C&O 355: Mathematical Programming
Fall 2010
Lecture 12 Notes

Nicholas Harvey
http://www.math.uwaterloo.ca/~harvey/

1 Zero-Sum Games

Let M be any $m \times n$ real matrix, which we use as the payoff matrix for a two-player, zero-sum game. Von Neumann's theorem states that

$$\max_x \min_y x^\mathsf{T} My \;=\; \min_y \max_x x^\mathsf{T} My,$$

where the max and min are over distributions $x \in \mathbb{R}^m$ and $y \in \mathbb{R}^n$. Recall that "distribution" means that $x \geq 0$, $\sum_{i=1}^m x_i = 1$. Consequently, there exist distributions $x^* \in \mathbb{R}^m$ and $y^* \in \mathbb{R}^n$ such that

$$\max_x \min_y x^\mathsf{T} My \;=\; {x^*}^\mathsf{T} My^* \;=\; \min_y \max_x x^\mathsf{T} My. \tag{1}$$

This quantity is called the ***value*** of the game and is denoted by v.

Observation 1. Note that for any fixed x, we have $\min_y x^\mathsf{T} My \leq v$. (In particular, $x^\mathsf{T} My^* \leq v$.) Similarly, for any particular y, we have $\max_x x^\mathsf{T} My \geq v$. (In particular, ${x^*}^\mathsf{T} My \geq v$.)

Observation 2. For any fixed x, there is a y achieving $\min_y x^\mathsf{T} My$ such that y has only one non-zero coordinate (which must have value 1). Such a y corresponds to Bob choosing a single action, rather than a randomized choice of actions.

Fix any desired error $\delta \in (0, 1)$. We will give a method to find distributions $\hat{x}$ and $\hat{y}$ such that

$$|\min_y \hat{x}^\mathsf{T} My - v| \;\leq\; \delta \qquad \text{and} \qquad |\max_x x^\mathsf{T} M\hat{y} - v| \;\leq\; \delta. \tag{2}$$

Due to Observation 1, we see that (2) is equivalent to

$$\min_y \hat{x}^\mathsf{T} My \;\geq\; v - \delta \qquad \text{and} \qquad \max_x x^\mathsf{T} M\hat{y} \;\leq\; v + \delta. \tag{3}$$

In other words, if Alice plays according to distribution $\hat{x}$, then no matter how Bob plays, she is guaranteed a payoff of at least $v - \delta$. Conversely, if Bob plays according to distribution $\hat{y}$, then no matter how Alice plays, he is guaranteed to pay her at most $v + \delta$.

2 The Multiplicative Weights Update Method

By scaling, we may assume that $M_{i,j} \in [-1, 1]$ for every i, j. Set $\epsilon = \delta/3$. Alice will assign some "weights" to each of her actions, then simulate the game by herself for $T = (\ln m)/\epsilon^2$ rounds, modifying her weights between each round. These weights are essentially a probability distribution, except they are not normalized to have sum 1.

Algorithm 1 Finding an approximate equilibrium in a zero-sum game by the multiplicative weights update method.

procedure FindEquilibrium($M,\ \delta$)

input: An $m \times n$ matrix M and desired error $\delta \in (0,1)$. Assume that $M_{i,j} \in [-1,1]$ for every i,j.

output: Distributions $\hat{x} \in \mathbb{R}^m$ and $\hat{y} \in \mathbb{R}^n$ satisfying (3).

Set $\epsilon = \delta/3$ and $T = (\ln m)/\epsilon^2$

Set $w_i^{(1)} = 1$ for every $i = 1, ..., m$

For $t = 1, ..., T$

  Set $x^{(t)} = w^{(t)} / \sum_{i=1}^m w_i^{(t)}$

  Let $j^{(t)}$ be a value of j minimizing $(x^{(t)} M)_j$

  Let $y^{(t)}$ be the vector with 1 in coordinate $j^{(t)}$ and other coordinates 0

  Set

$$w_i^{(t+1)} \;=\; \begin{cases} w_i^{(t)} \cdot (1+\epsilon)^{M_{i,j^{(t)}}} & (\text{if } M_{i,j^{(t)}} \geq 0) \\ w_i^{(t)} \cdot (1-\epsilon)^{-M_{i,j^{(t)}}} & (\text{if } M_{i,j^{(t)}} < 0) \end{cases}$$

Set $\hat{x} = \sum_{t=1}^T x^{(t)}/T$ and $\hat{y} = \sum_{t=1}^T y^{(t)}/T$

Return $\hat{x}$ and $\hat{y}$

Formally, in round t, Alice has non-negative weights $w^{(t)} \in \mathbb{R}^m$. In round 1, these weights are all initially 1. She normalizes them to get a distribution $x^{(t)} = w^{(t)} / \sum_i w_i^{(t)}$. Then she imagines what Bob would do if he knew she were using distribution $x^{(t)}$. Of course, he would choose a distribution $y^{(t)}$ achieving $\min_y x^{(t)\mathsf{T}} M y$. By Observation 2, Bob could even choose $y^{(t)}$ to have a single non-zero coordinate. In other words, there is an action $j^{(t)}$ such that, if Bob knows that Alice is using distribution $x^{(t)}$, his best choice is action $j^{(t)}$.

