Труды Международной конференции
"Проблемы управления и приложения (техника, производство, экономика)"
Минск, 16 – 20 мая 2005, с. 1–15

Метод динамического программирования в задачах синтеза управлений при разнотипных и двойных ограничениях*

А. Н. Дарьин , А. Б. Куржанский †

Аннотация. Рассматриваются две задачи синтеза управлений при двойном ограничением на управление. Решение задач достигается при помощи методов динамического программирования.

Ключевые слова: динамическое программирование, синтез управлений, неопределённость, импульсное управление.

Введение

В задачах синтеза управлений нередко целесообразно рассматривать системы с нестандартными ограничениями на управления. Так, в системах с неопределенностью традиционно принято считать, что управление и помеха принадлежат однотипным классам, например, одновременно стеснены или геометрическими, или интегральными ограничениями. Однако на практике возникают ситуации, когда необходимо выбирать для управления и помехи различные классы ограничений. Имеет смысл также налагать на управление одновременно несколько ограничений различных типов, что отражает приложения с одновременным ограничением на запас топлива и возможность маневра (см. [1]).

В настоящей работе рассматриваются две задачи с двойным ограничением на управление, решение которых достигается при помощи методов динамического программирования. Вначале рассматривается задача синтеза гарантирующих управлений в условиях неопределённости для линейной системы, где управление стеснено двойным ограничением (а именно, геометрическим и интегральным), а на помеху наложено только геометрическое ограничение:

$$\begin{cases} \dot{x}(t) = A(t)x(t) + B(t)u(t) + v(t), \\ \dot{k}(t) = -\langle u(t), R(t)u(t)\rangle, \\ u(t) \in \mathcal{P}(t), \quad v(t) \in \mathcal{Q}(t), \quad k(t) \geqslant 0, \\ (x(t_1), k(t_1)) \in \mathcal{M}. \end{cases} \tag{1}$$

Для решения задачи (1) предлагается подход [2], опирающийся на методы динамического программирования в сочетании со схемами теории альтернированного интеграла Л. С. Понтрягина [3] и теории игровых задач динамики Н. Н. Красовского [4]. Это позволяет разработать конструктивные методы, направленные на решение задачи «до конца», то есть до эффективных численных алгоритмов (которые могут быть основаны, например, на эллипсоидальном исчислении [5]).

* Работа выполнена при финансовой поддержке РФФИ (грант 03-01-00663), программы «Университеты России — фундаментальные исследования» (грант УР.03.02.521) и программы «Государственная поддержка ведущих научных школ» (грант НШ-1889.2003.1).

† Московский государственный университет им. М. В. Ломоносова, факультет вычислительной математики и кибернетики, кафедра системного анализа. 119992 Москва, Ленинские горы. daryin@cs.msu.su, kurzhanski@mail.ru

Двойное ограничение на управление может возникать при аппроксимации систем с импульсным управлением. Рассмотрим следующую задачу (см. [6, 7]):

$$\begin{cases} J(U(\cdot)) = \operatorname*{Var}_{[t,t_1]} U(t) \to \min \\ x(t_0 - 0) = x_0, \quad x(t_1 + 0) = x_1, \quad \mathrm{d}x(t) = A(t)x(t)\,\mathrm{d}t + B(t)\,\mathrm{d}U(t). \end{cases} \tag{2}$$

Здесь управление $U(\cdot)$ выбирается из класса функций ограниченной вариации и может иметь скачки. Для обеспечения непрерывности управления введём дополнительное геометрическое ограничение $u(t) = dU(t)/dt \in \mathcal{B}_1$. Далее к новой задаче применим методы динамического программирования. Соответствующая функция цены $V_\mu(t_0, x_0)$ может быть представлена как решение задачи конечномерной оптимизации. При стремлении μ к бесконечности V_μ поточечно сходится к функции цены V задачи (2). При этом уравнение Гамильтона–Якоби–Беллмана для в пределе даёт уравнение для V в форме квазивариационного неравенства.

1. Задача с двойным ограничением

В этом разделе приводится решение задачи синтеза гарантирующих управлений для линейной системы в условиях неопределённости, когда управление одновременно стеснено ограничениями двух типов — геометрическим и интегральным [8, 9]. Таким двойным ограничением можно описать как конструктивные особенности управляемой системы, так и ограниченный запас топлива. Помеха стеснена лишь геометрическим ограничением. Отсутствие интегрального ограничения для помехи может объясняться, например, следующими причинами: либо запас топлива у помехи настолько велик, что нет смысла вводить дополнительное ограничение; либо управлению неизвестен реальный запас топлива помехи, и для получения гарантирующего управления следует считать этот запас бесконечным.

Управляемая система описывается линейным дифференциальным уравнением

$$\dot{x}(t) = A(t)x(t) + B(t)u + C(t)v, \tag{3}$$

рассматриваемым на закреплённом отрезке времени $t \in [t_0, t_1]$. Здесь $x(t) \in \mathbb{R}^n$ — вектор положения системы, $u \in \mathbb{R}^{n_u}$ — управление, $v \in \mathbb{R}^{n_v}$ — помеха; $A(t) \in \mathbb{R}^{n \times n}$, $B(t) \in \mathbb{R}^{n \times n_u}$, $C(t) \in \mathbb{R}^{n \times n_v}$ — заданные непрерывные матричные функции.

Двойное ограничение на управление задаётся соотношениями

$$u(t) \in \mathcal{P}(t) \quad \text{и} \quad \int_{t_0}^{t_1} \langle u, R(t)u \rangle\,\mathrm{d}t \leqslant \nu. \tag{4}$$

Первое включение — геометрическое, или «жёсткое» ограничение, описывает возможности управления по манёвру. Здесь $\mathcal{P}(t)$ — заданное многозначное отображение, непрерывное по Хаусдорфу, имеющее непустые выпуклые компактные значения. Второе неравенство представляет собой интегральное, или «мягкое» ограничение, указывающее запас топлива, доступный управлению (величина ν). $R(t)$ — заданная непрерывная матричная функция, принимающая положительно определённые значения. Мы потребуем $0 \in \mathcal{P}(t)$, чтобы заведомо существовали управления, удовлетворяющие двойному ограничению.

Далее, в ходе решения задачи, для применения метода динамического программирования потребуется определить, какие переменные составляют позицию системы. При этом пара (t, x) не может быть выбрана в качестве позиции, поскольку для неё не будет выполнен принцип оптимальности. Это связано с тем, что в начальный момент t_0 точно известен запас топлива ν, однако по значениям t и x невозможно однозначно восстановить запас топлива в момент времени t. Чтобы преодолеть эту трудность, мы добавим к основному уравнению (3) ещё одно уравнение, отслеживающее изменения запаса топлива:

$$\dot{k}(t) = -\langle u, R(t)u\rangle, \quad k(t_0) = \nu. \tag{5}$$

Мы будем считать, что в любой момент времени управлению доступно точное значение переменной $k(t)$ — либо через измерение уровня топлива, либо путём учёта расхода и вычисления текущего остатка.

