Advanced Topics in Behavioral Game Theory

Modeling Strategic Behavior as a Machine Learning Problem

Kevin Leyton-Brown

University of British Columbia
Canada CIFAR AI Chair, Amii

James R. Wright

University of Alberta
Canada CIFAR AI Chair, Amii

Lecture Overview

Recap

Prediction in EFGs

Bayesian games

No-regret as a behavioral assumption

Recap: Behavioral Game Theory

- **Descriptive** models, not normative
- **QRE:** All agents quantally best respond to each other
- **CH:** Level-0 agents do something (uniform?), level-1 agents best respond to level-0, level-2 agents best respond to mix of level-0 and level-1, ...
- **QCH:** Level-0 agents do something (uniform?), level-1 agents **quantally** best respond to level-0, level-2 agents **quantally** best respond to mix of level-0 and level-1, ...
- **Linear4:** One story about the "something" that level-0 do: linear combination of simple rules.
- Every model has parameters that need to be set:
  - QRE, QCH: Precision parameter λ
  - CH, QCH: Distribution of levels $\alpha_0, \ldots, \alpha_K$
  - Linear4: Rule weights $w_{\text{unif}}, \ldots, w_{\text{maxmax}}$

Recap: Fitting BGT Models

- Parameterized behavioral game theory models can be fitted and compared using standard supervised learning techniques
- Parameters of cognitively-inspired models can be interesting for their own sake
- Black-box ML models (CNNs) do an even better job of predicting NFG behavior than BGT models
  - Some special domain-specific issues
  - Cognitive models and black-box models each have benefits and drawbacks

Lecture Overview

Recap

Prediction in EFGs

Bayesian games

No-regret as a behavioral assumption

Agent Form

- Behavioral strategies: Each agent decides on an action distribution for each of their infosets: $b_i = (b_i^1, b_i^2, \ldots, b_i^{m_i})$
- Agent form: Can equivalently imagine that each infoset is owned by a different agent
  - Agent for infoset I_i^j chooses b_i^j
  - All the imaginary agents for “real” agent i have the same utility over terminal nodes

Agent Form

- Behavioral strategies: Each agent decides on an action distribution for each of their infosets: $b_i = (b_i^1, b_i^2, \ldots, b_i^{m_i})$
- Agent form: Can equivalently imagine that each infoset is owned by a different agent
  - Agent for infoset I_i^j chooses b_i^j
  - All the imaginary agents for "real" agent i have the same utility over terminal nodes
- Recall: Every randomization over pure strategies (i.e., mixed strategy) has a corresponding behavioral strategy
  - And therefore, a corresponding agent-form strategy

Agent Form QRE

Definition (AQRE)

A profile b of behavioral strategies is a (logit) **agent quantal response equilibrium** with precision λ if

$$b_i^j(a) = QBR_i^j(b_i^{-j}, b_{-i}; \lambda)$$

for every agent i and infoset $I_i^j \in \mathcal{I}_i$.

- Interpretation: Treat “future selves” as entirely different people

Agent Form QRE

Definition (AQRE)

A profile b of behavioral strategies is a (logit) **agent quantal response equilibrium** with precision λ if

$$b_i^j(a) = QBR_i^j(b_i^{-j}, b_{-i}; \lambda)$$

for every agent i and infoset $I_i^j \in \mathcal{I}_i$.

- Interpretation: Treat "future selves" as entirely different people
- **Question:** Is this guaranteed to exist? (**why or why not?**)

Agent Form QRE

Definition (AQRE)

A profile b of behavioral strategies is a (logit) **agent quantal response equilibrium** with precision λ if

$$b_i^j(a) = QBR_i^j(b_i^{-j}, b_{-i}; \lambda)$$

for every agent i and infoset $I_i^j \in \mathcal{I}_i$.

- Interpretation: Treat “future selves” as entirely different people
- **Question:** Is this guaranteed to exist? (**why or why not?**)
- **Question:** Why is this not the same as a QRE of the induced normal form?

Agent Form QRE

Definition (AQRE)

A profile b of behavioral strategies is a (logit) **agent quantal response equilibrium** with precision λ if

$$b_i^j(a) = QBR_i^j(b_i^{-j}, b_{-i}; \lambda)$$

for every agent i and infoset $I_i^j \in \mathcal{I}_i$.

