

# 1 Introduction

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 competing interests: the row layer wants to minimize the payoff and the column player wants to maximize the payoff. In the game, the row player chooses a row  $i \in [n]$  and the column player choose a column  $j \in [m]$ . Then  $M(i, j)$  will be the payoff of this game. Imagine that the row player needs to pay  $M(i, j)$  dollars to the column player; thus the row player wants to minimize  $M(i, j)$  and the column player wants to maximize  $M(i, j)$ .

Consider the payoff matrix for the rock-paper-scissor game:

	rock	paper	scissors
rock	0	1	-1
paper	-1	0	1
scissors	1	-1	0

	rock	paper	scissors
rock	0	1	-1
paper	-1	0	1
scissors	1	-1	0

Table 1: The table on the left-side shows the gains of both players in a rock-paper-scissors game. Since the sum of the gains of the two players is always 0, we can simply use the gain for the column player to define the payoff matrix, as given by the table on the right-side.

## 1.1 Pure and Mixed Strategies

We call a row  $i$  (a column  $j$ , resp.) a pure strategy for the row (column, resp.) player. If the players are only allowed to use pure strategies, it is important to know which player plays first. For example, in the rock-paper-scissors game, the second player will always win the game if he knows the strategy of the first player. To make the problem interesting, we consider *mixed strategies* for both players:

**Definition 1.** A mixed strategy for the row player (resp., the column player) is defined as a distribution  $D$  (resp.,  $P$ ) over the rows (resp., columns) of the payoff matrix. If the strategies used by the two players are  $D$  and  $P$  respectively, then the payoff of game is defined as

$$M(D, P) := \mathbb{E}_{i \sim D, j \sim P} M(i, j).$$

Notice that a pure strategy can also be viewed as a special case of a mixed strategy.

Consider the rock-paper-scissors game where the row player plays the following mixed strategy  $D$ : rock with probability 1/2, paper with probability 1/4 and scissors with probability 1/4. Then we have  $M(D, \text{rock}) = 1/4 - 1/4 = 0$ ,  $M(D, \text{paper}) = 1/2 - 1/4 = 1/4$ , and  $M(D, \text{scissors}) = -1/2 + 1/4 = -1/4$ . So given that the row player plays  $D$ , the best pure strategy for the column player is paper. It is easy to see that this is also the best strategy (pure or mixed) for the column player.

However, the row player can do better: He can use mixed strategy  $D^*$  where the probabilities for rock, paper and scissors are all 1/3. Then even the column player knows the strategy, the payoff of the game is at most (indeed, exactly) 0. On the other hand, if the column player plays first, he can also use the uniform distribution  $P^*$  and then no matter what the row player plays, the payoff of the game is at least 0. Consider the situation where two players fix their strategies  $D^*$  and  $P^*$ . Then even if one player knows the strategy of the other player, he still does not have an incentive to change his own strategy. The min-max theorem, which is stated in the next section, that this is always the case for 0-sum games.

## 2 Min-max Theorem

The min-max theorem of 0-sum games is defined as follows.

**Theorem 2.**  $\inf_D \max_j M(D, j) = \sup_P \min_i M(i, P)$ .

We define the value of the game to be  $\lambda^* = \inf_D \max_j M(D, j) = \sup_P \min_i M(i, P)$ . Let  $D^* = \arg \inf_D \max_j M(D, j)$  and  $P^* = \arg \sup_P \min_i M(i, P)$  to be the optimum strategies for both players. Then notice that  $\lambda^* = M(D^*, P^*)$ .

**Example** Consider the following payoff matrix.

	a	b
A	30	-10
B	-10	20
C	20	-20

Then the optimum strategies for both players are defined as follows:

$$D^* : \begin{cases} A & 0 \\ B & \frac{4}{7} \\ C & \frac{3}{7} \end{cases} \quad P^* : \begin{cases} a & \frac{4}{7} \\ b & \frac{3}{7} \end{cases}$$

First, we fix the strategy  $D^*$  for the row player. Then  $M(D^*, a) = \frac{4}{7} \times (-10) + \frac{3}{7} \times 20 = \frac{20}{7}$  and  $M(D^*, b) = \frac{4}{7} \times 20 + \frac{3}{7} \times (-20) = \frac{20}{7}$ . So the payoff of the game is  $\frac{20}{7}$ .

