Blackwell Approachability

Constructing Decision-Calibrated Forecasts

Aaron Roth

University of Pennsylvania

Fall 2026

Overview

- Blackwell approachability is a general method for satisfying several long-run constraints in an online decision problem.

Overview

- Blackwell approachability is a general method for satisfying several long-run constraints in an online decision problem.
- It generalizes the minimax theorem from one scalar objective to a vector of objectives that must be controlled simultaneously.

Overview

- Blackwell approachability is a general method for satisfying several long-run constraints in an online decision problem.
- It generalizes the minimax theorem from one scalar objective to a vector of objectives that must be controlled simultaneously.
- Decision calibration fits this framework by recording all action-cell biases as coordinates of one vector.

Overview

- Blackwell approachability is a general method for satisfying several long-run constraints in an online decision problem.
- It generalizes the minimax theorem from one scalar objective to a vector of objectives that must be controlled simultaneously.
- Decision calibration fits this framework by recording all action-cell biases as coordinates of one vector.
- We will use approachability to construct approximately decision-calibrated forecasts even when outcomes are chosen adversarially over time.

Recall: Decision Calibration Uses Action Cells

On round t, the forecaster reports $p_t \in \Delta(\mathcal{Y})$, the user takes the utility-maximizing action

$$a_t = \mathrm{BR}_u(p_t),$$

and outcome y_t occurs. Write e_{y_t} for the point mass on that outcome.

Recall: Decision Calibration Uses Action Cells

On round t, the forecaster reports $p_t \in \Delta(\mathcal{Y})$, the user takes the utility-maximizing action

$$a_t = \mathrm{BR}_u(p_t),$$

and outcome y_t occurs. Write e_{y_t} for the point mass on that outcome.

The *action-cell bias* for action a is the cumulative forecast error on the rounds when the forecast induced that action:

$$B_a := \sum_{t : a_t = a} (e_{y_t} - p_t).$$

Recall: Decision Calibration Uses Action Cells

On round t, the forecaster reports $p_t \in \Delta(\mathcal{Y})$, the user takes the utility-maximizing action

$$a_t = \mathrm{BR}_u(p_t),$$

and outcome y_t occurs. Write e_{y_t} for the point mass on that outcome.

The *action-cell bias* for action a is the cumulative forecast error on the rounds when the forecast induced that action:

$$B_a := \sum_{t:a_t=a} (e_{y_t} - p_t).$$

The sequence is *decision calibrated* for this user when

$$B_a = 0 \qquad \text{for every } a \in \mathcal{A}.$$

The exact condition is norm-independent. Lecture 7 measured approximation by $\sum_a \|B_a\|_2$ and used it to bound swap regret.

The Forecast Determines Which Bias Is Updated

Decision calibration asks the forecaster to keep every action-cell bias small. But the forecast itself selects the cell that receives the next residual:

$$p_t \longrightarrow a_t = \mathrm{BR}_u(p_t) \longrightarrow \text{update } B_{a_t}.$$

The Forecast Determines Which Bias Is Updated

Decision calibration asks the forecaster to keep every action-cell bias small. But the forecast itself selects the cell that receives the next residual:

$$p_t \longrightarrow a_t = \mathrm{BR}_u(p_t) \longrightarrow \text{update } B_{a_t}.$$

The forecaster must choose p_t before seeing y_t, yet p_t determines which constraint the error $e_{y_t} - p_t$ will affect.

The Forecast Determines Which Bias Is Updated

Decision calibration asks the forecaster to keep every action-cell bias small. But the forecast itself selects the cell that receives the next residual:

$$p_t \longrightarrow a_t = \mathrm{BR}_u(p_t) \longrightarrow \text{update } B_{a_t}.$$

The forecaster must choose p_t before seeing y_t, yet p_t determines which constraint the error $e_{y_t} - p_t$ will affect.
How can the forecaster keep every action-cell bias small against an adversarial sequence of outcomes?

The Setting

Let $\mathcal{Y} = \{1, \ldots, d\}$ and let $\mathcal{P}_\eta \subseteq \Delta(\mathcal{Y})$ be a finite forecast grid. The construction will randomize over this grid. On round t:

1. the forecaster announces a distribution $\rho_t \in \Delta(\mathcal{P}_\eta)$;

The Setting

Let $\mathcal{Y} = \{1, \ldots, d\}$ and let $\mathcal{P}_\eta \subseteq \Delta(\mathcal{Y})$ be a finite forecast grid. The construction will randomize over this grid. On round t:

1. the forecaster announces a distribution $\rho_t \in \Delta(\mathcal{P}_\eta)$;
2. after seeing ρ_t, the environment commits to an outcome $y_t \in \mathcal{Y}$;

The Setting

Let $\mathcal{Y} = \{1, \ldots, d\}$ and let $\mathcal{P}_\eta \subseteq \Delta(\mathcal{Y})$ be a finite forecast grid. The construction will randomize over this grid. On round t:

1. the forecaster announces a distribution $\rho_t \in \Delta(\mathcal{P}_\eta)$;
2. after seeing ρ_t, the environment commits to an outcome $y_t \in \mathcal{Y}$;
3. a forecast $p_t \sim \rho_t$ is sampled and revealed to the user, who chooses $a_t = \mathrm{BR}_u(p_t)$;

The Setting

Let $\mathcal{Y} = \{1, \ldots, d\}$ and let $\mathcal{P}_\eta \subseteq \Delta(\mathcal{Y})$ be a finite forecast grid. The construction will randomize over this grid. On round t:

1. the forecaster announces a distribution $\rho_t \in \Delta(\mathcal{P}_\eta)$;
2. after seeing ρ_t, the environment commits to an outcome $y_t \in \mathcal{Y}$;
3. a forecast $p_t \sim \rho_t$ is sampled and revealed to the user, who chooses $a_t = \mathrm{BR}_u(p_t)$;
4. the outcome is revealed and the residual $e_{y_t} - p_t$ is recorded in the a_t cell.

The Setting

Let $\mathcal{Y} = \{1, \ldots, d\}$ and let $\mathcal{P}_\eta \subseteq \Delta(\mathcal{Y})$ be a finite forecast grid. The construction will randomize over this grid. On round t:

1. the forecaster announces a distribution $\rho_t \in \Delta(\mathcal{P}_\eta)$;
2. after seeing ρ_t, the environment commits to an outcome $y_t \in \mathcal{Y}$;
3. a forecast $p_t \sim \rho_t$ is sampled and revealed to the user, who chooses $a_t = \mathrm{BR}_u(p_t)$;
4. the outcome is revealed and the residual $e_{y_t} - p_t$ is recorded in the a_t cell.

