Approachability, Calibration, Adaptive Algorithms, and Sophisticated Learning

Duarte Gonçalves

University College London

Topics in Economic Theory

Overview

1. Learning in Games

2. Approachability

3. Calibration

4. Adaptive Algorithms

5. Sophisticated Learning

Overview

1. Learning in Games

2. Approachability

3. Calibration

4. Adaptive Algorithms

5. Sophisticated Learning

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

1. Learning in Games
2. Approachability
3. Calibration
4. Adaptive Algorithms
5. Sophisticated Learning

MSZ Ch. 14

Overview

1. Learning in Games
2. Approachability
3. Calibration
4. Adaptive Algorithms
5. Sophisticated Learning

Calibrated learning procedure: learning procedure such that in the long run each action is a best response to the frequency distribution of opponents' choices in all periods in which that action was played

Foster Vohra 1997 GEB, Calibrated Learning and Correlated Equilibrium

Foster Hart 2018 GEB, Smooth calibration, leaky forecasts, finite recall, and Nash dynamics

Foster Hart 2021 JPE, Forecast Hedging and Calibration

Overview

1. Learning in Games
2. Approachability
3. Calibration
4. Adaptive Algorithms
5. Sophisticated Learning

Papers: Hart Mas-Colell 2003 AER, Uncoupled Dynamics Do Not Lead to Nash Equilibrium
*Hart 2005 Ecta, Adaptive Heuristics
Papers on reinforcement learning and Q-learning

Sophisticated Learning

What is players are Bayesian wrt gameplay and engage in sophisticated learning?

Two papers:

Kalai and Lehrer (1993 Ecta) “Rational Learning Leads to Nash Equilibria”

Kalai and Lehrer (1993 Ecta) “Subjective Equilibrium in Repeated Games”

(Will favour Fudenberg and Levine’s “sophisticated learning” terminology.)

Stage Game and Repeated Interaction

Players $i \in I = \{1, \ldots, n\}$; actions A_i (finite). Profile $A = \times_i A_i$.

Payoffs $u_i : A \to \mathbb{R}$. One-period outcome $a^t = (a_i^t)_i \in A$.

Repeated game: infinite horizon, perfect monitoring, discounts $\delta_i \in (0, 1)$.

Histories $h^t = (a^0, \ldots, a^{t-1}) \in H^t := A^t$; $H = \cup_{t \geq 0} H^t$; $\emptyset$ at $t = 0$.

Behavioural strategies $\sigma_i = (\sigma_{i,t})_{t \geq 0}$, with $\sigma_{i,t} : H^t \to \Delta(A_i)$.

Strategy profile $\sigma = (\sigma_i)_i$. Outcome law μ^σ on $\Omega := A^{\mathbb{N}}$ (product σ-algebra).

History concatenation: $hh' \in H^{t+r} : h \in H^t, h' \in H^r$.

Continuation histories starting from h_t: $C(h_t) := \{h' \in H^\infty \mid (h_t h') \in H^\infty\}$.

Filtration $(\mathcal{F}_t)$, $\mathcal{F}_t := \sigma(\{h^t\})$.

Normalised expected discounted payoff:

$$U_i(\sigma) = (1 - \delta_i)\mathbb{E}_{\mu^\sigma}\left[\sum_{t \geq 0} \delta_i^t \, u_i(a^t)\right].$$

Beliefs, Absolute Continuity, and Payoffs

Player i's conjectures/degenerate beliefs about opponents' strategies σ^i_{-i}.

Induces belief $\mu_i = \mu^{\sigma^i_{-i}}$ on Ω.

Player i's prior ν_i on opponents' strategies σ_{-i} (Actual uncertainty).

Induces belief μ_i on Ω via $\tilde{\sigma}_{-i} \mapsto \mu^{(\sigma_i, \tilde{\sigma}_{-i})}$.

For ν_i, expected conjecture: $\sigma^i_{-i}(h)(a_{-i}) = \mathbb{E}_{\tilde{\sigma}_{-i} \sim \nu_i}[\tilde{\sigma}_{-i}(h)(a_{-i})]$.

Player i's **Subjective joint strategy:** $\sigma^i = (\sigma_i, \sigma^i_{-i})$.

Truth-compatibility (absolute continuity): $\mu^\sigma \ll \mu_i$ for all i.

(i.e., $\mu^\sigma(E) > 0 \implies \mu_i(E) > 0$ for any μ_i-measurable E.)

