CS272 - Theoretical Foundations of Learning, Decisions, and Games

Lecture 13: Correlated Equilibria

October 9, 2025

Lecturer: Nika Haghtalab *Readings: Chapter 17, Twenty Lectures on AGT (Roughgarden)*
Scribe: Juno Kim

1 Motivation

Recall that a set of distributions $(p_1, \cdots, p_n)$ with $p_i \in \Delta(A_i)$ is a mixed Nash equilibrium if for all $i \in [n]$ and $p'_i \in \Delta(A_i)$,

$$\mathbb{E}\left[u_i(a_1, \cdots, a_n)\right] \geq \mathop{\mathbb{E}}_{a'_i \sim p'_i}\left[u_i(a_1, \cdots, a_{i-1}, a'_i, a_{i+1}, \cdots, a_n)\right]$$

where $a_1 \sim p_1, \cdots, a_n \sim p_n$. In this formulation, crucially, each agent samples actions from their own distribution p_i independently; in other words, the set of actions is drawn from the product distribution $p_1 \times \cdots \times p_n$. This requires that there is no coordination, shared understanding or any other correlation between the agents. However, real-life scenarios often include various forms of correlation; examples include agents reading the same newspaper, watching the same stock market history, or sharing cultural norms.

We illustrate the importance of studying such correlations via the *traffic light game*. Consider the payoff matrix for two drivers defined as:

	go	stop
go	$-100, -100$	$5, 0$
stop	$0, 5$	$0, 0$

That is, if one driver goes while the other stops, only the former receives reward 5, while if both stop, both receive no reward; but if both choose to go, they will collide and receive a large negative reward. For this game, it is easy to see that there are two pure Nash equilibria (NE): (go, stop) and (stop, go). Perhaps less obviously, there also exists a mixed Nash equilibrium (MNE), which can be derived as follows. Suppose the row and column players choose 'go' with probability $p, q \in [0, 1]$, respectively. The utility function of the row and column players are

$$u_1(p, q) = -100pq + 5p(1-q) = 5p - 105pq,$$
$$u_2(p, q) = -100pq + 5q(1-p) = 5q - 105pq,$$

respectively. Then setting partial derivatives $\frac{\partial u_1}{\partial p} = 5 - 105q = 0$ and $\frac{\partial u_2}{\partial q} = 5 - 105p = 0$ yields the solution $p = q = \frac{5}{105}$.

However, all three equilibria are flawed or unsatisfactory. Both pure NE are inherently unfair since one player always proceeds while the other is stopped. On the other hand, the MNE has a large probability to stop both and a nonzero chance to select (go, go) and cause an accident. The ideal solution seems to be a 50-50 mixture of the two pure NE. Of course, this solution is no longer Nash as it is not a product distribution, which leads us to the concept of correlated equilibria.

2 Correlated and Coarse Correlated Equilibria

2.1 Correlated Equilibria

We define a **correlated equilibrium** (CE) as any joint distribution π where if a mediator draws an action recommendation $(a_1, \cdots, a_n) \sim \pi$, and any player i observes their recommendation a_i, their best move will still not deviate from a_i. Formally:

Definition 2.1 (Correlated Equilibrium)**.** Denote the action space for player i by A_i and their utility function by $u_i : A \to \mathbb{R}$, where $A = A_1 \times \cdots \times A_n$. A joint distribution $\pi \in \Delta(A)$ is a CE if for all $i \in [n]$ and $a_i' \in A_i$,

$$\mathop{\mathbb{E}}_{a\sim\pi}[u_i(a) \mid a_i] \geq \mathop{\mathbb{E}}_{a\sim\pi}[u_i(a_i', a_{-i}) \mid a_i].$$

Equivalently, for all $a_i \in A_i$,

$$\mathop{\mathbb{E}}_{a_{-i}\sim\pi(\cdot|a_i)}[u_i(a_i, a_{-i})] \geq \mathop{\mathbb{E}}_{a_{-i}\sim\pi(\cdot|a_i)}[u_i(a_i', a_{-i})].$$

In the traffic light game, $\frac{1}{2}$(go, stop) $+\frac{1}{2}$(stop, go) is a CE. To see this, suppose the row player is recommended $a_1 =$ 'go.' Then the column player must have been recommended 'stop' and the expected utility is 5. In comparison, the expected utility given $a_2 =$ 'stop' with $a_1' =$ 'go' or 'stop' is $5, 0$, respectively. Now suppose the row player is recommended $a_1 =$ 'stop.' Then the column player must have been recommended 'go' and the expected utility is 0. On the other hand, the expected utility given $a_2 =$ 'go' with $a_1' =$ 'go' or 'stop' is $-100, 0$, respectively. In both cases, the recommended action is still the best action even after taking the recommendation to the other player into account. The same holds for the column player due to symmetry.