Then Alice updates her weights by setting, for every i,

$$w_i^{(t+1)} \;=\; \begin{cases} w_i^{(t)} \cdot (1+\epsilon)^{M_{i,j^{(t)}}} & (\text{if } M_{i,j^{(t)}} \geq 0) \\ w_i^{(t)} \cdot (1-\epsilon)^{-M_{i,j^{(t)}}} & (\text{if } M_{i,j^{(t)}} < 0). \end{cases}$$

The key to analyzing this algorithm is the following lemma. In English, it says: the average payoff that Alice receives during the execution of the algorithm is not much worse than it would be if Bob continued to choose action $j^{(t)}$ in every round t, but Alice chose a fixed action i in every round.

Lemma 3. For every i, we have

$$\sum_{t=1}^{T} \frac{x^{(t)\mathsf{T}} M y^{(t)}}{T} \;\geq\; \sum_{t=1}^{T} \frac{M_{i,j^{(t)}}}{T} \;-\; 3\epsilon. \tag{4}$$

Proof. We require two inequalities that follow from convexity of the exponential function.

$$\begin{aligned} (1+\epsilon)^x \;&\leq\; (1+\epsilon x) \qquad \forall\, \epsilon \geq 0,\ x \in [0,1] \\ (1-\epsilon)^{-x} \;&\leq\; (1+\epsilon x) \qquad \forall\, \epsilon \in [0,1),\ x \in [-1,0] \end{aligned}$$

Let $W^{(t)}$ be the total weight in round t. Now let us see how the total weight differs between round t and $t+1$.

$$W^{(t+1)} \;=\; \sum_{i=1}^{m} w_i^{(t+1)}$$

$$
\begin{aligned}
&= \sum_{i\,:\,M_{i,j^{(t)}}\geq 0} w_i^{(t)}\cdot(1+\epsilon)^{M_{i,j^{(t)}}} \;+\; \sum_{i\,:\,M_{i,j^{(t)}}<0} w_i^{(t)}\cdot(1-\epsilon)^{-M_{i,j^{(t)}}}\\
&\leq \sum_{i=1}^{m} w_i^{(t)}(1+\epsilon M_{i,j^{(t)}})
\end{aligned}
$$

Now we use that $x^{(t)} = w^{(t)}/W^{(t)}$, and that $y^{(t)}$ has a 1 in coordinate $j^{(t)}$ and all other coordinates 0.

$$
\begin{aligned}
&= W^{(t)}(1+\epsilon x^{(t)\mathsf{T}} M y^{(t)})\\
&\leq W^{(t)}\exp(\epsilon x^{(t)\mathsf{T}} M y^{(t)}).
\end{aligned}
$$

Thus, after T rounds, we have the following upper bound on $W^{(T+1)}$.

$$
W^{(T+1)} \;\leq\; W^{(1)}\cdot\prod_{t=1}^{T}\exp\left(\epsilon x^{(t)\mathsf{T}} M y^{(t)}\right) \;=\; m\cdot\exp\left(\epsilon\sum_{t=1}^{T} x^{(t)\mathsf{T}} M y^{(t)}\right). \tag{5}
$$

Since the weights are non-negative, for any i we obtain a lower bound on $W^{(T+1)}$.

$$
W^{(T+1)} \;\geq\; w_i^{(T+1)} \;=\; \prod_{t\,:\,M_{i,j^{(t)}}\geq 0}(1+\epsilon)^{M_{i,j^{(t)}}} \;\cdot\; \prod_{t\,:\,M_{i,j^{(t)}}<0}(1-\epsilon)^{-M_{i,j^{(t)}}} \tag{6}
$$

So, combining (5) and (6), taking the logarithm, and using $W^{(1)} = m$, we have:

$$
\sum_{t\,:\,M_{i,j^{(t)}}\geq 0} M_{i,j^{(t)}}\ln(1+\epsilon) \;+\; \sum_{t\,:\,M_{i,j^{(t)}}<0} M_{i,j^{(t)}}\ln\left((1-\epsilon)^{-1}\right) \;\leq\; \ln m \;+\; \epsilon\sum_{t=1}^{T} x^{(t)\mathsf{T}} M y^{(t)}.
$$

Now using the inequalities $\ln(\frac{1}{1-\epsilon}) \leq \epsilon+\epsilon^2$ and $\ln(1+\epsilon)\geq\epsilon-\epsilon^2$ (which are valid for all $\epsilon\in(0,1/2)$), then dividing by ϵ, we get:

$$
\begin{aligned}
\sum_{t=1}^{T} x^{(t)\mathsf{T}} M y^{(t)} &\geq (1-\epsilon)\sum_{t\,:\,M_{i,j^{(t)}}\geq 0} M_{i,j^{(t)}} \;+\; (1+\epsilon)\sum_{t\,:\,M_{i,j^{(t)}}<0} M_{i,j^{(t)}} \;-\; \frac{\ln m}{\epsilon}\\
&\geq \sum_{t=1}^{T} M_{i,j^{(t)}} \;-\; 2\epsilon T \;-\; \frac{\ln m}{\epsilon},
\end{aligned}
$$

where the second inequality uses our assumption that $M_{i,j}\in[-1,1]$ for all i,j. Dividing by T and using the definition $T = (\ln m)/\epsilon^2$ proves the lemma. ■

