Machine Learning Theory (CSC 482A/581B) - Lecture 19

Nishant Mehta

1 Game theory

Online learning is a game between Learner and Nature; simultaneously, optimal strategies for two-player games can be found via online learning. In this lecture, we will see how von Neumann's celebrated minimax theorem can be proved using an algorithm called Multiplicative Weights, which is a special case of the Hedge algorithm when the loss vectors take a certain parametric form.

A *two-player game* is defined by a matrix $M \in [0,1]^{m \times n}$. The game is played between a row player and a column player, where:

- the row players selects a row i of M;
- the column player selects a column j of M.

The loss of the row player is equal to M_{ij}, and the row player's goal is to minimize their loss. On the other hand, the column player seeks to maximize the loss of the row player. If we define the loss of the column player to be the negation of the loss of the row player, then the column player seeks to minimize its own loss, and the game is a *zero-sum game*: the losses of the players sum to zero.

In the above, each player was restricted to a *pure strategy*, where they deterministically select a single row (or column) of M. Relaxing this, each player could instead play a *mixed strategy*; a mixed strategy randomizes over the rows (or columns) of M. In this more general setting, we talk about expected losses, where the expectation is taken with respect to the random indices i and j.

For any $n \in \mathbb{N}$, let Δ_n denote the $(n-1)$-simplex $\Delta^{n-1} := \left\{ \alpha \in \mathbb{R}^n_+ : \sum_{j=1}^n \alpha_j = 1 \right\}$. We use the former notation so that the ambient dimension of Δ_n (as embedded into $\mathbb{R}^n$) is clear. If the row player plays $p \in \Delta_m$ and the column player plays $q \in \Delta_n$, then the loss of the row player is $p^\top M q$.

Up until now, we have not discussed the sequence of play. Consider the case when the row player moves first, selecting some strategy $p \in \Delta_m$. For any such strategy, if the column player acts optimally, the row player's loss is

$$\max_{q \in \Delta_n} p^\top M q.$$

The optimal strategy for the row player is thus one which obtains loss

$$\min_{p \in \Delta_m} \max_{q \in \Delta_n} p^\top M q.$$

If instead the column player moves first and both players act optimally, the row player's loss is

$$\max_{q \in \Delta_n} \min_{p \in \Delta_m} p^\top M q.$$

What is the effect of the order of play? Intuitively, the row player wields more power by moving second, as they can respond to the strategy of the column player. Indeed, this is easily verified:

Define $\bar{q} \in \Delta_n$ and $\underline{p} \in \Delta_m$ such that

$$\min_{p \in \Delta_m} p^\top M \bar{q} = \max_{q \in \Delta_n} \min_{p \in \Delta_m} p^\top M q \qquad \max_{q \in \Delta_n} \underline{p}^\top M q = \min_{p \in \Delta_m} \max_{q \in \Delta_n} p^\top M q.$$

Then

$$\max_{q \in \Delta_n} \min_{p \in \Delta_m} p^\top M q = \min_{p \in \Delta_m} p^\top M \bar{q} \leq \underline{p}^\top M \bar{q} \leq \max_{q \in \Delta_n} \underline{p}^\top M q = \min_{p \in \Delta_m} \max_{q \in \Delta_n} p^\top M q. \tag{1}$$

However, when both players act optimally, does the row player truly suffer greater loss when they move first rather than second? The answer is no, and this is the content of von Neumann's minimax theorem[1]:

Theorem 1 (Von Neumann's Minimax Theorem)**.**

$$\min_{p \in \Delta_n} \max_{q \in \Delta_n} p^\top M q = \max_{q \in \Delta_n} \min_{p \in \Delta_m} p^\top M q.$$

When a minimax theorem like Theorem 1 holds, we call the common value on either side of the equality the *value* of the game. One additional observation which we often will employ is that, for sequential play as above, the second player loses no power by restricting to pure strategies. Therefore, rewriting the minimax theorem, we also have

$$\min_{p \in \Delta_n} \max_{j \in [n]} p^\top M e_j = \max_{q \in \Delta_n} \min_{i \in [m]} e_i^\top M q.$$

2 Repeated games

Suppose that the game matrix M is unknown, and that the row player is a learning algorithm that wishes to minimize its loss. If the game is played only once, this clearly is a hopeless situation since we learn something about M only once the game is over. However, if the game is played repeatedly over T rounds, the learning algorithm can hope to learn something about M and obtain low cumulative loss:

$$\sum_{t=1}^{T} p_t^\top M q_t.$$

However, just as in decision-theoretic online learning (DTOL), it is not possible in general to obtain low cumulative loss; instead, Learner might hope to achieve low regret, which we now define as

$$\sum_{t=1}^{T} p_t^\top M q_t - \min_{p \in \Delta_m} \sum_{t=1}^{T} p^\top M q_t,$$

the amount by which the cumulative loss of the learning algorithm exceeeds the cumulative loss of the best constant strategy p in hindsight of $q_1, \ldots, q_T$.