Moreover, it can also be seen that $\frac{1}{3}$(go, stop) $+\frac{1}{3}$(stop, go) $+\frac{1}{3}$(stop, stop) is a CE. In this case, if the row player is recommended $a_1 =$ 'go,' the expected utility is still 5 (the maximum). If $a_1 =$ 'stop,' however, the conditional distribution of a_2 is $\frac{1}{2}$'go'$+\frac{1}{2}$'stop' under which the expected utility is 0 for $a_1 =$ 'stop' and $\frac{-100+5}{2} = -47.5$ for the alternative $a_1 =$ 'go.'

2.2 Coarse Correlated Equilibria

We also define a closely related but distinct concept. A **coarse correlated equilibrium** (CCE) is a joint distribution π where if a mediator draws an action recommendation $(a_1, \cdots, a_n) \sim \pi$ and plays on the behalf of each player, no player will opt out. Formally:

Definition 2.2 (Coarse Correlated Equilibrium)**.** Denote the action space for player i by A_i and their utility function by $u_i : A \to \mathbb{R}$, where $A = A_1 \times \cdots \times A_n$. A joint distribution $\pi \in \Delta(A)$ is a CCE if for all $i \in [n]$ and $a_i' \in A_i$,

$$\mathop{\mathbb{E}}_{a \sim \pi}[u_i(a)] \geq \mathop{\mathbb{E}}_{a \sim \pi}[u_i(a_i', a_{-i})].$$

Note that this is a weaker notion than CE since the player does not see their recommended action; any CE is a CCE since for a CE, the player will not opt out even with more information (i.e., after conditioning on any a_i).

We illustrate the difference between CE and CCE with the classic rock-paper-scissors (RPS) game. This is a zero-sum game with the following payoff matrix for the row player:

	R	P	S
R	0	-1	$+1$
P	$+1$	0	-1
S	-1	$+1$	0

We claim that the distribution π defined as the uniform mixture of the 6 off-diagonal (non-drawing) cells, is a CCE but *not* a CE. First note that if $a \sim \pi$, each player will be recommended a uniform distribution of actions, so that $\mathbb{E}_{a\sim\pi}[u_1(a)] = 0$ and

$$\mathop{\mathbb{E}}_{a \sim \pi}[u_1(a_1', a_{-1})] = \frac{1}{3}u_1(a_1', R) + \frac{1}{3}u_1(a_1', P) + \frac{1}{3}u_1(a_1', S) = 0$$

for any a_1'; the same holds for the column player due to symmetry. Hence π is a CCE.

Now, we show that π is not a CE. Suppose the row player is recommended R. The column player must have been recommended either P or S, so that

$$\mathop{\mathbb{E}}_{a \sim \pi}[u_1(a) \mid a_1 = R] = \frac{1}{2}u_1(R, P) + \frac{1}{2}u_1(R, S) = 0.$$

However, the row player can safely swap to playing S and have positive expected utility:

$$\mathop{\mathbb{E}}_{a_2 \sim \pi}[u_1(S, a_2) \mid a_1 = R] = \frac{1}{2}u_1(S, P) + \frac{1}{2}u_1(S, S) = \frac{1}{2},$$

hence π is not a CE.

In fact, it is straightforward to verify the following string of inclusions:

$$\boxed{\text{PNE} \subsetneq \text{MNE} \subsetneq \text{CE} \subsetneq \text{CCE}}$$

3 Computing Correlated Equilibria with No-regret Dynamics

While an MNE always exists for finite games (or more generally, if the set of actions is compact and payoffs are continuous), it can generally be hard to compute. Daskalakis et al. [2009] addressed the complexity of computing Nash equilibria and prove that it is PPAD-complete — a notion of hardness weaker than NP-completeness that nevertheless implies the problem is very likely intractable. In contrast, we will see that CE and CCE are easily computed by using any no-regret algorithm. Specifically, we consider **no-regret dynamics**, formalized as follows.

Definition 3.1 (No-regret Dynamics)**.** For time $t = 1, \cdots, T$,

- Each agent i simultaneously and independently chooses a strategy $p_i^t \in \Delta(A_i)$ using a no-regret algorithm.
- For all i, the utility vector $u_i^t : A_i \to \mathbb{R}$ is defined such that $u_i^t(a)$ is the expected utility agent i recieves from playing a when all other agents play according to p_{-i}^t:

$$u_i^t(a) = \mathop{\mathbb{E}}_{a_{-i} \sim \pi_{-i}^t}[u_i(a, a_{-i})], \quad \pi_{-i}^t = \prod_{j \neq i} p_j^t.$$

- Each player observes u_i^t (full information) or $u_i^t(a^t)$ (partial information) after each round.

