

Distributed learning in potential games over large-scale networks

Fabio Fagnani, DISMA, Politecnico di Torino

joint work with

Giacomo Como, Lund University

Sandro Zampieri, DEI, University of Padova

Networking seminars

INRIA Sophia-Antipolis

Jan 29, 2014

Introduction

A central issue in game theory: provide **sound dynamical foundations for Nash equilibria**.

Possible approach: **evolutionary game theory** or the theory of **learning** in games. Agents compare their current performance with that of other agents and update consequently their state: **imitative best response**.

In the literature: **mean field** assumption, all agents can interact with each other.

In this talk: a social network constraining **interactions along the edges of a preassigned graph**.

Interacting agent networked model: **agents are doubly coupled** through the game rewards and through the interaction graph.

Outline

1. Potential game with n agents. Nash equilibria.
2. Imitative noisy best response dynamics over a graph.
3. An ergodic Markov chain. Analysis of the invariant probability measure.
4. Double limit: $n \rightarrow +\infty$ and noise to 0. Recovering the Nash equilibria as concentration point of the invariant probability.
5. Conclusions and further steps.

A case study

$\mathcal{V} = \{1, \dots, n\}$ finite set of agents; each agent has two possible options $\mathcal{X} = \{0, 1\}$.

$$x \in \mathcal{X}^n \quad z(x) := \frac{1}{n} |\{v \in \mathcal{V} \mid x_v = 1\}|$$

Reward for an agent playing option $0, 1 \in \mathcal{X}$:

$$\begin{aligned} r_0(z(x)) &= z(x) + \alpha_0 \\ r_1(z(x)) &= 1 - z(x) + \alpha_1 \end{aligned} \tag{1}$$

where $\alpha_0, \alpha_1 \in \mathbb{R}$ with $|\alpha_0 - \alpha_1| < 1$.

Reward of an agent decreases as more agents choose the same option, because this will make the resource less available

Nash equilibrium:

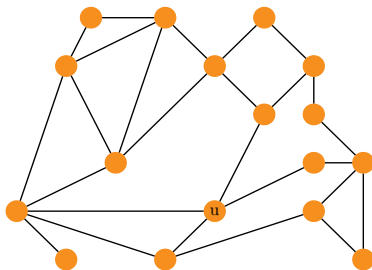
$$r_1(z) = r_0(z) \Leftrightarrow z = z_{\text{Nash}} := \frac{1 - \alpha_0 + \alpha_1}{2}$$

A case study

$\mathcal{G} = (\mathcal{V}, \mathcal{E})$ connected undirected graph.

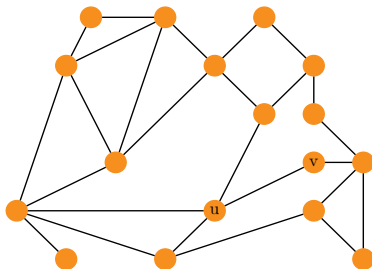
Dynamics: Markov chain $X(t)$ on \mathcal{X}^n . At each time t ,

1. With probability $\varepsilon > 0$ a **spontaneous mutation** occurs:
 - ▶ a node u is selected uniformly at random
 - ▶ his option changes from $X_u(t)$ to $\overline{X_u(t)} = 1 - X_u(t)$



A case study

2. With probability $1 - \varepsilon$, a **pairwise imitation step** occurs:
- ▶ a directed link (u, v) is selected uniformly at random
 - ▶ u copies the option of v with probability $\varphi(r_j(z(X(t))) - r_i(z(X(t))))$ where $X_u(t) = i$ and $X_v(t) = j$.
 - ▶ φ is strictly increasing and $\varphi(0) = 1/2$: the higher the difference between the rewards currently associated with the states j and i is, the larger the incentive is for an agent currently choosing state i to copy a neighbor and adopting his state j , upon observing his reward.



A case study

- ▶ When $\epsilon = 0$, namely when the spontaneous mutation term is absent, the chain is not ergodic. Indeed there are 2 pure configurations which are absorbing states for the Markov chain, i.e., the configurations where all the agents choose the same state 0 or 1.
- ▶ When $\epsilon > 0$, the chain is ergodic and there exists just one invariant probability μ .

Denote by μ_z the law of $z(X(t))$ when $X(t)$ is distributed according to μ .

