

Игры среднего поля первого порядка. Основные подходы

Ю.В. Авербух
ayv@imm.uran.ru

ИММ УрО РАН (Екатеринбург)
УрФУ (Екатеринбург)

Аннотация

Настоящий обзор представляет собой попытку дать описание основных направлений и результатов игр среднего поля первого порядка для русскоязычного читателя. В настоящей работе отражены основные теоретические результаты игр среднего поля первого порядка (т.е. игр среднего поля с динамикой, задаваемой обыкновенным дифференциальным уравнением). Кроме того, рассмотрена модель энергетических сетей, формализуемая в рамках теории игр среднего поля первого порядка.

1 Введение

Теория игр среднего поля – относительно молодая часть теории оптимального управления, начавшаяся с работ [31]–[34], [49]–[51]. Она ставит перед собой целью исследовать задачи управления для системы однотипных взаимодействующих субъектов путем рассмотрения предельной ситуации – системы бесконечного числа взаимодействующих субъектов.

Необходимость такого подхода вызвана экстремальной сложностью, возникающей при применении традиционных подходов (динамического программирования, методов, основанных на “народных” теоремах) к исследованию систем большого числа взаимодействующих субъектов, называемых также игроками.

Систему взаимодействующих субъектов можно описывать в рамках теории динамических игр. Предполагается, что изменение положения каждого игрока задается некоторой динамической системой. Обычно эта система описывается обыкновенным дифференциальным уравнением, стохастическим дифференциальным уравнением или марковской цепью. На динамику влияет как положение игрока и выбранное им управляющее воздействие, так и положение остальных игроков. Каждый игрок стремится к максимизации собственного выигрыша, который также зависит от траектории данного игрока, его управляющего воздействия и траектории всех игроков. Для широкого класса задач адекватным является определение решения в терминах равновесия по Нэшу – ситуации, когда изменение поведения одного игрока не ведет к увеличению его выигрыша.

Напомним, что в рамках подхода, основанного на динамическом программировании, задача поиска равновесия по Нэшу сводится к решению системы уравнений в частных производных [54, 23]. Количество уравнений в системе равно количеству игроков. Это делает почти невозможным применение динамического программирования для анализа систем большого количества взаимодействующих объектов. Второй подход основан на стратегиях наказания и идейно восходит к “народным” теоремам в теории повторяющихся игр [38, 12]. Его применение в теории игр с непрерывным временем приводит к необходимости решать большое количество вспомогательных антагонистических игр. Эта задача является весьма трудоемкой.

Copyright © by the paper's authors. Copying permitted for private and academic purposes.

In: A.A. Makhnev, S.F. Pravdin (eds.): Proceedings of the 47th International Youth School-conference “Modern Problems in Mathematics and its Applications”, Yekaterinburg, Russia, 02-Feb-2016, published at <http://ceur-ws.org>

Также стратегии, которые строятся в рамках этого подхода, основаны на угрозах, которые стоит признать неправдоподобными.

Идея теории игр среднего поля состоит в том, чтобы рассмотреть систему бесконечного числа игроков. При этом влияние каждого игрока становится пренебрежимо малым, а значение имеет только масса игроков. В получившейся системе каждый игрок оптимизирует свое поведение в соответствии с движением всей массы игроков, при этом поведение массы игроков задается суммой поведения игроков.

Отметим основные задачи, которые рассматриваются в теории игр среднего поля:

- исследование существования и единственности решения игры среднего поля; при этом игра среднего поля формализуется в виде системы двух уравнений в частных производных или задачи управления для некоторой динамической системы;
- построение приближенного равновесия в игре конечного числа лиц по заданному решению игры среднего поля;
- исследование сходимости равновесий в играх конечного числа лиц к решению игры среднего поля при стремлении числа игроков к бесконечности;
- анализ конкретных прикладных задач при помощи методологии игр среднего поля.

В настоящее время существует несколько подходов к определению и построению решения игры среднего поля. Исторически первый подход основан на прямом анализе системы двух уравнений: первое уравнение – уравнение Беллмана, описывающее выигрыш игроков (игроки предполагаются однотипными), второе уравнение – уравнение Колмогорова, описывающее распределение игроков. Отметим, что краевое условие для уравнения Беллмана задается в последний момент времени, а для уравнения Колмогорова – в начальный. Этот подход был развит в работах J.-M. Lasry и P.-L. Lions [49]–[51], M. Huang, R.P. Malhamé и P. Caines (см. [31]–[34]). В этих работах преимущественно рассматривались системы, заданные стохастическими дифференциальными уравнениями. В работах [51, 52] был рассмотрен случай динамики, задаваемой обыкновенным дифференциальным уравнением. Также здесь отметим работу [25], где рассмотрены игры среднего поля с динамикой каждого игрока, задаваемой марковской цепью с непрерывным временем и конечным числом состояний. Отметим также, что в рамках подхода, основанного на анализе системы дифференциальных уравнений, были изучены игры среднего поля с динамикой, задаваемой марковским процессом достаточно общего вида (см. [42, 39, 43]). К этому же подходу примыкает идея слабых решений для детерминированных игр среднего поля, предложенная в [13] (см. также [15, 29, 52]). В этом случае оба уравнения заменяются на специально построенные задачи оптимизации.

Альтернативный подход к играм среднего поля называется вероятностным и связан с исследованием задачи управления для динамической системы, описываемой нелинейным марковским процессом. В этом случае динамика и интегральная часть выигрыша зависят от распределения игроков в текущий момент времени. Распределение игроков определяется по решению задачи оптимизации. Этот подход был развит для стохастических игр среднего поля в серии работ [16]–[19]. В частности, в рамках этого подхода были получены аналоги принципа максимума Понтрягина для игр среднего поля. Для детерминированных игр среднего поля вероятностный подход был развит в [28] (см. также [27]). Некоторым недостатком вероятностного подхода является зависимость определения решения от выбора сопутствующего вероятностного пространства. Этот недостаток пытается преодолеть минимаксный подход, в рамках которого задача поиска решения игры среднего поля сводится к решению игры бесконечного числа лиц; при этом динамика определяется распределением оптимальных траекторий.

Игра среднего поля является предельным (“искусственным”) объектом. Поэтому необходимо исследовать вопрос о связи равновесия в игре конечного (но большого) числа игроков и решения игры среднего поля. Построение приближенных решений в играх конечного числа игроков рассматривалось для стохастических игр в работах [52, 43]. В этом случае строятся позиционные стратегии. Для детерминированных игр среднего поля приближенное равновесие по Нэшу рассматривалось в классе программных стратегий [4, 13].

Предел решений игр многих лиц при стремлении числа игроков к бесконечности как решение игры среднего поля исследовался в работах [22, 46]. Было показано, что если в игре N лиц с однотипными игроками и динамикой каждого игрока, определяемой положением игрока, его управлением и симметричной функцией от положения всех игроков, существует универсальное равновесие по Нэшу (т.е. набор стратегий, не зависящий от выбора начальных положений игроков), то все предельные точки эмпирических распределений игроков соответствуют решениям игры среднего поля в вероятностном смысле.

Отметим, что в настоящее время разработаны численные методы для игр среднего поля. Они преимущественно охватывают стохастические игры среднего поля [1]–[3].

Еще одно направление игр среднего поля связано с так называемым “основным уравнением” (master equation). В рамках этого подхода игра среднего поля описывается одним уравнением, записанным в терминах производных в функциональных пространствах или пространствах вероятностей. Исследование основного уравнения было инициировано Р.-Л. Lions [52]. Этой тематике посвящены работы [11, 14, 28, 41]. Для детерминированных игр среднего поля было введено понятие вязкостного решения и доказана теорема существования этого решения [24].

Многие задачи включают в себя модели взаимодействия множества “слабых” агентов с одним или несколькими могущественными игроками. Такие задачи формализованы в рамках игр среднего поля с ведущим игроком [10, 20, 30, 53].

Отметим, что теория игр среднего поля находит многочисленные применения в финансовом анализе [51], в экономических задачах [47, 26], в моделировании движения [48] и в моделировании социально-экономических задач, в частности вопросов коррупции и преступности [40, 44]. Также многочисленны технические приложения, в частности к задачам моделирования электрических и энергетических сетей [35]–[37], [8].

Настоящий обзор представляет собой попытку дать описание основных направлений и результатов игр среднего поля первого порядка, т.е. игр среднего поля, которые описываются системой дифференциальных уравнений первого порядка, что соответствует детерминированным динамическим системам, в рамках “вероятностного подхода”. Кроме того, рассмотрена модель энергетических сетей, формализуемая в рамках теории игр среднего поля первого порядка. Обзор основных результатов, касающихся игр среднего поля второго порядка, дан в [10, 27].

2 Эмпирический вывод системы уравнений для игры среднего поля

В большинстве работ, посвящённых играм среднего поля, задача ставится аксиоматически. Однако, представляется полезным получить постановки теории игр среднего поля, исходя из теории игр конечного числа лиц. Данный вывод является математически нестрогим. В целом вывод следует работам [39, 51].

