A General Theory for Discrete-Time Mean-Field Games

TAMER BAŞAR

Director, Center for Advanced Study
Dept of ECE and Coordinated Science Laboratory
University of Illinois at Urbana-Champaign

(based on joint work with Naci Saldi and Maxim Raginsky)

Workshop on Mean Field Games and Related Topics–4

Rome, Italy

14-16 June 2017

Nonzero-sum stochastic dynamic games

- Inherent difficulties in obtaining (and showing existence of) Nash equilibria in stochastic nonzero-sum dynamic games with asymmetric information and a finite number, N, of players
  - strategic interaction
  - non-standard dynamic optimization by each player for characterization of reaction functions
  - iterative second guessing
- Mean field game approach provides a way out
- Two recent papers this talk is based on
  - N. Saldi, T. Başar, M. Raginsky, "Markov-Nash equilibria in mean-field games with discounted cost," arXiv:1612.07878, 14 January 2017.
  - N. Saldi, T. Başar, M. Raginsky, "Approximate Nash equilibria in partially observed stochastic games with mean-field interactions," arXiv:1705.02036, 4 May 2017.

Outline

- **Part I** Fully-observed DT mean-field games
  - Formulation of the finite-agent game problem
  - Formulation of the mean-field game problem
  - Existence of mean-field equilibrium
  - Existence of approximate Markov-Nash equilibrium
- **Part II** Partially-observed DT mean-field games
  - Formulation of the finite-agent game problem
  - Formulation of the mean-field game problem
  - Existence of mean-field equilibrium
  - Existence of approximate Nash equilibrium
- Conclusions

Part I

Fully-observed mean-field games

Fully-observed mean-field game model

- For $i \in \{1, \ldots, N\}$, Agent i has the following state dynamics:

$$x_i^N(t+1) = F\big(x_i^N(t), a_i^N(t), e_t^{(N)}, v_i^N(t)\big),$$

  where the noise process $\{v_i^N(t)\}$ is an i.i.d. sequence and has the same distribution for each i and N.

- $x_i^N(t)$ is the X-valued state variable, $a_i^N(t)$ is the A-valued control action variable, and $e_t^{(N)}$ is the empirical distribution of the state configuration, i.e.

$$e_t^{(N)}(\,\cdot\,) := \frac{1}{N} \sum_{i=1}^{N} \delta_{x_i^N(t)}(\,\cdot\,) \in \mathcal{P}(\mathsf{X})$$

- Agent i selects its actions $\{a_i^N(t)\}$ to minimize a functional given in terms of the one-stage cost function $c : \mathsf{X} \times \mathsf{A} \times \mathcal{P}(\mathsf{X}) \to [0, \infty)$.

Control policies

- History spaces $\mathsf{H}_t = (\mathsf{X} \times \mathsf{A} \times \mathcal{P}(\mathsf{X}))^t \times (\mathsf{X} \times \mathcal{P}(\mathsf{X}))$, $t = 0, 1, 2, \ldots$.
- A *policy* for a generic agent is a sequence $\pi = \{\pi_t\}$ of stochastic kernels on A given H_t.
- A policy is *Markov* if each π_t is a Markov kernel on A given X.
- The set of all policies for Agent i is denoted by Π_i and the subset consisting of all Markov policies by M_i.
- Let $\mathbf{\Pi}^{(N)} = \prod_{i=1}^{N} \Pi_i$ and $\mathbf{M}^{(N)} = \prod_{i=1}^{N} \mathsf{M}_i$.

A special class of stochastic NZSDGs

- Dynamics of a generic agent i:

$$x_i^N(t+1) = \frac{1}{N}\sum_{j=1}^{N} f\big(x_i^N(t), a_i^N(t), x_j^N(t)\big) + g\big(x_i^N(t), a_i^N(t)\big) v_i^N(t)$$
$$= \int_{\mathsf{X}} f\big(x_i^N(t), a_i^N(t), y\big) e_t^{(N)}(dy) + g\big(x_i^N(t), a_i^N(t)\big) v_i^N(t)$$
$$=: F\big(x_i^N(t), a_i^N(t), e_t^{(N)}, v^i(t)\big),$$

  where $\{v_i^N(t)\}$ is a sequence of i.i.d. Gaussian random variables.

- One-stage cost function of a generic agent i:

$$c(x_i^N(t), a_i^N(t), e_t^{(N)}) = \frac{1}{N}\sum_{j=1}^{N} d\big(x_i^N(t), a_i^N(t), x_j^N(t)\big)$$
$$= \int_{\mathsf{X}} d\big(x_i^N(t), a_i^N(t), y\big) e_t^{(N)}(dy)$$

Cost functions

- For Agent i, the infinite-horizon discounted cost under a policy $\boldsymbol{\pi}^{(N)} \in \boldsymbol{\Pi}^{(N)}$ is given by

$$J_i^{(N)}(\boldsymbol{\pi}^{(N)}) = E^{\boldsymbol{\pi}^{(N)}}\left[\sum_{t=0}^{\infty} \beta^t c(x_i^N(t), a_i^N(t), e_t^{(N)})\right],$$

where $\beta \in (0, 1)$ is the discount factor.

