CS295 Introduction to Algorithmic Game Theory

Instructor: Ioannis Panageas — Scribed by: Jingming Yan, Nikolas Patris

Lecture 10. Other Notions of Equilibria.

1 Introduction

Nash equilibrium is the central concept in Game Theory. Much of its importance and attractiveness comes from its universality: by Nash's Theorem, every finite game has at least one. The result that finding a Nash equilibrium is PPAD-complete, and therefore intractable casts this universality in doubt, since it suggests that there are games whose Nash equilibria, though existent, are in any practical sense inaccessible. Therefore, for a meaningful equilibrium analysis of games, we need to enlarge the set of equilibria. We introduce two relaxations of Nash equilibria, each more permissive and computationally tractable than the previous one. Both of these relaxed equilibrium concepts are guaranteed to exist in all finite games, similarly to the case of mixed Nash equilibria.

2 Correlated Equilibrium and Coarse Correlated Equilibrium

First, we introduce an important class of equilibiria, namely corellated equilibiria, that may regarded as a relaxation of the strictest class of Nash equilibria.

Definition 2.1 (Correlated Equilibrium) *A joint distribution μ over $S = S_1 \times \cdots S_n$ is a* correlated equilibrium *if and only if for every agent i and (deviation) strategy $x_i' \in \Delta_i$,*

$$\mathop{\mathbb{E}}_{x \backsim \mu}[u_i(x_i; x_{-i})|x_i] \geq \mathop{\mathbb{E}}_{x \backsim \mu}[u_i(x_i'; x_{-i})|x_i] \tag{1}$$

There is also a useful equivalent definition of *CE* in terms of swapping function. A *swap function* is defined as $f : S_i \to S_i$, where S_i denotes the strategy space of player i. Then, an alternative definition of CE is the following:

$$\mathop{\mathbb{E}}_{x \backsim \mu}[u_i(x_i; x_{-i})] \geq \mathop{\mathbb{E}}_{x \backsim \mu}[u_i(f(x_i); x_{-i})] \tag{2}$$

Despite the positive results of the *CE*, we can enlarge the set of equilibria even further, to an even more tractable concept; we define the set of coarse correlated equilibria, *(CCE)*.

Definition 2.2 (Coarse Correlated Equilibrium) *A joint distribution μ over $S = S_1 \times \cdots S_n$ is a* correlated equilibrium *if and only if for every agent i and (deviation) strategy $x_i' \in \Delta_i$,*

$$\mathop{\mathbb{E}}_{x \backsim \mu}[u_i(x_i; x_{-i})] \geq \mathop{\mathbb{E}}_{x \backsim \mu}[u_i(x_i'; x_{-i})] \tag{3}$$

At first glance, Definition (2.2) is the same as that for a mixed Nash equilibrium, except without the restriction that μ is a product distribution. Compared to Definition (2.1), a CCE only protects against unconditional unilateral deviations, as opposed to the unilateral deviations conditioned on x_i. Put it differently, when agent i contemplates a deviation x'_i, they know only the distribution μ and not the component x_i of the realization.

In Lemma 2.1, we examine the relationships between different sets of equilibria, see Figure 1.

Lemma 2.1 *The set of Nash equilibria is a subset of the set of correlated equilibrium, which in turn is a subset of the set of coarse Correlated Equilibrium. Specifically,*

$$NE \subseteq CE \subseteq CCE$$

Proof: Note that if μ is a product distribution, then we have $\mathop{\mathbb{E}}_{x\backsim\mu}[u_i(x_i;x_{-i})|x_i] = \mathop{\mathbb{E}}_{x\backsim\mu}[u_i(x_i;x_{-i})]$ and $\mathop{\mathbb{E}}_{x\backsim\mu}[u_i(x'_i;x_{-i})|x_i] = \mathop{\mathbb{E}}_{x\backsim\mu}[u_i(x'_i;x_{-i})]$. Thus, it proves the left-side inclusion $NE \subseteq CE$. Moreover, by the *Law of Total Expectation* we conclude

$$\mathbb{E}\big[\mathop{\mathbb{E}}_{x\backsim\mu}[u_i(x_i;x_{-i})|x_i]\big] = \mathop{\mathbb{E}}_{x\backsim\mu}[u_i(x_i;x_{-i})]$$

and

$$\mathbb{E}\big[\mathop{\mathbb{E}}_{x\backsim\mu}[u_i(x'_i;x_{-i})|s_i = x_i]\big] = \mathop{\mathbb{E}}_{x\backsim\mu}[u_i(x'_i;x_{-i})]$$