Прежде всего напомним определение равновесия по Нэшу. Рассмотрим игру N лиц, пусть Z_i , $i = 1, \dots, N$, – множество стратегий i -го игрока. Выигрыш каждого игрока определяется функцией полезности $J_i(z_1, \dots, z_N)$. Если $z = (z_1, \dots, z_k, \dots, z_N) \in Z_1 \times \dots \times Z_N$ – профиль стратегий, $w_i \in Z_k$ – стратегия игрока i , то обозначим через $z|w_i$ профиль стратегий, полученный из профиля z путём замены стратегии z_i игрока i на стратегию w_i , т.е.

$$z|w_i = (z_1, \dots, z_{i-1}, w_i, z_{i+1}, \dots, z_N).$$

Профиль стратегий z^* называется равновесием по Нэшу, если для всех $i \in \overline{1, N}$ и $w_i \in Z_i$

$$J_i(z^*|w_i) \leq J_i(z^*).$$

Далее мы рассмотрим дифференциальные игры многих лиц с разделённой динамикой, т.е. будем предполагать, что i -й игрок непосредственно воздействует только на динамику вектора $x_i \in \mathbb{R}^d$:

$$\dot{x}_i = f_i(t, x_1, \dots, x_N, u_i), \quad t \in [0, T], \quad u \in U.$$

Более того, мы будем предполагать симметричную зависимость уравнений динамики от x_j , $j \neq i$. Для того чтобы упростить дальнейшее изложение, мы предполагаем, что правая часть динамики вектора x_i зависит от времени, текущего состояния, эмпирического распределения игроков и управления самого i -го игрока. Под эмпирическим распределением игроков мы понимаем вероятность

$$\delta_{x(t)}^N \triangleq \frac{1}{N} (\delta_{x_1(t)} + \dots + \delta_{x_N(t)}).$$

Таким образом, мы рассматриваем управляемую систему

$$\dot{x}_i = f(t, x_i, \delta_{x(t)}^N, u_i), \quad i = \overline{1, N}. \quad (1)$$

Предполагается, что каждый игрок стремится максимизировать свой выигрыш, состоящий из терминальной и интегральной частей. Мы будем предполагать, что выигрыш также зависит от эмпирического распределения игроков:

$$J_i(u_1, \dots, u_N) \triangleq \sigma(x_i(T), \delta_{x(T)}^N) + \int_0^T g(t, x_i(t), \delta_{x(t)}^N, u_i(t)) dt. \quad (2)$$

В качестве примера зависимости от эмпирического распределения можно рассматривать зависимость от среднего. В этом случае

$$f(t, x_i, \delta_x^N, u) = \hat{f}\left(t, x_i, \sum_{j=1}^N x_j, u_i\right) = \hat{f}\left(t, x_i, \int_{\mathbb{R}^d} y \delta_x^N(dy), u_i\right),$$

$$\sigma(x_i, \delta_x^N) = \hat{\sigma}\left(x_i, \sum_{j=1}^N x_j\right) = \hat{\sigma}\left(x_i, \int_{\mathbb{R}^d} y \delta_x^N(dy)\right),$$

$$g(t, x_i, \delta_x^N, u) = \hat{g}\left(t, x_i, \sum_{j=1}^N x_j, u_i\right) = \hat{g}\left(t, x_i, \int_{\mathbb{R}^d} y \delta_x^N(dy), u_i\right).$$

Отметим, что множество взвешенных сумм мер Дирака плотно во множестве $\mathcal{P}^1(\mathbb{R}^d)$, состоящем из вероятностных мер m таких, что конечен первый абсолютный момент, т.е.

$$\int_{\mathbb{R}^d} \|y\| m(dy) < \infty.$$

Поэтому можно считать, что функции f , g определены для всех t , x , вероятностных мер с конечным первым абсолютным моментом m и управлений u . Аналогично, функция σ определена для состояний x и вероятностей с конечным первым абсолютным моментом m .

Мы будем рассматривать концепцию равновесия по Нэшу для игры с динамикой (1) и показателем качества (2). Кроме того, мы предполагаем существование универсального равновесия в классе чистых позиционных стратегий.

Замечание 1. Вообще говоря, нельзя говорить о существовании универсального равновесия по Нэшу для дифференциальной игры общего вида в классе чистых позиционных стратегий. Для игр двух лиц построен пример, показывающий, что в чистых стратегиях нельзя построить универсальную (т.е. не зависящую от начальной позиции) оптимальную позиционную стратегию [57]. Из этого следует и невозможность построить универсальную равновесную стратегию уже в игре двух лиц. В настоящий момент доказано существование равновесия по Нэшу в классе позиционных стратегий, зависящих от параметра ε [38]. В этом случае движения строятся пошагово, на основе конструкций, предложенных в теории антагонистических дифференциальных игр Н.Н. Красовским и А.И. Субботиным [55] (см. также конструкцию универсальных стратегий в [55]). Отметим, что равновесие достигается лишь в предельном случае, когда длина шага и параметр ε стремятся к нулю. Необходимым условием является и согласованность параметров мелкости у стратегий игроков. Также универсальные в заданном компакте стратегии могут быть построены в классе управлений с поводырем [6, 7].

Поскольку динамика и функции выигрыша зависят от эмпирической меры, то логично считать, что стратегии игроков зависят от позиции и вероятности, описывающей распределение игроков. Пусть $u_i : [0, T] \times \mathbb{R}^d \times \mathcal{P}^1(\mathbb{R}^d) \rightarrow U$ – стратегия i -го игрока. Через u обозначим профиль стратегий. Пусть также $x_0 = (x_{0,1}, \dots, x_{0,N})$ – набор начальных позиций игроков. Обозначим через $x[\cdot, t_0, x_0, u]$ решение задачи Коши:

$$\dot{x}_i(t) = f(t, x_i(t), \delta_{x(t)}^N, u_i(t, x_i(t), \delta_{x(t)}^N)), \quad x_i(t_0) = x_{0,i}, \quad i = \overline{1, N}. \quad (3)$$

Напомним, что

$$\delta_{x(t)}^N = \frac{1}{N} (\delta_{x_1(t)} + \dots + \delta_{x_N(t)}).$$

Результат i -го игрока оценивается функционалом

$$J_i[t_0, x_0, u] = \sigma(x_i(T), \delta_{x(T)}^N) + \int_{t_0}^T g(t, x_i(t), \delta_{x(t)}^N, u_i(t, x_i(t), \delta_{x(t)}^N)) dt. \quad (4)$$

Введем понятие равновесия по Нэшу. Будем говорить, что профиль u^* доставляет равновесие по Нэшу, если для любого номера $i = 1, \dots, N$ и любой стратегии v_i выполнено неравенство

$$J_i[t_0, x_0, u^* | v_i] \leq J_i[t_0, x_0, u^*].$$

Рассмотрим лишь симметричные равновесия по Нэшу. В этом случае u_i^* не зависит от номера i и равна некоторой стратегии \hat{u} . В дальнейшем будем и набор из N стратегий \hat{u} обозначать через \hat{u} . Также можно считать, что \hat{u} есть функция из $[0, T] \times \mathbb{R}^d \times \mathcal{P}^1(\mathbb{R}^d)$ со значениями в U .

Также выпишем уравнение на изменение вероятности $\mu^N[t] \triangleq \delta_{x(t)}^N$. Пусть $\varphi(x)$ – функция класса C^1 , быстро убывающая на бесконечности вместе со своими частными производными. В силу динамики системы имеем, что

$$\begin{aligned} \frac{d}{dt} \varphi(x_i[t, t_0, x_0, \hat{u}]) &= \langle \nabla \varphi(x_i[t, t_0, x_0, \hat{u}]), \dot{x}_i[t, t_0, x_0, \hat{u}] \rangle = \\ &= \langle \nabla \varphi(x_i[t, t_0, x_0, \hat{u}]), f(t, x_i, \delta_{x(t)}^N, \hat{u}(t, x_i(t), \delta_{x(t)}^N)) \rangle. \end{aligned}$$

Отсюда,

$$\begin{aligned} \frac{d}{dt} \int_{\mathbb{R}^d} \varphi(x) \mu^N[t](dx) &= \frac{d}{dt} \left[\frac{1}{N} \sum_{i=1}^N \varphi(x_i[t, t_0, x_0, \hat{u}]) \right] = \\ \sum_{i=1}^N \langle \nabla \varphi(x_i[t, t_0, x_0, \hat{u}]), f(t, x_i, \delta_{x(t)}^N, \hat{u}(t, x_i(t), \delta_{x(t)}^N)) \rangle &= \int_{\mathbb{R}^d} \langle \nabla \varphi(x), f(t, x, \mu^N[t], \hat{u}(t, x, \mu^N[t])) \rangle \mu^N[t](dx). \quad (5) \end{aligned}$$

Отметим, что задача поиска равновесия в игре большого числа лиц чрезвычайно сложна. Ключевая идея теории игр среднего поля – перейти к пределу и исследовать систему бесконечного числа лиц. Прежде всего, отметим, что при больших N вклад каждого игрока в эмпирическое распределение игроков невелик (порядка $\sim 1/N$). В первом приближении им можно пренебречь. Тогда для каждого игрока распределение игроков становится внешним параметром. Обозначим распределение игроков в момент времени t через $\mu[t]$. Заметим, что в предельной игре каждый игрок решает задачу максимизации

$$\sigma(x(T), \mu[T]) + \int_{t_0}^T g(t, x(t), \mu[t], u(t)) dt$$

при условии того, что динамика задаётся уравнением

$$\dot{x} = f(t, x, \mu[t], u(t)), \quad x(t_0) = y.$$

При этом динамика распределения игроков в момент t – вероятности $\mu[t]$ – определяется по оптимальному управлению всех игроков. Необходимо отметить, что при фиксированном распределении игроков задача поиска равновесной по Нэшу стратегии сводится к задаче оптимального управления.