The environment may depend on the entire past and on ρ_t, but not on the current random draw.

Action-Cell Biases Form One Joint Vector

For a candidate forecast p and outcome y, define

$$g(p, y) := \big(\mathbf{1}\{\mathrm{BR}_u(p) = a\}(e_y - p)\big)_{a \in \mathcal{A}} \in \mathbb{R}^{|\mathcal{A}|d}.$$

Action-Cell Biases Form One Joint Vector

For a candidate forecast p and outcome y, define

$$g(p, y) := \big(\mathbf{1}\{\mathrm{BR}_u(p) = a\}(e_y - p)\big)_{a \in \mathcal{A}} \in \mathbb{R}^{|\mathcal{A}|d}.$$

Exactly one action block is nonzero. On the realized sequence,

$$\bar{g}_T := \frac{1}{T} \sum_{t=1}^{T} g(p_t, y_t) = \frac{1}{T}(B_a)_{a \in \mathcal{A}}.$$

Action-Cell Biases Form One Joint Vector

For a candidate forecast p and outcome y, define

$$g(p, y) := \big(\mathbf{1}\{\mathrm{BR}_u(p) = a\}(e_y - p)\big)_{a \in \mathcal{A}} \in \mathbb{R}^{|\mathcal{A}|d}.$$

Exactly one action block is nonzero. On the realized sequence,

$$\bar{g}_T := \frac{1}{T}\sum_{t=1}^{T} g(p_t, y_t) = \frac{1}{T}(B_a)_{a \in \mathcal{A}}.$$

Thus constructing a decision-calibrated forecast sequence is a vector-control problem: drive $\bar{g}_T$ toward a small neighborhood of zero.

Why Controlling the Joint Vector Is Enough

Lecture 7 proved the transfer bound

$$\mathrm{SwapReg}_T \le \frac{D_{u,2}}{T} \sum_{a \in \mathcal{A}} \|B_a\|_2, \qquad D_{u,2} := \max_{a,b \in \mathcal{A}} \|u_a - u_b\|_2.$$

Why Controlling the Joint Vector Is Enough

Lecture 7 proved the transfer bound

$$\mathrm{SwapReg}_T \le \frac{D_{u,2}}{T} \sum_{a \in \mathcal{A}} \|B_a\|_2, \qquad D_{u,2} := \max_{a,b \in \mathcal{A}} \|u_a - u_b\|_2.$$

Since $\bar{g}_T = T^{-1}(B_a)_{a \in \mathcal{A}}$, Cauchy–Schwarz across action blocks gives

$$\frac{1}{T} \sum_{a \in \mathcal{A}} \|B_a\|_2 \le \sqrt{|\mathcal{A}|}\, \|\bar{g}_T\|_2.$$

Therefore

$$\boxed{\mathrm{SwapReg}_T \le D_{u,2} \sqrt{|\mathcal{A}|}\, \|\bar{g}_T\|_2.}$$

Approaching zero controls every action remapping simultaneously.

Predict within a Finite Grid

Assume $\mathcal{P}_\eta$ is an η-net in Euclidean norm: for every $q \in \Delta(\mathcal{Y})$, some $p \in \mathcal{P}_\eta$ satisfies

$$\|p - q\|_2 \leq \eta.$$

Predict within a Finite Grid

Assume $\mathcal{P}_\eta$ is an η-net in Euclidean norm: for every $q \in \Delta(\mathcal{Y})$, some $p \in \mathcal{P}_\eta$ satisfies

$$\|p - q\|_2 \leq \eta.$$

We will drive the average action-cell bias vector $\bar{g}_T$ toward

$$C_\eta := \{z : \|z\|_2 \leq \eta\}.$$

The radius is the price of replacing an arbitrary outcome distribution by a nearby grid forecast.

Predict within a Finite Grid

Assume $\mathcal{P}_\eta$ is an η-net in Euclidean norm: for every $q \in \Delta(\mathcal{Y})$, some $p \in \mathcal{P}_\eta$ satisfies

$$\|p - q\|_2 \leq \eta.$$

We will drive the average action-cell bias vector $\bar{g}_T$ toward

$$C_\eta := \{z : \|z\|_2 \leq \eta\}.$$

The radius is the price of replacing an arbitrary outcome distribution by a nearby grid forecast.
For the moment, we abstract away from forecasting and ask how to approach a general nonempty closed convex target set C.

Abstract to a Finite Vector Game

For the general theorem, let the finite sets

$$\mathcal{X} = \{\text{learner actions}\}, \qquad \mathcal{Y} = \{\text{environment actions}\},$$

and let

$$g : \mathcal{X} \times \mathcal{Y} \longrightarrow \mathbb{R}^m$$

assign a vector—called the *vector payoff*—to every pair of actions. Its coordinates record the several quantities we want to control simultaneously.

Abstract to a Finite Vector Game

For the general theorem, let the finite sets

$$\mathcal{X} = \{\text{learner actions}\}, \qquad \mathcal{Y} = \{\text{environment actions}\},$$

and let

$$g : \mathcal{X} \times \mathcal{Y} \longrightarrow \mathbb{R}^m$$

assign a vector—called the *vector payoff*—to every pair of actions. Its coordinates record the several quantities we want to control simultaneously.

On round t, the learner announces $\rho_t \in \Delta(\mathcal{X})$; the environment commits to y_t after seeing ρ_t but before $x_t \sim \rho_t$ is drawn; then the vector $g(x_t, y_t)$ is recorded.

Abstract to a Finite Vector Game

For the general theorem, let the finite sets

$$\mathcal{X} = \{\text{learner actions}\}, \qquad \mathcal{Y} = \{\text{environment actions}\},$$

and let

$$g : \mathcal{X} \times \mathcal{Y} \longrightarrow \mathbb{R}^m$$

assign a vector—called the *vector payoff*—to every pair of actions. Its coordinates record the several quantities we want to control simultaneously.

On round t, the learner announces $\rho_t \in \Delta(\mathcal{X})$; the environment commits to y_t after seeing ρ_t but before $x_t \sim \rho_t$ is drawn; then the vector $g(x_t, y_t)$ is recorded.

For decision calibration, $\mathcal{X} = \mathcal{P}_\eta$, the learner's action x is a forecast p, and $g(x, y)$ is the action-cell residual vector already defined. For now we'll consider the general case.

Approachability Controls an Average Vector

A vector's Euclidean distance from a nonempty closed set C is

$$\operatorname{dist}(z, C) := \min_{c \in C} \|z - c\|_2.$$

A randomized strategy *approaches* C if, against every permitted environment strategy, it makes

$$\mathbb{E}\left[\operatorname{dist}(\bar{g}_T, C)^2\right] \longrightarrow 0, \qquad \bar{g}_T := \frac{1}{T} \sum_{t=1}^{T} g(x_t, y_t).$$

The expectation is over the learner's random draws.