The above two-player zero-sum game setting bears strong similarity to the DTOL setting. Indeed, if in the latter setting we constrain the loss vectors to take the parametric form $\ell_t = M q_t$ (where M is fixed and q_t is chosen by Nature), then DTOL recovers the two-player zero-sum game.

[1] Here, we only present von Neumann's minimax theorem in a simplified form. The full version holds in the more general situation where $p^\top M q$ is replaced by $f(x, y)$ for a function $f : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ that is convex in x (for fixed y), concave in y (for fixed x), and for $\mathcal{X}$ and $\mathcal{Y}$ compact convex subsets of $\mathbb{R}^m$ and $\mathbb{R}^n$ respectively.

It therefore is fitting that we will study a low-regret algorithm called Multiplicative Weights which is essentially equivalent to Hedge.

Multiplicative Weights (MW) operates as follows. The strategy p_1 is set to some initial value; a sensible choice that we take is to set $p_1 = \frac{1}{m}\mathbf{1}$, the uniform distribution over $[m]$. In round $t+1$, MW plays strategy p_{t+1}, where p_{t+1} is defined via the update

$$p_{j,t+1} = \frac{p_{j,t} \exp\left(-\eta e_j^\top M q_t\right)}{Z_t} \qquad \text{for } j = 1, \ldots, m;$$

here, e_j is the j th standard basis vector for $\mathbb{R}^m$, and Z_t is the normalization term

$$Z_t := \sum_{i=1}^m p_{i,t} \exp\left(-\eta e_i^\top M q_t\right).$$

Because the two-player zero-sum game is a special case of DTOL, we can apply Theorem 1 from the last lecture to obtain the following regret guarantee for MW.

Theorem 2. *Let MW be run with learning rate $\eta = \sqrt{\frac{8 \log m}{T}}$. Then, for any sequence $q_1, \ldots, q_T \in \Delta_n$,*

$$\sum_{t=1}^T p_t^\top M q_t \leq \min_{i \in [m]} \sum_{t=1}^T e_i^\top M q_t + \sqrt{\frac{T \log m}{2}}.$$

Also, it is easy to verify that for any $q_1, \ldots, q_T \in \Delta_n$,

$$\min_{p \in \Delta_m} \sum_{t=1}^T p^\top M q_t = \min_{i \in [m]} \sum_{t=1}^T e_i^\top M q_t.$$

Using this fact and further dividing through by T to get a bound on the average (per round) regret, we have the following corollary.

Corollary 1. *Let MW be run with learning rate $\eta = \sqrt{\frac{8 \log m}{T}}$. Then, for sequences $q_1, \ldots, q_T \in \Delta_n$,*

$$\frac{1}{T} \sum_{t=1}^T p_t^\top M q_t \leq \min_{p \in \Delta_m} \frac{1}{T} \sum_{t=1}^T p^\top M q_t + \sqrt{\frac{\log m}{2T}}.$$

Recall from (1) that we already know that

$$\max_{q \in \Delta_n} \min_{p \in \Delta_m} p^\top M q \leq \min_{p \in \Delta_m} \max_{q \in \Delta_n} p^\top M q. \tag{2}$$

Using Corollary 1, we can also show the other direction,

$$\min_{p \in \Delta_m} \max_{q \in \Delta_n} p^\top M q \leq \max_{q \in \Delta_n} \min_{p \in \Delta_m} p^\top M q, \tag{3}$$

and thus we will have proved von Neumann's minimax theorem via MW.

Proof of Theorem 1. We already have established (2); it remains to establish (3).

Let $p_1, \ldots, p_T$ be the strategies played by MW against $q_1, \ldots, q_T$, where, for each $t \in [T]$, the column player selects strategy $q_t = \max_{q \in \Delta_n} p_t^\top M q_t$. Also, define the average strategies $\bar{p} = \frac{1}{T} \sum_{t=1}^T p_t$ and $\bar{q} = \frac{1}{T} \sum_{t=1}^T q_t$.

