Lecture 08: Online Learning in Potential Games

Lecturer: Anas Barakat

October 09, 2025

Abstract

We introduce the celebrated class of potential games. The second part of the lecture is devoted to show how to learn approximate Nash equilibria in potential games using no-regret learning.

1 Preamble

In game theory, each agent has their own utility. The preferences of agents might be aligned or misaligned. There is a large spectrum of possible interactions, from full cooperation to full competition. In this lecture, we focus on one of the most fundamental classes of games: potential games. This class is the canonical class of common interest games. In the spectrum of games, this is the closest class to the world of optimization.

2 Potential Games

The goal of this section is to define potential games, starting with their simplest instance: identical-interest games.

2.1 Identical-interest games

All players share a common utility function, i.e.

$$\forall i, j \in \mathcal{I}, \quad \forall a \in \mathcal{A}, \quad u_i(a) = u_j(a)\,. \tag{1}$$

Example 1 (Identical interest game)**.** Consider the following 2-player identical interest game:

	b_1	b_2	b_3
a_1	$(0,0)$	$(1,1)$	$(3,3)$
a_2	$(4,4)$	$(8,8)$	$(-2,-2)$
a_3	$(6,6)$	$(7,7)$	$(2,2)$

This game has two pure NE: (a_1, b_3), (a_2, b_2) with payoffs 3 and 8 respectively.

- NE are not unique.
- NE have different payoffs, they are not equally efficient.
- Observe that (a_2, b_2) is the strategy maximizing the common payoff over $\mathcal{A}$.

Proposition 2. *In any identical interest game, if $a^\star \in \operatorname{argmax}_{a\in\mathcal{A}} u(a)$ (where u is the common payoff) then $a^\star$ is a pure NE.*

Proof. Since $a^\star \in \operatorname{argmax}_{a\in\mathcal{A}} u(a)$, we have for all $i \in \mathcal{I}, a_i \in \mathcal{A}_i$,

$$u_i(a_i, a^\star_{-i}) = u(a_i, a^\star_{-i}) \leq u(a^\star) = u_i(a^\star_i, a^\star_{-i})\,. \tag{2}$$

Therefore $a^\star$ is a NE by definition. □

2.2 From identical-interest to potential games

Consider the following modified version of the identical interest game introduced in the previous section.[1]

	y_1	y_2	y_3
x_1	$(1+a, 1+A)$	$(2+b, 2+A)$	$(4+c, 4+A)$
x_2	$(5+a, 5+B)$	$(9+b, 9+B)$	$(-1+c, -1+B)$
x_3	$(7+a, 7+C)$	$(8+b, 8+C)$	$(3+c, 3+C)$

Figure 4: An identical interest game with shifted payoffs.

Observe that this game is strategically equivalent to the previous identical interest game, i.e. both games have the exact same NE (e.g. relative preferences of players between their actions remain unchanged even if payoffs might be different).

2.3 Definition of potential games

Definition 3 (Potential game)**.** *A game $\Gamma = (\mathcal{I}, \{\mathcal{A}_i\}_{i\in\mathcal{I}}, \{\sqcap_i\}_{i\in\mathcal{I}})$ is said to be potential if there exists a function $\Phi : \mathcal{A} \to \mathbb{R}$ such that for every action profile $a \in \mathcal{A}$, agent $i \in N$, and alternative action choice $a_i' \in \mathcal{A}_i$,*

$$u_i(a_i', a_{-i}) - u_i(a_i, a_{-i}) = \Phi(a_i', a_{-i}) - \Phi(a_i, a_{-i}). \tag{3}$$

We provide a few comments regarding this definition:

- The potential game is not unique in general (just add a constant, only deviations matter).
- In a potential game, it is as if players are playing in a (hidden) identical interest game with a common payoff Φ (not necessarily known to the players). The relative preferences of any player are the same as they would be if their utility was simply Φ.
- Any profitable deviation for an agent increases the potential function. This draws an implicit link with optimization.

Remark 4. There are variants of potential games in the literature (e.g. weighted, ordinal).

Remark 5. Important examples of potential games are congestion games (Rosenthal, 1973). In fact, Monderer and Shapley (1996) have shown that potential games and congestion games are isomorphic.

