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
   - Multi-Utility Representation
   - Approachability
   - Application: Picking Experts
   - Universal Consistency

3. Calibration

4. Adaptive Algorithms

5. Sophisticated Learning

Approachability

Setup:

Two players face repeated game. m-dimensional goal/multi-utility representation.

Goal of each player: control average vector of attributes s.t. approaches/excludes (keeps away) target set $S \subset \mathbb{R}^m$.

Brief detour: rationalising multi-utility.

Overview

1. Learning in Games

2. Approachability
   - Multi-Utility Representation
   - Approachability
   - Application: Picking Experts
   - Universal Consistency

3. Calibration

4. Adaptive Algorithms

5. Sophisticated Learning

Multi-Utility Representation

$\succsim \subseteq X^2$: preorder (reflexive, transitive).

Want to allow for incompleteness.

x and y are **incomparable** if $\neg(x \succsim y \text{ or } y \succsim x)$. Ow, they are comparable.

E.g., choice by unanimity, incomparable attributes.

Multi-Utility Representation

$\succsim \subseteq X^2$: preorder (reflexive, transitive).

Want to allow for incompleteness.

x and y are **incomparable** if $\neg(x \succsim y \text{ or } y \succsim x)$. Ow, they are comparable.

E.g., choice by unanimity, incomparable attributes.

Definition

For any binary relation $\succsim$ on X with symmetric part $\sim$, for any $x \in X$, x's **equivalence class** is $[x] := \{y \in X | x \sim y\}$ and the set of equivalence classes $\hat{X} := \{[x], x \in X\}$.

Multi-Utility Representation

$\succsim \subseteq X^2$: preorder (reflexive, transitive).

Want to allow for incompleteness.

x and y are **incomparable** if $\neg(x \succsim y \text{ or } y \succsim x)$. Ow, they are comparable.

E.g., choice by unanimity, incomparable attributes.

Definition

For any binary relation $\succsim$ on X with symmetric part $\sim$, for any $x \in X$, x's **equivalence class** is $[x] := \{y \in X | x \sim y\}$ and the set of equivalence classes $\hat{X} := \{[x], x \in X\}$.

Remark

For any preorder $\succsim$ on X, let $\hat{\succsim}$ on $\hat{X} : \forall x, y \in X : [x] \hat{\succsim} [y]$ if $x \succsim y$. Then, $\hat{\succsim}$ is partial order.

Multi-Utility Representation

$\succsim \subseteq X^2$: preorder (reflexive, transitive).

Want to allow for incompleteness.

Multi-Utility Representation

$\succsim \subseteq X^2$: preorder (reflexive, transitive).

Want to allow for incompleteness.

x and y are **incomparable** if $\neg(x \succsim y \text{ or } y \succsim x)$. Ow, they are comparable.

E.g., choice by unanimity, incomparable attributes.

Multi-Utility Representation

$\succsim \subseteq X^2$: preorder (reflexive, transitive).

Want to allow for incompleteness.

x and y are **incomparable** if $\neg(x \succsim y \text{ or } y \succsim x)$. Ow, they are comparable.

E.g., choice by unanimity, incomparable attributes.

Set of all linear extensions of $\succsim$ on X: $\mathcal{L}(\hat{X}, \hat{\succsim})$.

Szpilrajn's Theorem: any partial order can be extended to a linear order.

Multi-Utility Representation

$\succsim \subseteq X^2$: preorder (reflexive, transitive).

Want to allow for incompleteness.

x and y are **incomparable** if $\neg(x \succsim y \text{ or } y \succsim x)$. Ow, they are comparable.

E.g., choice by unanimity, incomparable attributes.

Set of all linear extensions of $\hat{\succsim}$ on X: $\mathcal{L}(\hat{X}, \hat{\succsim})$.

Szpilrajn's Theorem: any partial order can be extended to a linear order.

Remark 1: $\implies \mathcal{L}(\hat{X}, \hat{\succsim}) \neq \emptyset$.

Remark 2: $\hat{\succsim} = \cap_{\tilde{\succsim} \in \mathcal{L}(\hat{X}, \hat{\succsim})} \tilde{\succsim}$.

Multi-Utility Representation

$\succsim \subseteq X^2$: preorder (reflexive, transitive).

Want to allow for incompleteness.

x and y are **incomparable** if $\neg(x \succsim y \text{ or } y \succsim x)$. Ow, they are comparable.

E.g., choice by unanimity, incomparable attributes.

Set of all linear extensions of $\hat{\succsim}$ on X: $\mathcal{L}(\hat{X}, \hat{\succsim})$.

Szpilrajn's Theorem: any partial order can be extended to a linear order.

Remark 1: $\implies \mathcal{L}(\hat{X}, \hat{\succsim}) \neq \emptyset$.

Remark 2: $\hat{\succsim} = \cap_{\tilde{\succeq} \in \mathcal{L}(\hat{X}, \hat{\succsim})} \tilde{\succeq}$.

Order dimension: $\dim(X, \succsim) := \min\{k \in \mathbb{N} | \succeq_i \in \mathcal{L}(\hat{X}, \hat{\succsim}), i = 1, ..., k : \hat{\succsim} = \cap_{i=1}^{k} \succeq_i\}$.

$\dim(X, \succsim)$: min number of linear extensions of $\hat{\succsim}$ whose intersection yields $\hat{\succsim}$.

Multi-Utility Representation

$\succsim \subseteq X^2$: preorder (reflexive, transitive).

Want to allow for incompleteness.

x and y are **incomparable** if $\neg(x \succsim y \text{ or } y \succsim x)$. Ow, they are comparable.

E.g., choice by unanimity, incomparable attributes.

Set of all linear extensions of $\hat{\succsim}$ on X: $\mathcal{L}(\hat{X}, \hat{\succsim})$.

Szpilrajn's Theorem: any partial order can be extended to a linear order.

Remark 1: $\Longrightarrow \mathcal{L}(\hat{X}, \hat{\succsim}) \neq \emptyset$.

Remark 2: $\hat{\succsim} = \cap_{\tilde{\succeq} \in \mathcal{L}(\hat{X}, \hat{\succsim})} \tilde{\succeq}$.

Order dimension: $\dim(X, \succsim) := \min\{k \in \mathbb{N} | \succeq_i \in \mathcal{L}(\hat{X}, \hat{\succsim}), i = 1, ..., k : \hat{\succsim} = \cap_{i=1}^{k} \succeq_i\}$.

$\dim(X, \succsim)$: min number of linear extensions of $\hat{\succsim}$ whose intersection yields $\hat{\succsim}$.

Examples:

$\hat{\succsim}$ is linear order on X iff $\dim(X, \succsim) = 1$.

Multi-Utility Representation

$\succsim \subseteq X^2$: preorder (reflexive, transitive).

Want to allow for incompleteness.

x and y are **incomparable** if $\neg(x \succsim y \text{ or } y \succsim x)$. Ow, they are comparable.

E.g., choice by unanimity, incomparable attributes.

Set of all linear extensions of $\hat{\succsim}$ on X: $\mathcal{L}(\hat{X}, \hat{\succsim})$.

Szpilrajn's Theorem: any partial order can be extended to a linear order.

Remark 1: $\Longrightarrow \mathcal{L}(\hat{X}, \hat{\succsim}) \neq \emptyset$.

Remark 2: $\hat{\succsim} = \cap_{\tilde{\geq} \in \mathcal{L}(\hat{X}, \hat{\succsim})} \tilde{\geq}$.

Order dimension: $\dim(X, \succsim) := \min\{k \in \mathbb{N} | \geq_i \in \mathcal{L}(\hat{X}, \hat{\succsim}), i = 1, ..., k : \hat{\succsim} = \cap_{i=1}^{k} \geq_i\}$.

$\dim(X, \succsim)$: min number of linear extensions of $\hat{\succsim}$ whose intersection yields $\hat{\succsim}$.

Examples:

$\hat{\succsim}$ is linear order on X iff $\dim(X, \succsim) = 1$.

If no distinct x, y are comparable ($\hat{\succsim}$ is antichain) and $\dim(X, \succsim) = 2$ since $\hat{\succsim} = \geq \cap \leq$.

Multi-Utility Representation

$\succsim \subseteq X^2$: preorder (reflexive, transitive).

Want to allow for incompleteness.

x and y are **incomparable** if $\neg(x \succsim y \text{ or } y \succsim x)$. Ow, they are comparable.

E.g., choice by unanimity, incomparable attributes.

Set of all linear extensions of $\hat{\succsim}$ on X: $\mathcal{L}(\hat{X}, \hat{\succsim})$.

Szpilrajn's Theorem: any partial order can be extended to a linear order.

Remark 1: $\implies \mathcal{L}(\hat{X}, \hat{\succsim}) \neq \emptyset$.

Remark 2: $\hat{\succsim} = \cap_{\tilde{\geq} \in \mathcal{L}(\hat{X}, \hat{\succsim})} \tilde{\geq}$.

Order dimension: $\dim(X, \succsim) := \min\{k \in \mathbb{N} | \geq_i \in \mathcal{L}(\hat{X}, \hat{\succsim}), i = 1, ..., k : \hat{\succsim} = \cap_{i=1}^{k} \geq_i\}$.

$\dim(X, \succsim)$: min number of linear extensions of $\hat{\succsim}$ whose intersection yields $\hat{\succsim}$.

Examples:

$\hat{\succsim}$ is linear order on X iff $\dim(X, \succsim) = 1$.

If no distinct x, y are comparable ($\hat{\succsim}$ is antichain) and $\dim(X, \succsim) = 2$ since $\hat{\succsim} = \geq \cap \leq$.

If $X = 2^A$ and $|A| = \infty$, then $\dim(X, \subseteq) = \infty$.

Multi-Utility Representation

Definition

$\succsim \subseteq X^2$ admits a multi-utility representation $u : X \to \mathbb{R}^m$ with $m \in \mathbb{N}$ iff $\forall x, y \in X, x \succsim y \iff u(x) \geq u(y)$.

Multi-Utility Representation

Definition

$\succsim \subseteq X^2$ admits a multi-utility representation $u : X \to \mathbb{R}^m$ with $m \in \mathbb{N}$ iff $\forall x, y \in X, x \succsim y \iff u(x) \geq u(y)$.

Proposition 1 (Ok 2002 JET)

Let $\succsim$ be preorder on X.
(1) $\succsim$ admits a multi-utility representation u only if $\dim(X, \succsim) < \infty$.
(2) If $\hat{X}$ countable, $\succsim$ admits a multi-utility representation u if and only if $\dim(X, \succsim) < \infty$.

Multi-Utility Representation

Definition

$\succsim \subseteq X^2$ admits a multi-utility representation $u : X \to \mathbb{R}^m$ with $m \in \mathbb{N}$ iff $\forall x, y \in X, x \succsim y \iff u(x) \geq u(y)$.

