Graphon Mean Field Games with a Representative Player: Analysis and Learning Algorithm

Fuzhong Zhou, Chenyu Zhang, Xu Chen, Xuan Di

COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

ICML 2024, Vienna Austria

Outline

- Graphon
- Graphon games with a representative player
- Analogous with classic MFGs
- Comparison with the continuum-player formulation

Graphon

Graphon

A graphon W is an L_1 integrable function : $[0,1]^2 \to \mathbb{R}_+$.

- Regarded as a graph with infinitely many vertices taking labels in $[0,1]$
- The edge weight connecting vertex u and v is given by $W(u,v)$

Graphon games with a representative player (I)

- Consider a game for a continuum of players, each player has two 'properties': a label (valued in $[0, 1]$) and a state (valued in $\mathbb{R}^d$)
- The label represents the type of the player, and there are a continuum types of players
- The interaction intensities of different types of players is given by a known graphon W
- We abstract all other players but the representative into a label-state joint measure $\mu_t \in \mathcal{P}([0, 1] \times \mathbb{R}^d)$ for $t = 0, 1, \ldots$
- We model the action and dynamic for only the single representative player, as follows

Graphon games with a representative player (II)

- At initial time 0, the rep is assigned uniformly a label $U \sim \mathrm{unif}[0,1]$ and a starting position X_0 according to some initial distrubtion
- At the following discrete time steps $t = 0, 1, 2, \ldots,$ the player adopts a policy π and update her next state

$$a_t \sim \pi_t, \qquad X_{t+1} \sim P_t(X_t, \mathrm{W}\mu_t(U), a_t),$$

- $\pi_t \in \mathcal{P}(A)$: a stochastic policy
  α_t: the random action she takes at time t
  P_t: the state dynamic transition kernel
  $\mathrm{W}\mu_t(U)$: the population influence on the representative player (explained in next slide)

Graphon games with a representative player (III)

Graphon Operator

Given any graphon W, the graphon operator W and any population measure $\mu_t \in \mathcal{P}([0,1] \times \mathbb{R}^d)$, let it admits the disintegration $\mu_t(dv, dx) = dv\mu_t(v, dx)$.
we define $\mathrm{W}\mu_t$ as a operator .mapping $[0,1]$ to a finite Borel measure:

$$\mathrm{W}\mu_t(u) := \int_{[0,1]} \left[W(u,v) \int_{\mathbb{R}^d} \delta_x \mu_t(v, dx) \right] dv,$$

- The inner integral is the state law of the v-labeled player
- The outer integral is an average of the state law of players over all labels $v \in [0,1]$, weighted by the interaction intensity of player v onto player u, which is fixed

Graphon games with a representative player (IV)

- The representative aims to optimize her objective function

$$\sup_{\pi} J_W(\mu, \pi)$$

where

$$J_W(\mu, \pi) := \mathbb{E}\left[\sum_{t=0}^{T-1} f_t(X_t^{\pi}, \mathrm{W}\mu_t(U), a_t^{\pi}) + g(X_T^{\pi}, \mathrm{W}\mu_T(U)) \right].$$

- A measure-policy pair $(\widehat{\mu}, \widehat{\pi}) \in \mathcal{P}_{\mathrm{unif}}([0,1] \times \mathbb{R}^d) \times \mathcal{V}_U$ is a Nash equilibrium if

$$J_W(\widehat{\mu}, \widehat{\pi}) = \sup_{\pi} J_W(\widehat{\mu}, \pi),$$
$$\widehat{\mu} = \mathcal{L}(U, X^{\widehat{\pi}}).$$

Analogous with classic MFGs

	Graphon game	MFG
representative player	A random label U her state X	State X
population measure	A label-state joint law $\mu \in \mathcal{P}([0,1] \times \mathbb{R}^d)$	A state law $\mu \in \mathcal{P}(\mathbb{R}^d)$

- Both graphon game and MFG abstract a continuum of players into only a distribution (measure), and considers the state dynamic transition, action, and objective for only one representative player.
- In particular, when the graphon is $W \equiv 1$, our formulation degenerate to the classic MFG

Comparison with the continuum-player formulation (I)

Let us recap the continuum-player formulation in piror works: for any label u, the player with label u has her state X^u, which follows the state transition

$$a_t^u \sim \pi_t^u(X_t^u), \qquad X_{t+1}^u \sim P_t(X_t^u, \mathrm{W}\mu_t(u), a_t),$$

and aim to optimize her own objective over her policy π^u

$$J^u(\mu, \pi^u) := \mathbb{E}\left[\sum_{t=0}^{T-1} f_t(X_t^{u,\pi^u}, \mathrm{W}\mu_t(u), a_t^u) + g(X_T^{u,\pi^u}, \mathrm{W}\mu_T(u))\right],$$

The equilibrium is defined as a pair $(\widehat{\boldsymbol{\mu}}, \widehat{\boldsymbol{\pi}}) := (\{\widehat{\mu}^u\}_{u\in[0,1]}, \{\widehat{\pi}^u\}_{u\in[0,1]})$ such that

$$J^u(\widehat{\boldsymbol{\mu}}, \widehat{\pi}^u) = \sup_{\pi\in\mathcal{A}} J^u(\widehat{\boldsymbol{\mu}}, \pi),$$
$$\widehat{\mu}^u = \mathcal{L}(X^{u,\widehat{\pi}^u}),$$

for almost every $u \in [0, 1]$.

Comparison with the continuum-player formulation (II)

Difference:

- Our formulation abstract all other players into only a distribution, and consider the state, action, objective of only one representative player;
- The prior works model the state, action and objective for every player, which is uncountably infinite.

Difficulty faced by continuum-player formulation:

- Philosophical inconsistency with classic MFGs
- The formulation face significant non-measurability issue in proofs as the joint measurability of state dynamics with respect to a continuum of players' labels and randomness under the usual product space σ-algebra is not compatible with the independence of their evolution, (Appendix C)

Our Contributions

- Present many discussion on various concepts (Appendix A)
- Under mild assumptions, we proved the following analysis properties:
  - Existence of Nash equilibrium
  - Uniqueness of Nash equilibrium
  - The equilibrium of finite player game with heterogenous interactions asymptotically converge to the graphon game equilibrium
- We proposed the first full online, oracle-free algorithm based on fixed point iteration, and showed the sample complexity analysis

Thank you for your attention.