8803 Connections between Learning, Game Theory, and Optimization

Maria-Florina Balcan | Lecture 13: October 5, 2010

Reading: Algorithmic Game Theory book, Chapters 17, 18 and 19.

Price of Anarchy and Price of Staility

We assume a (finite) game with n players, where player i's set of possible strategies is S_i. We let $s = (s_1, \ldots, s_n)$ denote the (joint) vector of strategies selected by players in the space $S = S_1 \times \cdots \times S_n$ of joint actions.

The game assigns utilities $u_i : S \to \mathbb{R}$ or costs $u_i : S \to \mathbb{R}$ to any player i at any joint action $s \in S$: any player maximizes his utility $u_i(s)$ or minimizes his cost $c_i(s)$.

As we recall from the introductory lectures, any finite game has a mixed Nash equilibrium (NE), but a finite game may or may not have pure Nash equilibria.

Today we focus on games with pure NE. Some NE are "better" than others, which we formalize via a social objective function $f : S \to \mathbb{R}$. Two classic social objectives are:

sum social welfare $f(s) = \sum_i u_i(s)$ measures social welfare – we make sure that the average satisfaction of the population is high

maxmin social utility $f(s) = \min_i u_i(s)$ measures the satisfaction of the most unsatisfied player

A social objective function quantifies the efficiency of each strategy profile. We can now measure how efficient a Nash equilibrium is in a specific game. Since a game may have many NE we have at least two natural measures, corresponding to the best and the worst NE.

We first define the best possible solution in a game

Definition 1. *Given a social objective f and a game inducing utilities, we define the optimal solution to be $OPT = \max_{s \in S} f(s)$ (for a game inducing costs, the optimal solution is the minimum of the costs).*

Definition 2. *The Price of Anarchy (PoA) of a game G with respect to social function f is*

$$\frac{\min_{s\ NE} f(s)}{OPT} \text{ for utilities} \qquad \text{and} \qquad \frac{\max_{s\ NE} f(s)}{OPT} \text{ for costs}$$

In both cases, the price of anarchy is the ratio between the quality of the worst *NE to the quality of the optimal solution.*

If PoA is (close to) 1 then any stable state (i.e. NE) reached by players is socially good.

Definition 3. *The Price of Stability (PoS) of a game G with respect to social function f is*

$$\frac{\max_{s\ NE} f(s)}{OPT} \text{ for utilities} \qquad \text{and} \qquad \frac{\min_{s\ NE} f(s)}{OPT} \text{ for costs}$$

In both cases, the price of stability is the ratio between the quality of the best *NE to the quality of the optimal solution.*

PoS is relevant for games with some objective authority that can influence players a bit, and maybe help them "converge" to a good NE.

Congestion Games and Potential Games

We now define a class of games modeling many real world phenomena that has been extensively studied in Algorithmic Game Theory.

Definition 4. *A* congestion game *is defined by a group of resources E and a group of players. A strategy for player i is to use a subset of resources. Thus $S_i \subseteq 2^E$. For each resource $e \in E$ there is a cost (as perceived by a player) $c_e : \mathbb{N} \to \mathbb{R}$ such that $c_e(x)$ is the cost of resource e when x players are using it.*

The cost function for each player is $c_i(s) = \sum_{e \in s_i} c_e(x_e)$ where x_e is the number of players using e in s. The social function is often the sum social function

For example, in a fair cost sharing game (a well-studied congestion game), each resource $e \in E$ has some base cost C_e that is shared fairly among the players that use it, i.e. $c_e(x) = \frac{C_e}{x}$.

We now define a different class of games that we show to be equivalent to congestion games.

Definition 5. *A game G is an* exact potential game *inducing costs $c_1, \ldots, c_n$ if there exists an* exact potential function *$\Phi : S \to \mathbb{R}$ such that for every player i, strategy profile $s = (s_i, s_{-i})$ and strategy s_i', we have*

$$c_i(s_i, s_{-i}) - c_i(s_i', s_{-i}) = \Phi(s_i, s_{-i}) - \Phi(s_i', s_{-i})$$

In other words, an exact potential game has a potential function which maps joint actions to real numbers such that when player i deviates from s_i to s_i', the change in the player's cost is exactly the same as the change in the potential function.

