Lecture 9

Wednesday, 21 September 2022 11:12 AM

Recall: CCE & interpretation of coordinating device

σ is a CCE if $\forall i, \forall s_i' \in S_i$,

$$\mathbb{E}_{s\sim\sigma}[u_i(s)] \geq \sum_{s \in S} u_i(s_i', s_{-i})\, \sigma(s) \quad s = (s_1, \dots, s_n)$$

Traffic game:

	S	G
S	0 \ 0	0 \ 10
G	10 \ 0	−100 \ −100

- can compute all CCE in poly-time by solving an LP

Q. Consider a game:

	L	R
U	0 \ 0	1 \ 1
D	1 \ 1	0 \ 0

It has 2 pure NE (D,L) & (U,R) & 1 mixed NE $((1/2, 1/2), (1/2, 1/2))$. Is the set of CCE equal to the set of MNE?

Regret Minimization (External Regret)

Single agent, n pure strategies, multiple time steps, played against an adversary.

- at time $t = 1 \dots T$
  - agent picks a distribution $p^t \in \Delta_n$
  - adversary sees p^t, picks a cost fn $c^t : [n] \to [0,1]$
  - nature picks $a^t \sim p^t$, agent gets cost $c^t(a^t)$
  - agent learns c^t

Objective: minimize external regret in expectation

$$\text{Regret}_T = \frac{1}{T}\left(\sum_{t=1}^{T} c^t(a^t) - \min_{a\in[n]} \sum_{t=1}^{T} c^t(a)\right)$$

i.e., regret is excess cost over best strategy in hindsight

- expectation taken over randomness of nature, agent, & adversary.
- at time $t+1$, agent knows $c^1, c^2, \dots, c^t$, adversary knows $p^1, \dots, p^t$

Agent has "no (external) regret" if $\lim_{T\to\infty} \text{Regret}_T \to 0$

Deterministic Algorithms:

- Since $\forall t$ p^t is a point distribution, a^t is known to adversary.
- let $c^t(a^t) = 1$, $c^t(a) = 0$ $\forall a \neq a^t$.
- then $\sum_{t=1}^{T} c^t(a^t) = T$, and $\sum_a \sum_{t=1}^{T} c^t(a) = T$. hence

$$\min_{a\in[n]} \sum_{t=1}^{T} c^t(a) = T/n$$

thus $\text{Regret}_T = \frac{1}{T}\left(T - \frac{T}{n}\right) = 1 - \frac{1}{n}$

hence, deterministic algos cannot have no regret.

Multiplicative Weight Algorithm

Define: $D^t(a) := \sum_{\tau \leq t} c^\tau(a)$

Suppose a^* has minimum cost at time t, would like to choose a^*. But deterministic algos have high regret

Idea: use soft min.

Given n nos. $x_1, \dots, x_n$

$$\max_i x_i \leq \ln \sum_i e^{x_i}$$

$$\& \quad \min_i x_i \geq -\ln \sum_i e^{-x_i}$$

(can replace e by any constant > 1)

$$\min_i x_i \geq -\ln_{1/(1-\varepsilon)} \sum_i \left(\frac{1}{1-\varepsilon}\right)^{-x_i} = -\ln_{1/(1-\varepsilon)} \sum_i (1-\varepsilon)^{x_i}$$

So in round $t+1$, we choose action a w.p. $\propto (1-\varepsilon)^{D^t(a)}$

Or, define $W^t(a) = (1-\varepsilon)^{D^t(a)}/n$

$W^t = \sum_a W^t(a)$

then $p^{t+1}(a) = W^t(a)/W^t$

So, the algo:

- in round 1, $p^1 = W^1 = (1/n, \dots, 1/n)$
- in round $t+1$, $W^t(a) = (1-\varepsilon)^{D^t(a)}/n$

  $W^t = \sum_a W^t(a)$

  $p^{t+1}(a) = W^t(a)/W^t$

Analysis:

Note:

- $W^t(a) = (1-\varepsilon)^{D^t(a)}/n = (1-\varepsilon)^{D^{t-1}(a) + c^t(a)}/n = W^{t-1}(a)\,(1-\varepsilon)^{c^t(a)}$
- if $c^t(a) > 0$, then $W^t(a) < W^{t-1}(a)$

  if $\exists a : c^{t-1}(a) > 0$, $W^t < W^{t-1}$
- if $c^t(a) = 0$, then $W^t(a) = W^{t-1}(a)$, and $p^{t+1}(a) = \frac{W^t(a)}{W^t} \geq \frac{W^{t-1}(a)}{W^{t-1}} = p^t(a)$

Will use:

(i) $\ln(1-x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \dots$ for $x \in [0,1)$

$\Rightarrow$ (ii) $-x - x^2 \leq \ln(1-x) \leq -x$ for $x \in [0, \frac{1}{2}]$

(iii) $(1-x)^y = 1 - xy + x^2 \frac{y(y-1)}{2} - x^3 \frac{y(y-1)(y-2)}{3\cdot 2} + \dots$

$\leq 1 - xy$ for $x, y \in [0,1]$

Now,

$$W^t = \frac{1}{n}\sum_a (1-\varepsilon)^{D^t(a)} = \frac{1}{n}\sum_a (1-\varepsilon)^{D^{t-1}(a) + c^t(a)}$$

$$= \sum_a W^{t-1}(a)(1-\varepsilon)^{c^t(a)}$$

$$\leq \sum_a W^{t-1}(a)(1 - \varepsilon c^t(a))$$

$$= W^{t-1} - \varepsilon \sum_a W^{t-1}(a) c^t(a)$$

$$\Rightarrow \frac{W^t}{W^{t-1}} \leq 1 - \varepsilon \frac{\sum_a c^t(a) W^{t-1}(a)}{W^{t-1}} = 1 - \varepsilon \sum_a p^t(a) c^t(a)$$

