Games of Incomplete Information Played by Statisticians

Annie Liang

Microsoft Research / UPenn

Becker Friedman Institute

Modeling Beliefs in Games

Classic approach:

- $\rightarrow$ payoff-relevant parameter $\theta \in \Theta$.
- $\rightarrow$ players observe (a sequence of) signals - data
- $\rightarrow$ common prior over combinations of parameters and data

Common data $\rightarrow$ same posterior beliefs.
Private data, CK of posterior beliefs $\rightarrow$ again same beliefs.

Modeling Beliefs in Games

Classic approach:

- $\rightarrow$ payoff-relevant parameter $\theta \in \Theta$.
- $\rightarrow$ players observe (a sequence of) signals - data
- $\rightarrow$ common prior over combinations of parameters and data

Common data $\rightarrow$ same posterior beliefs.
Private data, CK of posterior beliefs $\rightarrow$ again same beliefs.

These implications are:

- $\rightarrow$ descriptively wrong (politics, financial markets, etc.)
- $\rightarrow$ problematic for predictions in settings in which disagreement is important for behavior (trading, etc.).

Goal: relax the CPA, retain discipline on beliefs.

Introduce Model Uncertainty

Generalize CPA to a set of ways to form beliefs given data.

$\rightarrow$ set of priors over combinations of parameters and data

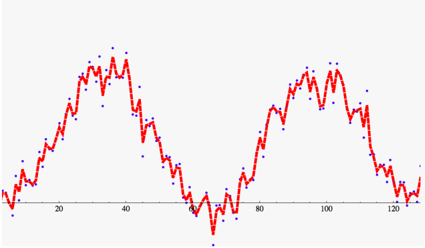
$\rightarrow$ set of frequentist estimators for inferring parameter from data

Introduce Model Uncertainty

Generalize CPA to a set of ways to form beliefs given data.

- $\rightarrow$ set of priors over combinations of parameters and data
- $\rightarrow$ set of frequentist estimators for inferring parameter from data

  example: observe common data $(\mathbf{x}_1, \theta_1), \ldots, (\mathbf{x}_n, \theta_n)$, use different (consistent) estimators to predict out-of-sample



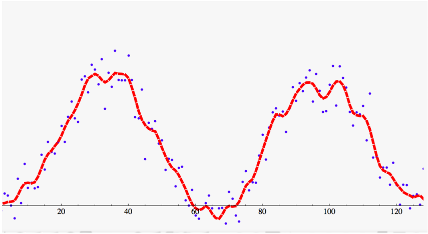
Introduce Model Uncertainty

Generalize CPA to a set of ways to form beliefs given data.

→ set of priors over combinations of parameters and data

→ set of frequentist estimators for inferring parameter from data

example: observe common data $(\mathbf{x}_1, \theta_1), \ldots, (\mathbf{x}_n, \theta_n)$, use different (consistent) estimators to predict out-of-sample



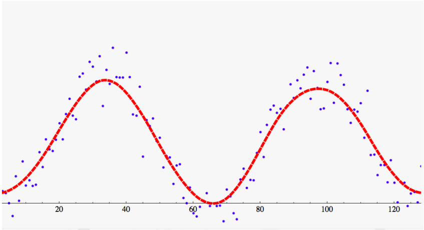
Introduce Model Uncertainty

Generalize CPA to a set of ways to form beliefs given data.

→ set of priors over combinations of parameters and data

→ set of frequentist estimators for inferring parameter from data

example: observe common data $(\mathbf{x}_1, \theta_1), \ldots, (\mathbf{x}_n, \theta_n)$, use different (consistent) estimators to predict out-of-sample



Approach

Discipline imposed by assuming:

- → all learning rules recover θ given sufficient data (statistical consistency).
- → players have common certainty in the (first-order) beliefs induced by the set of learning rules.

Under assumptions, every dataset induces a set of "plausible" hierarchies of beliefs.

