Lecture 7: von-Neumann min max. Proof and an application.

Theorem:- Let A be an $m \times n$ matrix. Then

$$\min_{\vec{r} \in \Delta_m} \max_{\vec{c} \in \Delta_n} \vec{r}^T A \vec{c} = \max_{\vec{c} \in \Delta_n} \min_{\vec{r} \in \Delta_m} \vec{r}^T A \vec{c}$$

(Equivalently

$$\min_{\vec{r} \in \Delta_m} \max_{1 \le j \le n} \vec{r}^T A \vec{e}_j = \max_{\vec{c} \in \Delta_n} \min_{1 \le i \le m} \vec{e}_i^T A \vec{c} \quad (*)$$

where $\{\vec{e}_i\}$ are the standard unit vectors)

where Δ_k is the set of probability distributions over a finite set of size k.

Interpretation:- $\vec{r}$: a 'mixed strategy' for the row player (who wants to minimize the payoff)

$\vec{c}$: a 'mixed strategy' for the column player (who wants to maximize the payoff).

$\vec{r}^T A \vec{c}$: Expected payoff when row player uses $\vec{r}$ and the column player uses $\vec{c}$.

Proof:- We will prove $(*)$.

Claim:- LHS of $(*)$ is equivalent to.

$$\min z$$
$$(\vec{r}^T A)_j = \sum_{i=1}^{m} r_i A_{ij} \le z \quad \forall\ 1 \le j \le n \qquad (LP\ 1)$$
$$\sum_{i=1}^{m} r_i = 1$$
$$r_i \ge 0.$$

Suppose r_* is the optimal r for the LHS of $*$. Then $r = r_*$ and $z = \max_{1 \le j \le n} (r_*^T A)_j$ is a feasible solution to (LP1) with the same objective value.

$$OPT(\text{LHS of } *) \geq OPT(\text{LP1}).$$

Now let (r_{LP}, z_{LP}) the optimal solution for LP1. Then $z_{LP} = \max_{1 \le j \le n} (r_{LP}^T A)_j$, so that we get.

$$OPT(\text{LHS of } *) \leq z_{LP} = OPT(\text{LP1}).$$

$$\text{Thus } OPT(\text{LP1}) = OPT(\text{LHS of } *)$$

$$\underline{\qquad} \times \underline{\qquad} \times \underline{\qquad}$$

We take duals by converting to standard form (Usually you would instead 'jump' to the dual directly, with some practice: the point is that the dual is supposed to represent all necessarily implied linear combinations of the constraints).

- To convert inequality constraints to equality variables we introduce slack variables.
- To remove 'unrestricted' variables, we express them as differences of non-negative variables.

  - Replace z by $v_1 - v_2$, $v_1, v_2 \geq 0$.
  - Introduce a slack variable $s_j \geq 0$ for each

$1 \leq j \leq m$. We write

Dual variables:

$$\min \quad v_1 - v_2$$

$-c_j$ $$\sum_{1 \leq i \leq m} r_i A_{ij} - v_1 + v_2 + s_j = 0 \qquad 1 \leq j \leq n.$$

u $$\sum_{1 \leq i \leq n} r_i = 1 \qquad (LP1')$$

$$r_i \geq 0 \qquad 1 \leq i \leq m$$
$$s_j \geq 0 \qquad 1 \leq j \leq m$$
$$v_1 \geq 0$$
$$v_2 \geq 0 .$$

Claim:- (LP1') and (LP) are 'equivalent': they have the same optimum and a feasible solution of one immediately gives a feasible solution of the other with the same objective value.

Dual is.

$$\max \quad u$$
$$\sum_{j=1}^{m} A_{ij} c_j \geq u. \qquad 1 \leq i \leq m.$$
$$\sum_{j=1}^{m} c_j = 1$$
$$c_j \geq 0 \qquad 1 \leq j \leq n$$

which is the same as

$$\max_{\vec{c} \in \Delta_n} \min_{1 \leq i \leq m} c_i^T A c \quad !!$$

(Check that in this case primal is feasible and bounded so that $p^* = d^*$ holds). This completes

the proof.

Application:- Boolean circuit: A directed acyclic graph with a unique sink node, and nodes labelled by either a boolean operation (AND, OR, NOT) or as an 'input node'. It's operation: Assign a bit value to each input node. All non-input nodes compute their 'label' function of the values of their parents, recursively, and assign this to be their value. The output of the circuit is the value computed by the sink node. [Size of the circuit = no of edges in the circuit.]

