Lecture 6

Lecturer: Bo Waggoner *Scribe: Aaron Roth, Bo Waggoner*

Zero Sum Games and the MinMax Theorem

In this lecture we study *zero sum games*, which have very special mathematical and computational properties. They are important for several reasons: first, they model *strictly adversarial games* – i.e. games in which the only way for player 1 to improve his payoff is to harm player 2, and vice versa. It turns out that zero sum games are much easier to play than "general sum" games (which do not have this property). Their equilibria are easy to compute, and avoid issues of needing to coordinate to select on an equilibrium. They also have the remarkable property that it is no disadvantage to have to *go first*, announce your (mixed) strategy, and then let your opponent best respond – you are guaranteed the same payoff as if you had the opportunity to best respond to your opponent. This is known as the minmax theorem, and is not satisfied by general sum games. Second, zero-sum games are closely related to linear programs and other optimization problems, and the minmax theorem in this context corresponds to the useful and powerful fact of *strong duality.*

Definition 1 *A* ***two player zero sum game*** *is any two player game such that for every $a \in A_1 \times A_2$, $u_1(a) = -u_2(a)$. That is, at every action profile, the two players' utilities sum to zero.*

Lets first think about how to reason about such games. Consider the "Battle of the Bands" game:

	Violin	Harmonica
Guitar	(3,-3)	(-1,1)
Drums	(-2,2)	(1,-1)

Since utilities sum to zero, it is typical to write zero sum games by specifying only player 1's utility. Equivalently, written this way, player 1 – the row player – wishes to maximize the payoff in the game. (He is the *maximization* player, call him Max). Player 2 wishes to minimize the payoff in the game. (She is the *minimization* player, named Min.)

	Violin	Harmonica
Guitar	3	-1
Drums	-2	1

We can write $u(a)$ for Max's utility in the game.

Lets first think about how the game should be played if Max has to go first, and announce his strategy, allowing Min to best respond. Suppose Max announces the mixed strategy $p = 1/2$, i.e. probability $\frac{1}{2}$ on each action. What should Min do?

Min should of course pick the action that minimizes her cost! She can compute:

$$\mathbb{E}\left[u\left(\text{Violin}, p\right)\right] = \frac{1}{2} \cdot 3 + \frac{1}{2} \cdot (-2) = \frac{1}{2}$$

$$\mathbb{E}\left[u\left(\text{Harmonica}, p\right)\right] = \frac{1}{2} \cdot (-1) + \frac{1}{2} \cdot 1 = 0$$

So Min should play Harmonica. More generally, if Max announces that he is going to play some strategy $(p, 1-p)$, then Min should play the strategy that minimizes her cost:

$$\min \begin{cases} p \cdot 3 - 2(1-p) & \text{(Min plays Violin)} \\ p \cdot (-1) + (1-p)) & \text{(Min plays Harmonica).} \end{cases}$$

So what should Max do, if he has to go first? Knowing what Min will do, he should play the the mixed strategy that *maximizes* the minimum payoff that Min will be able to inflict on him. That is, he should pick p to solve:

$$\max_p \min \begin{cases} p \cdot 3 - 2(1-p) & \text{(Min plays Violin)} \\ p \cdot (-1) + (1-p)) & \text{(Min plays Harmonica).} \end{cases}$$

Similarly, if Min has to go first, she should play the strategy that will minimize the maximum payoff that Max will be able to achieve when he best responds:

$$\min_q \max \begin{cases} q \cdot 3 - (1-q) & \text{(Max plays Guitar)} \\ q \cdot (-2) + (1-q) & \text{(Max plays Drums).} \end{cases}$$

Going first is clearly a disadvantage – it reveals your strategy to your opponent and lets them respond optimally. In this case, lets see how much of a disadvantage it is. What should Max do?