Preview of Results

Q: What can an analyst predict based on the proposed refinement alone (i.e. if he doesn't know exact beliefs)?

- → Strict solutions are good predictions if players observe a lot of data.
- → But how much data? Quantity required is increasing in
  - → "complexity" of informational environment (prev. slide: "roughness" and dimensionality of ϕ)
  - → richness of set of interpretations

  Can be impractically large.
- → Implausibility of equilibria/rationalizable actions in complex informational environments.

Preliminaries

Environment: Normal-Form Games

finite set of agents	$\mathcal{I}$
finite action sets	$(A_i)_{i \in \mathcal{I}}, \quad A := \prod_{i \in \mathcal{I}} A_i$
payoffs	$U = \mathbb{R}^{\lvert A \rvert \times \lvert \mathcal{I} \rvert}$

Strict Solution Concepts

Strict Nash Equilibrium. Action profile a is a strict NE if

$$u_i(a_i, a_{-i}) > u_i(a'_i, a_{-i}) \quad \forall\ a_i \neq a'_i$$

for every player i.

Strict Rationalizability. Action a_i is strictly rationalizable for player i if there is a family of sets $(R_j)_{j \in \mathcal{I}}$ such that every $a_j \in R_j$ is a strict best response to some mixed strategy with support in R_{-j}, and $a_i \in R_i$.

Strict Solution Concepts

δ-Strict Nash Equilibrium. Action profile a is a δ-strict NE if

$$u_i(a_i, a_{-i}) > u_i(a_i', a_{-i}) + \delta \quad \forall\ a_i \neq a_i'$$

for every player i.

δ-Strict Rationalizability. Action a_i is δ-strictly rationalizable for player i if there is a family of sets $(R_j)_{j \in \mathcal{I}}$ such that every $a_j \in R_j$ is a δ-strict best response to some mixed strategy with support in R_{-j}, and $a_i \in R_i$.

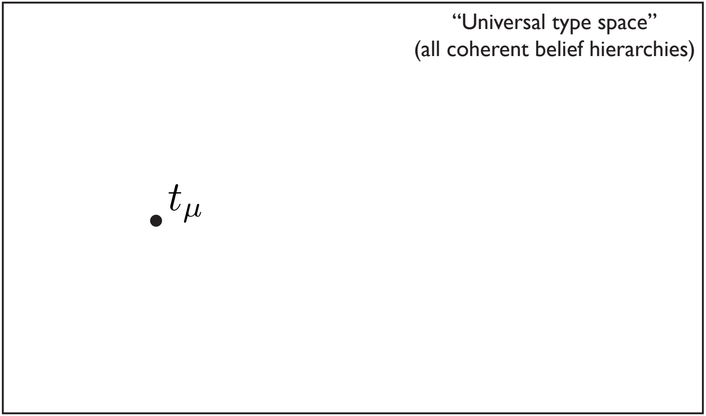
Incomplete Information

parameter space	$\Theta \subseteq \mathbb{R}^k$
	compact
map	$g : \underbrace{\Theta}_{\text{parameters}} \to \underbrace{U}_{\text{payoffs}}$
	bounded and Lipschitz continuous

