NETS 412: Algorithmic Game Theory February 2, 2017

Lecture 6

Lecturer: Aaron Roth *Scribe: Aaron Roth*

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 – 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 $a \in A_1 \times A_2$, $u_1(a) = -u_2(a)$.(i.e. at every action profile, the utilities sum to zero)*

Lets first think about how to reason about such games:

Consider the "Presidential Election Game":

	Morality	Tax-Cuts
Economy	(3,-3)	(-1,1)
Society	(-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 (named Max)). Player 2 wishes to minimize the payoff in the game. (She is the *minimization* player (named Min)):

	Morality	Tax-Cuts
Economy	3	-1
Society	-2	1

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 $(1/2, 1/2)$. What should Min do?

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

$$\text{E[Morality]} = \frac{1}{2} \cdot 3 + \frac{1}{2} \cdot (-2) = \frac{1}{2}$$

$$\text{E[Tax} - \text{Cuts]} = \frac{1}{2} \cdot (-1) + \frac{1}{2} \cdot 1 = 0$$

So Min should play Tax-Cuts. 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:

$$\arg\min(p \cdot 3 - 2(1-p), p \cdot (-1) + (1-p))$$

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 play:

$$\arg\max_p \min(p \cdot 3 - 2(1-p), p \cdot (-1) + (1-p))$$

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:

$$\arg\min_q \max(q \cdot 3 - (1-q), q \cdot (-2) + (1-q))$$

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...)

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.

Lets investigate this further. We use the notation $[n] = \{1, 2, \ldots, n\}$, and $\Delta[n]$ to denote the set of probability distributions over $[n]$:

Definition 2 *For an $n \times m$ matrix U (think about this as the payoff matrix in a two player zero sum game if you like):*

$$\max\min(U) = \max_{p \in \Delta[n]} \min_{y \in [m]} \sum_{i=1}^{n} p_i \cdot U(i, y)$$

$$\min\max(U) = \min_{q \in \Delta[m]} \max_{x \in [n]} \sum_{j=1}^{m} q_j \cdot U(x, j)$$

Note that if U is a zero sum game, then $\max\min(U)$ *represents the payoff that Max can guarantee if he goes first, and* $\min\max(U)$ *represents the payoff that he can guarantee if Min goes first. Note also that we have defined things so that the first player is optimizing over mixed strategies, but the second player is optimizing over pure strategies. We could have let both optimize over mixed strategies, but this is without loss, since any player always has a pure strategy among her set of best responses.*

It is apparent that playing second can only be an advantage: your strategy space is not limited by the first player's action, and you only have more information. In particular, this implies that for any game U:

$$\min\max(U) \geq \max\min(U)$$

What is not apparent is that going first is not a disadvantage:

Theorem 3 (Von Neumann) *In any zero sum game U:*

$$\min\max(U) = \max\min(U)$$

Corollary 4 *In any Nash equilibrium of a zero sum game, Max plays a maxmin strategy and Min plays a minmax strategy. Note that these can be computed without needing to reason about what the other player is doing.*

Corollary 5 *All Nash equilibria in Zero sum games have the same payoff – the* max min *value of the game.*

None of these things are true in general sum games (consider e.g. Battle of the Sexes). 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.

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 Suppose the theorem were false: That is, there is some game U for which $\min\max(U) > \max\min(U)$. Write $v_1 = \min\max(U)$ and $v_2 = \max\min(U)$ (And so $v_1 = v_2 + \epsilon$ for some constant $\epsilon > 0$). In other words, if Min has to go first, then Max can guarantee payoff at least v_1, but if Max is forced to go first, then Min can force Max to have payoff only v_2.

Lets consider what happens when Min and Max 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 w^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 the best response to Min's strategy. i.e. Max will play $x^t = \arg\max_x \mathrm{E}_{y \sim w^t}[U(x, y)]$.

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

On the one hand, we know from the guarantee of the polynomial weights algorithm that:

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T} \mathrm{E}[U(x^t, y^t)] &\le \frac{1}{T}\min_{y^*}\sum_{t=1}^{T} U(x^t, y^*) + \Delta(T) \\
&= \min_{y^*}\sum_{t=1}^{T} \frac{1}{T} U(x^t, y^*) + \Delta(T) \\
&= \min_{y^*} \mathrm{E}_{x\sim\bar{x}}[U(x, y^*)] + \Delta(T)
\end{aligned}$$

where $\bar{x}$ is the mixed strategy that puts weight $1/T$ on each action x^t. $\Delta(T)$ is the regret bound of the polynomial weights algorithm – recall:

$$\Delta(T) = 2\sqrt{\frac{\log n}{T}}.$$

But by definition, $\min_{y^*} \mathrm{E}_{x\sim\bar{x}} U(x, y^*) \le \max\min(U) = v_2$ and so we know:

$$\frac{1}{T}\sum_{t=1}^{T} \mathrm{E}[U(x^t, y^t)] \le v_2 + \Delta(T)$$

On the other hand, on each day t we know Max is best responding to Min's mixed strategy w^t. Thus:

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T} \mathrm{E}[U(x^t, y^t)] &\geq \frac{1}{T}\sum_{t=1}^{T} \max_{x^*} \mathrm{E}_{y\sim w^t}[U(x^*, y)] \\
&\geq \frac{1}{T}\sum_{t=1}^{T} v_1 \\
&= v_1
\end{aligned}$$

Combining these inequalities, we know:

$$v_1 \leq v_2 + \Delta(T)$$

Recall $v_1 = v_2 + \epsilon$, so:

$$\epsilon \leq \Delta(T)$$

but taking $T = \frac{16\ln(n)}{\epsilon^2}$ we get $\Delta(T) = \frac{\epsilon}{2}$ which implies:

$$\epsilon \leq \epsilon/2$$

Since ϵ is positive, this is a contradiction and concludes the proof. ■

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* $v = \max\min(U) = \min\max(U)$ that is the best value 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...