Approachability Controls an Average Vector

A vector's Euclidean distance from a nonempty closed set C is

$$\mathrm{dist}(z, C) := \min_{c \in C} \|z - c\|_2.$$

A randomized strategy *approaches* C if, against every permitted environment strategy, it makes

$$\mathbb{E}\big[\mathrm{dist}(\bar{g}_T, C)^2\big] \longrightarrow 0, \qquad \bar{g}_T := \frac{1}{T} \sum_{t=1}^{T} g(x_t, y_t).$$

The expectation is over the learner's random draws.
Unlike scalar regret, there is no single ordering of vector payoffs.
The first task is to identify a scalar objective at each round.

Response Satisfiability Is a One-Round Condition

For mixed actions $\rho \in \Delta(\mathcal{X})$ and $Q \in \Delta(\mathcal{Y})$, write

$$g(\rho, Q) := \mathbb{E}_{X\sim\rho,\, Y\sim Q}[g(X, Y)].$$

Response Satisfiability Is a One-Round Condition

For mixed actions $\rho \in \Delta(\mathcal{X})$ and $Q \in \Delta(\mathcal{Y})$, write

$$g(\rho, Q) := \mathbb{E}_{X \sim \rho,\, Y \sim Q}[g(X, Y)].$$

A target set C is *response satisfiable* if

$$\boxed{\forall Q \in \Delta(\mathcal{Y}), \quad \exists \rho \in \Delta(\mathcal{X}) : \quad g(\rho, Q) \in C.}$$

If the learner knew the environment's mixed action Q, it could respond with an expected vector already in the target.

Response Satisfiability Is a One-Round Condition

For mixed actions $\rho \in \Delta(\mathcal{X})$ and $Q \in \Delta(\mathcal{Y})$, write

$$g(\rho, Q) := \mathbb{E}_{X\sim\rho,\, Y\sim Q}[g(X, Y)].$$

A target set C is *response satisfiable* if

$$\boxed{\forall Q \in \Delta(\mathcal{Y}), \quad \exists \rho \in \Delta(\mathcal{X}) : \quad g(\rho, Q) \in C.}$$

If the learner knew the environment's mixed action Q, it could respond with an expected vector already in the target.
In the repeated game, however, the learner moves without being given Q. Blackwell's argument will use minimax to turn this wrong-order statement into an online approachability strategy.

Minimax Is One-Dimensional Approachability

Suppose $m = 1$ and the target is the halfline $C = (-\infty, v]$.
Response satisfiability says

$$\max_{Q \in \Delta(\mathcal{Y})} \min_{\rho \in \Delta(\mathcal{X})} \mathbb{E}_{X \sim \rho,\, Y \sim Q}[g(X, Y)] \leq v.$$

Lecture 5's minimax theorem says this is equivalent to

$$\min_{\rho \in \Delta(\mathcal{X})} \max_{y \in \mathcal{Y}} \mathbb{E}_{X \sim \rho}[g(X, y)] \leq v.$$

Minimax Is One-Dimensional Approachability

Suppose $m = 1$ and the target is the halfline $C = (-\infty, v]$. Response satisfiability says

$$\max_{Q \in \Delta(\mathcal{Y})} \min_{\rho \in \Delta(\mathcal{X})} \mathbb{E}_{X \sim \rho,\, Y \sim Q}[g(X, Y)] \leq v.$$

Lecture 5's minimax theorem says this is equivalent to

$$\min_{\rho \in \Delta(\mathcal{X})} \max_{y \in \mathcal{Y}} \mathbb{E}_{X \sim \rho}[g(X, y)] \leq v.$$

Thus one mixed action keeps the expected scalar payoff in C against every environment action. This one round guarantee does not extend to the multi-dimensional setting; but it almost does: we can *approach* response satisfiable convex bodies over time.

Minimax Is One-Dimensional Approachability

Suppose $m = 1$ and the target is the halfline $C = (-\infty, v]$. Response satisfiability says

$$\max_{Q \in \Delta(\mathcal{Y})} \min_{\rho \in \Delta(\mathcal{X})} \mathbb{E}_{X \sim \rho,\, Y \sim Q}[g(X, Y)] \leq v.$$

Lecture 5's minimax theorem says this is equivalent to

$$\min_{\rho \in \Delta(\mathcal{X})} \max_{y \in \mathcal{Y}} \mathbb{E}_{X \sim \rho}[g(X, y)] \leq v.$$

Thus one mixed action keeps the expected scalar payoff in C against every environment action. This one round guarantee does not extend to the multi-dimensional setting; but it almost does: we can *approach* response satisfiable convex bodies over time. Projection will select one relevant scalar direction on each round, and minimax will be applied in that direction.

Blackwell's Guarantee

Theorem (Finite Blackwell approachability)

Let $\mathcal{X}$ and $\mathcal{Y}$ be finite, and let $C \subseteq \mathbb{R}^m$ be nonempty, closed, and convex. Fix any $c^\circ \in C$, and define

$$G := \max_{x \in \mathcal{X},\, y \in \mathcal{Y}} \|g(x, y) - c^\circ\|_2.$$

If C is response satisfiable, then there is a randomized strategy such that

$$\mathbb{E}\left[\operatorname{dist}\left(\frac{1}{T}\sum_{t=1}^{T} g(x_t, y_t), C\right)^2\right] \le \frac{4G^2}{T}.$$

Blackwell's Guarantee

Theorem (Finite Blackwell approachability)

Let $\mathcal{X}$ and $\mathcal{Y}$ be finite, and let $C \subseteq \mathbb{R}^m$ be nonempty, closed, and convex. Fix any $c^\circ \in C$, and define

$$G := \max_{x \in \mathcal{X},\, y \in \mathcal{Y}} \|g(x, y) - c^\circ\|_2.$$

If C is response satisfiable, then there is a randomized strategy such that

$$\mathbb{E}\left[\operatorname{dist}\left(\frac{1}{T}\sum_{t=1}^{T} g(x_t, y_t), C\right)^2\right] \le \frac{4G^2}{T}.$$

The algorithm and analysis have three parts: select a scalar direction at each round, find a response in that direction, and show that these responses drive the average toward C.