Одним из способов решения задач оптимального управления является метод динамического программирования. В этом случае необходимо рассмотреть уравнение Беллмана

$$\frac{\partial V}{\partial t} + H(t, x, \mu[t], \nabla V) = 0, \quad V(T, x) = \sigma(x, \mu[T]), \quad (6)$$

где H – гамильтониан, вычисляемый по формуле

$$H(t, x, m, p) = \max_{u \in U} [\langle p, f(t, x, m, u) \rangle + g(t, x, m, u)].$$

Здесь m – вероятность, $p \in \mathbb{R}^d$ – сопряжённая переменная.

Решение уравнения (6) надо понимать в минимаксном/вязкостном смысле [56, 9]. Однако, для наших целей эмпирического вывода мы будем понимать решение этого уравнения в классическом смысле. Более того, мы считаем, что с классическим решением данного уравнения ассоциирована оптимальная стратегия $\hat{u}(t, x, m)$, которая удовлетворяет условию

$$\max_{u \in U} [\langle \nabla V(t, x), f(t, x, m, u) \rangle + g(t, x, m, u)] = \langle \nabla V(t, x), f(t, x, m, \hat{u}(t, x, m)) \rangle + g(t, x, m, \hat{u}(t, x, m)).$$

Заметим, что если гамильтониан H является дифференцируемым по сопряженной переменной, то

$$f(t, x, m, \hat{u}(t, x, m)) = \frac{\partial H(t, x, m, \nabla V(t, x))}{\partial p}.$$

В этих предположениях динамика распределения игроков (5) может быть переписана в виде

$$\begin{aligned} \frac{d}{dt} \int_{\mathbb{R}^d} \varphi(x) \mu[t](dx) &= \int_{\mathbb{R}^d} \langle \nabla \varphi(x), f(t, x, \mu[t], \hat{u}(t, x, \mu[t])) \rangle \mu[t](dx) = \\ &= \int_{\mathbb{R}^d} \left\langle \nabla \varphi(x), \frac{\partial H(t, x, \mu[t], \nabla V(t, x))}{\partial p} \right\rangle \mu[t](dx). \end{aligned}$$

В дальнейшем мы будем записывать последнее уравнение в формальном виде

$$\frac{d}{dt} \mu[t] = \left\langle \frac{\partial H(t, x, \mu[t], \nabla V(t, x))}{\partial p}, \nabla \right\rangle \mu[t], \quad \mu[0] = m_0. \quad (7)$$

Здесь m_0 – некоторое начальное распределение игроков, которое может быть получено как предел эмпирических распределений $\delta_{x_0}^N$.

Система (6), (7) и является системой уравнений для игр среднего поля.

Приведённый нами вывод является выводом на физическом уровне строгости и его строгое обоснование, по-видимому, невозможно. Другой способ вывода уравнений для игры среднего поля основан на исследовании равновесия по Нэшу в классе программных управлений [22]. Способ, предложенный в упомянутой работе, применяется к стохастическим дифференциальным играм и предполагает, что начальные позиции задаются случайно. Таким образом, в рамках этого подхода даже эмпирический вывод выходит из области детерминированных систем.

Связь игр среднего поля и игр конечного числа игроков может быть получена одним из двух способов.

1. Доказательством того, что решение системы для игры среднего поля является пределом (частичным пределом) симметричных равновесий в игре конечного числа игроков при стремлении числа игроков к бесконечности. Это сделано в работах [22, 46] для стохастических дифференциальных игр.
2. Построением приближённого равновесия по Нэшу в игре конечного числа игроков на основе решения игры среднего поля. Это сделано во многих работах. Некоторые подобные результаты будут приведены ниже.

3 Определения и обозначения

В предыдущем разделе мы эмпирически вывели систему уравнений

$$\frac{\partial V}{\partial t} + H(t, x, \mu[t], \nabla V) = 0, \quad V(T, x) = \sigma(x, \mu[T]), \quad (8)$$

$$\frac{d}{dt} \mu[t] = \left\langle \frac{\partial H(t, x, \mu[t], \nabla V(t, x))}{\partial p}, \nabla \right\rangle \mu[t], \quad \mu[0] = m_0. \quad (9)$$

Здесь $\sigma(x, m)$ – некоторая функция со значениями в \mathbb{R} , m_0 – некоторое начальное распределение игроков. Отметим, что в данном случае гамильтониан H , а также терминальное условие σ зависят от вероятности в целом.

Чтобы упростить запись, обозначим

$$H_p(t, x, m, \nabla V(t, x)) = \frac{\partial H(t, x, m, \nabla V(t, x))}{\partial p}.$$

Принципиально важным является случай, когда мера $\mu[t]$ при каждом t абсолютно непрерывна относительно меры Лебега. Обозначим через θ плотность меры $\mu[t]$, т.е. θ – это функция класса $L^1([0, T] \times \mathbb{R}^d)$ такая, что

$$\mu[t](A) = \int_A \theta(t, x) dx. \quad (10)$$

Отметим, что представление вероятности через ее плотность задает отображение функций θ из множества $L^1(\mathbb{R}^d)$ таких, что

$$\int_{\mathbb{R}^d} \|x\| \theta(x) dx < \infty$$

во множество вероятностей $\mathcal{P}^1(\mathbb{R}^d)$. Если вероятность $\mu[t]$ имеет плотность $\theta(t, \cdot)$, то, несмотря на некоторую нестрогость обозначений, будем считать, что гамильтониан H есть функция времени, позиции, плотности распределения и сопряженной переменной.

В этом случае уравнение (9) может быть преобразовано. Для всех $\phi \in C^1(\mathbb{R}^d)$, убывающих на бесконечности вместе со своими частными производными с экспоненциальной скоростью, имеем равенство

$$\begin{aligned} \int_{\mathbb{R}^d} \phi(x) \frac{\partial \theta(t, x)}{\partial t} dx &= \frac{d}{dt} \int_{\mathbb{R}^d} \phi(x) \mu[t] dt = \int_{\mathbb{R}^d} \langle H_p(t, x, \theta(t, \cdot), \nabla V(t, x)), \nabla \phi(x) \rangle \mu[t](dx) = \\ &= \int_{\mathbb{R}^d} \langle H_p(t, x, \theta(t, \cdot), \nabla V(t, x)), \nabla \phi(x) \rangle \theta(t, x) dx = - \int_{\mathbb{R}^d} \operatorname{div}(H_p(t, x, \theta(t, \cdot), \nabla V(t, x)) \theta(t, x)) \phi(x) dx. \end{aligned}$$

Оставляя лишь правую и левую части этого равенства, получаем систему уравнений в частных производных для игр среднего поля в предположении того, что распределение вероятностей имеет плотность:

$$\frac{\partial V}{\partial t} + H(t, x, \theta(t, \cdot), \nabla V) = 0, \quad V(T, x) = \sigma(x, \theta(T, \cdot)), \quad (11)$$

$$\frac{\partial \theta(t, x)}{\partial t} + \operatorname{div}(H_p(t, x, \theta(t, \cdot), \nabla V(t, x)) \theta(t, x)) = 0, \quad \theta(0, x) = \theta_0(x). \quad (12)$$

Отметим, что обе системы – (8), (9) и (11), (12) – описывают игры среднего поля с нелокальной связью, т.е. гамильтониан зависит от функции из класса L^1 “в целом”. Теми же методами, что и в предыдущем разделе, можно вывести систему уравнений среднего поля с локальной связью, в этом случае гамильтониан зависит от текущего значения плотности. В этом случае система уравнений для игры среднего поля имеет вид

$$\frac{\partial V}{\partial t} + H(t, x, \theta(t, x), \nabla V) = 0, \quad V(T, x) = \sigma(x, \theta(T, x)), \quad (13)$$

$$\frac{\partial \theta}{\partial t} + \operatorname{div}(H_p(t, x, \theta, \nabla V) \theta) = 0, \quad \theta(0, x) = \theta_0(x). \quad (14)$$

4 Игры среднего поля с нелокальной связью

4.1 Вязкостный подход

Вязкостный подход является исторически первым в теории игр среднего поля, он развит Lasry и Lions в работах [49]–[51]. Кратко опишем основные идеи этого подхода, следуя работе [51]. Мы предполагаем, что распределение игроков имеет плотность,

$$H(t, x, \theta, p) = \|p\|^2/2 + F(x, \theta). \quad (15)$$

Тогда система (8), (9) принимает вид

$$\frac{\partial V}{\partial t} + \frac{1}{2} \|\nabla V\|^2 + F(x, \theta(t, \cdot)) = 0, \quad V(T, x) = \sigma(x, \theta(T, \cdot)), \quad (16)$$

$$\frac{\partial \theta(t, x)}{\partial t} + \operatorname{div}(\nabla V(t, x)) \cdot \theta(t, x) = 0, \quad \theta(0, x) = \theta_0(x). \quad (17)$$

Первое уравнение соответствует задаче максимизации функционала

$$\sigma(x(T)) + \int_0^T \left[-\frac{u^2(t)}{2} + F(x(t), \theta(t, \cdot)) \right] dt$$

на множестве всех траекторий

$$\dot{x} = u, \quad u \in \mathbb{R}^d.$$

Отметим, что оптимальное управление задается равенством $\hat{u}(t, x) = \nabla V(t, x)$.

