CS6781 - Theoretical Foundations of Machine Learning

Lecture 20: Boosting, Online Learning, and Games

April 21, 2020

Lecturer: Nika Haghtalab *Readings: N/A*
Scribe: Michela Meister

1 Learning and Games

Two friends, Alice and Bob, are playing Rock, Paper, Scissors. Recall that rock beats scissors, scissors beats paper, and paper beats rock. The following matrices describe the loss of Alice and Bob when Alice is the row player and Bob is the column player. For example, when Alice chooses paper and Bob chooses rock, Alice wins and Bob loses, so Alice has a loss of -1 and Bob has a loss of 1. If Alice and Bob both choose paper, they tie and each have a loss of 0.

Table 1: Alice's Loss

	Rock	Paper	Scissors
Rock	0	1	-1
Paper	-1	0	1
Scissors	1	-1	0

Table 2: Bob's Loss

	Rock	Paper	Scissors
Rock	0	-1	1
Paper	1	0	-1
Scissors	-1	1	0

These loss matrices show that Rock, Paper, Scissors is a *zero-sum game*, that is, the sum of Alice's loss and Bob's loss is always 0. This means that minimizing Alice's loss is equivalent to maximizing Bob's loss, and vice versa. The same can be said if their losses always added to any other fixed constant c.

Let L be the loss matrix for the row player. Then if the row player chooses row r and the column player chooses row c, the row player receives loss $L(r, c)$ and the column player receives loss $-L(r, c)$. More generally, say the row player chooses a distribution P over rows and the column player chooses a distribution Q over columns. Then the expected loss for the row player is $P^\top LQ$ and the expected loss for the column player is $-P^\top LQ$.

There are a few different settings for two player games. In *deterministic game play*, the row player chooses a row r, the column player sees r, and then the column player chooses a column c. In *simultaneous game play* the row player picks a distribution P over rows and the column player picks a distribution Q over columns. Then the row player samples $r \sim P$ and the column player samples $c \sim Q$. In *randomized sequential game play* the row player announces their distribution P over rows, the column player sees P, and then the column player picks a distribution Q over columns. Then the row player samples $r \sim P$ and the column player samples $c \sim Q$.

In deterministic game play, the second player (say the column player) always has the advantage, because they know the first players (say row player's) action before choosing their own. What is the relationship between simultaneous game play and randomized sequential game play? Does the second player still have a strict advantage in these settings? The Min-Max Theorem says that these two settings are equivalent – the player that goes second does not have a strict advantage.

Theorem 1.1 (Min-Max Theorem for Zero Sum Games)**.** *Let $L \in \mathbb{R}^{m \times n}$ be the loss matrix for a finite, zero-sum game. Let $\mathcal{P}$ be the set of all distributions over m elements and let $\mathcal{Q}$ be the set of all distributions over n elements. Then*

$$\min_{P \in \mathcal{P}} \max_{Q \in \mathcal{Q}} L(P, Q) = \max_{Q \in \mathcal{Q}} \min_{P \in \mathcal{P}} L(P, Q).$$

Proof. The left hand side of the equation is the loss incurred by the row player when the row player goes first, and the right hand side is the loss incurred by the row player when the row player goes second. Taking the maximum of a minimum is always smaller than minimizing a maximum, so

$$\min_P \max_Q L(P, Q) \geq \max_Q \min_P L(P, Q).$$

(Another way to understand the above inequality is that going second should only help the row player minimze their loss.) Thus it is sufficient to show that

$$\min_P \max_Q L(P, Q) \leq \max_Q \min_P L(P, Q).$$

We will prove this via online learning, by analyzing the performance of a no-regret algorithm for the row player against a best-response adversary who takes the role of the column player. The row player and column player play a game at each time step. At time t, given the history of distributions $Q_1, \ldots, Q_{t-1}$ that the column player has played so far, the row player runs a no-regret algorithm to choose P_t. Then the column player chooses $Q_t = \arg\max_Q L(P_t, Q)$. Let $\bar{P} = (1/T) \sum_{t=1}^T P_t$ and let $\bar{Q} = (1/T) \sum_{t=1}^T Q_t$. Then we have

$$\begin{aligned}
\min_P \max_Q L(P, Q) &= \min_P \max_Q P^\top L Q && (1) \\
&\leq \max_Q \bar{P}^\top L Q && (2) \\
&= \max_Q \frac{1}{T} \cdot \sum_{t=1}^T P_t^\top L Q && (3)
\end{aligned}$$

$$\le \frac{1}{T} \cdot \sum_{t=1}^{T} \max_Q P_t^\top LQ \tag{4}$$

$$= \frac{1}{T} \cdot \sum_{t=1}^{T} P_t^\top LQ_t \tag{5}$$

$$\le \min_P \frac{1}{T} \cdot \sum_{t=1}^{T} P^\top LQ_t + \frac{\text{Regret}}{T} \tag{6}$$

