

## CMSC 39600: Online Algorithms

Lecture 6

Course Instructor: Adam Kalai

Date: October 11, 2004

### The min-max theorem and weighted majority.

Today we're going to prove the min-max theorem using the result from last week, and we're going to introduce the weighted majority algorithm.

## 1 Min-max

Say we have a  $n$  by  $n$  matrix  $M$  with  $M_{i,j} \in \mathbb{R}$ . Imagine playing a two-person zero-sum game where I choose  $i \in \{1, 2, \dots, n\}$  and you choose  $j \in \{1, 2, \dots, n\}$ , and then *you* pay *me*  $M_{i,j}$  dollars. In this game, clearly it's an advantage to go second.

Now imagine that we are allowed to pick distributions over rows and columns. That is I choose  $v \in \Delta$  and you choose  $w \in \Delta$ , where

$$\Delta = \{x \in \mathbb{R}^n \mid \sum_{i=1}^n x_i = 1, x_i \geq 0\}.$$

Then you pay me  $\sum_{1 \leq i, j \leq n} w_i v_j M_{i,j}$ , the expected value of what I would get if I played according to distribution  $v$  and you played according to distribution  $w$ . The min max theorem says that there is no advantage to going second in this game. In particular, it states that,

$$\begin{aligned} a &= \max_{v \in \Delta} \min_{w \in \Delta} \sum_{1 \leq i, j \leq n} w_i v_j M_{i,j} \\ b &= \min_{w \in \Delta} \max_{v \in \Delta} \sum_{1 \leq i, j \leq n} w_i v_j M_{i,j} \implies \\ a &= b. \end{aligned}$$

In the above,  $a$  is the value that I would get if I had to choose my vector  $v$  first, because for each  $v$  that I would choose, you would choose the  $w$  that is best response, and therefore I would choose the  $v$  that minimizes your best response. Similarly,  $b$  is the value that I would get if I got to go second, because clearly I would maximize my payoff after you chose  $w$ , and you would choose the  $w$  that minimizes this maximum payoff.

Hence we can see that  $a \leq b$  for the simple reason that it is not worse to go second in this game. The surprising thing that the min-max theorem says that it doesn't matter who goes first, i.e.  $a = b$ .

Let's use Zinkevich's result from last week to prove this. Without loss of generality, we can assume that  $M_{ij} \in [-1, 1]$ . Imagine the two players are playing a repeated game, where each player repeatedly chooses a row or column. Now, given your play  $w^t$  on period  $t$ , my payoff that period is a linear function of my vector  $v$ , i.e.,

$$f^t(v) = \sum_{1 \leq i \leq n} v_i \sum_{1 \leq j \leq n} w_j M_{ij}$$

Therefore, it is also a convex function. Note that it's gradient is

$$\nabla f^t(v) = \left( \sum_{1 \leq j \leq n} w_j M_{1j}, \sum_{1 \leq j \leq n} w_j M_{2j}, \dots, \sum_{1 \leq j \leq n} w_j M_{nj} \right) \in [-1, 1]^n.$$

The last step follows from the fact that  $M_{ij} \in [-1, 1]$  and each coordinate is just a weighted average. Thus,

$$\|\nabla f^t(v)\| \leq \sqrt{n}.$$

Also, the set  $S = \Delta$  is a convex set of diameter  $\sqrt{2}$ .

So, imagine that I think of my payoffs as a sequence of convex functions  $f^1, f_2, \dots, f^T$ . Then if I use Zinkevich's online gradient *ascent* algorithm (in this case, I am trying to maximize so I would go in the direction of the gradient), by his theorem I get the following guarantee about my average payoff,

$$\begin{aligned} \text{my average payoff} &\geq (\text{the best average payoff in hindsight of a single } v \in \Delta) - \sqrt{2n/T} \\ &\geq b - \sqrt{2n/T} \end{aligned}$$

We have taken the theorem from the previous lecture, divided by  $T$ , and changed it from maximization to minimization. The second line follows from the fact that the best average payoff in hindsight is at least what I could guarantee myself if I went second, because I would best respond to his average play. No matter what his average play is, I will get at least  $a$  by choosing the best in hindsight.

