Lecture 18

Lecturer: Aaron Roth *Scribe: Aaron Roth*

The Minimax Theorem and Equilibria of Zero Sum Games

In this lecture we'll take a brief interlude from traditional algorithm design to prove the fundamental *minimax* theorem for zero sum games, which turns out to be closely related to *linear programming duality.* We'll also show how to actually compute equilibria of zero sum games. Remarkably, all of this will follow once again from our analysis of the polynomial weights algorithm. Next lecture we'll use this theorem to derive more useful algortihms.

First, what is a "zero-sum game"? It is a model of a strictly adversarial interaction, in which one player's sole objective is to minimize some cost function, and their opponent's sole objective is to maximize it. The players have the ability to choose amongst a set of actions, which jointly determine the cost. One can define this somewhat more generally, but it will suffice for us to talk about players with finite sets of actions who are allowed to choose probability distributions over those actions (i.e. to randomize)

Definition 1 *A* zero sum game *is defined by an action set $A_1 = \{1, \ldots, m\}$ for the* minimization player*, an action set $A_2 = \{1, \ldots, n\}$ for the* maximization player*, and a cost function $C : A_1 \times A_2 \to \mathbb{R}$.*

We can represent a zero sum game by thinking of C as a matrix, in which the rows correspond to actions of the minimization player (lets call her Min), columns correspond to actions of the maximization player (lets call him Max), and the entries record the costs that result from the corresponding choices of actions by Min and Max. For example, you might recognize the following zero sum game as "Rock Paper Scissors":

	Rock	Paper	Scissors
Rock	1	2	0
Paper	0	1	2
Scissors	2	0	1

A (*mixed*) strategy for Min corresponds to a probability distribution over her actions: $p \in \Delta[m]$. Similarly, a mixed strategy for Max corresponds to a probability distribution over his actions: $q \in \Delta[n]$. When players randomize, we compute the *expected* cost of the resulting outcome:

$$C(p, q) = \sum_{i=1}^{m} \sum_{j=1}^{n} C[i, j] p_i q_j = q^T C p$$

If one of the players plays a *pure* strategy — i.e. does not randomize — for example, Max might deterministically play $y \in [n]$ we will abuse notation and write:

$$C(p, y) = \sum_{i=1}^{m} C[i, y] p_i = e_y^T C p$$

Normally, Rock Paper Scissors is played as a simultaneous move game. But what if Min were forced to play at a disadvantage, by having to first *announce* her strategy to Max, who would then get to best respond? If she played the strategy $p = (2/3, 1/3, 0)$, Max would exploit the fact that Min was playing Rock too frequently, and play paper in response, resulting in cost $2/3 \cdot 2 + 1/3 \cdot 1 = 5/3$. Instead, she should play so as to minimize the cost that results *after* Max best response. Similarly, if Max is handicapped by the need to go first and announce his strategy before Min gets an opportunity to best respond, what he should do is play so as to maximize the cost *after* Min best responds by choosing the action with minimum cost. We can define the corresponding MinMax and MaxMin values of the game:

Definition 2 *For an $n \times m$ matrix C:*

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

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

Here note that we have the player who goes second playing a single action, rather than a distribution over actions — but this is without loss of generality, since a player's *best response* need never be randomized. (because an average over a bunch of numbers can never be smaller than the minimum or larger than the maximum) Of course, in Rock Paper Scissors it doesn't matter who goes first: either player will randomize uniformly across Rock Paper and Scissors if they go first, which will make their opponent indifferent between their options: optimal play obtains cost $1/3 \cdot 1 + 1/3 \cdot 2 + 1/3 \cdot 0 = 1$ in both cases.

This turns out to be more generally true in zero-sum games: it doesn't matter who goes first. This is a surprisingly deep (and useful) fact known as Von Neumann's minimax theorem. For Min, going first is clearly only a disadvantage, since she is revealing information to Max, and so we know that $\min\max(C) \geq \max\min(C)$. The minimax theorem says that this inequality is in fact an equality:

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

$$\min\max(C) = \max\min(C)$$

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. In fact, what we'll do is give an algorithm that explicitely constructs strategies p, q for Min and Max respectively such that (p, q) form an ϵ-approximate *minimax equilibrium.* Von Neumann's minimax theorem will follow as a corollary.

Definition 4 *Vectors $p \in \Delta[m]$, $q \in \Delta[n]$ form an ϵ-approximate minimax equilibrium with respect to a game C if:*

$$\max_{y \in [n]} C(p, y) - \epsilon \leq C(p, q) \leq \min_{x \in [m]} C(x, q) + \epsilon$$

Algorithm 1 ComputeEQ(C, ϵ)

Let $T \leftarrow \frac{4 \log m}{\epsilon^2}$
Initialize a copy of polynomial weights to run over $w^t \in \Delta^m$.
for $t = 1$ to T **do**
  Let $y^t = \arg\max_{y \in [n]} C(w^t, y)$
  Let $\ell^t \in [0, 1]^m$ be such that $\ell^t_i = C[i, y_t]$.
  Pass ℓ^t to the PW algorithm.
end for
Let $\bar{x} = \frac{1}{T} \sum_{t=1}^{T} w^t$ and $\bar{y} = \frac{1}{T} \sum_{t=1}^{T} e_{y^t}$.
Return $(\bar{x}, \bar{y})$.

