6.883: Online Methods in Machine Learning

Alexander Rakhlin

LECTURE 17

The setting of "linearized experts" in the previous lecture serves as a stepping stone to the next subject: online convex optimization. We will make strong connections to gradient descent methods studied in the first part of the course. Before doing that, however, we need to learn a powerful tool – the minimimax theorem.

In online learning problems encountered so far, we either solved the "min max" on each round exactly (as in Cover's result, in collaborative filtering, etc) or came up with a strategy to approximately solve it in the particular situation of experts (the exponential weights strategy). In many problems of interest, we might not be able to solve the min max exactly, and might not be able to dream up a strategy that will approximately solve it. The minimax theorem may come to the rescue in such situations.

1. MINIMAX BASICS

Let us start by introducing the idea of games and minimax values. We only focus on two-person *zero-sum* games, where the loss of Player I is the reward of Player II and vice-versa. Let $\mathcal{A}$ be the set of moves of Player I and $\mathcal{B}$ the set of available moves of Player II. Let $\boldsymbol{\ell}(a,b)$ be the loss of the first player when she chooses an action $a \in \mathcal{A}$ while the opponent chooses $b \in \mathcal{B}$. Here, $\boldsymbol{\ell} : \mathcal{A} \times \mathcal{B} \to \mathbb{R}$ is a known loss function.

As an example, take the penny-matching game, with $\mathcal{A} = \mathcal{B} = \{0,1\}$ and $\boldsymbol{\ell}(a,b) = \mathbf{1}\{a \neq b\}$. That is, Player I suffers a loss of 1 if her chosen bit is not equal to the one chosen by the opponent. This game has a clear symmetry. If the players make their moves simultaneously, nobody has an advantage, and 1/2 should be the reasonable "expected" loss of either of the players. But how does 1/2 surface? To which question is this an answer?

Let us consider the optimal moves that the players should choose. Since Player I is trying to minimize the loss (and the opponent – maximize), we can write down an expression

$$\min_{a\in\mathcal{A}} \max_{b\in\mathcal{B}} \boldsymbol{\ell}(a,b).$$

Unfortunately, this expression is not quite what we want, and it has to do with the fact that the inner maximization can be done as a function of a. In the penny-matching game, b can always be chosen opposite of a to ensure $\boldsymbol{\ell}(a,b) = 1$. If we instead write $\max_b \min_a$, Player I now has the advantage and guarantees zero loss. This does not match our intuition that the game is symmetric.

The "min max" breaks simultaneity of the two moves and instead puts them in the order of their occurrence. Is there a way to write down simultaneous moves as a min max? The answer is yes, but under some conditions.

Let Q be the set of distributions on $\mathcal{A}$ and P the set of distributions on $\mathcal{B}$, and let us write the expression

$$V^+ \triangleq \min_{q\in Q} \max_{b\in\mathcal{B}} \mathbb{E}_{a\sim q} \boldsymbol{\ell}(a,b). \tag{1}$$

That is, the first player chooses her *mixed strategy* q and tells the second player "hey, you can choose b based on this q". It seems that we have gained nothing: the second player still has the advantage of responding to the first player move, and simultaneity is not preserved. It turns out that, under some conditions (which definitely hold for finite sets $\mathcal{A}$ and $\mathcal{B}$), the above minimax expression is equal to the maximin variant

$$V^- \triangleq \max_{p \in P} \min_{a \in \mathcal{A}} \mathbb{E}_{b \sim p} \boldsymbol{\ell}(a,b). \tag{2}$$

Moreover, both of the expressions are equal to

$$V = \mathbb{E}_{a \sim q^*, b \sim p^*} \boldsymbol{\ell}(a,b)$$

for some (q^*, p^*). It becomes irrelevant who makes the first move, as the minimax value is equal to the maximin value. In such a case, we say that the game *has a value* $V = V^+ = V^-$. We will be interested in conditions under which this happens. The "minimax swap", as we call it, will hold for infinite sets of moves $\mathcal{A}$ and $\mathcal{B}$.

