq i} \Delta(A_j)$, for each dimension Dirichlet with sum of parameters $\alpha_{i,0}$ and mean $\hat{\boldsymbol{\sigma}}^i_{-i,0}$.

For simplicity, assume that marginals of i, j wrt ℓ are the same.

Correlation vs independence

Different from prior $\boldsymbol{\mu}_i \in \Delta(A_{-i})$, Dirichlet with sum of parameters $\alpha_{i,0}$ and mean $\hat{\boldsymbol{\sigma}}^i_{-i,0}$.

Subjective beliefs about opponents; correlation reflecting strategic uncertainty (Fudenberg and Kreps 1993 GEB).

Correlation *has* important implications for gameplay.

Fictitious Play

Time $t = 0, 1, 2, \ldots$.

Empirical Frequencies $\bar{\sigma}_i^{t+1}(a_i) = \frac{1}{t+1} \sum_{\ell=0}^{t} 1_{\{s_{i,\ell}=a_i\}}$.

Dirichlet Prior + Categorical observations $\implies$ Dirichlet Posterior.

Mean $\hat{\sigma}^i_{j,t+1}(a_j) = \frac{\alpha_{i,t}}{\alpha_{i,t+1}} \hat{\sigma}^i_{j,t}(a_j) + \frac{1}{\alpha_{i,t+1}} 1_{\{a_{j,t}=a_j\}} = \frac{\alpha_{0,i}}{\alpha_{0,i}+t+1} \hat{\sigma}^i_{j,0} + \frac{t+1}{\alpha_{0,i}+t+1} \bar{\sigma}_{j,t}$.

Weights $\alpha_{i,t+1} = \alpha_{i,t} + 1 = \alpha_{i,0} + t + 1$.

Empirical frequencies when $\alpha_{i,0} = 0$ (limit case).

Best response $\arg\max_{\sigma_i} \mathbb{E}_{\sigma_{-i} \sim \mu_{i,t}} u_i(\sigma_i, \sigma_{-i}) = \arg\max_{a_i} u_i(\sigma_i, \hat{\sigma}^i_{-i,t})$.

$\sigma_i^*(\mu_{i,t})$ is selection from best-response correspondence (i.e., fix tie-breaking rule.)

Abuse notation $\sigma_i^*(\hat{\sigma}^i_{-i})$.

Induces dynamics of play: (σ_t).

Comments

Bayesian interpretation of fictitious play.

Players treat environment as stationary, but it's only stationary if start at steady state.

Overview

Convergence of Strategies

Definition

σ_t **converges** if $\exists \sigma_\infty : \sigma_t \to \sigma_\infty$.

σ_t **converges in time average** if $\exists \sigma_\infty : \bar{\sigma}_t \to \sigma_\infty$.

Note: empirical frequencies converge if and only if posterior means converge;

$\bar{\sigma}_t \to \sigma_\infty \iff \hat{\sigma}^i_{-i,t} \to \sigma_{-i,\infty}$ for all i.

Nash Equilibrium as a Limit of Fictitious Play

Proposition 3.0 (Fudenberg and Kreps 1993 GEB)

Suppose that $\boldsymbol{\sigma}$ is a strict NE. Under fictitious play, if $\boldsymbol{\sigma}_t = \boldsymbol{\sigma}$ for some $t > 0$, then $\boldsymbol{\sigma}_{t+h} = \boldsymbol{\sigma}$ for all $h > 0$.

Proof

Suppose $\boldsymbol{\sigma}_t = \boldsymbol{\sigma}$. Then $\forall i, \hat{\sigma}^i_{-i,t} : \sigma_i = \sigma_i^*(\hat{\sigma}^i_{-i,t}) \implies u_i(\sigma_i, \hat{\sigma}^i_{-i,t}) - u_i(\sigma_i', \hat{\sigma}^i_{-i,t}) \geq 0, \forall \sigma_i'$.

Strict NE implies $u_i(\sigma_i, \sigma_{-i}) - u_i(\sigma_i', \sigma_{-i}) > 0 \; \forall \sigma_i'$.

Then, $u_i(\sigma_i, \hat{\sigma}^i_{-i,t+1}) - u_i(\sigma_i', \hat{\sigma}^i_{-i,t+1}) = \frac{1}{\alpha_{t+1}}(u_i(\sigma_i, \sigma_{-i}) - u_i(\sigma_i', \sigma_{-i})) + \frac{\alpha_t}{\alpha_{t+1}}(u_i(\sigma_i, \hat{\sigma}^i_{-i,t}) - u_i(\sigma_i', \hat{\sigma}^i_{-i,t})) > 0$ for all σ_i'.

$\implies \sigma_i^*(\mu_{i,t+1}) = \sigma_i$.

Nash Equilibrium as a Limit of Fictitious Play

Proposition 3.0 (Fudenberg and Kreps 1993 GEB)

Suppose that $\boldsymbol{\sigma}$ is a strict NE. Under fictitious play, if $\boldsymbol{\sigma}_t = \boldsymbol{\sigma}$ for some $t > \mathbf{0}$, then $\boldsymbol{\sigma}_{t+h} = \boldsymbol{\sigma}$ for all $h > \mathbf{0}$.

Proposition 3.1 (Fudenberg and Kreps 1993 GEB)

Under fictitious play, if $\boldsymbol{\sigma}_t = \boldsymbol{\sigma}$ for all but a finite set of periods, then $\boldsymbol{\sigma}$ is a NE.

Proposition 3.2 (Fudenberg and Kreps 1993 GEB)

Under fictitious play, if $\bar{\boldsymbol{\sigma}}_t \to \boldsymbol{\sigma}_\infty$, then product of marginals $(\boldsymbol{\sigma}_{i,\infty}, i \in I)$ is a Nash equilibrium.

Nash Equilibrium as a Limit of Fictitious Play

Proposition 3.2 (Fudenberg and Kreps 1993 GEB)

Under fictitious play, if $\bar{\sigma}_t \to \sigma_\infty$, then product of marginals $(\sigma_{i,\infty}, i \in I)$ is a Nash equilibrium.

Proof

Suppose σ_∞ not NE. Then, $\exists a_i, a_i' : \sigma_i(a_i) > 0$ and $u_i(a_i', \sigma_\infty) > u_i(a_i, \sigma_\infty)$.

Hence, $\exists T > 0 : \forall t > T, u_i(a_i', \hat{\sigma}^i_{-i,t}) > u_i(a_i, \hat{\sigma}^i_{-i,t})$.

This implies that a_i not a best response $\forall t > T$ and so $\forall t > T$ $\sigma_{i,t}(a_i) = 0$.

Finally, $\bar{\sigma}_{i,t}(a_i) = \frac{T}{t}\sigma_{i,T}(a_i) \to 0 < \sigma_\infty(a_i)$, a contradiction.

Overview

(Non)-Convergence of Strategies

		Col Player	
		L	R
Row Player	L	0, 0	1, 1
	R	1, 1	0, 0

Symmetric. $\alpha_0 = 1 + \sqrt{2}, \hat{\sigma}_0(L) = 1/(1 + \sqrt{2})$.

Never indifference (irrational prior mean).

Period 0: $\hat{\sigma}_0(L) = 1/(2 + \sqrt{2}) < 1/2$, both play L.

Period 1: $\hat{\sigma}_1(L) = 2/(2 + \sqrt{2}) > 1/2$, both play R.

Period 2: $\hat{\sigma}_0(L) = 2/(3 + \sqrt{2}) < 1/2$, both play L.

Alternating sequence $(L, L), (R, R), (L, L), (R, R), \ldots$

Insights:

1. Strategies don't converge. σ_t never converges and keeps cycling.
2. Empirical frequencies converge but violate independence assumption. σ_t does converge in time average to σ_∞ and marginals of empirical distribution do converge to a NE (1/2,1/2), but σ_∞ is not NE due to correlation.

Limit Cycles

		Player R		
		a	b	c
	a	1, 0	0, 0	0, 1
Player C	b	0, 1	1, 0	0, 0
	c	0, 0	0, 1	1, 0

Version of Rock-Paper-Scissors from Shapley (1964).

Initial values: $\alpha_{C,0} = \alpha_{R,0} = \sqrt{2}10^{-10}$, $\hat{\sigma}^R_{C,0} = (3/4, 1/4, 0)$, and $\hat{\sigma}^C_{R,0} = (0, 3/4, 1/4)$.

Period 0: $a_t = (a, a)$. 1-period run.

Period 1-2: $a_t = (a, c)$. 1+1=2-period run.

Period 3-6: $a_t = (c, c)$. 2+1+1=4-period run.

Period 7-13: $a_t = (c, b)$. 4+2=6-period run.

Period 13-24: $a_t = (b, b)$. 6+4=10-period run.

