Machine Learning Theory (CSC 431/531) - Lecture 18

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 minimax theorem can be proved using Hedge in the special case 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 its 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 $n \in \mathbb{N}$, let Δ_n denote the simplex over n outcomes, defined as $\Delta_n := \left\{ \alpha \in \mathbb{R}_+^n : \sum_{j=1}^n \alpha_j = 1 \right\}$. If the row player plays mixed strategy $p \in \Delta_m$ and the column player plays mixed strategy $q \in \Delta_n$, then the loss of the row player is $p^\top M q$.

We have not yet discussed the order in which the players select their strategies. 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 can only be better off if it moves second since a player that moves second can respond to the other player's strategy. 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)**.**

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

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

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

2 Repeated games

There are various proofs of von Neumann's minimax theorem that use what are known as fixed-point theorems. Instead, we will see an elementary proof based on the existence of a no-regret learning algorithm for a certain online learning game. On the face of it, that online learning can be used to prove von Neumann's minimax theorem might seem strange: the minimax theorem is a game with only one round, whereas online learning takes place over many rounds. To form the connection, let us consider the row player as a learning algorithm that wishes to minimize its cumulative loss in a repeated game. If the learning algorithm (row player) plays $p_1, \ldots, p_T$ and the column player plays $q_1, \ldots, q_T$, then the learning algorithm has cumulative loss

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

As we observed in online learning, the learning algorithm cannot guarantee that its cumulative loss is low in general; however, the learning algorithm can aim 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 exceeds the cumulative loss of the best constant strategy p in hindsight of $q_1, \ldots, q_T$.

[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.

The above two-player zero-sum game setting bears strong similarity to the decision-theoretic online learning (DTOL) setting. In fact, if in DTOL we constrain the loss vectors to take the parametric form $\ell_t = Mq_t$ (where M is fixed and q_t is chosen by Nature), then the loss in the DTOL game is the same as the row player's loss in the two-player zero-sum game. Therefore, the row player can achieve low regret by using Hedge.

Let us look closely at the plays made by Hedge in our two-player zero-sum game setting. The strategy p_1 is set to some initial value; let us take $p_1 = \frac{1}{m}\mathbf{1}$, the uniform distribution over $[m]$. In round t, Hedge plays strategy p_t, where p_t is defined via

$$p_{j,t} = \frac{\exp\left(-\eta \sum_{s=1}^{t-1} e_j^\top M q_s\right)}{\sum_{i=1}^{K} \exp\left(-\eta \sum_{s=1}^{t-1} e_i^\top M q_s\right)} \qquad \text{for } j = 1, \ldots, m; \tag{2}$$

here, e_j is the j^{th} standard basis vector in $\mathbb{R}^m$.

Because the two-player zero-sum game is a special case of DTOL, we can apply Theorem 3 from the last lecture to obtain the following regret guarantee for the above application of Hedge.

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

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

The equality in the above theorem uses the basic fact that for any vector $v \in \mathbb{R}^m$, it holds that $\min_{p \in \Delta_m} p^\top v = \min_{i \in [m]} e_i^\top v$.

Recall from (1) that we already know that

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

Using Theorem 2, we can also show the other direction,

$$\min_{p \in \Delta_m} \max_{q \in \Delta_n} p^\top M q \leq \max_{q \in \Delta_n} \min_{p \in \Delta_m} p^\top M q, \tag{4}$$

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

Proof of Theorem 1. We already have established (3); it remains to establish (4).

Let $p_1, \ldots, p_T$ be the strategies played by Hedge 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 $\bar{q} = \frac{1}{T}\sum_{t=1}^{T} q_t$.

Let $\varepsilon_T = \sqrt{\frac{\log m}{2T}}$. With this setup, it holds that

$$\begin{aligned}
\min_{p \in \Delta_m} \max_{q \in \Delta_n} p^\top M q &\le \max_{q \in \Delta_n} \bar{p}^\top M q && (5) \\
&= \max_{q \in \Delta_n} \frac{1}{T} \sum_{t=1}^{T} p_t^\top M q \\
&\le \frac{1}{T} \sum_{t=1}^{T} \max_{q \in \Delta_n} p_t^\top M q \\
&= \frac{1}{T} \sum_{t=1}^{T} p_t^\top M q_t, \\
&\le \min_{p \in \Delta_m} \frac{1}{T} \sum_{t=1}^{T} p^\top M q_t + \varepsilon_T \\
&= \min_{p \in \Delta_m} p^\top M \bar{q} + \varepsilon_T && (6) \\
&\le \max_{q \in \Delta_n} \min_{p \in \Delta_m} p^\top M q + \varepsilon_T. && (7)
\end{aligned}$$

where the second equality holds by the definition of q_t and the line thereafter is from Theorem 2.

Finally, ε_T vanishes as we take $T \to \infty$, and so (4) 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

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

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

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

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 (5) until (7), we see that

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

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 (6) backwards to the left-hand side of (5), it holds that

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

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