Dynamics and Equilibria

Algorithmic Game Theory '26

ΑΛΜΑ, ΣΗΜΜΥ

Nash Equilibria

Pure Nash equilibrium (PNE). Strategy profile s on pure strategies where no player has incentive to deviate:

$$\forall i \in N, s_i' \in S_i : c_i(s) \leq c_i(s_i', s_{-i})$$

Mixed Nash equilibrium (MNE). Strategy profile s (mixed strategies allowed) where no player has incentive to deviate:

$$\forall i \in N, s_i' \in S_i : E[c_i(s)] \leq E[c_i(s_i', s_{-i})]$$

Strong Nash equilibrium. Strategy profile s on pure strategies where in no deviating coalition one player in the coalition benefits without some other in the coalition losing.

Correlated Equilibria

Correlated equilibrium (CorEq). Distribution σ on strategy profiles where no player has incentive to deviate from her (any) assigned pure strategy to any of her (pure) strategies if the others are playing according to the distribution:

$$\forall i \in N, s_i, s'_i \in S_i : E_{s\sim\sigma}[c_i(s)|s_i] \leq E_{s\sim\sigma}[c_i(s'_i, s_{-i})|s_i]$$

Interpretation:

- A central authority announces to the players a distribution over strategy profiles
- Then it draws a strategy profile according to that distribution and announces to every player her assigned strategy
- Given her strategy s_i the player has no incentive to deviate to an s'_i considering only the strategy profiles of the distribution where her strategy is s_i.

Correlated Equilibria

Correlated equilibrium (CorEq). Distribution σ on strategy profiles where no player has incentive to deviate from her (any) assingned strategy to any of her strategies if the others are playing according to the distribution:

$$\forall i \in N, s_i, s_i' \in S_i : E_{s\sim\sigma}[c_i(s)|s_i] \leq E_{s\sim\sigma}[c_i(s_i', s_{-i})|s_i]$$

Example: Traffic lights (costs inside the array)

	stop	*go*
stop	$1, 1$	$1, 0$
go	$0, 1$	$5, 5$

- Four profiles: {top,left} {top,right} {bottom,left} {bottom, right}.
- Correlated equilibrium: 1/2 to {top,right} 1/2 to {bottom,left}

(Pure Nash equilibria? Mixed Nash Equilibria?)

Coarse Correlated Equilibria

Coarse Correlated equilibrium (CCE). Distribution σ on strategy profiles where no player has incentive not to follow the central authority:

$$\forall i \in N, s_i' \in S_i : E_{s\sim\sigma}[c_i(s)] \leq E_{s\sim\sigma}[c_i(s_i', s_{-i})]$$

Connection to Correlated equilibria:

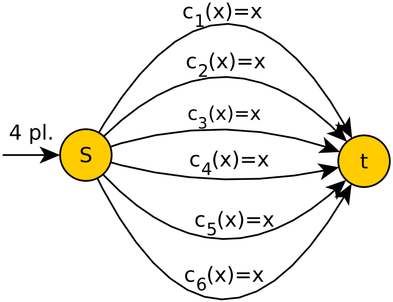
- Differing: Any player has no incentive not to follow the authority before seeing her assigned strategy.
- A Correlated equilibrium is Coarse Correlated since for all s_i:

$$E_{s\sim\sigma}[c_i(s)|s_i] \leq E_{s\sim\sigma}[c_i(s_i', s_{-i})|s_i]$$

  and multiplying each with the "correct" probability will imply

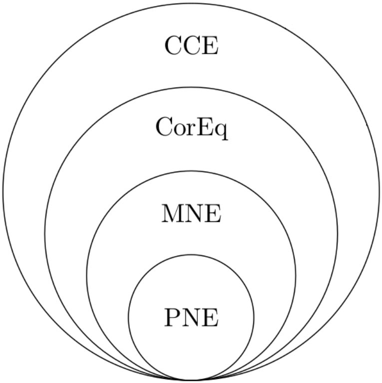
$$\begin{array}{rcl} sum_{s_i\in S_i} p_i E_{s\sim\sigma}[c_i(s)|s_i] & \leq & \sum_{s_i\in S_i} p_i E_{s\sim\sigma}[c_i(s_i', s_{-i})|s_i] \\ \Leftrightarrow E_{s\sim\sigma}[c_i(s)] & \leq & E_{s\sim\sigma}[c_i(s_i', s_{-i})] \end{array}$$

Example



- PNE: Four players in any four edges
- MNE: Each player plays the uniform distribution
- CorEq: Uniform distribution over strategy profiles where two players share an edge and each of the other two has her own.
- CCE: As above but only for profiles that use either edges 1, 3 and 5 or 2, 4 and 6

Equilibria (Strict) Hierarchy



A MNE is a CorEq. Why?