Period 25-43: $a_t = (b, a)$. 10+6=16-period run.

Period 44-: $a_t = (a, a)$. 16+10+1=27-period run.

Spend longer playing at the same action profile, but still switch quickly enough.

Also: Jordan (1993 GEB).

Insights:

3. Empirical frequencies may never converge; get stuck in limit cycle.

Non-Convergence

If fictitious play converges, we get a Nash equilibrium. But...

Insights from examples (Jordan 1991 GEB, Shapley 1964)

1. Strategies σ_t don't necessarily converge anc can cycle forever.
2. Empirical frequencies may converge to a but violate independence assumption. σ_t does converge in time average to σ_∞ and marginals of empirical distribution do converge to a NE (1/2,1/2), but σ_∞ is not NE due to correlation.
3. Empirical frequencies may never converge at all, and instead get stuck in limit cycle.

Bad news for learning foundations of NE.

Convergence of Strategies

Proposition 2.3 (Fudenberg and Levine, 1998)

Under fictitious play, if the stage game (i) is 2×2 with 1 or 3 NE (Robinson, 1951 AnnMath), or zero-sum (Miyasawa, 1961 WP), or dominance-solvable (Nachbar, 1951 IJGT), the $\bar{\sigma}_t$ converges to some σ_∞.

We'll see a proof later.

Independence vs Correlation

		Player R		
		a	b	c
	a	1, 0	1/2, 1/2	0, 1
Player C	b	0, 1	1, 0	1/2, 1/2
	c	1/2, 1/2	0, 1	1, 0

New version of Rock-Paper-Scissors from Shapley (1964). Example taken from Fudenberg and Levine (1998, §2.9).

Constant-sum game: empirical frequencies converge and need to converge to NE.

Exact same dynamics as before, but changes how long spend at each profile.

Add 3rd player O that gets to bet on the play:
if chooses In gets 10 if red outcomes arise ((a,b),(b,c),(c,a)), and otherwise gets -1;
if chooses Out gets 0.

Unique correlated equilibrium (hence ! NE): (1/3,1/3,1/3) for R, C, and In for O.

Marginals of empirical frequencies of R and C converge to uniform (1/3,1/3,1/3).

If O's prior doesn't allow for correlation, beliefs converge to (1/3,1/3,1/3) and chooses In. FP converges to NE.

If O's prior does allow for correlation, beliefs converge to assigning prob. zero to red outcomes and chooses Out. FP *converges* to *non*-NE.

Overview

Convergence of Payoffs

Definition

Fictitious play is ε-consistent along a history if $\exists T : \forall t \geq T$, $\bar{u}_{i,t} + \varepsilon \geq \bar{u}^*_{i,t}$, where $\bar{u}_{i,t} := \frac{1}{t}\sum_{\ell=0}^{t-1} u_i(a_{i,\ell}, a_{-i,\ell})$ and $\bar{u}^*_{i,t} := \max_{a_i} u_i(a_i, \bar{\sigma}_{-i,t})$.

Definition

Fictitious play exhibits infrequent switches along a history if $\forall \varepsilon > 0, \exists T : \forall t \geq T$, $\xi_{i,t} \leq \varepsilon$ for all i, where $\xi_{i,t}$ denotes the fraction of player i's action switches by time t, $\xi_{i,t} := \frac{1}{t}\sum_{\ell \leq t} \mathbf{1}_{\{a_{i,t} \neq a_{i,t-1}\}}$.

Proposition 2.4 (Fudenberg and Levine, 1998)

If fictitious play exhibits infrequent switches along a history, then it is ε-consistent along that history, for every $\varepsilon > 0$.

Convergence of Payoffs

Proof

Let $\bar{\sigma}^*_{i,t} := \sigma^*_i(\bar{\sigma}_{-i,t})$. Then,

$$\bar{u}^*_{i,t} = u_i(\bar{\sigma}^*_{i,t}, \bar{\sigma}_{-i,t}) \geq u_i(\bar{\sigma}^*_{i,t+1}, \bar{\sigma}_{-i,t}) = \frac{\alpha_{i,t+1} u_i(\bar{\sigma}^*_{i,t+1}, \bar{\sigma}_{-i,t+1}) - u_i(\bar{\sigma}^*_{i,t+1}, a_{-i,t})}{\alpha_{i,t}}$$

$$\iff \bar{u}^*_{i,t+1} = u_i(\bar{\sigma}^*_{i,t+1}, \bar{\sigma}_{-i,t+1}) \leq \frac{\alpha_{i,t}}{\alpha_{i,t+1}} u_i(\bar{\sigma}^*_{i,t}, \bar{\sigma}_{-i,t}) + \frac{1}{\alpha_{i,t+1}} u_i(\bar{\sigma}^*_{i,t+1}, a_{-i,t}) \leq \cdots$$

$$\leq \frac{1}{\alpha_{i,t+1}} u_i(\bar{\sigma}^*_{i,0}, \bar{\sigma}_{-i,0}) + \frac{1}{\alpha_{i,t+1}} \sum_{\ell=0}^{t} u_i(\bar{\sigma}^*_{i,\ell+1}, a_{-i,\ell})$$

$$\leq \frac{1}{\alpha_{i,t+1}} u_i(\bar{\sigma}^*_{i,0}, \bar{\sigma}_{-i,0}) + \frac{1}{\alpha_{i,t+1}} \sum_{\ell=0}^{t} u_i(\bar{\sigma}^*_{i,\ell}, a_{-i,\ell})$$

$$+ \frac{1}{\alpha_{i,t+1}} \sum_{\ell=0}^{t} [u_i(\bar{\sigma}^*_{i,\ell+1}, a_{-i,\ell}) - u_i(\bar{\sigma}^*_{i,t}, a_{-i,t})]$$

$$\leq \frac{1}{\alpha_{i,t+1}} \bar{u}^*_{i,0} + \frac{t+1}{\alpha_{i,t+1}} \bar{u}_{i,t+1} + \frac{t}{\alpha_{i,t+1}} \|u_i\|_\infty \xi_{i,t} \leq \bar{u}_{i,t+1} + \varepsilon_{i,t}, \qquad \varepsilon_{i,t} \downarrow 0$$

Convergence of Payoffs

Proposition 2.4 (Fudenberg and Levine, 1998)

If fictitious play exhibits infrequent switches along a history, then it is ε-consistent along that history, for every $\varepsilon > 0$.

Proposition 2.5 (Fudenberg and Levine, 1998)

$\exists \varepsilon_{i,t} \downarrow 0$ s.t. $\bar{u}^*_{i,t} \geq \bar{u}_{i,t} + \varepsilon_{i,t}$.

Note however, that $\bar{u}_t$ needn't converge to NE payoffs....

Convergence of Payoffs

		Col Player	
		L	R
Row Player	L	0, 0	1, 1
	R	1, 1	0, 0

Symmetric. $\alpha_0 = 1 + \sqrt{2}$, $\hat{\sigma}_0(L) = 1/(1 + \sqrt{2})$.

Never indifference (irrational prior mean).

- Period 0: $\hat{\sigma}_0(L) = 1/(2 + \sqrt{2}) < 1/2$, both play L.
- Period 1: $\hat{\sigma}_1(L) = 2/(2 + \sqrt{2}) > 1/2$, both play R.
- Period 2: $\hat{\sigma}_0(L) = 2/(3 + \sqrt{2}) < 1/2$, both play L.
- Alternating sequence $(L, L), (R, R), (L, L), (R, R), \ldots$
- σ_t never converges and keeps cycling, but it converges in time average.

Already seen this example, but note: payoffs are always exactly zero < payoffs from any NE.

Overview

Extensions of Fictitious Play

History $h_t = (a_\ell, \ell \leq t-1)$, $a_t = (a_{i,t}, i \in I)$. Histories $H_t = A^{t-1}$, $H_1 = \{\emptyset\}$, $H = \cup_t H_t$.

Behavioural strategy: $b_i = (b_{i,1}, b_{i,2}, \ldots)$ where $b_{i,t} : H_t \to \Sigma_i$.

Assessment at t $\hat{\sigma}_{-i,t} : H_t \to \Sigma_{-i}$.
(Think $\hat{\sigma}_{-i,t}(h_t) = \mathbb{E}_{\mu_i}[\sigma_{-i} \mid h_t]$. Also: note implied independence.)

