CS 6840 Algorithmic Game Theory March 2, 2020

Lecture 16: Coarse Correlated Equilibrium

Instructor: Eva Tardos *Scribes: Odysseas Drosis, Daniel Weber*

In today's lecture, we will review previous definitions of equilibrium in games and cover the properties of a correlated equilibrium and a coarse correlated equilibrium. Then, we will discuss how no regret learning is equivalent to a coarse correlated equilibrium. Finally, we start to discuss the complexity of finding these equilibria.

Review

In previous lectures we have covered the concepts of a pure Nash equilibrium and a mixed Nash equilibrium. In a pure Nash equilibrium, every player i has a single strategy s_i that they can play such that they could not be better off by switching to any other strategy s_i'. In a mixed Nash equilibrium, every player has a probability distribution over all possible strategies such that no player would be better off in expectation if they switched to a different distribution over strategies. Consider the standard rock paper scissors game, represented by the following payoff matrix:

		Player Y		
		R	P	S
Player X	R	$(0,0)$	$(-1,1)$	$(1,-1)$
	P	$(1,-1)$	$(0,0)$	$(-1,1)$
	S	$(-1,1)$	$(1,-1)$	$(0,0)$

As discussed before in class, this game does not have a pure strategy equilibrium. However, it does have a mixed strategy equilibrium where each player chooses to play rock, paper, or scissor uniformly at random (a probability of 1/3 for each strategy).

We also discussed a variation to the standard rock, paper, scissors game where there is a large negative payoff if both players play the same strategy. This game was defined with the following payoff matrix:

		Player Y		
		R	P	S
Player X	R	$(-9,-9)$	$(-1,1)$	$(1,-1)$
	P	$(1,-1)$	$(-9,-9)$	$(-1,1)$
	S	$(-1,1)$	$(1,-1)$	$(-9,-9)$

In the previous notes, it was shown that choosing every strategy uniformly at random is still a mixed strategy equilibrium. However, the expected social welfare is -6, which is much lower than the expected social welfare of 0 for the standard game.

We used this variation on rock. paper, scissors to talk about the concept of a correlated equilibrium (CE). In a CE there is a known probability distribution σ over all possible combinations of strategies for every player (i.e. each square in the payoff matrix has some probability associated with it). The reason that a CE is correlated is because we can think of there being a coordinator that pulls a strategy vector $s = (s_1, \ldots, s_n)$ (where n =the number of players) from the distribution σ and tells each player i only strategy s_i. All players know that their strategy came from a strategy vector which came from σ. The

CE is an equilibrium because if each player i assumes that every other player does what the coordinator tells them to do, then player i should also behave and play strategy s_i.

More concretely, we showed last class that for the modified rock, paper, scissors game, making all of the off diagonal entries have a probability of 1/6 (and all of the diagonal entries having a probability of 0) was a correlated equilibrium. If Player X is told to play paper by the coordinator, then player X knows that player Y was told to either play rock or scissors, but not paper. Assuming player Y cooperates with the coordinator, player X should also cooperate with the coordinator and play paper.

Formally, the condition for σ to be a correlated equilibrium can be expressed as follows:

$$\forall i, s_i, s_i' \qquad \mathop{\mathbb{E}}_{s\sim\sigma}(c_i(s) \mid s_i) \leq \mathop{\mathbb{E}}_{s\sim\sigma}(c_i(s_i', s_{-i}) \mid s_i)$$

This formula is saying that for all players, the expected cost of doing what the coordinator told you to do given that you were told to do s_i by the coordinator is less than or equal to expected cost of changing so any other strategy s_i'.

Using the modified rock, paper, scissors game as an example, we can replace costs c_i with utilities/values v_i. If $s_i = P$, then we have:

$$\mathop{\mathbb{E}}_{s\sim\sigma}(v_i(s) \mid s_i = P) = 0$$

because half of the time the other person plays R giving a payoff of 1 and the other half the other person plays S, giving a payoff of -1.

We can explore what happens if the player misbehaves and instead changes their strategy to R:

$$\mathop{\mathbb{E}}_{s\sim\sigma}(c_i(R, v_{-i}) \mid s_i = P) = \frac{1}{2}1 + \frac{1}{2}(-9) = -4$$

Similarly if the player instead changes their strategy to S:

$$\mathop{\mathbb{E}}_{s\sim\sigma}(c_i(S, v_{-i}) \mid s_i = P) = \frac{1}{2}(-1) + \frac{1}{2}(-9) = -5$$

So behaving clearly yields the best expected payoff.

