An approximation of the Bellman equation for the mean field type control problem

Yurii Averboukh

Krasovskii Institute of Mathematics and Mechanics
ayv@imm.uran.ru

November 7, 2022

Motivating example I. N-particle system

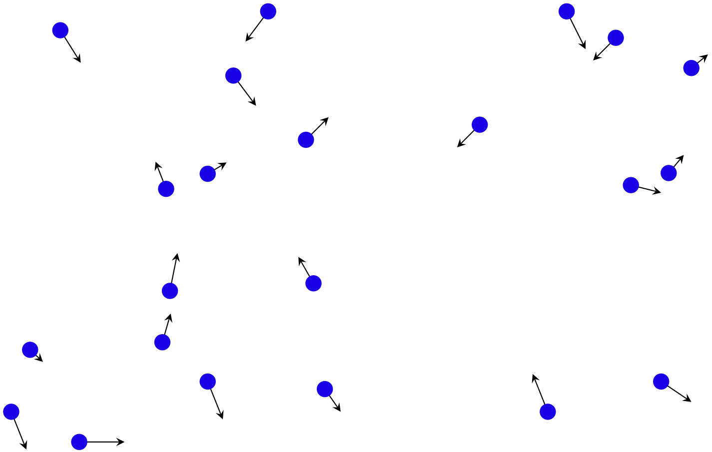
Mechanical system:

- N identical particles;
- the total mass is equal to 1;
- the interaction between particle placed in q' and q'' is given by the force $F(q'' - q')$.

The dynamics of the i-th particle is described by the equations:

$$\begin{aligned} \dot{q}_i &= v_i \\ \dot{v}_i &= \frac{1}{N} \sum_{j \neq i} F(q_j - q_i). \end{aligned}$$

N-particle systems



N-particle systems. Phase space

- State of each particle: $x_i = (q_i, v_i) \in \mathbb{R}^6$.
- State of the whole system: $(x_1, \ldots, x_N) \in \mathbb{R}^{6N}$. Often, this information is unavailable.
- Available information: the number of particles containing in each set $E \subset \mathbb{R}^6$. In fact, we know the distribution of the particles over the phase space.

N-particle systems. Dynamics of distribution

Distribution of particles: if $(x_1(t), \ldots, x_N(t))$ describes the states of the particles at time t, then

$$m_N(t) \triangleq \frac{1}{N}\sum_{j=1}^{N} \delta_{x_j(t)},$$

where δ_z stands for the Dirac measure concentrated at z.
Dynamics of each particle:

$$\begin{aligned} \dot{q}_i &= v_i \\ \dot{v}_i &= \int_{\mathbb{R}^6} F(q - q_i) m_N(t, dqdv). \end{aligned}$$

N-particle systems. Dynamics of distribution

$m(\cdot)$ satisfies in the distributional sense the equation

$$\frac{\partial}{\partial t} m_N(t) + \operatorname{div}(f(x, m_N(t)) m_N(t)) = 0,$$

i.e., for any $\varphi \in C_c((0, T) \times \mathbb{R}^6)$,

$$\int_0^T \int_{\mathbb{R}^6} \left[\frac{\partial \varphi}{\partial t}(t, x) + \nabla \varphi(t, x) f(x, m(t)) \right] m(t, dx) dt = 0.$$

Here $x = (q, v)$, $f(x, m) = (v, (F * m)(x))$,

$$(F * m)(q, v) \triangleq \int_{\mathbb{R}^6} F(q' - q) m(dq' dv').$$

Limiting system

We let $N \to \infty$ and consider the limit dynamics with

- phase variable is a probability m on $\mathbb{R}^6$;
- dynamics of the distribution obeys the Liouville equations:

$$\frac{\partial}{\partial t} m(t) + \operatorname{div}(f(x, m(t))m(t)) = 0.$$

Motivating example II. Opinion dynamics

Model:

- N participants;
- $x_i \in \mathbb{R}^d$ denotes the vector of opinions of the i-th participant;
- the dynamics of i-th participant's opinion is

$$\dot{x}_i = \frac{1}{N}\sum_{j=1}^{N} \xi(x_j - x_i)(x_j - x_i);$$

here $\xi : \mathbb{R}^d \to \mathbb{R}$ is nonnegative and radially symmetric.

Opinion dynamics. Phase space

- Only information of distribution of opinions is available.
- $$m_N(t) = \frac{1}{N}\sum_{j=1}^{N}\delta_{x_j(t)}.$$
- Dynamics of opinion of each participant:
$$\dot{x}_i = f(x_i, m_N(t)).$$
- Dynamics of the distribution of opinions:
$$\frac{\partial}{\partial t}m_N(t) + \operatorname{div}(f(x, m_N(t))m_N(t)) = 0.$$

Here
$$f(x, m) \triangleq \int_{\mathbb{R}^d}\xi(x' - x)(x' - x)m(dx').$$

Limiting system

We deal only with the distribution of opinions that is a probability over $\mathbb{R}^d$. Its dynamics satisfies

$$\frac{\partial}{\partial t} m(t) + \operatorname{div}(f(x, m(t))m(t)) = 0.$$

Nonlocal continuity equation

Let

- $\mathbb{R}^d$ be a phase space for each particle;
- $f(t, x, m)$, where $t \in [0, T]$, $x \in \mathbb{R}^d$, m is a probability on $\mathbb{R}^d$, be a nonlocal velocity field.

