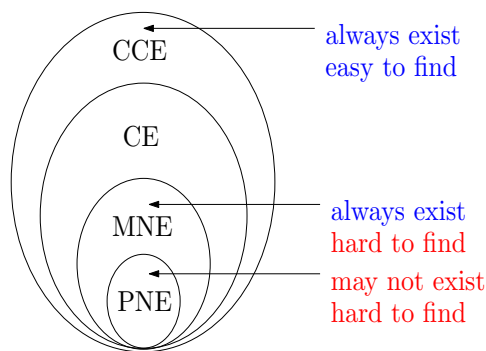


## Regret Minimization and Correlated Equilibria

## Overview

We have seen different type of equilibria and also considered the corresponding price of anarchy. These equilibria have different features:



In this lecture we show that coarse correlated equilibria (CCE) are easy to compute. We have also seen that the price of anarchy bounds obtained via the smooth framework extend to CCE equilibria. An interesting class of games are congestion games with affine delays: (1) The price of anarchy for PNE is  $5/2$ , but (2) computing PNE is such games is PLS-complete. Fortunately (3) the  $5/2$  bound on the price of anarchy holds also for CCE and (4) today we see that CCE can be computed in polynomial time (in any game).

### Structure of this lecture

- How to play against an adversary (regret minimization)
- From regret minimization to CCE (no-regret dynamics)

## 1 Regret Minimization

The next two sections will introduce the main ideas towards the general definition of regret-minimization and the algorithm. You can jump directly to Section 1.2 for the general results.

### 1.1 Experts Problem (warm up)

Consider this setting. We have  $m$  'experts' that tell us if tomorrow it will rain (R) or be sunny (S). One of these experts is a **real expert**, meaning that he/she is **never wrong**. We do not know who is the expert. Every day we make a prediction based on what the experts tell us. If our prediction is wrong, we have a cost equal to 1, otherwise we incur no cost. Here is one algorithm

**Majority Algorithm (MAJ):** Each day do the following:

- Take the majority of the experts' advice
- Every time an expert is wrong, discard him/her from future consideration;

**Claim 1.** *The number of mistakes is at most  $\log m$ , where  $m$  is the number of experts.*

*Proof.* Every mistake will half the number of experts that the algorithm takes into account.  $\square$

What if the best expert makes some mistakes?

We could *restart* the previous algorithm every time we run out of experts. If the (best) expert makes  $r^*$  errors, we are going to make at most  $r^* \log m$  errors: After each phase (we discarded all experts), the best expert must have done at least one mistake. So we cannot restart more than  $r^*$  times, and a phase will cost us at most  $\log m$  (as before).

The main idea of next algorithm is to keep a **weight** for each expert and reduce his/her weight whenever he/she was wrong.

**Weighted Majority (WM):** Each day do the following:

- $w^1(a) \leftarrow 1$  (initial weights)
- $w^{t+1}(a) \leftarrow w^t(a) \cdot \frac{1}{2}$  if  $a$  errs at step  $t$

Do weighted majority to decide  $S$  or  $R$  at step  $t$ ;

**Claim 2.** *The number of mistakes is at most*

$$2.41 C_{BEST} + 2.41 \log m,$$

where  $m$  is the number of experts and  $C_{BEST}$  is the number of mistakes of the best expert.

*Proof.* We work with the following quantities:

$$W^t := \sum_a w^t(a)$$

and show two things:

1. If the best expert does not make many mistakes, in the end  $W^T$  is not too small;
2. Every time we make an error, then  $W^t$  drops exponentially.