Интегрируя уравнение (5), можно видеть, что интегральное ограничение в (4) эквивалентно неравенству $k(t_1) \geqslant 0$, или, учитывая положительную определённость матрицы $R(t)$, — фазовому ограничению

$$k(t) \geqslant 0, \quad t \in [t_0, t_1]. \tag{6}$$

Прежде чем перейти к описанию класса допустимых управлений и формулировке основной задачи, сделаем несколько упрощений в уравнениях системы. Пусть $X(t,\tau)$ — фундаментальная матрица однородной системы $\dot{x}(t) = A(t)x(t)$, то есть решение задачи Коши

$$\frac{\mathrm{d}X(t,\tau)}{\mathrm{d}t} = X, \quad X(t,t) = I.$$

Тогда после замен $x = X(t,t_1)\bar{x}$, $\bar{B}(t) = X(t_1,t)B(t)R^{-\frac{1}{2}}(t)$, $\bar{u} = R^{\frac{1}{2}}(t)u$, $\bar{\mathcal{P}}(t) = R^{\frac{1}{2}}(t)\mathcal{P}(t)$, $\bar{\mathcal{Q}}(t) = X(t_1,t)C(t)\mathcal{Q}(t)$ мы получим (опуская черту над новыми функциями и переменными) систему

$$\begin{cases} \dot{x}(t) = B(t)u + v, \\ \dot{k}(t) = -\|u\|^2 \end{cases} \tag{7}$$

с ограничениями

$$u \in \mathcal{P}(t), \quad v \in \mathcal{Q}(t), \quad k(t) \geqslant 0. \tag{8}$$

Далее без ограничения общности мы будем вместо (3)–(5) рассматривать систему (7) с ограничениями (8).

Определение 1. *Позиционной стратегией, или синтезом управлений, называется многозначная функция вида $\mathcal{U}(t,x,k)$: $[t_0,t_1] \times \mathbb{R}^n \times \mathbb{R} \to \operatorname{conv}\mathbb{R}^{n_u}$, удовлетворяющая следующим требованиям:*

1. *функция $\mathcal{U}(t,x,k)$ измерима по переменной t;*
2. *функция $\mathcal{U}(t,x,k)$ полунепрерывна сверху по совокупности переменных (x,k);*
3. *выполнено геометрическое ограничение $\mathcal{U}(t,x,k) \subseteq \mathcal{P}(t)$, $t \in [t_0,t_1]$;*
4. *$\mathcal{U}(t,x,k) = \{0\}$, если $k < 0$, что гарантирует выполнение интегрального ограничения (напомним, что мы предполагаем $0 \in \mathcal{P}(t)$).*

Класс позиционных стратегий обозначается $\mathcal{U}_{CL}$.

Управления из класса $\mathcal{U}_{\mathrm{CL}}$ обеспечивают существование и продолжаемость решений дифференциального включения [10]

$$\begin{pmatrix} \dot{x}(t) \\ \dot{k}(t) \end{pmatrix} \in \left\{ \begin{pmatrix} B(t)u \\ -\|u\|^2 \end{pmatrix} \;\middle|\; u \in \mathcal{U}(t,x,k) \right\} + \mathcal{Q}(t) \times \{0\}. \tag{9}$$

которое возникает после подстановки в правую часть (7) позиционной стратегии $\mathcal{U}(t,x,k)$ вместо u и множества возможных значений помехи $\mathcal{Q}(t)$ вместо v. Отметим, что несмотря на линейность исходной системы (7), система с обратной связью (9) нелинейна, поскольку нелинейной является функция $\mathcal{U}(t, x, k)$.

Многозначные стратегии используются здесь для того, чтобы формально обеспечить существование и продолжаемость решений просинтезированной системы, поскольку в общем случае может не существовать непрерывных позиционных стратегий, решающих задачу. При этом практическая реализация подобного управления может быть однозначной (но разрывной) функцией, а дифференциальное включение (9) при фиксированной помехе $v(\cdot)$ превратится тогда в дифференциальное уравнение с разрывной правой частью.

Дифференциальное включение (9) описывает все возможные траектории системы при заданном синтезе управлений. Взятие выпуклой оболочки носит чисто технический характер и не влияет на возможности управления. Как было отмечено в [10], требование выпуклости правой части дифференциального включения может быть заменено на условие ее непрерывности, однако поскольку оптимальные позиционные стратегии являются лишь полунепрерывными сверху, но не непрерывными, предпочтительнее взять овыпукление правой части, чем требовать непрерывности позиционных стратегий.

Задача 1. *Указать такое* множество разрешимости $\mathcal{W}[t] \in \mathbb{R}^{n+1}$ *и такой синтез управлений* $\mathcal{U}(t, x, k) \in \mathcal{U}_{CL}$, *что все траектории системы* (7), *выпущенные в момент времени* t *из множества разрешимости* $\mathcal{W}[t]$, *достигнут заданное* целевое множество $\mathcal{M} \in \mathbb{R}^{n+1}$ *в момент времени* t_1, *несмотря на действия помехи.*

Иными словами, для всех траекторий дифференциального включения (9), выпущенных из точки $(x(t), k(t)) \in \mathcal{W}[t]$, должно выполняться включение $(x(t_1), k(t_1)) \in \mathcal{M}$. Последнее можно также сформулировать следующим образом: *множество достижимости* дифференциального включения (9) должно содержаться в целевом множестве $\mathcal{M}$.

Множество разрешимости также называют *попятным множеством достижимости при неопределённости* [11], или *мостом Н. Н. Красовского* [12].

Прежде чем перейти к решению задачи, необходимо ещё определить, из какого класса выбирается целевое множество $\mathcal{M}$. Поскольку нами применяются методы выпуклого анализа, то естественно было бы потребовать, чтобы целевое множество было выпуклым компактом (тогда множество разрешимости будет обладать аналогичными свойствами). Однако это требование можно ослабить, для чего вводится следующая вспомогательная конструкция.

