Lecture 20: Zero-Sum Game

Lecturer: Zongchen Chen

1 Two-Player Zero-Sum Game

Consider a two-player zero-sum game. Two players A and B play the following game:

- Player A picks an action $i \in [m] = \{1, \ldots, m\}$;
- Player B picks an action $j \in [n] = \{1, \ldots, n\}$;
- Then, player A pays an amount $M(i, j)$ to player B, called the *payoff*.

The goal of player A is to minimize the payoff, and the goal of player B is to maximize the payoff. The game is zero-sum since the loss (respectively, gain) of A always equals the gain (respectively, loss) of B. In particular, there is no win-win or lose-lose situation.

Table 1: Payoff matrix for rock paper scissors

	B plays rock	B plays paper	B plays scissors
A plays rock	0	1	-1
A plays paper	-1	0	1
A plays scissors	1	-1	0

A *pure strategy* is to choose a fixed/deterministic action. A *mixed strategy* is to choose a random action from some distribution over the action set. Thus, pure strategies are special cases of mixed strategies. Note that the payoff for a mixed strategy is random, and we shall always consider the expectation of the payoff.

Let $M \in \mathbb{R}^{m \times n}$ denote the payoff matrix, where we assume $M(i, j) \in [-1, 1]$ for all $i \in [m]$ and $j \in [n]$. Let

$$\Delta_m = \left\{ p \in \mathbb{R}^m_{\geq 0} : \sum_{i=1}^{m} p_i = 1 \right\}$$

be the set of probability distributions over the action set $[m]$ for player A. Every mixed strategy for A corresponds to a distribution $p \in \Delta_m$. In particular, a pure strategy which plays the action $i \in [m]$ corresponds to the basis vector $e_i \in \Delta_m$ whose i'th entry is 1 and all other entries 0. Similarly, let

$$\Delta_n = \left\{ q \in \mathbb{R}^n_{\geq 0} : \sum_{j=1}^{n} q_j = 1 \right\}$$

be the set of distributions over $[n]$ which represent mixed strategies for player B. A pure strategy for B which plays the action $j \in [n]$ corresponds to the basis vector $f_j \in \Delta_n$.

If A plays action i (i.e., pure strategy e_i) and B plays action j (i.e., pure strategy f_j), then the payoff is

$$M(i, j) = e_i^{\intercal} M f_j.$$

More generally, if A plays a mixed strategy $p \in \Delta_m$ and B plays a mixed strategy $q \in \Delta_n$, then the (expected) payoff is

$$\mathbb{E}_{i \sim p} \mathbb{E}_{j \sim q} M(i, j) = p^{\intercal} M q.$$

Thus, the goal of player A is to choose $p \in \Delta_m$ to minimize $p^\intercal Mq$, and the goal of player B is to choose $q \in \Delta_n$ to maximize $p^\intercal Mq$.

2 Minimax Theorem

In a fair game, player A chooses a strategy $p \in \Delta_m$ without knowing player B's strategy $q \in \Delta_n$, and the same for B. It is helpful to also consider the unfair versions.

- Game favoring B: Player A has to choose a strategy $p \in \Delta_m$ first and present it to player B (only the strategy p but not the actual action chosen from p). Then, player B chooses a strategy $q \in \Delta_n$ according to p. In this setting, play B can choose an optimal strategy q with respect to p which maximizes the payoff; namely, for a given p, play B solves

$$\max_{q \in \Delta_n} p^\intercal Mq = \max_{j \in [n]} p^\intercal M f_j.$$

Note, it suffices for B to play a pure strategy f_j given any p. Back to A, player A needs to choose an optimal strategy $p \in \Delta_m$ such that the payoff, after B chooses an optimal q accordingly, is minimized. Thus, the optimal payoff for both players is given by

$$\min_{p \in \Delta_m} \max_{q \in \Delta_n} p^\intercal Mq = \min_{p \in \Delta_m} \max_{j \in [n]} p^\intercal M f_j.$$

- Game favoring A: Player B has to choose $q \in \Delta_n$ first, and then player A chooses $p \in \Delta_m$ given q. Similarly as above, the optimal payoff in this setting is

$$\max_{q \in \Delta_n} \min_{p \in \Delta_m} p^\intercal Mq = \max_{q \in \Delta_n} \min_{i \in [m]} e_i^\intercal Mq.$$

From these observations, the following lemma is immediate. It is also straightforward to verify the lemma mathematically.

Lemma 1. *It holds*

$$\min_{p \in \Delta_m} \max_{q \in \Delta_n} p^\intercal Mq \geq \max_{q \in \Delta_n} \min_{p \in \Delta_m} p^\intercal Mq.$$

The following theorem is due to von Neumann, which establishes a surprising, non-trivial, and elegant equality.

Theorem 2 (Von Neumann's Minimax Theorem)**.** *It holds*

$$\min_{p \in \Delta_m} \max_{q \in \Delta_n} p^\intercal Mq = \max_{q \in \Delta_n} \min_{p \in \Delta_m} p^\intercal Mq. \tag{1}$$

The value in Eq. (1), denoted as λ^*, is the optimal payoff for both players, and is referred to as the value of the game.

3 Approximating the Value of the Game

Given the payoff matrix $M \in \mathbb{R}^{m \times n}$ where $M(i,j) \in [-1,1]$ for all $i \in [m]$ and $j \in [n]$, our goal is to estimate the value λ^* of the game, and find strategies of both players which approximate λ^*.