Definition

Given assessments $(\hat{\sigma}_{-i,t})$, b_i is **myopic** relative to $(\hat{\sigma}_{-i,t})$ if for every t and h_t, $b_{i,t}(h_t) \in \arg\max_{\sigma_i} u_i(\sigma_i, \hat{\sigma}_{-i,t}(h_t))$.
b_i is **asymptotically myopic** if for some $\varepsilon_{i,t} \downarrow 0$, for every t and h_t, $b_{i,t}$ is $\varepsilon_{i,t}$-optimal, i.e., $u_i(b_{i,t}(h_t), \hat{\sigma}_{-i,t}(h_t)) \geq \max_{\sigma_i} u_i(\sigma_i, \hat{\sigma}_{-i,t}(h_t))$.
b_i is **strongly asymptotically myopic** if for any selection $s_{i,t}(h_t) \in \operatorname{supp} b_{i,t}(h_t)$ of its support, $s_{i,t}$ is asymptotically myopic.

Myopia for repeated interaction hard to rationalise for small groups.

Large group interpretation: large populations of players with anonymous rematching (one each period + observable actions, all each period + aggregate statistics, different populations each period).

Extensions of Fictitious Play

Definition

Assessments $(\hat{\sigma}_{-i,t})$ are **adaptive** if $\forall \varepsilon > 0$ and t, there is $T(\varepsilon, t)$ s.t. $\forall t' > T(\varepsilon, t)$ and histories $h_{t'}$, $\hat{\sigma}_{-i,t'}(h_{t'})(a_{-i}) \leq \varepsilon$ for all pure strategies a_{-i} not played between t and t'.

Definition

Assessments $(\hat{\sigma}_{-i,t})$ are **asymptotically empirical** if $\lim_{t\to\infty} \|\hat{\sigma}_{-i,t}(h_t) - \bar{\sigma}_{-i,t}(h_t)\| = 0$.

Assign low probability to strategies not chosen for a long enough time.

Fictitious play has adaptive, asymptotically empirical assessments and is myopic.

Nash Equilibrium as a Limit of Extended Fictitious Play

Proposition 4.1 (Fudenberg and Kreps 1993 GEB)

Under adaptive assessments and strongly asymptotically myopic behaviour, if $a_t = a$ for all $t \geq T$, then a is a Nash equilibrium.

Proposition 4.2 (Fudenberg and Kreps 1993 GEB)

Under asymptotically empirical assessments and strongly asymptotically myopic behaviour, if empirical frequencies converge to $\boldsymbol{\sigma}_\infty$, then product of marginals $(\boldsymbol{\sigma}_{i,\infty}, i \in I)$ is a Nash equilibrium.

Overview

Potential Games

Cournot competition

Firms $i = 1, \dots, n$ choose $q_i \geq 0$; $Q = \sum_i q_i$.

Profit $u_i(q) = F(Q)q_i - c_i(q_i)$.

Two canonical potentials (Monderer and Shapley 1996 GEB):

Symmetric linear cost $c_i(q_i) = cq_i$: ordinal potential

$$P(q) = \left(\prod_{j=1}^{n} q_j\right)\left(F(Q) - c\right).$$

For each i and fixed q_{-i}, $u_i(q_i, q_{-i}) > u_i(q_i', q_{-i}) \iff P(q_i, q_{-i}) > P(q_i', q_{-i})$.

Linear inverse demand $F(Q) = a - bQ$, arbitrary c_i differentiable: exact potential

$$P^*(q) = a\sum_{j=1}^{n} q_j - b\sum_{j=1}^{n} q_j^2 - b\sum_{1\leq i<j\leq n} q_i q_j - \sum_{j=1}^{n} c_j(q_j).$$

Then, for each i and fixed q_{-i},

$$u_i(q_i, q_{-i}) - u_i(q_i', q_{-i}) = P^*(q_i, q_{-i}) - P^*(q_i', q_{-i}).$$

Consequences

Maximiser(s) of P^* are (pure) Cournot equilibria.

For the ordinal potential P, the set of pure equilibria is unchanged.

Potential Games

Definition (exact/ordinal/weighted potential game)

A game admits

(i) an **exact potential** $\Phi : A \to \mathbb{R}$ if for all i, a_{-i} and a_i, a_i',
$$u_i(a_i, a_{-i}) - u_i(a_i', a_{-i}) = \Phi(a_i, a_{-i}) - \Phi(a_i', a_{-i}).$$

(ii) an **ordinal potential** Φ if for all i, a_{-i} and a_i, a_i',
$$u_i(a_i, a_{-i}) > u_i(a_i', a_{-i}) \iff \Phi(a_i, a_{-i}) > \Phi(a_i', a_{-i}).$$

(iii) a w**/weighted-potential** Φ if for all i, a_{-i} and a_i, a_i',
$$u_i(a_i, a_{-i}) - u_i(a_i', a_{-i}) = w_i\big(\Phi(a_i, a_{-i}) - \Phi(a_i', a_{-i})\big), \text{ for some } w = (w_i)_{i \in I} \text{ with } w_i > 0.$$

Interpretation: w_i rescales player i's payoff units so that all players share the same potential scale.

Any exact potential is a w-potential with $w_i \equiv 1$; conversely, divide u_i by w_i to obtain an exact potential game with potential Φ.

Potential Games and Equilibria

Proposition (Lemma 2.1 and Corollary 2.2 Monderer and Shapley 1996 GEB)

Let Φ be an ordinal potential for a finite game (A_i, u_i). Then:

a is (pure) NE iff for every i, $\Phi(a) \geq \Phi(a_{-i}, a_i')$ for all $a_i' \in A_i$.

Hence, every finite ordinal potential game has a pure-strategy equilibrium.

Proof

If a is NE, no profitable unilateral deviation $\implies$ no unilateral move raises ordinal Φ.

If no unilateral move raises Φ, ordinality $\implies$ no profitable deviation $\implies$ a is NE.

Finite A $\implies$ Φ attains a maximiser $\hat{a}$; (1) yields that $\hat{a}$ is NE.

Fictitious Play Property

Definition (Fictitious Play Property)

A game has **fictitious play (FP) property** if, for all initial beliefs, any limit point of fictitious play empirical frequencies is NE.

Theorem 2.4 (Monderer and Shapley 1996 GEB; see also Monderer and Shapley 1996 JET)

Every finite *weighted potential game* has fictitious play property. In particular, finite identical-interest games ($u_i \equiv U$) have FP property and converge to maximisers of U.

Implications for convergence of FP

Empirical frequencies converge to the NE set; in identical interests, play settles at a pure maximiser of U under fixed tie-breaking.

Selection: among multiple maximisers, FP path/tie-breaking determines limiting equilibrium.

Finite Improvement Property

Definition (Finite Improvement Property)

A finite game has **FIP** if every path of unilateral *strict* improvements is finite.

Definition (generalised ordinal potential)

A function Φ is a **generalised ordinal potential** if for any unilateral move $a \to a' = (a'_i, a_{-i})$, $u_i(a') > u_i(a) \implies \Phi(a') > \Phi(a)$.

Proposition (Lemma 2.5 and Corollary 2.6 Monderer and Shapley 1996 GEB)

For finite games:

(i) FIP $\iff$ existence of generalised ordinal potential.

(ii) If, in addition, for all i and a_{-i} we have $u_i(a_i, a_{-i}) \neq u_i(a'_i, a_{-i})$ whenever $a_i \neq a'_i$, then existence generalised ordinal potential $\implies$ existence ordinal potential.

Finite Improvement Property

Proposition (Lemma 2.5 and Corollary 2.6 Monderer and Shapley 1996 GEB)

For finite games:

(i) FIP $\iff$ existence of generalised ordinal potential.

(ii) If, in addition, for all i and a_{-i} we have $u_i(a_i, a_{-i}) \neq u_i(a_i', a_{-i})$ whenever $a_i \neq a_i'$, then existence generalised ordinal potential $\implies$ existence ordinal potential.

Proof Idea

(FIP $\implies$ GOP) Define $x > y$ if there exists a finite improvement path from y to x. FIP $\implies$ transitive, acyclic. Extend to a ranking Φ with $x > y \implies \Phi(x) > \Phi(y)$.

(FIP $\impliedby$ GOP) Any strict improvement increases Φ; finiteness of the state space forbids infinite ascent.

Under no-indifference, u_i-increase $\iff$ Φ-increase $\implies$ Φ is ordinal.

Potential Games: Best-Reply Dynamics

Proposition (Lemma 2.5 and Corollary 2.6 Monderer and Shapley 1996 GEB)

For finite games:

(i) FIP $\iff$ existence of generalised ordinal potential.

(ii) If, in addition, for all i and a_{-i} we have $u_i(a_i, a_{-i}) \neq u_i(a'_i, a_{-i})$ whenever $a_i \neq a'_i$, then existence generalised ordinal potential $\implies$ existence ordinal potential.

Interpretation: In any finite ordinal potential game, every sequence of strict better replies is finite and terminates at a (pure) NE.