Proposition 1 (Ok 2002 JET)

Let $\succsim$ be preorder on X.
(1) $\succsim$ admits a multi-utility representation u only if $\dim(X, \succsim) < \infty$.
(2) If $\hat{X}$ countable, $\succsim$ admits a multi-utility representation u if and only if $\dim(X, \succsim) < \infty$.

Alternative (social) interpretation: $\exists U \subset \mathbb{R}^X$ such that $x \succsim y \iff u(x) \geq u(y) \forall u \in U$.

Multi-Utility Representation

Definition

$\succsim \subseteq X^2$ admits a multi-utility representation $u : X \to \mathbb{R}^m$ with $m \in \mathbb{N}$ iff $\forall x, y \in X, x \succsim y \iff u(x) \geq u(y)$.

Proposition 1 (Ok 2002 JET)

Let $\succsim$ be preorder on X.
(1) $\succsim$ admits a multi-utility representation u only if $\dim(X, \succsim) < \infty$.
(2) If $\hat{X}$ countable, $\succsim$ admits a multi-utility representation u if and only if $\dim(X, \succsim) < \infty$.

Alternative (social) interpretation: $\exists U \subset \mathbb{R}^X$ such that $x \succsim y \iff u(x) \geq u(y) \forall u \in U$.

Proof Idea

Take X finite. Let $u_x(y) = \mathbf{1}_{\{y \succsim x\}}$. $u(y) = (u_x(y))_{x \in X}$.

Let $x \succsim y$. (a) $\forall z \in X : (u_z(x) = 0) \iff (z \succsim x) \implies (z \succsim y) \iff (u_z(y) = 0)$

(b) $\forall z \in X : (u_z(y) = 1) \iff (y \succsim z) \implies (x \succsim z) \iff (u_z(x) = 1)$.

(a) + (b) $\implies u(x) \geq u(y)$.

Multi-Utility Representation

Definition

$\succsim \subseteq X^2$ admits a multi-utility representation $u : X \to \mathbb{R}^m$ with $m \in \mathbb{N}$ iff $\forall x, y \in X, x \succsim y \iff u(x) \geq u(y)$.

Proposition 1 (Ok 2002 JET)

Let $\succsim$ be preorder on X.
(1) $\succsim$ admits multi-utility representation u only if $\dim(X, \succsim) < \infty$.
(2) If $\hat{X}$ countable, $\succsim$ admits a multi-utility representation u if and only if $\dim(X, \succsim) < \infty$.

Proposition 2 (Ok 2002 JET)

(a) $X_0 = \times_{i=1}^{k} X_i$, with X_i be metric space and $\succsim_i$ be preorders on X_i, $i = 0, 1, ..., k$;

(b) Each X_i is s.t. $\{y_i \mid y_i \succ_i x_i\}$ is open for every $x_i \in X_i$ and $i = 1, ..., k$; and

(c) $x \succsim_0 y \iff x_i \succsim_i y_i \; \forall i = 1, ..., k$.

If X_0 admits a countable $\succsim_0$-dense subset, then $\succsim_0$ admits a multi-utility representation u which is continuous in the product topology.

Overview

1. Learning in Games

2. Approachability
   - Multi-Utility Representation
   - Approachability
   - Application: Picking Experts
   - Universal Consistency

3. Calibration

4. Adaptive Algorithms

5. Sophisticated Learning

Approachability

Setup:

Two players face repeated game. m-dimensional goal/multi-utility representation.

Goal of each player: control average vector of attributes s.t. approaches/excludes (keeps away) target set $S \subset \mathbb{R}^m$.

Back to approachability... Blackwell (1956). Good reference: Maschler, Solan, Zamir (2013 Book, Ch. 14).

Actions: A_i; **Stage-Game Payoffs:** $u_1 : A_1 \times A_2 \to \mathbb{R}^m$ $u_2 := -u_1$; (endowed with $d(x, y) = \|x - y\|_2$);

Histories: $H_t := A^t$, $H := \cup_t H_t$; **Strategies:** $\sigma_i : H \to \Delta(A_i)$; $\lambda_i \in \Delta(A_i)$.

Approachability

Setup:

Two players face repeated game. m-dimensional goal/multi-utility representation.

Goal of each player: control average vector of attributes s.t. approaches/excludes (keeps away) target set $S \subset \mathbb{R}^m$.

Back to approachability... Blackwell (1956). Good reference: Maschler, Solan, Zamir (2013 Book, Ch. 14).

Actions: A_i; **Stage-Game Payoffs:** $u_1 : A_1 \times A_2 \to \mathbb{R}^m$ $u_2 := -u_1$; (endowed with $d(x, y) = \|x - y\|_2$);

Histories: $H_t := A^t$, $H := \cup_t H_t$; **Strategies:** $\sigma_i : H \to \Delta(A_i)$; $\lambda_i \in \Delta(A_i)$.

Expected Payoffs: $u_i(\lambda_i, \lambda_{-i}) = \sum_{a_i} \sum_{a_{-i}} \lambda_i(a_i) u_i(a_i, a_{-i}) \lambda_{-i}(a_{-i}) \in \mathbb{R}^m$.

Approachability

Setup:

Two players face repeated game. m-dimensional goal/multi-utility representation.

Goal of each player: control average vector of attributes s.t. approaches/excludes (keeps away) target set $S \subset \mathbb{R}^m$.

Back to approachability... Blackwell (1956). Good reference: Maschler, Solan, Zamir (2013 Book, Ch. 14).

Actions: A_i; **Stage-Game Payoffs:** $u_1 : A_1 \times A_2 \to \mathbb{R}^m$ $u_2 := -u_1$; (endowed with $d(x, y) = \|x - y\|_2$);

Histories: $H_t := A^t$, $H := \cup_t H_t$; **Strategies:** $\sigma_i : H \to \Delta(A_i)$; $\lambda_i \in \Delta(A_i)$.

Expected Payoffs: $u_i(\lambda_i, \lambda_{-i}) = \sum_{a_i} \sum_{a_{-i}} \lambda_i(a_i) u_i(a_i, a_{-i}) \lambda_{-i}(a_{-i}) \in \mathbb{R}^m$.

Feasible Expected Payoffs for λ_i: $U_i(\lambda_i) := \{u_i(\lambda_i, \lambda_{-i}),\ \lambda_{-i} \in \Delta(A_{-i})\} \subseteq \mathbb{R}^m$.

Approachability

Setup:

Two players face repeated game. m-dimensional goal/multi-utility representation.

Goal of each player: control average vector of attributes s.t. approaches/excludes (keeps away) target set $S \subset \mathbb{R}^m$.

Back to approachability... Blackwell (1956). Good reference: Maschler, Solan, Zamir (2013 Book, Ch. 14).

Actions: A_i; **Stage-Game Payoffs:** $u_1 : A_1 \times A_2 \to \mathbb{R}^m$ $u_2 := -u_1$; (endowed with $d(x, y) = \|x - y\|_2$);

Histories: $H_t := A^t$, $H := \cup_t H_t$; **Strategies:** $\sigma_i : H \to \Delta(A_i)$; $\lambda_i \in \Delta(A_i)$.

Expected Payoffs: $u_i(\lambda_i, \lambda_{-i}) = \sum_{a_i} \sum_{a_{-i}} \lambda_i(a_i) u_i(a_i, a_{-i}) \lambda_{-i}(a_{-i}) \in \mathbb{R}^m$.

Feasible Expected Payoffs for λ_i: $U_i(\lambda_i) := \{u_i(\lambda_i, \lambda_{-i}),\ \lambda_{-i} \in \Delta(A_{-i})\} \subseteq \mathbb{R}^m$.

Average Payoff: $\bar{u}_{i,t} = \frac{1}{t} \sum_{\ell=1}^{t} u_i(a_t)$.

Feasible Avg Payoffs: $\text{co}(u_i) := \text{co}(\{u_i(a), a \in A\}) \subseteq \mathbb{R}^m$.

Approachability

Definition

$C \subseteq \mathbb{R}^m$ is

approachable by player i if $\exists \boldsymbol{\sigma}_i$ s.t. $\forall \boldsymbol{\varepsilon} > \mathbf{0}$, $\exists T : \forall \boldsymbol{\sigma}_{-i}$, $\mathbb{P}^{\boldsymbol{\sigma}}(d(\bar{u}_t, C) < \boldsymbol{\varepsilon}, \forall t \geq T) > \mathbf{1} - \boldsymbol{\varepsilon}$; in this case, $\boldsymbol{\sigma}_i$ approaches C for player i; and

excludable by player i if $\exists \boldsymbol{\delta}$ s.t. set $C_{\boldsymbol{\delta}}^c := \{x \mid d(x, C) \geq \boldsymbol{\delta}\}$ is approachable by player i; if strategy $\boldsymbol{\sigma}_i$ approaches $C_{\boldsymbol{\delta}}^c$, then it excludes C for player i.

Approachability

Definition

$C \subseteq \mathbb{R}^m$ is

approachable by player i if $\exists \boldsymbol{\sigma}_i$ s.t. $\forall \boldsymbol{\varepsilon} > \mathbf{0}$, $\exists T : \forall \boldsymbol{\sigma}_{-i}$, $\mathbb{P}^{\boldsymbol{\sigma}}(d(\bar{u}_t, C) < \boldsymbol{\varepsilon}, \forall t \geq T) > \mathbf{1} - \boldsymbol{\varepsilon}$; in this case, $\boldsymbol{\sigma}_i$ approaches C for player i; and

excludable by player i if $\exists \boldsymbol{\delta}$ s.t. set $C^c_{\boldsymbol{\delta}} := \{x \mid d(x, C) \geq \boldsymbol{\delta}\}$ is approachable by player i; if strategy $\boldsymbol{\sigma}_i$ approaches $C^c_{\boldsymbol{\delta}}$, then it excludes C for player i.

Approachable by a player if can guarantee that average payoff approaches the set wp1 uniformly over opponent's strategies: $\mathbb{P}^{\boldsymbol{\sigma}}(\lim_{t \to \infty} d(\bar{u}_t, C) = \mathbf{0}) = \mathbf{1}$.

Approachability

Definition

$C \subseteq \mathbb{R}^m$ is

approachable by player i if $\exists \boldsymbol{\sigma}_i$ s.t. $\forall \boldsymbol{\varepsilon} > \mathbf{0}$, $\exists T : \forall \boldsymbol{\sigma}_{-i}$, $\mathbb{P}^{\boldsymbol{\sigma}}(d(\bar{u}_t, C) < \boldsymbol{\varepsilon}, \forall t \geq T) > \mathbf{1} - \boldsymbol{\varepsilon}$; in this case, $\boldsymbol{\sigma}_i$ approaches C for player i; and

excludable by player i if $\exists \boldsymbol{\delta}$ s.t. set $C^c_{\boldsymbol{\delta}} := \{x \mid d(x, C) \geq \boldsymbol{\delta}\}$ is approachable by player i; if strategy $\boldsymbol{\sigma}_i$ approaches $C^c_{\boldsymbol{\delta}}$, then it excludes C for player i.