The intuition is that we cannot do too many mistakes, if the best expert does few mistakes. Here is the first step: every time the best expert  $a^*$  makes one mistake, we half its weight, therefore

$$W^{T+1} \geq w^{T+1}(a^*) = \underbrace{w^1(a^*)}_1 \cdot \left(\frac{1}{2}\right)^{C_{BEST}}.$$

We claim that every time we make a mistake at step  $t$ , we have

$$W^{t+1} \leq W^t \left(\frac{3}{4}\right)$$

because we will half the weights of  $W^t$  which were the weighted majority, leaving the weighted minority unchanged. Therefore, if  $r$  is the number of mistakes we make, then

$$W^{T+1} \leq \underbrace{W^1}_m \cdot \left(\frac{3}{4}\right)^r$$

Combining the two inequalities on  $W^{T+1}$  we get

$$\left(\frac{1}{2}\right)^{C_{BEST}} \leq m \cdot \left(\frac{3}{4}\right)^r$$

and taking the log on both sides we obtain

$$r \leq \underbrace{1/\log(4/3)}_{2.41} (C_{BEST} + \log m)$$

□

## 1.2 Minimizing External Regret (general setting)

Consider the following problem. 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  $m$  strategies (also termed actions here). After the player has committed to a probability distribution, the adversary picks a cost vector fixing the cost for each of the  $m$  strategies.

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

- The player picks a probability distribution  $p^t$  over his strategies.
- The adversary picks a cost vector  $c^t$ , specifying a cost  $c^t(a) \in [0, 1]$  for every strategy  $a$ .
- The player picks a strategy using his/her probability distribution  $p^t$ , and therefore has an expected cost of

$$\sum_a p^t(a) c^t(a).$$

At this point the player gets to know the entire cost vector  $c^t$ .

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

**Example 3.** Suppose  $m = 2$  and consider an adversary that chooses  $c^t = (1, 0)$  if  $p_1^t \geq 1/2$  and  $c^t = (0, 1)$  otherwise. Then the expected cost of the player is at least  $T/2$ , while the best action sequence in hindsight has cost 0.

We will instead compare with the *best fixed action* over the same period:

$$C_{BEST} := \min_a \sum_{t=1}^T c^t(a) ,$$

which is nothing but the *best fixed action in hindsight*. The algorithm  $\mathcal{A}$  used by the player to determine the distributions  $p^t$ 's has cost

$$C_{\mathcal{A}} := \sum_{t=1}^T \sum_a p^t(a) c^t(a)$$

**Definition 4.** The difference of this cost and the cost of the best single strategy in hindsight is called **external regret**,

$$R_{\mathcal{A}} := C_{\mathcal{A}} - C_{BEST}$$

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

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

$$\frac{C_{\mathcal{A}}}{T} \leq \frac{C_{BEST}}{T} + \epsilon .$$

The next example shows that there can be no deterministic no-external-regret algorithm.

**Example 5** (Randomization is necessary). Suppose there are  $m \geq 2$  actions. In each round  $t$  the algorithm commits to a strategy  $a$ . The adversary can set  $c^t(a) = 1$  and  $c^t(b) = 0$  for  $b \neq a$ . The total cost of the algorithm will be  $T$ , while the cost of the best fixed action in hindsight is at most  $T/m$ .

### 1.3 The Multiplicative-Weights Algorithm

In this section, we will get to know the *multiplicative-weights algorithm* (also known as randomized weighted majority or hedge).

**Multiplicative Weights Update Algorithm (MW):**

- $w^1(a) \leftarrow 1$ ;
- $w^{t+1}(a) \leftarrow w^t(a) \cdot (1 - \eta)^{c^t(a)}$

At time  $t$  choose strategy  $a$  with probability

$$p^t(a) = w^t(a) / W^t \quad \text{where } W^t = \sum_a w^t(a). \quad (1)$$

The algorithm maintains weights  $w^t(a)$ , which are proportional to the probability that strategy  $a$  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.

