CSE 632: Analysis of Algorithms II **Fall 2017**

Lecture 22 (11/10/2017): 0-sum game and Nash-equilibrium

Lecturer: Shi Li *Scribe: Minwei Ye*

22.1 Zero-sum game

A 2-player zero-sum(matrix) game is defined by a matrix $M \in \mathbb{R}^{m\times n}$, called the payoff matrix. There are two players with opposing interests: the row layer wants to minimize the payoff and the column player wants to maximize the payoff. At each step of the play, the row player choose a row $i \in [n]$ and the column player choose a column $j \in [m]$, the entry $M(i,j)$ will be the payoff of this game and the row player pays that many dollars to the column player. How to play such a game?

Consider the classic rock-scissor-paper game. The payoff matrix looks like this:

$$\begin{bmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{bmatrix}$$

In this game, if any of the players has a deterministic(pure) strategy and it is revealed to the other player, the other player can always gain advantage.
But if they are allowed with mixed strategy then the knowledge of such strategy might not help the other player. For example, if the row player play each of the three rows with equal probability 1/3, no matter what strategy the column player choose, it will always have expected payoff less or equal than 0(actually in this special case all possible column strategies result in expected payoff 0). If we exchange the roles of the column and the row player the same logic applies. The value 0 is called the *value* of the game. This result generalizes to any finite matrix game and it can be formally stated as below:
A mixed strategy for row player is a distribution D over rows. A mixed strategy for column player is a distribution P over columns.
Define $M(D,P) = \mathbb{E}_{i\sim D}\,\mathbb{E}_{j\sim P}\,M(i,j)$, then

Theorem 22.1 (Von Neumann (1928))**.**

$$\inf_D \max_{j\in[m]} M(D,j) = \sup_P \min_{i\in[n]} M(i,P)$$

$\lambda^* = \inf_D \max_{j\in[m]} M(D,j) = \sup_P \min_{i\in[n]} M(i,P)$ is called the *value* of the game. It is also referred as Nash equilibrium in other context.

Our goal is to approximate this value within an additive error. More specifically,
find $\hat{D}$ s.t. $\max_j M(\hat{D},j) \le \lambda^* + \epsilon$ and
find $\hat{P}$ s.t. $\min_i M(i,\hat{P}) \ge \lambda^* - \epsilon$
This can be done using the expert prediction game.

22.2 Approximate the value of the zero-sum game

Let's constraint the game matrix M to be in $[-1,1]^{m\times n}$.
Let the pure strategies for row player correspond to experts, and pure strategies for column player correspond to events. The algorithm is the following: To see why this algorithm return an approx-

---

for $t = 1 \cdots T$: **do**
  choose D^t using multiplicative weight update rule
  Adversary picks j^t that maximize $M(D^t, j)$ and penalty for expert i is $M(i, j^t)$
return $\hat{\lambda} = \frac{1}{T}\sum_{i=1}^{T} M(D^t, j^t)$

---

imation of λ^*, first note that for any distribution D

$$M(D, j^t) \geq \min_{i\in[n]} M(i, j^t)$$

since a distribution is just a weighted average of pure strategies. Then, by the multiplicative weight update theorem, when $T \geq 4\ln n/\epsilon^2$ and $\epsilon < 1/2$

$$\hat{\lambda} = \frac{1}{T}\sum_{t=1}^{T} M(D^t, j^t) \leq \min_i \frac{1}{T}\sum_{i\in[n]} M(i, j^t) + \epsilon \leq \frac{1}{T}\sum_{t=1}^{T} M(D, j^t) + \epsilon$$

for every D, especially, when $D = D^*$, the distribution that achieves the equality in theorem 22.1, we have

$$\hat{\lambda} = \frac{1}{T}\sum_{t=1}^{T} M(D^t, j^t) \leq \lambda^* + \epsilon$$

To find the distribution $\hat{D}$ that achieves $\hat{\lambda}$, we can simply record all the $M(D^t, j^t)$ and report the distribution $\hat{D}$ with minimum value. Then

$$\max_{j\in[m]} M(\hat{D}, j) = M(\hat{D}, j^t) \leq \lambda^* + \epsilon$$

We can even extract $\hat{P}$ directly from this algorithm. Let $\hat{P}$ be such that the probability of taking the j-th column is $\frac{1}{T}|\{t : j^t = j\}|$. Then $\hat{P}$ satisfies $\min_i M(i, \hat{P}) \geq \lambda^* - \epsilon$. The proof is left as an exercise.