Approachable by a player if can guarantee that average payoff approaches the set wp1 uniformly over opponent's strategies: $\mathbb{P}^{\boldsymbol{\sigma}}(\lim_{t \to \infty} d(\bar{u}_t, C) = \mathbf{0}) = \mathbf{1}$.

Remark

(1) If $\boldsymbol{\sigma}_i$ approaches (resp. excludes) C, then it approachers (resp. excludes) the closure of C.

(2) C cannot be approachable by one player and excludable by the other.

Hyperplanes

Definition

Hyperplane $H(a,b) := \{x \in \mathbb{R}^m \mid a \cdot x = b\}$.
Half-spaces: $H^+(a,b) := \{x \in \mathbb{R}^m \mid a \cdot x \geq b\}$, $H^-(a,b) := \{x \in \mathbb{R}^m \mid a \cdot x \leq b\}$.

$H^+(a,b) \cap H^-(a,b) = H(a,b)$.

$H^+(a,b) = H^-(-a,-b)$ and $H(a,b) = H(-a,-b)$.

Hyperplanes

Definition

Hyperplane $H(a,b) := \{x \in \mathbb{R}^m \mid a \cdot x = b\}$.
Half-spaces: $H^+(a,b) := \{x \in \mathbb{R}^m \mid a \cdot x \geq b\}$, $H^-(a,b) := \{x \in \mathbb{R}^m \mid a \cdot x \leq b\}$.

$H^+(a,b) \cap H^-(a,b) = H(a,b)$.

$H^+(a,b) = H^-(-a,-b)$ and $H(a,b) = H(-a,-b)$.

Definition

Hyperplane $H(a,b)$ separates x from C if (i) $x \in H^+(a,b) \setminus H(a,b)$ and $C \subseteq H^-(a,b)$, or (ii) $x \in H^-(a,b) \setminus H(a,b)$ and $C \subseteq H^+(a,b)$.

Hyperplanes

Remark

Hyperplane $H(x - y, (x - y) \cdot y)$:

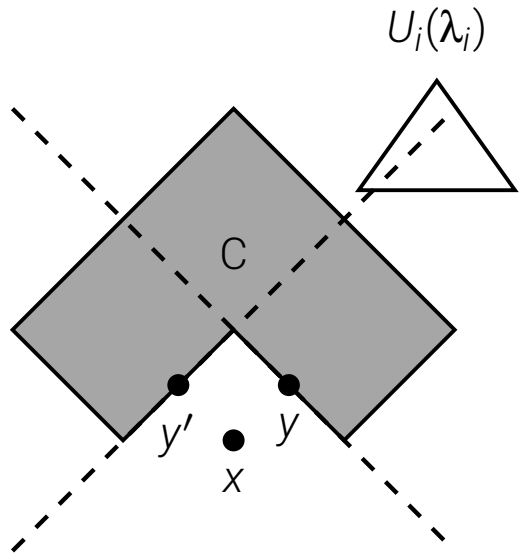
(i) $y \in H(x - y, (x - y) \cdot y)$.

(ii) Orthogonal/Perpendicular to line passing through x and y, i.e. $z \in H(x - y, (x - y) \cdot y) \iff (z - y) \cdot (x - y) = 0$.

(iii) y is point in $H(x - y, (x - y) \cdot y)$ closest to x: $\forall z \in H(x - y, (x - y) \cdot y)$, $\|x - z\|^2 = \|x - y\|^2 + \|y - z\|^2$. (Pythagorean theorem)

(iv) $\forall$ hyperplane H and $x \notin H$, if $y \in H$ is closest in H to x, then $H = H(x - y, (x - y) \cdot y)$.

Separating Hyperplanes in the *B*-set Condition



$H(y' - x, \langle y' - x, y' \rangle)$ $\qquad$ $H(y - x, \langle y - x, y \rangle)$

Hyperplane $H(y' - x, \langle y' - x, y' \rangle)$ separates x from $U_i(\lambda_i)$.

Approachability

Definition

Closed set C is **B-set** for player i if $\forall x \in \text{co}(u_i) \setminus C$, $\exists y \in C$ and $\boldsymbol{\lambda}_i \in \Delta(A_i)$ s.t.

(1) y is closest to x in C: $d(x,y) = d(x,C)$; and

(2) hyperplane $H(x-y,(x-y)\cdot y)$ separates x from $U_i(\boldsymbol{\lambda}_i)$:
(i) $x \in H^+(x-y,(x-y)\cdot y) \setminus H(x-y,(x-y)\cdot y)$ and (ii) $U_i(\boldsymbol{\lambda}_i) \subseteq H^-(x-y,(x-y)\cdot y)$.

(ii) is equiv. to $\forall \boldsymbol{\lambda}_{-i}$, $(u_i(\boldsymbol{\lambda}_i,\boldsymbol{\lambda}_{-i}) - y)\cdot(x-y) \leq 0$.

Approachability

Definition

Closed set C is **B-set** for player i if $\forall x \in \text{co}(u_i) \setminus C$, $\exists y \in C$ and $\boldsymbol{\lambda}_i \in \Delta(A_i)$ s.t.

(1) y is closest to x in C: $d(x,y) = d(x,C)$; and

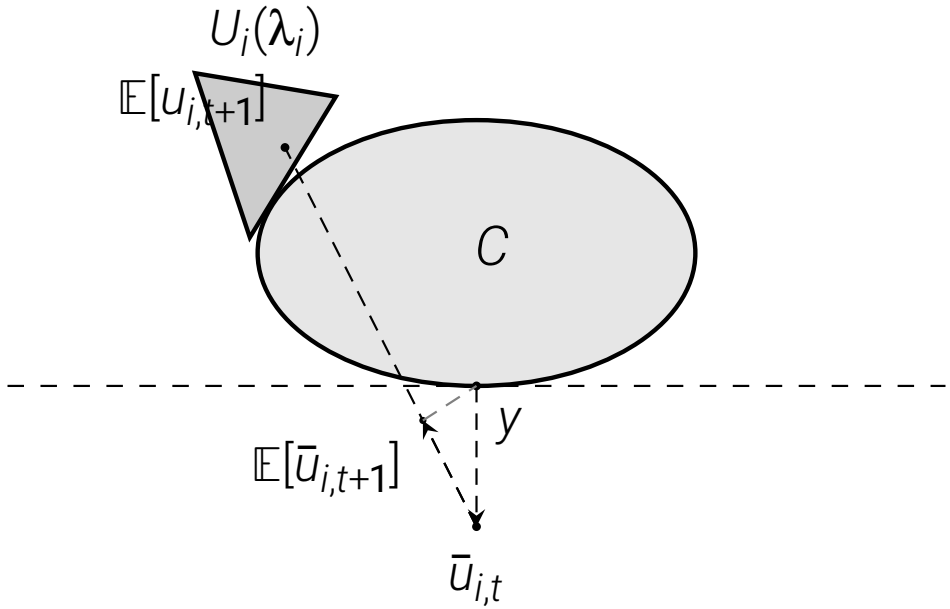
(2) hyperplane $H(x-y,(x-y)\cdot y)$ separates x from $U_i(\boldsymbol{\lambda}_i)$:
(i) $x \in H^+(x-y,(x-y)\cdot y) \setminus H(x-y,(x-y)\cdot y)$ and (ii) $U_i(\boldsymbol{\lambda}_i) \subseteq H^-(x-y,(x-y)\cdot y)$.

(ii) is equiv. to $\forall \boldsymbol{\lambda}_{-i}$, $(u_i(\boldsymbol{\lambda}_i,\boldsymbol{\lambda}_{-i}) - y)\cdot(x-y) \leq 0$.

Theorem (Blackwell 1956)

If C is B-set for player i, then it is approachable by player i.

Proving Approachability Theorem: One-Step Projection



Setup for One Step

(1) Fix $t \geq 1$. Let $\bar{u}_{i,t-1} \in \text{co}(u_i)$.

(2) If $\bar{u}_{i,t-1} \in C$, play anything (e.g., a_1). Otherwise, let $y_{t-1} \in C$ be a closest point: $d(\bar{u}_{i,t-1}, C) = \|\bar{u}_{i,t-1} - y_{t-1}\|$.

(3) By B-set, $\exists\, \lambda_{i,t} \in \Delta(A_i)$ s.t. hyperplane $H(\bar{u}_{i,t-1} - y_{t-1}, (\bar{u}_{i,t-1} - y_{t-1}) \cdot y_{t-1})$ separates $\bar{u}_{i,t-1}$ from $U_i(\lambda_{i,t})$:

$$(\bar{u}_{i,t-1} - y_{t-1}) \cdot \big(u_i(\lambda_{i,t}, \lambda_{-i}) - y_{t-1}\big) \leq 0, \quad \forall\, \lambda_{-i} \in \Delta(A_{-i}).$$

(4) Play $\lambda_{i,t}$ at stage t.

Proving Approachability Theorem: One-Step Projection Inequality

Key Inequality

Update: $\bar{u}_{i,t} = \bar{u}_{i,t-1} + \frac{1}{t}\left(u_i(a_t) - \bar{u}_{i,t-1}\right)$. Recall $d(\bar{u}_{t-1}, C) = \|\bar{u}_{i,t-1} - y_{t-1}\|$.

Proving Approachability Theorem: One-Step Projection Inequality

Key Inequality

Update: $\bar{u}_{i,t} = \bar{u}_{i,t-1} + \frac{1}{t}\left(u_i(a_t) - \bar{u}_{i,t-1}\right)$. Recall $d(\bar{u}_{t-1}, C) = \|\bar{u}_{i,t-1} - y_{t-1}\|$.

Identity:

$$\|\bar{u}_{i,t} - y_{t-1}\|^2 = \|\bar{u}_{i,t-1} - y_{t-1}\|^2 + \frac{2}{t}(\bar{u}_{i,t-1} - y_{t-1}) \cdot (u_i(a_t) - \bar{u}_{i,t-1}) + \frac{1}{t^2}\|u_i(a_t) - \bar{u}_{i,t-1}\|^2.$$

Proving Approachability Theorem: One-Step Projection Inequality

Key Inequality

Update: $\bar{u}_{i,t} = \bar{u}_{i,t-1} + \frac{1}{t}\left(u_i(a_t) - \bar{u}_{i,t-1}\right)$. Recall $d(\bar{u}_{t-1}, C) = \|\bar{u}_{i,t-1} - y_{t-1}\|$.

Identity:

$$\|\bar{u}_{i,t} - y_{t-1}\|^2 = \|\bar{u}_{i,t-1} - y_{t-1}\|^2 + \frac{2}{t}(\bar{u}_{i,t-1} - y_{t-1}) \cdot (u_i(a_t) - \bar{u}_{i,t-1}) + \frac{1}{t^2}\|u_i(a_t) - \bar{u}_{i,t-1}\|^2.$$

Write $(\bar{u}_{i,t-1} - y_{t-1}) \cdot (u_i(a_t) - \bar{u}_{i,t-1}) = (\bar{u}_{i,t-1} - y_{t-1}) \cdot (u_i(a_t) - y_{t-1}) - \|\bar{u}_{i,t-1} - y_{t-1}\|^2$.