Posteriors: after h^t, update $\mu_i(\cdot \mid h^t)$ by Bayes (well-defined by abs. cont.).

Rationality path: each period t, $\sigma_{i,t}$ is a best response to $\mu_i(\cdot \mid h^t)$.

Induced strategy: for histories $h, h' \in H$, denote $\sigma_h(h') := \sigma(hh')$ (strategy following h for h').

Closeness and "Plays ε-Like"

Definition (ε-close measures)

For $\varepsilon > 0$, μ is **ε-close** to $\tilde{\mu}$ if $\exists Q$ with $\mu(Q), \tilde{\mu}(Q) \geq 1 - \varepsilon$ s.t. $\forall$ measurable $A \subseteq Q$,

$$(1 - \varepsilon)\tilde{\mu}(A) \leq \mu(A) \leq (1 + \varepsilon)\tilde{\mu}(A).$$

Definition (plays ε-like)

A profile σ **plays ε-like** σ' if μ^{σ} is ε-close to $\mu^{\sigma'}$; equivalently, after any h^t, the conditional laws are ε-close on a large-probability subset.

Controls conditional probabilities on tails; prevents cumulative small-error blowup across time.

Learning to Predict Future Play

Theorem 1 (Learning to predict)

Fix actual strategy $\boldsymbol{\sigma}$ and player i's subjective joint strategy $\boldsymbol{\sigma}^i := (\boldsymbol{\sigma}_i, \boldsymbol{\sigma}^i_{-i})$. If $\boldsymbol{\mu}^{\boldsymbol{\sigma}} \ll \boldsymbol{\mu}^{\boldsymbol{\sigma}^i}$, then for every $\boldsymbol{\varepsilon} > \mathbf{0}$ and for $\boldsymbol{\mu}^{\boldsymbol{\sigma}}$-a.e. path $h \in H^\infty$, $\exists T$ s.t. $\forall t \geq T$, continuation $\boldsymbol{\sigma}_{h_t}$ plays $\boldsymbol{\varepsilon}$-like $\boldsymbol{\sigma}^i_{h_t}$.

Posterior forecasts of future play (conditional on realised history) merge with truth.

No optimality required here; this is a property of Bayesian updating under abs. cont.

Merging via Likelihood Ratios

Theorem 3 (Blackwell and Dubins, 1962)

If $\mu \ll \tilde{\mu}$, then with μ-probability 1, for every $\varepsilon > 0$ there exists random time $\tau(\varepsilon)$ such that for all $t \geq \tau(\varepsilon)$ the posteriors $\mu(\cdot \mid \mathcal{F}_t)$ and $\tilde{\mu}(\cdot \mid \mathcal{F}_t)$ are ε-close.

If people start off with compatible priors, posteriors become arbitrarily close after exposed to enough information.

Proof Idea

Radon-Nikodym derivative $\phi = \frac{d\mu}{d\tilde{\mu}}$ exists; set $M_t = \mathbb{E}_{\tilde{\mu}}[\phi \mid \mathcal{F}_t]$.

(M_t) is a nonnegative $\tilde{\mu}$-martingale; $M_t \to M_\infty$ a.s.

Control likelihood ratios on Q with $\mu(Q), \tilde{\mu}(Q) \approx 1$.

Translate bounds to conditionals on continuation histories $C(h^t)$; conclude ε-closeness.

Subjective ε-Equilibrium

Definition (Subjective ε-equilibrium)

A profile $\sigma = (\sigma_i)_i$ is a **subjective ε-equilibrium** if there exist beliefs $\sigma^i = (\sigma_i, \sigma^i_{-i})$ with:

σ_i is a best response to σ^i_{-i}, for every i;

σ plays ε-like σ^i, for every i.

Corollary 1

If each σ_i best responds to σ^i_{-i} and $\sigma \ll \sigma^i$ for all i, then for a.e. path h $\exists T$ s.t. $\forall t \geq T$, the continuation σ_{h_t} is a subjective ε-equilibrium.

Proof Idea

Fix $\varepsilon > 0$; for μ^σ-a.e. h $\exists T$ s.t. $\forall t \geq T$, σ_{h_t} plays ε-like $\sigma^i_{h_t}$ for each i (Theorem 1).

By rationality, at every t player i plays a best response to $\mu_i(\cdot \mid h_t)$.

Merging $\Longrightarrow$ those best responses are ε-best responses to true continuation $\mu^\sigma(\cdot \mid h_t)$.

Both (supporting beliefs & closeness) $\Longrightarrow$ subjective ε-equilibrium from time T.