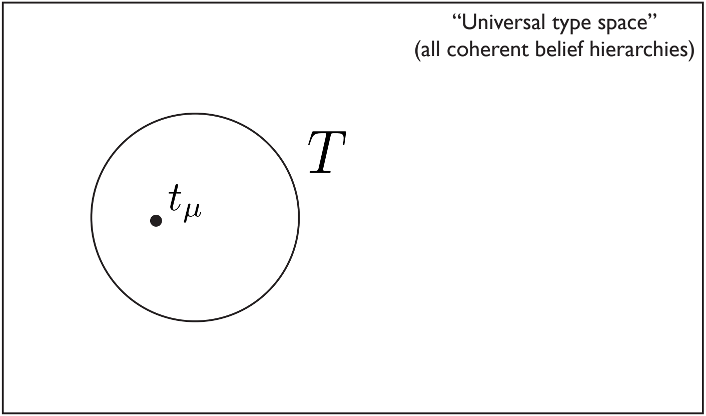
Incomplete Information

parameter space	$\Theta \subseteq \mathbb{R}^k$
	compact
map	$g : \underbrace{\Theta}_{\text{parameters}} \rightarrow \underbrace{U}_{\text{payoffs}}$
	bounded and Lipschitz continuous
first-order belief	$\Delta(\Theta)$
	uncertainty over parameter
second-order belief	$\Delta(\Theta \times \Delta(\Theta))$
	uncertainty over opponent's uncertainty
a "hierarchy" or "type"	$t \in \Delta(\Theta) \times \Delta(\Theta \times \Delta(\Theta)) \times \ldots$

Which Type Space?



Which Type Space?



Related Literature

Agents who Learn from Data: Gilboa, Postlewaite, Samuelson and Schmeidler (1995-2014), Cripps, Ely, Mailath, and Samuelson (2008), Al-Najjar (2009-2014), Acemoglu, Chernozhukov, and Yildiz (2015).

Robust Predictions: Carlsson and van Damme (1993), Kajii and Morris (1997), Weinstein and Yildiz (2007), Morris, Takahashi, and Tercieux (2012), Bergemann and Morris (2005-2016)

Hierarchies of Beliefs: Mertens and Zamir (1985), Brandenburger and Dekel (1993), Monderer and Samet (1989), Dekel, Fudenberg and Morris (2007), Chen, di Tello, Faingold, and Xiong (2010).

Related Literature

Agents who Learn from Data: Gilboa, Postlewaite, Samuelson and Schmeidler (1995-2014), Cripps, Ely, Mailath, and Samuelson (2008), Al-Najjar (2009-2014), Acemoglu, Chernozhukov, and Yildiz (2015).

Robust Predictions: Carlsson and van Damme (1993), Kajii and Morris (1997), Weinstein and Yildiz (2007), Morris, Takahashi, and Tercieux (2012), Bergemann and Morris (2005-2016)

Hierarchies of Beliefs: Mertens and Zamir (1985), Brandenburger and Dekel (1993), Monderer and Samet (1989), Dekel, Fudenberg and Morris (2007), Chen, di Tello, Faingold, and Xiong (2010).

Related Literature

Agents who Learn from Data: Gilboa, Postlewaite, Samuelson and Schmeidler (1995-2014), Cripps, Ely, Mailath, and Samuelson (2008), Al-Najjar (2009-2014), Acemoglu, Chernozhukov, and Yildiz (2015).

Robust Predictions: Carlsson and van Damme (1993), Kajii and Morris (1997), Weinstein and Yildiz (2007), Morris, Takahashi, and Tercieux (2012), Bergemann and Morris (2005-2016)

Hierarchies of Beliefs: Mertens and Zamir (1985), Brandenburger and Dekel (1993), Monderer and Samet (1989), Dekel, Fudenberg and Morris (2007), Chen, di Tello, Faingold, and Xiong (2010).

Related Literature

Agents who Learn from Data: Gilboa, Postlewaite, Samuelson and Schmeidler (1995-2014), Cripps, Ely, Mailath, and Samuelson (2008), Al-Najjar (2009-2014), Acemoglu, Chernozhukov, and Yildiz (2015).

Robust Predictions: Carlsson and van Damme (1993), Kajii and Morris (1997), Weinstein and Yildiz (2007), Morris, Takahashi, and Tercieux (2012), Bergemann and Morris (2005-2016)

Hierarchies of Beliefs: Mertens and Zamir (1985), Brandenburger and Dekel (1993), Monderer and Samet (1989), Dekel, Fudenberg and Morris (2007), Chen, di Tello, Faingold, and Xiong (2010).