Theorem 3.2. *Suppose the regret of each player over T rounds is at most ϵT in expectation. Let $\pi^t = \prod_{i \in [n]} p_i^t$ and $\pi = \frac{1}{T} \sum_{t=1}^T \pi^t$. Then π is an ϵ-approximate CCE, i.e., for all $i \in [n]$ and $a_i' \in A_i$,*

$$\mathop{\mathbb{E}}_{a \sim \pi}[u_i(a)] \geq \mathop{\mathbb{E}}_{a \sim \pi}[u_i(a_i', a_{-i})] - \epsilon.$$

Note that since we are averaging the product distributions π^t over time, π is no longer a product distribution, so this approach only yields a CCE (not an MNE).

Proof. Since $u_i^t(a) = \mathbb{E}_{a_{-i} \sim \pi_{-i}^t}[u_i(a, a_{-i})]$, the left-hand side is equal to

$$\mathop{\mathbb{E}}_{a \sim \pi}[u_i(a)] = \frac{1}{T} \sum_{t=1}^T \mathop{\mathbb{E}}_{a^t \sim \pi^t}[u_i(a^t)] = \sum_{t=1}^T \mathop{\mathbb{E}}_{a_i^t \sim p^t}[u_i^t(a^t)].$$

Similarly, the right-hand side can be rewritten as

$$\mathop{\mathbb{E}}_{a \sim \pi}[u_i(a_i', a_{-i})] = \frac{1}{T} \sum_{t=1}^T \mathop{\mathbb{E}}_{a^t \sim \pi^t}[u_i(a_i', a_{-i})] = \frac{1}{T} \sum_{t=1}^T \mathop{\mathbb{E}}_{a_{-i}^t \sim \pi_{-i}^t}[u_i(a_i', a_{-i})] = \frac{1}{T} \sum_{t=1}^T u_i^t(a_i').$$

Now due to the assumption on the regret, it must hold that

$$\frac{1}{T} \mathop{\mathbb{E}}_{\{a_i^t \sim p_i^t\}_{t \in [T]}} \left[\max_{a_i' \in A_i} \sum_{t=1}^T u_i^t(a_i') - \sum_{t=1}^T u_i^t(a_i^t) \right] \leq \frac{\epsilon T}{T} = \epsilon$$

from which the assertion follows. □

Next, we aim to prove a similar result for CE. To this end, we first give an alternative definition of CE using *swap functions*, which formalizes the notion of changing to another action: $\pi \in \Delta(A)$ is a CE if for all $i \in [n]$ and all swap functions $\delta : A_i \to A_i$,

$$\mathop{\mathbb{E}}_{a \sim \pi}[u_i(a)] \geq \mathop{\mathbb{E}}_{a \sim \pi}[u_i(\delta(a_i), a_{-i})].$$

Moreover, we define:

Definition 3.3 (Swap Regret)**.** The swap regret of actions $a^1, \cdots, a^T$ with respect to a sequence of utility functions $u^1, \cdots, u^T$ is

$$\max_{\delta: A \to A} \sum_{t=1}^{T} u^t(\delta(a^t)) - \sum_{t=1}^{T} u^t(a^t).$$

Then by using a *swap* no-regret algorithm, we obtain an analogous theorem which allows us to find CE instead of CCE.

Theorem 3.4. *Suppose the* swap regret *of each player over T rounds is at most ϵT in expectation, in the dynamics defined in Definition 3.1. Let $\pi^t = \prod_{i \in [n]} p_i^t$ and $\pi = \frac{1}{T} \sum_{t=1}^{T} \pi^t$. Then π is an ϵ-approximate CE, i.e., for all $i \in [n]$ and $\delta : A_i \to A_i$,*

$$\mathop{\mathbb{E}}_{a \sim \pi}[u_i(a)] \geq \mathop{\mathbb{E}}_{a \sim \pi}[u_i(\delta(a_i), a_{-i})] - \epsilon.$$

This shows that to compute (or learn) a Correlated equilibrium, it is sufficient to design algorithms that have no swap-regret. In the next two lectures, we study *swap*-regret and algorithms for guaranteeing no-swap regret.

References

Constantinos Daskalakis, Paul W. Goldberg, and Christos H. Papadimitriou. The Complexity of Computing a Nash Equilibrium. *SIAM Journal on Computing*, 39(1):195–259, 2009.