From Subjective to (Approximate) Nash

Proposition 1

For every $\varepsilon > 0$, $\exists \eta > 0$: if σ is a subjective η-equilibrium then $\exists \sigma^*$ s.t.

(i) σ plays ε-like σ^*;

(ii) σ^* is an ε-Nash equilibrium of the repeated game.

Idea: under perfect monitoring and known own payoffs, adjust off-path prescriptions to align incentives while preserving realisations up to ε.

Proof Idea

Fix $\eta > 0$ small. Given subjective η-equilibrium σ, modify off-path prescriptions s.t. unilateral deviations trigger responses that keep the deviator's continuation payoff within ε of best-reply payoff.

Perfect monitoring $\implies$ changes leave realisations ε-close.

Resulting σ^* is an ε-best reply for each player: σ^* is an ε-Nash equilibrium; and σ plays ε-like σ^*.

Main Theorem: Rational Learning $\Longrightarrow$ Nash Play

Theorem 2 (Kalai and Lehrer 1993)

Suppose each σ_i best responds to σ^i_{-i} and $\mu^\sigma \ll \mu^{\sigma^i}$ for all i. Then for every $\varepsilon > 0$ and for μ^σ-a.e. path h, $\exists T$ s.t. $\forall t \geq T$ there is an ε-Nash equilibrium σ^ε of the repeated game with σ_{h_t} playing ε-like σ^ε.

Proof Idea

1) Theorem 1 $\Longrightarrow$ eventually correct forecasts (merging).
2) Best responses to beliefs $\Longrightarrow$ ε-best responses to truth (large t).
3) Proposition 1 $\Longrightarrow$ approximate Nash play along the realised path.

Absolute Continuity and Bayesian Nash Equilibrium

Bayesian Nash equilibrium (BNE): in incomplete information (finite type space), each σ_i maximises expected utility given beliefs over types and strategies.

At a BNE of the repeated game, priors give a *grain of truth*: realised play has positive probability under beliefs $\implies$ absolute continuity holds.

Application: starting from a BNE, players eventually play (approximately) a Nash equilibrium of the *realised* complete-information repeated game.

Meaning and Interpretation

What converges? Not actions each period, but *forecasts* of future play; behaviour is best response to (nearly) correct forecasts.

Why it matters: ensures long-run play consistent with Nash discipline without common knowledge of rationality or equilibrium selection.

Learning vs commitment: players learn the environment they *face* (others' strategies), not a fixed state of nature.

Role of absolute continuity: bans dogmatic zero-probability beliefs about realised events; makes Bayes informative.

Learning: with merging, each player's beliefs about future play match the truth; subjective ε-equilibrium obtains on-path.

Incomplete Information on Payoffs

Bayesian Nash starting point

In a repeated game with finitely many payoff types, if play starts at a **Bayesian Nash equilibrium**, then eventually players play (approximately) a Nash equilibrium of the *realised* complete-information repeated game.

Grain of truth at BNE $\Longrightarrow$ abs. cont.; merging $\Longrightarrow$ correct forecasts; best responses $\Longrightarrow$ near-NE of realised environment.

Fudenberg and Levine (1998; 2009 ARE): Main Critique

Endogeneity of absolute continuity: abs. cont. must hold for the realised path under the true play; ensuring this is itself an equilibrium-like fixed-point problem.

Grain of truth: wanting priors that *always* put positive mass on the truth is impossible in rich (uncountable) environments; workable classes may be very restrictive.

Interpretation caution: Kalai and Lehrer (1993 Ecta) shows a *consistency* result conditional on abs. cont.; not a general path-to-equilibrium selection theory.

Comparative statics: results sensitive to prior support assumptions; small changes can break abs. cont. and merging conclusion.

Bottom line: powerful when abs. cont. holds (e.g., BNE start with finite types), but limited as a general behavioural foundation without specifying priors.

"Our interest here, however, is in "learning models," by which we mean that the allowed priors are exogenously specified, without reference to a fixed point problem." Fudenberg and Levine (1998)

Takeaways

Under absolute continuity, Bayesian learning merges beliefs with the truth along realised play.

Rational (best-reply) control with merged beliefs $\implies$ eventual (approximate) Nash play.

At BNE with finite types, eventual play tracks an NE of the realised complete-information game.

Abs. cont. is strong and endogenous; use with care as general foundation for learning in games.

Approachability, Calibration, Adaptive Algorithms, and Sophisticated Learning

Duarte Gonçalves

University College London

Topics in Economic Theory