Пусть в пространстве $\mathbb{R}^{n+1}$ переменных (x, k) задано множество $\mathcal{N}$. Будем называть *сечениями* множества $\mathcal{N}$ значения следующего многозначного отображения:

$$\mathcal{N}(k) = \{\, x \in \mathbb{R}^n \mid (x, k) \in \mathcal{N} \}.$$

Очевидно, что само множество $\mathcal{N}$ однозначно восстанавливается по своим сечениям, поскольку является графиком отображения $\mathcal{N}(\cdot)$:

$$\mathcal{N} = \left\{ (x, k) \in \mathbb{R}^{n+1} \,\middle|\, x \in \mathcal{N}(k) \right\}.$$

Из выпуклости (компактности) множества $\mathcal{N}$ следует выпуклость (компактность) его сечений, при этом обратное неверно.

Мы будем обозначать сечения целевого множества $\mathcal{M}$ через $\mathcal{M}(k)$, и потребуем выполнения следующих условий:

1. сечения монотонно возрастают по вложению: $\mathcal{M}(k_1) \subseteq \mathcal{M}(k_2)$, если $k_1 \leqslant k_2$;

2. $\mathcal{M}(k) = \varnothing$, если $k < 0$;

3. многозначное отображение $\mathcal{M}(k)$ непрерывно по Хаусдорфу для тех k, где $\mathcal{M}(k) \neq \varnothing$;
4. сечения $\mathcal{M}(k)$ — выпуклые компактные множества.

Сечения множества разрешимости мы будем обозначать через $\mathcal{W}[k,t]$. Предположения, сделанные выше относительно сечений целевого множества, уже не гарантируют, что $\mathcal{W}[k,t]$ будет выпуклым, поэтому мы оформим это требование в виде дополнительного предположения:

5. сечения $\mathcal{W}[k,t]$ — выпуклые множества.

Таким образом, использование сечений позволяет ослабить требование выпуклости и компактности целевого множества и множества разрешимости до требования выпуклости и компактности их сечений.

Потребуем от управления, чтобы оно минимизировало расстояние до сечения целевого множества в конечный момент времени. Этой задаче соответствует *функция цены*

$$V(t,x,k) = \inf_{\mathcal{U}\in\mathcal{U}_{\mathrm{CL}}} \sup_{z(\cdot)\in\mathcal{Z}_{\mathcal{U}}(\cdot)} d(x(t_1), \mathcal{M}(k(t_1))),$$

где $z(t) = (x(t), k(t))$, а $\mathcal{Z}_{\mathcal{U}}(\cdot)$ — трубка решений дифференциального включения (9) при позиционном управлении $\mathcal{U}$.

Мы воспользовались расстоянием до сечения целевого множества $d(x, \mathcal{M}(k))$, а не расстояние до самого целевого множества $d((x,k), \mathcal{M})$. Это связано с тем, что при сделанных предположениях множество $\mathcal{M}$ может не быть выпуклым, и тогда для вычисления последнего расстояния будет невозможно применить методы выпуклого анализа. Кроме того, это позволяет не смешивать при вычислении расстояния переменные x и k, имеющие, вообще говоря, различные размерности.

Удобно считать, что функция цены зависит не только от начального момента времени t и начального положения системы (x,k), но также и от конечного момента времени t_1 и заданного в этот момент терминального функционала, в данном случае $\varphi(x,k) = d(x, \mathcal{M}(k))$. Введём развёрнутое обозначение, подчеркивающее данную зависимость:

$$V(t,x,k) = V(t,x,k;t_1,\varphi(\cdot,\cdot)) = V(t,x,k;t_1,d(\cdot,\mathcal{M}(\cdot))).$$

Функция цены удовлетворяет *принципу оптимальности*, выражаемому полугрупповым свойством для обобщённой динамической системы:

$$V(t,x,k;t_1,\varphi(\cdot,\cdot)) = V(t,x,k;\tau,V(\tau,\cdot,\cdot;t_1,\varphi(\cdot,\cdot))),$$

при $t \leqslant \tau \leqslant t_1$. Выполнение принципа оптимальности означает, что в тройке (t,x,k) содержится вся информация о текущем состоянии системы.

Функция цены является решением уравнения Гамильтона–Якоби–Беллмана–Айзекса (Г–Я–Б–А)

$$\frac{\partial V}{\partial t} + \min_{u\in\mathcal{P}(t)} \max_{v\in\mathcal{Q}(t)} \left\{ \left\langle \frac{\partial V}{\partial x}, B(t)u + v \right\rangle - \frac{\partial V}{\partial k}\|u\|^2 \right\} = 0, \tag{10}$$

при $t_0 \leqslant t \leqslant t_1$, $k > 0$, $x \in \mathbb{R}^n$, с краевым условием

$$\frac{\partial V}{\partial t} + \max_{v\in\mathcal{Q}(t)} \left\langle \frac{\partial V}{\partial x}, v \right\rangle \Bigg|_{k=0} = 0, \quad t_0 \leqslant t \leqslant t_1, \quad x \in \mathbb{R}^n, \tag{11}$$

и начальным условием

$$V(t_1,x,k) = d(x,\mathcal{M}(k)), \quad k \geqslant 0, \quad x \in \mathbb{R}^n.$$

Вообще говоря, функция цены не является всюду гладкой по переменной x. В тех точках, где производная функции цены по переменной x терпит разрыв, слагаемое $\langle \partial V/\partial x, B(t)u + v\rangle$ в уравнении (10) следует понимать как производную в направлении $B(t)u + v$ (производные по направлению существуют в силу выпуклости по переменной x). Отметим, что причина негладкости не только в выборе недифференцируемой функции расстояния $d(x, \mathcal{M}(k))$ в качестве краевого условия. Даже при замене этой функции на гладкую $d^2(x, \mathcal{M}(k))$ функция цены становится негладкой при $k = 0$, $t < t_1$.

Краевое условие (11) отвечает фазовому ограничению (6): после исчерпания запаса топлива управление может принимать лишь нулевое значение. Уравнение (11) получается из (10), если положить $u = 0$.

Если каким-либо образом найдена функция цены, то из неё можно сразу же получить решение задачи 1. В самом деле, множество разрешимости является множеством уровня функции цены:

$$\mathcal{W}[k, t] = \{x \mid V(t, x, k) \leqslant 0\},$$

а искомый синтез управлений — множество минимизирующих векторов в (10):

$$\mathcal{U}^*(t, x, k) = \operatorname*{Arg\,min}_{u\in\mathcal{P}(t)} \left\{ \left\langle \frac{\partial V}{\partial x}, B(t)u \right\rangle - \frac{\partial V}{\partial k}\|u\|^2 \right\}$$

Например, в случае скалярного управления с геометрическим ограничением $|u| \leqslant 1$ синтез принимает следующий вид:

$$\mathcal{U}^*(t, x, k) = \begin{cases} [-1, 1], & V(t, x, k) = 0; \\ u^*, & |u^*| \leqslant 1; \\ -1, & u^* < -1; \\ 1, & u^* > 1, \end{cases} \qquad u^*(t, x, k) = -\frac{1}{2\frac{\partial V}{\partial k}} B'(t) \frac{\partial V}{\partial x}.$$

В общем случае нахождение функции цены сопряжено с существенными вычислительными затратами. В то же время, в формуле для оптимального управления (1) участвует не сама функция цены, а лишь её производные по переменным x и k. Это означает, что для вычисления управления достаточно найти множества уровня функции цены, и в частности множество разрешимости. Опишем, каким образом можно реализовать эту идею.