Coarse Correlated Equilibrium (CCE)

We can also ask if following the distribution σ is no worse than always following some fixed strategy s_i' no matter what the coordinator tells you. If this is true, then σ is a coarse correlated equilibrium. This condition is formally stated as:

$$\forall i, s_i' \qquad \mathop{\mathbb{E}}_{s\sim\sigma} c_i(s) \leq \mathop{\mathbb{E}}_{s\sim\sigma} c_i(s_i', s_{-i})$$

This looks almost exactly the same as the condition for correlated equilibrium, except we have removed the conditioning on the strategy s_i. If σ satisfies the above condition, due to the linearity of expectations it must be the case that σ satisfies the following inequality:

$$\forall i \qquad \mathop{\mathbb{E}}_{s\sim\sigma} c_i(s) \le \mathop{\mathbb{E}}_{\substack{s\sim\sigma \\ s'_i\sim\sigma'_i}} c_i(s'_i, s_{-i})$$

where σ'_i is a probability distribution on player i's strategies (a mixed strategy).

No regret learning converges to CCE

We will now assume that the players play some arbitrary game for multiple rounds using no regret learning to pick their strategies. Using the formula for no regret learning from a previous lecture, we get the following inequality after the game is player for T rounds:

$$\forall i, s'_i \qquad \sum_{t=1}^{T} c_i(s^t) \le \sum_{t=1}^{T} c_i(s'_i, s^t_{-i}) + 2\sqrt{T\log(n)}$$

Where s^t is the strategy vector at time t and $2\sqrt{T\log(n)}$ is the regret after T rounds for n players. The above inequality states that using no regret learning, each player i will do no worse than $2\sqrt{T\log(n)}$ more than the minimum cost fixed strategy s'_i.

We can take the above equation and divide both sides by T to get a bound on the average cost during no regret learning:

$$\forall i, s'_i \qquad \frac{1}{T}\sum_{t=1}^{T} c_i(s^t) \le \frac{1}{T}\sum_{t=1}^{T} c_i(s'_i, s^t_{-i}) + 2\sqrt{\frac{\log(n)}{T}}$$

We call σ^T the uniform random distribution over strategy vectors in the set $\{s_1, \ldots, s_T\}$. The two averages in the above inequality are equivalent to the expectation of strategy vectors draw from σ^T:

$$\forall i, s'_i \qquad \mathop{\mathbb{E}}_{s\sim\sigma^T} c_i(s^t) \le \mathop{\mathbb{E}}_{s\sim\sigma^T} c_i(s'_i, s^t_{-i}) + 2\sqrt{\frac{\log(n)}{T}}$$

As T grows, the above inequality shows that σ^T has the necessary properties to be a CCE. σ^T may not converge to a single distribution, however all of the limit points (which must exist because σ^T is a probability distribution) are CCE.

Corollary: A CCE must exist for any game as long as there are a finite number of players and a finite number of strategies for each player. This is because all players can run the no regret learning algorithm on the game as described above. All of the limit points are CCE.

Claim: CE and CCE are convex sets (for any two elements in the set, the line connecting the two elements is entirely in the set).

Proof: Given 2 CCE σ and σ', we want to show that $\alpha\sigma + (1-\alpha)\sigma'$ is also a CCE. Because σ and σ' are probability distributions, $\alpha\sigma + (1-\alpha)\sigma'$ is also a probability distribution. Next, we can unroll the expectation in the definition of CCE:

$$\forall i, s'_i \qquad \mathop{\mathbb{E}}_{s\sim\sigma} c_i(s) = \sum_s P_s * c_i(s) \le \sum_s P_s * c_i(s'_i, s_{-i})$$

Where P_s is the probability of choosing strategy s from σ. We can combine these inequalities for multiple σ and still get that the new inequality holds for the new σ, showing that $\alpha\sigma + (1-\alpha)\sigma'$ is a CCE

Linear Programs

We can also express CE and CCE as linear programs. We can express a CCE as finding an assignment to all variables P_s (where s is a strategy) such that:

$$\forall i, s'_i \qquad \mathop{\mathbb{E}}_{s\sim\sigma} c_i(s) = \sum_s P_s * c_i(s) \leq \sum_s P_s * c_i(s'_i, s_{-i})$$

and that the P_s's define a probability distribution (they sum to 1). The constraints look similar for CE. We can solve linear programs in polynomial time. However, the number of variables in both the CE and CCE constraints are k^n if there are k strategies and n players, meaning that it is not possible to find these equilibria using this method in a time polynomial in the description of the game. However, the number of constrains is polynomial, nk^2 for CE and nk for CCE, so we could maybe use the revised simplex method to maybe get a better runtime. We can always run the no regret learning algorithm to try and approximate solutions for CCE.