Potential Games: Logit QRE

Logit (quantal response) choice: $\sigma_i(a_i) \propto \exp\left(u_i(a_i, \sigma_{-i})/\lambda\right)$; $\lambda > 0$ noise.

Fact: In potential games, logit equilibria coincide with maximisers of

$$\Psi(\sigma) = \mathbb{E}_\sigma[\Phi(a)] + \lambda \sum_i H(\sigma_i),$$

where H is Shannon entropy.

Implication: as $\lambda \downarrow 0$, stationary points select maximisers of Φ; connects to stochastic fictitious play.

Overview

Brown's Original Fictitious Play (Berger 2007 JET)

Key point: what we called fictitious play is *simultaneous* FP (SimFP); Brown (1951) analysed *alternating* FP (AltFP).

AltFP: players update and best respond *alternately*; gives clean convergence in a broad class.

Main result: in nondegenerate ordinal potential games, AltFP converges to a (pure) NE.

Setup and Notation

Players $i \in \{1, 2\}$; actions A_i (finite).

Payoffs $u_i : A \to \mathbb{R}$; mixed strategies $\sigma_i \in \Delta(A_i)$.

Best replies: $\mathrm{BR}_1(\sigma_2) = \arg\max_{a_1 \in A_1} u_1(a_1, \sigma_2)$, $\mathrm{BR}_2(\sigma_1) = \arg\max_{a_2 \in A_2} u_2(\sigma_1, a_2)$.

NE $\sigma^* = (\sigma_1^*, \sigma_2^*)$: if $\sigma_1^*(a_1) > 0$ then $a_1 \in \mathrm{BR}_1(\sigma_2^*)$, and symmetrically for player 2.

Nondegeneracy and Improvement Steps

Definition (Nondegenerate Two-Player Game)

The game is **nondegenerate** if for each player i and any two distinct actions $a_i \neq a_i'$, we have $u_i(a_i, a_{-i}) \neq u_i(a_i', a_{-i})$ for every opponent's pure action $a_{-i} \in A_{-i}$.

Definition (improvement step/path/cycle; FIP)

$a \to a'$ is an **improvement step** if a' differs from a in one player's action and deviator's payoff strictly increases.
An **improvement path** is chain of such steps; a cycle returns to its start.
The game has **finite improvement property (FIP)** if there are no cycles.

Nondegenerate two-player games: ordinal potential $\iff$ FIP
(Monderer and Shapley 1996 GEB).

Alternating vs Simultaneous Fictitious Play

Definition (empirical beliefs)

Let $a_t = (a_{1,t}, a_{2,t})$ be play in period $t \in \{1, 2, \ldots\}$. Empirical frequencies are

$$\bar{\sigma}_{i,t+1}(a_i) = \frac{t}{t+1}\,\bar{\sigma}_{i,t}(a_i) + \frac{1}{t+1}\,1_{\{a_{i,t}=a_i\}}, \qquad i = 1, 2,$$

with some initial $\bar{\sigma}_{i,0} \in \Delta(A_i)$.

Definition (AltFP and SimFP)

AltFP (alternating FP): players best respond in turn using current empirical belief:
for odd t, $a_{1,t} \in \mathrm{BR}_1(\bar{\sigma}_{2,t})$ and $a_{2,t} = a_{2,t-1}$; for even t, $a_{2,t} \in \mathrm{BR}_2(\bar{\sigma}_{1,t})$ and $a_{1,t} = a_{1,t-1}$.

SimFP (simultaneous FP): at each t, both choose $a_{i,t} \in \mathrm{BR}_i(\bar{\sigma}_{-i,t})$ simultaneously.

Improvement Lemma

Lemma (Berger 2007 JET, Lemma 9)

In a nondegenerate game, if an AltFP process switches from $a = (a_1, a_2)$ to $a' = (a_1', a_2')$ at time t, then there exists an improvement path from a to a'.

Proof

Suppose t is odd. Then $a_1' \in \mathrm{BR}_1(\bar{\sigma}_{2,t})$ and $a_1 \in \mathrm{BR}_1(\bar{\sigma}_{2,t-1})$;
note $\bar{\sigma}_{2,t}(a_2) = \frac{t-1}{t}\,\bar{\sigma}_{2,t-1}(a_2) + \frac{1}{t}\,\mathbf{1}_{\{a_{2,t-1}=a_2\}}$.

Hence $u_1(a_1', a_{2,t-1}) \geq u_1(a_1, a_{2,t-1})$, and by nondegeneracy the inequality is strict when a switch occurs: $(a_1, a_2) \to (a_1', a_2)$.

Even t gives analogous step for player 2.
Concatenating yields an improvement path from a to a'.

Convergence of AltFP

Theorem (Berger 2007 JET, Theorem 10)

In nondegenerate ordinal potential games, every AltFP process converges to a (pure) Nash equilibrium.

Proof

Suppose not. Then AltFP switches infinitely often. Finiteness of A implies some profiles recur infinitely many times.

Lemma $\implies$ repeated switches among recurrent profiles generate improvement cycle.

Contradiction with FIP (equivalently, existence of an ordinal potential). Hence convergence to a pure NE.

Discussion and Contrast with SimFP

AltFP tracks improvement paths; ordinal potentials forbid cycles $\implies$ convergence.

SimFP may fail to track improvement paths when both players switch; global convergence under SimFP is not guaranteed in the ordinal class.

Weighted/identical-interest potential games: FP property also for SimFP (earlier slides); AltFP extends convergence to broader ordinal class under nondegeneracy.

Overview

Supermodular Games

Definition (Games with Strategic Complementarities; Section 5, Milgrom and Shannon 1994 Ecta)

A normal-form game $\Gamma = (I, A, u)$ has **(ordinal) strategic complementarities** if $\forall i$:

- A_i is non-empty compact complete lattice (order $\leq$). Write $A := \times_i A_i$ and $a = (a_i, a_{-i})$.
- u_i is upper semicontinuous in a_i for fixed a_{-i}, and continuous in a_{-i} for fixed a_i.
- u_i is quasisupermodular in a_i and has the single-crossing property in $(a_i; a_{-i})$.

Consequences (under the above): (i) best reply $B_i(a_{-i}) := \arg\max_{a_i \in A_i} u_i(a_i, a_{-i})$ is nonempty and nondecreasing; (ii) equilibrium set is non-empty complete lattice.

Smooth special case: if $A_i \subseteq \mathbb{R}$ and $u_i \in C^2$, single crossing in $(a_i; a_{-i})$ is implied by $\partial^2 u_i / \partial a_i \partial a_j \geq 0$ for $j \neq i$.

Extremal equilibria notation: $a_* = (a_{i,*})_{i \in I}$ (least), $a^* = (a_i^*)_{i \in I}$ (greatest).

Supermodular Games

Theorem 12 (Milgrom and Shannon 1994 Ecta)

In games with strategic complementarities, for each player i there exist smallest and largest serially undominated strategies $a_{i,*}$ and a_i^*. Profiles $a_* = (a_{i,*})_i$ and $a^* = (a_i^*)_i$ are (pure) Nash equilibria.

Theorem 13 (Milgrom and Shannon 1994 Ecta)

For parameterised family with strategic complementarities and $u_i(a_i, a_{-i}, \theta)$ having single-crossing property in $(a_i; a_{-i}, \theta)$, smallest and largest NE $a_*(\theta)$ and $a^*(\theta)$ are nondecreasing in t.

Proof Sketch (Refresher)

Monotone $B_i \implies$ isotone $B(a) := \times_i B_i(a_{-i})$ on product lattice.

Tarski $\implies$ complete lattice of fixed points; least/greatest fixed points a_*, a^* exist.

Single crossing in $(\cdot; t) \implies$ monotone comparative statics of a_*, a^* in θ.

Supermodular Games

Definition (Adaptive Learning)

Sequence of actions $\{a_t\}_t$ is **consistent with adaptive learning** if for all $\varepsilon > 0$ and all dates T there exists $T' > T$ such that for all $t > T'$ and all i and $a_i, a_i' \in A_i$,

$$u_i(a_i, a_{-i,s}) + \varepsilon < u_i(a_i', a_{-i,s}) \quad \forall s \in [T, t] \implies a_{i,t} \neq a_i.$$

Equivalently: eventually, each player avoids any strategy that is strictly and uniformly outperformed on the window $[T, t]$ by some alternative against every observed $a_{-i,s}$.

Theorem 14 (Milgrom and Shannon 1994 Ecta)

In finite games with strategic complementarities, if $\{a_t\}$ is consistent with adaptive learning, then for all large t, $a_* \leq a_t \leq a^*$. In finite or infinite games, if pure NE is unique, then $\{a_t\}$ is consistent with adaptive learning iff $a_t \to a^*$.

Implications for fictitious play

FP-type rules eliminate uniformly inferior actions on sliding windows $\implies$ consistency.

