Algorithmic Game Theory, Summer 2025	Lecture 7 (5 pages)
Minimizing External Regret	
Thomas Kesselheim	Last Update: April 30, 2025

Last time, we got to know correlated equilibria and coarse correlated equilibria. We showed that if all players use a no-external-regret algorithm to update their strategy choices, the average history of play will converge to a coarse correlated equilibrium. The only missing piece is: How do these algorithms work?

1 Problem Statement

There is a single player playing T rounds against an adversary, trying to minimize his cost. In each round, the player chooses a probability distribution over N strategies (also termed actions here). After the player has committed to a probability distribution, or mixed strategy as we will say, the adversary picks a cost vector fixing the cost for each of the N strategies.

In round $t = 1, \dots, T$, the following happens:

- The player picks a probability distribution $p^{(t)} = (p_1^{(t)}, \dots, p_N^{(t)})$ over his strategies.
- The adversary picks a cost vector $\ell^{(t)} = (\ell_1^{(t)}, \dots, \ell_N^{(t)})$, where $\ell_i^{(t)} \in [0,1]$ for all i.
- A strategy $a^{(t)}$ is chosen according to the probability distribution $p^{(t)}$. The player incurs this strategy's cost and gets to know the entire cost vector.

What is the right benchmark for an algorithm in this setting? The *best action sequence in hindsight* achieves a cost of $\sum_{t=1}^{T} \min_{i \in [N]} \ell_i^{(t)}$. However, getting close to this number is generally hopeless as the following example shows.

Example 7.1. *Suppose $N = 2$ and consider an adversary that chooses $\ell^{(t)} = (1,0)$ if $p_1^{(t)} \geq 1/2$ and $\ell^{(t)} = (0,1)$ otherwise. Then the expected cost of the player is at least $T/2$, while the best action sequence in hindsight has cost 0.*

Instead, we will swap the sum and the minimum, and compare to $L_{\min}^{(T)} = \mathbf{E}\left[\min_{i \in [N]} \sum_{t=1}^{T} \ell_i^{(t)}\right]$. That is, instead of comparing to the best action sequence in hindsight, we compare to the *best fixed action in hindsight.* The expected cost of some algorithm $\mathcal{A}$ is given as $L_{\mathcal{A}}^{(T)} = \mathbf{E}\left[\sum_{t=1}^{T} \ell_{a^{(t)}}^{(t)}\right]$. The difference of this cost and the cost of the best single strategy in hindsight is called *external regret.*

Definition 7.2. *The* expected external regret *of algorithm $\mathcal{A}$ is defined as $R_{\mathcal{A}}^{(T)} = L_{\mathcal{A}}^{(T)} - L_{min}^{(T)}$.*

Definition 7.3. *An algorithm is called* no-external-regret algorithm *if for any adversary and all T we have $R_{\mathcal{A}}^{(T)} = o(T)$.*

This means that the *average* cost per round of a no-external-regret algorithm approaches the one of the best fixed strategy in hindsight or even beats it.

2 The Multiplicative-Weights Algorithm

By the definition it is not even clear that there are no-external-regret algorithms. Fortunately, there are. In this section, we will get to know the *multiplicative-weights algorithm* (also known as randomized weighted majority or hedge).

The algorithm maintains weights $w_i^{(t)}$, which are proportional to the probability that strategy i will be used in round t. After each round, the weights are updated by a multiplicative factor, which depends on the cost in the current round.

Let $\eta \in (0, \frac{1}{2}]$; we will choose η later.

- Initially, set $w_i^{(1)} = 1$, for every $i \in [N]$.
- At every time t,
  - Let $W^{(t)} = \sum_{i=1}^{N} w_i^{(t)}$;
  - Choose strategy i with probability $p_i^{(t)} = w_i^{(t)}/W^{(t)}$;
  - Set $w_i^{(t+1)} = w_i^{(t)} \cdot (1-\eta)^{\ell_i^{(t)}}$.

