2/15/22:

Saw that if all players use no-regret learning algorithms, time-averaged distribution converges to a CCE.

Need to give a NR algorithm!

Easy no-regret algorithm: Multiplicative weights

(Randomized Weighted Majority, Hedge)

Assumptions:

- Experts setting: after t, find out $c^t(a)$ $\forall a$, not just action we chose
- Know T at beginning
- Adversary is oblivious

All can be removed. See book.

- Initialize $w^1(a) = 1$ $\forall a \in A$
- for $t = 1, 2, \dots, T$:
  - Use distribution $p^t = \frac{w^t}{\sum_{a \in A} w^t(a)}$, i.e.,

    choose each action a' with probability $\frac{w^t(a')}{\sum_{a \in A} w^t(a)}$

    (interpret weights as probabilities)
  - Given $c^t(a)$ $\forall a \in A$, update weights:

$$w^{t+1}(a) = w^t(a) \cdot (1-\varepsilon)^{c^t(a)} \quad \forall a \in A$$

Q: How should we set ε?

ε small: almost uniform, more "exploration"

ε big: focus on actions that have done well in past, more "exploitation"

Analysis:

Since adversary oblivious, before we start there is already some best action in hindsight:

$$a^* = \arg\min_{a \in A} \sum_{t=1}^{T} c^t(a)$$

$$OPT = \sum_{t=1}^{T} c^t(a^*)$$

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

$$\Rightarrow \Gamma^T \geq w^T(a^*) = w^1(a^*) \cdot \prod_{t=1}^{T} (1-\varepsilon)^{c^t(a^*)}$$

$$= \underline{1} \cdot (1-\varepsilon)^{\sum_{t=1}^{T} c^t(a^*)}$$

$$= (1-\varepsilon)^{OPT}$$

Expected cost of algorithm at time t:

$$v^t = \sum_{a \in A} \frac{w^t(a)}{T^t} \cdot c^t(a)$$

(want to bound $\sum_{t=1}^{T} v^t$ in terms of OPT)

Think about how T^t changes from t to $t+1$:

$$T^{t+1} = \sum_{a \in A} w^{t+1}(a) = \sum_{a \in A} w^t(a) \cdot (1-\varepsilon)^{c^t(a)}$$

$$\leq \sum_{a \in A} w^t(a) \cdot (1 - \varepsilon \cdot c^t(a))$$ ($(1-\varepsilon)^x \leq 1 - \varepsilon x \quad \forall \varepsilon \in (0, \frac{1}{2}], x \in [0,1]$)

$$= \sum_{a \in A} w^t(a) - \varepsilon \sum_{a \in A} w^t(a) c^t(a)$$

$$= T^t - \varepsilon T^t v^t$$ (def of T^t, v^t)

$$= T^t (1 - \varepsilon v^t)$$

Combine with previous bound:

$$(1-\varepsilon)^{OPT} \leq T^T \leq T^1 \prod_{t=1}^{T} (1 - \varepsilon v^t)$$

$$= n \prod_{t=1}^{T} (1 - \varepsilon v^t)$$ ($n = |A|$)

ln of both sides:

$$OPT \cdot \ln(1-\varepsilon) \leq \ln n + \sum_{t=1}^{T} \ln(1-\varepsilon v^t)$$

Claim: $-x-x^2 \leq \ln(1-x) \leq -x$ for $0 \leq x \leq \frac{1}{2}$

Pf: Taylor expansion

$$\Rightarrow OPT(-\varepsilon-\varepsilon^2) \leq \ln n + \sum_{t=1}^{T}(-\varepsilon v^t)$$

$$\Rightarrow \sum_{t=1}^{T} \varepsilon v^t \leq OPT(\varepsilon+\varepsilon^2) + \ln n$$

$$\Rightarrow \sum_{t=1}^{T} v^t \leq OPT(1+\varepsilon) + \frac{1}{\varepsilon}\ln n$$

$$\leq OPT + \varepsilon T + \frac{1}{\varepsilon}\ln n \qquad (OPT \leq T)$$

Finally set $\varepsilon = \sqrt{\frac{\ln n}{T}}$ (assumes we know T)

$$\Rightarrow \text{expected regret} = \frac{1}{T}\sum_{t=1}^{T} v^t - \frac{1}{T}\cdot OPT$$

$$\leq \frac{1}{T}\left(OPT + \varepsilon T + \frac{1}{\varepsilon}\ln n - OPT\right)$$

$$= \frac{1}{T}\left(\varepsilon T + \frac{1}{\varepsilon}\ln n\right)$$

$$= \frac{1}{T}\left(\sqrt{T\ln n} + \sqrt{T\ln n}\right) = 2\sqrt{\frac{\ln n}{T}}$$

$\Rightarrow$ no-regret!