Hence a_t enters the equilibrium band $[a_*, a^*]$; uniqueness $\implies$ convergence.

Connection with fictitious play

FP uses empirical beliefs; persistently inferior actions vanish under window rules $\implies$ adaptive-learning consistency.

Strategic complementarities yield order-preserving best replies $\implies$ convergence to $[a_*, a^*]$; unique equilibrium $\implies$ convergence to that equilibrium.

With extremal tie-breaking, monotone best-reply dynamics converge to a_* from below and to a^* from above.

Applications

Bertrand with differentiated products

Prices $p_i \in A_i \subseteq \mathbb{R}$; demands $D_i(p)$ decreasing in p_i; $\log D_i$ has SCP in $(p_i; p_{-i})$.

Profit $u_i(p) = (p_i - c_i)D_i(p)$ is quasisupermodular in p_i and has SCP in $(p_i; p_{-i})$.

Extremal equilibria p_*, p^* exist; upward demand shift $\theta \implies p_*(\theta), p^*(\theta)$ nondecreasing. Adaptive pricing $\to [p_*, p^*]$; uniqueness $\implies p_t \to p^*$.

Cournot with network effects

Outputs $q_i \in \mathbb{R}_+$; inverse demand $P(Q, \theta)$ with $\partial^2 P / \partial Q \partial \theta \geq 0$.

$u_i(q) = q_i P(Q, \theta) - C_i(q_i)$ has single crossing in $(q_i; (q_{-i}, \theta))$.

Extremal equilibria $q_*(\theta), q^*(\theta)$ monotone in θ; adaptive learning $\to [q_*, q^*]$.

Applications

Network adoption

$a_i \in \{0, 1\}$; payoff increasing in neighbours' adoption $\sum_{j \neq i} a_j \implies$ single crossing in $(a_i; \sum_{j \neq i} a_j)$ and quasisupermodularity in a_i.

Extremal equilibria a_*, a^*; subsidies shift thresholds monotonically; adaptive diffusion converges within $[a_*, a^*]$.

Gross substitutes exchange (market-maker game)

Prices as actions; gross substitutes: own-demand decreases in own price and is nondecreasing in other prices.

Single crossing holds; extremal price equilibria p_*, p^*; tatonnement-like adaptive rules converge into $[p_*, p^*]$; endowment shifts move p_*, p^* up.

Overview

Fictitious Play in Extensive-Form Games

Main punchline: with (generalised) fictitious play + only learn about reached information sets, stable steady states 'if and only if' self-confirming equilibria.

For rulying out non-NE steady states, need 'enough' experimentation.

Great summary: Fudenberg and Levine (1998 Chs. 6-7).

Key references: Fudenberg Kreps 1988 WP, 1995 GEB; Fudenberg and Levine 1993a Ecta, 1993b Ecta, 1998.

Overview

Overview

1. Learning in Games
2. Fictitious Play
3. Stability and Stochastic Approximation
   - Stability
   - Stochastic Approximation
   - Stochastic Fictitious Play
4. Evolutionary Game Theory and Replicator Dynamic

Definitions of Stability

$X \subseteq \mathbb{R}^n; F : \mathbb{R}^n \to \mathbb{R}^n$.

Dynamic system: discrete time $x_{t+1} = F(x_t)$; or continuous time $\dot{x} = F(x)$.

Steady state: discrete time $x = F(x)$; continuous time $F(x) = \mathbf{0}$.

Invariant set: $\forall x_0 \in S, x_t \in S \forall t$. Discrete time $F(S) \subseteq S$.

Periodic point: $\exists t : x_t = x_0 = x$.

Definitions of Stability

Definition

(Lyapunov) stable point x^*: $\forall \varepsilon > 0, \exists \delta > 0$ s.t. $x_0 \in N_\delta(x^*) \implies x_t \in N_\varepsilon(x^*) \forall t$.

Stable set of steady state x^*: $W(x^*) = \{x | x_0 = x, \ \lim_{t \to \infty} x_t = x^*\}$.

Locally asymptotically stable x^* (or sink/attractor): steady state x^* : $\exists \varepsilon > 0$ s.t. $N_\varepsilon(x^*) \subseteq W(x^*)$.

If start near enough steady state, will converge to steady state. Or: if perturb steady state, will return to it.

Globally asymptotically stable x^*: steady state x^* : $W(x^*) = X$.

Converge to x^* no matter where one starts from.

Unstable x^*: not stable point.

Repeller x^*: $\exists \varepsilon > 0$ s.t. $\forall x_0 \neq x^*, \exists t : x_t \notin N_\varepsilon(x^*)$.

Always move away from x^*.

Stability Criteria

Discrete-Time Stability

Steady state x^* is **locally asymptotically stable** if spectral radius (largest eigenvalue in absolute terms) of Jacobian of F at x^* is strictly smaller than 1, $\rho(J_F(x^*)) < \mathbf{1}$.
If, moreover, (i) F is linear or (ii) $X \subseteq \mathbb{R}$ and $|F'| < \mathbf{1}$, then x^* is **globally asymptotically stable**.

Continuous-Time Stability

Steady state x^* is **locally asymptotically stable** if all eigenvalues of the Jacobian of F at x^*, $J_F(x^*)$, have strictly negative real parts.
If, moreover, (i) F is linear or (ii) $X \subseteq \mathbb{R}^2$ and all eigenvalues of the Jacobian of F are have strictly negative real parts at any $x \in X$, then x^* is **globally asymptotically stable**.

Lyapunov Function

Lyapunov Direct Method

For $\dot{x} = F(x)$ and steady state x^*, V is **Lyapunov function** for x^* if for a neighbourhood U of x^*:

(1) $V(x) > V(x^*)$ for $x \in U \setminus \{x^*\}$ and $V(x^*) = \mathbf{0}$; and

(2) $\dot{V}(x) := \nabla V(x) F(x) \leq \mathbf{0}$ for $x \in U$.

Theorem

(i) If x^* is a steady state and there is a Lyapunov function V for x^*, then x^* is **Lyapunov stable**.

(ii) If, in addition to (i), $\dot{V} < \mathbf{0}$, then x^* is **locally asymptotically stable**.

(iii) If, in addition to (i) and (ii), V is a Lyapunov function for x^* with conditions (1) and (2) satisfied for $U = X$ and (a) V is radially unbounded, $V(x) \to \infty$ as $\|x\| \to \infty$, or (b) X is bounded, then x^* is **globally asymptotically stable**.

Overview

Relating Discrete- and Continuous-Time Systems

For FP, we used something like

$$\bar{\sigma}_{t+1} = \frac{\alpha_t}{\alpha_{t+1}}\bar{\sigma}_t + \frac{1}{\alpha_{t+1}}\sigma^*(\bar{\sigma}_t).$$

In many papers and textbooks you'll see a "continuous-time approximation"

$$\dot{\bar{\sigma}}_t = \sigma^*(\bar{\sigma}_t) - \bar{\sigma}_t.$$

Where is this coming from? When is this okay?

Stochastic Approximation

Definition

Let the **discrete-time system** be given by $x_{t+1} = x_t + a_t(F(x_t) + M_{t+1} + e_t)$ where

(1) $F : \mathbb{R}^n \to \mathbb{R}^n$ is Lipschitz continuous;

(2) a_t satisfy $\sum_t a_t = \infty$, $\sum_t a_t^2 < \infty$,

(3) M_t is a martingale wrt $\mathcal{F}_t = \sigma(\{x_s, M_s, s < t\})$, i.e., $\mathbb{E}[M_{t+1}|\mathcal{F}_t] = 0$ a.s. and $\mathbb{E}[\|M_{t+1}\|^2 \mid \mathcal{F}_t] \leq K(1 + \|x_t\|^2)$ a.s.

(4) $\sup_t \|x_t\| < \infty$ a.s.; and

(5) e_t is a deterministic or random bounded sequence with is $o(1)$.

Let the (associated) **continuous-time system** be given by $\dot{x}_t = F(x_t)$.

Define $\Phi_t : \mathbb{R}^n \to \mathbb{R}^n$ denote the map taking x_0 to x_t according to the continuous-time system.

F Lipschitz $\implies \forall t > 0$, Φ_t Lipschitz and, indeed, homeomorphism (continuous bijection with continuous inverse).

Stochastic Approximation

Definition

Closed set $A \subseteq \mathbb{R}^n$ is

(i) **invariant** if $x_0 \in A \implies x_t \in A$ for all t, i.e., $\Phi_t(A) = A\ \forall t$; and

(ii) **internally chain transitive** if it is invariant and, for any $x, y \in A$ and any $\varepsilon > 0$, $T > 0$, $\exists n \geq 1$ and elements $x_0 = x, x_1, ..., x_{n-1}, x_n = y$ in A such that $\forall t < T$, $\Phi_t(x_i) \in B_\varepsilon(x_{i+1})$, i.e., the trajectory initiated at x_i stays within ε-neighbourhood of x_{i+1}.