Let's build up some intuition for what this algorithm does. First suppose $\ell_i^{(t)} \in \{0,1\}$. Strategies with cost 0 maintain their weight, while the weight of strategies with cost 1 is multiplied by $(1-\eta)$. So the weight decays exponentially quickly in the number of 1's. Next consider the impact of η. Setting η to zero means that we pick a strategy uniformly at random and continue to do so, on the other hand the higher η the more we punish strategies which incurred a high cost. So we can think of η as controlling how quickly the algorithm adapts. It should neither be to low nor to high.

Theorem 7.4 (Littlestone and Warmuth, 1994)**.** *The multiplicative-weights algorithm, for any choices by the adversary of cost vectors from* $[0,1]$*, guarantees*

$$L_{MW}^{(T)} \leq (1+\eta) L_{min}^{(T)} + \frac{\ln N}{\eta} \quad .$$

Setting $\eta = \sqrt{\frac{\ln N}{T}}$ yields

$$L_{\mathrm{MW}}^{(T)} \leq L_{\mathrm{min}}^{(T)} + 2\sqrt{T \ln N} \quad .$$

Corollary 7.5. *The multiplicative-weights algorithm with* $\eta = \sqrt{\frac{\ln N}{T}}$ *has external regret at most* $2\sqrt{T \ln N} = o(T)$ *and hence is a no-external-regret algorithm.*

3 Analysis

It seems particularly difficult to analyze the algorithm because the adversary is allowed to react to the player's choices. So, effectively, $p_i^{(t)}$ and $\ell_i^{(t)}$ are random variables because they depend on which strategies were chosen before. Our first lemma carefully disentangles these effects. If $p_i^{(t)}$ and $\ell_i^{(t)}$ were not random variables, it would follow immediately from the definition of the expectation.

Lemma 7.6. *For the expected cost of the multiplicative-weights algorithm, it holds that*

$$L_{MW}^{(T)} = \mathbf{E}\left[\sum_{t=1}^{T} \sum_{i=1}^{N} p_i^{(t)} \ell_i^{(t)}\right] \quad .$$

Proof. By linearity of expectation, we have

$$L_{\mathrm{MW}}^{(T)} = \mathbf{E}\left[\sum_{t=1}^{T} \ell_{a^{(t)}}^{(t)}\right] = \sum_{t=1}^{T} \mathbf{E}\left[\ell_{a^{(t)}}^{(t)}\right]$$

Note that $\ell^{(t)}$ is the adaptive adversary's reaction to $p^{(t)}$. However, neither of the two depends on $a^{(t)}$, which is only drawn from $p^{(t)}$ after $\ell^{(t)}$ has been chosen. So, let us fix all randomization happening until exactly before $a^{(t)}$ is drawn. Let $\mathcal{E}$ denote any such event and let $p(\mathcal{E})_i^{(t)}$ and $\ell(\mathcal{E})_i^{(t)}$ be the respective realizations in $\mathcal{E}$. Then $\mathbf{E}\left[\ell_{a^{(t)}}^{(t)} \;\middle|\; \mathcal{E}\right] = \sum_{i=1}^N p(\mathcal{E})_i^{(t)} \ell(\mathcal{E})_i^{(t)}$. Taking now the expectation over $\mathcal{E}$, we get

$$\mathbf{E}\left[\ell_{a^{(t)}}^{(t)}\right] = \sum_{\mathcal{E}} \mathbf{Pr}\left[\mathcal{E}\right] \mathbf{E}\left[\ell_{a^{(t)}}^{(t)} \;\middle|\; \mathcal{E}\right] = \sum_{\mathcal{E}} \mathbf{Pr}\left[\mathcal{E}\right] \sum_{i=1}^N p(\mathcal{E})_i^{(t)} \ell(\mathcal{E})_i^{(t)} = \mathbf{E}\left[\sum_{i=1}^N p_i^{(t)} \ell_i^{(t)}\right] \quad . \qquad \square$$

Below, we will prove the following proposition, which holds regardless of the random draws inside the algorithm.