Then, we fix the strategy  $P^*$  for the column player. We have  $M(A, P^*) = \frac{4}{7} \times 30 + \frac{3}{7} \times -10 = \frac{90}{7}$ ,  $M(B, P^*) = \frac{4}{7} \times (-10) + \frac{3}{7} \times 20 = \frac{20}{7}$  and  $M(C, P^*) = \frac{4}{7} \times 20 + \frac{3}{7} \times -20 = \frac{20}{7}$ . So, the row player will choose either  $B$  or  $C$  and the payoff of the game is  $\frac{20}{7}$ . The value of the game is  $\lambda^* = \frac{20}{7}$ .

## 3 Multiplicative weight update algorithm to find optimum strategies

In this section, we show how to use the multiplicative weight update method to compute the value of the game, as well as to give the optimum strategies for both players. We assume the set pure strategies for the row player is  $[n]$  and that for the column player is  $[m]$ . So, the size of  $M$  is  $n \times m$ . By scaling, we assume every entry in  $M$  is in  $[-1, 1]$ .

The algorithm is as follows:

---

### Algorithm 1

---

Multiplicative weight update for 0-sum games

---

- 1: let  $w_1 = w_2 = w_3 = \dots = w_n = 1$
- 2: **for**  $t \leftarrow 1$  to  $T$ , where  $T = \lceil \frac{\ln n}{\epsilon^2} \rceil$  **do**
- 3:      $D^t = \frac{(w_1, w_2, \dots, w_n)}{w_1 + w_2 + \dots + w_n}$ .
- 4:     let  $j^t$  be the  $j$  maximize  $M(D^t, j)$
- 5:     define the penalty of every  $i$  to be  $M(i, j_t)$ , and thus we update update  $w_i \leftarrow w_i \cdot e^{-\epsilon M(i, j_t)}$  according to the MWU rule

---

$$\hat{\lambda} = \frac{1}{T} \sum_{T=1}^T M(D^t, j^T) \tag{1}$$

Since  $T \geq \frac{\ln(n)}{\epsilon^2}$ , we have

$$\frac{1}{T} \sum_{t=1}^T M(D^t, j^t) \leq \min_i \frac{1}{T} \sum_{t=1}^T M(i, j^t) + 2\epsilon \tag{2}$$

Let  $\hat{t}$  be the  $t \in [T]$  with the minimum  $M(D^t, j^t)$ . Let  $\hat{D} = D^{\hat{t}}$ . Let  $\hat{P}$  be the uniform distribution over the multi-set  $\{j^1, j_2, \dots, j^T\}$ . Then we have

$$\max_j M(\hat{D}, j) = M(\hat{D}, j^{\hat{t}}) \leq \frac{1}{T} \sum_{t=1}^T M(D^t, j^t) \leq \min_i \frac{1}{T} \sum_{t=1}^T M(i, j^t) + 2\epsilon = \min_i M(i, \hat{P}) + 2\epsilon.$$

Notice that  $\max_j M(\hat{D}, j) \geq \lambda^*$  and  $\min_i M(i, \hat{P}) \leq \lambda^*$ . Thus the above inequality implies  $\max_j M(\hat{D}, j) \leq \lambda^* + 2\epsilon$  and  $\min_i M(i, \hat{P}) \leq \lambda^* - 2\epsilon$ . This implies that up to the additive factor of  $2\epsilon$ ,  $\hat{D}$  and  $\hat{P}$  are optimum strategies for the row and column players respectively. We can use  $\hat{\lambda} = \max_j M(\hat{D}, j) = M(D^{\hat{t}}, j^{\hat{t}})$  to approximate the value  $\lambda^*$  of the game.