Translate the Anchor to the Origin

Define translated payoffs and a translated target by

$$\widetilde{g}(x, y) := g(x, y) - c^\circ, \qquad \widetilde{C} := C - c^\circ.$$

Then

$$0 \in \widetilde{C}, \qquad \|\widetilde{g}(x, y)\|_2 \leq G.$$

Translate the Anchor to the Origin

Define translated payoffs and a translated target by

$$\widetilde{g}(x, y) := g(x, y) - c^{\circ}, \qquad \widetilde{C} := C - c^{\circ}.$$

Then

$$0 \in \widetilde{C}, \qquad \|\widetilde{g}(x, y)\|_2 \leq G.$$

Response satisfiability is preserved because

$$g(\rho, Q) \in C \quad \Longleftrightarrow \quad \widetilde{g}(\rho, Q) = g(\rho, Q) - c^{\circ} \in \widetilde{C}.$$

The translated running average is $\bar{\widetilde{g}}_T = \bar{g}_T - c^{\circ}$, and its distance to the translated target is unchanged:

$$\operatorname{dist}(\bar{\widetilde{g}}_T, \widetilde{C}) = \operatorname{dist}(\bar{g}_T, C).$$

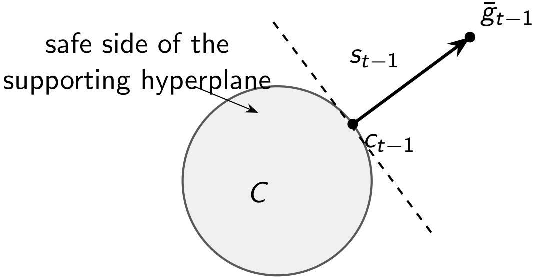
Thus it suffices to prove the theorem in translated coordinates. We now drop the tildes and assume $0 \in C$ and $\|g(x, y)\|_2 \leq G$.

Projection Identifies a Current Direction of Failure

The first step is to turn the current vector error into one scalar direction. Let

$$c_{t-1} := \Pi_C(\bar{g}_{t-1}), \qquad s_{t-1} := \bar{g}_{t-1} - c_{t-1}.$$

Thus c_{t-1} is the point of C closest to the current average, and s_{t-1} points from the target toward the current average.



Projection Creates a Supporting Halfspace

Fix any $z \in C$. Convexity implies that the entire segment

$$z_\alpha := c_{t-1} + \alpha(z - c_{t-1}), \qquad 0 \le \alpha \le 1,$$

lies in C. Since c_{t-1} is closest to $\bar{g}_{t-1}$, the function

$$f(\alpha) := \|\bar{g}_{t-1} - z_\alpha\|_2^2 = \|\bar{g}_{t-1} - c_{t-1} - \alpha(z - c_{t-1})\|_2^2$$

is minimized at $\alpha = 0$.

Projection Creates a Supporting Halfspace

Fix any $z \in C$. Convexity implies that the entire segment

$$z_\alpha := c_{t-1} + \alpha(z - c_{t-1}), \qquad 0 \le \alpha \le 1,$$

lies in C. Since c_{t-1} is closest to $\bar{g}_{t-1}$, the function

$$f(\alpha) := \|\bar{g}_{t-1} - z_\alpha\|_2^2 = \|\bar{g}_{t-1} - c_{t-1} - \alpha(z - c_{t-1})\|_2^2$$

is minimized at $\alpha = 0$.

Its right derivative at zero must therefore be nonnegative:

$$f'(0) = -2\,\langle \bar{g}_{t-1} - c_{t-1}, z - c_{t-1} \rangle \ge 0.$$

If this derivative were negative, a sufficiently small positive α would make z_α closer than c_{t-1}, contradicting the definition of the projection. Hence, using $s_{t-1} = \bar{g}_{t-1} - c_{t-1}$,

$$\boxed{\langle s_{t-1}, z - c_{t-1} \rangle \le 0 \quad \text{for every } z \in C.}$$

Every displacement $z - c_{t-1}$ from c_{t-1} into C has nonpositive dot product with the current outward direction s_{t-1}.

Projection Replaces the Target by One Halfspace

Define the halfspace

$$H_{t-1} := \{v : \langle s_{t-1}, v - c_{t-1} \rangle \leq 0\}.$$

The projection inequality says that $C \subseteq H_{t-1}$.

Projection Replaces the Target by One Halfspace

Define the halfspace

$$H_{t-1} := \{v : \langle s_{t-1}, v - c_{t-1} \rangle \le 0\}.$$

The projection inequality says that $C \subseteq H_{t-1}$.
A vector v lies in this halfspace exactly when

$$\langle s_{t-1}, v - c_{t-1} \rangle \le 0,$$

so its displacement from c_{t-1} has nonpositive dot product with the current outward direction s_{t-1}.

Projection Replaces the Target by One Halfspace

Define the halfspace

$$H_{t-1} := \{v : \langle s_{t-1}, v - c_{t-1} \rangle \leq 0\}.$$

The projection inequality says that $C \subseteq H_{t-1}$.
A vector v lies in this halfspace exactly when

$$\langle s_{t-1}, v - c_{t-1} \rangle \leq 0,$$

so its displacement from c_{t-1} has nonpositive dot product with the current outward direction s_{t-1}.
Our candidate per-round condition will ask only that the next expected vector lie in H_{t-1}, thereby controlling this one direction.
The potential argument will prove that repeatedly satisfying this weaker-looking condition approaches the full target. If $s_{t-1} = 0$, the current average is already in C and there is no outward direction to control.

The Next Expected Vector Should Stay on the Safe Side

For a fixed environment action y, randomizing the learner's action according to ρ_t gives the expected round vector

$$v_y := \mathbb{E}_{x \sim \rho_t}[g(x, y)].$$

We want $v_y \in H_{t-1}$ for every possible y. By the definition of the halfspace, this is equivalent to

The Next Expected Vector Should Stay on the Safe Side

For a fixed environment action y, randomizing the learner's action according to ρ_t gives the expected round vector

$$v_y := \mathbb{E}_{x\sim\rho_t}[g(x,y)].$$

We want $v_y \in H_{t-1}$ for every possible y. By the definition of the halfspace, this is equivalent to

$$\mathbb{E}_{x\sim\rho_t}\left[\langle s_{t-1}, g(x,y) - c_{t-1}\rangle\right] \le 0.$$

This is a candidate scalar condition for the current round. Two facts still need proof:

1. one ρ_t can satisfy it for every y, despite being chosen first;
2. enforcing it on every round makes the average vector approach C.

The Next Expected Vector Should Stay on the Safe Side

For a fixed environment action y, randomizing the learner's action according to ρ_t gives the expected round vector

$$v_y := \mathbb{E}_{x\sim\rho_t}[g(x, y)].$$

We want $v_y \in H_{t-1}$ for every possible y. By the definition of the halfspace, this is equivalent to

$$\mathbb{E}_{x\sim\rho_t}\left[\langle s_{t-1}, g(x, y) - c_{t-1}\rangle\right] \le 0.$$

This is a candidate scalar condition for the current round. Two facts still need proof:

1. one ρ_t can satisfy it for every y, despite being chosen first;
2. enforcing it on every round makes the average vector approach C.