- $E[c_i(s)] \leq E[c_i(s_i', s_{-i})]$
- Strategies on the support of s_i cost (on expectation) equal to $E[c_i(s)]$
- Authority's distribution implied by the MNE
- Any (pure) strategy s_i assigned to the player satisfies

$$E_{s\sim\sigma}[c_i(s)|s_i] \leq E_{s\sim\sigma}[c_i(s_i', s_{-i})|s_i]$$

Examples

Congestion Games

- Potential function exists: $\Phi(f) = \sum_{e \in E} \sum_{i=1}^{f_e} c_e(i)$
- Best response dynamics may have poor convergence rates
- PLS complete to compute a pure Nash equilibrium in general
- Easy for Network CGs with a single source or sink
- What about weighted Congestion Games?

Max-Cut Game

- Potential function exists: $\Phi(S) = \left| \{\{u,w\} \in E : u \in S, w \in V \setminus S\} \right|$
- Best response dynamics converge quickly $\Rightarrow$ efficient pure Nash equilibrium computation
- What about weighted Max-Cut?

Best Response Dynamics in Potential Games

Consider any finite potential game.

Best response dynamics converge to a minimizer of the potential.

- Consider the best response graph, a directed graph with all possible configurations as vertices
- An edge from one configuration points to another iff they differ in a single player's strategy who is in her best response in the destination-configuration
- Finite game implies finite number of vertices
- Existence of a potential implies no cycles
- Thus, bounded longest path $\Rightarrow$ from every initial configuration, best response dynamics converge.

The Framework

A single player, the Learner, having an action set $A = \{a_1, a_2, \ldots, a_n\}$ plays a game for T rounds.
At time t:

1. The Learner picks a distribution p^t on A as her mixed strategy.
2. An Adversary assigns a cost $c^t : A \to [0, 1]$ to the actions of A
3. The Learner draws an action a^t accordig to her distribution and incurs cost $c^t(a^t)$, yet she learns all the costs.

(informal) Goal:

Keep the Learner's cost as close to the optimal (in some sense)

But what can we hope for?

Gap between Learner's Cost and Optimal

Learner needs randomized strategies

- Learner: deterministic action a^t
- Adversary: $c(a^t) = 1$ and $c(a) = 0$ for all $a \neq a^t$
- In T timesteps there is a $a \in A$ with $c^t(a) = 1$ at most $\frac{T}{n}$ times
- Learner pays T, Adversary pays at most T/n

Cannot vanish gap if optimal switches strategies

- Learner: $A = \{a_1, a_2\}$ and always for some $a_j : p^t(a_j) \geq \frac{1}{2}$
- Adversary: $c^t(a_j) = 1$ while $c^t(a_{j+1^*}) = 0$
- Optimal with switching strategies=0
- Learner's cost at least $T/2$

Regret Minimization

We focus on cases where the Learner

- uses randomized strategies and
- compares to fixed actions.

Regret with respect to action a:

$$\frac{1}{T}\Big[\sum_{i=1}^{T} c^t(a^t) - \sum_{i=1}^{T} c^t(a)\Big]$$

Goal: $\underline{\text{Vanishing Regret as } T \to \infty \text{, for all } a}$

Good news: Simple algorithm with Regret$=O\left(\sqrt{\frac{\ln n}{T}}\right)$, w.r.t. any a.

Bad news: Regret is $\Omega\left(\sqrt{\frac{\ln n}{T}}\right)$

Lower Bound

Cosider a setting with action set $A = \{a_1, a_2\}$

- Adversary choses uniformly either $(1, 0)$ or $(0, 1)$ as $(c^t(a_1), c^t(a_2))$, at any t.
- Any action a_i at any t has expected cost $\frac{1}{2}$, independent of the Learner's choice
  $\Rightarrow$ Learner's expected cost always equals $\frac{T}{2}$
- Assigning costs to a_1 and a_2 is like putting balls in 2 bins.
- After T balls: min bin is expected to have $\frac{T}{2} - \Theta(\sqrt{T})$
  $\Rightarrow$ Optimal strategy's expected cost is $\frac{T}{2} - \Theta(\sqrt{T})$

<u>Thus, Learner's cost-OPT= $\Theta(\sqrt{T}) \Rightarrow Regret = \Theta(1/\sqrt{T})$</u>

Multiplicative Weights Update

The Multiplicative Weights Update (MWU) algorithm maintains and updates weights for the actions

- Initially $w^1(a) = 1$ for al $a \in A$
- At time t play action a with probability

$$\frac{w^t(a)}{\sum_{a \in A} w^t(a)}$$

- For some ϵ, update the weights using

$$w^{t+1}(a) = w^t(a) \cdot (1 - \epsilon)^{c^t(a)}$$

MWU has expected regret $O\left(\sqrt{\frac{\ln n}{T}}\right)$ w.r.t. any $a \in A$.

Seen differently: MWU has expected regret w.r.t. any $a \in A$ at most $\epsilon > 0$ after $O\left(\frac{\ln n}{\epsilon^2}\right)$ iterations.