Proving Approachability Theorem: One-Step Projection Inequality

Key Inequality

Update: $\bar{u}_{i,t} = \bar{u}_{i,t-1} + \frac{1}{t}\left(u_i(a_t) - \bar{u}_{i,t-1}\right)$. Recall $d(\bar{u}_{t-1}, C) = \|\bar{u}_{i,t-1} - y_{t-1}\|$.

Identity:

$$\|\bar{u}_{i,t} - y_{t-1}\|^2 = \|\bar{u}_{i,t-1} - y_{t-1}\|^2 + \frac{2}{t}(\bar{u}_{i,t-1} - y_{t-1}) \cdot (u_i(a_t) - \bar{u}_{i,t-1}) + \frac{1}{t^2}\|u_i(a_t) - \bar{u}_{i,t-1}\|^2.$$

Write $(\bar{u}_{i,t-1} - y_{t-1}) \cdot (u_i(a_t) - \bar{u}_{i,t-1}) = (\bar{u}_{i,t-1} - y_{t-1}) \cdot (u_i(a_t) - y_{t-1}) - \|\bar{u}_{i,t-1} - y_{t-1}\|^2$.

Conditional expectation: by separation,

$$\mathbb{E}\left[(\bar{u}_{i,t-1} - y_{t-1}) \cdot (u_i(a_t) - y_{t-1}) \mid H_{t-1}\right] \leq \max_{\lambda_{-i}} (\bar{u}_{i,t-1} - y_{t-1}) \cdot (u_i(\lambda_{i,t}, \lambda_{-i}) - y_{t-1}) \leq 0.$$

Proving Approachability Theorem: One-Step Projection Inequality

Key Inequality

Update: $\bar{u}_{i,t} = \bar{u}_{i,t-1} + \frac{1}{t}\left(u_i(a_t) - \bar{u}_{i,t-1}\right)$. Recall $d(\bar{u}_{t-1}, C) = \|\bar{u}_{i,t-1} - y_{t-1}\|$.

Identity:

$$\|\bar{u}_{i,t} - y_{t-1}\|^2 = \|\bar{u}_{i,t-1} - y_{t-1}\|^2 + \frac{2}{t}(\bar{u}_{i,t-1} - y_{t-1}) \cdot (u_i(a_t) - \bar{u}_{i,t-1}) + \frac{1}{t^2}\|u_i(a_t) - \bar{u}_{i,t-1}\|^2.$$

Write $(\bar{u}_{i,t-1} - y_{t-1}) \cdot (u_i(a_t) - \bar{u}_{i,t-1}) = (\bar{u}_{i,t-1} - y_{t-1}) \cdot (u_i(a_t) - y_{t-1}) - \|\bar{u}_{i,t-1} - y_{t-1}\|^2$.

Conditional expectation: by separation,

$$\mathbb{E}\left[(\bar{u}_{i,t-1} - y_{t-1}) \cdot (u_i(a_t) - y_{t-1}) \mid H_{t-1}\right] \leq \max_{\lambda_{-i}} (\bar{u}_{i,t-1} - y_{t-1}) \cdot (u_i(\lambda_{i,t}, \lambda_{-i}) - y_{t-1}) \leq 0.$$

Let $L := \max_{a \in A} \|u_i(a)\| < \infty$. Then $\mathbb{E}[\|u_i(a_t) - \bar{u}_{i,t-1}\|^2 \mid H_{t-1}] \leq 4L^2$.

Hence $\mathbb{E}\left[\|\bar{u}_t - y_{t-1}\|^2 \mid H_{t-1}\right] \leq \left(1 - \frac{2}{t}\right)\|\bar{u}_{i,t-1} - y_{t-1}\|^2 + \frac{4L^2}{t^2}$.

Proving Approachability Theorem: One-Step Projection Inequality

Key Inequality

Update: $\bar{u}_{i,t} = \bar{u}_{i,t-1} + \frac{1}{t}\left(u_i(a_t) - \bar{u}_{i,t-1}\right)$. Recall $d(\bar{u}_{t-1}, C) = \|\bar{u}_{i,t-1} - y_{t-1}\|$.

Identity:

$$\|\bar{u}_{i,t} - y_{t-1}\|^2 = \|\bar{u}_{i,t-1} - y_{t-1}\|^2 + \frac{2}{t}(\bar{u}_{i,t-1} - y_{t-1}) \cdot (u_i(a_t) - \bar{u}_{i,t-1}) + \frac{1}{t^2}\|u_i(a_t) - \bar{u}_{i,t-1}\|^2.$$

Write $(\bar{u}_{i,t-1} - y_{t-1}) \cdot (u_i(a_t) - \bar{u}_{i,t-1}) = (\bar{u}_{i,t-1} - y_{t-1}) \cdot (u_i(a_t) - y_{t-1}) - \|\bar{u}_{i,t-1} - y_{t-1}\|^2$.

Conditional expectation: by separation,

$$\mathbb{E}\left[(\bar{u}_{i,t-1} - y_{t-1}) \cdot (u_i(a_t) - y_{t-1}) \mid H_{t-1}\right] \leq \max_{\lambda_{-i}} (\bar{u}_{i,t-1} - y_{t-1}) \cdot (u_i(\lambda_{i,t}, \lambda_{-i}) - y_{t-1}) \leq 0.$$

Let $L := \max_{a \in A} \|u_i(a)\| < \infty$. Then $\mathbb{E}[\|u_i(a_t) - \bar{u}_{i,t-1}\|^2 \mid H_{t-1}] \leq 4L^2$.

Hence $\mathbb{E}\left[\|\bar{u}_t - y_{t-1}\|^2 \mid H_{t-1}\right] \leq \left(1 - \frac{2}{t}\right)\|\bar{u}_{i,t-1} - y_{t-1}\|^2 + \frac{4L^2}{t^2}$.

Since $d(\bar{u}_{i,t}, C) \leq \|\bar{u}_{i,t} - y_{t-1}\|$, setting $V_t := d(\bar{u}_{i,t}, C)^2$ gives

$$\mathbb{E}[V_t \mid H_{t-1}] \leq \left(1 - \frac{2}{t}\right)V_{t-1} + \frac{4L^2}{t^2}.$$

Convergence of Non-negative Almost Supermartingales

Theorem (Robbins and Siegmund 1971)

Let $(\mathcal{F}_t)$ be filtration and $(V_t)_{t\geq 0}$ be nonnegative, adapted. Suppose there are nonnegative, $\mathcal{F}_t$-adapted processes $(\xi_t), (\beta_t), (\zeta_t)$ s.t.

$$\mathbb{E}[V_{t+1} \mid \mathcal{F}_t] \;\leq\; (1+\xi_t)V_t \;-\; \zeta_t \;+\; \beta_t, \qquad t \geq 0,$$

with $\sum_{t=0}^{\infty} \xi_t < \infty$ and $\sum_{t=0}^{\infty} \beta_t < \infty$ a.s.

Then, V_t converges a.s. to a finite, nonnegative limit V_∞, and $\sum_{t=0}^{\infty} \zeta_t < \infty$ a.s.

Going beyond Doob's MCT: convergence for non-negative almost supermartingales.

Convergence of Non-negative Almost Supermartingales

Theorem (Robbins and Siegmund 1971)

Let $(\mathcal{F}_t)$ be filtration and $(V_t)_{t\geq 0}$ be nonnegative, adapted. Suppose there are nonnegative, $\mathcal{F}_t$-adapted processes $(\xi_t), (\beta_t), (\zeta_t)$ s.t.

$$\mathbb{E}[V_{t+1} \mid \mathcal{F}_t] \;\leq\; (1+\xi_t)V_t \;-\; \zeta_t \;+\; \beta_t, \qquad t \geq 0,$$

with $\sum_{t=0}^{\infty} \xi_t < \infty$ and $\sum_{t=0}^{\infty} \beta_t < \infty$ a.s.

Then, V_t converges a.s. to a finite, nonnegative limit V_∞, and $\displaystyle\sum_{t=0}^{\infty} \zeta_t < \infty$ a.s.

Corollary

If nonnegative (V_t) satisfies $\mathbb{E}[V_t \mid H_{t-1}] \leq (1-\alpha_t)V_{t-1}+\beta_t$ with $\sum_t \alpha_t = \infty$ and $\sum_t \beta_t < \infty$, then $V_t \to 0$ a.s.

Useful corollary: $\xi_t = 0$, $\zeta_t = \alpha_t V_t$ with $\alpha_t \in [0,1]$. If $\sum_t \alpha_t = \infty$, then $V_\infty = 0$ a.s.

Since $\sum_t \alpha_t V_t < \infty$; if $\sum_t \alpha_t = \infty$, only possible limit V_∞ is 0.

Proving Approachability Theorem: Concluding

Concluding the Proof of Blackwell's Approachability Theorem

(1) Use the one-step choice $\lambda_{i,t}$ from the B-set condition at $\bar{u}_{i,t-1}$.

(2) $\mathbb{E}[V_t \mid H_{t-1}] \leq (1 - \frac{2}{t})V_{t-1} + \frac{4L^2}{t^2}$.

(3) Take $\alpha_t = \frac{2}{t}$ (diverges) and $\beta_t = \frac{4L^2}{t^2}$ (summable).

(4) Robbins–Siegmund $\implies$ $V_t = d(\bar{u}_t, C)^2 \to 0$ a.s.

Proving Approachability Theorem: Concluding

Concluding the Proof of Blackwell's Approachability Theorem

(1) Use the one-step choice $\lambda_{i,t}$ from the B-set condition at $\bar{u}_{i,t-1}$.

(2) $\mathbb{E}[V_t \mid H_{t-1}] \leq (1 - \frac{2}{t})V_{t-1} + \frac{4L^2}{t^2}$.

(3) Take $\alpha_t = \frac{2}{t}$ (diverges) and $\beta_t = \frac{4L^2}{t^2}$ (summable).

(4) Robbins–Siegmund $\Longrightarrow$ $V_t = d(\bar{u}_t, C)^2 \to 0$ a.s.

Theorem (Blackwell 1956)

If C is B-set for player i, then it is approachable by player i.

Strategy (Blackwell's rule): at each t, project $\bar{u}_{i,t-1}$ onto C, pick $\lambda_{i,t}$ separating $\bar{u}_{i,t-1}$ from $U_i(\lambda_{i,t})$, play $\lambda_{i,t}$.

Generalisations and Variations

Theorem (Blackwell 1956)

If C is B-set for player i, then it is approachable by player i.

Generalisations and Variations:

Lehrer (2002 IJGT): generalises Blackwell's approachability theorem to infinite-dimensional spaces.

Hou (1971 AMS): A closed set C is approachable by player i if and only if it contains a B-set for player i.

Overview

1. Learning in Games

2. Approachability
   - Multi-Utility Representation
   - Approachability
   - Application: Picking Experts
   - Universal Consistency

3. Calibration

4. Adaptive Algorithms

5. Sophisticated Learning

Application: Picking Experts

S states of nature. A actions. Payoffs $u : A \times S \to \mathbb{R}$. Set of experts E.