Nash equilibria

The standard notion of optimality is a player-by-player one:

Definition

A policy $\boldsymbol{\pi}^{(N*)} = (\pi^{1*}, \ldots, \pi^{N*})$ constitutes a *Nash equilibrium* if

$$J_i^{(N)}(\boldsymbol{\pi}^{(N*)}) = \inf_{\pi^i \in \Pi_i} J_i^{(N)}(\boldsymbol{\pi}_{-i}^{(N*)}, \pi^i)$$

for each $i = 1, \ldots, N$, where $\boldsymbol{\pi}_{-i}^{(N*)} := (\pi^{j*})_{j \neq i}$.

Nash equilibria

- There are two challenges pertaining to Nash equilibria:
  - Almost decentralized nature of the information structure of the problem.
  - *Curse of dimensionality*: the solution of the problem becomes intractable when the numbers of states/actions and/or agents are large.
- Therefore, it is of interest to find an approximate decentralized equilibrium with reduced complexity.
- This is where the MFG approach comes in handy.

Approximate Markov-Nash equilibria

To that end, we adopt the following solution concept:

Definition

A policy $\boldsymbol{\pi}^{(N*)} \in \mathsf{M}^{(N)}$ is a *Markov-Nash equilibrium* if

$$J_i^{(N)}(\boldsymbol{\pi}^{(N*)}) = \inf_{\pi^i \in \mathsf{M}_i} J_i^{(N)}(\boldsymbol{\pi}_{-i}^{(N*)}, \pi^i)$$

for each $i = 1, \ldots, N$, and an ε-*Markov-Nash equilibrium* (for a given $\varepsilon > 0$) if

$$J_i^{(N)}(\boldsymbol{\pi}^{(N*)}) \leq \inf_{\pi^i \in \mathsf{M}_i} J_i^{(N)}(\boldsymbol{\pi}_{-i}^{(N*)}, \pi^i) + \varepsilon$$

for each $i = 1, \ldots, N$.

Highlights–Main steps

- A proof of existence of ε-Markov-Nash equilibria for games with sufficiently many agents.
- To this end, we first consider a mean-field game that arises in the infinite-population limit $N \to \infty$.
- We prove the existence of a mean-field equilibrium for this limiting mean-field game.
- Then, we show that the mean-field equilibrium is an approximate Markov-Nash equilibrium for the original game problem with sufficiently many agents.

Relevant literature

- History and literature on continuous-time mean-field games are well known to this audience.
- There are relatively few results on *discrete-time* mean-field games.
- Most earlier works (Gomes et al.'10, Adlakha et al.'15, Elliot et al.'13, Moon and Başar'15, Nourian and Nair'13) consider discrete or linear models.
- Biswas'15 considers average-cost setting with σ-compact Polish state space. But, in this work, agents are only coupled through the cost function.

Mean field game (MFG)

- Mean field game is the infinite-population limit $N \to \infty$ of the original game.
- We have a single agent and model the collective behavior of other agents by an exogenous *state-measure flow* $\boldsymbol{\mu} := (\mu_t)_{t \geq 0} \in \mathcal{P}(\mathsf{X})^\infty$ with a given initial condition μ_0.
- Given $\boldsymbol{\mu}$, a generic agent has the following state dynamics:

$$x(t+1) = F\big(x(t), a(t), \mu_t, v(t)\big).$$

- A generic agent selects its actions $\{a(t)\}$ to minimize a functional given in terms of the one-stage cost function c.

Control policies for MFG

- History spaces $\mathsf{G}_t = (\mathsf{X} \times \mathsf{A})^t \times \mathsf{X}$ for $t = 0, 1, \ldots$.
- A *policy* is a sequence $\pi = \{\pi_t\}$ of stochastic kernels on A given G_t. The set of all policies is denoted by Π.
- A *Markov* policy is a sequence $\pi = \{\pi_t\}$ of stochastic kernels on A given X. The set of Markov policies is denoted by M.

Mean field equilibria

- A policy $\pi^* \in \mathsf{M}$ is optimal for $\boldsymbol{\mu}$ if

$$J_{\boldsymbol{\mu}}(\pi^*) = \inf_{\pi \in \Pi} J_{\boldsymbol{\mu}}(\pi),$$

where

$$J_{\boldsymbol{\mu}}(\pi) := E^{\pi}\left[\sum_{t=0}^{\infty} \beta^t c(x(t), a(t), \mu_t)\right].$$

- Let $\mathcal{M} := \{\boldsymbol{\mu} \in \mathcal{P}(\mathsf{X})^{\infty} : \mu_0 \text{ is fixed}\}$.