Analysis

Update rule:

$$w^{t+1}(a) = w^t(a) \cdot (1-\epsilon)^{c^t(a)} = w^1(a)(1-\epsilon)^{\sum_{k=1}^{t} c^k(a)} = (1-\epsilon)^{\sum_{k=1}^{t} c^k(a)}$$

(ϵ used for the exploration vs exploitation tradeoff)

Normalizing factor in $\frac{w^t(a)}{\sum_{a \in A} w^t(a)}$:

$$\Gamma^t = \sum_{a \in A} w^t(a)$$

After T steps the best action a^* minimizes $\sum_{k=1}^{t} c^k(a)$ over $a \in A$.
Thus, with $OPT = \sum_{k=1}^{t} c^k(a^*)$

$$\Gamma^T \geq w^T(a^*) = (1-\epsilon)^{OPT}$$

Analysis (2)

Cost of Algorithm at time t, ν^t:

$$\sum_{a\in A} p^t(a)\cdot c^t(a) = \sum_{a\in A} \frac{w^t(a)}{\Gamma^t}\cdot c^t(a).$$

Γ^{t+1} as a function of Γ^t:

$$\begin{aligned}
\Gamma^{t+1} &= \sum_{a\in A} w^{t+1}(a) \\
&= \sum_{a\in A} w^t(a)\cdot(1-\epsilon)^{c^t(a)} \\
&\le \sum_{e\in A} w^t(a)\cdot(1-\epsilon c^t(a)) \\
&= \Gamma^t(1-\epsilon\nu^t),
\end{aligned}$$

for $\epsilon \in (0, \frac{1}{2})$, yielding for Γ^T

$$(1-\epsilon)^{OPT} \le \Gamma^T \le \underbrace{\Gamma^1}_{=n}\prod_{t=1}^{T}(1-\epsilon\nu^t)$$

$$OPT\cdot\ln(1-\epsilon) \le \ln n + \sum_{t=1}^{T}\ln(1-\epsilon\nu^t)$$

Using $\ln(1-x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \dots$

$$OPT\cdot\left[-\epsilon-\epsilon^2\right] \le \ln n + \sum_{t=1}^{T}(-\epsilon\nu^t)$$

$$\sum_{t=1}^{T}\nu^t \le OPT\cdot(1+\epsilon) + \frac{\ln n}{\epsilon} \le OPT + \epsilon T + \frac{\ln n}{\epsilon}$$

choose $\epsilon = \sqrt{\frac{\ln n}{T}}$ to equalize the last two terms

$$\sum_{t=1}^{T}\nu^t - OPT \le 2\sqrt{T\ln n}$$

No-Regret Dynamics

Consider a minimization game played repeatedly.
Players act simultaneously and at time $t = 1, 2, \ldots, T$:

1. Each player i uses a no-regret algorithm to decide on a mixed strategy p_i^t
2. Each player i receives a vector c_i^t of expected costs for her pure strategies

Player i at time t has distribution p_i^t.

- Let σ^t be the probability distribution on strategy profiles implied by the p_i^t 's
- Let $\sigma = \frac{1}{T} \sum_{i=1}^{T} \sigma^t$ be their time averaged distribution

Distribution σ will serve as an approxiamte CCE

Convergence to Approximate CCE

Distribution $\sigma = \frac{1}{T}\sum_{i=1}^{T} \sigma^t$ will serve as an approxiamte CCE

- For any $\epsilon > 0$ there exist a large enough T so that the expected regret for all players is at most ϵ
- For the cost of σ:

$$E_{s\sim\sigma}[c_i(s)] = \frac{1}{T}\sum_{t=1}^{T} E_{s\sim\sigma^t}[c_i(s)]$$

- For the cost of any deviation s'_i:

$$E_{s\sim\sigma}[c_i(s'_i, s_{-i})] = \frac{1}{T}\sum_{t=1}^{T} E_{s\sim\sigma^t}[c_i(s'_i, s_{-i})]$$

- Right Hand Sides differ by at most ϵ, thus:

$$E_{s\sim\sigma}[c_i(s)] \leq E_{s\sim\sigma}[c_i(s'_i, s_{-i})] + \epsilon$$

Swap-Regret Dynamics

Swap regret with respect to a function $\delta : A \to A$:

$$\frac{1}{T}\left[\sum_{i=1}^{T} c^t(a^t) - \sum_{i=1}^{T} c^t(\delta(a^t))\right]$$

Goal: Vanishing swap Regret as $T \to \infty$, for all δ

- Existence of no-regret algorithm implies existence of no swap regret algorithms
- No swap regret implies no regret: general vs constant functions δ

No swap-regret dynamics converge to approximate CorEq.

$$E_{s\sim\sigma}[c_i(s)] \leq E_{s\sim\sigma}[c_i(\delta(s_i'), s_{-i})] + \epsilon$$

(using notation from the no regret dynamics case)