A circuit is an explicitly given function that on input $x \in \{0,1\}^n$ (n is the number of input nodes) computes a bit x.

Ckt. Complexity theory:- 'Find' functions $f : \{0,1\}^n \to \{0,1\}$ which 'cannot' be computed by 'small' circuits.

One possible figure of merit: Let $\mathcal{C}$ be a set of circuits (e.g. all circuits of size $\leq n^3$) on n inputs bits. A function $f : \{0,1\}^n \to \{0,1\}$ said to be δ-hard for $\mathcal{C}$ if

$$\Pr_{x \sim_{un} \{0,1\}^n} [C(x) = f(x)] < 1 - \delta. \quad \forall C \in \mathcal{C}.$$

(Impagliazzo '95) Can write this in terms of 'advantage':

$$\mathrm{Adv}(C, x) := \begin{cases} 1 & \text{if } C(x) = f(x) \\ -1 & \text{if } C(x) \neq f(x). \end{cases}$$

$$Adv(C, S) := \frac{1}{|S|} \sum_{x \in S} Adv(C, x).$$

when $S \subseteq \{0,1\}^n$.

(A function f is δ-hard for $\mathcal{C}$ if

$$\forall C \in \mathcal{C} \quad Adv(C, \{0,1\}^n) < 1 - 2\delta$$)

Impagliazzo's question :- If f is δ-hard for $\mathcal{C}'$, is f really hard ($Adv \approx 0$) (on smaller sets $\theta(\delta \cdot 2^n)$) for some slightly smaller class $\mathcal{C}$?

Nisan's method :- $\mathcal{D}$ be the set of probability distribution over subsets of size exactly $\delta \cdot 2^n$ of $\{0,1\}^n$.

Suppose that

$$\textcircled{*} \quad \min_{D \in \mathcal{D}} \max_{C \in \mathcal{C}} \mathbb{E}_{S \sim D} [A(C,S)] \geq \varepsilon.$$

Set-player :- Has to play a set $S \subseteq \{0,1\}^n$ of size $\delta \cdot 2^n$

Circuit-player :- Has to play a ckt $C \in \mathcal{C}$. (want to maximize payoff)

Payoff : $A(C,S)$

So $\textcircled{*}$ is exactly the LHS of the min-max theorem. and thus we get that there is a distribution Γ over $\mathcal{C}$ s.t. $\forall S$ of size $\delta \cdot 2^n$

$$\mathbb{E}_{C \sim \Gamma} [A(C,S)] \geq \varepsilon. \quad \text{— (1)}$$

(from the RHS of min-max and $\textcircled{*}$)

(Weak duality not sufficient for this)

Now let $\alpha \in (0,1)$, and define

$$S' = \{x \mid \mathop{E}_{C\sim\Gamma}[A(C,x)] < \alpha\varepsilon\}.$$

From ①, this means $|S'| < \delta \cdot 2^n$. (as $\alpha < 1$).

In fact, let G be a subset of size $\delta \cdot 2^n$ s.t. $G \supseteq S'$.

Then. $\varepsilon \leq \mathop{E}_{C\sim\Gamma}[A(C,G)]$

$$= \frac{1}{\delta \cdot 2^n}\left[\sum_{x\in S'} \mathop{E}_{C\sim\Gamma}[A(C,x)] + \sum_{x\notin S'} \mathop{E}_{C\sim\Gamma}[A(C,x)]\right]$$

$$\leq \frac{1}{\delta\cdot 2^n}\left[\alpha\varepsilon \cdot |S'| + (\delta\cdot 2^n - |S|)\right]$$

$$= 1 - \frac{(1-\alpha\varepsilon)|S'|}{\delta\cdot 2^n}.$$

$$|S'| \leq \delta\cdot 2^n \cdot \left(\frac{1-\varepsilon}{1-\alpha\varepsilon}\right)$$

$$\leq \frac{\delta\cdot 2^n}{2-\alpha}. \quad [\text{Assume } \varepsilon \leq 1/2].$$

$$\forall x \notin S', \quad \mathop{E}_{C\sim\Gamma}[A(C,x)] \geq \alpha\varepsilon$$

Fix $x \notin S'$. We want a circuit (possibly larger than those in C, which has advantage close to 1 on x.

We know that $\Pr_{C\sim\Gamma}[C(x) = f(x)] \geq \frac{1}{2} + \frac{\alpha\varepsilon}{2}$.