Every period t,

(1) state s^t realises,

(2) each expert recommends action $a_{e,t} \in A$,

(3) DM chooses which expert to follow $e_t \in E$ and adopts their recommended action,

(4) payoffs realise, and DM observes s_t.

Application: Picking Experts

S states of nature. A actions. Payoffs $u : A \times S \to \mathbb{R}$. Set of experts E.

History: $h_t \in H_t := (S \times A^{|E|} \times E)^{t-1}$ (previous states, what each expert recommended, expert chosen). $H := \cup_t H_t$

Strategy: $\boldsymbol{\sigma} : H \to \Delta(E)$. Distribution of $s_t \sim \boldsymbol{\gamma}_t \in \Delta(S)$.

Average payoff: $\bar{u}_T(\boldsymbol{\sigma}, \boldsymbol{\gamma}) := \frac{1}{T} \sum_{t \leq T} \sum_e \boldsymbol{\sigma}(h_t)(e)\boldsymbol{\gamma}(s)u(a_{e,t}, s)$.

Payoff from following particular expert e: $\bar{u}_T(e, \boldsymbol{\gamma})$.

Application: Picking Experts

S states of nature. A actions. Payoffs $u : A \times S \to \mathbb{R}$. Set of experts E.

History: $h_t \in H_t := (S \times A^{|E|} \times E)^{t-1}$ (previous states, what each expert recommended, expert chosen). $H := \cup_t H_t$

Strategy: $\boldsymbol{\sigma} : H \to \Delta(E)$. Distribution of $s_t \sim \boldsymbol{\gamma}_t \in \Delta(S)$.

Average payoff: $\bar{u}_T(\boldsymbol{\sigma}, \boldsymbol{\gamma}) := \frac{1}{T} \sum_{t \leq T} \sum_e \boldsymbol{\sigma}(h_t)(e)\boldsymbol{\gamma}(s)u(a_{e,t}, s)$.

Payoff from following particular expert e: $\bar{u}_T(e, \boldsymbol{\gamma})$.

Small problem: DM doesn't know what experts actually know, whether have full info, partial, no info, biased, etc.

Definition

DM's $\boldsymbol{\sigma}$ is **no-regret strategy** if $\forall e \in E$ and each sequence $s_1, s_2, ...,$

$$\mathbb{P}^{\boldsymbol{\sigma}} \left(\liminf_{t \to \infty} \bar{u}_t(\boldsymbol{\sigma}, \boldsymbol{\gamma}) - \bar{u}_t(e, \boldsymbol{\gamma}) \geq 0 \right) = 1.$$

Does no-regret strategy even exist? Can we characterise it?

Application: Picking Experts

Theorem

The DM has a no-regret strategy.

Simplying assumption: $|E| = |A|$ and for each e, $a_{e,t} = a$ for some different a.

Application: Picking Experts

Theorem

The DM has a no-regret strategy.

Simplying assumption: $|E| = |A|$ and for each e, $a_{e,t} = a$ for some different a.

Proof

Opponent: nature, choosing $\boldsymbol{\sigma_0} : H \to \Delta(S)$. $C := \mathbb{R}^{|E|}_+$.

Let $v(\boldsymbol{\lambda}, \boldsymbol{\gamma}) := (u(\boldsymbol{\lambda}, \boldsymbol{\gamma}) - u(e, \boldsymbol{\gamma}))_{e \in E} \in \mathbb{R}^{|E|}$ and $\bar{v}_T := \frac{1}{T} \sum_{t \leq T} v(\boldsymbol{\sigma}(h_t), \boldsymbol{\sigma_0}(h_t))$. *Regret vector:* no-regret $\iff \liminf_t \bar{v}_t \in C$.

For $x \in \mathbb{R}^{|E|}$, projection onto C is $y := x^+$ (positive part), and the normal is $x^- := y - x$ (negative part).

Choose the *Blackwell action* at x: if $\sum_e x_e^- > \mathbf{0}$, set

$$\boldsymbol{\lambda}^x(e) := \frac{x_e^-}{\sum_{e'} x_{e'}^-}$$ (put weight on experts relative to which you are behind),

and any $\boldsymbol{\lambda}^x$ if $x \in C$.

Application: Picking Experts – Proof (via Blackwell)

Proof

$\boldsymbol{\sigma_0} : H \to \Delta(S)$. $C := \mathbb{R}^{|E|}_+$. $v(\boldsymbol{\lambda}, \boldsymbol{\gamma}) := (u(\boldsymbol{\lambda}, \boldsymbol{\gamma}) - u(e, \boldsymbol{\gamma}))_{e \in E} \in \mathbb{R}^{|E|}$; $\bar{v}_T := \frac{1}{T} \sum_{t \leq T} v(\boldsymbol{\sigma}(h_t), \boldsymbol{\sigma_0}(h_t))$.

For $x \in \mathbb{R}^{|E|}$, $y := x^+$, $x^- := y - x$. For $\neg(x \geq \mathbf{0})$, set $\boldsymbol{\lambda}^x(e) := \frac{x^-_e}{\sum_{e'} x^-_{e'}}$.

Application: Picking Experts – Proof (via Blackwell)

Proof

$\sigma_0 : H \to \Delta(S)$. $C := \mathbb{R}^{|E|}_+$. $v(\lambda, \gamma) := (u(\lambda, \gamma) - u(e, \gamma))_{e \in E} \in \mathbb{R}^{|E|}$; $\bar{v}_T := \frac{1}{T} \sum_{t \le T} v(\sigma(h_t), \sigma_0(h_t))$.

For $x \in \mathbb{R}^{|E|}$, $y := x^+$, $x^- := y - x$. For $\neg(x \ge 0)$, set $\lambda^x(e) := \frac{x^-_e}{\sum_{e'} x^-_{e'}}$.

For any opponent choice $\gamma \in \Delta(S)$,

(i) $\neg(x \ge 0) \implies \|x^-\| = \|x - y\| > 0 \implies x \in H^+(x - y, (x - y) \cdot y) \setminus H(x - y, (x - y) \cdot y)$;

Application: Picking Experts – Proof (via Blackwell)

Proof

$\sigma_0 : H \to \Delta(S)$. $C := \mathbb{R}^{|E|}_+$. $v(\lambda, \gamma) := (u(\lambda, \gamma) - u(e, \gamma))_{e \in E} \in \mathbb{R}^{|E|}$; $\bar{v}_T := \frac{1}{T} \sum_{t \le T} v(\sigma(h_t), \sigma_0(h_t))$.

For $x \in \mathbb{R}^{|E|}$, $y := x^+$, $x^- := y - x$. For $\neg(x \ge 0)$, set $\lambda^x(e) := \frac{x^-_e}{\sum_{e'} x^-_{e'}}$.

For any opponent choice $\gamma \in \Delta(S)$,

(i) $\neg(x \ge 0) \implies \|x^-\| = \|x - y\| > 0 \implies x \in H^+(x-y, (x-y)\cdot y) \setminus H(x-y, (x-y)\cdot y)$;

(ii) $0 \ge (v(\lambda^x, \gamma) - y)\cdot(x-y) = (x^+ - v(\lambda^x, \gamma))\cdot x^- = x^+ \cdot x^- - v(\lambda^x, \gamma)\cdot x^- = -v(\lambda^x, \gamma)\cdot x^-$.

Note that

$$x^- \cdot v(\lambda^x, \gamma) = \sum_e x^-_e \Big(\sum_{e'} \lambda^x(e') u(e', \gamma) - u(e, \gamma)\Big)$$

$$= \sum_{e'} \frac{\sum_e x^-_e}{\sum_{e''} x^-_{e''}} x^-_{e'} u(e', \gamma) - \sum_e x^-_e u(e, \gamma) = \sum_{e'} x^-_{e'} u(e', \gamma) - \sum_e x^-_e u(e, \gamma) = 0.$$

Application: Picking Experts – Proof (via Blackwell)

Proof

$\boldsymbol{\sigma_0} : H \to \Delta(S)$. $C := \mathbb{R}^{|E|}_+$. $v(\boldsymbol{\lambda}, \boldsymbol{\gamma}) := (u(\boldsymbol{\lambda}, \boldsymbol{\gamma}) - u(e, \boldsymbol{\gamma}))_{e \in E} \in \mathbb{R}^{|E|}$; $\bar{v}_T := \frac{1}{T} \sum_{t \leq T} v(\boldsymbol{\sigma}(h_t), \boldsymbol{\sigma_0}(h_t))$.

For $x \in \mathbb{R}^{|E|}$, $y := x^+$, $x^- := y - x$. For $\neg(x \geq \mathbf{0})$, set $\boldsymbol{\lambda}^x(e) := \frac{x_e^-}{\sum_{e'} x_{e'}^-}$.

For any opponent choice $\boldsymbol{\gamma} \in \Delta(S)$,

(i) $\neg(x \geq \mathbf{0}) \implies \|x^-\| = \|x - y\| > \mathbf{0} \implies x \in H^+(x-y, (x-y)\cdot y) \setminus H(x-y, (x-y)\cdot y)$;

(ii) $\mathbf{0} \geq (v(\boldsymbol{\lambda}^x, \boldsymbol{\gamma}) - y)\cdot(x-y) = (x^+ - v(\boldsymbol{\lambda}^x, \boldsymbol{\gamma}))\cdot x^- = x^+ \cdot x^- - v(\boldsymbol{\lambda}^x, \boldsymbol{\gamma})\cdot x^- = -v(\boldsymbol{\lambda}^x, \boldsymbol{\gamma})\cdot x^-$.

Note that

$$x^- \cdot v(\boldsymbol{\lambda}^x, \boldsymbol{\gamma}) = \sum_e x_e^- \Big(\sum_{e'} \boldsymbol{\lambda}^x(e') u(e', \boldsymbol{\gamma}) - u(e, \boldsymbol{\gamma})\Big)$$

$$= \sum_{e'} \frac{\sum_e x_e^-}{\sum_{e''} x_{e''}^-} x_{e'}^- u(e', \boldsymbol{\gamma}) - \sum_e x_e^- u(e, \boldsymbol{\gamma}) = \sum_{e'} x_{e'}^- u(e', \boldsymbol{\gamma}) - \sum_e x_e^- u(e, \boldsymbol{\gamma}) = \mathbf{0}.$$

Hence C is a B-set. Blackwell's theorem $\implies$ $\bar{v}_t$ approaches C a.s., i.e.,

$\liminf_{t \to \infty} \left(\bar{u}_t(\boldsymbol{\sigma}, \boldsymbol{\sigma_0}) - \bar{u}_t(e, \boldsymbol{\sigma_0})\right) \geq \mathbf{0}$ for all $e \in E$ and any of nature's moves $\boldsymbol{\sigma_0}$.