With this setup, it holds that

$$\begin{aligned}\min_{p\in\Delta_m}\max_{q\in\Delta_n} p^\top Mq &\le \max_{q\in\Delta_n} \bar{p}^\top Mq \\ &= \max_{q\in\Delta_n} \frac{1}{T}\sum_{t=1}^{T} p_t^\top Mq \\ &\le \frac{1}{T}\sum_{t=1}^{T}\max_{q\in\Delta_n} p_t^\top Mq \\ &= \frac{1}{T}\sum_{t=1}^{T} p_t^\top Mq_t,\end{aligned} \tag{4}$$

where the last equality holds by the definition of q_t. Now, from Corollary 1 and setting $\varepsilon_T = \sqrt{\frac{\log m}{2T}}$, the last line above is at most

$$\begin{aligned}\min_{p\in\Delta_m}\frac{1}{T}\sum_{t=1}^{T} p^\top Mq_t + \varepsilon_T &= \min_{p\in\Delta_m} p^\top M\bar{q} + \varepsilon_T \\ &\le \max_{q\in\Delta_n}\min_{p\in\Delta_m} p^\top Mq + \varepsilon_T.\end{aligned} \tag{5}$$

Finally, ε_T vanishes as we take $T \to \infty$, and so (3) does indeed hold. □

3 Approximate minimax and maximin optimal strategies

The above algorithmic proof of von Neumann's minimax theorem goes further than proving what was required. From the proof we can actually produce an approximately minimax strategy, i.e., a strategy $\bar{p}$ for which

$$v \le \max_{q\in\Delta_n} \bar{p}^T Mq \le v + \varepsilon,$$

as well as an approximately maximin strategy, i.e., a strategy $\bar{q}$ for which

$$v \ge \min_{p\in\Delta_m} p^\top M\bar{q} \ge v - \varepsilon,$$

where we recall that v is the value of the game. The first inequality of each of the above displays is trivial, since the value of the game v satisfies

$$v = \min_{p\in\Delta_n}\max_{q\in\Delta_n} p^\top Mq = \max_{q\in\Delta_n}\min_{p\in\Delta_m} p^\top Mq.$$

Let us verify that the second inequality in each display holds. Indeed, taking $\bar{p} = \frac{1}{T}\sum_{t=1}^{T} p_t$ and using the sequence of inequalities starting from the right-hand side of the first line of (4) until (5), we see that

$$\begin{aligned}\max_{q\in\Delta_n} \bar{p}^T Mq &\le \max_{q\in\Delta_n}\min_{p\in\Delta_m} p^\top Mq + \varepsilon_T \\ &= v + \varepsilon_T,\end{aligned}$$

where ε_T can be made as small as desired by increasing T.

Similarly, taking $\bar{q} = \frac{1}{T}\sum_{t=1}^{T} q_t$ and using the sequence of inequalities starting from the right-hand side of the first line of (5) backwards to the beginning of (4), it holds that

$$\begin{aligned}\min_{p \in \Delta_m} p^T M \bar{q} &\geq \min_{p \in \Delta_m} \max_{q \in \Delta_n} p^\top M q - \varepsilon_T \\ &= v - \varepsilon_T.\end{aligned}$$

Thus, $(\bar{p}, \bar{q})$ are ε_T-approximate solutions to the game defined by matrix M.

4 Connection to boosting

Finally, we see how boosting can be viewed as a two-player zero-sum game. In particular, we will study how this applies to a simplified version of AdaBoost whose final predictor is a simple majority rather than a weighted majority.

In this case, there are two players, Booster (the min player / row player) and Weak Learner (the max player / column player). In the game:

- Booster plays distributions over a finite set $\mathcal{X} = \{x_1, \ldots, x_n\}$; this corresponds to the training sample over which AdaBoost is run;
- Weak Learner plays hypotheses from a (possibly infinite) set of weak learning hypotheses $\mathcal{H}$.

Note that we could have allowed Weak Learner to play distributions over hypotheses, but this additional generality is not needed since the weak learning assumption only requires the weak learner to play a single hypothesis given some input.