Similarly to the multilinear mixed extension of the utility function, we define the mixed extension of the potential function, using again the same notation Φ (with the usual abuse of notation):

$$\Phi(x) := \mathbb{E}_{a\sim x}[\Phi(a)] = \sum_{a=(a_1,\cdots,a_N)\in\mathcal{A}} \left(\prod_{i=1}^{N} x_{i,a_i}\right) \Phi(a)\,. \tag{4}$$

We show that the potential game property of a given finite normal-form game Γ can be extended to the mixed extension of the game.

Lemma 6. *For any $x \in \prod_{i=1}^{N} \Delta(\mathcal{A}_i)$, and any $x_i' \in \Delta(\mathcal{A}_i)$,*

$$u_i(x_i', x_{-i}) - u_i(x_i, x_{-i}) = \Phi(x_i', x_{-i}) - \Phi(x_i, x_{-i})\,. \tag{5}$$

[1] The exposition and the example in this section are inspired from the lecture notes of Marden (2020).

Proof. It suffices to observe that for any $i \in \mathcal{I}, x \in \prod_{i=1}^N \Delta(\mathcal{A}_i)$, and any $x'_i \in \Delta(\mathcal{A}_i)$,

$$\begin{aligned} u_i(x'_i, x_{-i}) - u_i(x_i, x_{-i}) &= \mathbb{E}_{a'_i \sim x'_i, a_{-i} \sim x_{-i}}[u_i(a'_i, a_{-i})] - \mathbb{E}_{a_i \sim x_i, a_{-i} \sim x_{-i}}[u_i(a_i, a_{-i})] \\ &= \mathbb{E}_{a_i \sim x_i, a'_i \sim x'_i, a_{-i} \sim x_{-i}}[u_i(a'_i, a_{-i}) - u_i(a_i, a_{-i})] \\ &= \mathbb{E}_{a_i \sim x_i, a'_i \sim x'_i, a_{-i} \sim x_{-i}}[\Phi(a'_i, a_{-i}) - \Phi(a_i, a_{-i})] \\ &= \Phi(x'_i, x_{-i}) - \Phi(x)\,. \end{aligned}$$

□

2.4 Existence of pure Nash equilibria

Proposition 7. *Every potential game admits a pure NE.*

Proof. Let $a^\star \in \operatorname{argmax}_{a \in \mathcal{A}} \Phi(a)$. Then for all $i \in \mathcal{I}, a_i \in \mathcal{A}_i$,

$$u_i(a_i, a^\star_{-i}) - u_i(a^\star_i, a^\star_{-i}) = \Phi(a_i, a^\star_{-i}) - \Phi(a^\star_i, a^\star_{-i}) \le 0\,. \tag{6}$$

Hence $a^\star$ is a pure NE. □

3 Learning NE in potential games

3.1 Repeated game model

Consider a (static) finite normal-form game $\Gamma = (\mathcal{I}, \{\mathcal{A}_i\}_{i \in \mathcal{I}}, \{u_i\}_{i \in \mathcal{I}})$ where $|\mathcal{I}| = N \ge 1$. Each player i has an action space $\mathcal{A}_i$ where $\mathcal{A}_i$ is a finite set of actions, and a utility function $u_i : \prod_{i=1}^N \mathcal{A}_i \to [0,1]$ that maps an action profile $a = (a_1, \dots, a_n)$ to a utility $u_i(a)$.

We denote by $x = (x_1, \dots, x_n)$ a profile of mixed strategies, where $x_i \in \Delta(\mathcal{A}_i)$ and x_{i,a_i} is the probability of strategy $a_i \in \mathcal{A}_i$. Finally, recall the notation $u_i(x) = \mathbb{E}_{a \sim x}[u_i(a)]$, the expected utility of player i.

We consider the setting where the game Γ is played repeatedly for T time steps. At each time step t, each player i picks a mixed strategy $x_i^t \in \Delta(\mathcal{A}_i)$. At the end of the iteration, each player i observes the expected utility he would have received had he played any possible action $a_i \in \mathcal{A}_i$. More formally, let

$$u^t_{i,a_i} = \mathbb{E}_{a_{-i} \sim x^t_{-i}}[u_i(a_i, a_{-i})],$$

where x_{-i} is the set of strategies of all but the ith player, and let $u_i^t = (u^t_{i,a_i})_{a_i \in \mathcal{A}_i}$. At the end of each iteration, each player i observes u_i^t. Observe that the expected utility of a player at iteration t is the inner product $\langle x_i^t, u_i^t \rangle$.