Then, dynamics of distribution of particles satisfies in the distributional sense the nonlocal continuity equation:

$$\frac{\partial}{\partial t} m(t) + \operatorname{div}(f(t, x, m(t))m(t)) = 0.$$

In particular, the dynamics of each particle obeys the ODE:

$$\dot{x} = f(t, x, m(t)).$$

Control of systems consisting of infinite number of elements

- Individual (depending on state) control + individual aim = mean field games.
- Common control + common aim = control of continuity equation.
- Individual (depending on state) control + common aim = mean field type control.

Example. Control of charged variables

- Dynamics of particle:

$$\dot{q} = v$$
$$\dot{v} = \int_{\mathbb{R}^6} F(q - q_i)m(t, dqdv) + u(t, q, v).$$

- Dynamics of the distribution of particles:

$$\frac{\partial}{\partial t}m(t) + \mathrm{div}((f(x, m(t)) + u(t, x))m(t)) = 0,$$

where $x = (q, v)$, $f(x, m) = (v, (F * m)(x))$,
$(F * m)(q, v) \triangleq \int_{\mathbb{R}^6} F(q' - q)m(dq'dv')$.

- Aim: keep the system inside the set G spending minimal energy:

$$\int_0^T \int_{\mathbb{R}^6} [\mathbb{1}_G(q) - \mu u^2(t, q, v)]m(t, dqdv)dt \to \max.$$

Example. Control for consensus

- Dynamics of participant's opinion:

$$\dot{x} = \int_{\mathbb{R}^d} \xi(x' - x)(x' - x)m(t, dx') + \zeta(x)u(t, x)$$

- Dynamics of the distribution of opinions:

$$\frac{\partial}{\partial t} m(t) + \mathrm{div}((f(x, m(t)) + \zeta(x)u(t, x))m(t)) = 0,$$

where $f(x, m) \triangleq \int_{\mathbb{R}^d} \xi(x' - x)(x' - x)m(dx')$.

- Aim: maximize consensus at time T minimizing the efforts:

$$\int_{\mathbb{R}^d} \left[x - \int_{\mathbb{T}^d} x' m(T, dx')\right]^2 m(T, dx) \\ + \mu \int_0^T \int_{\mathbb{R}^d} u^2(t, dx)m(t, dx)dt \to \min .$$

Example. Control of swarm of robots

- Dynamics of each robot:

$$\dot{x} = f(x, u(t, x)).$$

- Dynamics of the whole swarm:

$$\frac{\partial}{\partial t} m(t) + \operatorname{div}(f(x, u(t, x)) m(t)) = 0.$$

- Aim: stir the system to the desired distribution m^* minimizing the efforts:

$$\text{squared distance}(m(T), m^*) \\ + \mu \int_0^T \int_{\mathbb{R}^d} u^2(t, dx) m(t, dx) dt \to \min.$$

Notation

- If (X, ρ_X) is a Polish space, then $\mathcal{B}(X)$ denotes the Borel σ-algebra on X.
- $\mathcal{P}(X)$ is the set of Borel probabilities on X.

Push-forward measure

Assume that

- $(\Omega, \mathcal{F})$, $(\Omega', \mathcal{F}')$ are measurable spaces,
- $\mathbb{P}$ is a probability on $\mathcal{F}$,
- $\xi : \Omega \to \Omega'$ is measurable function.

A probability $\xi\sharp\mathbb{P}$ on $\mathcal{F}'$ defined by the rule: for $E \in \mathcal{F}'$

$$(\xi\sharp\mathbb{P})(E) \triangleq \mathbb{P}(\xi^{-1}(E))$$

is called a push-forward measure.

Notation

- If (X, ρ_X) is a Polish space, $p \geq 1$, then $\mathcal{P}^p(X)$ is the set of probabilities on X with finite p-th moment, i.e., $m \in \mathcal{P}^p(X)$ iff, for some (equivalently, any) $x_* \in X$, $\int_X \rho_X^p(x, x_*) m(dx) < \infty$.
- Distance on $\mathcal{P}^p(X)$: if $m_1, m_2 \in \mathcal{P}^p(X)$, then

$$W_p(m_1, m_2) \triangleq \inf \Big[\int_{X \times X} \rho_X^p(x_1, x_2) \pi(dx_1 dx_2) : \\ \pi \in \Pi(m_1, m_2) \Big]^{1/p},$$

where $\Pi(m_1, m_2)$ is the set of probabilities π on $X \times X$ such that, for any measurable $E \subset X$, $\pi(E \times X) = m_1(X)$, $\pi(X \times E) = m_2(E)$.