Proposition 7.7. *For every sequence $\ell^{(1)}, \ldots, \ell^{(T)}$ of cost vectors from $[0,1]$ generated by the adversary, the vectors $p^{(1)}, \ldots, p^{(T)}$ generated by MW always fulfill*

$$\sum_{t=1}^T \sum_{i=1}^T p_i^{(t)} \ell_i^{(t)} \leq (1+\eta) \min_i \sum_{t=1}^T \ell_i^{(t)} + \frac{\ln N}{\eta} \quad .$$

Before we get to the proof of this key technical proposition, let us see how it enables us to prove Theorem 7.4.

Proof of Theorem 7.4. By Lemma 7.6, we can write

$$L_{\mathrm{MW}}^{(T)} = \mathbf{E}\left[\sum_{t=1}^T \sum_{i=1}^N p_i^{(t)} \ell_i^{(t)}\right] \quad .$$

By Proposition 7.7 we furthermore always have (regardless of the outcome of any random draws)

$$\sum_{t=1}^T \sum_{i=1}^N p_i^{(t)} \ell_i^{(t)} \leq \min_i \left((1+\eta) \sum_{t=1}^T \ell_i^{(t)} + \frac{\ln N}{\eta} \right) \quad .$$

So, plugging this bound in the expression for $L_{\mathrm{MW}}^{(T)}$ and using linearity of expectation, we get

$$L_{\mathrm{MW}}^{(T)} \leq \mathbf{E}\left[\min_i \left((1+\eta) \sum_{t=1}^T \ell_i^{(t)} + \frac{\ln N}{\eta} \right)\right] = (1+\eta) \mathbf{E}\left[\min_i \sum_{t=1}^T \ell_i^{(t)}\right] + \frac{\ln N}{\eta} = (1+\eta) L_{\min}^{(T)} + \frac{\ln N}{\eta} \quad .$$

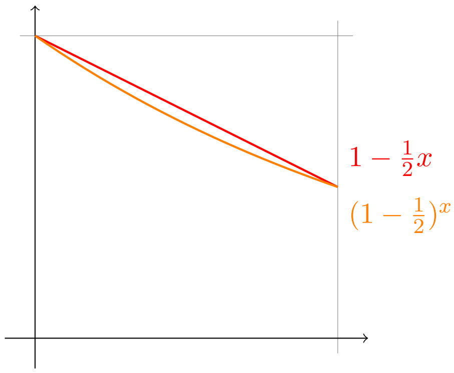
$\square$

4 Proof of Proposition 7.7

It remains to prove Proposition 7.7. To this end, we analyze how the sum of weights $W^{(t)}$ decreases over time. It holds that

$$W^{(t+1)} \;=\; \sum_{i=1}^N w_i^{(t+1)} \;=\; \sum_{i=1}^N w_i^{(t)} (1-\eta)^{\ell_i^{(t)}} \quad .$$

Observe that $(1-\eta)^x \;=\; (1-\eta x)$, for both $x = 0$ and $x = 1$. Furthermore, $x \mapsto (1-\eta)^x$ is a convex function. For $x \in [0,1]$ this implies $(1-\eta)^x \;\leq\; (1-\eta x)$.



This gives us

$$W^{(t+1)} \leq \sum_{i=1}^{N} w_i^{(t)}(1 - \ell_i^{(t)}\eta) = W^{(t)} - \eta \sum_{i=1}^{N} w_i^{(t)} \ell_i^{(t)} .$$

Let $\ell_p^{(t)} = \sum_{i=1}^{n} p_i^{(t)} \ell_i^{(t)}$. By the definition of $p_i^{(t)}$, we have $\ell_p^{(t)} = \sum_{i=1}^{N} \ell_i^{(t)} w_i^{(t)} / W^{(t)}$. Substituting this into the bound for $W^{(t+1)}$ gives

$$W^{(t+1)} \leq W^{(t)} - \eta \ell_p^{(t)} W^{(t)} = W^{(t)}(1 - \eta \ell_p^{(t)}) .$$