Related Literature

Agents who Learn from Data: Gilboa, Postlewaite, Samuelson and Schmeidler (1995-2014), Cripps, Ely, Mailath, and Samuelson (2008), Al-Najjar (2009-2014), Acemoglu, Chernozhukov, and Yildiz (2015).

Robust Predictions: Carlsson and van Damme (1993), Kajii and Morris (1997), Weinstein and Yildiz (2007), Morris, Takahashi, and Tercieux (2012), Bergemann and Morris (2005-2016)

Hierarchies of Beliefs: Mertens and Zamir (1985), Brandenburger and Dekel (1993), Monderer and Samet (1989), Dekel, Fudenberg and Morris (2007), Chen, di Tello, Faingold, and Xiong (2010).

Related Literature

Agents who Learn from Data: Gilboa, Postlewaite, Samuelson and Schmeidler (1995-2014), Cripps, Ely, Mailath, and Samuelson (2008), Al-Najjar (2009-2014), Acemoglu, Chernozhukov, and Yildiz (2015).

Robust Predictions: Carlsson and van Damme (1993), Kajii and Morris (1997), Weinstein and Yildiz (2007), Morris, Takahashi, and Tercieux (2012), Bergemann and Morris (2005-2016)

Hierarchies of Beliefs: Mertens and Zamir (1985), Brandenburger and Dekel (1993), Monderer and Samet (1989), Dekel, Fudenberg and Morris (2007), Chen, di Tello, Faingold, and Xiong (2010).

Approach

Agents Form Beliefs by Learning From Data

Data is a random sequence of observations from a set $\mathcal{Z}$

$$\mathbf{z}_n = (z^1, \ldots, z^n) \sim_{\text{i.i.d.}} P.$$

Notation: Z_n denotes random sequence of length n.

A learning rule is any map

$$f : \underbrace{\bigcup_{n=1}^{\infty} \mathcal{Z}^n}_{\text{datasets}} \quad \rightarrow \quad \underbrace{\Delta(\Theta)}_{\text{first-order beliefs}}$$

Fix $\mathcal{F}$ to be a set of learning rules.

Example Learning Rules

Bayesian Updating:

- $\rightarrow$ $\mu \in \Delta(\Theta)$ is a prior
- $\rightarrow$ conditional on θ, stochastic process ξ^{θ} generates a sequence of signals $\mathbf{z}$
- $\rightarrow$ f maps data $\mathbf{z}$ to the posterior induced by μ and $(\xi_{\theta})_{\theta \in \Theta}$.
- $\rightarrow$ $\mathcal{F}$ identified with different priors

Confidence Intervals

- $\rightarrow$ data is a sequence $\mathbf{z} = (\mathbf{x}_1, \theta_1), \ldots (\mathbf{x}_n, \theta_n)$
- $\rightarrow$ $\phi_{\mathbf{z}} : \mathcal{X} \rightarrow \Theta$ is the best linear fit to the data, $\Theta_{\mathbf{z}}$ is the 95% confidence interval for out-of-sample prediction at fixed $\mathbf{x}^*$.
- $\rightarrow$ f maps $\mathbf{z}$ to a uniform distribution over $\Theta_{\mathbf{z}}$
- $\rightarrow$ $\mathcal{F}$ identified with different distributions

Set of Plausible Hierarchies

For every dataset $\mathbf{z}$,

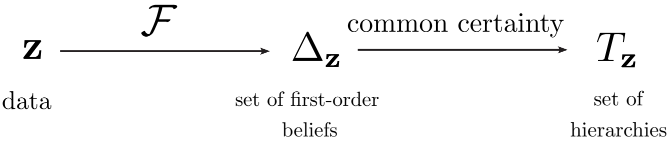
$$\Delta_{\mathbf{z}} = \{f(\mathbf{z}) : f \in \mathcal{F}\} \subseteq \Delta(\Theta)$$

is the set of "plausible" first-order beliefs.