Накладываются условия непрерывности функций F и σ , их дважды непрерывной дифференцируемости по x . Предполагаем, что сами функции F и σ , их первые и вторые производные по x равномерно ограничены. Также считаем, что θ_0 финитна.

Решение уравнения (16) понимается в вязкостном смысле [56, 9]. Напомним, что функция $\zeta(t, x)$ является вязкостным решением уравнения

$$\frac{\partial \zeta}{\partial t} + h(t, x, \nabla \zeta) = 0, \quad \zeta(T, x) = \zeta_T(x)$$

если ζ удовлетворяет краевому условию и

$$a^- + h(t, x, b^-) \leq 0, \quad \forall (a^-, b^-) \in D^- \zeta(t, x);$$

$$a^+ + h(t, x, b^+) \geq 0, \quad \forall (a^+, b^+) \in D^+ \zeta(t, x).$$

Здесь

$$D^- \zeta(t, x) = \left\{ (a, b) \in \mathbb{R} \times \mathbb{R}^d : \forall (\tau, \xi) \liminf_{\delta \rightarrow 0, \tau' \rightarrow \tau, \xi' \rightarrow \xi} \frac{\zeta(t + \delta\tau', x + \delta\xi') - \zeta(t, x) - a\tau - \langle b, \xi \rangle}{\delta} \geq 0 \right\},$$

$$D^+ \zeta(t, x) = \left\{ (a, b) \in \mathbb{R} \times \mathbb{R}^d : \forall (\tau, \xi) \limsup_{\delta \rightarrow 0, \tau' \rightarrow \tau, \xi' \rightarrow \xi} \frac{\zeta(t + \delta\tau', x + \delta\xi') - \zeta(t, x) - a\tau - \langle b, \xi \rangle}{\delta} \leq 0 \right\}.$$

Решение уравнения (17) понимается в смысле распределений, т.е. мы считаем, что $\theta \in L^1([0, T] \times \mathbb{R}^d)$ – решение уравнения (17), если для всех гладких финитных функций $\phi \in C^1([0, T] \times \mathbb{R}^d)$, обращающихся в ноль при $t = 0$ и $t = T$, выполнено равенство

$$\int_{[0, T] \times \mathbb{R}^d} \left[\frac{\partial \phi(t, x)}{\partial t} + \langle \nabla \phi(t, x), \nabla V(t, x) \rangle \right] \theta(t, x) dt dx.$$

Теорема существования для такого решения имеет два доказательства. Первое доказательство основано на следующих идеях.

1. Оператор Φ , который ставит в соответствие функции θ решение уравнения

$$\frac{\partial \zeta}{\partial t} + \frac{1}{2} \|\nabla \zeta\|^2 + F(t, \theta(t, \cdot)) = 0, \quad (18)$$

является непрерывным. Более того, решения уравнения (18) липшицевы, имеют ограниченные смешанные производные по t и x_i , также в смысле распределений ограничены и вторые частные производные по x .

2. Благодаря явной формуле для вычисления функции θ по заданной ζ , доказывается липшицевость зависимости $t \mapsto \theta(t, \cdot)$ (как элемента $L^1(\mathbb{R}^d)$) и непрерывная зависимость θ от ζ (т.е. локальная равномерная сходимость ζ_n к ζ влечет слабую сходимость соответствующих плотностей вероятностей).
3. Композиция операторов, сопоставляющих потоку вероятностей функцию цены и функции цены – поток вероятностей, задает вполне непрерывный оператор на пространстве потоков вероятностей. Неподвижная точка этого оператора определяет решение системы (16), (17).

Второе доказательство основано на предельном переходе от стохастических игр с гамильтонианом вида (15) среднего поля при стремлении коэффициента вязкости к нулю.

Также имеет место единственность решения системы (16), (17) при дополнительных условиях:

$$\int_{\mathbb{R}^d} [F(x, \theta') - F(x, \theta'')](\theta'(x) - \theta''(x))dx \neq 0 \quad \theta' \neq \theta'',$$

$$\int_{\mathbb{R}^d} [F(x, \theta') - F(x, \theta'')](\theta'(x) - \theta''(x))dx \cdot \int_{\mathbb{R}^d} [\sigma(x, \theta') - \sigma(x, \theta'')](\theta'(x) - \theta''(x))dx \geq 0.$$

Вопрос о том, единственно ли решение без этих условий, остается открытым.

Как видно, конструкция решения в рамках вязкостного подхода существенно опирается на структуру гамильтониана $H(t, x, p, \theta) = \|p\|^2 + F(x, \theta)$. Прежде всего используется явная формула вычисления потока вероятности по заданной почти всюду дифференцируемой функции, эта формула является следствием уравнения (17). Перенос ее на более общий случай невозможен. Это требует других подходов к определению решения.

4.2 Вероятностный подход

В настоящем разделе мы рассматриваем полную систему (8), (9). Вероятностный подход основан на исследовании игры бесконечного числа игроков. А именно, пусть Ω – множество игроков. На этом множестве необходимо ввести некоторую вероятность (измеримую структуру также предполагаем заданной) ϱ . Предположим, что каждый игрок в начальный момент находится в позиции $X_0(\omega)$. От функции X_0 необходимо потребовать согласованности с распределением m_0 , т.е. чтобы для каждого измеримого множества $A \subset \mathbb{R}^d$ $\varrho(X_0^{-1}(A)) = m_0(A)$. Движение каждого игрока описывается уравнением

$$\frac{d}{dt}X(t, \omega) = f(t, X(t, \omega), \mu[t], u(t, \omega)), \quad X(0, \omega) = X_0(\omega), \quad u(t, \omega) \in P. \quad (19)$$

Каждый игрок стремится к максимизации своего выигрыша, который задается величиной

$$\sigma(X(t, \omega), \mu[T]) + \int_0^T g(t, X(t, \omega), \mu[t], u(t, \omega))dt. \quad (20)$$

При этом $\mu[t]$ является законом распределения $X(t, \omega)$ т.е. для любого измеримого множества $A \subset \mathbb{R}^d$ выполнено равенство

$$\mu[t](A) = \varrho(X^{-1}(t, A)).$$

Далее в рамках этого раздела (несколько нестрого) будем считать, что функции f , g , σ и H зависят не от вероятности m , а от функции $Y : \Omega \rightarrow \mathbb{R}^d$, которая определяет вероятность m по правилу

$$m(A) = \varrho(Y^{-1}(A)).$$

Данные предположения позволяют переписать уравнение (8) в виде

$$\frac{\partial V}{\partial t} + H(t, x, X(t, \cdot), \nabla V) = 0, \quad V(T, x) = \sigma(x, X(t, \cdot)). \quad (21)$$

Изменение $X(t, \omega)$ описывается уравнением

$$\frac{d}{dt}X(t, \omega) = \frac{\partial H(t, X(t, \omega), X(t, \cdot), \nabla V(t, X(t, \omega)))}{\partial p}, \quad X(0, \omega) = X_0(\omega). \quad (22)$$

Последнее уравнение получено из (19) и (20), исходя из предположения, что в достаточно гладком случае оптимальное управление игрока $\hat{u}(t, x)$, находящегося в позиции (t, x) , таково, что выполнено соотношение

$$f(t, x, \hat{u}(t, x), X(t, \cdot)) = \frac{\partial H(t, X(t, \omega), X(t, \cdot), \nabla V(t, X(t, \omega)))}{\partial p}.$$

Решение системы (21), (22) будем понимать в следующем смысле:

- уравнение (21) выполняется в минимаксном/вязкостном смысле;
- система уравнений (22) выполнена в смысле Каратеодори, т.е. $X(\cdot, \omega)$ абсолютно непрерывна и почти всюду удовлетворяет дифференциальному уравнению.

Данный подход развит в работе [28] не только для игр среднего поля, но и для расширенных игр среднего поля; в этом случае предполагается зависимость гамильтониана H (и функций f и g) не только от времени t , положения игрока x , положения всех игроков $X(t, \cdot)$ и сопряженной переменной p , но и от скоростей игроков $\dot{X}(t, \cdot)$.

