Finite state mean field games. Control theory approach

Yurii Averboukh

Krasovskii Institute of Mathematics and Mechanics
Higher School of Economics
averboukh@gmail.com

Joint work with Aleksey Volkov (Krasovskii Institute of Mathematics and Mechanics)

October 14, 2022

Mean field games

- are proposed by Lions, Lasry and (independently) Huang, Malhamé, Caines in 2006;
- study the system of many identical agents who optimize their rewards in the limiting case when the number of agents tends to infinity;
- (key assumptions) the agents interacts via some external media, i.e., the dynamics and reward of each agent depends on his/her action, his/her state and the distribution of all other agents.

Finite-state mean field games

Theory:

Basna, Hilbert, Kolokoltsov, 2014;

Gomes, Mohr, Souza, 2013;

Bayraktar, Cohen, 2018;

Cecchin, Fischer, 2018;

Cecchin, Pelino, 2019

Bayraktar, Cecchin, Cohen, Delarue, 2019.

Applications:

Kolokoltsov, Malafeyev, 2019, 2017;

Katsikas, Kolokoltsov, 2019,

Kolokoltsov, Bensoussan, 2016.

Motivating example. Epidemic MFG model

- States of the individual: Susceptible, Infectious, Recovered.
- Distribution of states: m_S, m_I, m_R ($m_S + m_I + m_R = 1$).
- Control: $u = 1$ the individual observe the quarantine, $u = 0$ does not observe.
- The probability of disease on $[t, t + \Delta t]$ is $Q_{S,I}(m_I, u)\Delta t + o(\Delta t)$; $\quad Q_{S,I}(m_I, 1) < Q_{S,I}(m_I, 0)$.
- The probability of recover is $Q_{I,R}(m_I)\Delta t + o(\Delta t)$.
- The reward of healthy individual is $g(S, u)\Delta t$, $g(S, 0) > g(S, 1) \geq 0$.
- The reward of sicken individual is negative.

We wish to find all scenario of the epidemic!

Problem

Given a finite state mean field game, describe all its solution.

Continuous-time Markov chain

Let

- $\{1, \ldots, d\}$ be the set of states;
- $Q_{i,j}(t)$ be a transition rate from i to j;
- $Q_{i,i}(t) = -\sum_{j \neq i} Q_{i,j}(t)$.

Continuous-time Markov chain

Let

- $\{1, \ldots, d\}$ be the set of states;
- $Q_{i,j}(t)$ be a transition rate from i to j;
- $Q_{i,i}(t) = -\sum_{j \neq i} Q_{i,j}(t)$.

On the time interval $[t, t + \Delta t]$

conditional probability of transition from i to j is

$$Q_{i,j}(t)\Delta t + o(\Delta t),$$

condition probability of remaining at i is

$$1 + Q_{i,i}(t)\Delta t + o(\Delta t).$$

Dynamics of probabilities

Denote

- $m_i(t)$ be a probability of being at i at time t;
- $m(t) = (m_1(t), \ldots, m_d(t))$;
- $\Sigma^d = \{m = (m_1, \ldots, m_d) : m_i \geq 0, \quad m_1 + \ldots + m_d = 1\}$.

Kolmogorov equation

$$\frac{d}{dt} m(t) = m(t) Q(t), \quad m(t_0) = m_0,$$

where

- $Q(t) = (Q_{i,j}(t))_{i,j=1}^{d}$ is the Kolmogorov matrix,
- m_0 is the initial distribution.

Markov decision problem

- Transition rate: $Q_{i,j}(t, u)$, $u \in U$;
- Reward functional:

$$\mathbb{E}\left[\sigma(X(T)) + \int_{t_0}^{T} g(t, X(t), u(t))dt\right],$$

where $X(t)$ is the state of the system at time t; $X(\cdot)$ is a stochastic process.

Assumptions

- U is compact metric space;
- the functions g, σ, $Q_{i,j}$ are continuous.