Mean field equilibria

- Define the set-valued mapping $\Phi : \mathcal{M} \to 2^{\mathsf{M}}$ as

$$\Phi(\boldsymbol{\mu}) = \{\pi \in \mathsf{M} : \pi \text{ is optimal for } \boldsymbol{\mu}\}.$$

- Conversely, define a mapping $\Lambda : \mathsf{M} \to \mathcal{M}$ as follows: given $\pi \in \mathsf{M}$, $\boldsymbol{\mu} := \Lambda(\pi)$ is constructed recursively as

$$\mu_{t+1}(\,\cdot\,) = \int_{\mathsf{X}\times\mathsf{A}} p(\,\cdot\,|x(t), a(t), \mu_t)\pi_t(da(t)|x(t))\mu_t(dx(t)).$$

Mean field equilibria

We now introduce the notion of an equilibrium for the mean-field game:

Definition

A pair $(\pi, \boldsymbol{\mu}) \in \mathsf{M} \times \mathcal{M}$ is a *mean-field equilibrium* if $\pi \in \Phi(\boldsymbol{\mu})$ and $\boldsymbol{\mu} = \Lambda(\pi)$.

Notation

- Let $w : \mathsf{X} \to \mathbb{R}_+$ be a continuous moment function and $\alpha > 0$.
  (We assume $w(x) \geq 1 + d_X(x, x_0)^p$ for some $p \geq 1$ and $x_0 \in \mathsf{X}$; w can be replaced by 1 if c is bounded.)
- For any $g : \mathsf{X} \to \mathbb{R}$ and $\mu \in \mathcal{P}(\mathsf{X})$, define

$$\|g\|_w := \sup_{x \in \mathsf{X}} \frac{|g(x)|}{w(x)}$$

$$\|\mu\|_w := \sup_{\|g\|_w \leq 1} \left| \int_{\mathsf{X}} g(x)\mu(dx) \right|.$$

- For each $t \geq 0$, define

$$\mathcal{P}_w^t(\mathsf{X}) := \left\{ \mu \in \mathcal{P}(\mathsf{X}) : \int_{\mathsf{X}} w(x)\mu(dx) \leq \alpha^t M \right\},$$

  where $\int_{\mathsf{X}} w(x)\mu_0(dx) =: M < \infty$.

Assumption 1

(a) The cost function c is continuous.

(b) A is compact and X is locally compact.

(c) The stochastic kernel p is weakly continuous.

(d) We have

$$\sup_{(a,\mu)\in\mathsf{A}\times\mathcal{P}(\mathsf{X})} \int_{\mathsf{X}} w(y)p(dy|x,a,\mu) \leq \alpha w(x).$$

(e) The function $\int_{\mathsf{X}} w(y)p(dy|x,a,\mu)$ is continuous in (x,a,μ).

(f) There exist $\gamma \geq 1$ and a positive real number R such that for each $t \geq 0$, if we define $M_t := \gamma^t R$, then

$$\sup_{(a,\mu)\in\mathsf{A}\times\mathcal{P}_w^t(\mathsf{X})} c(x,a,\mu) \leq M_t w(x).$$

(g) $\alpha\beta\gamma < 1$.

Revisiting the special class of stochastic NZSDGs

- Dynamics of a generic agent i:

$$x_i^N(t+1) = \frac{1}{N}\sum_{j=1}^{N} f\left(x_i^N(t), a_i^N(t), x_j^N(t)\right) + g\left(x_i^N(t), a_i^N(t)\right) v_i^N(t)$$
$$= \int_{\mathsf{X}} f\left(x_i^N(t), a_i^N(t), y\right) e_t^{(N)}(dy) + g\left(x_i^N(t), a_i^N(t)\right) v_i^N(t)$$
$$=: F\left(x_i^N(t), a_i^N(t), e_t^{(N)}, v^i(t)\right),$$

where $\{v_i^N(t)\}$ is a sequence of i.i.d. Gaussian random variables.

- One-stage cost function of a generic agent i:

$$c(x_i^N(t), a_i^N(t), e_t^{(N)}) = \frac{1}{N}\sum_{j=1}^{N} d\left(x_i^N(t), a_i^N(t), x_j^N(t)\right)$$
$$= \int_{\mathsf{X}} d\left(x_i^N(t), a_i^N(t), y\right) e_t^{(N)}(dy)$$

Assumption 1 vis à vis the special class

Assumption 1 holds for the special class of NZSDGs, with $w(x) = 1 + x^2$, under the following conditions:

- A is compact.
- g is continuous, and f is bounded and continuous.
- $\sup_{a \in \mathsf{A}} g^2(x, a) \leq Lx^2$ for some $L > 0$.
- $\sup_{(x,a,y) \in K \times \mathsf{A} \times \mathsf{X}} d(x, a, y) < \infty$ for any compact $K \subset \mathsf{X}$.
- $d(x, a, y) \leq Rw(x)w(y)$ for some $R > 0$.
- $\omega_d(r) \to 0$ as $r \to 0$, where

$$\omega_d(r) = \sup_{y \in \mathsf{X}} \sup_{|x - x'| + |a - a'| \leq r} |d(x, a, y) - d(x', a', y)|$$

- $\alpha^2 \beta < 1$, where $\alpha = \max\{1 + \|f\|, L\}$

Existence of MF equilibrium

Theorem 1

Under Assumption 1, the mean-field game admits a mean-field equilibrium $(\pi, \boldsymbol{\mu})$.