Теорема существования решения игры среднего поля в рамках вероятностного подхода доказана при целом ряде предположений. Для того чтобы сформулировать эти предположения, необходимо ввести сопряженную функцию к гамильтониану H по правилу

$$H^*(t, x, Y, \xi) \triangleq \sup_{p \in \mathbb{R}^d} [\langle p, \xi \rangle - H(t, x, Y, p)].$$

Функция H^* определена на $[0, T] \times \mathbb{R}^d \times L^q(\Omega; \mathbb{R}^d) \times \mathbb{R}^d$. Предполагается, что функция H^* удовлетворяет свойству коэрцитивности, т.е. величина

$$\lim_{\|\xi\| \rightarrow \infty} \frac{|H^*(t, x, Y, \xi)|}{\|\xi\|} \rightarrow \infty$$

равномерно по t и x , равномерно ограничена снизу величиной $c_0(1 + \|Y\|^q)$ и ограничена, если вместо ξ подставлена некоторая функция $\xi_0(Y)$, имеет первые и вторые производные, также ограниченные снизу величиной $(c_1 L + c_2)\|Y\|^q$. Сам гамильтониан H непрерывен, локально липшицев по фазовой и сопряженной переменным, а также по функции Y . При этом производная $\partial H / \partial x$ непрерывна и глобально липшицева по всем переменным, кроме времени. От терминальной части функции платы требуется непрерывность, липшицевость по x и ограниченность. Также m_0 предполагается абсолютно непрерывной относительно меры Лебега.

Доказательство теоремы существования основано на принципе неподвижной точки Шаудера, примененном к системе принципа максимума Понтрягина для задачи оптимизации

$$\inf_{x(\cdot)} \left[\int_0^T H^*(t, x(t), X(t, \cdot), \dot{x}(t)) dt + \sigma(x(T), X(T, \cdot)) \right], \quad x(0) = X_0(\omega)$$

с дополнительным условием: кривые $X(\cdot, \omega)$ являются экстремальями.

Отметим, что единственность решения в рамках вероятностного подхода доказана при дополнительных предположениях. Первый способ доказательства предполагает, что $H(t, x, Y, p) = h_0(x, p) + F(x, Y)$, где h_0 выпукло по p , а F и σ удовлетворяют условиям монотонности:

- для всех $Y', Y'' \in L^q(\Omega; \mathbb{R}^d)$

$$\begin{aligned} & \int_{\Omega} [F(Y'(\omega), Y'(\cdot)) - F(Y'(\omega), Y''(\cdot)) + F(Y''(\omega), Y''(\cdot)) - F(Y''(\omega), Y'(\cdot))] \varrho(d\omega) \cdot \\ & \cdot \int_{\Omega} [\sigma(Y'(\omega), Y'(\cdot)) - \sigma(Y'(\omega), Y''(\cdot)) + \sigma(Y''(\omega), Y''(\cdot)) - \sigma(Y''(\omega), Y'(\cdot))] \varrho(d\omega) \leq 0, \end{aligned}$$

- и при $Y' \neq Y''$

$$\int_{\Omega} [F(Y'(\omega), Y'(\cdot)) - F(Y'(\omega), Y''(\cdot)) + F(Y''(\omega), Y''(\cdot)) - F(Y''(\omega), Y'(\cdot))] \varrho(d\omega) \neq 0.$$

В рамках второго подхода условие накладывается на H^* . А именно, единственность имеет место, когда

- для всех $Y', Y'' \in L^q(\Omega; \mathbb{R}^d)$

$$\begin{aligned} & \int_{\Omega} [H^*(t, Y'(\omega), Y'(\cdot), Z'(\omega)) - H^*(t, Y''(\omega), Y'(\cdot), Z''(\omega)) + \\ & + H^*(t, Y''(\omega), Y''(\cdot), Z''(\omega)) - H^*(t, Y'(\omega), Y''(\cdot), Z''(\omega))] \varrho(d\omega) \cdot \\ & \cdot \int_{\Omega} [\sigma(Y'(\omega), Y'(\cdot)) - \sigma(Y'(\omega), Z(\cdot)) + \sigma(Z(\omega), Z(\cdot)) - \sigma(Z(\omega), Y'(\cdot))] \varrho(d\omega) \leq 0, \end{aligned}$$

- и при $Y' \neq Y''$

$$\int_{\Omega} [H^*(t, Y'(\omega), Y'(\cdot), Z'(\omega)) - H^*(t, Y''(\omega), Y'(\cdot), Z''(\omega)) + H^*(t, Y''(\omega), Y''(\cdot), Z''(\omega)) - H^*(t, Y'(\omega), Y''(\cdot), Z''(\omega))] \varrho(d\omega) \neq 0.$$

Отметим, что решение в рамках вероятностного подхода определяется выбором множества игроков Ω , вероятностью ϱ и функцией, определяющей начальное положение игрока $X_0 : \Omega \rightarrow \mathbb{R}^d$. Вообще говоря, доказанная в рамках этого подхода единственность – это единственность при заданном выборе вероятностного пространства.

4.3 Минимаксный подход

Минимаксный подход [4] очень близок к вероятностному. В то же время, он позволяет рассматривать случай сублинейного гамильтониана, при этом требует минимальной гладкости от гамильтониана (не предполагается даже существование $\partial H/\partial p$). Главное отличие от вероятностного подхода состоит в том, что мы фиксируем множество, на котором задаем вероятность, а саму вероятность делаем частью решения.

Прежде всего, напомним [56], что в случае выпуклого по сопряженной переменной гамильтониана можно ввести эквивалентное определение минимаксного/вязкостного решения уравнения Гамильтона–Якоби. В этом случае функция ζ – минимаксное/вязкостное решение уравнения

$$\frac{\partial \zeta}{\partial t} + h(t, x, \nabla \zeta) = 0$$

тогда и только тогда, когда график ζ слабо инвариантен относительно дифференциального включения

$$(\dot{x}(t), \dot{z}(t)) \in \{(\xi, \eta) : \xi \in \text{dom } h^*(t, x, \cdot), \eta \geq h^*(t, x, \xi)\}. \quad (23)$$

Здесь

$$h^*(t, x, \xi) \triangleq \sup_{p \in \mathbb{R}^d} [\langle \xi, p \rangle - h(t, x, p)],$$

$$\text{dom } h^*(t, x, \cdot) = \{\xi : h^*(t, x, \xi) < \infty\}.$$

При достаточно широких предположениях решения дифференциального включения (23) продолжимы.

В рамках минимаксного подхода решение системы (8), (9) определяется следующим образом. Прежде всего, напомним, что носитель вероятности χ ($\text{supp}(\chi)$), определенной на некотором банаховом пространстве A , есть наименьшее замкнутое множество B такое, что для всех измеримых $C \subset A \setminus B$ выполнено $\chi(C) = 0$.

Далее, если V – некоторая функция, определенная на $[0, T] \times \mathbb{R}^d$, со значениями в \mathbb{R} , μ – поток вероятностей (т.е. $\mu[t]$ – вероятность на \mathbb{R}^d), то обозначим через $\mathcal{S}[V, \mu]$ множество кривых $(x(\cdot), z(\cdot)) : [0, T] \rightarrow \mathbb{R}^{d+1}$, удовлетворяющих дифференциальному включению

$$(\dot{x}(t), \dot{z}(t)) \in \{(\xi, \eta) : \xi \in \text{dom } H^*(t, x, \mu[t], \cdot), \eta \geq h^*(t, x, \xi)\},$$

таких, что для всех $t, s \in [0, T]$

$$V(x(t)) - z(t) = V(x(s)) - z(s).$$

Здесь

$$H^*(t, x, m, \xi) = \sup_{p \in \mathbb{R}^d} [\langle \xi, p \rangle - H(t, x, m, p)],$$

$$\text{dom } H^*(t, x, m, \cdot) = \{\xi : H^*(t, x, m, \xi) < \infty\}.$$

Определение 1. Будем говорить, что пара (V, μ) , где $V : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$, μ – поток вероятностей, является решением системы (8), (9), если

- V – минимаксное/вязкостное решение уравнения (8);

- существует вероятность χ , определенная на $C([0, T]; \mathbb{R}^{d+1})$, такая, что ее носитель лежит в множестве $\mathcal{S}[V, \mu]$, и для всех измеримых множеств $A \subset \mathbb{R}$

$$\mu[t](A) = \chi\{(x(\cdot), z(\cdot)) : x(t) \in A\}.$$

Теорема существования решения в рамках минимаксного подхода может быть получена при следующих предположениях:

- H представим в виде

$$H(t, x, m, p) = \max_{u \in U} [\langle p, f(t, x, m, u) \rangle + g(t, x, m, u)];$$

- функции f , g и σ непрерывны (при этом по m предполагается топология слабой сходимости);
- функция f локально липшицева по x ;
- начальное распределение m_0 имеет первый абсолютный момент, т.е.

