CS-352 Theoretical Computer Science	October 4, 2012
Lecture 6	
Lecturer: Aleksander Mądry	*Scribes: Siddhartha Brahma, Milan Korda*

1 Introduction

Today we continue with the topic of the previous lecture – game theory. Last time we introduced a notion of Nash equilibrium and showed that all the examples of games we considered posses them. This begs to wonder: was it a coincidence or maybe every game has a Nash equilibrium? Answering this question is the main topic of this lecture. In particular, we prove two fundamental theorems in this regard: the MinMax theorem and the Nash's (Existence) theorem.

2 Existence of Nash Equilibria

Let us start with an example that shows that not every game has a Nash equilibrium. Consider the situation in Figure 1, where Apple and Microsoft want to sell their products to three clients. There are some restrictions, however. Apple can sell only to Client 1 and Client 2, whereas Microsoft can sell only to Client 2 and Client 3. Let us (rather unrealistically) assume that the clients do not care about the product vendor and will always go with the lower price. In addition, we require that there is no price discrimination, i.e., each company has to charge all their clients the same price.

Now, it is not hard to see that there is no pure[1] Nash equilibrium here. Indeed, for any two prices p_M and p_A posted by the companies, we can have two cases. Either these prices are equal, or one of them (say Microsoft's one) is strictly lower than the other. In the former case, the company that lost Client 2 due to tie-breaking, has an incentive to just slightly reduce its price to win this client over (while losing only a little bit on reduction of the price for the other client). Similarly, in the latter case, Microsoft has an incentive to bring its price even closer to (but still below) p_A to get a further increase of its profit (this increase can be minuscule, but still positive). A more involved analysis reveals that a mixed Nash equilibrium does not exist here either.

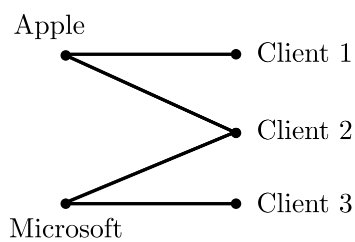

Figure 1: An illustration for the game without Nash equilibrium.

This example is a bit worrisome, as it suggests that there might be a large class of games that do not possess Nash equilibrium and thus the applicability of the theory we just developed might be severely limited. So, is it only something wrong with this particular game, or is it really that only some special type of games has Nash equilibria?

Fortunately, it turns out that the culprit here is just peculiarity of our game. Specifically, one can show that once we make our game more realistic by discretizing the range of prices – e.g., by insisting that they have to be multiples of, say, 1 centime – the reasoning we applied above does not work anymore. And, if we additionally, impose some absolute upper bound on the possible prices (which again is completely reasonable), this game will have Nash equilibrium.

[1] Recall that a Nash equilibrium is pure if it consists only of pure (i.e., deterministic) strategies.

In fact, the following fundamental theorem of game theory shows that every game has a Nash equilibrium as long as there is only *finitely many actions* and *finitely many players.*

Theorem 1 (Nash 1951) *Any game with a finite number of players and a finite set of actions has a (mixed) Nash equilibrium*

We postpone the proof of this theorem till later (see Section 2.3).

2.1 Two-Player Zero-Sum Games

Before we prove the general Nash's theorem, we focus on a simpler (but very important) case of two-player zero-sum games. Such games model a direct competition, where any gain of one player is the loss of the other one. Formally, a two-player game is *zero-sum* if the utility functions of the two players satisfy

$$\forall s \in \overline{\mathcal{S}} \;\; u_1(s) + u_2(s) = 0,$$

where we recall that $\overline{\mathcal{S}}$ is the set of the possible outcomes of the game corresponding to (mixed) strategies of both players.

Note that any such game can be conveniently represented as an $n \times m$ *payoff matrix* A, where n is the number of actions available to the first player – we will call him the *row player* (as he/she chooses rows) – and m is the number of actions offered to the second player that we will call *column player.* Now, when the row player is playing action i and column player chooses action j, the entry A_{ij} encodes the gain of the row player, and $-A_{ij}$ is the gain of the column player.

So, for example, the Penalty Kick game from the previous lecture (which is zero-sum), can be represented by a matrix

$$A = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix},$$

where the goalkeeper is the row player and the shooter is the column player.

Furthermore, one can see that in this notation, any mixed strategy[2] of the row player corresponds to a vector $x = [x_1, \ldots, x_n]^T$. On the other hand, $y = [y_1, \ldots, y_m]^T$ can encode the mixed strategy of the column player. Therefore, the expected utility of the row player corresponding to playing these two strategies is simply $x^T Ay$ for the row player and $-x^T Ay$ for the column player, as

$$x^T Ay = \sum_{i,j} x_i y_j A_{ij} = \sum_{i,j} p_{ij} A_{ij},$$

where $p_{ij} := x_i y_j$ is the probability of getting the outcome (i, j).

