No-Regret Learning

CMPUT 654: Modelling Human Strategic Behaviour

Hart & Mas-Colell (2000)
Nekipolov, Syrgkanis, and Tardos (2015)

Lecture Outline

1. Recap
2. Hart & Mas-Colell (2000)
3. Coarse Correlated Equilibrium
4. Nekipolov, Syrgkanis, and Tardos (2015)

Hart & Mas-Colell (2000)

Why:

- A no-regret algorithm (**regret matching**) that converges to **correlated equilibrium**
- **Influential:** This paper is always cited in this area

1. Defines regret matching algorithm and argues for its plausibility
2. Proves that it converges to correlated equilibrium

Correlated Equilibrium

Definition:
Given an n-agent game G=(N,A,u), a **correlated equilibrium** is a tuple (v, π, σ), where

- $v = (v_1, \ldots, v_n)$ is a tuple of random variables with domains $(D_1, \ldots, D_n)$,
- π is a joint distribution over v,
- $\sigma = (\sigma_1, \ldots, \sigma_n)$ is a vector of mappings $\sigma_i : D_i \to A_i$, and
- for every agent i and mapping $\sigma_i' : D_i \to A_i$,

$$\sum_{d \in D_1 \times \cdots \times D_n} \pi(d) u_i(\sigma_1(d_1), \ldots, \sigma_n(d_n)) \geq \sum_{d \in D_1 \times \cdots \times D_n} \pi(d) u_i(\sigma_1(d_1), \ldots, \sigma_i'(d_i), \ldots, \sigma_n(d_n))$$

Correlated Equilibrium (simplified)

Definition:
Given an n-agent game $G=(N,A,u)$, a **correlated equilibrium** is a distribution $\sigma \in \Delta(A)$ such that for every $i \in N$ and actions $a'_i, a''_i \in A_i$,

$$\sum_{a \in A : a_i = a'_i} \sigma(a)[u_i(a''_i, a_{-i}) - u_i(a)] \leq 0$$

Repeated Setting

- A game G=(N,A,u) is played repeatedly over t=1,2,...
- At time t, agent i selects action a^i_t
- Each agent i receives utility $u_i(a_t)$

Regret Matching

- For every pair of strategies j,k, let $W_{i,t}(j,k)$ be the utility that i would have received at time t by playing k instead of j
    - Unchanged from $u_i(a^t)$ if i didn't play j
- $D_{i,t}(j,k)$ is the average of $W_{i,t}(j,k)$ - $u_i(a^t)$ up until time t
- At each time step, each agent chooses between actions with **positive** D(j,k), where j is the **most-recent action**, and the most-recent action j

Convergence of Regret Matching

Theorem:
If all players play according to regret matching, then the empirical distributions of play converge to the set of correlated equilibria.

Coarse Correlated Equilibrium

- Instead of getting to replace each action with an arbitrary action, compare to the case where we play a single action:

Definition:
Given an n-agent game G=(N,A,u), a **coarse correlated equilibrium** is a distribution $\sigma \in \Delta$(A) such that for every i $\in$ N and action $a'_i \in A_i$,

$$\sum_{a_{-i} \in A_{-i}} \sigma(a) u_i(a'_i, a_{-i}) - \sum_{a \in A} \sigma(a) u_i(a) \leq 0$$

Convergence of Multiagent No-Regret Learning

Proposition:
If every agent plays a no-regret learning algorithm, then the empirical distribution of play will converge to a coarse correlated equilibrium.

Nekipolov, Syrgkanis, and Tardos (2015)

Why:
Application of a non-equilibrium behavioural rule to econometrics

1. Define rationalizable set NR

2. Prove properties of NR for sponsored search auctions

3. Apply to value estimation

Setting: Sponsored Search

- There are k slots
- Each agent submits a bid b_i
- Highest bid gets first slot, etc.
- Each agent pays bid of next-highest slot
- Payments are **per-click** rather than **per-impression**

Problem: Estimating Types

- Each agent has a **value** v_i for a click
- We want to **estimate** what those values are, based on bids
- *Previously:* Assume **equilibrium**
- *Now:* Assume **no-regret learning**

Rationalizable Set

Definition:
The **rationalizable set** *NR* is the set 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.

Data Analysis

Claims:

1. Bids are highly **shaded** (only 60% of value)

2. Almost all accounts have a few keywords with very small error, and others with large error

Epilogue

Some questions:

1. Regret matching includes a notion of **inertia**. How closely related to I-SAW is it?

2. Why do we think that the **smallest** rationalizable error is the one to use for point estimates?