$$\int_{\mathbb{R}^d} \|y\| m_0(dy) < \infty.$$

Доказательство теоремы существования такого решения основано на доказательстве существования равновесия по Нэшу в игре бесконечного числа лиц и в этой части смыкается с вероятностным подходом. Предполагается, что задано множество игроков Ω и неатомическая мера ϱ на Ω . Также предполагается выбранной функция $X_0 : \Omega \rightarrow \mathbb{R}^d$, физический смысл которой состоит в том, что она назначает каждому игроку его положение в начальный момент. Естественным условием является условие согласованности вероятности ϱ и функции X_0 с вероятностью m_0 : для всех измеримых множеств $A \subset \mathbb{R}^d$

$$m_0(A) = \varrho(X_0^{-1}(A)). \quad (24)$$

Отметим, что всегда можно подобрать множество Ω , вероятность ϱ на нем и функцию X_0 так, чтобы условие (24) выполнялось. Далее рассматривается игра бесконечного числа лиц, в которой положение каждого игрока $X(t, \omega)$ задается системой

$$\frac{d}{dt} X(t, \omega) = f(t, X(t, \omega), \mu[t], u(t)), \quad u(t) \in P, \quad X(0, \omega) = X_0(\omega),$$

а цель каждого игрока состоит в максимизации величины

$$\int_0^T g(t, X(t, \omega), \mu[t], u(t)) dt + \sigma(X(T, \omega), \mu[T]).$$

Здесь μ – поток вероятностей, связанный с динамикой игроков правилом:

$$\mu[t](A) = \varrho\{\omega : X(t, \omega) \in A\}.$$

В данной игре существует равновесие по Нэшу в классе обобщенных программных управлений (меруправлений). Полученное равновесие по Нэшу определяет вероятность χ и соответствующую функцию цены для пробного игрока, начавшего движение в некоторой позиции (t_*, x_*) .

Вопрос о единственности в рамках минимаксного подхода пока не рассматривался. Однако было построено приближенное равновесие по Нэшу в игре конечного числа игроков в классе рандомизированных программных стратегий.

5 Игры среднего поля с локальной связью

Для игр среднего поля с локальной связью (т.е. для системы (13), (14)) можно построить вязкостный подход, это сделано, в частности в [51]. Однако этот подход имеет те же ограничения, что и в случае нелокальной связи. В этом случае обобщенные решения строятся на основе концепции слабых решений [13] (см. также [15, 29]). Она развита в предположении, что

$$H(t, x, \theta, p) = h_1(x, \theta) - h_2(x, p),$$

а терминальная функция σ зависит исключительно от x . Также для простоты рассматривается случай, когда $x \in \mathbb{T}^d$, здесь \mathbb{T}^d обозначает d -мерный тор.

В этом случае система (13), (14) имеет вид

$$\frac{\partial V}{\partial t} + h_1(t, \theta) - h_2(x, \nabla V) = 0, \quad V(T, x) = \sigma(x), \quad (25)$$

$$\frac{\partial \theta}{\partial t} - \operatorname{div}(h_{2,p}(x, \nabla V)\theta) = 0, \quad \theta(0, x) = \theta_0(x). \quad (26)$$

Здесь и далее

$$h_{2,p}(x, p) = \frac{\partial h_2(x, p)}{\partial p}.$$

Кроме этого, предполагается, что:

1. h_1 непрерывна, возрастает по θ (напомним, что $\theta \in \mathbb{R}$), также выполнено условие роста по θ : существует показатель $q > 1$ такой, что

$$\frac{1}{C_1}|\theta|^{q-1} - C_1 \leq h_1(x, \theta) \leq C_1|\theta|^{q-1} + C_1;$$

2. h_1 непрерывна, а также выпукла и дифференцируема по p ; $\partial h_1/\partial p$ непрерывна; сама же функция h_1 имеет суперлинейный рост по p , т.е. существуют показатель $r > \max\{d(q-1), 1\}$ и константа C_2 такие, что

$$\frac{1}{rC_2}|\theta|^r - C_2 \leq h_1(x, \theta) \leq \frac{C_2}{r}|\theta|^r + C_2;$$

3. σ непрерывно дифференцируема, θ_0 непрерывна, неотрицательна и

$$\int_{\mathbb{T}^d} \theta_0(x) dx = 1.$$

Вводится функция $F(x, \theta)$ по правилу

$$F(x, \theta) = \begin{cases} \int_0^\theta h_1(x, \tau) d\tau, & \theta \geq 0 \\ +\infty, & \theta < 0. \end{cases}$$

В дальнейшем для функции ψ двух переменных через ψ^* будем обозначать сопряженную по второй переменной, т.е. F^* – это функция, сопряженная к F относительно переменной θ .

Понятие слабого решения основано на двух (дуальных) задачах оптимизации. Прежде всего, введем множество

$$\mathcal{K}_0 = \{\zeta \in C_1([0, T] \times \mathbb{T}^d) : \zeta(T, x) = \sigma(x)\}.$$

Пусть

$$\mathcal{A}_0(\zeta) = \int_0^T \int_{\mathbb{T}^d} F^*(x, -\zeta_t(t, x) + h_2(t, \nabla \zeta(t, x))) dx dt - \int_{\mathbb{T}^d} \zeta(0, x) \theta_0(x) dx.$$

Первая оптимизационная задача:

$$\inf_{\zeta \in \mathcal{K}_0} \mathcal{A}_0(\zeta). \quad (27)$$

С исходным уравнением (25) данная задача связана следующим образом: обозначим

$$\alpha(t, x) = -\zeta_t(t, x) + h_2(t, \nabla \zeta(t, x)). \quad (28)$$

Далее рассматривается задача минимизации функционала

$$\int_0^T \int_{\mathbb{T}^d} F^*(x, \alpha(t, x)) dx dt - \int_{\mathbb{T}^d} \zeta(0, x) \theta_0(x) dx.$$

Вторая оптимизационная задача связана с уравнением (26). Рассмотрим множество \mathcal{K}_1 , состоящее из пар $(\theta, w) \in L^1([0, T] \times \mathbb{T}^d) \times L^1([0, T] \times \mathbb{T}^d, \mathbb{R}^d)$ таких, что $\theta \geq 0$,

$$\int_{\mathbb{T}^d} \theta(t, x) dx = 1$$

для почти всех $t \in [0, T]$ и пара (θ, w) удовлетворяет в смысле распределений уравнению неразрывности

$$\zeta_t + \operatorname{div}(w) = 0, \quad \theta(0, x) = \theta_0(x).$$

Рассматривается функционал

$$\mathcal{B}(\theta, w) = \int_0^T \int_{\mathbb{T}^d} \left[\theta(t, x) h_2^* \left(t, -\frac{w(t, x)}{\theta(t, x)} \right) + F(x, \theta(t, x)) \right] dx dt + \int_{\mathbb{T}^d} \sigma(x) \theta(T, x) dx.$$

Отметим, что величина

$$\frac{w(t, x)}{\theta(t, x)}$$

может рассматриваться как скорость потока в момент времени t в позиции x . Вторая оптимизационная задача состоит в минимизации величины $\mathcal{B}(\theta, w)$ при $(\theta, w) \in \mathcal{K}_1$.

Имеет место следующее свойство (дуальность оптимизационных задач):

$$\inf_{\zeta \in \mathcal{K}_0} \mathcal{A}_0(\zeta) = - \min_{(\theta, w) \in \mathcal{K}_1} \mathcal{B}(\theta, w). \quad (29)$$

Доказательство этого факта основано на теореме Фенхеля–Рокафеллара [21]. Кроме того, минимум во второй оптимизационной задаче достигается на единственном элементе таком, что $\theta \in L^q([0, T] \times \mathbb{T}^d)$.

Для того чтобы найти минимизирующий элемент в первой оптимизационной задаче, проведем ее релаксацию. Для этого рассмотрим множество \mathcal{K} , состоящее из пар $(\zeta, \alpha) \in BV([0, T] \times \mathbb{T}^d) \times L^1([0, T] \times \mathbb{T}^d)$ таких, что

$$\alpha(t, x) \geq -\zeta_t(t, x) + h_2(t, \nabla \zeta(t, x)).$$

Последнее неравенство понимается в смысле распределений.

Отметим, что если $\zeta \in \mathcal{K}_0$, α определена по правилу (28), то пара $(\zeta, \alpha) \in \mathcal{K}$. Таким образом, \mathcal{K} – расширение \mathcal{K}_0 . Расширение первой оптимизационной задачи основано на функционале

$$\mathcal{A}(\zeta, \alpha) = \int_0^T \int_{\mathbb{T}^d} F^*(x, \alpha(t, x)) dx dt - \int_{\mathbb{T}^d} \zeta(0, x) \theta_0(x) dx.$$

Ставится задача поиска

$$\min_{(\zeta, \alpha) \in \mathcal{K}} \mathcal{A}(\zeta, \alpha).$$

Имеет место свойство:

$$\inf_{(\zeta, \alpha) \in \mathcal{K}} \mathcal{A}(\zeta, \alpha) = \inf_{\zeta \in \mathcal{K}_0} \mathcal{A}_0(\zeta) = - \min_{(\theta, w) \in \mathcal{K}} \mathcal{B}(\theta, w).$$

Принципиальная польза введенного расширения состоит в том, что минимум в левой части достигается. Более того, если (ζ, α) – минимизирующий элемент, то

- ζ удовлетворяет в вязкостном смысле неравенству

$$\zeta_t - h_1(t, \nabla \zeta) \leq 0,$$

т.е. для всех $(a, b) \in D^- \zeta(t, x)$ выполнено неравенство

$$a - h_1(t, b) \leq 0;$$

- $\alpha = \max\{-\zeta_t^{ac} + h_2(t, \nabla \zeta), 0\}$.