Feedback strategies

The decision maker observes both time and state, i.e., the he/she plays by a feedback strategy:

$$u_{\text{fb}}(t) = (u_1(t), \ldots, u_d(t)).$$

New control space:

$$U^d = \{u_{\text{fb}} = (u_1, \ldots, u_d) : \ u_i \in U\}.$$

Feedback strategies

Kolmogorov matrix:

$$Q(t, u_{\text{fb}}) = \begin{pmatrix} Q_{1,1}(t, u_1), & Q_{1,2}(t, u_1), & \dots, & Q_{1,d}(t, u_1) \\ Q_{2,1}(t, u_2), & Q_{2,2}(t, u_2), & \dots, & Q_{2,d}(t, u_2) \\ \dots & \dots & \dots & \dots \\ Q_{d,1}(t, u_d), & Q_{d,2}(t, u_d), & \dots, & Q_{d,d}(t, u_d) \end{pmatrix}.$$

If $t \mapsto u_{\text{fb}}(t) \in U^d$ is a feedback strategy, then the Kolmogorov equation takes the form:

$$\frac{d}{dt} m(t) = m(t) Q(t, u_{\text{fb}}(t)).$$

Terminal and running rewards

$$\sigma = \begin{pmatrix} \sigma(1) \\ \vdots \\ \sigma(d) \end{pmatrix}, \quad g(t, u_{\mathsf{fb}}) = \begin{pmatrix} g(t, 1, u_1) \\ \vdots \\ g(t, d, u_d) \end{pmatrix}.$$

ODE formulation

The Markov decision problem can be reformulated as

$$\text{Maximize } m(T)\sigma + \int_{t_0}^{T} m(t)g(t, u(t))dt$$

$$\text{subject to } \frac{d}{dt}m(t) = m(t)Q(t, u(t)), \quad m(t_0) = m_0.$$

Bellman equation

$$\frac{d}{dt}\phi(t) = -H(t, \phi(t)), \quad \phi(T) = \sigma.$$

Here

- $\phi(t) = (\phi_1(t), \ldots, \phi_d(t))^T$ is the value function;
- $H(t, \phi) = (H_1(t, \phi), \ldots, H_d(t, \phi))^T$ is the Hamiltonian;
- $$H_i(t, \phi) \triangleq \max_{u \in U} \left[\sum_{j=1}^{d} Q_{i,j}(t, u)\phi_j + g(t, i, u) \right].$$

Optimal control

If $\phi : [0, T] \to \mathbb{R}$ solves the Bellman equation for the MDP, then

- $m_0\phi(t_0)$ is the optimal outcome for the initial distribution $m(t_0) = m_0$;
- the strategy $u^*(t) = (u_1^*(t), \ldots, u_d^*(t))$ satisfying

$$u_i^*(t) \in \operatorname*{Argmax}_{u \in U} \left[\sum_{j=1}^{d} Q_{i,j}(t,u)\phi_j + g(t,i,u) \right]$$

is the optimal feedback strategy.

Existence of the optimal control

- The Bellman equation for the MDP admits solution due to the existence and uniqueness theorem for ODEs.
- Given solution of the Bellman equation, optimal feedback exists by the Measurable Maximum Theorem.

Relaxation

- At the state i the decision maker chooses his/her control according to measure $\nu_i(du) \in \mathcal{P}(U)$.
- $\mathcal{P}(U)$ denotes the set of probabilities on U.
- A profile of feedback relaxed strategies is

$$\nu = (\nu_1, \ldots, \nu_d) \in (\mathcal{P}(U))^d.$$

Dynamics and rewards

$$\mathcal{Q}(t,m,\nu)$$

$$\triangleq \begin{pmatrix} \int_U Q_{1,1}(t,u_1)\nu_1(du_1), & \dots, & \int_U Q_{1,d}(t,u_1)\nu_1(du_1) \\ \dots & \dots & \dots \\ \int_U Q_{d,1}(t,u_d)\nu_d(du_d), & \dots, & \int_U Q_{d,d}(t,u_d)\nu_d(du_d) \end{pmatrix}.$$

