Algorithmic Game Theory — Summer 2016, Week 3

Mixed and Correlated Equilibria and Regret Minimization

ETH Zürich — Paolo Penna

In this lecture we consider more general equilibrium concepts, namely, *mixed* and *(coarse) correlated equilibria*. We extend the definition of price of anarchy to these equilibria and study under which conditions the results on pure Nash equilibria still hold. Our main motivation is

> Pure Nash equilibria may not exist in some games, and even when they exist they are hard to compute (so it is unlikely that players will be able to always converge to one). In these cases, the bounds on the price of anarchy may be not meaningful.

This lecture resolves this 'contradiction' by showing that coarse correlated equilibria can be computed efficiently (by the players themselves).

1 Pure, Mixed, and Correlated Equilibria

We are going to study extensions of the pure Nash equilibria introduced in the previous lecture. Before giving the formal definitions, we start building some intuition by looking at simple games.

> **Pure Nash Equilibria (PNE)**: Each player chooses **one strategy** and no player has a reason to deviate.

Matching Pennies

	A	B
A	-1 / 1	1 / -1
B	1 / -1	-1 / 1

NO PNE (best response cycle)

Coordination Game

	A	B
A	1 / 2	0 / 0
B	0 / 0	1 / 1

PNE: (AA) and (BB)

> **Mixed Nash Equilibria (MNE)**: Each player chooses a **probability distribution** over his/her strategies, and no player has a reason to switch to another strategy.

The numbers in brackets are the probabilities that the player chooses the corresponding strategy.

	A $(\frac{1}{2})$	B $(\frac{1}{2})$
A $(\frac{1}{2})$	1, -1	-1, 1
B $(\frac{1}{2})$	-1, 1	1, -1

MNE (flip fair coins)

	A $(\frac{1}{3})$	B $(\frac{2}{3})$
A $(\frac{1}{2})$	2, 1	0, 0
B $(\frac{1}{2})$	0, 0	1, 1

MNE

Consider the **row player** given the probabilities used by the other player. The row player is indifferent between the two strategies that he/she is choosing randomly:

- Switch to A: utility $= 2 \times \frac{1}{3}$
- Switch to B: utility $= 1 \times \frac{2}{3}$
- Play mixed strategy '1/2-1/2': utility $= \frac{2}{3}$

More precisely, these are *expected utilities*:

$$u_i(p) := \sum_{s \in S} p(s) \cdot u_i(s) = \mathbf{E}_{s \sim p}\left[u_i(s)\right] \quad .$$

In the games above, we can say that

$$u_i(p) \geq u_i(s'_i, p_{-i})$$

for $s'_i \in \{A, B\}$, where (s'_i, p_{-i}) is the probability distribution in which i plays s'_i with probability 1.

Coarse Correleted Equilibria: A **trusted device** chooses **randomly** one state (one strategy per player), and no player has a reason to switch to another strategy:

$$u_i(p) \geq u_i(s'_i, p_{-i})$$

For cost-minimization games,

$$c_i(p) \leq c_i(s'_i, p_{-i})$$

The **trusted device** uses this distribution over the four states (numbers in brackets):

	A	B
A	2, 1 (1/3)	0, 0 (0)
B	0, 0 (0)	1, 1 (2/3)

CCE

Again, if a player decides a priory to play A (or B), his/her expected utility is not going to improve. Given that the other player(s) agree to accept the device choice, there is no reason to not do so.

Definition 1. *An* ϵ-approximate coarse correlated equilibrium *(or* ϵ-coarse correlated equilibrium*) of a cost-minimization game is a probability distribution p on the set of states S such that for every player i and every deviation $s'_i \in S_i$ we have*

$$\mathbf{E}_{s\sim p}\left[c_i(s)\right] \leq \mathbf{E}_{s\sim p}\left[c_i(s'_i, s_{-i})\right] + \epsilon \ .$$

The case of $\epsilon = 0$ is called coarse correlated equilibrium*.*

Note that the distribution p in the above definition need not be a product distribution like in mixed Nash equilibria.