The next minimax argument proves the first fact. The later squared-distance argument proves the second.

Response Satisfiability Gives a Reply to Every Q

Define the scalar payoff

$$A_t(x, y) := \langle s_{t-1}, g(x, y) - c_{t-1} \rangle \,.$$

Response Satisfiability Gives a Reply to Every Q

Define the scalar payoff

$$A_t(x, y) := \langle s_{t-1}, g(x, y) - c_{t-1} \rangle \, .$$

Fix any hypothetical $Q \in \Delta(\mathcal{Y})$. Response satisfiability gives a mixed response ρ_Q for which

$$z_Q := g(\rho_Q, Q) \in C.$$

The projection inequality therefore gives

$$\mathbb{E}_{X\sim\rho_Q,\, Y\sim Q}[A_t(X, Y)] = \langle s_{t-1}, z_Q - c_{t-1} \rangle \leq 0.$$

Thus response satisfiability has supplied

$$\forall Q \in \Delta(\mathcal{Y}), \quad \exists \rho_Q \in \Delta(\mathcal{X}) : \quad \mathbb{E}_{X\sim\rho_Q,\, Y\sim Q}[A_t(X, Y)] \leq 0.$$

Response Satisfiability Gives a Reply to Every Q

Define the scalar payoff

$$A_t(x, y) := \langle s_{t-1}, g(x, y) - c_{t-1} \rangle .$$

Fix any hypothetical $Q \in \Delta(\mathcal{Y})$. Response satisfiability gives a mixed response ρ_Q for which

$$z_Q := g(\rho_Q, Q) \in C.$$

The projection inequality therefore gives

$$\mathbb{E}_{X \sim \rho_Q, \, Y \sim Q}[A_t(X, Y)] = \langle s_{t-1}, z_Q - c_{t-1} \rangle \leq 0.$$

Thus response satisfiability has supplied

$$\forall Q \in \Delta(\mathcal{Y}), \quad \exists \rho_Q \in \Delta(\mathcal{X}) : \quad \mathbb{E}_{X \sim \rho_Q, \, Y \sim Q}[A_t(X, Y)] \leq 0.$$

The quantifiers still point the wrong way: the learner needs one ρ_t chosen before y that works for every environment action.

Minimax Reverses the Quantifiers

View A_t as a finite zero-sum game. The learner minimizes by choosing $x \in \mathcal{X}$; the environment maximizes by choosing $y \in \mathcal{Y}$. Applying the minimax theorem to the current projection direction gives

$$\begin{aligned}\min_{\rho \in \Delta(\mathcal{X})} \max_{y} \mathbb{E}_{X \sim \rho}[A_t(X, y)] &= \max_{Q \in \Delta(\mathcal{Y})} \min_{\rho \in \Delta(\mathcal{X})} \mathbb{E}_{\substack{X \sim \rho \\ Y \sim Q}}[A_t(X, Y)] \\ &\leq 0.\end{aligned}$$

The inequality is exactly the response-satisfiability conclusion from the previous slide.

Minimax Reverses the Quantifiers

View A_t as a finite zero-sum game. The learner minimizes by choosing $x \in \mathcal{X}$; the environment maximizes by choosing $y \in \mathcal{Y}$. Applying the minimax theorem to the current projection direction gives

$$\min_{\rho \in \Delta(\mathcal{X})} \max_{y} \mathbb{E}_{X\sim\rho}[A_t(X, y)] = \max_{Q \in \Delta(\mathcal{Y})} \min_{\rho \in \Delta(\mathcal{X})} \mathbb{E}_{\substack{X\sim\rho \\ Y\sim Q}}[A_t(X, Y)] \\ \leq 0.$$

The inequality is exactly the response-satisfiability conclusion from the previous slide.

Therefore there is a distribution ρ_t such that

$$\mathbb{E}_{X\sim\rho_t}\left[\langle s_{t-1}, g(X, y) - c_{t-1}\rangle\right] \leq 0 \qquad \text{for every } y.$$

This is the *halfspace response* needed by the projection argument.

Blackwell's Approachability Algorithm

On round t:

1. Compute the average payoff $\bar{g}_{t-1}$, starting from $\bar{g}_0 = 0$.

Blackwell's Approachability Algorithm

On round t:

1. Compute the average payoff $\bar{g}_{t-1}$, starting from $\bar{g}_0 = 0$.
2. Project it onto the target: $c_{t-1} = \Pi_C(\bar{g}_{t-1})$.

Blackwell's Approachability Algorithm

On round t:

1. Compute the average payoff $\bar{g}_{t-1}$, starting from $\bar{g}_0 = 0$.
2. Project it onto the target: $c_{t-1} = \Pi_C(\bar{g}_{t-1})$.
3. Use $s_{t-1} = \bar{g}_{t-1} - c_{t-1}$ to form the scalar game $A_t(x, y)$.

Blackwell's Approachability Algorithm

On round t:

1. Compute the average payoff $\bar{g}_{t-1}$, starting from $\bar{g}_0 = 0$.
2. Project it onto the target: $c_{t-1} = \Pi_C(\bar{g}_{t-1})$.
3. Use $s_{t-1} = \bar{g}_{t-1} - c_{t-1}$ to form the scalar game $A_t(x, y)$.
4. Solve that game for a distribution ρ_t whose expected scalar payoff is nonpositive against every environment action.

Blackwell's Approachability Algorithm

On round t:

1. Compute the average payoff $\bar{g}_{t-1}$, starting from $\bar{g}_0 = 0$.
2. Project it onto the target: $c_{t-1} = \Pi_C(\bar{g}_{t-1})$.
3. Use $s_{t-1} = \bar{g}_{t-1} - c_{t-1}$ to form the scalar game $A_t(x, y)$.
4. Solve that game for a distribution ρ_t whose expected scalar payoff is nonpositive against every environment action.
5. Announce ρ_t, let the environment commit to y_t, sample $x_t \sim \rho_t$, and update the average with $g(x_t, y_t)$.

Blackwell's Approachability Algorithm

On round t:

1. Compute the average payoff $\bar{g}_{t-1}$, starting from $\bar{g}_0 = 0$.
2. Project it onto the target: $c_{t-1} = \Pi_C(\bar{g}_{t-1})$.
3. Use $s_{t-1} = \bar{g}_{t-1} - c_{t-1}$ to form the scalar game $A_t(x, y)$.
4. Solve that game for a distribution ρ_t whose expected scalar payoff is nonpositive against every environment action.
5. Announce ρ_t, let the environment commit to y_t, sample $x_t \sim \rho_t$, and update the average with $g(x_t, y_t)$.