Kolmogorov equation:

$$\frac{d}{dt}m(t) = m(t)\mathcal{Q}(t,\nu(t))$$

$$g(t,\nu) = \begin{pmatrix} \int_U g(t,1,u_1)\nu_1(du_1) \\ \vdots \\ \int_U g(t,d,u_d)\nu_d(du_d) \end{pmatrix}.$$

ODE formulation in relaxed strategies

$$\text{Maximize } m(T)\sigma + \int_{t_0}^{T} m(t) g(t, \nu(t)) dt$$

$$\text{subject to } \frac{d}{dt} m(t) = m(t)\mathcal{Q}(t, \nu(t)), \quad m(t_0) = m_0,$$

$$t \mapsto \nu(t) \in (\mathcal{P}(U))^d \text{ is weakly measurable.}$$

Relaxed strategies vs feedbacks. I

$$
\begin{aligned}
H_i(t,\phi) &= \max_{u\in U}\left[\sum_{j=1}^{d} Q_{i,j}(t,u)\phi_j + g(t,i,u)\right] \\
&= \max_{\nu\in\mathcal{P}(U)}\left[\sum_{j=1}^{d}\int_U Q_{i,j}(t,u)\nu(du)\phi_j + \int_U g(t,i,u)\nu(du)\right].
\end{aligned}
$$

Thus, the value function in the class of usual feedback strategies and in the class of relaxed feedback strategies coincide!

Relaxed strategies vs feedbacks. II

If $[t_0, T] \ni t \mapsto \nu(t) \in (\mathcal{P}(U))^d$ there exists a sequence of feedback strategies $u_n(\cdot)$, where $u_n(\cdot) : [t_0, T] \to U^d$, such that

- $\lambda \otimes \delta_{u_n(t)}$ converges weakly to $\lambda \otimes \nu(t)$ (here λ stands for the Lebesgue measure on $[t_0, T]$);
- if
$$\frac{d}{dt} m(t) = m(t)\mathcal{Q}(t, \nu(t)), \quad m(t_0) = m_0,$$
$$\frac{d}{dt} m_n(t) = m_n(t) Q(t, u_n(t)), \quad m_n(t_0) = m_0,$$
then $m_n(\cdot)$ converges uniformly to $m(\cdot)$.

Finite state mean field game

- Game with infinitely many players.
- $m(t)$ is the distribution of players;
- Transition rate: $Q_{i,j}(t, m, u)$;
- Reward function:

$$\mathbb{E}\left[\sigma(X(T), m(T)) + \int_{t_0}^{T} g(t, X(t), m(t), u(t))dt\right].$$

Relaxed control

- $\nu_i \in \mathcal{P}(U)$ is the control applied at the state i.
- $\nu = (\nu_1, \ldots, \nu_d) \in (\mathcal{P}(U))^d$;
- $$\mathcal{Q}_{i,j}(t, m, \nu) \triangleq \int_U Q_{i,j}(t, m, u)\nu_i(du);$$
- $$g_i(t, m, \nu) = \int_U g(t, i, m, u)\nu_i(du).$$

Why relaxation?

- If the usual feedback strategies are assumed, then all players placed at the state i are to play with the control $u_i(t)$.
- If the relaxed strategies are allowed, then the players placed at i can share their controls according to the measure $\nu_i(t, du)$.