Thus we see that, from this point of view, the goal of the row player is to choose strategy x that leads to *maximization* of $x^T Ay$, while the goal of the column player is to choose y that *minimizes* this quantity.

Now, before the (more general) Nash's theorem was proved, the following MinMax theorem due to von Neumann was established for two-player zero-sum games.

Theorem 2 (MinMax Theorem, von Neuman, 1928) *For any two-player zero-sum game given by a matrix $A \in R^{n \times m}$, let us define*

$$V_R := \max_x \min_y x^T Ay \text{ and } V_C := \min_y \max_x x^T Ay.$$

We then have that $V_R = V_C$.
Moreover, there exists a pair of mixed strategies (x^*, y^*) which achieves the common optimum V (which is called the value *of the game) and it is a Nash equilibrium of that game.*

[2]This is a randomized strategy, where entry x_i denotes the probability of choosing i-th action, i.e., $x_i \geq 0$, and $\sum_i x_i = 1$.

Although the statement of this theorem can look a bit mysterious at first, it actually turns out to have a very intuitive interpretation in game-theoretic terms.

To see this, observe that V_R describes the best expected utility of the row player in a situation when he/she has to reveal its mixed strategy x first. That is, when the column player can choose his/her strategy y after seeing x. (Recall from the discussion above that the goal of the column player is to choose y that minimizes $x^T Ay$.) On the other hand, V_C denotes the analogous best possible expected utility of the row player if it is the column player that goes first.

Therefore, in the above context, what the MinMax theorem is telling us is that in two-person zero-sum games, it does not matter who has to reveal his/her strategy first. (It is important, however, to note here that this is true only as long as we allow declaration of mixed strategy. It is no longer so if one had required the players to reveal their pure actions after choosing them randomly based on their mixed strategy.)

At first, this statement might seem to be of only modest interest, but it actually has some very deep and important consequences that go far beyond game theory. For one, one of its ramifications is existence of Nash equilibrium in two-player zero-sum games. (On the other hand, one can also show that, more generally, the existence of Nash equilibria implies that not having to reveal one's strategy first does not give any benefit.)

2.2 Proof of the MinMax Theorem

Assume for the sake of contradiction that the theorem is not true, i.e., that there exists a two-person zero-sum game that is described by an $n \times m$ payoff matrix A and has $V_C \neq V_R$. Note that as games (and the statement of the theorem) is invariant under scaling by positive scalars and shifting of all the utilities by the same additive factor, we can assume wlog that $A_{ij} \in [0, 1]$ for all i and j.

Now, clearly, if $V_C \neq V_R$, we can't have $V_C < V_R$, as being able to reveal one's strategy after the column player does, can be only a benefit to row player. So, we just need to focus on proving that $V_C > V_R$ cannot be the case either.

To derive our desired contradiction, we will use the learning-from-expert-advice framework we introduced in Lecture 1 (see Section 6 in the notes from that lecture) to capture a process of repeated playing of our two-player zero-sum game described by A, from the perspective of the row player.

To this end, let us have n experts – one expert per each pure action of the row player – and work with gains instead of losses. (Note that our framework developed in Lecture 1 can simply model gains as negative losses.) For a given T, let us consider the following T-round instance of the learning-from-expert advice framework.

In each round $t = 1, \ldots, T$:

- We output a probability distribution $\bar{p}^t = (p_1^t, \ldots, p_n^t)$ over the experts (actions of row player).
- This distribution $\bar{p}^t$ is treated as a mixed strategy of the row player. Then, let $j_t \in \{1, \ldots, m\}$ be the (pure) strategy of the column player that is his/her best response action to row player's commitment to play $\bar{p}^t$, i.e.,
$$j^t = \arg\min_{1 \leq j \leq n} \left(\bar{p}^t\right)^T Ae_j,$$
where e_j is the m-dimensional vector having 1 at coordinate j, and 0 elsewhere.
- For each expert $1 \leq i \leq n$, his/her gain in this round to be $g_i^t := A_{ij^t}$. As a result, our gain in round t is
$$g^t := \sum_i p_i^t A_{ij^t} = \left(\bar{p}^t\right)^T Ae_j.$$

Observe that our gain g^t, in each round t, corresponds exactly to the utility of the row player when playing the mixed strategy $\bar{p}^t$ and having the column player play the pure action j^t in response. So, we can directly relate our total gain in the understanding of learning-from-expert-advice framework to the total utility we get by repeated playing of our two-player zero-sum game as a row player while going first.