$$= \min_P P^\top L\bar{Q} + \frac{\text{Regret}}{T} \tag{7}$$

$$\le \max_Q \min_P P^\top LQ + \frac{\text{Regret}}{T} \tag{8}$$

$$= \max_Q \min_P L(P, Q) + \frac{\text{Regret}}{T}. \tag{9}$$

In (1) we re-write the expected loss as a matrix-vector multiplication. The loss with the minimizing P is at most the loss with the average ($\bar{P}$), which implies (2). In (3) we plug in the defintion for $\bar{P}$. The loss when Q is maximized over all time steps is at most the loss when Q is maximized at each timestep, giving (4). In (5) we plug in the definition of Q_t. In (6) the total loss accumulated by time T, $\sum_{t=1}^{T} P_t^\top LQ_t$, is bounded by the sum of the loss due to the best fixed choice of P, which is $\min_P \sum_{t=1}^{T} P^\top LQ_t$, plus the regret of the row player's algorithm. We plug in the defintion of $\bar{Q}$ in (7). In (8) the loss due to the average $\bar{Q}$ at most the loss due to the maximizing choice of Q. In (9) we substitute $L(P, Q) = P^\top LQ$.

Since the row player is playing a no-regret algorithm, as $T \to \infty$, $\frac{\text{Regret}}{T} \to 0$, in which case $\min_P \max_Q L(P, Q) \le \max_Q \min_P L(P, Q)$. □

2 Boosting and Minmax Theorem

As we saw in Section 1, online learning provides a constructive proof to the Min-Max theorem. In this section, we dig deeper at the connection between minmax theorem and other learning paradigms and see how the Min-Max Theorem helps explain why boosting is even plausible. Let $S = \{(x_1, y_1), \ldots, (x_n, y_n)\}$ be a set of n labeled data points and let $\mathcal{H}$ be a set of m hypotheses produced by a weak learner trained on different distributions over S. Consider the loss matrix M where the rows are hypotheses $h_1, \ldots, h_m$ and the columns are datapoints $x_1, \ldots, x_n$, and $M(h_j, x_i) = 1$ if $h_j(x_i) \neq y_i$. That is, the row player has a loss of 1 for each mislabeled point. The row player and column player play the following game: The row player chooses a distribution P over hypotheses and the column player chooses a distribution Q over datapoints. The row player wants to minimize the loss $P^\top MQ$, while the column player wants to maximize $P^\top MQ$.

Let (P^*, Q^*) be the Min-Max pair of strategies and let

$$v_1 := \max_{i \in [n]} \sum_{j=1}^{m} P^*(j) M(h_j, x_i).$$

That is, given a set of hypotheses distributed as P^*, the weighted error of these hypotheses on any datapoint x_i is at most v_1. Furthermore, let v_2 be such that

$$v_2 := \min_{j\in[m]} \sum_{i=1}^{n} Q^*(i)M(h_j, x_i) = \min_{h_j} \mathrm{err}_{Q^*}(h_j)$$

This means that, given data points weighted by Q^*, the error on Q^* of the best hypothesis is v_2. By the Min-Max theorem $v_1 = v_2$. Since $h_1, \ldots, h_m$ are returned by a weak learner on distributions over S, there is a h_j such that $\mathrm{err}(h_j) \leq 1/2 - \gamma$. This implies that $v_2 \leq 1/2 - \gamma$, and therefore $v_1 \leq 1/2 - \gamma$. So for any data point x_i, given a set of hypotheses distributed as P^*, the majority vote of the hypotheses is correct, which shows that in principle boosting a weak learner to a strong learner is possible.

3 Boosting and No-Regret Algorithms

Indeed, we can formulate boosting and AdaBoost's gaurantees as no-regret interactions between a learner and an adversary. Consider another loss matrix M' where the rows are datapoints $x_1, \ldots, x_n$ and the columns are hypotheses $h_1, \ldots, h_m$, and $M'(x_i, h_j) = 1$ if $h_j(x_i) = y_i$. That is, we have swapped the rows and columns of the matrix in Section 2 and changed the sign of losses, so that our matrix demonstrates row player's losses. Consider an setting where the row player and column player play a sequence of games where at time t the row player uses a no-regret algorithm to pick a distribution P_t that received a low accuracy predictions from historical hypothesis the column player has playes. The column player picks a hypothesis h_t that has accuracy at least $1/2 + \gamma$. Then we have

$$\frac{1}{2} + \gamma \leq M'(P_t, h_t) \qquad \text{(Accuracy of } h_t \text{ on } P_t\text{)}$$

So,

$$\begin{aligned} \frac{1}{2} + \gamma &\leq \frac{1}{T} \cdot \sum_{t=1}^{T} M'(P_t, h_t) \qquad \text{(Average accuracy)} \\ &\leq \min_{x_i} \frac{1}{T} \cdot \sum_{t=1}^{T} M'(x_i, h_t) + \frac{\text{Regret}}{T}. \end{aligned}$$

So for all x_i, we have that

$$\frac{1}{T} \cdot \sum_{t=1}^{T} M'(x_i, h_t) \geq \frac{1}{2} + \gamma - \frac{\text{Regret}}{T}.$$

Note that when $\frac{\text{Regret}}{T} < \gamma$, $(1/T)\sum_{t=1}^{T} M'(x, h_t) > 1/2$, so the majority vote for everyx_i is correct. So we need to set

$$\frac{\text{Regret}}{T} = \frac{\sqrt{T \ln(m)}}{T} \leq \frac{\gamma}{2}.$$

So we get $T = \Theta\left((1/\gamma^2)\ln(m)\right)$, which matches the Adaboost guarantee.