Hamiltonian

$$H_i(t, m, \phi) = \max_{\nu_i \in \mathcal{P}(U)} \left[\sum_{j=1}^{q} \int_U Q(t, m, u)\nu_i(du) \cdot \phi_j + g_i(t, m, \nu_i) \right].$$

$$H(t, m, \phi) \triangleq (H_1(t, m, \phi), \ldots, H_d(t, m, \phi))^T,$$

$$g(t, m, \nu) \triangleq (g_1(t, m, \nu), \ldots, g_d(t, m, \nu))^T,$$

$$\sigma(m) \triangleq (\sigma(1, m), \ldots, \sigma(d, m))^T.$$

Mean field game system

- Kolmogorov equation:

$$\frac{d}{dt}m(t) = m(t)\mathcal{Q}(t, m(t), \hat{\nu}(t)), \quad m(t_0) = m_0.$$

- Bellman equation:

$$\frac{d}{dt}\phi(t) = -H(t, m(t), \phi(t)), \quad \phi(T) = \sigma(m(T)).$$

- Optimality condition:

$$\hat{\nu}_i(t) \in \underset{\nu_i \in U}{\mathrm{Argmax}} \Big[\sum_{j=1}^{d} \int_U Q_{i,j}(t, m, u)\nu_i(du) \cdot \phi_j(t) + \int_U g(t, i, m, u)\nu_i(du)\Big].$$

Existence

Theorem. There exists at least one solution of the finite state mean field game.

Key idea of the proof: use the convexity of the set of relaxed feedback strategies $t \mapsto \nu(t) \in (\mathcal{P}(U))^d$ and the Kakutani fixed-point theorem.

Existence

Theorem. There exists at least one solution of the finite state mean field game.

Key idea of the proof: use the convexity of the set of relaxed feedback strategies $t \mapsto \nu(t) \in (\mathcal{P}(U))^d$ and the Kakutani fixed-point theorem.

Remark. Generally, one cannot prove the existence of the equilibrium strategy in the class of usual feedbacks. However, if H_i are strictly convex, then there exists an usual feedback equilibrium.

Equivalent control problem

Minimize

$$J(m(\cdot),\phi(\cdot),\mu(\cdot),\nu(\cdot)) \triangleq \mu_0\phi(t_0) - \mu(T)\sigma(m(T)) - \int_{t_0}^{T}\mu(t)g(t,m(t),\nu(t))dt$$

subject to

$$\dot{m}(t) = m(t)\mathcal{Q}(t,m(t),\nu(t)), \quad m(t_0) = m_0$$

(Kolmogorov eq.),

$$\dot{\phi}(t) = -H(t,m(t),\phi(t)), \quad \phi(T) = \sigma(m(T))$$

(Bellman eq.),

$$\dot{\mu}(t) = \mu(t)\mathcal{Q}(t,m(t),\nu(t)), \quad \mu(t_0) = \mu_0$$

(complementary Kolmogorov eq.).

Admissible control processes

- $\nu(t) \in (\mathcal{P}(U))^d$ is a relaxed feedback control;
- $m(t) \in \Sigma^d$, $\phi(t) \in \mathbb{R}^d$, $\mu(t) \in \Sigma^d$ are state variables.

Admissible control processes

- $\nu(t) \in (\mathcal{P}(U))^d$ is a relaxed feedback control;
- $m(t) \in \Sigma^d$, $\phi(t) \in \mathbb{R}^d$, $\mu(t) \in \Sigma^d$ are state variables.

A control process $(m(\cdot), \phi(\cdot), \mu(\cdot), \nu(\cdot))$ is admissible iff

$$\dot{m}(t) = m(t)\mathcal{Q}(t, m(t), \nu(t)), \quad m(t_0) = m_0,$$
$$\dot{\phi}(t) = -H(t, m(t), \phi(t)), \quad \phi(T) = \sigma(m(T)),$$
$$\dot{\mu}(t) = \mu(t)\mathcal{Q}(t, m(t), \nu(t)), \quad \mu(t_0) = \mu_0.$$

Property

If $(m(\cdot), \phi(\cdot), \mu(\cdot), \nu(\cdot))$ is admissible control process, then

$$J(m(\cdot), \phi(\cdot), \mu(\cdot), \nu(\cdot)) \geq 0.$$

Key idea

Consider

$$\frac{d}{dt}\mu^k(t) = \mu^k(t)\mathcal{Q}(t, m(t), \nu(t)), \quad \mu^k(t_0) = \vartheta^k.$$