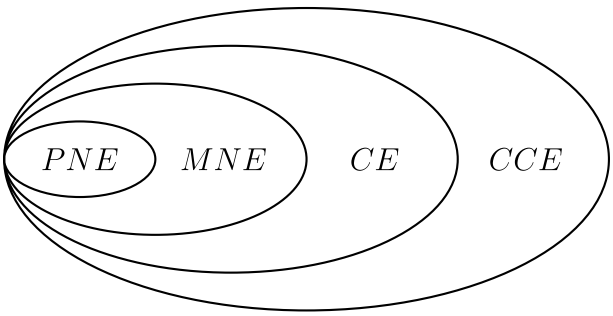
A stronger notion is that of correlated equilibria (below). Intuitively, it requires that a player still does not want to deviate even after receiving a 'signal' s_i by the trusted device (like in a traffic light, if we see 'Red" we know that the other cars crossing our street received 'Green').

Definition 2. *An* ϵ-approximate correlated equilibrium *(or* ϵ-correlated equilibrium*) of a cost-minimization game is a probability distribution p on the set of states S such that for every player i, every strategy $s_i \in S_i$, and every deviation $s'_i \in S_i$ we have*

$$\mathbf{E}_{s\sim p}\left[c_i(s) \mid s_i\right] \leq \mathbf{E}_{s\sim p}\left[c_i(s'_i, s_{-i}) \mid s_i\right] + \epsilon \ .$$

The case of $\epsilon = 0$ is called correlated equilibrium*.*

Every mixed Nash equilibrium is also a correlated equilibrium, and every correlated equilibrium is also a coarse correlated equilibrium. This leaves us with the following hierarchy of equilibrium concepts:



Unlike pure Nash equilibria, mixed Nash equilibria always exist:

Theorem 3 (Nash). *Every finite game has a mixed Nash equilibrium.*

We next extend the price of anarchy to these equilibria concepts. Because finding a mixed Nash equilibrium is also computationally hard, we will derive a natural algorithm for computing the most general equilibria (coarse correlated).

2 Price of Anarchy

We consider *cost-minimization* games like in the previous lecture. That is, each player i has a cost $c_i(s)$ and the *social cost* of a state s is the sum of all players' costs

$$cost(s) = \sum_i c_i(s).$$

When dealing with mixed and correlated equilibria, it is natural to consider the *expected social cost*:

$$cost(p) := \sum_{s \in S} p(s) cost(s) \;=\; \mathbf{E}_{s \sim p}[cost(s)] \;. \tag{1}$$

The Price of Anarchy compares the **worst equilibrium** with the **optimum**. In particular, we will take the worst equilibrium of a **certain type** and consider its expected cost:

Definition 4 (Price of Anarchy). *For a cost-minimization game, the* price of anarchy for Eq *is defined as*

$$PoA_{\mathsf{Eq}} = \frac{\max_{p \in \mathsf{Eq}} cost(p)}{\min_{s \in S} cost(s)} \;,$$

where $cost(p)$ is the expected social cost (1) and Eq is a set of probability distributions over the set of states S.

Observation 5. *Take $\mathsf{Eq} = \mathsf{PNE}$ and observe that this is the Price of Anarchy for pure Nash equilibria in the previous lecture.*

Observation 6. *The hierachy of equilibrium concepts says that* the price of anarchy can get worst *when we consider more general notions of equilibria:*

$$PoA_{\mathsf{PNE}} \leq PoA_{\mathsf{MNE}} \leq PoA_{\mathsf{CE}} \leq PoA_{\mathsf{CCE}} \;.$$

Recall that for congestion games with affine latency functions we have proven

$$PoA_{\mathsf{PNE}} = 5/2$$

but we also know that pure Nash equilibria are hard to compute.

What do we do with the bounds from previous lecture?

It turns out that whatever bounds we obtained with the "smooth framework", automatically extend to *all* equilibria in the hierarchy above. Recall the definition of smooth game from last lecture:

Definition 7. *A game is called* (λ, μ)-smooth *for* $\lambda > 0$ *and* $\mu < 1$ *if, for every pair of states* $s, s^* \in S$*, we have*

$$\sum_i c_i(s_i^*, s_{-i}) \leq \lambda \cdot cost(s^*) + \mu \cdot cost(s) \ .$$

Observe that this condition needs to hold for *all* states $s, s^* \in S$, as opposed to only pure Nash equilibria or only social optima. The following theorem says that the bounds for PNE obtained via this technique extend to all equilibria (in particular to the most general ones):