Слабым решением системы (25), (26) являются первые элементы минимизирующих пар. А именно, если $(\zeta^*, \hat{\alpha})$ – решение задачи

$$\min_{(\zeta, \alpha) \in \mathcal{K}} \mathcal{A},$$

а пара (θ^*, w^*) минимизирует функционал $\mathcal{B}(\theta, w)$, то пара $(\hat{\zeta}, \theta^*)$ и будет слабым решением системы (25), (26).

Дадим точное определение слабого решения, суммировав то, что мы знаем о решениях обеих задач оптимизации. Прежде всего, напомним, что если ϖ является зарядом на $[0, T] \times \mathbb{T}^d$, то ϖ может быть представлена в виде суммы $\varpi = \varpi^{ac} + \varpi^s$, где ϖ^{ac} – часть меры ϖ , абсолютно непрерывная относительно меры Лебега, а ϖ^s – сингулярная часть. Если $\phi : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$ является функцией ограниченной вариации, то через ϕ_t^{ac} и ϕ_t^s будем обозначать абсолютно непрерывную и сингулярную относительно меры Лебега части заряда $\partial\phi/\partial t$.

Определение 2. Пара $(\zeta, \theta) \in C([0, T] \times \mathbb{T}^d) \times L^1([0, T] \times \mathbb{T}^d)$ является слабым решением системы (25), (26), если

- ζ имеет ограниченную вариацию, $\nabla\zeta \in L^r([0, T] \times \mathbb{T}^d)$;
- $\theta(\cdot, \cdot)h_{2,p}(\cdot, \nabla\zeta(\cdot, \cdot)) \in L^1([0, T] \times \mathbb{T}^d)$;
- $(\zeta_t^{ac}(\cdot, \cdot) - \langle \nabla\zeta(\cdot, \cdot), h_2(\cdot, \nabla\zeta(\cdot, \cdot)) \rangle)\theta \in L^1([0, T] \times \mathbb{T}^d)$;
- почти всюду на области $\theta > 0$ выполнено равенство

$$\zeta_t^{ac} + h_1(t, \theta) - h_2(x, \nabla\zeta) = 0,$$

а неравенство

$$\zeta_t^{ac} + h_1(t, \theta) - h_2(x, \nabla\zeta) \geq 0$$

выполнено в смысле распределений; также $\zeta(T, x) = \sigma(x)$;

- уравнение

$$\frac{\partial\theta}{\partial t} - \operatorname{div}(h_{2,p}(x, \nabla V)\theta) = 0, \quad \theta(0, x) = \theta_0(x)$$

выполнено в смысле распределений;

-

$$\int_0^T \int_{\mathbb{T}^d} (\zeta_t^{ac}(t, x) - \langle \nabla\zeta(t, x), h_2(t, \nabla\zeta(t, x)) \rangle)\theta(t, x) dx dt = \int_{\mathbb{T}^d} (\theta(T, x)\sigma(x) - \theta_0(x)\zeta(0, x)) dx.$$

Отметим, что данное определение согласуется с классическим решением, т.е. классическое решение – слабое, а гладкое слабое решение является классическим.

Также единственность минимизирующего элемента во второй задаче влечет единственность слабого решения.

Отметим, что по слабому решению может быть построено приближенное решение в игровой вариационной задаче N лиц с показателем качества i -го игрока

$$J_i^N(\gamma^1, \dots, \gamma^N) = \int_0^T \left(L(\gamma_i(t), \dot{\gamma}_i(t)) + h_1^{\delta', \delta''} \left(\gamma_i(t), \frac{1}{N-1} \sum_{j=1, N, j \neq i} \delta_{\gamma_j(t)} \right) \right) dt + \sigma(\gamma(T)).$$

Здесь $L(x, \xi)$ – функция, сопряженная к функции $p \mapsto h_2(t, -p)$, $h_1^{\delta', \delta''}$ – результат двойного сглаживания функции h^1 (вначале сглаживаем второй аргумент путем свертки с некоторой гладкой функцией при параметре, равном δ' , потом саму функцию свертываем с тем же ядром при параметре δ'').

6 Одна модель энергетических сетей

Данная модель была предложена и исследовалась в работе [8]. Рассматривается сеть большого (в пределе бесконечного) количества одинаковых устройств. Каждое устройство характеризуется своей температурой $x(t)$. Предполагается, что изменение температуры подчиняется уравнению

$$\dot{x} = \begin{cases} -\alpha(x - x_{on}), & \text{устройство выключено,} \\ -\alpha(x - x_{off}), & \text{устройство включено.} \end{cases}$$

Если считать, что устройство управляется включением/выключением, то динамика температуры может быть переписана в виде

$$\dot{x} = -\alpha x + \beta u + \gamma,$$

где α, β и γ – некоторые константы, $\alpha > 0$, $u \in [0, 1]$ – управляющее воздействие.

Если m – некоторая вероятность на \mathbb{R} , то через $E[m]$ обозначим среднюю по вероятности m :

$$E[m] \triangleq \int_{\mathbb{R}} xm(dx).$$

Далее обозначим через $E_+[m] \triangleq \max\{E[m] - E_*, 0\}$, $E_-[m] = \max\{-(E[m] - E_*), 0\}$. Здесь E_* – номинальная температура сети.

Пусть $\mu[t]$ – распределение устройств по температуре в момент времени t . Интегральная часть функции платы каждого игрока определяется функцией

$$\begin{aligned} -g &= u(t)(r + qx^2(t) + hE + [\mu[t]]) + (1 - u(t))(qx^2(t) + kE_-[\mu[t]]) = \\ &= ru(t) + qx^2(t) + hE_+[\mu[t]]u(t) + kE_-[\mu[t]](1 - u(t)). \end{aligned}$$

Предполагается, что r, q, h и k – положительные константы. Смысл этой функции платы следующий: каждый игрок стремится к минимизации времени включения устройства, температуры устройства, также каждый игрок наказывается за включение устройства, когда температура сети выше номинала, и за невключение устройства, если температура ниже номинала.

Как видно, мы имеем ситуацию, когда поведение каждого игрока определяется средним состоянием системы. При этом гамильтониан равен

$$H(t, x, m, u) = \max_{u \in [0, 1]} \{-\alpha x + \beta u + \gamma + ru + qx^2 + hE_+[m]u + kE_-[m](1 - u)\}.$$

Данная задача может быть исследована в рамках описанных выше подходов для игр среднего поля с нелокальной связью или с использованием специальных методов, предложенных в [8].

Благодарности

Работа частично поддержана в рамках Программы Президиума РАН “Математические задачи современной теории управления” и гранта РФФИ 15-01-07909.

Список литературы

- [1] Y. Achdou. Finite difference methods for mean field games. *In Hamilton-Jacobi Equations: Approximations, Numerical Analysis and Applications*:1-47. Springer, 2013.
- [2] Y. Achdou, F. Camilli, I. Capuzzo-Dolcetta. Mean field games: numerical methods for the planning problem. *SIAM J. Control Optim.*, 50(1):77-109, 2012.
- [3] Y. Achdou and I. Capuzzo-Dolcetta. Mean field games: numerical methods. *SIAM J. Numer. Anal.*, 48(3):1136-1162, 2010.
- [4] Yu. V. Averboukh. A minimax approach to mean field games. *Sbornik: Mathematics*, 206(7):893–920, 2015 (in Russian). = Ю. В. Авербух. Минимаксный подход к играм среднего поля. *Матем. сб.*, 206(7):3-32, 2015.