Since Min will always play the lower cost of her two actions, Max should play to equalize the cost of these two actions – i.e. he should pick p such that:

$$3p - 2(1-p) = -p + (1-p) \Leftrightarrow 5p - 2 = 1 - 2p \Leftrightarrow p = \frac{3}{7}$$

So Max should play $(3/7, 4/7)$, and when Min best responds, he gets payoff $1 - 2p = 1/7$ (and since it is a zero sum game, Min suffers cost $1/7$.)

(Note one way to see this is to plot Min's expected utility for each pure strategy, as a function of p. Min will get the minimum of the two lines at each point, and Max can maximize this by playing the place where they intersect.)

And if Min goes first? Again, since Max will best respond, Min should play so as to equalize the payoff of Max's two options. She should pick q such that:

$$3q - (1-q) = -2q + (1-q) \Leftrightarrow 4q - 1 = -3q + 1 \Leftrightarrow q = \frac{2}{7}$$

So Min should play $(2/7, 5/7)$. In this case, when Max best responds, Min suffers cost $4q - 1 = 1/7$ (and Max gets payoff $1/7$)...

Hm – So in this case, it was in fact no disadvantage to have to go first! We have the max min value of the game exactly equal to the min max value of the game. Whats more, this immediately implies that $(3/7, 4/7), (2/7, 5/7)$ is a Nash equilibrium of the game, since both players are getting payoff/cost $1/7$, and we have just derived that both are best responding to one another.

More formally, in a 2-player zero-sum game with action spaces A_1 and A_2, let $\Delta[A_1]$ be the set of distributions on A_1, so that player 1's mixed strategy is some $p \in \Delta[A_1]$; and similarly, player 2's mixed strategy is some $q \in \Delta[A_2]$. Let $u(p, q)$ be the utility of player 1 when player 1 plays according to p and player 2 plays according to q; so player 2's utility is $-u(p, q)$.

Notice that for a fixed p, player 2's best response will result in $\min_q u(p, q)$ for player 1. So:

1. If player 1 goes first and player 2 best-responds, then player 1 should play the p that solves $\max_p \min_q u(p, q)$.

2. Analogously, if player 2 goes first and player 1 best-responds, then player 2 should play the q that solves $\min_q \max_p u(p, q)$.

Now, if player 1 goes first, then he will certainly do worse than if he goes second: if he goes first, he must pick p without knowing which q he will face, while if he goes second, he gets to observe q and then best-respond.

Theorem 2 *In any 2-player zero-sum game,*

$$\max_p \min_q u(p, q) \leq \min_q \max_p u(p, q).$$

Proof Let p^* solve $\max_p \min_q u(p,q)$. Then

$$\begin{aligned}\max_p \min_q u(p,q) &= \min_q u(p^*,q)\\ &\le \min_q \max_p u(p,q).\end{aligned}$$

The last line follows because, for any particular q, $\max_p u(p,q) \ge u(p^*,q)$. So the minimum over all $u(p^*,q)$ is smaller than the minimum over all $\max_p u(p,q)$. ■

Theorem 3 (von Neumann Min-Max Theorem) *In any 2-player zero-sum game,*

$$\max_p \min_q u(p,q) = \min_q \max_p u(p,q).$$

Before proving this, what interesting things does it imply?

Corollary 4 *All Nash equilibria of a 2-player zero-sum game have the same expected utility of player 1; this is called the **value** of the game. In all of these, player 1 plays a max-min strategy and player 2 plays a min-max strategy.*

Proof Player 1 can guarantee himself $\max_p \min_q u(p,q)$ by playing a min-max strategy, and only by playing a max-min strategy, i.e. solving for p in that expression. On the other hand, player 2 can guarantee that he gets no more than $\min_q \max_p u(p,q)$ by playing a min-max strategy. The min-max theorem says these are equal, so player 1 must get that value in any equilibrium, and the players must play such strategies. ■

None of these things are true in general sum games. The minmax theorem should also be surprising: it means that (if they are smart), there is no way of taking advantage of knowing your opponent's strategy and forcing them to commit to it. (Nor, for that matter, can you be "coerced" into an undesirable equilibrium by your opponent committing to a strategy, as in Battle of the Sexes.)