Sample t ckts $C_1, \ldots, C_t$ <u>independently</u> from Γ. Define $H = \text{Maj}(C_1, \ldots, C_t)$. (Assume t is odd.)

(Distribution of $H =: \mathcal{H}$)

$$\Pr_{H\sim\mathcal{H}}\left[H(x) \neq f(x)\right] \leq \Pr\left[\sum_{i=1}^{t} X_i \leq t/2\right]$$

where X_i are independent Bernoulli random variables with $E[X_i] \geq \frac{1}{2} + \frac{\alpha\varepsilon}{2}$ $\forall i$.

$$\leq \exp\left(-\frac{t\alpha^2\varepsilon^2}{4}\right)$$

(Chernoff bound).

Thus for any ~~fixed~~ $x \notin S'$.

$$\Pr_{H\sim\mathcal{H}}\left[H(x) = f(x)\right] \geq 1 - \exp\left(-\frac{\alpha^2\varepsilon^2 t}{8}\right).$$

What is the expected no. of x's on which $H(x) = f(x)$, when $H \sim \mathcal{H}$? This is

$$\sum_{x\notin S'} E_{H\sim\mathcal{H}}\left[I[H(x) = f(x)]\right] + \sum_{x\in S'} E_{H\sim\mathcal{H}}\left[I[H(x) = f(x)]\right]$$

$$\geq \sum_{x\notin S'} \Pr_{H\sim\mathcal{H}}\left[H(x) = f(x)\right]$$

Thus,

$$E_{H\sim\mathcal{H}}\left[\frac{1}{2^n}\left|\{x \mid H(x) = f(x)\}\right|\right] \geq \left(1 - \exp\left(-\frac{t\alpha^2\varepsilon^2}{8}\right)\right)\frac{(2^n - |S'|)}{2^n}$$

$$\geq \left(1 - \exp\left(-\frac{t\alpha^2\varepsilon^2}{8}\right)\right)\left(1 - \frac{\delta}{2-\alpha}\right)$$

$$\geq \left(1 - \exp\left(-\frac{t\alpha^2\varepsilon^2}{8}\right) - \frac{\delta}{2-\alpha}\right).$$

Let $\alpha = 0.1$, $t = \frac{c}{\alpha^2\varepsilon^2}\log\left(\frac{1}{\delta}\right)$ with c chosen so that $\exp\left(-\frac{t\alpha^2\varepsilon^2}{8}\right) \leq 0.05\,\delta$.

So, with this choice of α and t,

$$\mathop{E}_{H\sim\mathcal{H}}\left[\Pr_{x\sim \mathrm{Unif}(\{0,1\}^n)}[H(x)=f(x)]\right] \geq 1 - \underbrace{0.8\,\delta}$$

could be made closer to 0.5 by choosing α and t appropriately

So we have shown (so far) [Impagliazzo]

If for all distributions D on sets of size $\delta \cdot 2^n$ G has non-negligible (i.e. at least ε (probability of success $\approx 1/2$)) advantage, then there is a distribution $\mathcal{H}$ on 'slightly larger' circuits s.t.

$$\mathop{E}_{H\sim\mathcal{H}}\left[\Pr_{x\sim \mathrm{Unif}(\{0,1\}^n)}[H(x)=f(x)]\right] \geq 1 - 0.8\,\delta.$$

these are majorities of $O\left(\frac{1}{\varepsilon^2}\log\left(\frac{1}{\delta}\right)\right)$ circuits from G

['Probabilistic method': If $E[X] = \delta$, then we must have $\Pr[X \geq \delta] > 0$.]

and therefore, there exists H, of the form

$$\mathrm{Maj}(C_1, \ldots C_t), \quad C_i \in G, \text{ i.e.}$$

where $t = \Theta\left(\frac{1}{\varepsilon^2}\log\left(\frac{1}{\delta}\right)\right)$ is as above

s.t.

$$\Pr_{x \in_{\mathrm{unif}} \{0,1\}^n}[H(x)=f(x)] \geq 1 - 0.8\,\delta.$$

(i.e. f is not δ-hard for any class G' of circuits that includes all t-sized Majorities of circuits in G)

Note:- (δ-hard fns. 'exist' for reasonable ckt. classes, as can be proved by the probabilistic method, but finding explicit ones is an important open problem in several interesting settings.).

(Explicit functions are known for bdd depth families.)