- [5] Yu. Averboukh. *Deterministic limit of mean field games associated with nonlinear markov processes*. Preprint, 2016.
- [6] Yu. V. Averboukh. Universal Nash equilibria for differential games with many persons *Tr. IMM UrO RAN*, 20(3):26-40, 2014 (in Russian). = Ю. В. Авербух. Универсальные равновесия по Нэшу в дифференциальных играх многих лиц. *Тр. ИММ УрО РАН*, 20(3):26-40, 2014.
- [7] Yu. Averboukh. Universal Nash Equilibrium Strategies for Differential Games. *Journal of Dynamical and Control Systems*. 21(3):329-350, July 2015.
- [8] F. Bagagiolo, D. Bauso. Mean-field games and dynamic demand management in power grids. *Dyn. Games Appl.*, 4(2):155-176, 2014.
- [9] M. Bardi, I. Capuzzo-Dolcetta. *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*. Birkhäuser Boston Inc., Boston, MA, 1997. With appendices by Maurizio Falcone and Pierpaolo Soravia.
- [10] A. Bensoussan, J. Frehse, and P. Yam. *Mean field games and mean field type control theory*. Springer Briefs in Mathematics. Springer, New York, 2013.
- [11] A. Bensoussan, J. Frehse, and S. Yam. The master equation in mean field theory. *Journal de Mathématiques Pures et Appliquées*, 103(6):1441-1474, 2015.
- [12] R. Buckdahn, P. Cardaliaguet, M. Quincampoix. Some Recent Aspects of Differential Game Theory. *Dyn. Games Appl.* 1:74–114, 2011
- [13] P. Cardaliaguet. *Weak solutions for first order mean-field games with local coupling*. Preprint, 2013.
- [14] P. Cardaliaguet, F. Delarue, J.-M. Lasry, P.-L. Lions. *The master equation and the convergence problem in mean field games*. arXiv:1509.02505, 2015
- [15] P. Cardaliaguet, P. J. Graber. Mean field games systems of first order. *ESAIM Control Optim. Calc. Var.*, 21(3):690-722, 2015.
- [16] R. Carmona, F. Delarue. Mean field forward-backward stochastic differential equations. *Electron. Commun. Probab.*, 18(68):1–15, 2013.
- [17] R. Carmona, F. Delarue. Probabilistic analysis of mean-field games. *SIAM J. Control Optim.*, 51(4):2705-2734, 2013.
- [18] R. Carmona, F. Delarue, D. Lacker. *Mean field games with common noise*. arXiv:1407.6181, 2013.
- [19] R. Carmona, D. Lacker. A probabilistic weak formulation of mean field games and applications. *Ann. Appl. Probab.*, 25(3):1189-1231, 2015.
- [20] R. Carmona, X. Zhu. *A Probabilistic Approach to Mean Field Games with Major and Minor Players*. arXiv:1409.7141, 2014
- [21] I. Ekeland, R. Témam. *Convex analysis and variational problems, english ed., vol. 28 of Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1999.
- [22] M. Fischer. *On the connection between symmetric N-player games and mean field games*. arXiv:1405.1345.
- [23] A. Friedman. *Differential games*. Wiley, New York, 1971
- [24] W. Gangbo, A. Świąch. Existence of a solution to an equation arising from the theory of Mean Field Games. *J. Differential Equations*, 259(11):6573-6643, 2015.
- [25] D. Gomes, J. Mohr, R. R. Souza. Continuous time finite state mean-field games. *Appl. Math. and Opt.*, 68(1):99-143, 2013.

- [26] D. Gomes, L. Nurbekyan, E. Pimentel. Economic models and mean-field games theory. *Publicações Matemáticas do IMPA. [IMPA Mathematical Publications]. Instituto Nacional de Matemática Pura e Aplicada (IMPA), Rio de Janeiro, 2015. 30o Colóquio Brasileiro de Matemática. [30th Brazilian Mathematics Colloquium].*
- [27] D. Gomes and J. Saúde. Mean field games models-a brief survey. *Dyn. Games Appl.*, 4(2):110-154, 2014.
- [28] D. Gomes and V. Voskanyan. *Extended deterministic mean-field games*. Preprint, 2013.
- [29] J. Graber. *Weak solutions for mean field games with congestion*. Preprint, 2015.
- [30] M. Huang. Large-population LQG games involving a major player: the Nash certainty equivalence principle. *SIAM J. Control Optim.* 48:3318-3353, 2010.
- [31] M. Huang, P. E. Caines, and R. P. Malhamé. Large-population cost-coupled LQG problems with nonuniform agents: individual-mass behavior and decentralized ε -Nash equilibria. *IEEE Trans. Automat. Control*, 52(9):1560-1571, 2007.
- [32] M. Huang, P. E. Caines, and R. P. Malhamé. The NCE (mean field) principle with locality dependent cost interactions. *IEEE Trans. Automat. Control*, 55(12):2799-2805, 2010.
- [33] M. Huang, P. E. Caines, and R. P. Malhamé. Social optima in mean field LQG control: centralized and decentralized strategies. *IEEE Trans. Automat. Control*, 57(7):1736-1751, 2012.
- [34] M. Huang, R. P. Malhamé, and P. E. Caines. Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Commun. Inf. Syst.*, 6(3):221- 251, 2006.
- [35] A.C. Kizilkale and R.P. Malhame. A class of collective target tracking problems in energy systems: Cooperative versus non-cooperative mean field control solutions. *53rd IEEE Conference on Decision and Control*:3493- 3498, 2014.
- [36] A.C. Kizilkale and R.P. Malhame. Collective target tracking mean field control for electric space heaters. *Control and Automation (MED), 2014 22nd Mediterranean Conference of*:829-834, 2014.
- [37] A.C. Kizilkale and R.P. Malhame. Collective target tracking mean field control for markovian jumpdriven models of electric water heating loads. *IFAC World Congress* 19:1867–1872, 2014.
- [38] A.F. Klejmenov. *Nonzero-sum positional differential games*. Nauka, Ural'skoe otdelenie, Ekaterinburg, 1993 (in Russian). = А.Ф. Клейменов. *Неантагонистические позиционные дифференциальные игры*. Екатеринбург: Наука, Уральское отделение, 1993.
- [39] V.N. Kolokoltsov, J.J. Li, W. Yang. *Mean field games and nonlinear Markov processes*. arXiv:1112.3744v2, 2011.
- [40] V. N. Kolokoltsov and O. Malafeyev. *Mean field game model of corruption*. arXiv:1507.03240, 2015.
- [41] V. N. Kolokoltsov and M. Troeva. On the mean field games with common noise and the McKean-Vlasov SPDEs. arXiv:1506.04594, 2015.
- [42] V. N. Kolokol'tsov, M. S. Troeva, Wei Yang. Mean field games based on stable processes. *МТИП*, 5:4 (2013), 33-65 (in Russian). = В. Н. Колокольцов, М. С. Троева, Вей Янг. Игры среднего поля, связанные с процессами устойчивого типа. *МТИП*, 5:4 (2013), 33-65
- [43] K. Kolokoltsov and W. Yang. *Sensitivity analysis for HJB equations with an application to a coupled backward-forward system*. arXiv:1303.6234, 2013.
- [44] V. N. Kolokoltsov and W. Yang. *Inspection games in a mean field setting*. arXiv:1507.08339, 2015.
- [45] N.N. Krasovskii. *Control of dynamical system*. Nauka. M., 1985 (in Russian). = Н.Н. Красовский. *Управление динамической системой*. М.: Наука, 1985.

- [46] D. Lacker. *A general characterization of the mean field limit for stochastic differential games.* arXiv:1408.2708, 2014.
- [47] A. Lachapelle, J. Salomon and G. Turinic. Computation of mean field equilibria in economics. *Mathematical Models and Methods in Applied Sciences*, 20(4):567-588, 2010.
- [48] A. Lachapelle, M-T. Wolfram. On a mean field game approach modeling congestion and aversion in pedestrian crowds. *Transportation Research Part B: Methodological*, 45(10):1572-1589, 2011.
- [49] J.-M. Lasry, P.-L. Lions. Jeux à champ moyen. I. Le cas stationnaire. *C. R. Math. Acad. Sci. Paris*, 343(9):619-625, 2006.
- [50] J.-M. Lasry, P.-L. Lions. Jeux à champ moyen. II. Horizon fini et contrôle optimal. *C. R. Math. Acad. Sci. Paris*, 343(10):679-684, 2006.
- [51] J.-M. Lasry, P.-L. Lions. Mean field games. *Jpn. J. Math.*, 2(1):229-260, 2007.
- [52] P.-L. Lions. *College de france course on mean-field games.* 2007-2011.
- [53] M. Nourian, P.E. Caines. ε -Nash mean field game theory for nonlinear stochastic dynamical systems with major and minor agents. *SIAM J. Control Optim.* 51(4):3302-3331, 2013.
- [54] L.A.Petrosyan, N.A. Zenkevich, E.V. Shevkoplyas. *Game theory.* Izd-vo «BHV-Peterburg». SPb. 2012 (in Russian). = Л.А.Петросян, Н.А. Зенкевич, Е.В. Шевкопляс. *Теория Игр.* СПб: БХВ-Петербург, 2012.
- [55] N.N. Krasovskii, A.I. Subbotin. *Game-Theoretical Control Problems.* Springer-Verlag. New York-Berlin. 1988.
- [56] A.I. Subbotin. *Generalized Solutions of First-Order PDEs. The Dynamical Optimization Perspective.* Systems & Control: Foundations & Applications, Boston: Birkhäuser, 1995.
- [57] N.N. Subbotina. *Universal optimal strategies in positional differential games.* *Differential Equations* 19:1377-1382, 1983.

First order mean field games. Main directions

Yurii Averboukh

Krasovskii Institute of Mathematics and Mechanics (Yekaterinburg, Russia),
Ural Federal University (Yekaterinburg, Russia)

Keywords: mean field games, differential games, N -person games, Nash equilibrium.

This survey attempts to describe the main directions of the first order mean field games theory. We present the general theoretical results concerning first-order MFG. Additionally, the model of power-grid is discussed which is studied within MFG point of view.