In the algorithm above, $e_i \in [0, 1]^n$ refers to the i'th standard basis vector — i.e. it has a 1 in its i'th index and a 0 in every other index.

Theorem 5 *For any $\epsilon > 0$, and any $n \times m$ 0-sum game C, ComputeEQ(C, ϵ) returns vectors $(\bar{x}, \bar{y})$ that form an ϵ-approximate minimax equilibrium.*

To see that Von Neumann's theorem follows as a corollary, note that this implies that for any $\epsilon > 0$, we can find $\bar{x}, \bar{y}$ such that:

$$\min\max(C) - \epsilon \leq \max_{y \in [n]} C(\bar{x}, y) - \epsilon \leq C(\bar{x}, \bar{y}) \leq \min_{x \in [m]} C(x, \bar{y}) + \epsilon \leq \max\min(C) + \epsilon$$

Thus we have $\min\max(C) \leq \max\min(C) + 2\epsilon$ for every ϵ, so it must be that $\min\max(C) = \max\min(C)$.

Now to prove the theorem:

Proof We begin with a useful observation coming from linearity. Let $x^* \in [m]$ be any fixed action. Then:

$$\begin{aligned} \frac{1}{T}\sum_{t=1}^{T} C(x^*, y^t) &= C\left(x^*, \frac{1}{T}\sum_{t=1}^{T} e_{y^t}\right) \\ &= C(x^*, \bar{y}) \end{aligned}$$

Similarly, for any fixed y^* $\frac{1}{T}C(w^t, y^*) = C(\bar{x}, y^*)$. Now suppose Min and Max are playing using $\bar{x}$ and $\bar{y}$ respectively. Let $x^* = \arg\min_x C(x, \bar{y})$ and $y^* = \arg\max_y C(\bar{x}, y)$ be their best responses. By definition:

$$C(x^*, \bar{y}) \leq C(\bar{x}, \bar{y}) \leq C(\bar{x}, y^*)$$

We also know that by the guarantee of the polynomial weights algorithm that on the one hand:

$$\begin{aligned} \frac{1}{T}\sum_{t=1}^{T} C(w^t, y^t) &\leq \frac{1}{T}\sum_{t=1}^{T} C(x^*, y^t) + \sqrt{\frac{4\log m}{T}} \\ &= C(x^*, \bar{y}) + \sqrt{\frac{4\log m}{T}} \end{aligned}$$

And on the other hand, by definition of the choice of the y^t:

$$\begin{aligned} \frac{1}{T}\sum_{t=1}^{T} C(w^t, y^t) &\geq \frac{1}{T}\sum_{t=1}^{T} C(w^t, y^*) \\ &= C(\bar{x}, y^*) \end{aligned}$$

Subtracting the second inequality from the first, we have:

$$0 \leq C(x^*, \bar{y}) - C(\bar{x}, y^*) + \sqrt{\frac{4\log m}{T}}.$$

Adding and subtracting $C(\bar{x}, \bar{y})$ and multiplying by -1 we get:

$$(C(\bar{x}, \bar{y}) - C(x^*, \bar{y})) + (C(\bar{x}, y^*) - C(\bar{x}, \bar{y})) \leq \sqrt{\frac{4\log m}{T}}$$

Finally recall that by definition of x^* and y^*, we have that both terms on the left hand side are non-negative. Thus we have that individually:

$$(C(\bar{x}, \bar{y}) - C(x^*, \bar{y})) \leq \sqrt{\frac{4\log m}{T}} \quad (C(\bar{x}, y^*) - C(\bar{x}, \bar{y})) \leq \sqrt{\frac{4\log m}{T}}$$

which implies that $(\bar{x}, \bar{y})$ form a $\sqrt{\frac{4\log m}{T}}$-approximate minimax equilibrium. By our choice of T, $\sqrt{\frac{4\log m}{T}} = \epsilon$. ■

Let us make some important remarks about this algorithm:

1. We had the minimization player play according to polynomial weights, and the maximization player best respond — but the situation is symmetric, and we could have reversed these roles.
2. The convergence bound depends logarithmically on the number of actions of the polynomial weights player (inherited from PWs regret bound), but no dependence at all on the number of actions of the best response player.
3. The per-round running time of the algorithm is linear in the number of actions of the polynomial weights player since we need to maintain the PW distribution. But all we need to be able to do for the best response player is efficiently compute a best response, which we might sometimes be able to do even when his action space is exponentially large.

When applying the algorithm, we can try and take advantage of this property by framing the computation correctly.