A case study: main result

1. Connected graphs $\mathcal{G}_n = (\mathcal{V}_n, \mathcal{E}_n)$ where $\mathcal{V}_n = \{1, \dots, n\}$.
2. d_n^* and \bar{d}_n : the max and average degree of the nodes in \mathcal{G}_n
3. γ_n **Cheeger constant** of \mathcal{G}_n :
4. Let $\mu^{(n)}$ be the corresponding invariant probability over \mathcal{X}^n .

The following is our main result:

Theorem

Assume that

$$\sup_n \frac{d_n^*}{d_n} < +\infty, \quad \inf_n \gamma_n > 0 \text{ (expander graphs)}.$$

Then, for every $\delta > 0$, weakly

$$\lim_{\epsilon \downarrow 0} \lim_{n \rightarrow \infty} \mu_z = \delta_{z_{\text{Nash}}}$$

Expander graphs

$\mathcal{G}_n = (\mathcal{V}_n, \mathcal{E}_n)$ where $\mathcal{V}_n = \{1, \dots, n\}$.

Cheeger constant:

$$\gamma_n := \min_{\mathcal{U} \subseteq \mathcal{V}, |\mathcal{U}| \leq n/2} \frac{|\{(u, v) \in \mathcal{E}_n \mid u \in \mathcal{U}, v \notin \mathcal{U}\}|}{|\mathcal{U}|}$$

Expander graphs: $\inf_n \gamma_n > 0$

- ▶ Complete graphs are expanders.
- ▶ Grid-like graphs are not expanders.
- ▶ The class of expander graphs encompasses **important examples** typically considered in socio-economic networks like **Erdos-Renji graphs, configuration models, small world**.
- ▶ The theorem shows that through the gossip interactions, a **learning** process is taking place: **the population reaches the Nash equilibrium**.

A case study: considerations on the proof

Result was known (Sandholm) in the **mean field** case (graph complete). In this case $z(X(t))$ is **Markovian**

A birth and death chain for which the invariant probability μ_z can be explicitly computed.

Hydrodynamic limit for $n \rightarrow +\infty$ (and $t = n\tau$)

$$z' = \epsilon(1 - 2z) + (1 - \epsilon)z(1 - z)(2p(z) - 1)$$

where $p(z) = \varphi(r_1(z)) - r_0(z)$. Notice that $p(z_{\text{Nash}}) = 1/2$.

Just one stable equilibrium point converging to z_{Nash} for $\epsilon \rightarrow 0$.

A case study: considerations on the proof

When the graph is not complete $z(X(t))$ is no longer Markovian!

Imposing the balanced equation (μ is invariant)

$$\mathbb{E}_\mu[g(z(X(1)))] = \mathbb{E}_\mu[g(z(X(0)))]$$

for a generic C^1 observable $g : [0, 1] \rightarrow \mathbb{R}$ one gets, for $n \rightarrow +\infty$, that

$$\int_0^1 \left[\epsilon(1 - 2z) + \frac{(1 - \epsilon)}{2} (2p(z) - 1) \frac{d\eta}{d\mu_z}(z) \right] g'(z) d\mu_z(z) = 0$$

where $\eta(z)$ is the (mean value in the macro state z of the) **fraction of disagreement links**.

This implies that μ_z must be supported where

$$\epsilon(1 - 2z) + \frac{(1 - \epsilon)}{2} (2p(z) - 1) \frac{d\eta_\mu}{d\mu_z}(z) = 0$$

A case study: considerations on the proof

Analysis of the equation:

$$\epsilon(1 - 2z) + \frac{(1 - \epsilon)}{2}(2p(z) - 1) \frac{d\eta}{d\mu_z}(z) = 0$$

Complete case: $\frac{d\eta}{d\mu_z}(z) = 2z(1 - z)$.

For expander graphs $\frac{d\eta_\mu}{d\mu_z}(z) \geq \delta z(1 - z)$.

This yields the result

Potential games

\mathcal{X} finite set of options, $\mathcal{V} = \{1, \dots, n\}$ finite set of agents.

For $x \in \mathcal{X}^n$ and $i \in \mathcal{X}$, $\theta_i(x) = n^{-1} |\{v \in \mathcal{V} \mid x_v = i\}|$.

$r_i(\theta(x))$ the **reward** for an agent playing option $i \in \mathcal{X}$ when the rest of population options are distributed as $\theta(x)$.

Potential game: there exists a potential ψ such that

$$\theta_i > 0 \quad \implies \quad r_j(\theta) - r_i(\theta) = \frac{\partial}{\partial \theta_j} \psi(\theta) - \frac{\partial}{\partial \theta_i} \psi(\theta),$$

The variation of reward for an infinitesimal fraction of agents moving from state i to state j equals the corresponding variation in the potential.