Application: Picking Experts – Proof (via Blackwell)

Proof

$\boldsymbol{\sigma_0} : H \to \Delta(S)$. $C := \mathbb{R}^{|E|}_+$. $v(\boldsymbol{\lambda}, \boldsymbol{\gamma}) := (u(\boldsymbol{\lambda}, \boldsymbol{\gamma}) - u(e, \boldsymbol{\gamma}))_{e \in E} \in \mathbb{R}^{|E|}$; $\bar{v}_T := \frac{1}{T} \sum_{t \leq T} v(\boldsymbol{\sigma}(h_t), \boldsymbol{\sigma_0}(h_t))$.

For $x \in \mathbb{R}^{|E|}$, $y := x^+$, $x^- := y - x$. For $\neg(x \geq \mathbf{0})$, set $\boldsymbol{\lambda}^x(e) := \frac{x^-_e}{\sum_{e'} x^-_{e'}}$.

For any opponent choice $\boldsymbol{\gamma} \in \Delta(S)$,

(i) $\neg(x \geq \mathbf{0}) \implies \|x^-\| = \|x - y\| > \mathbf{0} \implies x \in H^+(x-y, (x-y)\cdot y) \setminus H(x-y, (x-y)\cdot y)$;

(ii) $\mathbf{0} \geq (v(\boldsymbol{\lambda}^x, \boldsymbol{\gamma}) - y)\cdot(x-y) = (x^+ - v(\boldsymbol{\lambda}^x, \boldsymbol{\gamma}))\cdot x^- = x^+ \cdot x^- - v(\boldsymbol{\lambda}^x, \boldsymbol{\gamma})\cdot x^- = -v(\boldsymbol{\lambda}^x, \boldsymbol{\gamma})\cdot x^-$.

Note that

$$x^- \cdot v(\boldsymbol{\lambda}^x, \boldsymbol{\gamma}) = \sum_e x^-_e \Big(\sum_{e'} \boldsymbol{\lambda}^x(e') u(e', \boldsymbol{\gamma}) - u(e, \boldsymbol{\gamma})\Big)$$

$$= \sum_{e'} \frac{\sum_e x^-_e}{\sum_{e''} x^-_{e''}} x^-_{e'} u(e', \boldsymbol{\gamma}) - \sum_e x^-_e u(e, \boldsymbol{\gamma}) = \sum_{e'} x^-_{e'} u(e', \boldsymbol{\gamma}) - \sum_e x^-_e u(e, \boldsymbol{\gamma}) = \mathbf{0}.$$

Hence C is a B-set. Blackwell's theorem $\implies$ $\bar{v}_t$ approaches C a.s., i.e.,

$\liminf_{t \to \infty} \left(\bar{u}_t(\boldsymbol{\sigma}, \boldsymbol{\sigma_0}) - \bar{u}_t(e, \boldsymbol{\sigma_0})\right) \geq \mathbf{0}$ for all $e \in E$ and any of nature's moves $\boldsymbol{\sigma_0}$.

Strategy (implementable): compute current average regrets $x := \bar{v}_{t-1}$; if $x \notin C$ play $\boldsymbol{\lambda}^x$.

Overview

1. Learning in Games

2. Approachability
   - Multi-Utility Representation
   - Approachability
   - Application: Picking Experts
   - Universal Consistency

3. Calibration

4. Adaptive Algorithms

5. Sophisticated Learning

Hannan (1957): Setting

Repeated decision problem

Actions $A = \{1, \ldots, |A|\}$; states $S = \{1, \ldots, |S|\}$.

Bounded stage payoff $u : A \times S \to [0, 1]$.

At t: DM chooses $a_t \in A$; Nature reveals $s_t \in S$; payoff $u(a_t, s_t)$.

Hannan (1957): Setting

Repeated decision problem

Actions $A = \{1, \ldots, |A|\}$; states $S = \{1, \ldots, |S|\}$.

Bounded stage payoff $u : A \times S \to [0, 1]$.

At t: DM chooses $a_t \in A$; Nature reveals $s_t \in S$; payoff $u(a_t, s_t)$.

Beliefs and empirical distribution

$\hat{p}_t \in \Delta(S)$: empirical distribution of $(s_\ell)_{\ell \leq t}$.

Bayes payoff of action a against $p \in \Delta(S)$: $U(a, p) := \mathbb{E}_{s \sim p}[u(a, s)]$.

Hannan (1957): Setting

Repeated decision problem

Actions $A = \{1, \ldots, |A|\}$; states $S = \{1, \ldots, |S|\}$.

Bounded stage payoff $u : A \times S \to [0, 1]$.

At t: DM chooses $a_t \in A$; Nature reveals $s_t \in S$; payoff $u(a_t, s_t)$.

Beliefs and empirical distribution

$\hat{p}_t \in \Delta(S)$: empirical distribution of $(s_\ell)_{\ell \leq t}$.

Bayes payoff of action a against $p \in \Delta(S)$: $U(a, p) := \mathbb{E}_{s \sim p}[u(a, s)]$.

Benchmark: (External) Regret

For any fixed $a \in A$, $\bar{u}_T(a) := \dfrac{1}{T} \sum_{t \leq T} u(a, s_t)$.

Average payoff: $\bar{u}_T(\sigma) := \dfrac{1}{T} \sum_{t \leq T} u(a_t, s_t)$.

No (External) Regret: $\liminf_{T \to \infty} \left(\bar{u}_T(\sigma) - \max_a \bar{u}_T(a) \right) \geq 0$ a.s.

External regret: comparison relative to swapping to fixed action.

Hannan's Procedure (1957)

Smoothed fictitious play

Fix a full-support $\nu_t \in \Delta(S)$ iid and a sequence (γ_t) with $\gamma_t \downarrow 0$ and $\sum_t \gamma_t < \infty$.

At $t \geq 1$, form the **smoothed empirical belief**

$$p_t := (1 - \gamma_t)\,\hat{p}_{t-1} + \gamma_t\,\nu_t.$$

Choose a Bayes action (or mixed Bayes rule) against p_t:

$$a_t \in \arg\max_{a \in A} u(a, p_t) \qquad \left(\text{or } \lambda_t \in \arg\max_{\lambda \in \Delta(A)} \sum_a \lambda(a) u(a, p_t)\right).$$

Ties are broken by a fixed rule; γ_t prevents zero exploration.

Interpretation: best respond to slightly perturbed empirical model; perturbations vanish.

Hannan Consistency

Theorem (Hannan 1957)

Under SFP with bounded payoffs and (γ_t) as above,

$$\liminf_{T\to\infty}\left(\bar{u}_T(\sigma) - \max_{a\in A}\bar{u}_T(a)\right) \geq 0 \qquad \text{almost surely.}$$

Equivalently, for every $a \in A$, $\liminf_{T\to\infty}\left(\bar{u}_T(\sigma) - \bar{u}_T(a)\right) \geq 0$ a.s.

Full-support ν_t ensures absolute continuity of beliefs; (γ_t) controls approximation error.

Proof Idea

Key steps

Let $v_t(a) := u(a_t, s_t) - u(a, s_t)$; regret against a is $\bar{v}_T(a) = \bar{u}_T(\sigma) - \bar{u}_T(a)$.

Conditional on H_{t-1}, a_t maximises $u(\cdot, p_t)$, hence for all a,

$$\mathbb{E}[v_t(a) \mid H_{t-1}] = u(a_t, p_t) - u(a, p_t) + \underbrace{\Big(\mathbb{E}[u(a_t, s_t) - u(a, s_t) \mid H_{t-1}] - [u(a_t, p_t) - u(a, p_t)]\Big)}_{\text{model error}}.$$

The first term ≥ 0 by optimality of a_t for p_t.

The model-error term is $O(\gamma_t)$ since $p_t - (1-\gamma_t)\hat{p}_{t-1}$ has mass γ_t and payoffs are bounded in $[0, 1]$.

Summing and dividing by T gives $\mathbb{E}[\bar{v}_T(a)] \geq -\frac{1}{T}\sum_{t \leq T} c\,\gamma_t$; with $\sum_t \gamma_t < \infty$, $\liminf_T \mathbb{E}[\bar{v}_T(a)] \geq 0$.

A standard martingale SLLN upgrades to a.s. statements (bounded differences).

Direct No-Regret via Approachability

History $H = \cup_t H_t$, $H_t := (A \times S)^{t-1}$. Strategy $\boldsymbol{\sigma} : H \to \Delta(A)$; Nature $\boldsymbol{\sigma_0} : H \to \Delta(S)$.

Define **regret vector** $v(\boldsymbol{\lambda}, \boldsymbol{\gamma}) \in \mathbb{R}^A$ with $v(\boldsymbol{\lambda}, \boldsymbol{\gamma}) := (u(\boldsymbol{\lambda}, \boldsymbol{\gamma}) - u(a, \boldsymbol{\gamma}))_{a \in A}$.

$\bar{v}_T := \frac{1}{T} \sum_{t \leq T} v(\boldsymbol{\sigma}(h_t), \boldsymbol{\sigma_0}(h_t))$.

Target set $C = \mathbb{R}^A_+$ (no external regret).

Let $x \in \mathbb{R}^{|A|}$, $y := x^+$, $x^- := y - x$, and choose

$$\boldsymbol{\lambda}^x(a) = \frac{x^-_a}{\sum_{a'} x^-_{a'}} \quad \text{if } x \notin C, \qquad \text{any } \boldsymbol{\lambda}^x \text{ if } x \in C.$$

Then $x^- \cdot \mathbb{E}[r_t \mid H_{t-1}] = \mathbf{0} \leq \mathbf{0}$; C is a B-set $\implies \bar{v}_T \to C$ a.s.

Consequence: $\liminf_T \left(\bar{u}_T(\boldsymbol{\sigma}) - \bar{u}_T(a)\right) \geq \mathbf{0}$ for all a.

Direct No-Regret via Approachability

History $H = \cup_t H_t$, $H_t := (A \times S)^{t-1}$. Strategy $\boldsymbol{\sigma} : H \to \Delta(A)$; Nature $\boldsymbol{\sigma_0} : H \to \Delta(S)$.

Define **regret vector** $v(\boldsymbol{\lambda}, \boldsymbol{\gamma}) \in \mathbb{R}^A$ with $v(\boldsymbol{\lambda}, \boldsymbol{\gamma}) := (u(\boldsymbol{\lambda}, \boldsymbol{\gamma}) - u(a, \boldsymbol{\gamma}))_{a \in A}$.

$\bar{v}_T := \frac{1}{T} \sum_{t \leq T} v(\boldsymbol{\sigma}(h_t), \boldsymbol{\sigma_0}(h_t))$.

Target set $C = \mathbb{R}^A_+$ (no external regret).

Let $x \in \mathbb{R}^{|A|}$, $y := x^+$, $x^- := y - x$, and choose

$$\boldsymbol{\lambda}^x(a) = \frac{x^-_a}{\sum_{a'} x^-_{a'}} \quad \text{if } x \notin C, \qquad \text{any } \boldsymbol{\lambda}^x \text{ if } x \in C.$$

Then $x^- \cdot \mathbb{E}[r_t \mid H_{t-1}] = \mathbf{0} \leq \mathbf{0}$; C is a B-set $\implies \bar{v}_T \to C$ a.s.