Formally, for each player i, the regret after T time steps is equal to the maximum achievable gain by deviating to any other fixed strategy:

$$\mathrm{Reg}_i^T = \max_{\mathbf{x}_i \in \Delta(\mathcal{A}_i)} \sum_{t=1}^T \langle x_i - x_i^t, u_i^t \rangle.$$

3.2 Online gradient ascent for learning in games

At each timestep $t \ge 1$:

1. Each agent i chooses their mixed strategy $x_i^t \in \mathcal{X}_i := \Delta(\mathcal{A}_i)$.
2. Each agent i receives a gradient feedback $u_i^t = \nabla_{x_i} u_i(x_i^t, x_{-i}^t)$.

3. Each agent i updates their mixed strategy as follows:

$$x_i^{t+1} := \Pi_{\mathcal{X}_i}\left(x_i^t + \eta\, \nabla_{x_i} u_i(x_i^t, x_{-i}^t)\right), \tag{7}$$

where $\eta > 0$ is a step size.

3.3 Regret analysis

Most of the results and analysis in the remaining sections of these notes appeared in Anagnostides et al. (2022) for instance.

Proposition 8 (Potential improvement)**.** *Suppose that each player $i \in \mathcal{I}$ runs online gradient ascent (see section 3.2) with a step size $\eta = \frac{1}{L}$, where L is defined as in Lemma 9 below. Then, for any $t \geq 1$,*

$$\Phi(x^{t+1}) - \Phi(x^t) \geq \frac{1}{2\eta} \sum_{i=1}^{n} \|x_i^{t+1} - x_i^t\|_2^2 \geq 0. \tag{8}$$

Proof. The proof follows a standard optimization argument based on the smoothness of the potential function Φ.

Lemma 9. *The multilinear potential extension is L-smooth with smoothness constant $L = \Phi_{\max} \sum_{i=1}^{N} |\mathcal{A}_i|$ where $\Phi_{\max} := \max_{x \in \mathcal{X}} \Phi(x)$ and $\mathcal{X} = \prod_{i=1}^{N} \mathcal{X}_i$.*

Proof. The proof is left as an exercise. Hint: Prove that the operator norm of the Hessian of Φ is bounded by L, using the multilinearity of Φ. □

It follows from the definition of the potential function that:

$$\nabla_{x_i} u_i(x_i, x_{-i}) = \nabla_{x_i} \Phi(x_i, x_{-i}).$$

As a consequence, the OGD update rule can be equivalently rewritten as follows:

$$x^{t+1} := \Pi_{\mathcal{X}}\left(x^t + \eta \nabla \Phi(x^t)\right).$$

Using L-smoothness of Φ (Lemma 9), we can write:

$$\Phi(x^{t+1}) \geq \Phi(x^t) + \langle \nabla \Phi(x^t), x^{t+1} - x^t \rangle - \frac{L}{2}\|x^{t+1} - x^t\|^2.$$

By the characterization of the projection, we have:

$$\forall x \in \mathcal{X}, \quad \langle x - x^{t+1}, x^t + \eta \nabla \Phi(x^t) - x^{t+1} \rangle \leq 0. \tag{9}$$

Setting $x = x^t$ and rearranging the inequality yields:

$$\langle \nabla \Phi(x^t), x^{t+1} - x^t \rangle \geq \frac{1}{\eta}\|x^{t+1} - x^t\|^2. \tag{10}$$

Using the above inequality in (3.3) yields:

$$\Phi(x^{t+1}) \geq \Phi(x^t) + \left(\frac{1}{\eta} - \frac{L}{2}\right)\|x^{t+1} - x^t\|^2. \tag{11}$$

Setting $\eta = 1/L$ in the above inequality concludes the proof. □

We now control the regret incurred by each player.