Rather than using standard matrix notation, we now switch to the notation

$$M(x, h)$$

to indicate the loss of Booster (or reward of Weak Learner) when some pure strategy $x \in \mathcal{X}$ is played against some pure strategy $h \in \mathcal{H}$. Similarly, for a mixed strategy P that is a distribution over $\mathcal{X}$ and pure strategy $h \in \mathcal{H}$, we denote the loss of Booster as

$$M(P, h).$$

The remaining ingredient is to define the game matrix itself. For each $x \in \mathcal{X}$ and $h \in \mathcal{H}$, we set

$$M(x, h) = \mathbf{1}\left[h(x) = c(x)\right].$$

Thus, for a mixed strategy P and pure strategy h, it holds that

$$M(P, h) = \Pr_{X \sim P}\left(h(X) = c(X)\right).$$

With this specification of the game, the goal of Booster is to minimize classification accuracy while the goal of Weak Learner is to maximize classification accuracy.

We now derive a simplified version of AdaBoost by using MW specialized to this game. As usual, the initial distribution p_1 is the uniform distribution: $p_1(x_j) = \frac{1}{n}$ for all $x \in \mathcal{X}$. In round t, the update is[2]

$$p_{t+1}(x) = \frac{p_t(x) e^{-\eta\, \mathbf{1}[h_t(x) = c(x)]}}{Z_t} \qquad \text{for } x \in \mathcal{X}.$$

[2] Note that, unlike AdaBoost, the unnormalized weights of the misclassified examples are not increased; still, their unnormalized weight *is* increased relative to the correctly classified examples, and so the effect is the same as in AdaBoost.

Finally, we set the learning rate to $\eta = \sqrt{\frac{8 \log n}{T}}$, as suggested by Corollary 1. Note that unlike in AdaBoost, here the learning rate is constant over the rounds.

Now, we use the proof of Theorem 1 to show that our boosting algorithm outputs a simple majority classifier consistent with the correct concept c, when run for enough rounds. The last line of (4) together with the left-hand side of the first line of (5) implies that

$$\frac{1}{T}\sum_{t=1}^{T} M(P_t, h_t) \leq \min_P \frac{1}{T}\sum_{t=1}^{T} M(P, h_t) + \varepsilon_T = \min_{x \in \mathcal{X}} \frac{1}{T}\sum_{t=1}^{T} M(x, h_t) + \varepsilon_T, \tag{6}$$

where the equality holds because for fixed $h_1, \ldots, h_T$ the minimum over all distributions is obtained by a pure strategy. Now, from the weak learning assumption, for each t it holds that

$$\frac{1}{2} + \gamma \leq \Pr_{X \sim P_t}(h_t(X) = c(X)) = M(P_t, h_t);$$

Taken together with (6), it follows that for all $x \in \mathcal{X}$:

$$\frac{1}{2} + \gamma - \varepsilon_T \leq \frac{1}{T}\sum_{t=1}^{T} \mathbf{1}\left[h_t(x) = c(x)\right].$$

Here, $\varepsilon_T = \sqrt{\frac{\log n}{2T}}$, and so if we take $T > \frac{\log n}{2\gamma^2}$, we have

$$\frac{1}{T}\sum_{t=1}^{T} \mathbf{1}\left[h_t(x) = c(x)\right] > \frac{1}{2}.$$

Therefore, the simple majority classifier based on $h_1, \ldots, h_T$ classifies every example $x \in \mathcal{X}$ correctly.

5 Connection between the edge of the weak learner and margins

One implication of the minimax theorem is that

$$\frac{1}{2} + \gamma \leq \min_P \max_{h \in \mathcal{H}} \Pr_{X \sim P}(h(X) = c(X)) = \max_Q \min_{x \in \mathcal{X}} \Pr_{h \sim Q}(h(x) = c(x)),$$

where the first inequality is from the weak learning assumption. The left-hand side and right-hand side together imply that there exists a distribution Q over $\mathcal{H}$ for which, for all $x \in \mathcal{X}$,

$$\Pr_{h \sim Q}(h(X) = c(X)) \geq \frac{1}{2} + \gamma.$$

Therefore, for this distribution (call it Q^*), we have

$$\Pr_{h \sim Q^*}(h(X) = c(X)) - \Pr_{h \sim Q^*}(h(X) \neq c(X)) \geq 2\gamma.$$

From this we immediately conclude that the weak learnability assumption with edge γ implies the existence of a majority vote classifier that obtains margin at least 2γ.