There is an alternate definition for an *ordinal* potential game which is a bit weaker

Definition 6. *A game G inducing costs $c_1, \ldots, c_n$ is an* ordinal potential game *if there exists an* ordinal potential function *$\Phi : S \to \mathbb{R}$ such that for every player i, strategy profile $s = (s_i, s_{-i})$ and strategy s_i', we have*

$$c_i(s_i, s_{-i}) > c_i(s_i', s_{-i}) \iff \Phi(s_i, s_{-i}) > \Phi(s_i', s_{-i})$$

In other words, player i decreases its cost by deviating from s_i to s_i' if and only if the ordinal potential function also decreases (but not necessarily by the same amount).

Clearly, any exact potential game is an ordinal potential game but not the other way around.

A first appealing property of ordinal potential games is that they always have pure NE.

Theorem 1. *Every (ordinal) potential game has at least a pure Nash equilibrium, namely any joint strategy s minimizing $\Phi(s)$.*

Proof: Let s be a joint strategy minimizing $\Phi(s)$, that must exist since S is finite. If s was not an NE then there would exist some player i that can strictly lower its cost and thus strictly lowers Φ, contradiction. $\blacksquare$

A more powerful property is the convergence of best-response dynamics

Theorem 2. *In any finite potential game, best response dynamics always converge to a NE.*

Proof: The pure NE of a potential game coincide with the local minima of its potential Φ and improving moves decrease Φ. $\blacksquare$

Monderer and Shapley proved in 1996 that exact potential games and congestion games are equivalent. We prove just one direction.

Theorem 3. *Every congestion game is an* exact *potential game.*

Proof: Given a congestion game G, we will construct an exact potential function Φ for it. The difference in Φ must match the difference in any player i's utility when deviating from s_i to s'_i (assuming a fixed strategy vector s_{-i} for the other players).

Let $E^+ = \{e \in E : e \in s'_i, e \notin s_i\}$ and $E^- = \{e \in E : e \notin s'_i, e \in s_i\}$.

Since resources in both or none of E^+ and E^- do not affect cost,

$$c_i(s'_i, s_{-i}) - c_i(s_i, s_{-i}) = \sum_{e \in E^+} c_e(x_e + 1) - \sum_{e \in E^-} c_e(x_e)$$

We define $\Phi(s) = \sum_{e \in E} \sum_{j=1}^{x_e} c_e(j)$ and we can see that the difference in Φ matches the difference in payoffs of player i when switching from s_i to s'_i. $\blacksquare$

We use a potential function argument to upper bound the PoS of fair cost sharing games

Theorem 4. *The price of stability of fair cost sharing is $H(n) = 1 + \frac{1}{2} + \cdots + \frac{1}{n} = \Theta(\log n)$.*

Proof: Iterate best-response dynamics starting from an optimal solution s^*: while there is a player that can improve, pick an arbitrary such player and let him do best-response. Note that the potential always decreases and since there are finitely many states, we must reach a pure NE s_T. Since $\Phi(s) = \sum_{e \in E \text{ used}} \sum_{j=1}^{n_e(s)} \frac{C_e}{j}$, it is immediate to show that $cost(s) \leq \Phi(s) \leq cost(s) H_n$, $\forall\, s$ i.e.

$$cost(s_T) \leq \Phi(s_T) \leq \Phi(s^*) \leq cost(s^*) H_n$$

One can show that this upper bound is tight, i.e. that there exists a fair cost sharing game in which the cost of the best NE is a $\Theta(\log n)$ factor higher than that of the optimum. $\blacksquare$

One can also show that in fair cost sharing games, the price of anarchy is $\Theta(\log n)$.