Mean field type control problem. Informal setting

Dynamics of each agent

$$\dot{x} = f(t, x, m(t), u(t, x)),$$

where

- $t \in [0, T]$,
- $x \in \mathbb{T}^d$, $\mathbb{T}^d \triangleq \mathbb{R}^d/\mathbb{Z}^d$,
- $m(t) \in \mathcal{P}^2(\mathbb{T}^d)$ is the distribution of agents,
- $u(t, x) \in U$ is the control.

The aim is to minimize the payoff of all agents that is equal to

$$\sigma(m(T)).$$

Assumptions

- U is a convex compact subset of some Banach space;
- f is continuous and Lipschitz continuous w.r.t. x and m;
- σ is continuous;
- f is affine w.r.t. u, f_0 is convex w.r.t. u.

Eulerian approach

- Control process: $(m(\cdot), u_E)$, where $m(t)$ is a probability on $\mathbb{T}^d$, $u_E : [0, T] \times \mathbb{T}^d \to U$.
- Dynamics: $m(\cdot)$ is a distributional solution of the nonlocal continuity equation:

$$\partial_t m(t) + \mathrm{div}(v_E(t,x)m(t)) = 0,$$

for $v_E(t,x) = f(t,x,m(t),u_E(t,x))$.
- Initial condition: $m(0) = m_0$.
- Payoff:

$$J_E(\mu, u_E) \triangleq \sigma(m(T))$$

Lagrangian approach

- $(\Omega, \mathcal{F}, \mathbb{P})$ is a standard probability space.
- Control process: (X, u_L), where $X : [0, T] \times \Omega \to \mathbb{T}^d$, $u_L : [0, T] \times \Omega \to U$.
- Dynamics:

$$\frac{d}{dt}X(t,\omega) = f(t, X(t,\omega), X(t)\sharp\mathbb{P}, u(t,\omega)).$$

- Initial condition: $X(0)\sharp\mathbb{P} = m_0$.
- Payoff:

$$J_L(X, u_L) \triangleq \sigma(X(T)\sharp\mathbb{P}).$$

Kantorovich approach

- Space of curves: $\Gamma = C([0,T];\mathbb{T}^d)$
- Control process: (η, u_K), where $\eta \in \mathcal{P}^2(\Gamma)$, $u_K : [0,T] \times \Gamma \to U$.
- Feasibility: for η-a.e. $\gamma \in \Gamma$,

$$\frac{d}{dt}\gamma(t) = f(t, \gamma(t), e_t \sharp \eta, u_K(t, \gamma)),$$

where $e_t(\gamma) = \gamma(t)$, $(e_t \sharp \eta)(E) = \eta\{\gamma \in \Gamma : \gamma(t) \in E\}$.
- Initial condition: $e_0 \sharp \eta = m_0$.
- Payoff:

$$J_K(\eta, u_K) \triangleq \sigma(e_T \sharp \eta)$$

Value function

- Eulerian approach:

$$\mathrm{Val}_E(m_0) \triangleq \inf\{J_E(m(\cdot), u_E) : (m(\cdot), u_E) \text{ is an Eulerian process}, m(0) = m_0\}.$$

- Lagrangian approach:

$$\mathrm{Val}_L(m_0) \triangleq \inf\{J_E(X, u_L) : (X, u_L) \text{ is a Lagrangian process}, X(0)\sharp\mathbb{P} = m_0\}.$$

- Kantorovich approach:

$$\mathrm{Val}_K(m_0) \triangleq \inf\{J_E(\eta, u_K) : (\eta, u_K) \text{ is a Kantorovich process}, e_0\sharp\eta = m_0\}.$$

Equivalence of approaches

Theorem (Cavagnari et al, 2022)

- $\mathrm{Val}_E(m_0) = \mathrm{Val}_L(m_0) = \mathrm{Val}_K(m_0)$.
- *the function* Val *is continuous.*

Existence of minimizer

Theorem (Cavagnari et al, 2022)

- *There exist optimal Eulerian and Kantorovich processes.*
- *If $\mathbb{P}$ is atomless, then there exists an optimal Lagragian process.*

Bellman equation

The value function Val should satisfy the following Bellman equation:

$$\frac{\partial\varphi}{\partial t} + \mathcal{H}(t, m, \nabla_m\varphi) = 0, \quad \varphi(T, m) = \sigma(m),$$

where, for $p \in L^2(\mathbb{R}^d, m; \mathbb{R}^d)$,

$$\mathcal{H}(t, m, p) \triangleq \int_{\mathbb{T}^d} \min_{u\in U}\langle p(x), f(t, x, m, u)\rangle m(dx).$$