- Interpretation: Treat "future selves" as entirely different people
- **Question:** Is this guaranteed to exist? (**why or why not?**)
- **Question:** Why is this not the same as a QRE of the induced normal form?
  - Quantal distribution over pure strategies corresponds to a particular behavioral strategy
  - But in general does not correspond to quantal distribution over actions at each infoset, given the randomization at the other infosets

Agent Form QRE

Definition (AQRE)

A profile b of behavioral strategies is a (logit) **agent quantal response equilibrium** with precision λ if

$$b_i^j(a) = QBR_i^j(b_i^{-j}, b_{-i}; \lambda)$$

for every agent i and infoset $I_i^j \in \mathcal{I}_i$.

- Interpretation: Treat “future selves” as entirely different people
- **Question:** Is this guaranteed to exist? (**why or why not?**)
- **Question:** Why is this not the same as a QRE of the induced normal form?
  - Quantal distribution over pure strategies corresponds to a particular behavioral strategy
  - But in general does not correspond to quantal distribution over actions at each infoset, given the randomization at the other infosets
- **Question:** Would an “Agent Form Cognitive Hierarchy” model make sense?

“Quantal Response Equilibria for Extensive Form Games” [McKelvey & Palfrey, 1998]

What kinds of claims do M&P make with this model?

1. Normative: AQRE selects a unique sequential equilibrium in generic EFGs
2. Descriptive: AQRE predicts patterns of behavior in a set of experimental data
3. Explaining Anomalies: AQRE can account for behavior (going “Across” in Centipede Game) that was previously explained using altruism

Lecture Overview

Recap

Prediction in EFGs

Bayesian games

No-regret as a behavioral assumption

Level-k for Bayesian Games

- Level-k model assumes agents respond to next level below
- Bayesian games: every agent has a type that determines preferences
- These are straightforwardly combined:

$$\pi_{i,0}(\theta) = f(\theta)$$
$$\pi_{i,k}(\theta) = \arg\max_a \sum_{\theta_{-i} \in \Theta_{-i}} p(\theta_{-i} \mid \theta_i) u_i(a, \pi_{-i,k-1}(\theta_{-i}); \theta_i)$$

$$\pi_i = \sum_{\theta_i \in \Theta_i} p(\theta_i) \sum_{k=1}^{K} a_k \pi_{i,k}(\theta_i).$$

Level-k for Bayesian Games

- Level-k model assumes agents respond to next level below
- Bayesian games: every agent has a type that determines preferences
- These are straightforwardly combined:

$$\pi_{i,0}(\theta) = f(\theta)$$
$$\pi_{i,k}(\theta) = \arg\max_a \sum_{\theta_{-i} \in \Theta_{-i}} p(\theta_{-i} \mid \theta_i) u_i(a, \pi_{-i,k-1}(\theta_{-i}); \theta_i)$$

$$\pi_i = \sum_{\theta_i \in \Theta_i} p(\theta_i) \sum_{k=1}^{K} a_k \pi_{i,k}(\theta_i).$$

- **Question:** Would this approach work for QRE?

"Level-k Auctions" [Crawford & Iriberri, 2007]

- "Stylized fact:" People tend to overbid in first-price auctions (relative to equilibrium bids)
- "Winner's curse" explains this phenomenon for common-value auctions
  - i.e., auctions where everyone has the same value for the good
  - people who over-estimate the value for the good will tend to win the auction if they don't condition on the event of their bid being the winning bid
- BUT: Winner's curse does not explain this phenomenon for **individual value** auctions
- And yet this phenomenon is observed in individual value auctions
- This paper: Do level-k bidding strategies imply overbidding?

Lecture Overview

Recap

Prediction in EFGs

Bayesian games

No-regret as a behavioral assumption

No-regret learning

Definition ((external) regret)

Suppose that a set of players repeatedly play a normal-form game (N, A, u). The (external) **regret** R_T for player $i \in N$ of a sequence of action profiles $a^{(1)}, a^{(2)}, \ldots, a^{(T)}$ is the difference between the utility of the best, in hindsight, single action $a_i^* \in A_i$ that i could have played, and the utility that i actually incurred. Formally,

$$R_T = \max_{a_i^* \in A_i} \sum_{t=1}^{T} u_i(a_i^*, a_{-i}^{(t)}) - u_i(a^{(t)}).$$

No-regret learning

Definition ((external) regret)