Here ϑ^k is the k-th coordinate row-vector.
Notice that

$$\phi_k(t_0) \geq \mu^k(T)\sigma(m(T)) + \int_{t_0}^{T} \mu^k(t)g(t, m(t), \nu(t))dt.$$

Theorem

For $(m(\cdot), \phi(\cdot))$ satisfying the Kolmogorov and Bellman equations and conditions $m(t_0) = m_0$, $\phi(T) = \sigma(m(T))$, the following statements are equivalent:

(i) The pair $(m(\cdot), \phi(\cdot))$ solves the mean field game whereas $\nu(\cdot)$ is the corresponding equilibrium profile of relaxed feedback strategies.

(ii) There exists $\mu(\cdot)$ satisfying the complementary Kolmogorov equation with nonzero coordinates of $\mu_0 \in \Sigma^d$ such that $(m(\cdot), \phi(\cdot), \mu(\cdot), \nu(\cdot))$ provides the solution of the optimal control problem with data (m_0, σ, μ_0).

(iii) For every $\mu(\cdot)$ satisfying the complementary Kolmogorov equation such that $\mu_0 \in \Sigma^d$, the control process $(m(\cdot), \phi(\cdot), \mu(\cdot), \nu(\cdot))$ is the solution of the optimal control problem with data (m_0, σ, μ_0).

Theorem. Part II

For $(m(\cdot), \phi(\cdot))$ satisfying the Kolmogorov and Bellman equations and conditions $m(t_0) = m_0$, $\phi(T) = \sigma(m(T))$, the following statements are equivalent:

(i) The pair $(m(\cdot), \phi(\cdot))$ solves the mean field game whereas $\nu(\cdot)$ is the corresponding equilibrium profile of relaxed feedback strategies.

(iv) The triple $(m(\cdot), \phi(\cdot), \nu(\cdot))$ is such that, for some $\mu_0 \in \Sigma^d$ with nonzero coordinates and $\mu(\cdot)$ satisfying the complementary Kolmogorov equation,

$$J(m(\cdot), \phi(\cdot), \mu(\cdot), \nu(\cdot)) = 0.$$

(v) The triple $(m(\cdot), \phi(\cdot), \nu(\cdot))$ is such that, for every $\mu(\cdot)$ satisfying the complementary Kolmogorov equation,

$$J(m(\cdot), \phi(\cdot), \mu(\cdot), \nu(\cdot)) = 0.$$

Dependence of solutions of MFG on initial distribution

Problem. Given $t_0 \in [0, T]$ and $m_0 \in \Sigma^d$, find all ψ such that $\psi = \phi(t_0)$, $m_0 = m(t_0)$ for some $(m(\cdot), \phi(\cdot))$ solving the MFG.

Generally, such ψ is nonunique.

Master equation

If $\psi(t,m) = (\psi_1(t,m),\dots,\psi_d(t,m))^T$ solves

$$\frac{\partial \psi_i}{\partial t} + H_i(t,m,\psi) + m\mathcal{Q}(t,m,\hat{\nu}(t,m,\psi(t,m)))\nabla_m \psi_i = 0,$$

$$\hat{\nu}(t,m,\phi) = (\hat{\nu}_1(t,m,\phi),\dots,\hat{\nu}_d(t,m,\phi))$$

$$\hat{\nu}_i(t,m,\phi) \in \operatorname*{Argmax}_{\nu_i \in \mathcal{P}(U)} \left[\sum_{j=1}^{d} \mathcal{Q}_{i,j}(t,m,\nu_i)\phi_j + g(t,i,m,\nu_i) \right],$$

then, for every t_0, m_0, there exists a solution of MFG $(m(\cdot),\phi(\cdot))$ such that

$$m(t_0) = m_0, \quad \phi(t_0) = \psi(t_0,m_0).$$

Master equation. II

Under uniqueness condition for MFG (e.g., monotonicity condition) the solution of the master equation provides the unique solution of MFG.