If μ is a CE, then

$$\mathop{\mathbb{E}}_{x\backsim\mu}[u_i(x_i;x_{-i})|s_i = x_i] \geq \mathop{\mathbb{E}}_{x\backsim\mu}[u_i(x'_i;x_{-i})|s_i = x_i] \quad \text{for all } x'_i$$

Finally, we conclude $\mathop{\mathbb{E}}_{x\backsim\mu}[u_i(x_i;x_{-i})] \geq \mathop{\mathbb{E}}_{x\backsim\mu}[u_i(x'_i;x_{-i})]$ which satisfy the requirement of CCE.

■

The classic example of Chicken game [1] could be proven quite useful to illustrate that notion of CCE even better and indicate a case where a CCE is not necessarily a Nash equilibrium for the game.

Example 2.2 *Consider a Chicken game defined as the following:*

	Dare	*Chicken out*
Dare	*0, 0*	*7, 2*
Chicken out	*2, 7*	*6, 6*

In each entry, the first number denotes the utility for player$_1$*, whereas the second number denotes the utility for player*$_2$*. Note that it is not a case of two-player zero-sum game. Let* μ *be a joint distribution, such that* $\mu(C,C) = \mu(C,D) = \mu(D,C) = \frac{1}{3}$, $\mu(D,D) = 0$ *is a correlated equilibrium. We could readily verify that* μ *is a CCE for the Chicken game.*

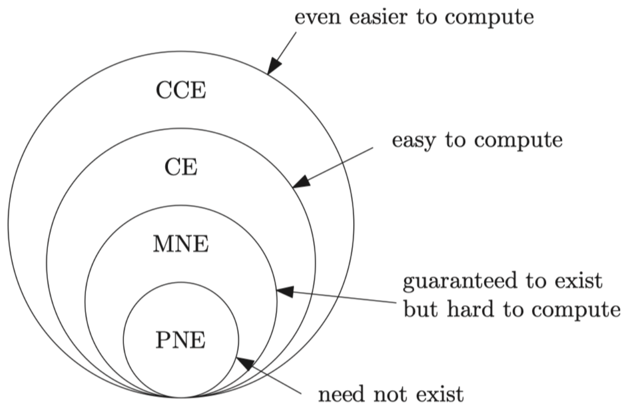

Figure 1: A hierarchy of equilibrium concepts: pure Nash equilibria (PNE), mixed Nash equilibria (MNE), correlated equilibria (CE), and coarse correlated equilibria (CCE).

Suppose that player$_1$ *is assigned with Chicken out, based on the joint distribution* χ*, player*$_2$ *is equally likely to play Chicken out or Dare if he or she follows the joint distribution, given that player*$_1$ *is assigned with Chicken out. Then, we can calculate the expected utility for player*$_1$*,*

$$\mathbb{E}[u_1(C, x_2) \,|\, C] = 2 \times \frac{1}{2} + 6 \times \frac{1}{2} = 4$$

If player$_1$ *chooses to deviate and play Dare instead, we have*

$$\mathbb{E}[u_1(D, x_2) \,|\, C] = 0 \times \frac{1}{2} + 7 \times \frac{1}{2} = 3.5$$