Suppose that a set of players repeatedly play a normal-form game (N, A, u). The (external) **regret** R_T for player $i \in N$ of a sequence of action profiles $a^{(1)}, a^{(2)}, \ldots, a^{(T)}$ is the difference between the utility of the best, in hindsight, single action $a_i^* \in A_i$ that i could have played, and the utility that i actually incurred. Formally,

$$R_T = \max_{a_i^* \in A_i} \sum_{t=1}^{T} u_i(a_i^*, a_{-i}^{(t)}) - u_i(a^{(t)}).$$

Definition (no-regret learning)

Let a learning algorithm $f : A^* \to \Delta(A_i)$ be a mapping from finite histories of action profiles to a distribution over actions for player i. We say that f is a **no-regret** learning algorithm if $\mathbb{E}[R_T/T] \to 0$ as $T \to \infty$ in any infinitely repeated game in which $a_i^{(t)} \sim f(a^{(1:T)})$.

No regret as a behavioral assumption

- Lots of algorithms have the no-regret property (regret matching, Hedge, follow-the-regularized-leader, etc.)
- They largely boil down to just playing the action you most wish you had played in hindsight with high probability
- Instead of assuming that people follow a specific procedure for choosing, you can instead assume that they will do some *unspecified* thing that has the no-regret property

No regret as a behavioral assumption

- Lots of algorithms have the no-regret property (regret matching, Hedge, follow-the-regularized-leader, etc.)
- They largely boil down to just playing the action you most wish you had played in hindsight with high probability
- Instead of assuming that people follow a specific procedure for choosing, you can instead assume that they will do some *unspecified* thing that has the no-regret property
- **Question:** Is this a reasonable assumption?

"Econometrics for Learning Agents" [Nekipolov *et al.*, 2015]

- Problem: Given observed bidding behavior in an ad auction, can we estimate the value that individual bidders have for clicks on a given keyword
  - Wrinkle: not the same bidders in every instance of the auction

"Econometrics for Learning Agents" [Nekipolov *et al.*, 2015]

- Problem: Given observed bidding behavior in an ad auction, can we estimate the value that individual bidders have for clicks on a given keyword
  - Wrinkle: not the same bidders in every instance of the auction
- Standard approach: Assume all agents best respond to their preferences
  - Find an assignment of values to players that satisfies that constraint
  - Problem: What if there is no such assignment?
  - Problem: Why should we believe that agents are all best-responding (i.e., in Nash equilibrium)?
- This paper: Assume only that player are doing some sort of no-regret learning
  - Every value assignment to a bidder implies a specific regret for the observed sequence of bids

“Econometrics for Learning Agents” [Nekipolov *et al.*, 2015] #2

Definition (Rationalizable set)

The **rationalizable set** for a bidder i is the set NR of pairs (v_i, ϵ_i) such that i's sequence of bids has regret less than ϵ_i if i's value is v_i.

This paper choose point estimate $(\hat{v}_i, \hat{\epsilon}_i) \in \arg\min_v \min_\epsilon (v, \epsilon) \in NR$

Descriptive claims:

1. Bids are highly **shaded** (only 60% of value)
2. Almost all bidders have a few keywords with a very small error $\hat{\epsilon}_i$, and others with large error

Quantal regret

- The min-regret point estimate implicitly assumes strict regret minimization
- Another approach: **quantal regret** [Nisan & Noti, 2017]
- Point estimate: weighted average over all possible values
- Weights are proportional to exponential of inverse regret:

$$\hat{v}_i = \sum_v \frac{v \exp[-\lambda R(v)]}{\sum_{v'} \exp[-\lambda R(v')]}$$

  where $R(v)$ is the regret implied for player i by a value of v.
- By comparison: Nekipolov et al.'s scheme is something like

$$\hat{v}_i = \lim_{\lambda \to \infty} \sum_v \frac{v \exp[-\lambda R(v)]}{\sum_{v'} \exp[-\lambda R(v')]}$$

Summary

1. Examples of how to extend BGT models to more complex settings
   - Extensive form games
   - Bayesian games
2. Examples of additional assumptions that can be relaxed
   - Specific decision procedure
   - Utility of outcomes vs. value of changes
3. Examples of different kinds of questions BGT can bear on
   - Normative, reasons that specific models have desirable properties
   - Descriptive, predictions of decisions
   - Explanation of anomalies, implications of models