Sketch of proof (for the case of bounded c):

For any $\boldsymbol{\nu} \in \mathcal{P}(\mathsf{X} \times \mathsf{A})^{\infty}$, let $J^{\boldsymbol{\nu}}_{*,t}$ denote the value function at time t of the nonhomogeneous Markov decision process that a generic agent is faced with:

$$J^{\boldsymbol{\nu}}_{*,t}(x) := \inf_{\pi \in \mathsf{M}} E^{\pi}\left[\sum_{k=t}^{\infty} \beta^k c(x(t), a(t), \nu_{t,1}) \middle| x(t) = x\right].$$

Hence, using dynamic programming,

$$J^{\boldsymbol{\nu}}_{*,t}(x) = \min_{a \in \mathsf{A}}\left[c(x, a, \nu_{t,1}) + \beta \int_{\mathsf{X}} J^{\boldsymbol{\nu}}_{*,t+1}(y) p(dy|x, a, \nu_{t,1})\right] =: T^{\boldsymbol{\nu}}_t J^{\boldsymbol{\nu}}_{*,t+1}(x).$$

Existence of MF equilibrium

Adopting the technique of *Jovanovic-Rosental (JME'88)*, introduce a set-valued mapping $\Gamma : \mathcal{P}(\mathsf{X} \times \mathsf{A})^\infty \to 2^{\mathcal{P}(\mathsf{X}\times\mathsf{A})^\infty}$ by

$$\Gamma(\boldsymbol{\nu}) = C(\boldsymbol{\nu}) \cap B(\boldsymbol{\nu}),$$

where

$$C(\boldsymbol{\nu}) := \left\{ \boldsymbol{\nu}' : \nu'_{0,1} = \mu_0, \, \nu'_{t+1,1}(\,\cdot\,) = \int_{\mathsf{X}\times\mathsf{A}} p(\,\cdot\,|x, a, \nu_{t,1})\nu_t(dx, da) \right\}$$

and

$$B(\boldsymbol{\nu}) := \Bigg\{ \boldsymbol{\nu}' : \forall t \geq 0, \; \nu'_t \Bigg(\Bigg\{ (x, a) : c(x, a, \nu_{t,1}) + \beta \int_{\mathsf{X}} J^{\boldsymbol{\nu}}_{*,t+1} p(dy|x, a, \nu_{t,1}) = T^{\boldsymbol{\nu}}_t J^{\boldsymbol{\nu}}_{*,t+1}(x) \Bigg\} \Bigg) = 1 \Bigg\}.$$

Existence of MF equilibrium

Suppose that Γ has a fixed point $\boldsymbol{\nu} = (\nu_t)_{t \geq 0}$. Construct a Markov policy $\pi = (\pi_t)_{t \geq 0}$ by disintegrating each ν_t as $\nu_t(dx, da) = \nu_{t,1}(dx)\pi_t(da|x)$, and let $\boldsymbol{\nu}_1 = (\nu_{t,1})_{t \geq 0}$. Then the pair $(\pi, \boldsymbol{\nu}_1)$ is a mean-field equilibrium. Hence, to complete the proof it suffices to prove that Γ has a fixed point.

Let us define

$$\mathcal{P}_w^t(\mathsf{X} \times \mathsf{A}) := \left\{ \nu \in \mathcal{P}(\mathsf{X} \times \mathsf{A}) : \nu_1 \in \mathcal{P}_w^t(\mathsf{X}) \right\}.$$

Note that $\mathcal{P}_w^t(\mathsf{X} \times \mathsf{A})$ is compact. Let $\Xi := \prod_{t=0}^{\infty} \mathcal{P}_w^t(\mathsf{X} \times \mathsf{A})$, which is convex and compact with respect to the product topology. One can prove that for any $\boldsymbol{\nu} \in \Xi$, we have $\Gamma(\boldsymbol{\nu}) \subset \Xi$.

Existence of MF equilibrium

The final piece we need in order to deduce the existence of a fixed point of Γ by an appeal to Kakutani's fixed point theorem is the following: The graph of Γ, i.e., the set

$$\mathrm{Gr}(\Gamma) := \{(\boldsymbol{\nu}, \boldsymbol{\xi}) \in \Xi \times \Xi : \boldsymbol{\xi} \in \Gamma(\boldsymbol{\nu})\},$$

is closed.

Existence of approximate Markov-Nash equilibrium

- Define the following moduli of continuity:

$$\omega_p(r) := \sup_{(x,a)\in \mathsf{X}\times\mathsf{A}} \sup_{\substack{\mu,\nu: \\ \rho_w(\mu,\nu)\leq r}} \|p(\,\cdot\,|x,a,\mu) - p(\,\cdot\,|x,a,\nu)\|_w$$

$$\omega_c(r) := \sup_{(x,a)\in \mathsf{X}\times\mathsf{A}} \sup_{\substack{\mu,\nu: \\ \rho_w(\mu,\nu)\leq r}} |c(x,a,\mu) - c(x,a,\nu)|.$$

- For any function $g : \mathcal{P}_w(\mathsf{X}) \to \mathbb{R}$, we define the w-norm of g as follows:

$$\|g\|_w^* := \sup_{\mu\in\mathcal{P}_w(\mathsf{X})} \frac{|g(\mu)|}{\int_{\mathsf{X}} w d\mu}.$$

Assumption 2

(h) $\omega_p(r) \to 0$ and $\omega_c(r) \to 0$ as $r \to 0$. Moreover, for any $\mu \in \mathcal{P}_w(\mathsf{X})$, the functions

$$\omega_p(\rho_w(\,\cdot\,,\mu)) : \mathcal{P}_w(\mathsf{X}) \to \mathbb{R}$$
$$\omega_c(\rho_w(\,\cdot\,,\mu)) : \mathcal{P}_w(\mathsf{X}) \to \mathbb{R}$$

have finite w-norm.