Similarly, if you choose  $v$  according to Zinkevich's algorithm, you get the guarantee that,

$$\text{my average payoff} \leq a + \sqrt{2n/T}$$

Combining the above gives,

$$b - a \leq 2\sqrt{2n/T}.$$

Since this holds for all  $T$ , we have  $b - a \leq 0$ . Since we already had  $a \leq b$ , we have  $b = a$ .

## 2 Expert advice

Suppose there are  $n$  experts. Each period, each expert incurs a cost in  $[0, 1]$ . (In class, we did rewards, but costs are practically the same.) Each period, we must choose a distribution over experts, and our (expected) cost is the weighted average cost of these experts, weighted according to our distribution. Again, this could be viewed as an online convex optimization problem, and we could get a guarantee that, over  $T$  periods,

$$\text{our total cost} \leq \text{total cost of best expert} + \sqrt{2nT}.$$

This is a nice guarantee and has an average regret that approaches zero as  $T$  grows without bound. However, if the best expert has a zero total cost or there are very many experts, we can significantly improve on the bound.

A nice algorithm for this problem is the Weighted Majority algorithm. Here, we assign initial weights  $w_i^1 = 1$  for  $i = 1, 2, \dots, n$ . Each period we choose an expert at random with probability

proportional to  $w_i$ , i.e. the probability of taking expert  $i$  is  $w_i/W$  where  $W = \sum_{i=1}^n w_i$ . To update the weights, when expert  $i$  receives cost  $c_i^t$  on period  $t$ , we set its weight to be

$$w_i^{t+1} = (1 - \epsilon c_i^t) w_i^t.$$

Here,  $\epsilon$  is a parameter that we will set later. (The standard update is actually  $w_i^{t+1} = (1 + \epsilon)^{c_i^t} w_i^t$ , but the one we're using makes the analysis a bit easier.)

**Theorem 1** For any  $T \in \{1, 2, \dots\}$ ,  $\epsilon < 1/2$  and for any sequence of expert costs, the expected cost of the Weighted Majority algorithm is,

$$E[\text{cost of WM}] \leq (1 + \epsilon) \text{cost of best expert} + \frac{\log n}{\epsilon}.$$

**Proof.** Let  $W^t$  be the sum of the weights  $\sum_{i=1}^n w_i$  at period  $t$ . Let  $F^t$  be the cost that we receive at period  $t$ . Notice that our cost on period  $t$  is  $F^t = \sum \frac{w_i^t}{W^t} c_i^t$ , where  $c_i$  is the cost of the  $i$ th expert during that period. Also notice that if the sum of the weights was  $W^t$ , the new sum of weights is  $W^{t+1} = \sum w_i^t (1 - \epsilon c_i^t)$ . Thus we have

$$W^{t+1} = W^t (1 - \epsilon F^t).$$

Therefore, after  $T$  periods,

$$W^{T+1} = W^1 \prod (1 - \epsilon F^t).$$

In the following analysis we will use the fact that  $-x - x^2 \leq \log(1 - x) \leq -x$  for  $x \in [0, 1/2]$ , which can be easily verified (by graphing, for example).

Using the fact that  $W^1 = n$  and taking logs gives,

$$\log W^{T+1} = \log n + \sum_{t=1}^T \log(1 - \epsilon F^t) \leq \log n - \sum_{t=1}^T \epsilon F^t = \log n - \epsilon(\text{our total cost}).$$

For any expert  $i$

$$W^{T+1} \geq w_i^{T+1} = \prod_{t=1}^T (1 - \epsilon c_i^t)$$

Taking logs,

$$\log W^{T+1} \geq \sum \log(1 - \epsilon c_i^t) \geq - \sum \epsilon c_i^t + (\epsilon c_i^t)^2.$$

Since  $c_i^t \leq 1$ ,

$$\log W^{T+1} \geq - \sum \epsilon c_i^t + \epsilon^2 c_i^t = -(\epsilon + \epsilon^2)(\text{total cost of expert } i).$$

Putting things together, for  $i$  being the best expert in hindsight,

$$\log n - \epsilon(\text{our total cost}) \geq \log W^{T+1} \geq -(\epsilon + \epsilon^2)(\text{total cost of best expert}).$$

Dividing by  $-\epsilon$  gives the theorem. □