Справедлива следующая оценка для функции цены:

$$V(t, x, k) \leqslant d(x, \mathcal{W}[k, t]). \tag{12}$$

Подставляя эту оценку вместо самой функции в уравнение Г–Я–Б–А, мы получаем дифференциальное неравенство

$$\min_{u\in\mathcal{P}(t)} \max_{v\in\mathcal{Q}(t)} \frac{\mathrm{d}d^2(x, \mathcal{W}[k, t])}{\mathrm{d}t} \leqslant 0. \tag{13}$$

Здесь мы возвели расстояние в квадрат, чтобы обеспечить достаточную гладкость и избежать использования производных по направлению. Использование оценки (12) можно интерпретировать и как замену множеств уровня функции цены $\{x \mid V(t, x, k) \leqslant \alpha\}$ на меньшие множества $\mathcal{W}[k, t] + \alpha\mathcal{B}_1$.

Определим стратегию, экстремальную к множеству разрешимости $\mathcal{W}[k, t]$, как множество минимизирующих векторов в (13):

$$\mathcal{U}_{\mathcal{W}}(t, x, k) = \operatorname*{Arg\,min}_{u\in\mathcal{P}(t)} \left\{ \left\langle \frac{\partial d^2(x, \mathcal{W}[k, t])}{\partial x}, B(t)u \right\rangle - \frac{\partial d^2(x, \mathcal{W}[k, t])}{\partial k}\|u\|^2 \right\}.$$

Последнее может быть представлено в следующем виде:

$$\mathcal{U}_{\mathcal{W}}(t,x,k) = \operatorname*{Arg\,min}_{u\in\mathcal{P}(t)} \left\{ \langle \ell_0, B(t)u\rangle - \frac{\partial\rho(\,\ell_0 \mid \mathcal{W}[k,t])}{\partial k}\|u\|^2 \right\},$$

где ℓ_0 — единственный максимизатор в задаче

$$\langle \ell, x\rangle - \rho(\,\ell \mid \mathcal{W}[k,t]) - \frac{1}{4}\|\ell\|^2 \to \max,$$

а через $\rho(\,\ell \mid \mathcal{W}[k,t])$ обозначена опорная функция [13] множества разрешимости.

В силу (13) для траекторий дифференциального включения (9) при управлении $\mathcal{U}_{\mathcal{W}}$ выполняется неравенство

$$d(x(t_1), \mathcal{M}(k(t_1))) = d(x(t_1), \mathcal{W}[k(t_1), t_1]) \leqslant d(x(t), \mathcal{W}[k(t), t]).$$

Если начальная точка $(x(t), k(t))$ принадлежит множеству разрешимости $\mathcal{W}[k(t), t]$, то конечная точка траектории $(x(t_1), k(t_1))$ лежит в целевом множестве. Следовательно, управление $\mathcal{U}_{\mathcal{W}}$ является решением задачи 1.

Вычислительные схемы для нахождения множества разрешимости могут быть основаны на следующем *эволюционном уравнении*:

$$\lim_{\sigma\downarrow 0} \sigma^{-1} h_+ \left(\mathcal{W}[k,t] + \sigma\mathcal{Q}(t), \bigcup_{0\leqslant\gamma\leqslant k} \left[\mathcal{W}[\gamma, t+\sigma] - (\sigma\mathcal{P}(t)) \cap \sqrt{\sigma(k-\gamma)}\mathcal{B}_1 \right] \right) = 0, \qquad (14)$$

являющемся аналогом обыкновенных дифференциальных уравнений для многозначных отображений. Разностная схема для эволюционного уравнения (14) может рассматриваться как операция многозначного интегрирования — аналог *альтернированного интеграла* Л. С. Понтрягина [3, 14].

Множество разрешимости — не единственное решение эволюционного уравнения. Среди его решений находятся и внутренние эллипсоидальные аппроксимации множества разрешимости. Использование эллипсоидального исчисления позволяет приблизить множество разрешимости многозначными функциями, зависящими от малого числа параметров, являющихся решениями обыкновенных дифференциальных уравнений, и тем самым получить эффективные численные алгоритмы решения задачи [5].

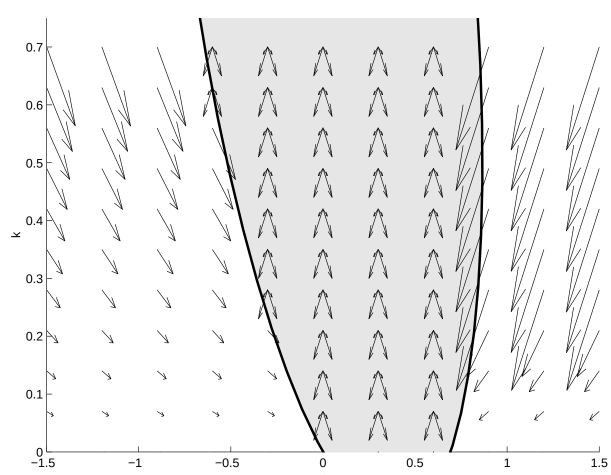

Рис. 1: Пример синтеза управлений

Пример синтеза управлений приведён на рис. 1. Позиционная стратегия построена для задачи со скалярной переменной x. Закрашенная область представляет собой множество разрешимости в начальный момент времени. Стрелками показаны направления векторного поля $(u, -\|u\|^2)$. Группы стрелок внутри множества разрешимости указывают, что управление может выбирать любое значение из $\mathcal{P}(t)$.

2. Задача импульсного управления

Рассмотрим задачу минимизации обобщённого функционала типа Майера–Больца на траекториях импульсной управляемой системы:

$$\begin{cases} J(u(\cdot)) = \operatorname*{Var}_{[t_0,t_1]} U(\cdot) + \varphi(x(t_1+0)) \to \inf, \\ \mathrm{d}x(t) = A(t)x(t)\,\mathrm{d}t + B(t)\,\mathrm{d}U(t), \quad t \in [t_0, t_1], \\ x(t_0 - 0) = x_0. \end{cases} \tag{15}$$

Здесь $x(t) \in \mathbb{R}^n$ — фазовый вектор, $U(\cdot) \in BV([t_0,t_1];\mathbb{R}^m)$ — обобщённое управление, $BV([t_0,t_1];\mathbb{R}^m)$ — пространство функций ограниченной вариации со значениями в $\mathbb{R}^m$. Предполагается, что матричные функции $A(t) \in \mathbb{R}^{n\times n}$, $B(t) \in \mathbb{R}^{n\times m}$ непрерывны. Конечный момент времени t_1 зафиксирован. Терминальное слагаемое $\varphi : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ является замкнутой выпуклой функцией, присутствие которой в выражении для $J(u(\cdot))$ позволит далее сформулировать принцип оптимальности.