Theorem 8. *In a* (λ, μ)*-smooth game, the PoA for coarse correlated equilibria* (PoA_{CCE}) *is at most*

$$\frac{\lambda}{1-\mu} \ .$$

Proof Idea. The proof for pure Nash equilibria (lecture 2) can be adapted. Let s be a coarse correlated equilbrium and s^* be an optimum solution, which minimizes social cost. Then:

$$cost(p) = \mathbf{E}_{s \sim p}[cost(s)] = \mathbf{E}_{s \sim p}\left[\sum_i c_i(s)\right] \qquad \text{(definition of social cost)}$$

$$\vdots \qquad \textbf{(Exercise!)}$$

$$\leq \lambda \cdot cost(s^*) + \mu \cdot cost(p)$$

and by rearranging the terms we get

$$\frac{cost(s)}{cost(s^*)} \leq \frac{\lambda}{1-\mu}$$

for any $s \in \mathsf{CCE}$ and any social optimum s^*. That is, $PoA_{\mathsf{CCE}} \leq \frac{\lambda}{1-\mu}$. □

For congestion games with affine delay functions, PNE are hard to compute.

The above result says that the price of anarchy for coarse correlated equilibria is still 5/2 for these games. In the following sections we show that coarse correlated equilibria are easy to compute instead.

3 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 3.2 for the general results.

3.1 Experts Problem

Consider this setting. We have m 'experts' that tell us if tomorrow it will rain (R) or be sunny (S). One of these expert is 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 9. *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. □

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 10. *The number of mistakes is at most*

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

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

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

□

3.2 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 $t = 1, \ldots, T$, 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

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

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 11. *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 ha cost

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

Definition 12. *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 13 (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.*

3.3 The Multiplicative-Weights Algorithm

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

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.

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

$$p^t(a) = w^t(a)/W^t \qquad \text{where } W^t = \sum_a w^t(a). \tag{2}$$

3.4 Analysis

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

$$W^T \geq (1-\eta)^{C_{BEST}} \tag{3}$$

Here is the proof of (3): if a^* denotes the best fixed action for the costs, $C_{BEST} = \sum_{t=1}^T c^t(a^*)$, then

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

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

$$W^{t+1} \leq W^t(1-\eta \cdot C^t_{MW}) \tag{4}$$

The expected cost of the algorithm at step t is

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

Now observe that

$$\begin{aligned} W^{t+1} = \sum_a w^{t+1}(a) &= \sum_a w^t(a) \cdot (1-\eta)^{c^t(a)} \\ &\leq \sum_a w^t(a) \cdot (1-\eta \cdot c^t(a)) && (5) \\ &= W^t - \eta W^t C^t_{MW} && (6) \end{aligned}$$

where (5) 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:

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

Take the logarithm on both sides

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

Now we use Taylor expansion:

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

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

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

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 η knowing T.

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

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

To summarize we have proven the following results.

Theorem 14 (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 15. *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.*

4 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.

In each time step $t = 1, \ldots, T$:

1. Each player i simultaneously and independently chooses a mixed strategy σ_i^t using a no-external-regret algorithm.
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 16. *Let $\sigma^1, \ldots, \sigma^T$ be generated by no-external-regret dynamics such that each player's external regret is at most ϵT. 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 σ_i^t. Then p is an ϵ-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 and the second term is bounded by the cost achieved by the best fixed cost in hindsight. □

Notice that a player that uses the multiplicative-weights algorithm needs only $O(\frac{\ln m}{\epsilon^2})$ iterations to achieve the required bound on the external regret.

Recommended Literature

- Tim Roughgarden's lecture notes, http://theory.stanford.edu/~tim/f13/f13.pdf (General reference)
  - Chapter 13 for definitions and hierachy of equilibrium concepts;
  - Chapter 17 for regret-minimization algorithm;
- A. Blum and Y. Mansour. Learning, Regret Minimization, and Equilibria. In: Algorithmic Game Theory, N. Nisan et al., pages 79–101, 2007. (General reference)

- T. Roughgarden. Intrinsic Robustness of the Price of Anarchy. STOC 2009. (Smoothness Framework and PoA)

A significant part of this notes is from last year's notes by Paul Dütting available here:

- http://www.cadmo.ethz.ch/education/lectures/HS15/agt_HS2015/

The analysis of the MW algorithm follows Roughgarden's lecture notes, while the introduction via the experts problem is from here:

- http://www.cs.cmu.edu/afs/cs.cmu.edu/academic/class/15859-f11/www/notes/lecture16.pdf