(i) There exists a positive real number B such that

$$\sup_{(a,\mu)\in\mathsf{A}\times\mathcal{P}_w(\mathsf{X})} \int_{\mathsf{X}} v^2(y)p(dy|x,a,\mu) \leq Bv^2(x).$$

Revisiting the special class of stochastic NZSDGs

- Dynamics of a generic agent i:

$$\begin{aligned} x_i^N(t+1) &= \frac{1}{N}\sum_{j=1}^{N} f\big(x_i^N(t), a_i^N(t), x_j^N(t)\big) + g\big(x_i^N(t), a_i^N(t)\big) v_i^N(t) \\ &= \int_{\mathsf{X}} f\big(x_i^N(t), a_i^N(t), y\big) e_t^{(N)}(dy) + g\big(x_i^N(t), a_i^N(t)\big) v_i^N(t) \\ &=: F\big(x_i^N(t), a_i^N(t), e_t^{(N)}, v^i(t)\big), \end{aligned}$$

where $\{v_i^N(t)\}$ is a sequence of i.i.d. Gaussian random variables.

- One-stage cost function of a generic agent i:

$$\begin{aligned} c(x_i^N(t), a_i^N(t), e_t^{(N)}) &= \frac{1}{N}\sum_{j=1}^{N} d\big(x_i^N(t), a_i^N(t), x_j^N(t)\big) \\ &= \int_{\mathsf{X}} d\big(x_i^N(t), a_i^N(t), y\big) e_t^{(N)}(dy) \end{aligned}$$

Assumption 2 vis à vis the special class

Assumption 2 holds for the special class of NZSSDGs under the following conditions:

- $d(x, a, y)$ is (uniformly) Hölder continuous in y with exponent p and Hölder constant K_d.
- $f(x, a, y)$ is (uniformly) Hölder continuous in y with exponent p and Hölder constant K_f.
- g is bounded and $\inf_{(x,a) \in \mathsf{X} \times \mathsf{A}} |g(x, a)| > 0$.

Main result

Theorem 2

Suppose that Assumptions 1 and 2 hold. Then, for any $\varepsilon > 0$, there exists a positive integer $N(\varepsilon)$, such that, for each $N \geq N(\varepsilon)$, the policy $\boldsymbol{\pi}^{(N)} = \{\pi, \pi, \ldots, \pi\}$ is an ε-Markov-Nash equilibrium for the game with N agents.

Logic of the proof

- We first show that as $N \to \infty$ the empirical distribution of the agents' states at each time t converges to a deterministic limit given by the mean-field equilibrium distribution of the state at time t.
  - This allows us to deduce that the evolution of the state of a generic agent closely tracks the equilibrium state-measure flow in the infinite-population limit.
- We then show that the infinite population limit is insensitive to individual-agent deviations from the mean-field equilibrium policy.

Sketch of the proof (for bounded c)

Let $(\pi, \boldsymbol{\mu})$ denote the mean-field equilibrium. One can prove that for all $t \geq 0$, $\mathcal{L}(e_t^{(N)}) \to \delta_{\mu_t}$ weakly in $\mathcal{P}(\mathcal{P}(\mathsf{X}))$, as $N \to \infty$. Using this, we obtain

$$\lim_{N \to \infty} J_1^{(N)}(\boldsymbol{\pi}^{(N)}) = J_{\boldsymbol{\mu}}(\pi) = \inf_{\pi' \in \Pi} J_{\boldsymbol{\mu}}(\pi'). \tag{1}$$

Sketch of the proof (*continued*)

Let $\{\tilde{\pi}^{(N)}\}_{N \geq 1} \subset \mathsf{M}_1$ be an arbitrary sequence of Markov policies for Agent 1 and let $\tilde{e}_t^{(N)}$ denote the corresponding empirical distribution. Then, we can prove that for each $t \geq 0$, $\mathcal{L}(\tilde{e}_t^{(N)}) \to \delta_{\mu_t}$ weakly in $\mathcal{P}(\mathcal{P}(\mathsf{X}))$, as $N \to \infty$. Using this, we obtain

$$\lim_{N \to \infty} \left| J_1^{(N)}(\tilde{\pi}^{(N)}, \pi, \ldots, \pi) - J_{\boldsymbol{\mu}}(\tilde{\pi}^{(N)}) \right| = 0. \tag{2}$$

Sketch of the proof (*continued*)

Given $\varepsilon > 0$, for each $N \geq 1$, let $\tilde{\pi}^{(N)} \in \mathsf{M}_1$ be such that

$$J_1^{(N)}(\tilde{\pi}^{(N)}, \pi, \ldots, \pi) < \inf_{\pi' \in \mathsf{M}_1} J_1^{(N)}(\pi', \pi, \ldots, \pi) + \frac{\varepsilon}{3}.$$

Then, by (1) and (2), we have

$$\begin{aligned}
\lim_{N\to\infty} J_1^{(N)}(\tilde{\pi}^{(N)}, \pi, \ldots, \pi) &= \lim_{N\to\infty} J_{\boldsymbol{\mu}}(\tilde{\pi}^{(N)}) \\
&\geq \inf_{\pi'} J_{\boldsymbol{\mu}}(\pi') \\
&= J_{\boldsymbol{\mu}}(\pi) \\
&= \lim_{N\to\infty} J_1^{(N)}(\pi, \pi, \ldots, \pi),
\end{aligned}$$

which completes the proof.