Intrinsic derivative

Definition

Let $\varphi : \mathcal{P}^2(\mathbb{T}^d) \to \mathbb{R}$. A function $\frac{\delta\varphi}{\delta m} : \mathcal{P}^2(\mathbb{T}^d) \times \mathbb{R}^d \to \mathbb{R}$ is a flat derivative iff, for any $m' \in \mathcal{P}^2(\mathbb{R}^d)$,

$$\lim_{s\downarrow 0} \frac{\varphi((1-s)m + sm') - \varphi(m)}{s} = \int_{\mathbb{T}^d} \frac{\delta\varphi}{\delta m}(m,y)[m'(dy) - m(dy)].$$

Intrinsic derivative

Definition

Let $\varphi : \mathcal{P}^2(\mathbb{T}^d) \to \mathbb{R}$. A function $\frac{\delta\varphi}{\delta m} : \mathcal{P}^2(\mathbb{T}^d) \times \mathbb{R}^d \to \mathbb{R}$ is a flat derivative iff, for any $m' \in \mathcal{P}^2(\mathbb{R}^d)$,

$$\lim_{s\downarrow 0} \frac{\varphi((1-s)m + sm') - \varphi(m)}{s} = \int_{\mathbb{T}^d} \frac{\delta\varphi}{\delta m}(m,y)[m'(dy) - m(dy)].$$

Definition

The function $\nabla_m \varphi$ defined by the rule

$$\nabla_m \varphi(m,y) \triangleq \nabla_y \frac{\delta\varphi}{\delta m}(m,y)$$

is called an intrinsic derivative of the function φ.

Lower directional derivative

Let

- $c > 0$, $\mathbb{B}_c$ stand for the ball of radius c,
- $s \in [0, T]$, $m \in \mathcal{P}^2(\mathbb{T}^d)$,
- $\zeta \in \mathcal{P}(\mathbb{T}^d \times \mathbb{B}_c)$,
- $\Theta^\tau(x, v) \triangleq x + \tau v$,
- $\mathrm{p}^{i_1,\dots,i_k}(x_1, \dots, x_n) \triangleq (x_{i_1,\dots,x_{i_k}})$.

$$\mathrm{d}_c^- \varphi(s, \zeta) \triangleq \liminf_{\substack{\zeta' \in \mathcal{P}(\mathbb{T}^d \times \mathbb{B}_c), \quad \mathrm{p}^1 \sharp \zeta = m \\ \tau \downarrow 0, \quad W_2(\zeta', \zeta) \downarrow 0}} \frac{\varphi(s + \tau, \Theta^\tau \sharp \zeta') - \varphi(s, m)}{\tau}.$$

Upper directional derivative

Let

- $c > 0$, $\mathbb{B}_c$ stand for the ball of radius c,
- $s \in [0, T]$, $m \in \mathcal{P}^2(\mathbb{T}^d)$,
- $\alpha \in \mathcal{P}(\mathbb{T}^d \times U)$, $\mathrm{p}^1 \sharp \alpha = m$.
- $\eta \in \mathcal{P}(\mathbb{T}^d \times U \times \mathbb{B}_c)$,
- $\Theta^\tau(x, u, v) \triangleq x + \tau v$.

$$\mathrm{d}_c^+ \varphi(s, \eta) \triangleq \liminf_{\substack{\eta' \in \mathcal{P}(\mathbb{T}^d \times U \times \mathbb{B}_c), \quad \mathrm{p}^{1,2} \sharp \eta = \alpha \\ \tau \downarrow 0, \quad W_2(\eta', \eta) \downarrow 0}} \frac{\varphi(s + \tau, \Theta^\tau \sharp \zeta') - \varphi(s, m)}{\tau}.$$

Admissible distributions

Let

- $s \in [0, T]$,
- $m \in \mathcal{P}^2(\mathbb{T}^d)$.

$$\mathcal{F}(s, m) \triangleq \{\zeta \in \mathcal{P}(\mathbb{T}^d \times \mathbb{R}^d) : \mathrm{supp}(\zeta) \subset F(s, m)\},$$

where

$$F(s, m) \triangleq \big\{(x, v) \in \mathbb{T}^d \times \mathbb{R}^d : \ v \in \mathrm{co}\{f(t, x, m, u) : u \in U\}\big\}.$$

Minimax solution of the Bellman equation

$$\frac{\partial \varphi}{\partial t} + \mathcal{H}(t, m, \nabla_m \varphi) = 0, \quad \varphi(T, m) = \sigma(m),$$

A function φ is a minimax solution to the Bellman equation if

- $\varphi(T, m) =$
- there exists $c > 0$ such that, for any $s \in [0, T]$, $m \in \mathcal{P}^2(\mathbb{T}^d)$,

$$\inf\{\mathrm{d}_c^- \varphi(s, \zeta) : \zeta \in \mathcal{F}(s, m)\} \leq 0;$$