Recall that $y^{(t)}$ was Bob's optimal stategy when Alice chooses her action according to $x^{(t)}$. So,

$$
x^{(t)\mathsf{T}} M y^{(t)} \;=\; \min_y x^{(t)\mathsf{T}} M y \;\leq\; v \qquad \forall\, t, \tag{7}
$$

by Observation 1.

Corollary 4. For any distribution $x\in\mathbb{R}^m$, we have

$$
v \;\geq\; \sum_{t=1}^{T}\frac{x^{(t)\mathsf{T}} M y^{(t)}}{T} \;\geq\; \sum_{t=1}^{T}\frac{x^{\mathsf{T}} M y^{(t)}}{T} \;-\; 3\epsilon.
$$

Proof. The upper bound follows from Eq. (7). To obtain the lower bound, we simply average Eq. (4) over all i, using coefficients x_i:

$$
\frac{\sum_{t=1}^{T} x^{(t)\mathsf{T}} M y^{(t)}}{T} \;=\; \Big(\sum_{i=1}^{m} x_i\Big)\frac{\sum_{t=1}^{T} x^{(t)\mathsf{T}} M y^{(t)}}{T} \;\geq\; \sum_{i=1}^{m} x_i\Big(\sum_{t=1}^{T}\frac{M_{i,j^{(t)}}}{T} \;-\; 3\epsilon\Big) \;=\; \sum_{t=1}^{T}\frac{x^{\mathsf{T}} M y^{(t)}}{T} \;-\; 3\epsilon.
$$

Here we have used that $y^{(t)}$ has a 1 in coordinate $j^{(t)}$ and all other coordinates 0. ■

Corollary 5.

$$\sum_{t=1}^{T} \frac{x^{(t)\mathsf{T}} M y^{(t)}}{T} \;\geq\; v - 3\epsilon.$$

Proof. Since x^* is an optimal strategy for Alice, we have $x^{*\mathsf{T}} M y^{(t)} \geq v$ for every t. Applying Corollary 4 with $x = x^*$ proves the claim. ■

Let $\hat{x} = \sum_{t=1}^{T} x^{(t)}/T$. Let y' achieve the minimum in $\min_y \hat{x}^{\mathsf{T}} M y$. Then

$$\min_y \hat{x}^{\mathsf{T}} M y \;=\; \hat{x}^{\mathsf{T}} M y' \;=\; \sum_{t=1}^{T} \frac{x^{(t)} M y'}{T} \;\geq\; \sum_{t=1}^{T} \frac{x^{(t)} M y^{(t)}}{T} \;\geq\; v - 3\epsilon,$$

since $y^{(t)}$ is Bob's optimal response to $x^{(t)}$, and by Corollary 5.

Now let $\hat{y} = \sum_{t=1}^{T} y^{(t)}/T$. Let x' achieve the maximum in $\max_x x^{\mathsf{T}} M \hat{y}$. Then Corollary 4 shows that

$$\max_x x^{\mathsf{T}} M \hat{y} \;=\; x'^{\mathsf{T}} M \hat{y} \;\leq\; v + 3\epsilon.$$

Since $\epsilon = \delta/3$, we have proven (3).

3 Proof of Von Neumann's Theorem

We used Von Neumann's Theorem above only to define the value of the game v. This was unnecessary: we could simply define $v = \max_x \min_y x^{\mathsf{T}} M y$ and the argument remains valid.

In fact, the analysis of Algorithm 1 provides a proof of Von Neumann's Theorem. Let $\hat{x}(\delta)$ and $\hat{y}(\delta)$ denote the outputs of this algorithm when run with parameter δ. Since the sets of distributions for Alice and Bob are both polytopes, they are both compact. The Bolzano-Weierstrass theorem implies that there exist limit points x^* and y^* for the sequences $\{\, \hat{x}(\delta) \,:\, \delta \to 0 \,\}$ and $\{\, \hat{y}(\delta) \,:\, \delta \to 0 \,\}$. Basic arguments with limits show that x^* and y^* satisfy Eq. (1).

Acknowledgements

These notes are based on [1]. Related papers include [2] and [3].

References

[1] S. Arora, E. Hazan, S. Kale. "The Multiplicative Weights Update Method: A Meta-Algorithm and Applications". Manuscript, 2005.

[2] Y. Freund, R. E. Schapire. "Adaptive game playing using multiplicative weights". *Games and Economic Behavior*, 29:79–103, 1999.

[3] M. Grigoriadis and L. Khachiyan. "A sublinear-time randomized approximation algorithm for matrix games". *Operations Research Letters*, 18:53–58, 1995.