Overview

Repeated Games in Computer Systems

Internet routing

advertising auctions

No-Regret converges to efficient outcomes:

Cost $\leq r$ Opt + Rate.

- Standard algorithms: Rate $= O(1/T)$.
- Fast Convergence: Rate $\approx O(1/T)$ [e.g. Serebinski et al.'15]

But seemingly only for restricted algorithms/settings.

Main result:
Fast Convergence of Learning is Robust!

Low Approximate Regret

Cost $\leq (r + \epsilon) Opt + ApsRate.$

Feedback
Prior work: Expectation feedback (information/computation issues).
Our work: Realized (costs given actions of the other players).

Range of Algorithms
Prior work: Specialized algorithms.
Our work: Broad class of standard algorithms.

Bandit Feedback
Fast convergence even when players only lose own action.

Dynamic Population Games
Fast convergence even when players come and go with time.

Prior work:
- [Zinkevich et al.'08, Stoltz et al.'05, Stoltz et al.'06, AdaNormalHedge '15]

Low Approximate Regret

- New feedback game with a players repeat for 2 time steps.
- Player $i$ has action $A_{i,t}$ with $A_{i,t} \in [1 : d]$. $A_{i,t} \to [1 : d]$.
- Each action is distributed $\text{Cost}_{i,t}(\cdot)$ and sampled action $A_{i,t}$.
- Observable cost vector $< \text{Cost}_{i_1,t}, \ldots, \text{Cost}_{i_n,t} >$ given other players sampled actions.

Define Social Cost: $\text{SC}_{i}(t) = \sum_{t} \text{Cost}_{i,t}$ (observe performance with Average Social Cost $\sum_{t} \text{Cost}_{i,t}$)

How does avg. cost approximate Opt $\approx \text{SC}_{i}(t)$?

Efficiency of Low Approximate Regret

Definition: Low Approximate Regret

Player $i$ satisfies the Low Approximate Regret property for parameter $\epsilon > 0$ and function $\text{Apopt}(\cdot)$ if for all action distributions $\text{Opt}(\cdot)$, $t = 1, \ldots, T$,

$$\left(\sum_{t} \text{Cost}_{i,t} - \text{Apopt}(\cdot)\right) \leq \epsilon \text{SC}_{i}(t).$$

Multiplicative approximations in causal regret

Definition: Smooth Games

Cost minimization game $\text{Cost}_{i}(x) \leq \epsilon$ for all strategy profiles $x$. $\epsilon$-smoothness on uniform distribution.

In causal games or after functions are $\epsilon$-smooth (Spence and Verge '12),

$$\sum_{t} \text{Cost}_{i,t} \leq \epsilon \text{SC}_{i}(t).$$

Hybrid Feedback

Bandit feedback: Player $i$ only sees own costs, not whole vector cost.

If Low Approximate Regret is satisfied (in expectation) under bandit feedback, efficiency guarantees in full information setting.

Dynamic Population Games

Beyond static population

Definition: Dynamic Population Games [Lykouris et al.'15]

Introduces between stochastic and worst-case changes in population.

- At every round $t$, every player is replaced with probability $\mu$.
- The types of the replacing players are adversarially selected.

Behavioral assumption
Players receive Low Approximate Regret feedback compared to a shifting benchmark.

Defining Low Approximate Regret feedback in Dynamic Population Games

In any $(\epsilon, \delta)$-no-regret game with a sequence $\{i_1, \ldots, i_T\}$ of players that is not Low Approximate Regret for an algorithm, for every $t$ $\epsilon_{t} = \epsilon_{i_t}$.

Algorithm with Low Approximate Regret for shifting experts

Navy Hinge: Hinge update that mix $\epsilon = \frac{1}{\mu}$ uniform noise at every round.

AdaptHinge [Luo-Schapire '15]:

$$\text{ApxReg}(\cdot) \approx \frac{\epsilon}{\mu} \log(\text{Cost}(\cdot)).$$

Proof for Navy Hinge

For $\epsilon = 1$ the comparator at round $t$ is $\epsilon = \frac{1}{\mu}$ too.

Algorithm with Low Approximate Regret for shifting experts

AdaptHinge [Luo-Schapire '15]:

$$\text{Reg}(\cdot) \approx \frac{\epsilon}{\mu} \log(\text{Cost}(\cdot)) + \epsilon\Theta(1).$$

Comparison to Previous work

- Broader range of algorithms [AdaptHinge '15]: Small regret esp. [AdaptHinge '15] typically requires high memory.
- Bregman: No regret includes regret bounds [Navy Hinge - sweeping & adaptive]
- Improved rates of turnover

Learning in Games: Robustness of Fast Convergence

Dylan Foster, Zhiyuan Li, Thodoris Lykouris, Karthik Sridharan, and Éva Tardos

$d$, $T$, $\mu$, $\epsilon$ $\rightarrow$

$\text{Cost}_{i,t}(\cdot)$

$\text{Reg}(\cdot)$

$\epsilon_{i_t}$

$\log(\text{Cost}(\cdot)) + \epsilon\Theta(1)$

$\epsilon = \frac{1}{\mu}$

$\Theta(1)$

$\text{Opt}$

$\text{ApxReg}$

$\text{Apopt}$

$\text{SC}$

$\epsilon_{t}$

$\mu$

$\epsilon$

$\epsilon_{i_t}$

$\text{Cost}_{i,t}$

$\text{Cost}$

$\mu$

$\epsilon$

$\epsilon_{i_t}$

$\text{Cost}_{i,t}$

$\text{Cost}$

$\mu$

$\epsilon$

$\epsilon_{i_t}$

$\text{Cost}_{i,t}$

$\text{Cost}$

$\mu$

$\epsilon$

$\epsilon_{i_t}$

$\text{Cost}_{i,t}$

$\text{Cost}$

$\mu$

$\epsilon$

$\epsilon_{i_t}$

$\text{Cost}_{i,t}$

$\text{Cost}$

$\mu$

$\epsilon$

$\epsilon_{i_t}$

$\text{Cost}_{i,t}$

$\text{Cost}$

$\mu$

$\epsilon$

$\epsilon_{i_t}$

$\text{Cost}_{i,t}$

$\text{Cost}$

$\mu$