Suppose the two players A and B play the game for multiple rounds. We focus on player A and design an algorithm to find a strategy for A in each round. Our plan is to use the prediction with expert advice framework and apply the multiplicative weight update algorithm. We regard every pure strategy e_i where

$i \in [m]$ as an expert. Then, player A using a mixed strategy $p \in \Delta_m$ corresponds to the algorithm following a random expert chosen from p. We still need to come up with the penalty for each expert (i.e., pure strategy). Suppose player A plays a strategy $p \in \Delta_m$, and let $j' \in \arg\max_{j\in[n]} p^\intercal M f_j$ be the optimal (pure) strategy for player B with respect to p. Then, the penalty for expert i (i.e., pure strategy e_i) is set to be

$$e_i^\intercal M f_{j'} = M(i, j'),$$

which is the payoff when A plays action i and B plays action j'. Under this setup, we obtain the following multiplicative weight update algorithm.

Algorithm 1 Multiplicative weight update algorithm for the value of the game

$w^1 \leftarrow (1, \ldots, 1)$
for $t = 1$ to T **do**
  $Z^t \leftarrow \sum_{i=1}^m w_i^t$
  for $i = 1$ to m **do**
    $p_i^t \leftarrow \frac{w_i^t}{Z^t}$
  end for
  $j^t \leftarrow \arg\max_{j\in[n]} (p^t)^\intercal M f_j$
  for $i = 1$ to m **do**
    $w_i^{t+1} \leftarrow w_i^t e^{-\varepsilon M(i,j^t)}$
  end for
end for
$\hat{p} \leftarrow \frac{1}{T}\sum_{t=1}^T p^t$
$\hat{q} \leftarrow \frac{1}{T}\sum_{t=1}^T f_{j^t}$
$\hat{\lambda} \leftarrow \hat{p}^\intercal M \hat{q}$
return $\hat{p}$, $\hat{q}$, $\hat{\lambda}$

Algorithm 1 can be understood as follows. Players A and B play the unfair game favoring B for T rounds. Player A (the algorithm) maintains a weight vector w^t for all actions, and in every round plays a strategy by normalizing the weight vector w^t into a distribution p^t. Player B then chooses an optimal action j^t against p^t. Finally, player A updates the weight vector via the multiplicative update rule, using the payoff $M(i, j^t)$ for each action $i \in [m]$.

We need the following theorem for the multiplicative weight update algorithm.

Theorem 3. *Suppose $T \geq \log m$. If we set $\varepsilon = \sqrt{\frac{\log m}{T}}$, then it holds*

$$\frac{1}{T}\sum_{t=1}^T (p^t)^\intercal M f_{j^t} \leq \min_{i\in[m]} \left\{\frac{1}{T}\sum_{t=1}^T e_i^\intercal M f_{j^t}\right\} + 2\sqrt{\frac{\log m}{T}}.$$

Theorem 4. *If $T = \lceil \frac{4\log m}{\delta^2} \rceil$, then*

1. *$\hat{p}$ is approximately optimal: $\max_{q\in\Delta_n} \hat{p}^\intercal M q \leq \lambda^* + \delta$;*
2. *$\hat{q}$ is approximately optimal: $\min_{p\in\Delta_m} p^\intercal M \hat{q} \geq \lambda^* - \delta$;*
3. *$\hat{\lambda}$ approximates the value of the game: $\lambda^* - \delta \leq \hat{\lambda} \leq \lambda^* + \delta$.*

Proof. Since $T \geq \frac{4\log m}{\delta^2}$, we deduce from Theorem 3 that

$$\frac{1}{T}\sum_{t=1}^T (p^t)^\intercal M f_{j^t} \leq \min_{i\in[m]} \left\{\frac{1}{T}\sum_{t=1}^T e_i^\intercal M f_{j^t}\right\} + \delta.$$

For the right-hand side, we have

$$\begin{aligned}
\min_{i\in[m]}\left\{\frac{1}{T}\sum_{t=1}^{T} e_i^\intercal M f_{j^t}\right\} &= \min_{i\in[m]}\left\{e_i^\intercal M\left(\frac{1}{T}\sum_{t=1}^{T} f_{j^t}\right)\right\}\\
&= \min_{i\in[m]} e_i^\intercal M\hat{q}\\
&= \min_{p\in\Delta_m} p^\intercal M\hat{q}\\
&\le \max_{q\in\Delta_n}\min_{p\in\Delta_m} p^\intercal M q = \lambda^*.
\end{aligned}$$

For the left-hand side, we have

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T}(p^t)^\intercal M f_{j^t} &\ge \max_{j\in[n]}\left\{\frac{1}{T}\sum_{t=1}^{T}(p^t)^\intercal M f_j\right\} && (j^t \text{ is optimal with respect to } p^t)\\
&= \max_{j\in[n]}\left\{\left(\frac{1}{T}\sum_{t=1}^{T} p^t\right)^\intercal M f_j\right\}\\
&= \max_{j\in[n]}\hat{p}^\intercal M f_j\\
&= \max_{q\in\Delta_n}\hat{p}^\intercal M q\\
&\ge \min_{p\in\Delta_m}\max_{q\in\Delta_n} p^\intercal M q = \lambda^*.
\end{aligned}$$

Therefore, we have established that

$$\lambda^* \le \max_{q\in\Delta_n}\hat{p}^\intercal M q \le \min_{p\in\Delta_m} p^\intercal M\hat{q} + \delta \le \lambda^* + \delta,$$

which shows items 1 and 2. For item 3, observe that

$$\lambda^* - \delta \le \min_{p\in\Delta_m} p^\intercal M\hat{q} \le \hat{p}^\intercal M\hat{q} \le \max_{q\in\Delta_n}\hat{p}^\intercal M q \le \lambda^* + \delta,$$

as wanted. $\square$