Специальный выбор $\varphi(x) = \mathcal{I}(x \mid \{x_1\})$[1] приводит к известной задаче о переводе управляемой системы из точки x_0 в момент t_0 в точке x_1 в момент t_1 при минимальной вариации управления:

$$\begin{cases} \operatorname*{Var}_{[t_0,t_1]} U(\cdot) \to \inf, \\ \mathrm{d}x(t) = A(t)x(t)\,\mathrm{d}t + B(t)\,\mathrm{d}U(t), \quad t \in [t_0, t_1], \\ x(t_0 - 0) = x_0, \quad x(t_1 + 0) = x_1. \end{cases} \tag{16}$$

Первые программные решения в задаче импульсного управления были получены Н. Н. Красовским [15]. В работе [16] было показано, что в линейной импульсной задаче число скачков оптимального управления не превышает размерность фазового пространства.

Задачи подобного типа хорошо изучены (см. [6, 7, 16, 17, 18, 19, 20]) и могут быть решены применением методов функционального и выпуклого анализа. Решением при этом является программное управление. Напротив, в этой статье нас будет интересовать решение задачи методами динамического программирования, дающее управление в виде стратегии с обратной связью. Рассматриваются линейные системы, что позволяет сочетать классическую теорию распределений с теорией обобщённых (вязкостных) решений ([21, 22, 23] для соответствующих квазивариационных неравенств типа Гамильтона–Якоби–Беллмана (Г–Я–Б) (см. [24]). Предлагаемый подход позволяет также рассматривать задачи с высшими производными дельта-функций (см. [25]).

2.1. Идеальная схема

Будем называть *функцией цены* $V(t_0, x_0)$ задачи (15) оптимальное значение функционала $J(U(\cdot))$ при фиксированной начальной позиции (t_0, x_0). Иногда будет использоваться

[1] Через $\mathcal{I}(x \mid A)$ обозначена индикаторная функция множества A (равная нулю на множестве A и $+\infty$ вне A).

развёрнутое обозначение $V(t_0, x_0; t_1, \varphi(\cdot))$, чтобы подчеркнуть зависимость оптимального значения $V(t_0, x_0)$ от конечного момента времени t_1 и терминальной функции $\varphi(\cdot)$.

Обозначим через $W(t_0, x_0) = W(t_0, x_0; t_1, x_1)$ минимальную вариацию управления в задаче (16). Как уже было указано ранее, $W(t_0, x_0; t_1, x_1) = V(t_0, x_0; t_1, \mathcal{I}(\cdot \mid \{x_1\}))$.

Задачу (15) можно разбить на две части:

- найти оптимальную конечную точку траектории x_1;
- найти оптимальное управление $U(\cdot)$ в задаче (16) при условии $x(t_1 + 0) = x_1$.

В следующих утверждениях кратко сформулировано решение второй подзадачи в форме программного управления, подробно описанное в [6, 25]:

Утверждение 1. *Оптимальное значение в задаче* (16) *может быть представлено в виде*

$$W(t_0, x_0; t_1, x_1) = \sup_{p \in \mathbb{R}^n} \frac{\langle p, x_1 - X(t_1, t_0)x_0 \rangle}{\|B'(\cdot)X'(t_1, \cdot)p\|_{C[t_0,t_1]}}. \tag{17}$$

Здесь $X(t, \tau)$ — фундаментальная матрица, то есть решение задачи Коши для матричного дифференциального уравнения $\partial X/\partial t = A(t)X$, $X(\tau, \tau) = I$.

Утверждение 2. *Если $W(t_0, x_0) < \infty$, то существует оптимальное программное управление $U(\cdot)$ специального вида:*

$$\mathrm{d}U(t) = \sum_{i=1}^{n} h_i \delta(t - \tau_i)\, \mathrm{d}t, \tag{18}$$

$$t_0 \leqslant \tau_1 < \tau_2 < \ldots < \tau_n \leqslant t_1, \quad u_i \in \mathbb{R}^m,$$

$$\operatorname*{Var}_{[t_0,t_1]} U(\cdot) = \sum_{i=1}^{n} \|h_i\| = V(t_0, x_0).$$

Другими словами, если возможно перевести управляемую систему из позиции (t_0, x_0) в (t_1, x_1), то существует управление, на котором достигается оптимальное значение вариации $W(t_0, x_0)$. Более того, существует оптимальное управление, составленное из не более чем n скачков, где n — размерность вектора x.

Используя выражение (17), можно записать функцию цены задачи (15) в виде

$$V(t_0, x_0) = \min_{x_1 \in \mathbb{R}^n} \{\varphi(x_1) + W(t_0, x_0; t_1, x_1)\}. \tag{19}$$

Это позволяет сформулировать следующий результат.

Теорема 1. *Если $V(t_0, x_0) < \infty$, то существует такое программное управление $U(\cdot)$ вида* (18)*, для которого выполнено равенство $J(U(\cdot)) = V(t_0, x_0)$.*

Введём следующую полунорму на $\mathbb{R}^n$:

$$\|p\|_{[t_0,t_1]} = \left\|B'(t)X'(t_1, \cdot)p\right\|_{C[t_0,t_1]}$$

тогда после ряда преобразований возможно представить функцию цены в виде

$$V(t_0, x_0) = \sup_{p \in \mathbb{R}^n} \Big[\langle p, X(t_1, t_0)x_0 \rangle - \varphi^*(p) - \mathcal{I}\Big(p \;\Big|\; \mathcal{B}_{\|\cdot\|_{[t_0,t_1]}}\Big)\Big]. \tag{20}$$

Здесь $\mathcal{B}_{\|\cdot\|_{[t_0,t_1]}}$ — единичный шар в указанной полунорме, а $\varphi^*(p)$ — сопряжённая функция по Фенхелю к $\varphi(x)$ [13]. Мы получили следующее утверждение:

Теорема 2. *Функция цены $V(t_0, x_0)$ является выпуклой по переменной x, и сопряжённая к ней функция задаётся формулой*

$$V^*(t_0, p) = \varphi^*(X'(t_0, t_1)p) + \mathcal{I}\Big(X'(t_0, t_1)p \;\Big|\; \mathcal{B}_{\|\cdot\|_{[t_0,t_1]}} \Big). \tag{21}$$

Используя (21), можно установить следующий результат:

Теорема 3. *Функция цены $V(t, x; t_1, \varphi(\cdot))$ задачи* (15) *удовлетворяет* принципу оптимальности*, который может быть записан в форме полугруппового свойства: для каждого момент времени $\tau \in [t_0, t_1]$ верно*

$$V(t_0, x_0; t_1, \varphi(\cdot)) = V(t_0, x_0; \tau, V(\tau, \cdot; t_1, \varphi(\cdot))).$$

Заметим, что в отличие от задач с ограниченными управлениями, в общем случае здесь имеет место неравенство $V(t_1, x; t_1, \varphi(\cdot)) \leqslant \varphi(x)$, которое может быть строгим: из (21) следует, что

$$V^*(t_1, p) = \varphi^*(p) + \mathcal{I}(\, B(t_1)p \mid \mathcal{B}_1).$$

Например, если $\varphi(x) = \mathcal{I}(x \mid \{x_1\})$ и $B(t_1) = I$, то $V(t_1, x; t_1, \varphi(\cdot)) = \|x - x_1\| \leq \varphi(x)$.

Теорема 4. *Функция цены $V(t, x)$ является вязкостным решением [21] уравнения Гамильтона–Якоби–Беллмана:*

$$\min\{H_1(t, x, V_t, V_x),\; H_2(t, x, V_t, V_x)\} = 0, \tag{22}$$

$$V(t_1, x) = V(t_1, x; t_1, \varphi(\cdot)).$$

$$H_1(t, x, \xi_t, \xi_x) = \xi_t + \langle \xi_x, A(t)x \rangle, \qquad H_2(t, x, \xi_t, \xi_x) = \min_{u \in S_1} \langle \xi_x, B(t)u \rangle + 1.$$

В соответствии с (22) в любой позиции (t, x) управление одну из двух возможностей. Либо $H_1 = 0$, и тогда можно выбрать $\mathrm{d}U(t) = 0$; либо $H_1 > 0$, тогда обязательно $H_2 = 0$, и управление должно совершить скачок в направлении $-B'(t)V_x$. Величина скачка должна быть выбрана так, чтобы после скачка снова было выполнено равенство $H_1 = 0$.

Однако, подобное рассуждение не приводит к полноценному определению синтезирующей стратегии, поскольку всё ещё не ясно, что будет представлять собой управляемая система, в которую в качестве управления подставлен синтез. Один из подходов к преодолению этой трудности заключается в использовании расширенной пространственно-временной системы [20, 19, 27]:

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}s} = A(t(s))x(s) \cdot u^t(s) + B(t(s))u^x(s), \\[2mm] \dfrac{\mathrm{d}t}{\mathrm{d}s} = u^t(s). \end{cases} \tag{23}$$

Здесь переменная s отвечает за параметризацию траекторий переменных x и t, $s \in [0, S]$, при этом правый конец S не закреплён. На расширенное управление $u(s) = (u^x(s), u^t(s)) \in \mathbb{R}^m \times \mathbb{R}$ наложено жёсткое (геометрическое) ограничение $u(s) \in \mathcal{B}_1 \times [0, 1]$. Исходной задаче импульсного управления (15) соответствует следующая задача для системы (23):

$$\begin{cases} J(u(\cdot)) = \displaystyle\int_0^S \|u^x(s)\|\, \mathrm{d}s + \varphi(x(S)) \to \inf, \\[2mm] t(0) = t_0, \quad t(S) = t_1. \end{cases} \tag{24}$$

Известно [20], что любое импульсное управлении и соответствующая ему траектория исходной системы (15) могут быть представлены как регулярное управление и непрерывная траектория расширенной системы (23), и что множество траекторий системы (15) всюду плотно во множестве траекторий (23).

Функция цены задачи (24) является вязкостным решением следующего уравнения Гамильтона–Якоби–Беллмана:

$$\min_{\substack{u^t\in[0,1]\\ u^x\in\mathcal{B}_1}} H(t,x,V_t,V_x,u^t,u^x)=0, \tag{25}$$

$$H(t,x,\tau,\xi,u^t,u^x)=\big\{[\tau+\langle\xi,A(t)x\rangle]u^t+[\langle\xi,B(t)u^x\rangle+\|u^x\|]\big\}=0,$$

что совпадает с уравнением Г–Я–Б (22) в задаче импульсного управления.

Теперь, используя (25), возможно определить синтез управлений в задаче (24) как множество минимизаторов в уравнении (25):

$$\mathcal{U}^*(t,x)=\bigcup_{(\tau,\xi)\in\partial_C V}\big\{u\mid H(t,x,\tau,\xi,u^t,u^x)=0\big\}, \tag{26}$$

Здесь $\partial_C V$ — дифференциал Кларка [28] функции цены по обеим переменным (t,x).

Поскольку система (23) описывает все траектории в задаче (15), управление (26) можно рассматривать как синтез управлений для задачи (15).

После подстановки управления (26) система принимает вид дифференциального включения

$$\frac{\mathrm{d}}{\mathrm{d}s}\begin{pmatrix}x\\ t\end{pmatrix}\in\left\{\begin{pmatrix}A(t)x & B(t)\\ 1 & 0\end{pmatrix}u\;\middle|\;u\in\mathcal{U}^*(t,x)\right\}. \tag{27}$$

В силу того, что синтез $\mathcal{U}^*(t,x)$ является полунепрерывным сверху по переменным (t,x) многозначным отображением с непустыми выпуклыми компактными значениями (что следует из свойств дифференциала Кларка), решения системы (27) существует и являются продолжаемыми на область $(t,x)\in[t_0,t_1]\times\mathbb{R}^n$ (см. [10]). Любая оптимальная траектория в задаче (15) удовлетворяет включению (27). Другим словами, дифференциальное включение (27) порождает все возможные оптимальные траектории. Тем не менее, остаётся открытым вопрос о том все ли траектории (27), достигающие плоскости $t=t_1$, являются оптимальными.

2.2. Схема двойного ограничения

В предыдущем разделе рассматривалась идеальная схема, в которой решением задачи управления являлись обобщённые функции. Теперь мы представим «реалистичный» подход, в котором управлениями являются ограниченные функции, однако число, ограничивающее норму управления может быть сколь угодно большим и стремиться к бесконечности.