Consequence: $\liminf_T \left(\bar{u}_T(\boldsymbol{\sigma}) - \bar{u}_T(a)\right) \geq \mathbf{0}$ for all a.

Remark

Hannan (1957) and Blackwell (1956) yield same external no-regret guarantee; update rules differ (smoothed best reply vs projection-based mixture).

Implementation, Interpretation and Links

Choice of Smoothing and Rates:

Any full-support ν_t works; e.g. ν uniform on S.

Typical schedules: $\gamma_t = t^{-1-\varepsilon}$ (summable) for a.s. convergence via almost-supermartingale; or $\gamma_t \asymp t^{-1/2}$ for $O(1/\sqrt{T})$ expected regret.

Boundedness of u and tie-breaking rules ensure measurability and martingale applicability.

Implementation, Interpretation and Links

Choice of Smoothing and Rates:

Any full-support ν_t works; e.g. ν uniform on S.

Typical schedules: $\gamma_t = t^{-1-\varepsilon}$ (summable) for a.s. convergence via almost-supermartingale; or $\gamma_t \asymp t^{-1/2}$ for $O(1/\sqrt{T})$ expected regret.

Boundedness of u and tie-breaking rules ensure measurability and martingale applicability.

Exploration: $\gamma_t > 0$ avoids being trapped by early noise; $\gamma_t \downarrow 0$ removes bias.

Interpretation: "Fictitious play" against the environment with vanishing perturbations.

Implementation, Interpretation and Links

Choice of Smoothing and Rates:

Any full-support ν_t works; e.g. ν uniform on S.

Typical schedules: $\gamma_t = t^{-1-\varepsilon}$ (summable) for a.s. convergence via almost-supermartingale; or $\gamma_t \asymp t^{-1/2}$ for $O(1/\sqrt{T})$ expected regret.

Boundedness of u and tie-breaking rules ensure measurability and martingale applicability.

Exploration: $\gamma_t > 0$ avoids being trapped by early noise; $\gamma_t \downarrow 0$ removes bias.

Interpretation: "Fictitious play" against the environment with vanishing perturbations.

Connections: Approachability (Blackwell), SFP.

Scope: applies to arbitrary state sequences (adversarial or stochastic).

Overview

1. Learning in Games

2. Approachability

3. Calibration
   - Foster and Vohra (1997): Calibrated Learning & CE

4. Adaptive Algorithms

5. Sophisticated Learning

Overview

1. Learning in Games

2. Approachability

3. Calibration
   - Foster and Vohra (1997): Calibrated Learning & CE

4. Adaptive Algorithms

5. Sophisticated Learning

Learning in Games

Convergence issues of the learning in games:

Generalised FP: *if* converge, asymptotic behaviour is Nash-like; but convergence not assured.

Similar issues with replicator dynamic and other models.

Foster and Vohra (1997 GEB): different learning basis – *calibration* – yields different solution concept – correlated equilibrium.

Calibration: from learning literature (Dawid 1982 JASA)

Suppose that, in a long conceptually infinite sequence of weather forecasts, we look at all those days for which the forecast probability of precipitation was, say, close to some given value p and assuming these form an infinite sequence determine the long run proportion ρ *of such days on which the forecast event rain in fact occurred. The plot of* ρ *against p is termed the forecaster's empirical calibration curve. If the curve is the diagonal* $\rho = p$*, the forecaster may be termed well calibrated.*

Calibrated Learning: Setup

Stage game

Players $i \in \{1, 2\}$; actions A_i finite; $A = A_1 \times A_2$.

Payoffs $u_i : A \to \mathbb{R}$.

Repeated play ($t = 1, 2, \ldots$)

History $H_t := A^{t-1}$, $H := \cup_t H_t$.

Strategies $\sigma_i : H \to \Delta(A_i)$; realised actions $a_t = (a_{1,t}, a_{2,t})$.

Empirical distribution $\bar{\sigma}_t \in \Delta(A)$: $\bar{\sigma}_t(a) := \frac{1}{t} \sum_{s \leq t} 1_{\{a_s = a\}}$.

Forecasts and behaviour

Player i's forecasting rule $f^i_{-i} : H \to \Delta(A_{-i})$; issues forecasts $f^i_{-i}(h_t) = \sigma^i_{-i,t} \in \Delta(A_{-i})$.

Myopic best replies: $a_{i,t} \in \arg\max_{a'_i} u_i(a'_i, \sigma^i_{-i})$

Fix a deterministic tie-breaking rule.

Calibration (Partition Version)

Definition

Fix finite partition $\Pi_i = \{B^k_{-i}\}_{k=1}^K$ of $\Delta(A_{-i})$ and representative $\sigma^k_{-i} \in B^k$. For $t \geq 1$ set

$$N^k_{i,t} := \sum_{s \leq t} 1_{\{\sigma^j_{-i,s} \in B^k_{-i}\}}, \qquad \bar{\sigma}^k_{-i,t}(a_{-i}) := \frac{1}{N^k_{i,t}} \sum_{s \leq t} 1_{\{\sigma^j_{-i,s} \in B^k_{-i}\}} 1_{\{a_{-i,s} = a_{-i}\}} \text{ if } N^k_{i,t} > 0; \text{ ow } = 0.$$

The forecasting rule is **calibrated (wrt** $(a_{-i,t})$ **on** Π_i**)** if for every k

$$\lim_{t \to \infty} \|\bar{\sigma}^k_{-i,t} - \sigma^k_{-i}\| \frac{N_{i,t^k}}{t} = 0.$$

Intuition: on subsequence when $\sigma^j_{-i,t} \in B^k_{-i}$, empirical conditional frequency of $a_{-i,t}$ matches σ^k_{-i}.

Refining Π_i (mesh $\downarrow 0$) yields full calibration on $\Delta(A_{-i})$.

Existence of Calibrated Forecasters

Proposition (Foster and Vohra 1997 GEB)

For any finite partition $\mathbf{\Pi}_i$ there exists a (possibly randomised) forecasting scheme $f^i_{-i} : H \to \Delta(A_{-i})$ that is calibrated on $\mathbf{\Pi}_i$ a.s.

Proof idea (via Blackwell approachability)

Build *calibration vector* $z_t \in \mathbb{R}^{K|A_{-i}|}$ with components

$$z_t(k, a_{-i}) := \mathbf{1}_{\{\sigma^i_{-i,t} \in B^k_{-i}\}} \left(\mathbf{1}_{\{a_{-i,t}=a_{-i}\}} - \sigma^k_{-i}(a_{-i})\right).$$

Calibration vector average: $\bar{z}_t := \sum_{s \leq t} z_s / t$

Target set $C := \mathbb{R}^{k|A_{-i}|}_-$.

Note that $d(\bar{z}_t, C) \to 0 \iff \bar{z}_\infty(k, a_{-i}) \leq \sigma^k_{-i}(a_{-i})\ \forall a_{-i}$ and $k : \lim_{t\to\infty} N^k_{i,t}/t > 0$ $\iff \bar{z}_\infty(k, a_{-i}) = \sigma^k_{-i}(a_{-i})$.

At step t, choose e forecast cell k (i.e. $\sigma^i_{-i,t} \in B^k$) to satisfy Blackwell's separation for the current average $\bar{z}_{t-1}$.

Blackwell $\implies \bar{z}_t \to 0$ a.s. on each active cell $k \implies$ calibration on $\mathbf{\Pi}_i$.

Calibration via Blackwell Approachability: Vector Game

Fix finite forecast menu $S = \{\sigma_{-i}^1, \ldots, \sigma_{-i}^K\} \subset \Delta(A_{-i})$ that *contains all pure actions* $\{e_{a_{-i}}\}$.

At t, forecaster chooses index $k_t \in \{1, \ldots, K\}$ (announces $\sigma_{-i}^{k_t}$); Nature chooses $a_{-i,t} \in A_{-i}$.

Stage vector payoff $z_t \in \mathbb{R}^{K|A_{-i}|}$ with blocks $z_t(k, \cdot) \in \mathbb{R}^{|A_{-i}|}$:

$$z_t(k, a_{-i}) \;:=\; 1_{\{k_t = k\}} \left(1_{\{a_{-i,t} = a_{-i}\}} - \sigma_{-i}^k(a_{-i}) \right).$$

Averaging gives (for each k)

$$\bar{z}_t(k, \cdot) \;=\; \frac{N_t(k)}{t} \left(\bar{\sigma}_t^k - \sigma_{-i}^k \right), \quad N_t(k) := \sum_{s \le t} 1_{\{k_s = k\}}, \quad \bar{\sigma}_t^k(a_{-i}) := \frac{\sum_{s \le t} 1_{\{k_s = k\}} 1_{\{a_{-i,s} = a_{-i}\}}}{N_t(k)}.$$

Target set $C := \{0\} \subset \mathbb{R}^{K|A_{-i}|}$.

Goal: $\bar{z}_t \to 0$ a.s. $\iff$ for each k, $\|\bar{\sigma}_t^k - \sigma_{-i}^k\| \cdot N_t(k)/t \to 0$ (partition calibration on S).

Separation and Control Rule (choose k_t)

Separation at current average

Let $x := \bar{z}_{t-1} \in \mathbb{R}^{K|A_{-i}|}$ and write its k-block as $x_k \in \mathbb{R}^{|A_{-i}|}$.

If we announce cell k and Nature uses $\boldsymbol{\lambda}_{-i} \in \Delta(A_{-i})$, then

$$\mathbb{E}[z_t \mid H_{t-1}, k, \boldsymbol{\lambda}_{-i}] \cdot x \;=\; (\boldsymbol{\lambda}_{-i} - \boldsymbol{\sigma}_{-i}^{k}) \cdot x_k.$$

Choose an index k_t such that *its fixed representative is a maximiser of the linear form* $q \mapsto x_{k_t} \cdot q$ *over the whole simplex*:

$$\boldsymbol{\sigma}_{-i}^{k_t} \in \arg\max_{q \in \Delta(A_{-i})} x_{k_t} \cdot q.$$

This exists because a linear form on $\Delta(A_{-i})$ is maximised at an extreme point; since all pure actions are in S, we can take $\boldsymbol{\sigma}_{-i}^{k_t} = e_{j^*}$ where $j^* \in \arg\max_{a_{-i}} x_{k_t}(a_{-i})$.

Then for every $\boldsymbol{\lambda}_{-i}$,

$$\sup_{\boldsymbol{\lambda}_{-i} \in \Delta(A_{-i})} \mathbb{E}[z_t \mid H_{t-1}, k_t, \boldsymbol{\lambda}_{-i}] \cdot x = \Big(\max_{\boldsymbol{\lambda}_{-i}} x_{k_t} \cdot \boldsymbol{\lambda}_{-i}\Big) - x_{k_t} \cdot \boldsymbol{\sigma}_{-i}^{k_t} \;\le\; 0.$$

Conclusion

The origin $C = \{\mathbf{0}\}$ satisfies Blackwell's separation condition for this control rule. Hence C is a B-set and the forecast-[illegible]

From Approachability to Calibration (on S)

Blackwell's theorem (bounded vectors) $\Longrightarrow \bar{z}_t \to 0$ a.s.