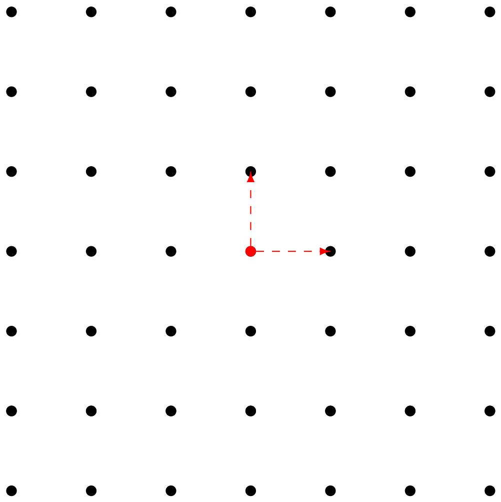
- there exists $c > 0$ such that, for any $s \in [0, T]$, $m \in \mathcal{P}^2(\mathbb{T}^d)$, $\alpha \in \mathcal{P}(\mathbb{T}^d \times U)$, $\mathrm{p}^1 \sharp \alpha = m$, $\eta = (\mathrm{Id}, f(s, \cdot, m, \cdot))\sharp \alpha$,

$$\mathrm{d}_c^+ \varphi(s, \eta) \geq 0.$$

Minimax solution and value function

Theorem. The value function of the mean field type control problem satisfies the Bellamn equation in the minimax sense.

Lattice approximation



Markov chains

Let

- $\mathcal{S}$ be a finite set;
- $\mathcal{S} \subset G$;
- Σ be a simplex on $\{1, \dots, |\mathcal{S}|\}$:

$$\Sigma \triangleq \left\{ \mu = (\mu_{\bar{x}})_{\bar{x} \in \mathcal{S}} : \mu_{\bar{x}} \geq 0, \sum_{x \in \mathcal{S}} \mu_{\bar{x}} = 1 \right\};$$

- $\mathbb{1}_{\bar{y}} = (\mathbb{1}_{\bar{y},\bar{x}})_{\bar{x} \in \mathcal{S}}$ be a pure state; here

$$\mathbb{1}_{\bar{y},\bar{x}} = \begin{cases} 1, & \bar{x} = \bar{y}, \\ 0, & \bar{x} \neq \bar{y}. \end{cases}$$

Σ vs $\mathcal{P}(\mathcal{S})$

- $\Sigma \subset \mathbb{R}^{|\mathbb{S}|}$;
- $\mu^1 = (\mu^1_{\bar{x}})_{\bar{x}\in\mathcal{S}}, \mu^2 = (\mu^2_{\bar{x}})_{\bar{x}\in\mathcal{S}} \in \mathbb{R}^{|\mathcal{S}|}$,

$$\|\mu^1 - \mu^2\|_p \triangleq \left[\sum_{x\in\mathcal{S}} |\mu^1_{\bar{x}} - \mu^2_{\bar{x}}|^p\right]^{1/p};$$

- Isomorphism between Σ and $\mathcal{P}(\mathcal{S})$

$$(\mu_{\bar{x}})_{\in\mathcal{S}} = \mu \mapsto \tilde{\mu} = \sum_{\bar{x}\in\mathcal{S}} \mu_{\bar{x}}\delta_{\bar{x}}.$$

Σ vs $\mathcal{P}(\mathcal{S})$

There exists constants C_1 and C_2 such that

$$\|\mu^1 - \mu^2\|_p \leq C_1 W_p(\widetilde{\mu^1}, \widetilde{\mu^2}),$$

$$W_p(\widetilde{\mu^1}, \widetilde{\mu^2}) \leq C_2(\|\mu^1 - \mu^2\|_p)^{1/p}.$$

Continuous-time Markov chain

Let

- $\mathcal{S}$ be the set of states;
- $Q_{\bar{x},\bar{y}}(t)$ be the transition rate from $\bar{x}$ to $\bar{y}$;
- $Q_{\bar{x},\bar{y}}(t) \geq 0$ if $\bar{x} \neq \bar{y}$;
- $Q_{\bar{x},\bar{x}}(t) = -\sum_{\bar{y}\neq\bar{x}} Q_{\bar{x},\bar{y}}(t)$.

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

- conditional probability of transition from $\bar{x}$ to $\bar{y}$ is

$$Q_{\bar{x},\bar{y}}(t)\Delta t + o(\Delta t),$$

- condition probability of remaining at $\bar{x}$ is

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

Dynamics of probabilities

Denote

- the probability of being at $\bar{x}$ at time t by $\mu_{\bar{x}}(t)$;
- $\mu(t) = (\mu_{\bar{x}}(t))_{\bar{x} \in \mathcal{S}} \in \Sigma$.