Part II

Partially-observed mean-field games

Partially-observed mean-field game model

- For $i \in \{1, \ldots, N\}$, Agent i has the following state and observation dynamics:

$$x_i^N(t+1) = F\big(x_i^N(t), a_i^N(t), e_t^{(N)}, v_i^N(t)\big),$$
$$y_i^N(t) = H\big(x_i^N(t), e_t^{(N)}, w_i^N(t)\big),$$

  where the noise processes $\{v_i^N(t)\}$ and $\{w_i^N(t)\}$ are i.i.d. sequences and independent of each other.

- $y_i^N(t)$ is the Y-valued observation variable.
- Agent i selects its actions $\{a_i^N(t)\}$ based on its observations to minimize discounted cost given in terms of the one-stage cost function $c : \mathsf{X} \times \mathsf{A} \times \mathcal{P}(\mathsf{X}) \to [0, \infty)$.

Control policies

- History spaces $\mathsf{H}_t = (\mathsf{Y} \times \mathsf{A})^t \times \mathsf{Y}$, $t = 0, 1, 2, \ldots$.
- A *policy* for a generic agent is a sequence $\pi = \{\pi_t\}$ of stochastic kernels on A given H_t.
- The set of all policies for Agent i is denoted by Π_i.
- Let $\mathbf{\Pi}^{(N)} = \prod_{i=1}^{N} \Pi_i$.

The special class with noisy measurements

- Dynamics of a generic agent i:

$$x_i^N(t+1) = \frac{1}{N}\sum_{j=1}^{N} f\left(x_i^N(t), a_i^N(t), x_j^N(t)\right) + g\left(x_i^N(t), a_i^N(t)\right)v_i^N(t)$$

$$=: F\left(x_i^N(t), a_i^N(t), e_t^{(N)}, v_i^N(t)\right),$$

$$y_i^N(t) = \frac{1}{N}\sum_{j=1}^{N} h\left(x_i^N(t), x_j^N(t)\right) + w_i^N(t)$$

$$=: H\left(x_i^N(t), e_t^{(N)}, w_i^N(t)\right),$$

  where $\{v_i^N(t)\}$ and $\{w_i^N(t)\}$ are sequences of i.i.d. Gaussian random variables.

- One-stage cost function of a generic agent i is same as in fully-observed case.

Nash equilibria

- Establishing the existence of Nash equilibria for PO-MFG is difficult due to:
  - Almost decentralized and noisy nature of the information structure of the problem.
  - *Curse of dimensionality*: the solution of the problem becomes intractable when the numbers of states/actions/observations and/or agents are large.
- Therefore, it is of interest to find an approximate decentralized equilibrium with reduced complexity.
- Note that if the number of players is small, it is all but impossible to show even the existence of approximate Nash equilibria for this general class (it is possible for the LQG case – TB (TAC'78)).

Approximate Nash equilibria

Definition

A policy $\boldsymbol{\pi}^{(N*)} \in \boldsymbol{\Pi}^{(N)}$ is an ε-*Nash equilibrium* (for a given $\varepsilon > 0$) if

$$J_i^{(N)}(\boldsymbol{\pi}^{(N*)}) \leq \inf_{\pi^i \in \Pi_i} J_i^{(N)}(\boldsymbol{\pi}_{-i}^{(N*)}, \pi^i) + \varepsilon$$

for each $i = 1, \ldots, N$.

Relevant literature

- In the literature relatively few results are available on PO-MFGs.
- This one appears to be the first one that studies discrete-time PO-MFGs with such generality.
- Most earlier works (Huang et al.'06, Şen and Caines'14,15,16,16, Tang and Meng'16, Huang and Wang'14) consider continuous time models.

Partially-observed mean field game (MFG)

- Similar to fully-observed case, mean field game is the infinite-population limit $N \to \infty$ of the original partially-observed game.
- We have a single agent and model the collective behavior of other agents by an exogenous *state-measure flow* $\boldsymbol{\mu} := (\mu_t)_{t \geq 0} \in \mathcal{P}(\mathsf{X})^\infty$ with a given initial condition μ_0.
- Given $\boldsymbol{\mu}$, a generic agent has the following state and observation dynamics:

$$x(t+1) = F\big(x(t), a(t), \mu_t, v(t)\big),$$
$$y(t) = H\big(x(t), \mu_t, w(t)\big).$$

- A generic agent selects its actions $\{a(t)\}$ based on its observations to minimize discounted cost given in terms of the one-stage cost function c.