Thus, $W^t \leq \prod_{\tau=1}^{t}\left(1 - \varepsilon \sum_a p^\tau(a) c^\tau(a)\right)$

now, at time t, let a^* be the minimum cost action

Then $(1-\varepsilon)^{D^t(a^*)} \leq n W^t \leq n \prod_{\tau=1}^{t}\left(1 - \varepsilon \sum_a p^\tau(a) c^\tau(a)\right)$

or, $D^t(a^*) \ln(1-\varepsilon) \leq \sum_{\tau=1}^{t} \ln\left(1 - \varepsilon \sum_a p^\tau(a) c^\tau(a)\right) + \ln n$

$$\leq \ln n + \sum_{\tau=1}^{t} -\varepsilon \cdot \sum_a p^\tau(a) c^\tau(a)$$

$$= \ln n - \varepsilon \sum_{\tau=1}^{t} \sum_a p^\tau(a) c^\tau(a)$$

& $\text{OPT}(-\varepsilon - \varepsilon^2) \leq \ln n - \varepsilon\, \text{ALG}$

$\varepsilon(\text{ALG} - \text{OPT}) \leq \ln n + \varepsilon^2\, \text{OPT}$

$\Rightarrow \text{Regret}_T = \frac{1}{T}(\text{ALG} - \text{OPT}) \leq \frac{1}{T}\left(\varepsilon\,\text{OPT} + \frac{\ln n}{\varepsilon}\right)$

choose ε: $\frac{\ln n}{\varepsilon} = \varepsilon T \Rightarrow \varepsilon = \sqrt{\frac{\ln n}{T}}$

then $\text{Regret}_T \leq \frac{1}{T}\left(\sqrt{T \ln n} + \sqrt{T \ln n}\right)$

$$= \sqrt{\frac{\ln n}{T}} \to 0 \text{ as } T \text{ grows}$$

hence, this MW algo has no regret.

(convergence: regret $\leq \varepsilon$ in $\frac{4 \ln n}{\varepsilon^2}$ steps)

Note:

(i) requires T known in advance

(ii) if only $c^t(a^t)$ known, this is the multi-armed bandit problem.

Playing No-Regret Dynamics leads to CCE

Consider a 2-player game, where each player uses no-regret dynamics to pick a mixed strategy at each time step σ_1^t, σ_2^t. Let $\sigma^t = (\sigma_1^t, \sigma_2^t)$.

Define $\hat{\sigma}$ as follows:

- pick $t \in [T]$ u.a.r.
- pick a pure strategy acc. to the prob. distribution $\sigma_1^t \times \sigma_2^t$

Lemma: If after T steps both players have at most ε regret, then $\hat{\sigma}$ is an ε-CCE

Defn.: $\sigma \in \Delta_{|S|}$ is an ε-CCE if $\forall i, \forall s_i' \in S_i$,

$$\mathbb{E}_{s\sim\sigma}[c_i(s)] \leq \mathbb{E}_{s\sim\sigma,\ s=(s_1,s_2)}[c_i(s_i', s_{-i})] + \varepsilon$$

Proof: For P1, $\text{Regret}_T \leq \varepsilon$

$$\Rightarrow s_1' \in S_1, \ \frac{1}{T}\sum_{i=1}^{T}\sum_{s\in S} c_1(s_1,s_2)\,\sigma_1^t(s_1)\,\sigma_2^t(s_2) - \frac{1}{T}\sum_{i=1}^{T}\sum_{s_2\in S_2} c_1(s_1', s_2)\,\sigma_2^t(s_2) \leq \varepsilon$$

$$\Rightarrow \forall s_1' \in S_1, \ \mathbb{E}_{s\sim\hat{\sigma}}[c_1(s)] - \mathbb{E}_{s\sim\hat{\sigma}}[c_1(s_1', s_2)] \leq \varepsilon \quad \blacksquare$$

Correlated Equilibria:

Consider the following game:

	L	R
U	2 \ 2	−2 \ −2
M	1 \ 1	−1 \ −1
D	−1 \ −1	1 \ 1

Here $\sigma(M,L) = 1/2$, $\sigma(D,R) = \frac{1}{2}$ is a CCE.

(verify)

However if correlating device tells P1 to play M, P1 knows that P2 will play L.

Hence it could improve its payoff by playing U (and continue to play D when told to do so)

(but this deviation is not allowed, hence CCE)

Defn: Distribution $\sigma \in \Delta_{|S|}$ is a Correlated Equilibrium if $\forall i, \forall s_i, \forall s_i'$,

$$\mathbb{E}_{s\sim\sigma}[c_i(s) \mid s_i] \leq \mathbb{E}_{s\sim\sigma}[c_i(s_i', s_{-i}) \mid s_i]$$

Claim σ is a CE if for every $\delta : [n] \to [n]$,

$$\mathbb{E}_{s\sim\sigma}[c_i(s)] \leq \mathbb{E}_{s\sim\sigma}[c_i(\delta(s_i), s_{-i})]$$

Claim: Every MNE is a CE