Kolmogorov equation

$$\frac{d}{dt}\mu_{\bar{y}}(t) = \sum_{\bar{x} \in \mathcal{S}} \mu_{\bar{x}} Q_{\bar{x},\bar{y}}(t)$$

or in the vector form

$$\frac{d}{dt}\mu(t) = \mu(t)Q(t), \quad \mu(t_0) = \mu_0,$$

where

- $Q(t) = (Q_{\bar{x},\bar{y}}(t))_{\bar{x},\bar{y} \in \mathcal{S}}$ is the Kolmogorov matrix,
- μ_0 is the initial distribution.

Nonlinear Markov chains

Assume that the transition rates depend on the current distribution of the agents.

- Kolmogorov matrix: $Q(t, \mu) = (Q_{\bar{x},\bar{y}}(t, \mu))_{\bar{x},\bar{y} \in \mathcal{S}}$;
- Kolmogorov equation:

$$\frac{d}{dt}\mu(t) = \mu(t)Q(t, \mu(t)).$$

Mean field type finite state control problem

- a decision maker controls infinitely many agents;
- distribution of agents $\mu = (\mu_{\bar{x}})_{\bar{x} \in \mathcal{S}} \in \Sigma$;
- initial distribution of agents is μ_0;
- dynamics of each agents is given by the Markov chain with the Kolmogorov matrix $Q(t, \mu, u) = (Q_{\bar{x},\bar{y}}(t, \mu, u))_{\bar{x},\bar{y} \in \mathcal{S}}$, $u \in U$;
- $X(\cdot)$ is a stochastic process describing the state of agents;
- The decision maker tries to minimize

$$\hat{\sigma}(\mu(T)).$$

Mean field type finite state control problem

- Dynamics: (Kolmogorov equation)

$$\frac{d}{dt}\mu(t) = \mu(t)Q(t, \mu(t), u(t)).$$

- Payoff:

$$\hat{\sigma}(\mu(T)).$$

Markov decision problem. Assumptions

- for every $(t, \mu, u) \in [0, T] \times \Sigma \times U$, $Q_{\bar{x},\bar{y}}(t, \mu, u) \geq 0$ when $\bar{x} \neq \bar{y}$ and
$$\sum_{\bar{y} \in \mathcal{S}} Q_{\bar{x},\bar{y}}(t, \mu, u) = 0;$$
- the functions $Q_{\bar{x},\bar{y}}$ and $\hat{\sigma}$ are continuous;
- there exists a constant L' such that, for any $t \in [0, T]$, $\bar{x}, \bar{y} \in \mathcal{S}$, $\mu^1, \mu^2 \in \Sigma$, $u \in U$,
$$|Q_{\bar{x},\bar{y}}(t, \mu^1, u) - Q_{\bar{x},\bar{y}}(t, \mu^2, u)| \leq L' \|\mu^1 - \mu^2\|_2.$$

Feedback controls

- We assume that the control depends on the time t and the state $\bar{x}$.
- Profile of controls: $u_{\mathcal{S}}(t) \triangleq (u_{\bar{x}}(t))_{\bar{x}\in\mathcal{S}}$.
- Set of profile of controls: $U^{\mathcal{S}}$.
- Kolmogorov matrix: if $u_{\mathcal{S}} \in U^{\mathcal{S}}$, then

$$\mathcal{Q}(t,\mu,u_{\mathcal{S}}) = (\mathcal{Q}_{\bar{x},\bar{y}}(t,\mu,u_{\bar{x}}))_{\bar{x},\bar{y}\in\mathcal{S}}.$$

Control problem

- control $\xi_{\mathcal{S}}(\cdot) = (\xi_{\bar{x}}(\cdot))_{\bar{x}\in\mathcal{S}}$;
- dynamics:

$$\frac{d}{dt}\mu_{\bar{y}}(t) = \sum_{\bar{x}\in\mathcal{S}} \mu_{\bar{x}}(t)\mathcal{Q}_{\bar{x},\bar{y}}(t,\mu(t),u_{\bar{x}}(t)), \quad \bar{y}\in\mathcal{S}$$

  or in the vector form

$$\frac{d}{dt}\mu(t) = \mu(t)\mathcal{Q}(t,\mu(t),u_{\mathcal{S}}(t)),$$

- payoff function:

$$\mathcal{I}(\mu(\cdot),\xi_{\mathcal{S}}(\cdot)) = \hat{\sigma}(\mu(T)).$$

Bellman equation

Hamiltonian

For $t \in [0, T]$, $\mu = (\mu_{\bar{x}})_{\bar{x} \in \mathcal{S}} \in \Sigma$, $w = (w_{\bar{x}})_{\bar{x} \in \mathcal{S}} \in \mathbb{R}^{\mathcal{S}}$,

$$\mathcal{H}^{\mathcal{Q}}(t, \mu, w) \triangleq \sum_{\bar{x} \in \mathcal{S}} \mu_{\bar{x}} \min_{u_{\bar{x}} \in U} \left[\sum_{\bar{y} \in \mathcal{S}} \mathcal{Q}_{\bar{x},\bar{y}}(t, \mu, \xi_{\bar{x}}) w_{\bar{y}} \right].$$

Bellman equation

$$\frac{\partial \varphi}{\partial t} + \mathcal{H}^{\mathcal{Q}}(t, \mu, \nabla \varphi) = 0, \quad \varphi(T, \mu) = \hat{\sigma}(\mu),$$

where $\nabla \varphi = (\partial \varphi / \partial \mu_{\bar{x}})_{\bar{x} \in \mathcal{S}}$.