The vector approachability problem is handled one scalar minimax problem at a time.

Track the Running Average's Distance from the Target

The last step is to show that the halfspace responses approach C. Let

$$\bar{g}_t := \frac{1}{t}\sum_{r=1}^{t} g(x_r, y_r), \qquad \delta_t := \mathrm{dist}(\bar{g}_t, C), \qquad \bar{g}_0 := 0.$$

Our goal is to prove

$$\mathbb{E}[\delta_T^2] \le \frac{4G^2}{T}.$$

Rewrite the Average Around the Previous Projection

The definition of the running average gives

$$\bar{g}_t = \left(1 - \frac{1}{t}\right) \bar{g}_{t-1} + \frac{1}{t} g(x_t, y_t).$$

Because the two coefficients sum to one, subtracting c_{t-1} gives

$$\begin{aligned} \bar{g}_t - c_{t-1} &= \left(1 - \frac{1}{t}\right) (\bar{g}_{t-1} - c_{t-1}) + \frac{1}{t} \big(g(x_t, y_t) - c_{t-1}\big) \\ &= \left(1 - \frac{1}{t}\right) s_{t-1} + \frac{1}{t} \big(g(x_t, y_t) - c_{t-1}\big), \end{aligned}$$

where the final equality uses $s_{t-1} := \bar{g}_{t-1} - c_{t-1}$.

Rewrite the Average Around the Previous Projection

The definition of the running average gives

$$\bar{g}_t = \left(1 - \frac{1}{t}\right) \bar{g}_{t-1} + \frac{1}{t} g(x_t, y_t).$$

Because the two coefficients sum to one, subtracting c_{t-1} gives

$$\begin{aligned}
\bar{g}_t - c_{t-1} &= \left(1 - \frac{1}{t}\right) (\bar{g}_{t-1} - c_{t-1}) + \frac{1}{t} \big(g(x_t, y_t) - c_{t-1}\big) \\
&= \left(1 - \frac{1}{t}\right) s_{t-1} + \frac{1}{t} \big(g(x_t, y_t) - c_{t-1}\big),
\end{aligned}$$

where the final equality uses $s_{t-1} := \bar{g}_{t-1} - c_{t-1}$.
This is the point at which the halfspace response enters the potential proof.

Expanding the Square Isolates the Halfspace Term

Because $c_{t-1} \in C$, the definition of distance and the preceding average update give

$$\begin{aligned}
\delta_t^2 &= \min_{c \in C} \|\bar{g}_t - c\|_2^2 \\
&\leq \|\bar{g}_t - c_{t-1}\|_2^2 \\
&= \left\| \left(1 - \frac{1}{t}\right) s_{t-1} + \frac{1}{t}\big(g(x_t, y_t) - c_{t-1}\big) \right\|_2^2 \\
&= \left(1 - \frac{1}{t}\right)^2 \|s_{t-1}\|_2^2 \\
&\quad + \frac{2}{t}\left(1 - \frac{1}{t}\right) \langle s_{t-1}, g(x_t, y_t) - c_{t-1} \rangle \\
&\quad + \frac{1}{t^2} \|g(x_t, y_t) - c_{t-1}\|_2^2.
\end{aligned}$$

Expanding the Square Isolates the Halfspace Term

Because $c_{t-1} \in C$, the definition of distance and the preceding average update give

$$\begin{aligned}
\delta_t^2 &= \min_{c \in C} \|\bar{g}_t - c\|_2^2 \\
&\leq \|\bar{g}_t - c_{t-1}\|_2^2 \\
&= \left\| \left(1 - \frac{1}{t}\right) s_{t-1} + \frac{1}{t}\big(g(x_t, y_t) - c_{t-1}\big) \right\|_2^2 \\
&= \left(1 - \frac{1}{t}\right)^2 \|s_{t-1}\|_2^2 \\
&\quad + \frac{2}{t}\left(1 - \frac{1}{t}\right) \langle s_{t-1}, g(x_t, y_t) - c_{t-1} \rangle \\
&\quad + \frac{1}{t^2} \|g(x_t, y_t) - c_{t-1}\|_2^2.
\end{aligned}$$

The three terms are, respectively, the old squared distance, movement in the current outward direction, and a bounded quadratic remainder.

The Halfspace Response Removes the Cross Term

Condition on the past, on ρ_t, and on the outcome y_t to which the environment has already committed. The only remaining randomness is $x_t \sim \rho_t$. Write $\mathbb{E}_t$ for expectation conditional on this information.

The Halfspace Response Removes the Cross Term

Condition on the past, on ρ_t, and on the outcome y_t to which the environment has already committed. The only remaining randomness is $x_t \sim \rho_t$. Write $\mathbb{E}_t$ for expectation conditional on this information.
The minimax response gives

$$\mathbb{E}_t\big[\langle s_{t-1}, g(x_t, y_t) - c_{t-1}\rangle\big] \leq 0.$$

The Halfspace Response Removes the Cross Term

Condition on the past, on ρ_t, and on the outcome y_t to which the environment has already committed. The only remaining randomness is $x_t \sim \rho_t$. Write $\mathbb{E}_t$ for expectation conditional on this information.
The minimax response gives

$$\mathbb{E}_t \left[\langle s_{t-1}, g(x_t, y_t) - c_{t-1} \rangle \right] \leq 0.$$

Taking conditional expectations in the expanded square therefore removes its middle term:

$$\mathbb{E}_t[\delta_t^2] \leq \left(1 - \frac{1}{t}\right)^2 \delta_{t-1}^2 + \frac{1}{t^2} \mathbb{E}_t \left[\| g(x_t, y_t) - c_{t-1} \|_2^2 \right].$$

The Origin Bounds the Quadratic Remainder

In our translated coordinates, $0 \in C$ and $\|g(x,y)\|_2 \le G$. Because $\bar{g}_{t-1}$ is an average of payoff vectors,

$$\|\bar{g}_{t-1}\|_2 \le G.$$

Put $z = 0$ in the projection inequality:

$$\begin{aligned} 0 &\ge \langle \bar{g}_{t-1} - c_{t-1}, -c_{t-1} \rangle \\ &= \|c_{t-1}\|_2^2 - \langle \bar{g}_{t-1}, c_{t-1} \rangle. \end{aligned}$$

Therefore, by Cauchy–Schwarz,

$$\|c_{t-1}\|_2^2 \le \langle \bar{g}_{t-1}, c_{t-1} \rangle \le G\|c_{t-1}\|_2.$$

If $c_{t-1} = 0$, then $\|c_{t-1}\|_2 \le G$ is immediate; otherwise, divide by $\|c_{t-1}\|_2$. The triangle inequality now gives

$$\|g(x_t, y_t) - c_{t-1}\|_2 \le G + G = 2G,$$

which bounds the quadratic remainder in the recursion.