As a consequence,

$$W^{(T+1)} \leq W^{(1)} \prod_{t=1}^{T} (1 - \eta \ell_p^{(t)}) = N \prod_{t=1}^{T} (1 - \eta \ell_p^{(t)}) .$$

This means that the sum of weights after step T can be *upper bounded* in terms of $\ell_p^{(t)} = \sum_{i=1}^{n} p_i^{(t)} \ell_i^{(t)}$. On the other hand, the sum of weights after step T can be *lower bounded* in terms of the costs of the best strategy as follows:

$$W^{(T+1)} \geq \max_{1 \leq i \leq N} w_i^{(T+1)} = \max_{1 \leq i \leq N} \left(w_i^{(1)} \prod_{t=1}^{T} (1 - \eta)^{\ell_i^{(t)}} \right) = \max_{1 \leq i \leq N} \left((1 - \eta)^{\sum_{t=1}^{T} \ell_i^{(t)}} \right) = (1 - \eta)^{\min_i \sum_{t=1}^{T} \ell_i^{(t)}} .$$

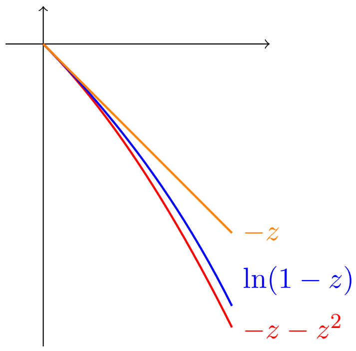
Combining the bounds and taking the logarithm on both sides gives us

$$\min_i \sum_{t=1}^{T} \ell_i^{(t)} \ln(1 - \eta) \leq (\ln N) + \sum_{t=1}^{T} \ln(1 - \eta \ell_p^{(t)}) .$$

In order to simplify, we will now use the following estimation

$$-z - z^2 \leq \ln(1 - z) \leq -z ,$$

which holds for every $z \in [0, \frac{1}{2}]$.



This gives us

$$\begin{aligned} \min_i \sum_{t=1}^{T} \ell_i^{(t)}(-\eta - \eta^2) &\leq (\ln N) + \sum_{t=1}^{T}(-\eta \ell_p^{(t)}) \\ &= (\ln N) - \eta \left(\sum_{t=1}^{T} \ell_p^{(t)} \right) . \end{aligned}$$

Finally, solving for $\sum_{t=1}^{T} \ell_p^{(t)}$ gives

$$\sum_{t=1}^{T} \ell_p^{(t)} \leq (1+\eta) \min_i \sum_{t=1}^{T} \ell_i^{(t)} + \frac{\ln N}{\eta} .$$

5 Outlook

The multiplicative-weights algorithm is only one of many no-regret algorithms. For example, they work with reduced feedback and only get to know the cost of the strategy they have chosen or the notions of regret are stronger. A much broader introduction into the topic is given in the class on *Algorithms and Uncertainty*.

Coming back to the game-theory perspective, it is interesting to see that these details do not matter too much: Our result on convergence to correlated equilibria holds regardless of which algorithm is being applied. Therefore, it is probably reasonable to assume that players indeed play an (approximate, coarse) correlated equilibrium.

Acknowledgments

Parts of these lecture notes are based on earlier versions by Berthold Vöcking, Martin Hoefer and Paul Dütting.

References and Further Reading

- Chapter 4 in the AGT book.
- Tim Roughgarden's lecture notes `http://timroughgarden.org/f13/l/l17.pdf` and lecture video `https://youtu.be/ssAEgJKRe9o`
- N. Littlestone, M. Warmuth. The Weighted Majority Algorithm. Information and Computation 108(2):212–261, 1994.