So no-regret algorithms exist, and if players use them then we converge to a CCE

Q: What about correlated equilibria?

Rest of today and tomorrow: "stronger" notion of no-regret (no swap-regret) gives "stronger" equilibria (correlated equilibria)!

Def: A distribution σ over S is a correlated equilibrium if

$$\underset{s \sim \sigma}{E}\left[C_i(s) \mid s_i\right] \leq \underset{s \sim \sigma}{E}\left[C_i(s_{-i}, s_i') \mid s_i\right]$$

$$\forall i \in [k], \ \forall s_i, s_i' \in S_i$$

Interpretation: trusted third party draws $s \sim \sigma$, tells s_i to player i.
Then player i does not want to switch to s_i'

Rewrite in terms of "switching" rather than conditioning

Thm: σ is a correlated equilibrium if and only if

$$\underset{s\sim\sigma}{E}[C_i(s)] \le \underset{s\sim\sigma}{E}[C_i(s_{-i}, \delta(s_i))] \quad \forall i \in [k],$$
$$\forall \delta : S_i \to S_i$$

switching function

Pf:

only if: Sps σ correlated equilibrium.

Let $i \in [k]$, let $\delta : S_i \to S_i$

$$\underset{s\sim\sigma}{E}[C_i(s)] = \sum_{a\in S_i}\left(\underset{s\sim\sigma}{Pr}[s_i = a] \cdot \underset{s\sim\sigma}{E}[C_i(s) \mid s_i = a]\right)$$

$$\le \sum_{a\in S_i}\left(\underset{s\sim\sigma}{Pr}[s_i = a] \cdot \underset{s\sim\sigma}{E}[C_i(s_{-i}, \delta(a)) \mid s_i = a]\right) \quad \text{(def of (E))}$$

$$= \underset{s\sim\sigma}{E}[C_i(s_{-i}, \delta(s_i))]$$

if: Sps $\underset{s\sim\sigma}{E}[C_i(s)] \le \underset{s\sim\sigma}{E}[C_i(s_{-i}, \delta(s_i))]$

$\forall i \in [k]$, $\forall \delta : S_i \to S_i$

Let $i \in [k]$, $a, b \in S_i$. Want to show

$$\mathop{\mathbb{E}}_{s\sim\sigma}[C_i(s) \mid s_i = a] \le \mathop{\mathbb{E}}_{s\sim\sigma}[C_i(s_{-i}, b) \mid s_i = a]$$

Let $\delta(x) = \begin{cases} b & \text{if } x = a \\ x & \text{otherwise} \end{cases}$

$$\Rightarrow \mathop{\mathbb{E}}_{s\sim\sigma}[C_i(s)] \le \mathop{\mathbb{E}}_{s\sim\sigma}[C_i(s_{-i}, \delta(s_i))]$$

$$\Rightarrow \sum_{x \in S_i} \left(\Pr_{s\sim\sigma}[s_i = x] \cdot \mathop{\mathbb{E}}_{s\sim\sigma}[C_i(s) \mid s_i = x] \right)$$

$$\le \sum_{x \in S_i} \left(\Pr_{s\sim\sigma}[s_i = x] \cdot \mathop{\mathbb{E}}_{s\sim\sigma}[C_i(s_{-i}, \delta(s_i)) \mid s_i = x] \right)$$

$$\Rightarrow \cancel{\Pr_{s\sim\sigma}[s_i = a]} \cdot \mathop{\mathbb{E}}_{s\sim\sigma}[C_i(s) \mid s_i = a] \le \cancel{\Pr_{s\sim\sigma}[s_i = a]} \cdot \mathop{\mathbb{E}}_{s\sim\sigma}[C_i(s_{-i}, b) \mid s_i = a]$$

$$\Rightarrow \mathop{\mathbb{E}}_{s\sim\sigma}[C_i(s) \mid s_i = a] \le \mathop{\mathbb{E}}_{s\sim\sigma}[C_i(s_{-i}, b) \mid s_i = a]$$

$\checkmark$

So from now on, will use switching definition:

<u>Def</u>: σ is a correlated equilibrium if for all $i \in [k]$ and for all $\delta : S_i \to S_i$,

$$\mathop{E}_{s \sim \sigma}\left[C_i(s) \right] \leq \mathop{E}_{s \sim \sigma}\left[C_i(s_{-i}, \delta(s_i)) \right]$$