Elements of internally chain transitive set can be connected by sequence of points ("chain") within that set which can be made arbitrarily close to actual system's trajectories.

Stable points are internally chain transitive sets; but limit cycles also define internally chain transitive sets...

Stochastic Approximation

Theorem (Borkar 2008 Theorem 2; Benaim 1996)

(1) Sequence $\{x_t\}$ generated by dynamic-time system converges a.s. to compact, connected and internally chain transitive invariant set of continuous-time dynamic system.

(2) Furthermore, if internally chain transitive invariant sets correspond to isolated equilibrium points, then $\{x_t\}$ a.s. converges to (possibly sample path-dependent) steady state.

Corollary

Let $W(x^*) := \{x \in X | \lim_{t\to\infty} \Phi_t(x) = x^*\}$ and suppose $\exists$ finite X^* s.t. $X = \cup_{x^* \in X^*} W(x^*)$. Then sequence $\{x_t\}$ generated by dynamic-time system converges a.s. to some $x^* \in X^*$.

Incredibly useful: can learn about discrete time asymptotic behaviour via continuous-time system.

Also allows for further convergence results via stochastic FP.

Stochastic Approximation

What about FP? F is not continuous there...

Can relax F continuous with X compact, modulo dealing adequately with tie-breaking. Bottom line: it works.

Can use continuous-time approx. of FP ($\dot{\bar{\boldsymbol{\sigma}}}_t = \boldsymbol{\sigma}^*(\bar{\boldsymbol{\sigma}}_t) - \bar{\boldsymbol{\sigma}}_t$) to examine asymptotic behaviour!

Technicalities follow:

- Replace F with Filippov convexification $G(x) := \cap_{\delta>0}\overline{\text{co}}(F(B_\delta(x) \setminus N))$, where N is a set of measure zero containing the discontinuities of F.
- Study *differential inclusion* $\dot{x}_t \in G(x_t)$.
- See Benaim, Hofbauer, Sorin (2005 SIAMCON, 2006 MOR).

Overview

Stochastic Fictitious Play

Motivation for stochastic fictitious play (Fudenberg and Levine 2016 JEL): Exact best response of fictitious play

(i) easier to exploitation by a clever opponent (Blackwell 1956; Fudenberg and Kreps 1993)

(ii) small change in beliefs can lead to discontinuous change in response probabilities (not good for stochastic approximation).

Stochastic Fictitious Play

APU: Additive iid perturbed payoffs $\eta_{i,t}$, absolutely continuous and full support on real line; $q_i(\sigma_{-i})(a_i) := \mathbb{P}(a_i \in \arg\max_{a_i'} u_i(a_i', \sigma_{-i}) + \lambda\eta_{i,t}(a_i'))$.

Logit QR: $q_i(\sigma_{-i})(a_i) := \frac{\exp(u_i(a_i,\sigma_{-i})/\lambda)}{\sum_{a_i'} \exp(u_i(a_i',\sigma_{-i})/\lambda)}$. $\lambda \downarrow 0$, convergence to NE.

Gives rise to $\bar{\sigma}_{t+1} = \bar{\sigma}_t + \frac{1}{\alpha_{t+1}} Q(\bar{\sigma}_t) + M_{t+1} + e_t$,

with Q smooth over compact set (hence Lipschitz), $M_{t+1} = \frac{1}{\alpha_{t+1}}(a_{t+1} - Q(\bar{\sigma}_t))$ (as $a_{t+1} \sim Q(\bar{\sigma}_t)$), and

$e_t = -\frac{1}{\alpha_{t+1}}\bar{\sigma}_t +$ differences arising from considering empirical frequencies rather than posterior beliefs (vanishing with large t).

Note Q inherits nice properties (**regular quantal response function**; Goeree, Holt, and Palfrey 2005 EE):

Monotonicity: $u_i(a_i, \sigma_{-i}) > u_i(a_i', \sigma_{-i}) \implies q_i(\sigma_{-i})(a_i) > q_i(\sigma_{-i})(a_i')$.

Responsiveness: $\frac{\partial}{\partial u_i(a_i,\sigma_{-i})} q_i(\sigma_{-i})(a_i) > 0$.

Interiority: $1 > q_i(\sigma_{-i}) > 0$.

(Lipschitz) continuity in σ_{-i}.

Stochastic Fictitious Play

Some results in the literature using SFP:

(1) Fudenberg and Kreps (1993 GEB): If ! NE in 2×2 game global, then for any η, as $\lambda \downarrow 0$ limit of SFP empirical frequencies converges to ! NE.

(2) Hofbauer and Sandholm (2002 Ecta): stability of limit SFP in zero-sum games and potential games via Lyapunov function.

(3) But... can get limit cycle even with smooth (logit) stochastic fictitious play. (Hommes and Ochea, 2012 GEB)

SFP useful, but not going to solve fundamental nonconvergence problems of learning in games.

Overview

Evolutionary Models

Originally used a simplified model of biological evolution (Maynard Smith 1974).

Basic Setup

- Large anonymous populations; random matching each period.
- Phenotypes = pure actions; inheritance of behaviour.
  - Each individual lives one period; genetically programmed to play single action.
  - Individuals leave offspring inheriting same phenotype (action) as parent.
- **Selection rule:** relative payoff $\mapsto$ differential growth/reproduction/imitation.
  - Survival of the fittest.
- Payoff-monotonicity: if *a* yields higher payoff than *b*, then growth rate of *a* higher than *b*.
- (Can be expanded with random mutations/experimentation.)

Objectives:

- Steady states: surviving traits and their shares in population.
- Stability: selection among equilibria; degree of resistence relative to 'invaders'/innovations.

Applications: diffusion/adoption of valuable innovation, imitation of successful agents.

Replicator Dynamic

Setup (symmetric population game)

Action space A. Payoffs $u : A^2 \to \mathbb{R}$.

Measure of agents playing action a: $q(a)$. Fraction $\boldsymbol{\sigma}(a) = q(a)/\sum_{a'} q(a')$.

Expected payoff: $u(a, \boldsymbol{\sigma})$. Average payoff in population: $\bar{u}(\boldsymbol{\sigma}) \equiv u(\boldsymbol{\sigma}, \boldsymbol{\sigma})$.

Fitness (material payoff) $u(a, \boldsymbol{\sigma}) - u(\boldsymbol{\sigma}, \boldsymbol{\sigma})$.

Discrete-time replicator

If $u > 00$, update by proportional fitness:

$$\boldsymbol{\sigma}_{t+1}(a) = \boldsymbol{\sigma}_t(a) \frac{u(a, \boldsymbol{\sigma}_t)}{u(\boldsymbol{\sigma}, \boldsymbol{\sigma})}.$$

Exact difference form:

$$\boldsymbol{\sigma}_{t+1}(a) - \boldsymbol{\sigma}_t(a) = \boldsymbol{\sigma}_t(a) \frac{u(a, \boldsymbol{\sigma}_t) - u(\boldsymbol{\sigma}, \boldsymbol{\sigma})}{u(\boldsymbol{\sigma}, \boldsymbol{\sigma})}.$$

Steady state (discrete): $\boldsymbol{\sigma}_{t+1} = \boldsymbol{\sigma}_t =: \boldsymbol{\sigma}$.

Continuous-time replicator

$$\dot{\boldsymbol{\sigma}}(a) = \boldsymbol{\sigma}(a)\big(u(a, \boldsymbol{\sigma}) - u(\boldsymbol{\sigma}, \boldsymbol{\sigma})\big).$$

Steady state (continuous): $\dot{\boldsymbol{\sigma}} = \mathbf{0}$.

Discrete → Continuous Approximation

Time-change (Fudenberg and Levine 1998):

From $\sigma_{t+1} - \sigma_t = \sigma_t \dfrac{u - \bar{u}}{\bar{u}}$, set $s_T := \sum_{t<T} \bar{u}(\sigma_t)^{-1}$.

Interpolate $\sigma(s_T) \approx \sigma_T$; difference quotient $\dfrac{\sigma_{T+1} - \sigma_T}{s_{T+1} - s_T} \to \sigma\,(u - \bar{u})$.

Limit ODE: $\dot{\sigma} = \sigma\,(u - \bar{u})$.

Small-step (Nachbar 1990 IJGT):

Take $\sigma_{t+1} = \sigma_t + \eta_t\,\sigma_t\big(u(\cdot,\sigma_t) - \bar{u}(\sigma_t)\mathbf{1}\big)$, $\eta_t \downarrow 0$, $\sum_t \eta_t = \infty$.