Theorem 10. *The regret of each player $i \in [n]$ is such that $\mathrm{Reg}_i^T = \mathcal{O}(\sqrt{T})$.*

Proof. Using Proposition 8 (potential improvement), we have

$$\|x^{t+1} - x^t\|^2 \leq \frac{2(\Phi(x^{t+1}) - \Phi(x^t))}{L}.$$

Summing this inequality for $t = 1, \cdots, T$, we obtain:

$$\sum_{t=1}^{T} \|x^{t+1} - x^t\|^2 \leq \frac{2}{L}\left(\Phi_{\max} - \Phi(x^1)\right). \tag{12}$$

It follows from the OGD per-player update rule (see eq. (7)) and the characterization of the projection that for all $t \geq 1$:

$$\forall x_i \in \mathcal{X}_i, \quad \langle x_i - x_i^t, x_i^t + \eta\, u_i^t - x_i^{t+1}\rangle \leq 0\,, \tag{13}$$

which can be rewritten as follows:

$$\forall x_i \in \mathcal{X}_i, \quad \langle x_i - x_i^{t+1}, u_i^t\rangle \leq \frac{1}{\eta}\langle x_i - x_i^{t+1}, x_i^{t+1} - x_i^t\rangle\,, \tag{14}$$

The utility deviation under unilateral strategy deviation for player $i \in \mathcal{I}$ can be upper bounded as follows:

$$\begin{aligned}
u_i(x_i, x_{-i}^t) - u_i(x_i^t, x_{-i}^t) &= \langle x_i - x_i^t, u_i^t\rangle \\
&= \langle x_i - x_i^{t+1}, u_i^t\rangle + \langle x_i^{t+1} - x_i^t, u_i^t\rangle \\
&\leq \frac{1}{\eta}\langle x_i - x_i^{t+1}, x_i^{t+1} - x_i^t\rangle + \langle x_i^{t+1} - x_i^t, u_i^t\rangle \\
&\leq \left(\frac{\Omega}{\eta} + U_{\max}\right)\|x_i^{t+1} - x_i^t\|\,,
\end{aligned}$$

where we used (14) in the first inequality and where $\Omega := \max_{i\in\mathcal{I}} \operatorname{diam}(\mathcal{X}_i), \operatorname{diam}(\mathcal{X}_i) := \max_{x_i, x_i' \in \mathcal{X}_i} \|x_i - x_i'\|$ and $U_{\max}$ is a uniform bound on $\max_{i\in\mathcal{I}} \max_{t\in\{1,\ldots,T\}} \|u_i^t\|$ (independent of T) which follows from boundedness of the utility functions.

Summing up the above inequality for $t = 1, \ldots, T$ and using the path length bound (12) together with the Cauchy Schwarz inequality, we obtain:

$$\begin{aligned}
\mathrm{Reg}_i^T &= \max_{\mathbf{x}_i \in \Delta(\mathcal{A}_i)} \sum_{t=1}^{T} \langle x_i - x_i^t, u_i^t\rangle \\
&\leq \left(\frac{\Omega}{\eta} + U_{\max}\right)\sum_{t=1}^{T}\|x_i^{t+1} - x_i^t\| \\
&\leq \left(\frac{\Omega}{\eta} + U_{\max}\right)\sqrt{T}\sqrt{\sum_{t=1}^{T}\|x^{t+1} - x^t\|^2} \\
&\leq \left(\frac{\Omega}{\eta} + U_{\max}\right)\sqrt{\frac{2}{L}\left(\Phi_{\max} - \Phi(x^1)\right)}\sqrt{T}\,,
\end{aligned}$$

which concludes the proof. $\square$

3.4 Convergence to approximate Nash equilibria

Theorem 11. *Suppose that each player $i \in \mathcal{I}$ runs online gradient ascent (see section 3.2) with a step size $\eta = \frac{1}{L}$. Then for any $\varepsilon > 0$, after $T = \mathcal{O}(\varepsilon^{-2})$ iterations, there exists a strategy x^t (for some $t \in \{1, \cdots, T\}$) which is an ε-approximate NE.*

Proof. The proof follows similar lines to the proof of the regret bound of Theorem 10. We provide a complete proof. Using Proposition 8 (potential improvement), we have

$$\|x^{t+1} - x^t\|^2 \le \frac{2(\Phi(x^{t+1}) - \Phi(x^t))}{L}.$$

Summing this inequality for $t = 1, \cdots, T$, we obtain:

$$\sum_{t=1}^{T} \|x^{t+1} - x^t\|^2 \le \frac{2}{L}\left(\Phi_{\max} - \Phi(x^0)\right).$$