Note that $\mathbb{E}[u_1(C, x_2) \,|\, C] \geq \mathbb{E}[u_1(D, x_2) \,|\, C]$*. Now suppose player*$_1$ *is assigned with Dare, based on the joint distribution* μ*, player*$_2$ *can only play Chicken out if he or she follows the joint distribution. Then the utility for player*$_1$ *is*

$$\mathbb{E}[u_1(D, x_2) \,|\, D] = 7 \times 1 = 7$$

Now suppose player$_1$ *deviates and plays Chicken, we have*

$$\mathbb{E}[u_1(C, x_2) \,|\, D] = 6 \times 1 = 6$$

Again we have $\mathbb{E}[u_1(D, x_2) \,|\, D] \geq \mathbb{E}[u_1(C, x_2) \,|\, D]$*. Similar argument also apply to player*$_2$*. Therefore we claim that the joint distribution* μ *is a Correlated Equilibrium for the game. Note that since* μ *is a joint distribution, it is not a Nash Equilibrium. However, there exists Nash Equilibrium for this game, for example,* $x_1 = (1, 0)$ *and* $x_2 = (0, 1)$ *is a Nash Equilibrium.*

3 Algorithms

In this section we discuss an important family of algorithms in online learning: no-regret algorithms. Furthermore we demonstrate why no-regret algorithms efficiently compute a certain class of equilibria. To introduce no-regret algorithm, we first need to show the definition of online learning algorithm in games. Generally, an online learning algorithm can be defined as follows:

Definition 3.1 (Online Learning) *At each time step $t = 1, 2, \ldots, T$:*

- *Each player i chooses $x_i^{(t)} \in \Delta_i$, where Δ_i is the simplex over the strategies of player i.*
- *Each player experience utility $u_i(x^{(t)})$ and observe all player's strategies $x_j^{(t)}$.*

In online learning algorithm, each player i chooses $x_i^{(t)}$ in order to minimize their regret, which can be defined as following:

$$\text{Regret} = \frac{1}{T}\left[\max_{x\in\Delta_i}\sum_{t=1}^{T} u_i(x, x_{-i}^{(t)}) - \sum_{t=1}^{T} u_i(x^{(t)})\right]. \tag{4}$$

Note that the best strategy x is fixed for all $t = 1, 2, \ldots, T$, so the regret can be treated as utility gap between the chosen strategy and the best pure strategy of all time. An online decision-making algorithm $\mathcal{A}$ has no regret if for every $\epsilon > 0$ there exists a sufficiently large time horizon $T = T(\epsilon)$ such that the regret 4 is at most ϵ.

Bellow, we introduce the Projected Gradient Descent algorithm.

Definition 3.2 (Projected Gradient Descent) *Let $\ell_t : \mathcal{D} \to \mathbb{S}$ be a family of convex, differentiable, and L-Lipschitz functions over the compact set $\mathcal{S}$ with diameter D. We define* Projected Gradient Descent *as the following:*

Algorithm 1 Projected Gradient Descent

Initialize x_0
For $t = 1, 2, \ldots, T$:
  $y_t = x_t - \alpha_t \nabla \ell_t(x_t)$
  $x_{t+1} = \Pi_{\mathcal{X}}(y_t)$

---

where α_t is the step size at t, $\Pi_{\mathcal{S}}$ denotes the projection operator over $\mathcal{S}$. In the setting of online learning, we frequently adopt the notation that $\ell_t = -u_i(x^{(t)})$. We prove that Projected Gradient Descent is a no-regret algorithm.

Theorem 3.1 (Projected Gradient Descent) *For the algorithm defined in 3.2, it holds that*

$$\frac{1}{T}\left[\sum_{t=1}^{T} \ell_t(x_t) - \min_{x}\sum_{t=1}^{T} \ell_t(x)\right] \leq \frac{3}{2}\frac{LD}{\sqrt{T}}.$$

where L is the Lipschitz constant for ℓ_t, D is the diameter of set $\mathcal{S}$, and T is the total number of iterations.