Let $T_{\mathbf{z}}$:= set of types with common certainty in $\Delta_{\mathbf{z}}$:

- ↪ first-order belief is in $\Delta_{\mathbf{z}}$.
- ↪ believes with probability 1 that every other agent's first-order belief is in $\Delta_{\mathbf{z}}$, etc.

Summary of Approach



Statistical Consistency

Let μ be a point mass on the true parameter θ^*.

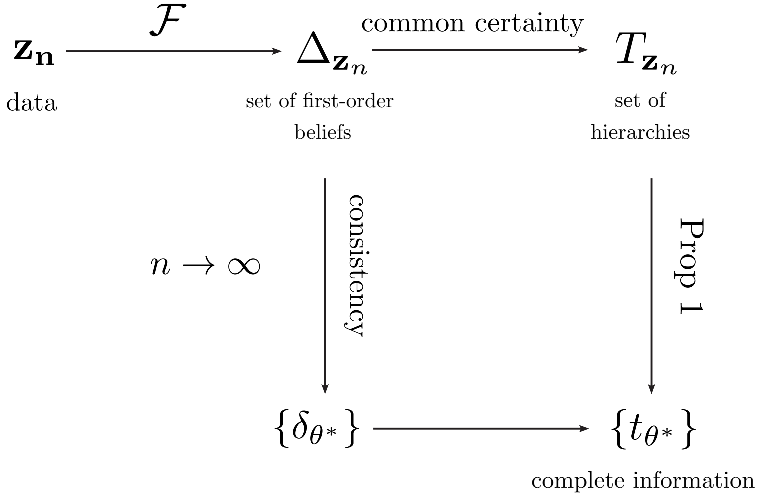
Focus on θ^*-uniform consistency families $\mathcal{F}$:

$$\sup_{f \in \mathcal{F}} \underbrace{d(f(Z_n)}_{\text{belief induced by } f}, \underbrace{\mu)}_{\text{“correct” belief}} \rightarrow 0 \text{ a.s.}$$

(Implies common learning - see Prop 1).

Limit as $n \rightarrow \infty$ is a complete information game.

Limit is Complete Information



Asymptotic Behavior

Behavior Given n Random Observations

Should behavior predicted at $n = \infty$ also be predicted for large n?

→ Let $p_n^{NE}(a)$ be the probability (over datasets $\mathbf{z}_n$) that

$$(\sigma_i)_{i \in \mathcal{I}}, \quad \text{with } \sigma_i(t_i) = a_i \quad \forall\, i, \forall\, t_i \in T_{\mathbf{z}_n}$$

is an (interim Bayesian Nash) equilibrium.

→ Let $p_n^R(a_i, i)$ be the probability (over datasets $\mathbf{z}_n$) that a_i is (interim correlated) rationalizable $\forall\ t_i \in T_{\mathbf{z}_n}$.

Eq property of a is robust to inference if $p_n^{NE}(a) \to 1$ as $n \to \infty$.

→ "Probability that a is guaranteed to be an equilibrium is arbitrarily close to 1 when players observe enough data."

Behavior Given n Random Observations

Should behavior predicted at $n = \infty$ also be predicted for large n?

- $\rightarrow$ Let $p_n^{NE}(a)$ be the probability (over datasets $\mathbf{z}_n$) that

$$(\sigma_i)_{i \in \mathcal{I}}, \quad \text{with } \sigma_i(t_i) = a_i \quad \forall\, i, \forall\, t_i \in T_{\mathbf{z}_n}$$

  is an (interim Bayesian Nash) equilibrium.

- $\rightarrow$ Let $p_n^R(a_i, i)$ be the probability (over datasets $\mathbf{z}_n$) that a_i is (interim correlated) rationalizable $\forall\ t_i \in T_{\mathbf{z}_n}$.

Rationalizability of a_i is robust to inference if $p_n^R(a_i, i) \rightarrow 1$ as $n \rightarrow \infty$.