Thus, there exists $t^\star \in \{1, \cdots, T\}$ such that

$$\|x_i^{t^\star+1} - x_i^{t^\star}\| \le \|x^{t^\star+1} - x^{t^\star}\| \le \sqrt{\frac{2\left(\Phi_{\max} - \Phi(x^0)\right)}{LT}}.$$

It follows from the OGD per-player update rule (see eq. (7)) and the characterization of the projection that:

$$\forall x_i \in \mathcal{X}_i, \quad \langle x_i - x_i^{t^\star+1}, x_i^{t^\star} + \eta \nabla_{x_i} u_i(x^{t^\star}) - x_i^{t^\star+1}\rangle \le 0\,, \tag{15}$$

which can be rewritten as follows:

$$\forall x_i \in \mathcal{X}_i, \quad \langle x_i - x_i^{t^\star+1}, \nabla_{x_i} u_i(x^{t^\star})\rangle \le \frac{1}{\eta}\langle x_i - x_i^{t^\star+1}, x_i^{t^\star+1} - x_i^{t^\star}\rangle\,, \tag{16}$$

The utility deviation under unilateral strategy deviation for player $i \in \mathcal{I}$ can be upper bounded as follows:

$$\begin{aligned}
u_i(x_i, x_{-i}^{t^\star}) - u_i(x_i^{t^\star}, x_{-i}^{t^\star}) &= \langle x_i - x_i^{t^\star}, u_i(\cdot, x_{-i}^{t^\star})\rangle \\
&= \langle x_i - x_i^{t^\star+1}, u_i(\cdot, x_{-i}^{\star})\rangle + \langle x_i^{t^\star+1} - x_i^{t^\star}, u_i(\cdot, x_{-i}^{\star})\rangle \\
&\le \frac{1}{\eta}\langle x_i - x_i^{t^\star+1}, x_i^{t^\star+1} - x_i^{t^\star}\rangle + \langle x_i^{t^\star+1} - x_i^{t^\star}, u_i(\cdot, x_{-i}^{\star})\rangle \\
&\le \left(\frac{\Omega}{\eta} + U_{\max}\right)\|x_i^{t^\star+1} - x_i^{t^\star}\| \\
&\le \left(\frac{\Omega}{\eta} + U_{\max}\right)\sqrt{\frac{2\left(\Phi_{\max} - \Phi(x^0)\right)}{LT}}\,,
\end{aligned}$$

where $\Omega = \max_{i\in\mathcal{I}} \mathrm{diam}(\mathcal{X}_i), \mathrm{diam}(\mathcal{X}_i) = \max_{x_i, x_i' \in \mathcal{X}_i} \|x_i - x_i'\|$ and $U_{\max} := \max_{i\in\mathcal{I}} \|u_i(\cdot, x_{-i}^{t^\star})\|$.

It remains to select T such that the above upper bound is smaller than ε to obtain an ε-NE, i.e. $T = \mathcal{O}(\varepsilon^{-2})$. □

Remark 12. The proof can be adapted to the case of continuous games with concave utilities.

4 Next lecture

2-player zero-sum games, minmax theorem, online learning proof, online learning in zero-sum games for approximate NE computation.

References

Ioannis Anagnostides, Ioannis Panageas, Gabriele Farina, and Tuomas Sandholm. On last-iterate convergence beyond zero-sum games. In *Proceedings of the 39th International Conference on Machine Learning*, pages 536–581, 2022. URL https://proceedings.mlr.press/v162/anagnostides22a.html.

Jason R. Marden. Game theory and multiagent systems (ece 149), lectures 11, 12, 14. https://web.ece.ucsb.edu/~jrmarden/ece149.html, 2020.

Dov Monderer and Lloyd S Shapley. Potential games. *Games and economic behavior*, 14(1): 124–143, 1996.

Robert W Rosenthal. A class of games possessing pure-strategy nash equilibria. *International Journal of Game Theory*, 2(1):65–67, 1973.