Affine reparametrisation of time yields the same ODE: $\dot{\sigma} = \sigma\,(u - \bar{u})$.

Remark: equal-payoff characterisation of steady states is identical in discrete and continuous time.

Hawk-Dove Game

		Player R	
		H	D
Player C	H	$\frac{V-C}{2}, \frac{V-C}{2}$	$V, 0$
	D	$0, V$	$\frac{V}{2}, \frac{V}{2}$

Parameters: $C > V > 0$.

Mixed NE: $p^* = \sigma^*(H) = \frac{V}{C} \in (0,1)$.

$u(H, \sigma^*) = \frac{V(C-V)}{2C} = u(D, \sigma^*)$.

Replicator Dynamic:

Discrete: $p_{t+1} - p_t = \dfrac{p_t(1-p_t)\big(u(H,\sigma_t) - u(D,p_t)\big)}{\bar{u}(p_t)}$. p^* is steady state.

Continuous: $\dot{p}_t = p_t(1-p_t)\big(u(H,p_t) - u(D,p_t)\big)$. p^* globally stable.

Overview

Replicator Dynamic Steady States

Proposition 3.1 (Fudenberg and Levine 1998)

For the discrete or continuous replicator dynamic:

(i) $\boldsymbol{\sigma}$ is a steady state iff $u(a, \boldsymbol{\sigma}) = \bar{u}(\boldsymbol{\sigma})$ for all $a \in \operatorname{supp} \boldsymbol{\sigma}$ and $u(b, \boldsymbol{\sigma}) \leq \bar{u}(\boldsymbol{\sigma})$ for all $b \notin \operatorname{supp} \boldsymbol{\sigma}$.

(ii) If $\boldsymbol{\sigma}$ is stable steady state, then $\boldsymbol{\sigma}$ is NE.

(iii) If $\boldsymbol{\sigma}$ is strict symmetric NE, then $\boldsymbol{\sigma}$ is locally asymptotically stable steady state.

Proof Sketch

(i) Immediately implied by fixed point identity: $\forall a'$, $\boldsymbol{\sigma}_{t+1}(a') - \boldsymbol{\sigma}_t(a') = \boldsymbol{\sigma}_t(a')(u(a', \boldsymbol{\sigma}_t) - u(\boldsymbol{\sigma}_t, \boldsymbol{\sigma}_t))/u(\boldsymbol{\sigma}_t, \boldsymbol{\sigma}_t) = \mathbf{0}$ or $\dot{\boldsymbol{\sigma}}_t(a') = \boldsymbol{\sigma}_t(a')(u(a, \boldsymbol{\sigma}_t) - u(\boldsymbol{\sigma}_t, \boldsymbol{\sigma}_t)) = \mathbf{0}$.

(ii) If it is not NE, then $\exists a : u(a, \boldsymbol{\sigma}) > u(\boldsymbol{\sigma}, \boldsymbol{\sigma})$. Perturbing $\boldsymbol{\sigma}$ a bit will see it move away from $\boldsymbol{\sigma}$, not stable.

(iii) Let $a^* : \boldsymbol{\sigma} = \mathbf{1}_{\{a=a^*\}}$. Strictness implies $\exists \boldsymbol{\varepsilon} > \mathbf{0} : \forall \boldsymbol{\sigma}' \in B_{\boldsymbol{\varepsilon}}(\boldsymbol{\sigma})$, $u(a^*, \boldsymbol{\sigma}') > u(a', \boldsymbol{\sigma}')$, $\forall a' \neq a^* \implies$ local contraction.

Replicator Dynamic Steady States

Proposition 3.1 (Fudenberg and Levine 1998)

For the discrete or continuous replicator dynamic:

(i) $\boldsymbol{\sigma}$ is a steady state iff $u(a, \boldsymbol{\sigma}) = \bar{u}(\boldsymbol{\sigma})$ for all $a \in \operatorname{supp} \boldsymbol{\sigma}$ and $u(b, \boldsymbol{\sigma}) \leq \bar{u}(\boldsymbol{\sigma})$ for all $b \notin \operatorname{supp} \boldsymbol{\sigma}$.

(ii) If $\boldsymbol{\sigma}$ is stable steady state, then $\boldsymbol{\sigma}$ is NE.

(iii) If $\boldsymbol{\sigma}$ is strict symmetric NE, then $\boldsymbol{\sigma}$ is locally asymptotically stable steady state.

Proof Sketch (v2)

(iii) Using Lyapunov in continuous-time for asymptotic stability of strict NE: Near strict NE $\boldsymbol{\sigma}^* = \mathbf{1}_{\{a=a^*\}}$: $V(\boldsymbol{\sigma}) := D_{KL}(\boldsymbol{\sigma}^* \| \boldsymbol{\sigma}) = \sum_a \boldsymbol{\sigma}^*(a) \ln\left(\frac{\boldsymbol{\sigma}^*(a)}{\boldsymbol{\sigma}(a)}\right)$ is Lyapunov function with $V(\boldsymbol{\sigma}^*) = \mathbf{0}$ and $V(\boldsymbol{\sigma}) > \mathbf{0}\ \forall \boldsymbol{\sigma} \neq \boldsymbol{\sigma}^*$, $\dot{V}(\boldsymbol{\sigma}) \propto -(u(a^*, \boldsymbol{\sigma}) - u(\boldsymbol{\sigma}, \boldsymbol{\sigma}))/\boldsymbol{\sigma}(a^*) < \mathbf{0}$ in neighbourhood of $\boldsymbol{\sigma}^*$.

Replicator Dynamic Steady States

Proposition 3.2 (Fudenberg and Levine 1998; Bomze 1986 IJGT)

Under the discrete or continuous replicator dynamic, any asymptotically stable steady state corresponds to an isolated THPE.

Proof Sketch

2-player game: THPE iff no weakly dominated strategies. If not THPE, cannot be stable (small perturbation leads to moving farther away). If not isolated, then cannot be asymptotically stable (won't come back after small perturbation).

Note: Generically, asymptotically stable steady states are strict NE in discrete-time version of replicator dynamic.

Overview

Evolutionary Stable Strategies

Definition

$\sigma \in \Delta(A)$ is an **evolutionary stable strategy (ESS)** if $\forall \sigma' \neq \sigma$, $\exists \bar{\varepsilon} > 0$ s.t. $\forall \varepsilon \in (0, \bar{\varepsilon})$,

$$u\big(\sigma, (1-\varepsilon)\sigma + \varepsilon\sigma'\big) > u\big(\sigma', (1-\varepsilon)\sigma + \varepsilon\sigma'\big).$$

Intuition: biological resistance against invaders; require equilibrium to resist against 'invaders'/mutants/disruptors.

Remark

σ is ESS if and only if $\forall \sigma' \neq \sigma$, (a) $u(\sigma, \sigma) > u(\sigma', \sigma)$ or (b) $u(\sigma, \sigma) = u(\sigma', \sigma)$ and $u(\sigma, \sigma') > u(\sigma', \sigma')$.

Definition

$\sigma \in \Delta(A)$ is a **weak ESS** if $\forall \sigma' \neq \sigma$, (a) $u(\sigma, \sigma) > u(\sigma', \sigma)$ or (b) $u(\sigma, \sigma) = u(\sigma', \sigma)$ and $u(\sigma, \sigma') \geq u(\sigma', \sigma')$.

Intuition: invader is not driven out, but does not grow either.

Evolutionary Stable Strategies

Application ESS: (outside learning proper)

- Take simple 2×2 symmetric coordination game with two PSNE, Pareto dominant and Pareto inferior.
- Allow for cheap talk à la Crawford-Sobel (1982 Ecta). Turns game into extensive form, but standard refinements can't rule out inferior outcome.
- ESS suggest tendency for coordinating on Pareto superior (see e.g., Blume, Kim, and Sobel 1993 GEB)

Evolutionary Stable Strategies

Proposition 3.3 (Fudenberg and Levine 1998)

If σ^* is ESS, then it is asymptotically stable steady state of continuous replicator dynamic.

Proof Sketch

By definition, ESS implies $\exists \varepsilon > 0 : \forall \sigma \in B_\varepsilon(\sigma^*), u(\sigma, \sigma) - u(\sigma^*, \sigma) < 0$.

Let $V(\sigma) = D_{KL}(\sigma^* || \sigma)$. Then, $\dot{V}(\sigma) = -\sum_a \sigma^*(a) \frac{\mathrm{d} \ln \sigma(a)}{\mathrm{d}t} = -\sum_a \sigma^*(a) \frac{\sigma(a)(u(a,\sigma) - u(\sigma,\sigma))}{\sigma(a)} = u(\sigma, \sigma) - u(\sigma^*, \sigma) < 0$.