Proof: Set $x^* = \arg\min(\sum \ell_t(x))$, since ℓ_t is convex, for any x_t, we have

$$\ell_t(x_t) - \ell_t(x^*) \leq \nabla \ell_t(x_t)^T(x_t - x^*). \tag{5}$$

In the Projected Gradient Descent algorithm, we have $y_t = x_t - \alpha_t \nabla \ell_t(x_t)$, therefore

$$\nabla \ell_t(x_t)^T(x_t - x^*) = \frac{1}{\alpha_t}(x_t - y_t)^T(x_t - x^*). \tag{6}$$

By Law of Consines, $||y_t - x^*||_2^2 = ||x_t - y_t||_2^2 + ||x_t - x^*||_2^2 - 2(x_t - y_t)^T(x_t - x^*)$. Therefore

$$\frac{1}{\alpha_t}(x_t - y_t)^T(x_t - x^*) = \frac{1}{2\alpha_t}(||x_t - y_t||_2^2 + ||x_t - x^*||_2^2 - ||y_t - x^*||_2^2). \tag{7}$$

Since $y_t = x_t - \alpha_t \nabla \ell_t(x_t)$,

$$\frac{1}{2\alpha_t}(||x_t - y_t||_2^2 + ||x_t - x^*||_2^2 - ||y_t - x^*||_2^2) = \frac{1}{2\alpha_t}(||x_t - x^*||_2^2 - ||y_t - x^*||_2^2) + \frac{\alpha_t}{2}||\nabla \ell_t(x_t)||_2^2. \tag{8}$$

Combine equation (5), (6), (7), (8), we have

$$\ell_t(x_t) - \ell_t(x^*) \leq \nabla \ell_t(x_t)^T(x_t - x^*) = \frac{1}{2\alpha_t}(||x_t - x^*||_2^2 - ||y_t - x^*||_2^2) + \frac{\alpha_t}{2}||\nabla \ell_t(x_t)||_2^2. \tag{9}$$

Moreover, since ℓ_t is L-Lipschitz,

$$\frac{\alpha_t}{2}||\nabla \ell_t(x_t)||_2^2 \leq \frac{\alpha_t L^2}{2}. \tag{10}$$

Since $\Pi_{\mathcal{S}}$ is the projection operator, we have

$$||y_t - x^*||_2^2 \geq ||x_{t+1} - x^*||_2^2. \tag{11}$$

From equation (9), (10), and (11) we conclude that

$$\ell_t(x_t) - \ell_t(x^*) \leq \frac{1}{2\alpha_t}(||x_t - x^*||_2^2 - ||x_{t+1} - x^*||_2^2) + \frac{\alpha_t L^2}{2}.$$

Taking the sum and consider the fact that $||x_t - x^*||_2^2 \leq D^2$, we get

$$\begin{aligned}
\sum_{t=1}^{T}(\ell_t(x_t) - \ell_t(x^*)) &\leq \sum_{t=1}^{T}(\frac{1}{2\alpha_t}(||x_t - x^*||_2^2 - ||x_{t+1} - x^*||_2^2) + \frac{\alpha_t L^2}{2}) \\
&= \sum_{t=1}^{T} ||x_t - x^*||_2^2(\frac{1}{2\alpha_t} - \frac{1}{2\alpha_{t-1}}) + \frac{L^2}{2}\sum_{t=1}^{T}\alpha_t \\
&\leq \frac{D^2}{2}\sum_{t=1}^{T}(\frac{1}{\alpha_t} - \frac{1}{\alpha_{t-1}}) + \frac{L^2}{2}\sum_{t=1}^{T}\alpha_t \\
&\leq \frac{D^2}{2\alpha_T} + \frac{L^2}{2}\sum_{t=1}^{T}\alpha_t.
\end{aligned}$$