Control policies for MFG

- History spaces $\mathsf{G}_t = (\mathsf{Y} \times \mathsf{A})^t \times \mathsf{Y}$ for $t = 0, 1, \ldots$.
- A *policy* is a sequence $\pi = \{\pi_t\}$ of stochastic kernels on A given G_t. The set of all policies is denoted by Π.

Mean field equilibria

- A policy $\pi^* \in \Pi$ is optimal for $\boldsymbol{\mu}$ if

$$J_{\boldsymbol{\mu}}(\pi^*) = \inf_{\pi \in \Pi} J_{\boldsymbol{\mu}}(\pi).$$

- Let $\mathcal{M} := \{\boldsymbol{\mu} \in \mathcal{P}(\mathsf{X})^\infty : \mu_0 \text{ is fixed}\}$.

Mean field equilibria

- Define the set-valued mapping $\Phi : \mathcal{M} \to 2^{\Pi}$ as

$$\Phi(\boldsymbol{\mu}) = \{\pi \in \Pi : \pi \text{ is optimal for } \boldsymbol{\mu}\}.$$

- Conversely, define a mapping $\Lambda : \Pi \to \mathcal{M}$ as follows: given $\pi \in \Pi$, $\boldsymbol{\mu} := \Lambda(\pi)$ is constructed recursively as

$$\mu_{t+1}(\,\cdot\,) = \int_{\mathsf{X}\times\mathsf{A}} p(\,\cdot\,|x(t), a(t), \mu_t) P^{\pi}(da(t)|x(t))\mu_t(dx(t)),$$

where $P^{\pi}(da(t)|x(t))$ is the conditional distribution of $a(t)$ given $x(t)$, under π.

Mean field equilibria

Definition

A pair $(\pi, \boldsymbol{\mu}) \in \Pi \times \mathcal{M}$ is a *mean-field equilibrium* if $\pi \in \Phi(\boldsymbol{\mu})$ and $\boldsymbol{\mu} = \Lambda(\pi)$.

Assumption 3

(a) The cost function c is bounded and continuous.

(b) A is compact and X is locally compact.

(c) The stochastic kernel p is weakly continuous.

(d) The observation kernel r is continuous in (x, μ) with respect to total variation distance.

(e) For some $\alpha > 0$ and a continuous moment function $w : \mathsf{X} \to [0, \infty)$,

$$\sup_{(a,\mu) \in \mathsf{A} \times \mathcal{P}(\mathsf{X})} \int_{\mathsf{X}} w(y) p(dy|x, a, \mu) \leq \alpha w(x).$$

Assumption 3 vis à vis the special class

Assumption 3 holds for the special case with $w(x) = 1 + x^2$ and $\alpha = \max\{1 + \|f\|^2, L\}$ under the following conditions:

- A is compact.
- d is continuous and bounded.
- g is continuous, and f is bounded and continuous.
- $\sup_{a \in \mathsf{A}} g^2(x, a) \leq Lx^2$ for some $L > 0$.
- h is continuous and bounded.

Existence of MF equilibrium

Theorem 3

Under Assumption 3, the partially-observed mean-field game admits a mean-field equilibrium $(\pi, \boldsymbol{\mu})$.

- To establish the existence of mean-field equilibrium, we use fully observed reduction of partially observed control problems and the dynamic programming principle.

Sketch of proof

- Given any measure flow $\boldsymbol{\mu}$, the optimal control problem for the mean-field game $(\mathsf{X}, \mathsf{A}, \mathsf{Y}, p, r, c)$ reduces to partially observed Markov decision process (POMDP).
- We know that this POMDP can be reduced to a fully observed MDP (belief-state MDP) $(\mathsf{Z}, \mathsf{A}, \{\eta_t^{\boldsymbol{\mu}}\}_{t \geq 0}, \{C_t^{\boldsymbol{\mu}}\}_{t \geq 0})$, where $\mathsf{Z} = \mathcal{P}(\mathsf{X})$.
- Hence, we can use the technique developed for fully-observed case to prove the existence of mean-field equilibrium.
- However, there is a crucial difference between this problem and fully-observed version: We do not have explicit analytical expression describing the relation between $\boldsymbol{\mu}$ and $\eta_t^{\boldsymbol{\mu}}$ from which we can deduce the continuity of $\eta_t^{\boldsymbol{\mu}}$ with respect to $\boldsymbol{\mu}$.

Sketch of proof

- Define the mapping $\mathsf{B} : \mathcal{P}(\mathsf{Z}) \to \mathcal{P}(\mathsf{X})$ as follows:

$$\mathsf{B}(\nu)(\,\cdot\,) = \int_{\mathsf{Z}} z(\,\cdot\,)\,\nu(dz).$$

- Using this definition, for any $\boldsymbol{\nu} \in \mathcal{P}(\mathsf{Z} \times \mathsf{A})^{\infty}$, we define the measure flow $\boldsymbol{\mu}^{\boldsymbol{\nu}} \in \mathcal{P}(\mathsf{X})^{\infty}$ as follows:

$$\boldsymbol{\mu}^{\boldsymbol{\nu}} = \big(\mathsf{B}(\nu_{t,1})\big)_{t \geq 0}.$$

- For any $\boldsymbol{\nu} \in \mathcal{P}(\mathsf{Z} \times \mathsf{A})^{\infty}$, let $J_{*,t}^{\boldsymbol{\nu}}$ denote the value function at time t of the belief-state MDP that a generic agent is faced with for $\boldsymbol{\mu}^{\boldsymbol{\nu}}$.