For each k:

$$\bar{z}_t(k,\cdot) = \frac{N_t(k)}{t}\left(\hat{\lambda}_t^k - \sigma_{-i}^k\right) \xrightarrow[t\to\infty]{} 0,$$

hence $\|\hat{\lambda}_t^k - \sigma_{-i}^k\| \cdot \frac{N_t(k)}{t} \to 0$.

Calibration (partition version): the forecast is calibrated w.r.t. the (finite) forecast menu S actually used.

If both players best reply to their calibrated forecasts, standard Foster–Vohra argument $\Longrightarrow$ the empirical distribution of play is a (coarse) correlated equilibrium; with vanishing mesh grids, obtain correlated equilibrium.

Implementation Recipe (what to compute each period)

At the start of period t

Maintain $x = \bar{z}_{t-1}$ and its blocks x_k.

For each k whose representative is pure e_a, compute $x_k(a)$.

Pick k_t with representative $\boldsymbol{\sigma}_{-i}^{k_t} = e_{a^*}$ where $a^* \in \arg\max_a x_{k_t}(a)$.

Announce forecast $\boldsymbol{\sigma}_{-i,t}^{i} := \boldsymbol{\sigma}_{-i}^{k_t}$; play a myopic best reply to this forecast (for the "learning $\rightarrow$ CE" part).

Update z_t and the running average.

Remark

To move from calibration on a finite menu S to ε-calibration on $\Delta(A_{-i})$, take S to be a fine grid (mesh $\leq \varepsilon$) *together with all pure actions*. The same separation rule works.

Existence of Calibrated Forecasters

Separation step and control rule

Let $x := \bar{z}_{t-1} \in \mathbb{R}^{K|A_{-i}|}$. Target set $C := \mathbb{R}^{K|A_{-i}|}_{-}$. Projection $y := \mathbf{0} \wedge x$; normal $x - y = x^{+}$.

If we announce cell k (i.e. forecast $\boldsymbol{\sigma}^{k}_{-i}$) and Nature uses $\boldsymbol{\lambda}_{-i} \in \Delta(A_{-i})$, then

$$\mathbb{E}[z_t \mid H_{t-1}, k, \boldsymbol{\lambda}_{-i}] = e_k \otimes (\boldsymbol{\lambda}_{-i} - \boldsymbol{\sigma}^{k}_{-i}),$$

where e_k is k-th basis vector and $\otimes$ concatenates $|A_{-i}|$-block at k.

For Blackwell's approachability, need: $\mathbb{E}[z_t \mid H_{t-1}, k, \boldsymbol{\lambda}_{-i}] \cdot x^{+} = (\boldsymbol{\lambda}_{-i} - \boldsymbol{\sigma}^{k}_{-i}) \cdot x^{+}_{k} \leq \mathbf{0}$, $\forall\, \boldsymbol{\lambda}_{-i} \in \Delta(A_{-i})$, where x^{+}_{k} is the k-block of x^{+}. This holds iff $\boldsymbol{\sigma}^{k}_{-i} \in \arg\max_{q \in \Delta(A_{-i})} x^{+}_{k} \cdot q$.

Rule at time t: compute x^{+}_{k} for each k and choose

$$k_t \in \arg\max_{k} x^{+}_{k} \cdot \boldsymbol{\sigma}^{k}_{-i}, \qquad \text{then forecast } \boldsymbol{\sigma}^{i}_{-i,t} := \boldsymbol{\sigma}^{k_t}_{-i}.$$

Then for every $\boldsymbol{\lambda}_{-i}$, $(\boldsymbol{\lambda}_{-i} - \boldsymbol{\sigma}^{k_t}_{-i}) \cdot x^{+}_{k} \leq 0 \implies$ Blackwell separation at x.

Implementation remark

Include all pure actions among representatives $\{\boldsymbol{\sigma}^{k}_{-i}\}$; since $\boldsymbol{\lambda}_{-i} \mapsto x^{+}_{k} \cdot \boldsymbol{\lambda}_{-i}$ is linear, maximiser can be taken pure, ensuring argmax is available.

Correlated Equilibrium

Definition (correlated equilibrium)

$\pi \in \Delta(A)$ is a **correlated equilibrium** if for all maps $F_1 : A_1 \to A_1$ and $F_2 : A_2 \to A_2$,

$$\sum_a \pi(a)\left[u_1(a) - u_1(F_1(a_1), a_2)\right] \geq 0, \qquad \sum_a \pi(a)\left[u_2(a) - u_2(a_1, F_2(a_2))\right] \geq 0.$$

Equivalent to Aumann's "no profitable deviation conditional on signal".

Denote the set by $\mathcal{CE}$.

Calibrated Learning $\Longrightarrow$ Correlated Equilibrium

Theorem (Foster and Vohra 1997)

Suppose each player uses a calibrated forecasting scheme (on arbitrarily fine partitions) and in each t plays a myopic best reply to their forecast (fixed tie-breaking). Then every limit point of (D_t) lies in $\mathcal{CE}$.

Proof

For player **1**, define best-reply regions $M_1(a_1) := \{q \in \Delta(A_2) : a_1 \in \arg\max_{x_1} \sum_{a_2} q(a_2)u_1(x_1,$

By tie-breaking, whenever $p_t \in B^k \subseteq M_1(a_1)$, player **1** plays a_1.

Calibration on $\Pi \implies$ conditional empirical law of $a_{2,t}$ given $\{p_t \in B^k\}$ converges to q^k; hence any limit π satisfies $\pi(\cdot \mid a_1) \in M_1(a_1)$ whenever $\pi(a_1) > 0$.

Symmetrically for player **2**: $\pi(\cdot \mid a_2) \in M_2(a_2)$ if $\pi(a_2) > 0$.

These conditions are equivalent to the CE inequalities $\implies \pi \in \mathcal{CE}$.

Attainability of Any Correlated Equilibrium

Theorem (Foster and Vohra 1997)

For any $\pi^\star \in \mathcal{C}E$ there exist calibrated forecasters and myopic best replies such that $D_t \to \pi^\star$ (all limit points equal $\pi^\star$).

Construct partitions and representatives matching $\pi^\star$'s conditionals; calibrate to them.

Best replies implement the recommended supports; empirical play tracks $\pi^\star$.

From Calibration to CE: Details

Write $D_t(a_1, \cdot)$ for row-a_1. If $\liminf_t \sum_{a_2} D_t(a_1, a_2) > 0$, then along any convergent subsequence $D_{t_k} \to \pi$,

$$\frac{D_{t_k}(a_1, \cdot)}{\sum_{a_2} D_{t_k}(a_1, a_2)} \to q^{a_1} \in M_1(a_1).$$

Hence $\sum_{a_2} \pi(a_1, a_2)\left[u_1(a_1, a_2) - u_1(x_1, a_2)\right] \geq 0$ for all x_1.

Do the same for player 2; collect the inequalities to obtain the CE conditions.

Constructing Calibrated Forecasts (Procedure)

Fix partition $\mathbf{\Pi} = \{B^k\}$ with representatives (q^k).

Maintain running average $\bar{z}_{t-1} \in \mathbb{R}^{K|A_2|}$ of calibration vectors $z_t(k, a_2)$.

Forecast rule (Blackwell step):

If $\bar{z}_{t-1} \in C$ (all active components near $\mathbf{0}$), pick any k.

Else let $x^- := (\bar{z}_{t-1})^-$; choose k that minimises $\sum_{a_2} x^-(k, a_2)\, q^k(a_2)$; set $p_t \in B^k$.

Guarantees: $\bar{z}_t \to \mathbf{0}$ a.s. on all active $k \implies$ calibration on $\mathbf{\Pi}$.

Refining $\mathbf{\Pi}$ over time (mesh $\downarrow \mathbf{0}$) yields full calibration.

Interpretation and Links

Meaning of calibration: whenever you forecast p, reality looks like p on that subsequence.

Behavioural content: minimal discipline on beliefs + myopic optimality $\Longrightarrow$ CE.

Internal vs external regret: no internal regret also leads to $\mathcal{C}E$ (contrast with calibration-based route).

Design/selection: by designing calibrated grids, any target $\pi^\star \in \mathcal{C}E$ can be attained.

Uniform Calibration (Refinement Limit)

Definition (uniform ε-calibration)

A forecast sequence (p_t) is **ε-calibrated** if there exist points $q^1, \ldots, q^K$ with $\max_{a_2} \min_k |p(a_2) - q^k(a_2)| \leq \varepsilon$ for all $p \in \Delta(A_2)$ and

$$\limsup_{t \to \infty} \sum_{k=1}^{K} \frac{N^k(t)}{t} \max_{a_2} \left| r_t^k(a_2) - q^k(a_2) \right| \leq \varepsilon.$$

Existence for all $\varepsilon > 0$; letting $\varepsilon \downarrow 0$ yields full calibration.

Compatible with the approachability construction by refining the partition.

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

The first study of no-regret strategies was conducted by Hannan [1957]. The connection between no-regret strategies and the concept of approachable sets was first made by Hart and Mas-Colell [2000]. Several studies, including Foster and Vohra [1997] and Fudenberg and Levine [1999], define no-regret in a stronger form than the one presented here. Rustichini [1999], Lugosi, Mannor, and Stoltz [2007], and Lehrer and Solan [2007] studied no-regret strategies under which the decision maker does not know the true state of nature, but receives information that depends on the state of nature and the chosen action.

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
Foster Young 2006 TE, Regret testing. learning to play Nash equilibrium without knowing you have an opponent
Papers on reinforcement learning and Q-learning

Sophisticated Learning

What is players are Bayesian wrt gameplay and engage in sophisticated 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 $a^t = (a_i^t)_i \in A$.

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).

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 $\boldsymbol{\sigma}^i_{-i}$.

Induces belief $\boldsymbol{\mu}_i = \boldsymbol{\mu}^{\boldsymbol{\sigma}^i_{-i}}$ on $\boldsymbol{\Omega}$.

Player i's prior $\boldsymbol{\nu}_i$ on opponents' strategies $\boldsymbol{\sigma}_{-i}$ (Actual uncertainty).

Induces belief $\boldsymbol{\mu}_i$ on $\boldsymbol{\Omega}$ via $\tilde{\boldsymbol{\sigma}}_{-i} \mapsto \boldsymbol{\mu}^{(\boldsymbol{\sigma}_i, \tilde{\boldsymbol{\sigma}}_{-i})}$.

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

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

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.).

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 $A \subseteq Q$,

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

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.

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 $\boldsymbol{\sigma}^i_{h_t}$.

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.

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.

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.

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.

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).

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 $\mu_i(\cdot \mid h_t)$.

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 $\mu^\sigma(\cdot \mid h_t)$.

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.

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.

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

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