- $\rightarrow$ "Probability that a_i is guaranteed to be rationalizable is arbitrarily close to 1 when players observe enough data."

Examples: Robustness to Inference Trivially Met

	a_1	a_2
a_1	θ, θ	$0, 0$
a_2	$0, 0$	$\frac{1}{2}, \frac{1}{2}$

true value $\theta^* > 0$

(1) Take $\mathcal{F} = \{f\}$, where $f(\mathbf{z})$ is a point mass on θ^* for all $\mathbf{z}$.

Examples: Robustness to Inference Trivially Met

	a_1	a_2
a_1	θ, θ	$0, 0$
a_2	$0, 0$	$\frac{1}{2}, \frac{1}{2}$

true value $\theta^* > 0$

(1) Take $\mathcal{F} = \{f\}$, where $f(\mathbf{z})$ is a point mass on θ^* for all $\mathbf{z}$.

Assumption (Nontrivial Inference): For every n sufficiently large, $\delta_{\theta^*} \in \text{Interior}(\Delta_{\mathbf{z}_n})$ for a γ-measure of datasets $\mathbf{z}_n$.

Examples: Robustness to Inference Trivially Met

	a_1	a_2
a_1	θ, θ	$0, 0$
a_2	$0, 0$	$\frac{1}{2}, \frac{1}{2}$

true value $\theta^* > 0$

(1) Take $\mathcal{F} = \{f\}$, where $f(\mathbf{z})$ is a point mass on θ^* for all $\mathbf{z}$.

Assumption (Nontrivial Inference): For every n sufficiently large, $\delta_{\theta^*} \in \text{Interior}(\Delta_{\mathbf{z}_n})$ for a γ-measure of datasets $\mathbf{z}_n$.

(2) Set $\Theta := [0, \infty)$.

Examples: Robustness to Inference Trivially Met

	a_1	a_2
a_1	θ, θ	$0, 0$
a_2	$0, 0$	$\frac{1}{2}, \frac{1}{2}$

true value $\theta^* > 0$

(1) Take $\mathcal{F} = \{f\}$, where $f(\mathbf{z})$ is a point mass on θ^* for all $\mathbf{z}$.

Assumption (Nontrivial Inference): For every n sufficiently large, $\delta_{\theta^*} \in \text{Interior}(\Delta_{\mathbf{z}_n})$ for a γ-measure of datasets $\mathbf{z}_n$.

(2) Set $\Theta := [0, \infty)$.

Assumption (Richness): For every i and $a_i \in A_i$, $\exists\ \theta \in \Theta$ such that a_i is strictly dominant in the game $g(\theta)$. (e.g. WY, CvD)

Strict Equilibrium are Robust

Theorem 1. Assume nontrivial inference and richness. Then, the equilibrium property of action profile a^* is robust to inference if and only if a^* is a strict Nash equilibrium in the complete information game with payoffs $u^* = g(\theta^*)$.

Strict-Rationalizability

Conjecture: robustness to inference for rationalizability characterized by strict rationalizability.

Simple counterexample:

	a_3	a_4
a_1	$\theta, 0$	$\theta, 0$
a_2	$0, 0$	$0, 0$

true value $\theta^* = 1$

Action a_1 is robust to inference (strictly dominant at all nearby payoffs), but not strictly rationalizable for player 1.

Strict-Rationalizability

Conjecture: robustness to inference for rationalizability characterized by strict rationalizability.

Simple counterexample:

	a_3	a_4
a_1	$\theta, \mathbf{0}$	$\theta, \mathbf{0}$
a_2	$0, \mathbf{0}$	$0, \mathbf{0}$

true value $\theta^* = 1$

Action a_1 is robust to inference (strictly dominant at all nearby payoffs), but not strictly rationalizable for player 1.

Strict-Rationalizability

Conjecture: robustness to inference for rationalizability characterized by strict rationalizability.