Interestingly, no-regret online learning methods can be used to prove the minimax theorem and its generalizations. In turn, we will use the minimax theorem to come up with online learning methods. There is no circular logic here, as we will see.

1.1 A few useful facts

Let us mention a few facts which are easy to verify. First, as already mentioned, the inner maximization (minimization) is achieved at a pure strategy if it is followed by an expectation:

$$\min_{q \in Q} \max_{p \in P} \mathbb{E}_{a \sim q, b \sim p} \boldsymbol{\ell}(a,b) = \min_{q \in Q} \max_{b \in \mathcal{B}} \mathbb{E}_{a \sim q} \boldsymbol{\ell}(a,b)$$

and

$$\max_{p \in P} \min_{q \in Q} \mathbb{E}_{a \sim q, b \sim p} \boldsymbol{\ell}(a,b) = \max_{p \in P} \min_{a \in \mathcal{A}} \mathbb{E}_{b \sim p} \boldsymbol{\ell}(a,b)$$

because the minimum (maximum) of a linear functional is achieved at a vertex of the simplex (or set of distributions). In fact, we will go in the opposite direction: we will start with the inner player using a deterministic strategy, and we will **linearize** the problem by passing to the set of distributions over her choices. The expected cost function $\mathbb{E}_{a \sim q, b \sim p} \boldsymbol{\ell}(a,b)$ can be thought of as a loss function of two moves q and p. Furthermore, this function is bilinear.

The second observation is:

Lemma 1. *If $\boldsymbol{\ell}(a,b)$ is convex in a and $\mathcal{A}$ is a convex set, then the outer minimization*

$$\min_{q \in Q} \max_{b \in \mathcal{B}} \mathbb{E}_{a \sim q} \boldsymbol{\ell}(a,b) = \min_{a \in \mathcal{A}} \max_{b \in \mathcal{B}} \boldsymbol{\ell}(a,b)$$

is achieved at a pure strategy.

Proof. We have

$$\min_{a' \in \mathcal{A}} \max_{b \in \mathcal{B}} \boldsymbol{\ell}(a',b) = \min_{q \in Q} \max_{b \in \mathcal{B}} \boldsymbol{\ell}(\mathbb{E}_{a \sim q} a, b) \leq \min_{q \in Q} \max_{b \in \mathcal{B}} \mathbb{E}_{a \sim q} \boldsymbol{\ell}(a,b) \leq \min_{a \in \mathcal{A}} \max_{b \in \mathcal{B}} \boldsymbol{\ell}(a,b) \ .$$

□

As a rather trivial observation, note that *upper bounds* on the minimax value in (1) can be obtained by taking a particular choice q for Player I, while *lower bounds* arise from any choice $b \in \mathcal{B}$. We take the side of Player I, associating her with the Statistician (player, learner), whereas Player II will is Nature (or, adversary, opponent).

1.2 Minimax theorems

In the simplest (bilinear) form, the minimax theorem is due to von Neumann:

Theorem 2. *Let $M \in \mathbb{R}^{p\times m}$ for some $p, m \geq 1$, and let Δ_p and Δ_m denote the set of distributions over the rows and columns, respectively. Then*

$$\min_{a\in\Delta_p}\max_{b\in\Delta_m} a^\top M b = \max_{b\in\Delta_m}\min_{a\in\Delta_p} a^\top M b \tag{3}$$

Since von Neumann, there have been numerous extensions of this theorem, generalizing the statement to uncountable sets of rows/columns, and relaxing the assumptions on the payoff function (which is bilinear in the above statement). To make our discussion more precise, consider the following definition:

Definition 1. Consider sets $\mathcal{A}$ and $\mathcal{B}$ and a function $\boldsymbol{\ell} : \mathcal{A}\times\mathcal{B} \mapsto \mathbb{R}$. We say that *the minimax theorem holds* for the triple $(\mathcal{A}, \mathcal{B}, \boldsymbol{\ell})$ if

$$\inf_{a\in\mathcal{A}}\sup_{b\in\mathcal{B}} \boldsymbol{\ell}(a,b) = \sup_{b\in\mathcal{B}}\inf_{a\in\mathcal{A}} \boldsymbol{\ell}(a,b) \tag{4}$$

With this definition, von Neumann's result says that the minimax theorem holds for the triple $(\Delta_p, \Delta_m, \boldsymbol{\ell})$ with $\boldsymbol{\ell}(a,b) = a^\top M b$ for $a \in \Delta_p, b \in \Delta_m$.

We now give two examples for which the minimax theorem does not hold. The first is the famous "Pick the Bigger Integer" game, attributed to Wald.

Example 1. *The two players in the zero sum game pick positive integers, and the one with the larger number wins. It is not hard to see that the player making the first move has a disadvantage. As soon as her mixed strategy is revealed, the opponent can simply choose a number in the* 99*th percentile of this distribution in order to win most of the time. The same holds when the order is switched, now to the advantage of the other player, and so the minimax theorem does not hold. More precisely, let $\mathcal{A} = \mathcal{B} = \Delta(\mathbb{N})$ be the two sets, and $\boldsymbol{\ell}(a,b) = \mathbb{E}_{w\sim a, v\sim b}\mathbf{1}\{w \leq v\}$. The minimax theorem does not hold for the triple $(\Delta(\mathbb{N}), \Delta(\mathbb{N}), \boldsymbol{\ell})$, as the gap between $\inf_a \sup_b \boldsymbol{\ell}(a,b)$ and $\sup_b \inf_a \boldsymbol{\ell}(a,b)$ can be made as close to 1 as we want.*

The above example might lead us to believe that the reason the minimax theorem does not hold is because we consider distributions over an unbounded set $\mathbb{N}$. However, the above game can be embedded into the $[0,1]$ interval as follows:

Example 2. *The two players in the zero sum game pick real numbers in the interval $[0,1]$. Any player that picks 1 loses (it is a tie if both pick 1); otherwise, the person with the largest real number wins. The setup forces the players into the same situation as in the "Pick the Bigger Integer" example, as the larger number wins while the limit point 1 should be avoided. Let $\mathcal{A} = \mathcal{B} = \Delta([0,1])$ and $\boldsymbol{\ell}(a,b) = \mathbb{E}_{w\sim a, v\sim b}\tilde{\boldsymbol{\ell}}(w,v)$ with*

$$\tilde{\boldsymbol{\ell}}(w,v) = \begin{cases} 1 & \text{if } w = 1, v \neq 1, \\ 0 & \text{if } w \neq 1, v = 1, \\ 0.5 & \text{if } w = v = 1 \\ \mathbf{1}\{w \leq v\} & \text{otherwise} \end{cases}$$

We can use the same reasoning as in Example 1 to show that the minimax theorem does not hold for the triple $(\Delta([0,1]), \Delta([0,1]), \boldsymbol{\ell})$.

Theorem 3 (Theorem 7.1 in [CBL06]). *Suppose $\boldsymbol{\ell} : \mathcal{A} \times \mathcal{B} \to \mathbb{R}$ is a bounded function, where $\mathcal{A}$ and $\mathcal{B}$ are convex and $\mathcal{A}$ is compact. Suppose that $\boldsymbol{\ell}(\cdot, b)$ is convex and continuous for each fixed $b \in \mathcal{B}$ and $\boldsymbol{\ell}(a, \cdot)$ is concave for each $a \in \mathcal{A}$. Then the minimax theorem holds for $(\mathcal{A}, \mathcal{B}, \boldsymbol{\ell})$.*