Dynamical system

- Kolmogorov equation:

$$\frac{d}{dt}m(t) = m(t)\mathcal{Q}(t, m(t), \nu(t)),$$

- Bellman equation:

$$\frac{d}{dt}\phi(t) = -H(t, m(t), \phi(t)),$$

- complementary Kolmogorov equation:

$$\frac{d}{dt}\mu(t) = \mu(t)\mathcal{Q}(t, m(t), \nu(t)),$$

Dynamics of residual

$$\frac{d}{dt}z(t) = -\mu(t)\mathcal{Q}(t, m(t), \nu(t))\phi(t) \\ + \mu(t)H(t, m(t), \phi(t)) - \mu(t)g(t, m(t), \nu(t)).$$

Lemma. If $(m(\cdot), \phi(\cdot), \mu(\cdot), \nu(\cdot))$ is a control process, then, for every $s, r \in [0, T]$, $s < r$,

$$z(r) - z(s) = \mu(s)\phi(s) - \mu(r)\phi(r) - \int_s^r \mu(t)g(t, m(t), \nu(t))dt.$$

In particular, the function $t \mapsto z(t)$ is nondecreasing.

Extended dynamic system

- $\nu(t) \in (\mathcal{P}(U))^d$ is a control;
- $m(t) \in \Sigma^d$, $\phi(t) \in \mathbb{R}^d$, $\mu(t) \in \Sigma^d$, $z(t) \in \mathbb{R}$ are state variables.

A 5-tuple $(m(\cdot), \phi(\cdot), \mu(\cdot), z(\cdot), \nu(\cdot))$ is a control process if it satisfies

- Kolmogorov equation;
- Bellman equation;
- complementary Kolmogorov equation;
- equation on residual.

Viability

A set $\mathcal{V} \subset [0, T] \times \Sigma^d \times \mathbb{R}^d \times \Sigma^d \times \mathbb{R}$ is viable with respect, if, for any $(s, m_*, \phi_*, \mu_*, z_*) \in \mathcal{V}$, there exist $r \in (s, T]$ and an admissible control process $(m(\cdot), \phi(\cdot), \mu(\cdot), z(\cdot), \nu(\cdot))$ satisfying initial conditions $m(s) = m_*$, $\phi(s) = \phi_*$, $\mu(s) = \mu_*$, $z(\cdot) = z_*$ such that

$$(r, m(r), \phi(r), \mu(r), z(r)) \in \mathcal{V}.$$

Theorem

Assume that

1. the closed set $\mathcal{V}$ is viable;
2. the inclusion $(T, m, \phi, \mu, z) \in \mathcal{V}$ implies that $\phi = \sigma(m)$ and $z = 0$.

Then, for any $t_0, m_0, \phi_0 \in [0, T] \times \Sigma^d \times \mathbb{R}^d$ such that $(t_0, m_0, \phi_0, \hat{\vartheta}, 0) \in \mathcal{V}$, there exists at least one solution of the mean field game $(m(\cdot), \phi(\cdot))$ satisfying

$$\phi(t_0) = \phi_0, \quad m(t_0) = m_0.$$

Here

$$\hat{\vartheta} \triangleq (1/\sqrt{d}, \ldots, 1/\sqrt{d}).$$

Corollary. Backward propagation of MFG

Let $\mathcal{W}$ be the set of 5-tuples $(s, m(s), \phi(s), \mu(s), z(s))$ where $(m(\cdot), \phi(\cdot), \mu(\cdot), z(\cdot), \nu(\cdot))$ is an admissible control process and

$$\phi(T) = \sigma(m(T)), \quad z(T) = 0.$$

Then, given $t_0 \in [0, T]$, the set

$$\{(m_0, \phi_0) : (t_0, m_0, \phi_0, \hat{\vartheta}, 0) \in \mathcal{W}\}$$

is equal to

$$\{(m(t_0), \phi(t_0)) : (m(\cdot), \phi(\cdot)) \text{ solves the finite state mean field game}\}.$$