The Recursion Telescopes

Taking total expectations in the preceding distance bound and using the $2G$ remainder bound gives

$$\mathbb{E}[\delta_t^2] \leq \left(1 - \frac{1}{t}\right)^2 \mathbb{E}[\delta_{t-1}^2] + \frac{4G^2}{t^2},$$

$$\text{and hence} \qquad t^2\mathbb{E}[\delta_t^2] \leq (t-1)^2\mathbb{E}[\delta_{t-1}^2] + 4G^2.$$

The Recursion Telescopes

Taking total expectations in the preceding distance bound and using the $2G$ remainder bound gives

$$\mathbb{E}[\delta_t^2] \le \left(1 - \frac{1}{t}\right)^2 \mathbb{E}[\delta_{t-1}^2] + \frac{4G^2}{t^2},$$

and hence $$t^2\mathbb{E}[\delta_t^2] \le (t-1)^2\mathbb{E}[\delta_{t-1}^2] + 4G^2.$$

Summing this inequality from $t = 1$ through T cancels every intermediate term and gives

$$T^2\mathbb{E}[\delta_T^2] \le 4G^2 T.$$

Dividing by T^2 gives

$$\boxed{\mathbb{E}[\mathrm{dist}(\bar{g}_T, C)^2] = \mathbb{E}[\delta_T^2] \le \frac{4G^2}{T}.}$$

This proves the approachability theorem.

Decision Calibration Is Response Satisfiable

Return to $\mathcal{X} = \mathcal{P}_\eta$ and the action-cell payoff $g(p, y)$. Fix any hypothetical distribution $Q \in \Delta(\mathcal{Y})$ over outcomes, and choose a grid forecast p_Q satisfying

$$\|p_Q - Q\|_2 \leq \eta.$$

The pure response p_Q has expected payoff

$$z_Q := \mathbb{E}_{Y \sim Q}[g(p_Q, Y)] = \big(\mathbf{1}\{\mathrm{BR}_u(p_Q) = a\}(Q - p_Q)\big)_{a \in \mathcal{A}}.$$

Decision Calibration Is Response Satisfiable

Return to $\mathcal{X} = \mathcal{P}_\eta$ and the action-cell payoff $g(p, y)$. Fix any hypothetical distribution $Q \in \Delta(\mathcal{Y})$ over outcomes, and choose a grid forecast p_Q satisfying

$$\|p_Q - Q\|_2 \le \eta.$$

The pure response p_Q has expected payoff

$$z_Q := \mathbb{E}_{Y \sim Q}[g(p_Q, Y)] = \big(\mathbf{1}\{\mathrm{BR}_u(p_Q) = a\}(Q - p_Q)\big)_{a \in \mathcal{A}}.$$

Exactly one action block is active, so

$$\|z_Q\|_2 = \|Q - p_Q\|_2 \le \eta.$$

Hence $z_Q \in C_\eta$ for every Q: the decision-calibration target is response satisfiable.

For Action Cells the Vector Payoff Has Norm at Most $\sqrt{2}$

For the target C_η, choose the anchor $c^\circ = 0$. Since exactly one action block is active,

$$G = \max_{p,y} \|g(p, y) - c^\circ\|_2 = \max_{p,y} \|e_y - p\|_2 \le \sqrt{2}.$$

Response satisfiability and Blackwell's theorem therefore give

$$\mathbb{E}\left[\mathrm{dist}(\bar{g}_T, C_\eta)^2\right] \le \frac{8}{T}.$$

For Action Cells the Vector Payoff Has Norm at Most $\sqrt{2}$

For the target C_η, choose the anchor $c^\circ = 0$. Since exactly one action block is active,

$$G = \max_{p,y} \|g(p, y) - c^\circ\|_2 = \max_{p,y} \|e_y - p\|_2 \le \sqrt{2}.$$

Response satisfiability and Blackwell's theorem therefore give

$$\mathbb{E}\left[\mathrm{dist}(\bar{g}_T, C_\eta)^2\right] \le \frac{8}{T}.$$

Because C_η is the radius-η ball, $\|\bar{g}_T\|_2 \le \eta + \mathrm{dist}(\bar{g}_T, C_\eta)$. Taking expectations and applying Cauchy–Schwarz gives

$$\begin{aligned} \mathbb{E}[\|\bar{g}_T\|_2] &\le \eta + \mathbb{E}[\mathrm{dist}(\bar{g}_T, C_\eta)] \\ &\le \eta + \sqrt{\mathbb{E}[\mathrm{dist}(\bar{g}_T, C_\eta)^2]} \\ &\le \eta + \sqrt{\frac{8}{T}}. \end{aligned}$$

Approachability Gives the Swap-Regret Guarantee

Combining the approachability bound with the transfer inequality recalled near the start of the lecture gives

$$\boxed{\mathbb{E}[\mathsf{SwapReg}_T] \leq D_{u,2}\sqrt{|\mathcal{A}|}\left(\eta + \sqrt{\frac{8}{T}}\right).}$$

One Forecast Can Serve Many Users

For N users, concatenate all of their action-cell vectors. Every user sees the same forecast p, which activates one action block per user.

One Forecast Can Serve Many Users

For N users, concatenate all of their action-cell vectors. Every user sees the same forecast p, which activates one action block per user. Use the product target

$$C_\eta^{\mathrm{many}} := \prod_{i=1}^{N} \{z^i : \|z^i\|_2 \leq \eta\}.$$

One grid forecast within η of Q places every user's expected block in its factor, so this target is response satisfiable.

One Forecast Can Serve Many Users

For N users, concatenate all of their action-cell vectors. Every user sees the same forecast p, which activates one action block per user. Use the product target

$$C_\eta^{\text{many}} := \prod_{i=1}^{N} \{z^i : \|z^i\|_2 \leq \eta\}.$$

One grid forecast within η of Q places every user's expected block in its factor, so this target is response satisfiable.