1.3 Multi-stage games

Now, let us consider a two-stage game where both players make the first moves $a_1 \in \mathcal{A}$ and $b_1 \in \mathcal{B}$, learn each other's moves, and then proceed to the second round, choosing $a_2 \in \mathcal{A}$ and $b_2 \in \mathcal{B}$. Suppose the payoff is a function $\boldsymbol{\ell}(a_1, b_1, a_2, b_2)$. We can then write down the minimax expression

$$\min_{q_1 \in Q} \max_{b_1 \in \mathcal{B}} \mathbb{E}_{a_1 \sim q_1} \min_{q_2 \in Q} \max_{b_2 \in \mathcal{B}} \mathbb{E}_{a_2 \sim q_2} \boldsymbol{\ell}(a_1, b_1, a_2, b_2) \tag{5}$$

Notice that the first-stage moves are written first, followed by the second-stage. This is indeed correct: by writing the terms in such an order we allow, for instance, the choice of q_2 to depend on both a_1 and b_1. This corresponds to the protocol that was just described: both players observe each other's moves. In fact, any minimum or maximum in the sequence is calculated with respect to all the variables that have been "chosen" so far. Informally, by looking at the above expression, we can say that when Player II makes the second move, she "knows" the moves a_1 and b_1 at the first time step, as well as the mixed strategy q_1 of Player I at the second step. Based on these, the maximization over b_2 can be performed. As the expressions in future lectures involve many more than two stages, it is good to go through the sequence and make sure each player "knows" not more and not less than he is supposed to.

Another way to think of the long minimax expression is via operators. To illustrate, $\boldsymbol{\ell}(a_1, b_1, a_2, b_2)$ is a function of 4 variables. The inner expectation over a_2 is an operator (defined by q_2), mapping the function ℓ to a function of 3 variables. The maximization over b_2 maps it to a function of 2 variables, and so forth. When each stage involves the same sequence of operators (e.g. $\min \max \mathbb{E}$), we collapse the long sequence of operators and write

$$\left\langle\!\!\left\langle \min_{q_i \in Q} \max_{b_i \in \mathcal{B}} \mathbb{E}_{a_i \sim q_i} \right\rangle\!\!\right\rangle_{i=1,2} \{\boldsymbol{\ell}(a_1, b_1, a_2, b_2)\} \ .$$

There is yet another way to write the multiple-stage minimax value, such as (5). Consider the second move of the two players. For different a_1, b_1, the best moves a_2, b_2 might be different, and so we will write the second-stage choice as a function of a_1, b_1. To avoid the issue of simultaneity, assume that both players choose mixed strategy $q_2 = \pi_2(a_1, b_1)$ and $p_2 = \tau_2(a_1, b_1)$, respectively. The first-stage decisions are trivial constant mappings $q_1 = \pi_1(), p_1 = \tau_1()$, but the later-stage decisions depend on the past. We can write $\pi = (\pi_1, \pi_2)$ and $\tau = (\tau_1, \tau_2)$ as two-stage *strategies* of the players. We can then rewrite (5) as

$$\min_{\pi} \max_{\tau} \mathbb{E}_{a_1 \sim \pi_1} \mathbb{E}_{b_1 \sim \tau_1} \mathbb{E}_{a_2 \sim \pi_2(a_1, b_1)} \mathbb{E}_{b_2 \sim \tau_2(a_1, b_1)} \boldsymbol{\ell}(a_1, b_1, a_2, b_2) \ ,$$

or, more succinctly as

$$\min_{\pi} \max_{\tau} \mathbb{E} \boldsymbol{\ell}(a_1, b_1, a_2, b_2) \ . \tag{6}$$

We say that the above expression is written in terms of *strategies*. As we already observed, the maximization over τ will be achieved at a non-randomized strategy, resulting in the simplified stochastic process that evolves according to the randomized choice of Player I and the deterministic choice of Player II. Finally, we remark that π can be thought of as a joint distribution over sequences (a_1, a_2).

References

[CBL06] N. Cesa-Bianchi and G. Lugosi. *Prediction, Learning, and Games.* Cambridge University Press, 2006.