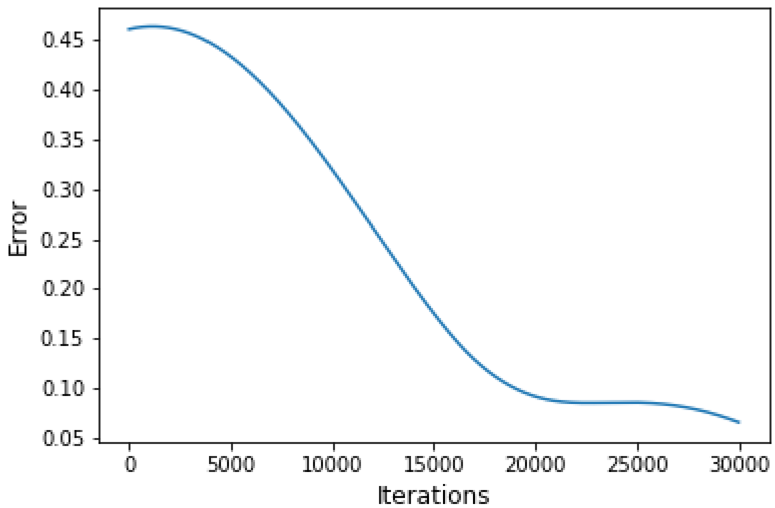

Figure 2: Error between output policy of projected gradient descent and Nash Equilibrium

Set $\alpha_t = \frac{D}{\sqrt{t}L}$ and use the fact that $\sum \frac{1}{\sqrt{t}} \leq 2\sqrt{T}$, we have

$$\begin{aligned}\sum_{t=1}^{T}(\ell_t(x_t) - \ell_t(x^*)) &\leq \frac{LD}{2}\sqrt{T} + \sum_{t=1}^{T}\frac{LD}{2\sqrt{t}} \\ &\leq \frac{LD}{2}\sqrt{T} + \frac{LD}{2}2\sqrt{T} \\ &= \frac{3}{2}\frac{LD}{\sqrt{T}}.\end{aligned}$$

■

Usually, $D = \sqrt{n}$ and we set $\alpha_t = \frac{\sqrt{n}}{\sqrt{t}L}$, we have Regret $= \frac{3}{2}\frac{L\sqrt{n}}{\sqrt{T}}$. Observe that Regret $\to 0$ as $T \to \infty$. We conclude that Projected Gradient Descent algorithm is no-regret algorithm. Moreover, let $\ell_t = -u_i(x^{(t)})$, let σ^t be the product distribution of $x^{(t)}$, let σ denote the uniform distribution over $\{\sigma^1, \sigma^2, \ldots, \sigma^T\}$, then σ is a joint distribution. Specifically, we have

$$\frac{1}{T}\sum_{t=1}^{T}\mathop{\mathbb{E}}_{s\backsim\sigma^t}[u_i(s)] = \mathop{\mathbb{E}}_{s\backsim\sigma}[u_i(s)],$$

and

$$\max_{s'\in S}\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}_{s\backsim\sigma^t}[u_i(s', s_{-i})] = \max_{s'\in S}\mathop{\mathbb{E}}_{s\backsim\sigma}[u_i(s', s_{-i})].$$

Using Theorem 3.1, we have that

$$\max_{s'\in S}\mathop{\mathbb{E}}_{s\backsim\sigma}[u_i(s', s_{-i})] - \mathop{\mathbb{E}}_{s\backsim\sigma}[u_i(s)] = \frac{3}{2}\frac{LD}{\sqrt{T}} \to 0 \text{ as } T \to \infty.$$

Note this matches with the definition of CCE, therefore we conclude that running Projected Gradient Descent algorithm defined in 3.2 on MARL games will result a CCE.

Remark 3.3 *It is worth noting that if we run projected gradient descent algorithm in two-player zero-sum games, the output policy will converge to Nash equilibrium. In order to verify this, we implemented an experiment using projected gradient descent over the game defined in example* **??**. *The implementation of projection follows the code from [2] and the output policy converges to a Nash Equilibrium. The error to Nash Equilibirum is plotted in figure [2].*

References

[1] Robert Sugden et al. *The economics of rights, co-operation and welfare*. Springer, 2004.

[2] Yunmei Chen and Xiaojing Ye. *Projection Onto A Simplex*. 2011. DOI: `10.48550/ARXIV.1101.6081`. URL: `https://arxiv.org/abs/1101.6081`.