Sketch of proof

- Similar to fully-observed case, define the set-valued mapping $\Gamma : \mathcal{P}(\mathsf{Z} \times \mathsf{A})^\infty \to 2^{\mathcal{P}(\mathsf{Z}\times\mathsf{A})^\infty}$ by

$$\Gamma(\boldsymbol{\nu}) = C(\boldsymbol{\nu}) \cap B(\boldsymbol{\nu}),$$

where

$$C(\boldsymbol{\nu}) := \left\{ \boldsymbol{\nu}' : \nu'_{0,1} = \delta_{\mu_0}, \nu'_{t+1,1}(\,\cdot\,) = \int_{\mathsf{Z}\times\mathsf{A}} \eta_t^{\boldsymbol{\nu}}(\,\cdot\,|z,a)\nu_t(dz,da) \right\}$$

and

$$B(\boldsymbol{\nu}) := \left\{ \boldsymbol{\nu}' : \forall t \geq 0,\ \nu'_t\left(\left\{ (z,a) : C_t^{\boldsymbol{\nu}}(z,a) + \beta \int_{\mathsf{Z}} J^{\boldsymbol{\nu}}_{*,t+1} \eta_t^{\boldsymbol{\nu}}(dy|z,a) = T_t^{\boldsymbol{\nu}} J^{\boldsymbol{\nu}}_{*,t+1}(z) \right\} \right) = 1 \right\}.$$

Sketch of proof

- One can prove that $\eta_t^{\boldsymbol{\nu}}$ is continuous with respect to $\boldsymbol{\nu}$.
- This is a key element of the proof.
- The rest of the proof is the same as the proof in fully-observed case.

Existence of approximate Markov-Nash equilibrium

- Let $(\pi', \boldsymbol{\mu})$ be a mean-field equilibrium.
- Define the following moduli of continuity:

$$\omega_p(r) := \sup_{(x,a)\in \mathsf{X}\times\mathsf{A}} \sup_{\substack{\mu,\nu: \\ d_{BL}(\mu,\nu)\leq r}} \|p(\,\cdot\,|x,a,\mu) - p(\,\cdot\,|x,a,\nu)\|_w$$

$$\omega_c(r) := \sup_{(x,a)\in \mathsf{X}\times\mathsf{A}} \sup_{\substack{\mu,\nu: \\ d_{BL}(\mu,\nu)\leq r}} |c(x,a,\mu) - c(x,a,\nu)|.$$

Assumption 4

(f) $\omega_p(r) \to 0$ and $\omega_c(r) \to 0$ as $r \to 0$.

(g) For each $t \geq 0$, $\pi'_t : \mathsf{G}_t \to \mathcal{P}(\mathsf{A})$ is deterministic; that is, $\pi'_t(\,\cdot\,|g(t)) = \delta_{f_t(g(t))}(\,\cdot\,)$ for some measurable function $f_t : \mathsf{G}_t \to \mathsf{A}$.

(h) The observation kernel $r(\,\cdot\,|x)$ does not depend on the mean-field term.

Assumption 4 vis à vis the special class

Assumption 4 holds for the special case under the following conditions:

- $d(x, a, y)$ is (uniformly) Lipschitz in y with Lipschitz constant K_d.
- $f(x, a, y)$ is (uniformly) Lipschitz in y with Lipschitz constant K_f.
- g is bounded and $\inf_{(x,a) \in \mathsf{X} \times \mathsf{A}} |g(x, a)| > 0$.
- h is only a function of x.

Main result

Theorem 4

Suppose that Assumptions 3 and 4 hold. Then, for any $\varepsilon > 0$, there exists a positive integer $N(\varepsilon)$, such that, for each $N \geq N(\varepsilon)$, the policy $\boldsymbol{\pi}^{(N)} = \{\pi, \pi, \ldots, \pi\}$ is an ε-Nash equilibrium for the game with N agents.

Logic of the proof

- Since π' is deterministic, we can construct another deterministic policy π, which only uses observations, such that $(\pi, \boldsymbol{\mu})$ is also a mean-field equilibrium.
- Note that π is not necessarily Markovian.
- We construct an equivalent game model whose states are the state of the original model plus the current and past observations.
- In this new model, π automatically becomes Markov.
- Then, we use the proof technique in fully-observed case to show the existence of an approximate Nash equilibrium.

Conclusion

- Existence of approximate Nash equilibria for finite-population game problems of the mean-field type for both fully-observed and partially-observed MFGs
  - First established the existence of a Nash equilibrium in the limiting mean-field game problem.
  - Then applied this policy to the finite population game and demonstrated that it constitutes an ε-Nash equilibrium for games with a sufficiently large number of agents.
- Extensions
  - Non-homogeneous MFGs
  - Average cost MFGs
  - Risk-sensitive MFGs