The theorem is not obvious... Von Neumann proved it in 1928, and said ""As far as I can see, there could be no theory of games without that theorem I thought there was nothing worth publishing until the Minimax Theorem was proved". Previously, Borell had proven it for the special case of 5×5 matrices, and thought it was false for larger matrices.

However. Now that we know of the polynomial weights algorithm, we can provide a very simple, constructive proof.

Proof [Min-Max Theorem] Let:

$$\begin{aligned}v_1 &:= \min_q \max_p u(p,q),\\ v_2 &:= \max_p \min_q u(p,q).\end{aligned}$$

By Theorem 2, we have $v_1 \ge v_2$. Let $v_1 = v_2 + \epsilon$ for some $\epsilon > 0$ (strictly positive); we will show a contradiction.

Lets consider what happens when Min (player 2) and Max (player 1) repeatedly play against each other as follows, for T rounds:

1. Min will play using the polynomial weights algorithm. i.e. at each round t, the weights q^t of the polynomial weights algorithm will form her mixed strategy, and she will sample an action at random from this distribution, updating based on the losses she experiences at that round.

2. Max will play a best response to Min's strategy. i.e. Max will play $p^t = \arg\max_p u(p, q^t)$.

Consider what we know about each of their average payoffs when they play in this manner.

Each round, Max plays a best response, so at each round t, he gets

$$\begin{aligned} u(p^t, q^t) &= \max_p u(p, q^t) \\ &\geq \min_q \max_p u(p, q) \\ &= v_1. \end{aligned}$$

So Max gets average payoff at least v_1, which makes sense as he always gets to best-respond to Min.

On the other hand, consider the loss of the best action for Min in hindsight. This is like seeing Max's strategies at all the time steps, then best-responding, so intuitively it should have loss at most v_2. Let $\hat{p}$ be the distribution obtained by picking one of the T time steps uniformly at random, then playing according to p^t. Then the best response of Min in hindsight has average loss

$$\begin{aligned} \min_{a_2 \in A_2} \frac{1}{T} \sum_{t=1}^{T} u(p^t, a_2) &= \min_{a_2 \in A_2} u(\hat{p}, a_2) \\ &= \min_q u(\hat{p}, q) \\ &\leq \max_p \min_q u(p, q) \\ &= v_2. \end{aligned}$$

Min has average regret at most $2\sqrt{\frac{\log |A_2|}{T}}$ by the guarantee of the polynomial weights algorithm. So Min has average loss at most

$$v_2 + 2\sqrt{\frac{\log |A_2|}{T}}.$$

Now if we choose $T = \frac{16 \log |A_2|}{\epsilon^2}$, then Min has average loss at most

$$\begin{aligned} &v_2 + 2\sqrt{\frac{\epsilon^2 \log |A_2|}{16 \log |A_2|}} \\ &= v_2 + \frac{\epsilon}{2}. \end{aligned}$$

Now Max's average gain equals Min's average loss. We first showed above that this number is at least v_1, and then we just showed it is at most $v_2 + \frac{\epsilon}{2}$. So

$$v_1 \leq \text{Min's average loss} \leq v_2 + \frac{\epsilon}{2}.$$

So

$$v_1 \leq v_2 + \frac{\epsilon}{2}.$$

But we said above that $v_1 = v_2 + \epsilon$, which is a contradiction. ■

This proof is worth a bit of reflection. In particular, it has highlighted the particularly amazing feature of the polynomial weights algorithm: it guarantees that no matter what happens, you do as well *as if you had gotten to observe your opponent's strategy, and then best respond after the fact.* In particular, we have proven that every zero sum game U has a unique *value* that is the best payoff a player can hope for, if her opponent is playing optimally. Using the polynomial weights algorithm guarantees that the player using gets payoff quickly approaching the value of the game. What's more, it does so *without needing to know what the game is.* Note that at no point is the game matrix input to the PW algorithm! The only information it needs to know is what the realized payoffs are for its actions, as it actually plays the game. As such, it is an attractive algorithm to use in an interaction that you don't know much about...