Viscosity supersolution of Bellman equation

A lower semicontinuous function φ is a viscosity supersolution of the Bellman equation if

$$\varphi(T,\mu) \geq \hat{\sigma}(\mu) \text{ for every } \mu \in \Sigma$$

and

$$a + \mathcal{H}^{\mathcal{Q}}(t,\mu,w_{\mathcal{S}}) \leq 0,$$
$$s \in [0,T],\ \mu \in \Sigma,\ (a,w_{\mathcal{S}}) \in D^{-}\varphi(t,\mu),$$

where

$$D^{-}\varphi(t,\mu) = \{(a,w_{\mathcal{S}}) : \varphi(\tau,\vartheta) - \varphi(t,\mu) \geq a(\tau - t) + \sum_{\bar{x}\in\mathcal{S}} w_x(\vartheta_{\bar{x}} - \mu_{\bar{x}}) + o(|t-s| + \|\vartheta - \mu\|_2)\}.$$

Minimax supersolution of Bellman equation

A lower semicontinuous function φ is a minimax supersolution of the Bellman equation if, for every $s \in [0, T]$, $\mu \in \Sigma$,

$$\varphi(T, \mu) \geq \hat{\sigma}(\mu)$$

and

$$\inf\{\mathrm{d}^- \varphi(s, \mu, 1, w) : w \in \mathcal{G}(s, \mu)\} \leq 0,$$

where

$$\mathrm{d}^- \varphi(s, \mu, a, w) \triangleq \liminf_{\substack{\tau \downarrow 0, \\ (a', w') \to (a, w)}} \frac{\varphi(s + \tau a', \mu + \tau w') - \varphi(s, \mu)}{\tau},$$

$$\mathcal{G}(s, \mu) \triangleq \mathrm{co}\{\mu \mathcal{Q}(t, \mu, u_{\mathcal{S}}) : u_{\mathcal{S}} \in U^{\mathcal{S}}\}.$$

Viscosity subsolution of Bellman equation

An upper semicontinuous function φ is a viscosity subsolution of the Bellman equation if

$$\varphi(T,\mu) \leq \hat{g}(\mu) \text{ for every } \mu \in \Sigma$$

and

$$a + \mathcal{H}^{\mathcal{Q}}(t,\mu,w_{\mathcal{S}}) \geq 0,$$
$$s \in [0,T],\ \mu \in \Sigma,\ (a,w_{\mathcal{S}}) \in D^+\varphi(t,\mu)$$

$$D^+\varphi(t,\mu) = \{(a,w_{\mathcal{S}}) : \varphi(\tau,\vartheta) - \varphi(t,\mu) \leq a(\tau - t) + \sum_{\bar{x}\in\mathcal{S}} w_x(\vartheta_{\bar{x}} - \mu_{\bar{x}}) + o(|t-s| + \|\vartheta - \mu\|_2)\}.$$

Minimax subsolution of Bellman equation

An upper semicontinuous function φ is a minimax subsolution of the Bellman equation if, for every $s \in [0, T]$, $\mu \in \Sigma$,

$$\varphi(T, \mu) \leq \hat{\sigma}(\mu)$$

and

$$\mathrm{d}^+ \varphi(s, \mu, 1, w) \geq 0,$$

for each $w = \mu \mathcal{Q}(s, \mu, u_{\mathcal{S}})$, $u_{\mathcal{S}} \in U^{\mathcal{S}}$.
Here

$$\mathrm{d}^+ \varphi(s, \mu, a, w) \triangleq \limsup_{\substack{\tau \downarrow 0, \\ (a', w') \to (a, w)}} \frac{\varphi(s + \tau a', \mu + \tau w') - \varphi(s, \mu)}{\tau}.$$

Minimax/viscosity solution of Bellman equation

A continous function φ is called a minimax/viscosity solution of the Bellman equation if it is sub- and supersolution simultaneously.

Subtraction on $\mathbb{T}^d$

Let $\ell : \mathbb{T}^d \times \mathbb{T}^d \to \mathbb{R}^d$ be a measurable function assigning to a pair of elements $x, y \in \mathbb{T}^d$ a vector $z' \in x - y$ of the minimal norm.

