Online non-Bayesian Decision-Making

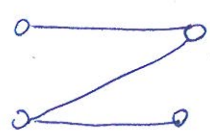
- Basic setup
  - Input set $\mathcal{I}_{[1,T]} = \prod_{t=1}^{T} \mathcal{I}_t$
  - Output / Action set $\mathcal{A}_{[1,T]} = \prod_{t=1}^{T} A_t$
  - Reward / Cost fn $C_\theta : \mathcal{I}_{[1,T]} \times \mathcal{A}_{[1,T]} \rightarrow \mathbb{R}$

- Online algorithm - Sequence of fns $F_t : \mathcal{I}_{[1,t]} \times \mathcal{A}_{[1,t-1]} \rightarrow A_t$

- Adversary - Sequence of fns $G_t : \mathcal{I}_{[1,t]} \times \mathcal{A}_{[1,t]} \rightarrow \mathcal{I}_{t+1}$
  (adaptive adversary)
  - Easier adversary. G_t is independent of history (oblivious adv) (i.e., constant / fixed element of $\mathcal{I}_{[1,T]}$)

- Can view this as a zero-sum game between algo and adversary. As a result, randomization may be needed

- Eg - Online bipartite matching.



Zero-Sum Games

rows/columns ≡ strategies

- Real valued $m \times n$ matrix $A \equiv$ Payoff matrix of row player

A_{ij} = Reward of row player if row i and col j are played

$-A_{ij}$ = Reward of col player " " " "

- Strategy profile ≡ distn x over rows, y over cols

(zero-sum) - Reward of row player $= x^T A y$

Thm (VonNeumann Minimax Thm) - for every 2-player zero-sum game A

$$\max_x \left(\min_y x^T A y \right) = \min_y \left(\max_x x^T A y \right)$$

row player chooses first = col player chooses first

Pf - Suppose row player fixes x

Claim - column player has a deterministic strategy

$$\max_x \left(\min_y x^T A y \right) = \max_x \left(\min_{j=1}^{n} (x^T A) \cdot e_j \right)$$

$$= \max_x \left(\min_{j=1}^{n} \sum_{i=1}^{m} A_{ij} x_i \right)$$

- Now we can write this as an LP

- LP for ~~col~~ row player (if she goes first) ~~for given strategy x~~

$$\max \quad v$$
$$\text{s.t.} \quad v - \sum_{i=1}^{m} A_{ij} x_i \leq 0 \quad \forall\, j \in \{1, 2, \dots, n\}$$
$$\sum_{i=1}^{m} x_i = 1$$
$$x_i \geq 0 \quad \forall\, i \in [n]$$

- Thus we have $v^* = \max_x \left(\min_y x^T A y \right)$

- Similarly for col player, going first, we have

$$\min \quad w$$
$$\text{st} \quad w - \sum_{j=1}^{n} A_{ij} y_j \geq 0 \quad \forall\, i \in [m]$$
$$\sum y_j = 1$$
$$y_j \geq 0 \quad \forall\, j \in [n]$$

and $w^* = \min_y \left(\max_x x^T A y \right)$

- These are dual LPs $\Rightarrow$ $v^* = w^*$

Yao's Lemma

We can now use the minimax thm to get bounds on any online algorithm's performance

- recall cost fn $C : \mathcal{I}_{[1,T]} \times \mathcal{A}_{[1,T]} \to \mathbb{R}$
- Let $\Delta_{\mathcal{I}}, \Delta_{\mathcal{A}}$ denote distrib on inputs/outputs

Lemma -

value of game

$$\max_{D \in \Delta_{\mathcal{I}}} \min_{a \in \mathcal{A}} \mathbb{E}\left[C(D, a)\right] = \max_{D \in \Delta_{\mathcal{I}}} \min_{A \in \Delta_{\mathcal{A}}} \mathbb{E}\left[C(D,A)\right] = V$$

$$= \min_{A \in \Delta_{\mathcal{A}}} \max_{D \in \Delta_{\mathcal{I}}} \mathbb{E}\left[C(D, A)\right] = \min_{A \in \Delta_{\mathcal{A}}} \max_{i \in \mathcal{I}} \mathbb{E}\left[C(i,A)\right]$$

moreover $$\min_{A \in \Delta_{\mathcal{A}}} \max_{i \in \mathcal{I}} \mathbb{E}\left[C(i,A)\right] \geq \min_{a \in \mathcal{A}} \mathbb{E}\left[C(D,a)\right]$$

for any $D \in \Delta_{\mathcal{I}}$