### 1.4 Analysis

The first step is to show that if the optimum has ‘large cost’ the weight  $W^T$  is also large:

$$W^T \geq (1 - \eta)^{C_{BEST}} \tag{2}$$

Here is the proof of (2): if  $a^*$  denotes the best fixed action for the costs,  $C_{BEST} = \sum_{t=1}^T c^t(a^*)$ , then

$$W^T \geq w^T(a^*) = w^1(a^*)(1 - \eta)^{c^1(a^*)}(1 - \eta)^{c^2(a^*)} \dots (1 - \eta)^{c^T(a^*)} .$$

The second step is to relate  $W^{t+1}$  to the expected cost of the algorithm at time  $t$ :

$$W^{t+1} \leq W^t(1 - \eta \cdot C_{MW}^t) \tag{3}$$

The expected cost of the algorithm at step  $t$  is

$$C_{MW}^t := \sum_a p^t(a) \cdot c^t(a) = \sum_a \frac{w^t(a)}{W^t} \cdot c^t(a) .$$

Now observe that

$$W^{t+1} = \sum_a w^{t+1}(a) = \sum_a w^t(a) \cdot (1 - \eta)^{c^t(a)} \tag{4}$$

$$\leq \sum_a w^t(a) \cdot (1 - \eta \cdot c^t(a)) \tag{5}$$

$$= W^t - \eta W^t C_{MW}^t .$$

where (4) follows from the fact that  $(1 - \eta)^x \leq (1 - \eta x)$  for  $\eta \in [0, \frac{1}{2}]$  and  $x \in [0, 1]$ .

This step of the proof gives the hypothesis:  $\eta \in [0, \frac{1}{2}]$  and costs  $c^t(a)$  in  $[0, 1]$ .

Now we compare the cost of the algorithm to the optimum:

$$(1 - \eta)^{C_{BEST}} \leq W^T \leq W^1 \prod_{t=1}^T (1 - \eta \cdot C_{MW}^t)$$

Take the logarithm on both sides

$$C_{BEST} \cdot \ln(1 - \eta) \leq \ln m + \sum_{t=1}^T \ln(1 - \eta \cdot C_{MW}^t)$$

Now we use Taylor expansion:

$$\ln(1 - x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \dots$$

in particular,  $\ln(1 - \eta) \geq -\eta - \eta^2$  because  $\eta \leq 1/2$ , and  $\ln(1 - \eta \cdot C_{MW}^t) \leq -\eta \cdot C_{MW}^t$ , thus obtaining

$$C_{BEST} \cdot (-\eta - \eta^2) \leq \ln m + \sum_{t=1}^T -\eta C_{MW}^t = \ln m - \eta \cdot C_{MW}$$

that is

$$C_{MW} \leq (1 + \eta)C_{BEST} + \frac{\ln m}{\eta} \leq C_{BEST} + \eta T + \frac{\ln m}{\eta}$$

where the inequality uses a crude upper bound  $C_{BEST} \leq T$  because  $c^t(a) \leq 1$ . Now we can optimize our parameter  $\eta$  knowing  $T$ .

For  $\eta = \sqrt{\ln m/T}$  the cost of MW satisfies

$$\frac{C_{MW}}{T} \leq \frac{C_{BEST}}{T} + 2\sqrt{\frac{\ln m}{T}}$$

To summarize we have proven the following results.

**Theorem 6** (Littlestone and Warmuth, 1994). *The multiplicative-weights algorithm, for any sequence of cost vectors from  $[0, 1]$ , guarantees*

$$C_{\mathcal{A}} \leq (1 + \eta)C_{BEST} + \frac{\ln m}{\eta} .$$

**Corollary 7.** *The multiplicative-weights algorithm with  $\eta = \sqrt{\frac{\ln m}{T}}$  has external regret at most  $2\sqrt{T \ln m} = o(T)$  and hence is a no-external-regret algorithm.*

## 2 Connection to Coarse Correlated Equilibria

Let us now connect this back to cost-minimization games. For this fix a cost-minimization game. Without loss of generality, assume that all costs are in  $[0, 1]$ . We consider *no-external-regret dynamics* defined as follows.

At each time step  $t = 1, \dots, T$ :

1. Each player  $i$  simultaneously and independently chooses a mixed strategy  $\sigma_i^t$  using a no-external-regret algorithm  $\mathcal{A}$ .
2. Each player  $i$  receives a cost vector  $c_i^t$ , where  $c_i^t(s_i)$  is the expected cost of strategy  $s_i$  when the other players play their chosen mixed strategies:

$$c_i^t(s_i) := \mathbf{E}_{s_{-i} \sim \sigma_{-i}} [c_i(s_i, s_{-i})].$$

Do such dynamics converge to Nash equilibria? Not necessarily. However, “on average” the players play according to an approximate coarse correlated equilibrium.