Nash equilibria:

$$\theta_i > 0 \quad \implies \quad r_i(\theta) \geq r_j(\theta), \quad \forall j \in \mathcal{X}.$$

\mathcal{N} set of Nash equilibria. $\mathcal{N} \subseteq$ the set of stationary points of ψ .

Congestion games

$r_i(\theta) = r_i(\theta_i)$ with r_i strictly decreasing.

Potential $\psi(\theta) = \sum_i \int_0^{\theta_i} r_i(s) ds$ is strictly concave which implies uniqueness of the Nash equilibrium.

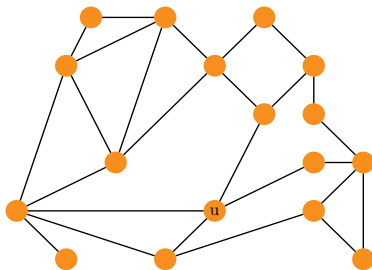
The case study considered fits in this framework.

Learning dynamics

$\mathcal{G} = (\mathcal{V}, \mathcal{E})$ connected undirected graph.

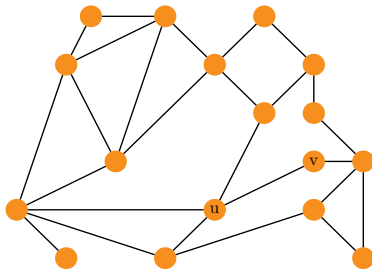
Dynamics: Markov chain $X(t)$ on \mathcal{X}^n . At each time t ,

1. With probability $\varepsilon > 0$ a spontaneous mutation occurs:
 - ▶ a node u is selected uniformly at random
 - ▶ his option changes from $X_v(t) = i$ to some $j \in \mathcal{X}$ with probability $P_{ij} > 0$.



Learning dynamics

1. With probability $1 - \varepsilon$, a pairwise imitation step occurs:
 - ▶ a directed link (u, v) is selected uniformly at random
 - ▶ u copies the option of v with probability $\varphi(r_j(\theta(X(t))) - r_i(\theta(X(t))))$ where $X_u(t) = i$ and $X_v(t) = j$.
 - ▶ φ is strictly increasing and $\varphi(0) = 1/2$: the higher the difference between the rewards currently associated with the states j and i is, the larger the incentive is for an agent currently choosing state i to copy a neighbor and adopting his state j , upon observing his reward.



Markov chain analysis

1. When $\epsilon = 0$, namely when the spontaneous mutation term is absent, the chain is not ergodic. Indeed there are $|\mathcal{X}|$ pure configurations which are absorbing states for the Markov chain, i.e., the configurations where all the agents choose the same state.
2. When $\epsilon > 0$, the chain is ergodic and there exists just one invariant probability μ .

Denote by μ_θ the law of $\theta(X(t))$ when $X(t)$ is distributed according to μ .

Main result

1. Connected graphs $\mathcal{G}_n = (\mathcal{V}_n, \mathcal{E}_n)$ where $\mathcal{V}_n = \{1, \dots, n\}$.
2. d_n^* and \bar{d}_n : the max and average degree of the nodes in \mathcal{G}_n
3. γ_n Cheeger constant of \mathcal{G}_n :
4. Let $\mu^{(n)}$ be the corresponding invariant probability over \mathcal{X}^n .

The following is our main result:

Theorem

Assume that

$$\sup_n \frac{d_n^*}{d_n} < +\infty, \quad \inf_n \gamma_n > 0 \text{ (expander graphs)}.$$

Then, for every $\delta > 0$,

$$\lim_{\epsilon \downarrow 0} \lim_{n \rightarrow \infty} \mu_{\theta}^{(n)}(\mathcal{N}_{\delta}) = 0$$

where $\mathcal{N}_{\delta} := \{\theta \in \mathcal{P}(\mathcal{X}) \mid \exists \theta^* \in \mathcal{N}, \text{ with } d(\theta, \theta^*) \leq \delta\}$

Conclusions and further steps

We have proposed an interacting agent model where agents are engaged in a potential game and compare experiences through gossip interaction.

Our main result shows that in expander graphs, when the number of units is large and the spontaneous mutation is small, they dynamically reach the Nash equilibrium.

Many open issues:

- ▶ Analyze the behavior on graphs which are not expanders
- ▶ Analyze the fraction of active links η at equilibrium. Which configurations are privileged?
- ▶ Study different models where local and global influences are present.