Set of solutions of mean field game

1. Find the backward attainability set for the dynamical system

$$\begin{aligned}
\dot{m}(t) &= m(t)\mathcal{Q}(t, m(t), \nu(t)),\\
\dot{\phi}(t) &= H(t, m(t), \phi(t)),\\
\dot{\mu}(t) &= \mu(t)\mathcal{Q}(t, m(t), \nu(t)),\\
\dot{z}(t) &= -\mu(t)\mathcal{Q}(t, m(t), \nu(t))\phi(t)\\
&\qquad\quad + \mu(t)H(t, m(t), \phi(t)) - \mu(t)g(t, m(t), \nu(t)).
\end{aligned}$$

   from the seed $\{(T, m_T, \sigma(m_T), \mu_T, 0) : m_T, \mu_T \in \Sigma^d\}$.
   Denote it by $\mathcal{W}$.

2. Given t_0, m_0,

$$\{\phi(t_0) : (m(\cdot), \phi(\cdot)) \text{ solves MFG with } m(t_0) = m_0\}$$
$$= \{\phi_0 : (t_0, m_0, \phi_0, \hat{\vartheta}, 0) \in \mathcal{W}\}.$$

Planning problem

The planning problem for mean field games implies that one wishes to transfer the distribution of players from measure m_0 to the m_T choosing a terminal payoff σ.

Planning problem

- Kolmogorov equation:

$$\frac{d}{dt}m(t) = m(t)\mathcal{Q}(t, m(t), \hat{\nu}(t)), \quad m(0) = m_0, \quad m(T) = m_T.$$

- Bellman equation:

$$\frac{d}{dt}\phi(t) = -H(t, m(t), \phi(t)).$$

- Optimality condition:

$$\hat{\nu}_i(t) \in \operatorname*{Argmax}_{\nu_i \in U}\Big[\sum_{j=1}^{d} \int_U Q_{i,j}(t, m, u)\nu_i(du) \cdot \phi_j(t) + \int_U g(t, i, m, u)\nu_i(du)\Big].$$

Attainability assumption

There exists $\nu(\cdot)$ such that, for $m(\cdot)$ satisfying

$$\frac{d}{dt}m(t) = m(t)\mathcal{Q}(t, m(t), \hat{\nu}(t))$$

one has $m(0) = m_0$, $m(T) = m_T$.

Nonexistence of the solution of the planning problem

- $d = 3, \quad T = 1;$
- $m_0 = (1, 0, 0),\ m_T = (e^{-1/3}, 1 - e^{-1/3}, 0);$
- $$Q(t, m, u) = \begin{pmatrix} -u & u & 0 \\ 0 & -\rho(t) & \rho(t) \\ 0 & 0 & 0 \end{pmatrix},$$

  where

  $$\rho(t) = \begin{cases} 1, & 0 \leq t \leq \frac{1}{3} \\ -3t + 2, & \frac{1}{3} \leq t \leq \frac{2}{3}, \\ 0, & \frac{2}{3} \leq t \leq 1. \end{cases}$$
- $$\tilde{g}(t, i, m, u) = \begin{cases} -u^2 & , i = 1, \\ 0 & , i = 2, 3, \end{cases}$$

The only admissible control

$$\tilde{u}(t) = \begin{cases} 0, & 0 \leq t \leq \frac{2}{3}, \\ 1, & \frac{2}{3} \leq t \leq 1. \end{cases}$$

Analysis of optimal controls

- The Bellman equation for $\phi_1(\cdot)$ takes the form:

$$\dot{\phi}_1(t) = -\max_{u \in U} \left\{ u(\phi_2(t) - \phi_1(t)) - \frac{u^2}{2} \right\}$$