Наложим на управление в задаче (15) дополнительное геометрическое ограничение, $u(t)\in\mathcal{B}_\mu$, и рассмотрим соответствующую задачу управления:

$$\begin{cases} J(u(\cdot))=\displaystyle\int_{t_0}^{t_1}\|u(t)\|\,\mathrm{d}t+\varphi(x(t_1))\to\inf,\\ \quad\dot{x}(t)=A(t)x(t)+B(t)u(t),\quad t\in[t_0,t_1],\\ \quad x(t_0)=x_0,\quad \|u(t)\|\leqslant\mu. \end{cases} \tag{28}$$

Решение в задаче (28) существует в силу теоремы Вейерштрасса: множество допустимых управлений слабо компактно (оно ограничено, замкнуто и выпукло в пространстве $L_2([t_0,t_1];\mathbb{R}^m)$), а минимизируемый функционал $J(u(\cdot))$ слабо полунепрерывен снизу (так как он является выпуклым и полунепрерывным снизу в $L_2([t_0,t_1];\mathbb{R}^m)$).

Функция цены $V_\mu(t_0, x_0) = V_\mu(t_0, x_0; t_1, \varphi(\cdot))$ в этой задаче является вязкостным решением [21] уравнения Гамильтона–Якоби–Беллмана

$$\frac{\partial V_\mu}{\partial t} + \min_{u \in \mu \mathcal{B}_1} \left\{ \left\langle \frac{\partial V_\mu}{\partial x}, A(t)x(t) + B(t)u \right\rangle + \|u\| \right\} = 0 \tag{29}$$

с начальным условием $V_\mu(t_1, x) = \varphi(x)$. Следовательно, кроме некоторых вырожденных случаев оптимальное управление должно либо принимать значения из сферы S_μ, либо равняться нулю.

Функция цены задачи (28) может быть найдена как оптимальное значение в конечномерной задаче оптимизации:

$$V_\mu(t,x) = \sup_{p \in \mathbb{R}^n} \Big\{ \langle p, X(t_1, t_0) x_0 \rangle - \mu \int_t^{t_1} \big(\big\| B'(\tau) X'(t_1, \tau) p \big\| - 1 \big)_+ \,\mathrm{d}\tau - \varphi^*(p) \Big\}, \tag{30}$$

а сопряжённая к ней функция по переменной x задаётся формулой

$$V_\mu^*(t,p) = \varphi^*(p) + \mu \int_t^{t_1} \big(\big\| B'(\tau) X'(t_1, \tau) p \big\| - 1 \big)_+ \,\mathrm{d}\tau. \tag{31}$$

Здесь $a_+ = \max\{a, 0\}$.

Заметим, что при стремлении μ к бесконечности выражения (30), (31) превращаются соответственно в (20) и (21). Для случая $x \in \mathbb{R}^1$ можно получить следующую оценку скорости сходимости: для каждой позиции (t,x) существует такая константа $C > 0$, что выполняется

$$0 \leqslant V(t,x) - V_\mu(t,x) = C\mu^{-1}$$

Уравнение Г–Я–Б (22) также может быть формально получено из (29) в пределе при $\mu \to \infty$.

Оптимальный синтез управлений представляет собой множество минимизаторов в (29), и в точках дифференцируемости $V_\mu(t,x)$ может быть записан следующим образом:

$$\mathcal{U}_\mu^*(t,x) = \begin{cases} \{0\}, & \|\zeta\| < 1; \\ [0, -\mu\zeta], & \|\zeta\| = 1; \\ \left\{ -\mu \dfrac{\zeta}{\|\zeta\|} \right\}, & \|\zeta\| > 1, \end{cases} \tag{32}$$

где $\zeta = B'(t) \frac{\partial V_\mu}{\partial x}$. Стратегия (32) удовлетворяет условиям существования и продолжаемости траекторий просинтезированной системы, записанной в форме дифференциального включения [10]:

$$\dot{x}(t) \in A(t)x(t) + B(t)\mathcal{U}_\mu^*(t,x). \tag{33}$$

Пример 1. *Рассмотрим следующую задачу об успокоении затухающих колебаний маятника импульсным управлением:*

$$\begin{cases} \operatorname*{Var}_{[t_0,t_1]} U(\cdot) \to \inf, \\ \begin{cases} \mathrm{d}x_1(t) = x_2(t)\,\mathrm{d}t, \\ \mathrm{d}x_2(t) = -\omega_0^2 x_1(t)\,\mathrm{d}t - 2\alpha x_2(t) + \mathrm{d}U(t), \end{cases} & t_0 \leqslant t \leqslant t_1, \\ x_1(t_0 - 0) = x_1^0, \quad x_2(t_0 - 0) = x_2^0, \\ x_1(t_1 + 0) = 0, \quad x_2(t_1 + 0) = 0. \end{cases}$$

В качестве значений параметров выберем $\omega_0 = 5$, $\alpha = 3$, $t_0 = 0$, $t_1 = \pi/8$.

Приведём точное решение этой задачи. В отсутствие управления ($\mathrm{d}U(t) = 0$) движения системы представляют собой затухающие колебания с периодом $\omega = 4$ и коэффициентом затухания $\alpha = 3$:

$$x_1(t) = \mathrm{e}^{-3t}(C_1 \cos 4t + C_2 \sin 4t).$$

Вследствие затухания управлению следует ждать до тех пор, пока маятник не окажется в нижней точке ($x_1 = 0$) и затем одним импульсом остановить движение. Однако не из всех начальных позиций маятник достигнет нижнего положения до конечного момента t_1. Если в текущей позиции (t, x) выполняется неравенство

$$h_1 = x_2 + x_1(3 + 4\,\mathrm{tg}\,4t) > 0,$$

то оптимальное управление имеет скачок в этот момент с амплитудой h_1. Далее управление должно ждать того момента, когда $x_1 = 0$ (такой момент обязательно настанет) и совершать скачок с амплитудой $h_2 = -x_2$, приводящий систему прямо в начало координат. Оптимальные траектории, выпущенные в момент $t = 0$ показаны на рис. 2. Сплошными линиями указаны непрерывные части траекторий, тонким пунктиром — скачок в начальный момент, жирным пунктиром — скачок в конечный момент.

Соответствующий синтез управлений, полученный при двойном ограничении, определяемый формулой (32) (для $t = 0$) приведён на рис. 3. Отметим, что фазовое пространство распадается здесь на четыре области: три области R_0, $R_{-\mu}$, R_μ отвечают значениям управления 0, μ, $-\mu$, а внешняя область $R_\varnothing$ содержит те начальные позиции, из которых невозможно достичь начала координат за установленное время (то есть задача не имеет решения).

При стремлении μ к бесконечности участок R_0 заполняет те области, где импульсное управление не имело скачка, а участки $R_{-\mu}$ и R_μ заполняют те области, где импульсное управление имело скачок в соответствующую сторону (рис. 4).