Approximation condition

$$\max_{x \in \mathbb{T}^d} \min_{\bar{y} \in \mathcal{S}} \|x - \bar{y}\| \leq \varepsilon;$$

$$\max_{t \in [0,T], \bar{x} \in \mathcal{S}, \mu \in \Sigma, u \in U} \Bigg\| f(t, \bar{x}, \tilde{\mu}, u) - \sum_{\bar{y} \in \mathcal{S}, \bar{y} \neq \bar{x}} \ell(\bar{y}, \bar{x}) Q_{\bar{x},\bar{y}}(t, \mu, u) \Bigg\| \leq \varepsilon,$$

$$\max_{t \in [0,T], \bar{x} \in \mathcal{S}, \mu \in \Sigma, u \in U} \sum_{\bar{y} \in \mathcal{S}} \|\bar{y} - \bar{x}\|^2 Q_{\bar{x},\bar{y}}(t, \mu, u) \leq \varepsilon^2.$$

Proximal elements

For $m \in \mathcal{P}^2(\mathbb{T}^d)$, denote by $\mathrm{pr}_{\mathcal{S}}(m)$ an element of Σ such that $\mathrm{pr}_{\mathcal{S}}(m)$ is a proximal to m element of $\mathcal{P}^2(\mathcal{S})$, i.e., $\mathrm{pr}_{\mathcal{S}}(m)$ minimize the function

$$\Sigma \ni \mu = (\mu_{\bar{x}})_{\bar{x}\in\mathcal{S}} \mapsto W_2(\tilde{\mu}, m) = W_2\left(\sum_{\bar{x}\in\mathcal{S}} \mu_{\bar{x}}\delta_{\bar{x}}, m\right).$$

Designation

- Payoff function:
$$\hat{\sigma}(\mu) \triangleq \sigma(\tilde{\mu}).$$
Here, given $\mu = (\mu_{\bar{x}})_{\bar{x}\in\mathcal{S}} \in \Sigma$,
$$\tilde{\mu} \triangleq \sum_{\bar{x}\in\mathcal{S}} \mu_{\bar{x}}\delta_{\bar{x}},$$
$\delta_{\bar{x}}$ is the Dirac measure concentrated in $\bar{x}$.
- Modulus of continuity: Let $\varsigma(\cdot)$ be a modulus of continuity of the function $\sigma(m)$.
- Constant: $C^* \triangleq \sqrt{1+2T}e^{(2L+1/2)T}$, where L is the Lipschitz constant for f w.r.t. to x and m.

Approximation theorem

Let φ_Q be a solution of the Bellman equation for the mean field type finite state control problem with state space $\mathcal{S}$ and Kolmogorov matrix Q. Then, for every $t_0 \in [0, T]$, $m_0 \in \mathcal{P}^2(\mathbb{T}^d)$,

$$|\operatorname{Val}(t_0, m_0) - \varphi_Q(t_0, \operatorname{pr}_{\mathcal{S}}(m_0))| \leq \varsigma(C^* \varepsilon).$$

Here Val denotes the value function for the original mean field type control problem.

Lattice Markov chain

Let

- $h > 0$ be such that $1/h \in \mathbb{N}$
- $\mathcal{S}_h \triangleq h\mathbb{Z}^d \cap \mathbb{T}^d$.
- $f(t, x, m, u) = (f_1(t, x, m, u), \ldots, f_d(t, x, m, u))$,
- e^i stand for the i-th coordinate vector.

$$Q^h_{\bar{x},\bar{y}}(t, \mu, u) \triangleq \begin{cases} \frac{1}{h}|f_i(t, x, \tilde{\mu}, u)|, & \bar{y} = \bar{x} + h \\ & \cdot \operatorname{sgn}(f_i(t, x, \tilde{\mu}, u))e^i, \\ -\frac{1}{h}\sum_{j=1}^d |f_j(t, x, \tilde{\mu}, u)|, & \bar{x} = \bar{y}, \\ 0, & \text{otherwise.} \end{cases}$$

Distance between lattice Markov chain and original system

Let

$$\|f(t, x, m_1, u) - f(t, x, m_2, u)\| \leq L'' W_1(m_1, m_2)$$

for some constant L''.

If the matrix Q is the lattice Markov chain defined as above, then it approximates the origanl system with

$$\varepsilon = \sqrt{h} \cdot \max\left\{\sqrt{R}\sqrt[4]{d}, \frac{\sqrt[4]{d}}{\sqrt{2}}\right\}.$$

Hamiltonian for the lattice Markov chain

$$\mathcal{H}^{\mathcal{Q}}(t,\mu,w) = \frac{1}{h}\sum_{\bar{x}\in\mathcal{S}_h}\min_{u_{\bar{x}}\in U}\sum_{i=1}^{d}|f_i(t,\bar{x},\mu,u)| \left(w_{\bar{x}+he^i\,\mathrm{sgn}(f_i(t,\bar{x},\mu,u))} - w_{\bar{x}}\right).$$

Conclusion

- Viscosity solution of the Bellman equation for mean field type control problems.
- Construction of feedback strategies for mean field type control problems based on Bellman equation.
- Unified form of minimax solutions.

Thank you for your attention!