- Thus, the optimal control is

$$u^*(t) = \begin{cases} 0, & \phi_2(t) - \phi_1(t) \leq 0 \\ \phi_2(t) - \phi_1(t), & 0 \leq \phi_2(t) - \phi_1(t) \leq 1 \\ 1, & 1 \leq \phi_2(t) - \phi_1(t) \end{cases}.$$

- $u^*(\cdot)$ is continuous.

$$\tilde{u}(\cdot) \neq u^*(\cdot).$$

Equivalent optimization problem

Let $\mu_0 = (\mu_{0,1}, \ldots, \mu_{0,d}) \in \Sigma^d$ be such that $\mu_{0,i} > 0$.

The planning problem is equivalent to the following:
minimize

$$J_p(m(\cdot), \phi(\cdot), \mu(\cdot), \nu(\cdot)) = \mu_0\phi_0 - \mu(T)\phi(T) - \\ - \int_{t_0}^{T} \mu(\tau) g(\tau, m(\tau), \nu(\tau)) d\tau,$$

subject to

$$0 = \min_{\phi(T),\nu} J_p(t, m(t), \phi(t), \mu(t), \nu(t)),$$

$$\dot{m}(t) = m(t)\mathcal{Q}(t, m(t), \nu(t)), \quad m(t_0) = m_0, m(T) = m_T,$$

$$\dot{\phi}(t) = -H(t, m, \phi),$$

$$\dot{\mu}(t) = \mu(t)\mathcal{Q}(t, m(t), \nu(t)), \quad \mu(t_0) = \mu_0,$$

Weakened optimization problem

minimize

$$J_p(m(\cdot), \phi(\cdot), \mu(\cdot), \nu(\cdot)) = \mu_0\phi_0 - \mu(T)\phi(T) - \\ - \int_{t_0}^{T} \mu(\tau) g(\tau, m(\tau), \nu(\tau)) d\tau,$$

subject to

$$\dot{m}(t) = m(t)\mathcal{Q}(t, m(t), \nu(t)), \quad m(t_0) = m_0, m(T) = m_T,$$
$$\dot{\phi}(t) = -H(t, m, \phi),$$
$$\dot{\mu}(t) = \mu(t)\mathcal{Q}(t, m(t), \nu(t)), \quad \mu(t_0) = \mu_0,$$
$$\|\phi(T)\| \leq \alpha.$$

Minimal regret solution

Let $\mu_0 = (\mu_{0,1}, \ldots, \mu_{0,d}) \in \Sigma^d$ with $\mu_{0,i} > 0$ be fixed.

We say that $\gamma^*(\cdot)$ is a μ_0-minimal regret solution, if there exist sequences $\{\alpha^n\}$ and $\{(m^n(\cdot), \phi^n(\cdot), \nu^n(\cdot), \mu^n(\cdot))\}$ such that

- each 4-tuple $(m^n(\cdot), \phi^n(\cdot), \nu^n(\cdot), \mu^n(\cdot))$ solve the weakened optimization problem for MFG under condition $\|\phi^n(T)\| \leq \alpha^n$;
- $\alpha^n \uparrow \infty$;
- $\nu^n(\cdot) \to \gamma^*(\cdot)$.

Theorem

Assume that μ_0 is given.

- There exists at least one μ_0-minimal regret solution;
- The set of μ_0-minimal regret solution is closed.
- If there exists at least one classical solution to the planning problem, then the set of μ_0-minimal regret solution is the closure of the set of classical solutions.

References

- Yu. Averboukh. Control theory approach to continuous-time finite state mean field games. Mathematical Control and Related Fields (Online first), arXiv:2103.07493.
- Yu. Averboukh, A. Volkov. Planning problem for continuous-time finite state mean field game with compact action space. arXiv:2204.08736.

Thank you for your attention!