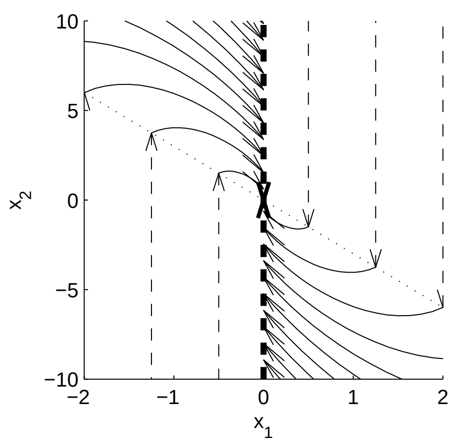

Рис. 2: Идеальные траектории системы

Список литературы

[1] *Красовский Н. Н.* К задаче об успокоении линейной системы при минимальной интенсивности управления // ПММ. 1965. Т. 29. № 2. С. 218–225.

[2] *Куржанский А. Б.* Альтернированный интеграл Понтрягина в теории синтеза управлений // Труды МИАН. 1999. Т. 224. С. 234–248.

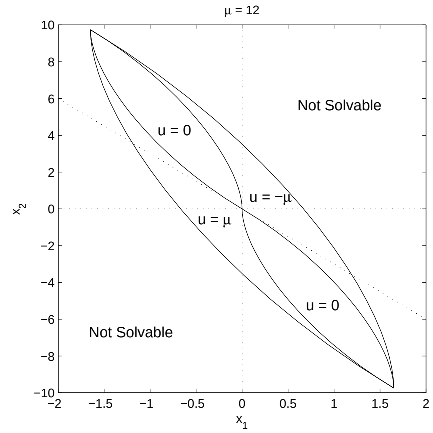

Рис. 3: Синтез управлений при двойном ограничении

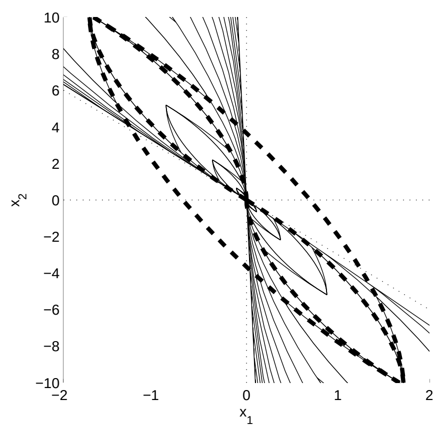

Рис. 4: Сходимость управления при двойном ограничении к импульсному управлению

[3] *Понтрягин Л. С.* О линейных дифференциальных играх II // Доклады АН СССР. 1967. Т. 175. № 4. С. 910–912.

[4] *Красовский Н. Н.* Игровые задачи о встрече движений. М.: Наука, 1970.

[5] *Kurzhanski A. B., Vályi I.* Ellipsoidal Calculus for Estimation and Control. SCFA. Boston: Birkhäuser, 1997.

[6] *Красовский Н. Н.* Теория управления движением. М.: Наука, 1968.

[7] *Куржанский А. Б.* Управление и наблюдение в условиях неопределенности. М.: Наука, 1977.

[8] *Дарьин А. Н., Куржанский А. Б.* Нелинейный синтез управления при двойных ограничениях // Дифференциальные уравнения. 2001. Т. 37. № 11. С. 1476–1484.

[9] *Дарьин А. Н., Куржанский А. Б.* Управление в условиях неопределенности при двойных ограничениях // Дифференциальные уравнения. 2003. Т. 39. № 11. С. 1474–1486.

[10] *Филиппов А. Ф.* Дифференциальные уравнения с разрывной правой частью. М.: Наука, 1985.

[11] *Kurzhanski A. B., Varaiya P.* On reachability under uncertainty // SIAM Journal on Control. 2002. V. 41. N. 1. P. 181–216.

[12] *Красовский Н. Н., Субботин А. И.* Позиционные дифференциальные игры. М.: Наука, 1974.

[13] *Рокафеллар Р.* Выпуклый анализ. М.: Мир, 1973.

[14] *Понтрягин Л. С.* Линейные дифференциальные игры преследования // Математический сборник. 1980. Т. 112 (154). № 3 (7). С. 307–330.

[15] *Красовский Н. Н.* Об одной задаче оптимального регулирования // ПММ. 1957. Т. 21. № 5. С. 670–677.

[16] *Neustadt L. W.* Optimization, a moment problem and nonlinear programming // SIAM Journal on Control. 1964. V. 2. N. 1.

[17] *Куржанский А. Б.* Оптимальные системы с импульсными управлениями // In *Дифференциальные игры и задачи управления* [29].

[18] *Гусев М. И.* Оптимальное управление обобщёнными процессами при невыпуклых фазовых ограничениях // In *Дифференциальные игры и задачи управления* [29].

[19] *Дыхта В. А., Самсонюк О. Н.* Оптимальное импульсное управление с приложениями. М.: Физматлит, 2003.

[20] *Motta M., Rampazzo F.* Space-time trajectories of nonlinear systems driven by ordinary and impulsive controls // Differential and Integral Equations. 1995. V. 8. P. 269–288.

[21] *Crandall M. G., Lions P.-L.* Viscosity solutions of Hamilton–Jacobi equations // Transactions of American Mathematical Society. 1983. V. 277. P. 1–41.

[22] *Fleming W. H., Soner H. M.* Controlled Markov Processes and Viscosity Solutions. N.Y.: Springer, 1993.

[23] *Субботин А. И.* Обобщенные решения уравнений в частных производных первого порядка. Перспективы динамической оптимизации. М., И.: Институт компьютерных исследований, 2003.

[24] *Bensoussan A., Lions J.-L.* Contrôle impulsionnel et inéquations quasi variationnelles. Paris, 1982.

[25] *Куржанский А. Б., Осипов Ю. С.* К управлению линейной системой обобщёнными воздействиями // Дифференциальные уравнения. 1969. Т. 5. № 8. С. 1360–1370.

[26] *Ky F.* Minimax theorems // Proc. Nat. Acad. of Sci. USA. 1953. V. 39. N. 1. P. 42–47.

[27] *Miller B. M., Ya. Rubinovich E.* Impulsive Control in Continuous and Discrete-Continuous Systems. N.Y.: Kluwer, 2003.

[28] *Clarke F. H.* Generalized gradients and applications // Transactions of American Mathematical Society. 1975. V. 205. P. 247–262.

[29] Дифференциальные игры и задачи управления. Свердловск: УНЦ АН СССР, 1975.