In particular, the above point of view implies that we have

$$g^t \leq V_R,$$

for each t, as that's the best utility/gain we can hope for while playing our game and having to go first. (Note that from the perspective of column player, there is no benefit in randomization once he/she knows what is the strategy of the row player. So, insisting that he/she uses a pure action is not restricting him/her in any way.)

By summing over all the T rounds, we get that our total gain g, no matter how well we play, is at most

$$g := \sum_{t=1}^{T} g^t \leq T V_R. \tag{1}$$

Now, we want to compare our gain to the total gain of the best expert in the hindsight. To this end, let us define $\hat{y} \in \mathbb{R}^m$ to be

$$\hat{y}_j := \frac{\#\text{ of times } j^t = j}{T},$$

for each $1 \leq j \leq m$.

Note that this definition of $\hat{y}$ implies $\hat{y}_j \geq 0$, for every $1 \leq j \leq m$, and $\sum_j \hat{y}_j = 1$. That is, $\hat{y}$ is a probability distribution over the actions of the column player. In fact, we can view $\hat{y}$ as the empirical estimation of the mixed strategy played by the column player repeatedly throughout the whole game.

Using $\hat{y}$, we can relate the gain of the best expert to the value of V_C. Namely, we have that

$$g^* := \max_i \sum_{t=1}^{T} g_i^t = \max_i \sum_{t=1}^{T} A_{ij^t} \frac{T}{T} = T \max_i e_i^T A \hat{y} \geq T V_C, \tag{2}$$

where the last inequality follows from noticing that $\max_i e_i^T A\hat{y}$ is just the best-response utility of the row player when it is the column player that has to go first (and commits to playing $\hat{y}$), and thus it is always at least V_C.

Once we established (1) and (2), the key remaining step is to interpret these bounds appropriately. Namely, what (1) is telling us is that no matter what algorithm we use to play our game in the above learning-from-expert-advice framework, our average gain per round will be never bigger than V_R. On the other hand, (2) states that the average gain per round of the best expert in hindsight is at least V_C.

However, if we just use the multiplicative-weights-update algorithm (see Lecture 1) to repeatedly play our game in our learning-from-expert-advice setup, then the performance guarantee of this algorithm are contradicting the fact that $V_C > V_R$. Namely, recall that we proved in Lecture 1 that the performance of the MWU algorithm asymptotically achieves the performance of the best expert in hindsight.

More precisely, once we transfer the bounds from Theorem 6 in Lecture 1 from loss to gain setting (by just multiplying both its sides by -1), we have that, for any $0 < \varepsilon \leq \frac{1}{2}$, the total gain g_{MWU} of this algorithm is at least

$$g_{MWU} \geq (1 - \varepsilon) g^* - \frac{\ln n}{\varepsilon},$$

where $g^* := \max_i \sum_{t=1}^{T} g_i^t$ is the performance of the best expert in hindsight. (Note that, as $A_{ij} \in [0, 1]$, we have that $\rho = 1$ here.)

So, if we divide both sides of this performance bound by T, we will get that the average per-round gain $\bar{g}_{MWU}$ of this algorithm is at least

$$\bar{g}_{MWU} := \frac{g_{MWU}}{T} \geq (1-\varepsilon)\frac{g^*}{T} - \frac{\ln n}{\varepsilon T} = (1-\varepsilon)\bar{g}^* - \frac{\ln n}{\varepsilon T} \underset{\varepsilon \to 0}{\overset{T\to\infty}{\longrightarrow}} \bar{g}^*. \tag{3}$$

That is, it approaches the average per-round gain $\bar{g}^*$ of the best expert in hindsight, once ε tends to 0 and T tends to ∞ appropriately.

However, from (1) we know that $\bar{g}_{MWU} \leq V_C$, while from (2) we know that $\bar{g}^* \geq V_R$. So, (3) gives us a contradiction with the fact that $V_C > V_R$ and thus indeed we need to have $V_R = V_C$, as desired.

So, it just remains to formalize the above reasoning by plugging the right values of ε and T. To this end, let us define $\delta := V_C - V_R > 0$ and observe that as $A_{ij} \in [0,1]$, $V_C \in [\delta, 1]$ and $V_R \in [0, 1-\delta]$. Let us then take $\varepsilon = \frac{\delta}{2}$ and $T > \frac{4\ln n}{\delta^2}$.

Plugging these values into the bound in (3), as well as, using the fact that $V_R \leq 1-\delta$ and inequalities (1) and (2), we get that

$$V_R \geq \bar{g}_{MWU} \geq (1-\varepsilon)\bar{g}^* - \frac{\ln n}{\varepsilon T} \geq (1-\varepsilon)V_C - \frac{\ln n}{\varepsilon T} > (1-\frac{\delta}{2})(V_R+\delta) - \frac{\delta}{2} \geq V_R,$$

which is the desired contradiction.