Using the anchor $c^\circ = 0$, and since one block per user is active,

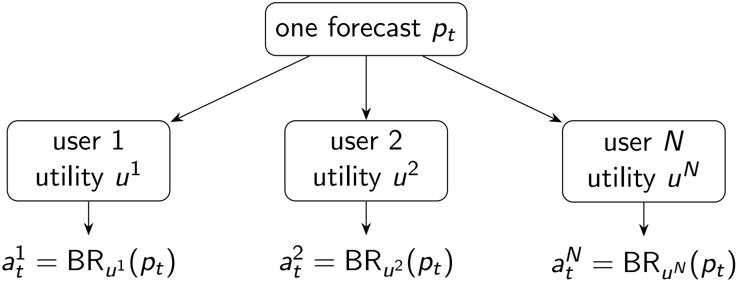
$$\|g(p, y)\|_2 \leq \sqrt{2N}.$$

If $\bar{g}_T^i$ is user i's block of the running average, distance to the product target bounds its excess norm. Thus

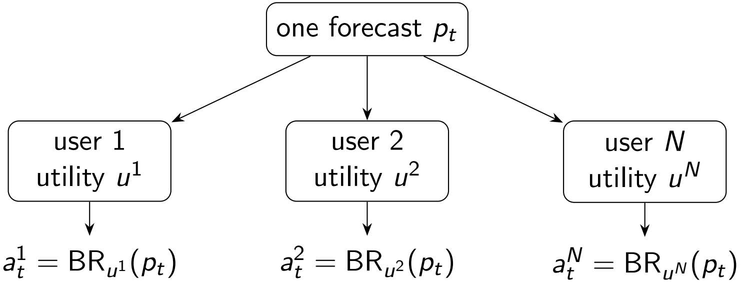
$$\mathbb{E}[\|\bar{g}_T^i\|_2] \leq \eta + 2\sqrt{\frac{2N}{T}}.$$

Hence all N users obtain vanishing swap regret.

One Forecast Becomes a Common Decision Interface

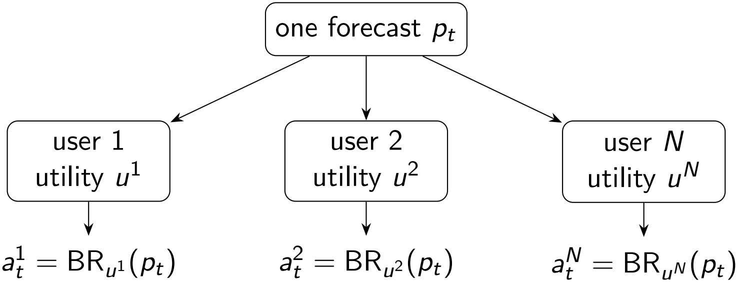


One Forecast Becomes a Common Decision Interface



The users need not share an objective or take the same action. Decision calibration makes the forecast errors balance within *each user's own* action cells, so each user obtains low swap regret relative to their own utility.

One Forecast Becomes a Common Decision Interface



The users need not share an objective or take the same action. Decision calibration makes the forecast errors balance within *each user's own* action cells, so each user obtains low swap regret relative to their own utility.

This is the alignment payoff: one forecasting model can report its beliefs through a common interface, while many users safely translate those reports into decisions serving different goals — i.e. simultanious alignment with many users.

Hedge Can Replace $\sqrt{N}$ by $\sqrt{\log N}$

The Euclidean construction concatenates the N users' payoff vectors. Since $\|g(p, y)\|_2 = O(\sqrt{N})$, its finite-horizon error is

$$O\left(\sqrt{\frac{N}{T}}\right).$$

This dependence on N is a feature of the Euclidean potential, not a limit of simultaneous decision calibration.

Hedge Can Replace $\sqrt{N}$ by $\sqrt{\log N}$

The Euclidean construction concatenates the N users' payoff vectors. Since $\|g(p, y)\|_2 = O(\sqrt{N})$, its finite-horizon error is

$$O\left(\sqrt{\frac{N}{T}}\right).$$

This dependence on N is a feature of the Euclidean potential, not a limit of simultaneous decision calibration.

A different construction based on Hedge can reduce the dependence on the number of users to $\sqrt{\log N}$. So it is possible to align to a very large number of users.

Summary

- Blackwell approachability generalizes minimax: it gives a strategy for satisfying many long-run constraints at once.

Summary

- Blackwell approachability generalizes minimax: it gives a strategy for satisfying many long-run constraints at once.
- Calibration and decision calibration can be expressed as approachability problems, so they can be guaranteed online without assuming that outcomes follow a stochastic model.

Summary

- Blackwell approachability generalizes minimax: it gives a strategy for satisfying many long-run constraints at once.
- Calibration and decision calibration can be expressed as approachability problems, so they can be guaranteed online without assuming that outcomes follow a stochastic model.
- Decision calibration turns forecasts into decision guarantees: users who best respond to the forecasts obtain low swap regret.

Summary

- Blackwell approachability generalizes minimax: it gives a strategy for satisfying many long-run constraints at once.
- Calibration and decision calibration can be expressed as approachability problems, so they can be guaranteed online without assuming that outcomes follow a stochastic model.
- Decision calibration turns forecasts into decision guarantees: users who best respond to the forecasts obtain low swap regret.
- One forecast can serve users with different utility functions. The AI supplies a common prediction interface that can be aligned with many user objectives.

Summary

- Blackwell approachability generalizes minimax: it gives a strategy for satisfying many long-run constraints at once.
- Calibration and decision calibration can be expressed as approachability problems, so they can be guaranteed online without assuming that outcomes follow a stochastic model.
- Decision calibration turns forecasts into decision guarantees: users who best respond to the forecasts obtain low swap regret.
- One forecast can serve users with different utility functions. The AI supplies a common prediction interface that can be aligned with many user objectives.
- The Euclidean construction has error $O(\sqrt{N/T})$ for N users; a Hedge-based construction improves this to $O(\sqrt{\log N/T})$, allowing one model to serve a very large user class.

Thanks!

See you next class!

References

- David Blackwell. "An Analog of the Minimax Theorem." *Pacific Journal of Mathematics*, 1956.
- Dean P. Foster and Rakesh V. Vohra. "Calibrated Learning and Correlated Equilibrium." *Games and Economic Behavior*, 1997.
- Vianney Perchet. "Approachability, Regret and Calibration." *Ph.D. thesis*, Université Pierre et Marie Curie, 2011.
- Aaron Roth and Mirah Shi. "Forecasting for Swap Regret for All Downstream Agents." *ACM EC*, 2024.
- Shengjia Zhao, Michael P. Kim, Roshni Sahoo, Tengyu Ma, and Stefano Ermon. "Calibrating Predictions to Decisions: A Novel Approach to Multi-Class Calibration." *NeurIPS*, 2021.
- Georgy Noarov, Ramya Ramalingam, Aaron Roth, and Stephan Xie. "High-Dimensional Prediction for Sequential Decision Making." *ICML*, 2025.