V Lyapunov function with $\dot{V} < 0$ in neighbourhood of σ^*. Done.

Evolutionary Stable Strategies

Remark

Not all asymptotically stable steady state of continuous replicator dynamic is ESS.

		Player R		
		T	M	B
	T	0	1	1
Player C	M	-2	0	4
	B	1	1	0

Counterexample from van Damme (1987) via Fudenberg and Levine (1998).

$F(\boldsymbol{\sigma}) = (\boldsymbol{\sigma}(a)(u(a, \boldsymbol{\sigma}) - u(\boldsymbol{\sigma}, \boldsymbol{\sigma})), a \in A)$.

! symNE: $\boldsymbol{\sigma}^* = (\mathbf{1/3}, \mathbf{1/3}, \mathbf{1/3})$; payoff $\mathbf{2/3}$.

$J_F(\boldsymbol{\sigma}^*)$ has only negative eigenvalues, hence $\boldsymbol{\sigma}^*$ asymptotically stable.

But $\boldsymbol{\sigma}^*$ not ESS: can be invaded by strategy $\boldsymbol{\sigma} = (\mathbf{0}, \mathbf{1/2}, \mathbf{1/2})$ with payoff 2/3 when matched with $\boldsymbol{\sigma}^*$ and payoff 5/4 when matched with itself.

Evolutionary Stable Strategies

Proposition 3.3 (Fudenberg and Levine 1998)

If σ^* is ESS, then it is asymptotically stable steady state of continuous replicator dynamic.

Remark

Not all asymptotically stable steady state of continuous replicator dynamic is ESS.

Absence of converse due to replicator dynamic only allowing inheritance of pure stratgies.

If extend replicator to mixed strategies, then ESS is equivalent to asymptotic stability under replicator dynamic (Bomze 1986 IJGT).

Overview

Asymmetric Replicator Dynamic

Populations $i \in I$; states $\boldsymbol{\sigma}_i \in \Delta(A_i)$.

Replicator for population i:

$$\dot{\boldsymbol{\sigma}}_i(a_i) = \boldsymbol{\sigma}_i(a_i)\big(u_i(a_i, \boldsymbol{\sigma}_{-i}) - \bar{u}_i(\boldsymbol{\sigma})\big), \qquad \bar{u}_i(\boldsymbol{\sigma}) := \sum_{a_i} \boldsymbol{\sigma}_i(a_i) u_i(a_i, \boldsymbol{\sigma}_{-i}).$$

Issue: non-degenerate strategies generically cannot be asymptotically stable... See Fudenberg and Levine §3.5.

Big problem when ! NE is in mixed strategies.

Example: symmetric BoS: if agree get 0; if disagree get 1. Unique symmetric NE is unique interior NE.

One population: no asymmetry between players; no way to coordinate on one of the pure strategy equilibria. $\dot{p} = p(u(L, p) - u(p, p)) = p(1 - p)(1 - 2p)$. $V(p) = (1 - 2p)^2$. $\dot{V}(p) = -4(1 - 2p)\dot{p} = -4p(1 - p)(1 - 2p)^2 < 0$ for any $p \neq 1/2$. $p^* = 1/2$ is globally asymptotically stable.

Two populations: $p^* = (1/2, 1/2)$ is not asymptotically stable; perturbing p^* toward (0,1) or (1,0) won't be self-correcting. Players use labels as coordinating device.

Overview

Generalisations of the Replicator Dynamic

Samuelson and Zhang (1992 JET); Hofbauer and Weibull (1996 JET).

2 player game.

Dynamic system $\dot{\boldsymbol{\sigma}} = \boldsymbol{\sigma}F(\boldsymbol{\sigma})$ is regular (i) F is Lipschitz continuous, (ii) $\sum_{a_i} \dot{\boldsymbol{\sigma}}_i(a_i) = \mathbf{0}$, and (iii) $\boldsymbol{\sigma}_i(a_i) = \mathbf{0} \implies \dot{\boldsymbol{\sigma}}_i(a_i) \geq \mathbf{0}$.

Write $\dot{\boldsymbol{\sigma}}_i(a_i) = \boldsymbol{\sigma}_i(a_i) f_{i,a_i}(\boldsymbol{\sigma})$.

Note: $f_{i,a_i}(\boldsymbol{\sigma}) = \frac{\dot{\boldsymbol{\sigma}}_i(a_i)}{\boldsymbol{\sigma}_i(a_i)}$, growth rate.

Definition (Swinkels 1993 GEB)

Dynamic system is **myopic adjustment dynamic** if $\sum_{a_i} u_i(a_i, \boldsymbol{\sigma}_{-i})\dot{\boldsymbol{\sigma}}_i(a_i) \geq \mathbf{0}$.

Share of more successful strategies expands in population (weakly) more than share of less successful strategies.

Nests replicator dynamic and best-response dynamics, i.e. $\dot{\boldsymbol{\sigma}}_i = \mathrm{BR}_i(\boldsymbol{\sigma}_{-i}) - \boldsymbol{\sigma}_i$

Generalisations of the Replicator Dynamic

Definition (Samuelson and Zhang 1992 JET)

Dynamic system is **payoff monotone** if for $\forall$ interior points, $u_i(a_i, \boldsymbol{\sigma}_{-i}) \geq (>) u_i(a_i', \boldsymbol{\sigma}_{-i}) \implies f_{i,a_i}(\boldsymbol{\sigma}) \geq (>) f_{i,a_i'}(\boldsymbol{\sigma})$.

Weak but rules out best-response dynamics $\because$ growth-rate for non-best responses identical at -1 regardless of payoffs.

Proposition (Samuelson and Zhang 1992 JET; via FL98 Proposition 3.4)

Under any regular, payoff monotone dynamic system, if a_i does not survive iterated pure-strategy strict dominance elimination, then, starting in any interior point, in any path, $\boldsymbol{\sigma}_i(a_i)$ converges to 0 asymptotically.

Note: weaker than IESDS: require strictly dominated by *pure* strategy to be eliminated.

Examples exist s.t. payoff monotone dynamics yields strict superset of IESDS.

Generalisations of the Replicator Dynamic

Definition (Samuelson and Zhang 1992 JET)

Dynamic system is **aggregate monotone** if $\forall$ interior points, $u_i(\sigma_i, \sigma_{-i}) > u_i(\sigma_i', \sigma_{-i}) \implies \sum_{a_i} (\sigma_i(a_i) - \sigma_i'(a_i)) f_{i,a_i}(\sigma) > 0$.

Definition (Hofbauer and Weibull 1996 JET)

Dynamic system is **convex monotone** if $\forall$ interior points, $u_i(\sigma_i', \sigma_{-i}) > u_i(a_i', \sigma_{-i}) \implies \sum_{a_i} \sigma_i'(a_i) f_{i,a_i}(\sigma) > f_{i,a_i'}(\sigma)$.

Aggregate monotone $\implies$ Convex monotone $\implies$ Payoff monotone $\implies$ Myopic adjustment.

Proposition (Hofbauer and Weibull 1996 JET; via FL98 Proposition 3.5)

Under any regular, convex monotone dynamic system, if a_i is eliminated by IESDS, then, starting in any interior point, in any path, $\sigma_i(a_i)$ converges to 0 asymptotically.

Further Generalisations of the Replicator Dynamic

Easy to consider further extensions:

- Embedding noise in dynamics, random mutations/innovations (e.g., Kandori Mailath Rob 1993 Ecta; Young 1993 Ecta).
- Death and birth rate/ evolving population size (Weibull 1995).
- Non-random matching in population, e.g., assortative matching/affiliation/homophily (e.g., Alger and Weibull 2013 Ecta).
- Types as payoffs $\neq$ fitness function (e.g., Dekel, Ely and Yilankaya 2007 REStud).

Further Generalisations of the Replicator Dynamic

Some reference textbooks:

More introductory but thought-provoking:
Bowles (2006), *Microeconomics: Behavior, Institutions, and Evolution*;
Gintis (2009), *Game Theory Evolving*.

Weibull (1995), *Evolutionary Game Theory*.

Vega-Redondo (1996), *Evolution, Games, and Economic Behaviour*.

Samuelson (1997), *Evolutionary Games and Equilibrium Selection*.

Hofbauer and Sigmund (1998), *Evolutionary Games and Population Dynamics*.

Young (1998), *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions*.

Sandholm (2009), *Population Games and Evolutionary Dynamics*.

My take: Huge literature with lots of fundamental results out there...
but somehow feels like there is lots of space for **applications**.
(Especially since fundamental results semi-forgotten.)

Fictitious Play and Replicator Dynamic

Duarte Gonçalves

University College London

Topics in Economic Theory