Now, once we established that $V_R = V_C$, we prove the second part of the theorem. To this end, consider the following two mixed strategies

$$x^* = \arg\max_x \min_y x^T A y \quad \text{and} \quad y^* = \arg\min_y \max_x x^T A y.$$

Clearly, $V = V_C \leq {x^*}^T A y^* \leq V_R = V$ and thus (x^*, y^*) must be a Nash equilibrium of the game.

We remark in passing that our proof of the MinMax theorem can be easily adapted to provide us with an efficient algorithm that computes δ-approximate Nash equilibrium for two-player zero-sum games in $O(\frac{\ln n}{\delta^2})$ rounds of MWU algorithm execution. (The runtime dependence on δ is not the best possible as linear programming can be used to get an $O(\ln \frac{1}{\delta})$ dependence.)

2.3 Proof of Nash's Theorem

We now proceed to the proof of the Nash's theorem. To this end, we will need the following powerful topological result.

Theorem 3 (Brouwer's Fixed Point theorem) *Let $f : C \to C$ be a continuous function and C a convex and compact set, then there exists $x \in C$ such that $f(x) = x$.*

The point x for which $f(x) = x$ is called a *fixed point* of f.

Proof of Nash theorem. (We will prove this theorem only for two-player games. However, a version for larger number of players is just a simple extension of exactly the same approach.)

Define $C := \overline{\mathcal{S}}$, where $\overline{\mathcal{S}} = \overline{\mathcal{S}}_1 \times \overline{\mathcal{S}}_2$ is the Cartesian product of the spaces of mixed strategies each of the players. Note that the set of mixed strategies can be embedded into $\mathbb{R}^n$ where it forms a compact simplex. Since the Cartesian product of (finitely many) compact and convex sets is compact and convex, our set C satisfies the conditions from Brouwer's Fixed Point theorem.

Now, we want to find a suitable function $f : C \to C$ that is continuous and whose fixed points correspond to Nash equilibria of the underlying game. A tempting choice for such function could be a one that is defined as

$$f((s_1, s_2)) = (s_1', s_2'),$$

where s_1' is the best response strategy of Player 1 to the strategy s_2 of Player 2 and vice versa.

Clearly, if some (s_1^*, s_2^*) is a fixed point of such function it must be a Nash equilibrium. Unfortunately, this function is not well-defined. To see that, recall the Penalty Shot game that we discussed. If the row player chooses $s_1 = (1/2, 1/2)$ (left and right with the same probability) then any strategy of the column player is the best response. Furthermore, even fixing this problem would not make this function suitable, as such f is also not continuous. This is so as if we take again the example of Penalty Shot game and look at strategies $s_1 = (1/2 - \varepsilon, 1/2 + \varepsilon)$ and $s_1 = (1/2 + \varepsilon, 1/2 - \varepsilon)$ of the row player, for any $\varepsilon > 0$, then the best response to the former is $(1, 0)$, while the best response to the latter is $(0, 1)$.

Fortunately, there is an easy remedy for the above problems. It just suffices to add to f a dampening term that prevents it from deviating too rapidly. Formally, let us define f as

$$f((s_1, s_2)) = (s'_1, s'_2),$$

where

$$s'_1 := \arg \max_{s''_1 \in \overline{S}_1} u_1(s''_1, s_2) - ||s_1 - s''_1||_2^2$$

and

$$s'_2 := \arg \max_{s''_2 \in \overline{S}_2} u_2(s_1, s''_2) - ||s_2 - s''_2||_2^2.$$

It is not hard to see that fixed points of this function are still Nash equilibria, as whenever there is a strictly better response to a given strategy, one is always able to move (by some small but positive amount) in that direction. Also, now this function is continuous as the quadratic dampening terms ensures that.

So, by the Brouwer' fixed point theorem f has a fixed point and thus the underlying game has at least one Nash equilibrium. $\square$

2.4 Discussion

It should be emphasized that the Nash's theorem only asserts the *existence* of a Nash equilibrium. The proof of this theorem is highly non-constructive and does not give any hint on how to efficiently find them. As it turns out, there is a strong evidence that make us believe that finding Nash equilibria in arbitrary games is a problem that is computationally very hard. As we already mentioned in the last lecture, this is very troubling if one thinks about the underlying belief of game theory that interactions of rational agents always converge to a corresponding Nash equilibrium. After all, as Kamal Jain (a prominent researcher in algorithmic game theory) said "If your laptop can't find it, then neither can the market".