**Proposition 8.** *Let  $\sigma^1, \dots, \sigma^T$  be generated by no-external-regret dynamics such that each player’s external regret is at most  $\epsilon T$ . Let  $p$  be the probability distribution that first selects a single  $t \in [T]$  uniformly at random and then chooses for every player  $i$  one  $s_i$  according to  $\sigma_i^t$ . Then  $p$  is an  $\epsilon$ -coarse correlated equilibrium.*

*Proof.* By definition, for each player  $i$ ,

$$\mathbf{E}_{s \sim p}[c_i(s)] - \mathbf{E}_{s \sim p}[c_i(s'_i, s_{-i})] = \frac{1}{T} \sum_{t=1}^T (\mathbf{E}_{s \sim \sigma^t}[c_i(s)] - \mathbf{E}_{s \sim \sigma^t}[c_i(s'_i, s_{-i})]) \leq \epsilon,$$

where the inequality follows by observing that the first term in the summation is the expected cost achieved by the regret-minimization algorithm  $\mathcal{A}$  and the second term is bounded by the cost achieved by the best fixed cost in hindsight:

$$\sum_{t=1}^T \mathbf{E}_{s \sim \sigma^t}[c_i(s)] = C_{\mathcal{A}} \quad \text{and} \quad \sum_{t=1}^T \mathbf{E}_{s \sim \sigma^t}[c_i(s'_i, s_{-i})] \geq C_{BEST}. \quad (6)$$

(Note that  $C_{\mathcal{A}}$  and  $C_{BEST}$  are defined with respect to the costs  $c_i^t(\cdot)$  of the “adversary” of  $i$ , that is, the distributions of all other players.)  $\square$

**Exercise 1.** *Verify that (6) indeed holds by looking at the definition of  $C_{\mathcal{A}}$  and  $C_{BEST}$  given above.*

**Exercise 2.** *Show that an  $\epsilon$ -CCE can be computed in  $O(\frac{\ln m}{\epsilon^2})$  iterations of the dynamics above. **Hint:** use multiplicative weights update algorithm.*

## References

The material of this lecture can be also found here:

- Tim Roughgarden, *Twenty Lectures on Algorithmic Game Theory*, Cambridge University Press, 2016 (Chapter 17 and references therein).

Alternatively, see Tim Roughgarden’s lecture notes, <http://theory.stanford.edu/~tim/f13/f13.pdf>

A significant part of this notes is from previous years’ notes by Paul Dütting available here:

- [http://www.cadmo.ethz.ch/education/lectures/HS15/agt\\_HS2015/](http://www.cadmo.ethz.ch/education/lectures/HS15/agt_HS2015/)

## Exercises

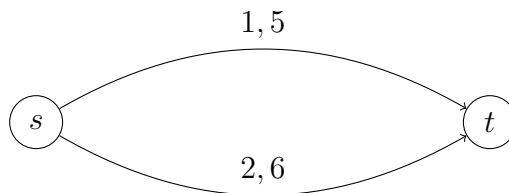
(during the exercise class - 15.10.2019)

We shall discuss and solve together these exercises.

**Exercise 3.** Each of the following statements is **false**. Your task is to disprove them (give a counterexample):

1. Suppose we have a game with no pure Nash equilibria. Then there is a mixed Nash equilibrium in which each player assigns strictly positive probability to every strategy.
2. Suppose best response converge to a pure Nash equilibrium, no matter the starting state. Then the game is a potential game.

**Exercise 4.** Consider this symmetric network congestion game with two players:



- (a) What is the price of anarchy for pure Nash equilibria?
- (b) What is the price of anarchy for mixed Nash equilibria?
- (c) What is the best price-of-anarchy bound that can be shown via smoothness?