Simple counterexample:

a_1 true value $\theta^* = 1$

a_2

Action a_1 is robust to inference (strictly dominant at all nearby payoffs), but not strictly rationalizable for player 1.

Weak Strict-Rationalizability

Alternative definition of strict rationalizability:

Recursively eliminate (up to) one action a_i that is not a strict best reply to any surviving opponent action.

	a_3	a_4
a_1	$1, 0$	$1, 0$
a_2	$0, 0$	$0, 0$

	a_3	a_4
a_1	$1, 0$	$1, 0$
a_2	$0, 0$	$0, 0$

Action a_i is weakly strict-rationalizable if survives every possible order of one-at-a-time elimination.

Weak Strict-Rationalizability

Alternative definition of strict rationalizability:

Recursively eliminate (up to) one action a_i that is not a strict best reply to any surviving opponent action.

	a_3
a_1	$1, 0$
a_2	

	a_3	a_4
a_1	$1, 0$	$1, 0$
a_2	$0, 0$	$0, 0$

Action a_i is weakly strict-rationalizable if survives every possible order of one-at-a-time elimination.

Weak Strict-Rationalizability

Alternative definition of strict rationalizability:

Recursively eliminate (up to) one action a_i that is not a strict best reply to any surviving opponent action.

	a_3
a_1	$1, 0$
a_2	

	a_4
a_1	$1, 0$
a_2	

Action a_i is weakly strict-rationalizable if survives every possible order of one-at-a-time elimination.

Rationalizability

Theorem 2. Assume nontrivial inference and richness. Then, the rationalizability of a_i^* is robust to inference if a_i^* is strictly rationalizable in the complete information game with payoffs u^*, and only if it is weakly strict-rationalizable.

Recap

Theorems 1 and 2 show that strict solutions of the limit game are solutions with high probability when players observe sufficient data:

$$p_n^{NE}(a) \to 1 \quad \text{and} \quad p_n^R(i, a_i) \to 1$$

for strict equilibria a and strict rationalizable a_i.

How large are $p_n^{NE}(a)$ and $p_n^R(i, a_i)$ for n far from the limit?

Finite Data Behavior

Lower Bound on $p_n^{NE}(a)$

$\exists$ constant $c > 0$ such that for every strict Nash Equilibrium a,

$$p_n^{NE}(a) \geq 1 - \frac{c}{\delta_a^{NE}} \mathbb{E}_{P^n}\left(\sup_{f \in \mathcal{F}} d(f(Z_n), \mu)\right)$$

Lower Bound on $p_n^{NE}(a)$

$\exists$ constant $c > 0$ such that for every strict Nash Equilibrium a,

$$p_n^{NE}(a) \geq 1 - \frac{c}{\delta_a^{NE}} \mathbb{E}_{P^n}\left(\sup_{f \in \mathcal{F}} d(f(Z_n), \mu)\right)$$

degree of "strictness" of equilibrium

Degree of Strictness

Equilibrium:

$\delta_a^{\mathsf{NE}} =$ largest δ s.t. a is a δ-strict Nash equilibrium.

Rationalizability:

$\delta_{a_i}^{\mathsf{R}} =$ largest δ s.t. a_i is δ-strictly rationalizable.

Lower Bound on $p_n^{NE}(a)$

$\exists$ constant $c > 0$ such that for every strict Nash Equilibrium a,

$$p_n^{NE}(a) \geq 1 - \frac{c}{\delta_a^{NE}} \mathbb{E}_{P^n}\left(\sup_{f \in \mathcal{F}} d(f(Z_n), \mu)\right)$$

Lower Bound on $p_n^{NE}(a)$

$\exists$ constant $c > 0$ such that for every strict Nash Equilibrium a,

$$p_n^{NE}(a) \geq 1 - \frac{c}{\delta_a^{NE}} \mathbb{E}_{P^n}\left(\sup_{f \in \mathcal{F}} d(f(Z_n), \mu)\right)$$

rate at which rules in $\mathcal{F}$ jointly learn parameter

Rate of Joint Learning

Decompose into:

(1) Rate of individual learning: how fast does

$$\mathbb{E}_{P^n}\left(d(f(Z_n), \mu)\right) \to 0$$

for each $f \in \mathcal{F}$? Depends on "complexity" of learning problem.

e.g. OLS: number of covariates, kernel regression: "smoothness" of underlying function, classification: VC dimension

Rate of Joint Learning

(2) Opinion diversity across $\mathcal{F}$

$\rightarrow$ nature of correlation across beliefs induced by rules in $\mathcal{F}$

Proposition: For every finite $\mathcal{F} = \{f_1, \ldots, f_K\}$,

$$\underbrace{1 - \sum_{k=1}^{K} p_k^{NE}}_{\text{"counter-monotonic"}} \leq p_n^{NE}(a) \leq \underbrace{1 - \min_{k \in \{1,\ldots,n\}} p_k^{NE}(a)}_{\text{"co-monotonic"}}$$

where p_k^{NE} is defined for $\mathcal{F}_k = \{f_k\}$. (Frechet-Hoeffding bound.)

Lower Bound on $p_n^R(i, a_i)$

$\exists\ c > 0$ such that for every strictly rationalizable action a_i,

$$p_n^R(i, a_i) \geq 1 - \frac{c}{\delta_{a_i}^R} \mathbb{E}_{P^n} \left(\sup_{f \in \mathcal{F}} d(f(Z_n)), \mu) \right).$$

Example: Obfuscation using Irrelevant Data

Example: Risky Joint Investment

→ Two banks decide whether or not to invest in a risky project

→ Payoffs are

	Invest	Don't Invest
Invest	θ, θ	$\theta - c, 0$
Don't Invest	$0, \theta - c$	$0, 0$

where $c > 0$ and $\theta \in \mathbb{R}$ is the unknown return to the invest.

→ Is Invest rationalizable?

Central Bank Releases Data

- → Central bank observes n past projects and their returns.
- → Each project described by $(x_k^1, \ldots, x_k^p) \in \mathbb{R}^p$, and its return is

$$\theta_k = \beta_0 + \beta_1 x_k^1 + \cdots + \beta_{p^*} x_k^{p^*} + \epsilon_k, \quad \epsilon_k \sim \mathcal{N}(0, \sigma^2)$$

  where $p^* < p$.
- → Bank reports $\{(x_k^1, \ldots, x_k^{p'}, \theta_k)\}_{k=1}^n$, where $p' \geq p^*$.
- → Players observe data, find best linear fit, predict return for current project (described by $(x_1^*, \ldots, x_{p'}^*)$).
- → $\mathcal{F}$ consists of all maps from data into a belief over the 95% confidence interval for the prediction.

Extraneous Data Decreases Investment

Corollary. Suppose $\theta^* > 0$. Then for every $n \geq 1$,

$$p_n^R(i, \text{Invest}) \geq 1 - \frac{1}{|\theta^*|}\phi(p')$$

where ϕ is monotonically increasing in the number of reported features p'.

$\rightarrow$ reporting extraneous variables reduces probability that investment is rationalizable.

Concluding Remarks

Conclusion

Proposed a learning-based refinement for the universal type space.

Weak solutions are not good predictions.

Strict solutions are good predictions if players observe a lot of data.

Non-equilibrium/rationalizable play may be expected in complex informational environments with small quantities of data.

Thank You

Discussion

Misspecification

- can relax uniform consistency to uniform convergence to δ_a^{NE}-ball (or $\delta_{a_i}^{R}$-ball) around point mass on θ^*.

Private Data

- unmodeled: how players form beliefs over other players' data.
- without restriction, results need not hold (email-game style counterexamples).

Limit Uncertainty

- in slides, beliefs converge to point mass on θ^*.
- can generalize to convergence to a "limit common prior" on Θ