

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМ. М. В. ЛОМОНОСОВА

На правах рукописи



КУРЕННОЙ АЛЕКСЕЙ СВЯТОСЛАВОВИЧ

**НЬЮТОНОВСКИЕ МЕТОДЫ
РЕШЕНИЯ ЗАДАЧ ОПТИМИЗАЦИИ
С ЛИПШИЦЕВЫМИ ПРОИЗВОДНЫМИ**

Специальность 01.01.09 — дискретная математика
и математическая кибернетика

АВТОРЕФЕРАТ
ДИССЕРТАЦИИ НА СОИСКАНИЕ УЧЁНОЙ СТЕПЕНИ
КАНДИДАТА ФИЗИКО-МАТЕМАТИЧЕСКИХ НАУК

Москва
2014

Работа выполнена на кафедре исследования операций факультета вычислительной математики и кибернетики Московского государственного университета имени М. В. Ломоносова.

Научный руководитель: доктор физико-математических наук,
профессор
Измаилов Алексей Феридович.

Официальные оппоненты: доктор физико-математических наук,
Жадан Виталий Григорьевич;
кандидат физико-математических наук,
Жуковский Сергей Евгеньевич.

Ведущая организация: **Тамбовский государственный
университет им. Г. Р. Державина.**

Защита диссертации состоится 16 мая 2014 г. в 11 ч. 00 мин. на заседании диссертационного совета Д 501.001.44 в Московском государственном университете имени М. В. Ломоносова по адресу: 119991, ГСП-1, Москва, Ленинские горы, 2-й учебный корпус, факультет ВМК, аудитория 685. Желающие присутствовать на заседании диссертационного совета должны сообщить об этом за два дня по тел. (495)9393010 (для оформления заявки на пропуск).

С диссертацией можно ознакомиться в Научной библиотеке МГУ, а также на официальном сайте ВМК МГУ <http://cs.msu.ru> в разделе «Диссертации».

Автореферат разослан 13 марта 2014 г.

Ученый секретарь
диссертационного совета
к.т.н., с.н.с.

В. А. Костенко

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

В работе рассматриваются задачи оптимизации, производные целевой функции и ограничений которых являются локально липшицевыми. Работа посвящена исследованию локальной сходимости и скорости сходимости ряда эффективных численных методов применительно к этому классу задач в различных предположениях регулярности. Для этой цели в диссертации разрабатываются абстрактные итерационные схемы решения обобщенных уравнений, которые представляют собой богатый и удобный набор средств для теоретического анализа численных методов решения оптимизационных и вариационных задач в различных предположениях.

Актуальность темы. Задачи оптимизации с липшицевыми производными имеют широкий спектр приложений. К ним относятся кусочно-квадратичные задачи, возникающие, например, в машинном обучении, и обобщенные линейно-квадратичные задачи, находящие применение в оптимальном управлении. Помимо этого, задачи оптимизации с липшицевыми производными возникают в методах поиска обобщенного равновесия Нэша. Также к этому классу задач относятся поднятые переформулировки задач оптимизации с комплементарными ограничениями. Несмотря на многочисленные приложения, задачи оптимизации с липшицевыми производными до недавнего времени оставались малоизученными, особенно в плане обоснования соответствующих численных методов оптимизации, в связи с чем тема работы является актуальной.

Кроме того, актуальной является решаемая в работе задача создания универсального инструментария для тонкого анализа локальной сходимости методов решения вариационных задач и задач условной оптимизации в различных наборах предположений. В определенном смысле, эти построения суммируют и развивают разнообразные конструкции абстрактных итерационных схем для вариационных задач, известные ранее.

Целью диссертационного исследования является построение единой теории локальной сходимости ряда эффективных численных методов оптимизации, пригодных для решения задач с липшицевыми производными, в как можно более слабых предположениях регулярности.

Предмет и объект исследования. Объектом исследования в диссертационной работе являются задачи оптимизации с липшицевыми производными. Предмет исследования состоит в тонком анализе сходимости ряда эффективных численных методов оптимизации применительно к задачам этого класса в различных предположениях регулярности.

Методы исследования. При выполнении диссертационного исследования использовались средства современного негладкого анализа, вариационного анализа, теории оптимизации, а также подходы современной численной оптимизации. Исследование локальной сходимости численных методов оптимизации осуществляется в диссертации путем представления их в виде частных случаев предварительно разработанных абстрактных итерационных схем и применения их теории локальной сходимости, которая также разработана в диссертации.

Научная новизна. В диссертации разработаны новые абстрактные итерационные схемы решения обобщенных уравнений и теория их локальной сходимости, обобщающая и развивающая известные ранее конструкции такого рода. Эта теория находит многочисленные приложения, в том числе ее применение позволяет получить новые результаты о локальной сходимости метода модифицированных функций Лагранжа и метода множителей с линеаризованными ограничениями для задач с липшицевыми производными. При этом некоторые из полученных результатов о локальной сходимости являются новыми и для задач оптимизации, целевая функция и ограничения которых дважды непрерывно дифференцируемы. Кроме того, в работе доказаны новые необходимые и достаточные условия прямой сверхлинейной сходимости полугладкого метода последовательного квадратичного программирования, установлены неизвестные ранее соотношения между важнейшими условиями регулярности для смешанных комплементарных задач, получена новая оценка расстояния до множества решений системы Каруша–Куна–Таккера задачи оптимизации с липшицевыми производными, а также установлены неизвестные ранее полезные соотношения между рядом обобщенных дифференциальных объектов.

Достоверность научных положений обусловлена строгостью математических доказательств и использованием апробированных научных методов.

Теоретическая и практическая значимость работы. Разработанные в диссертации абстрактные схемы решения обобщенных уравнений представляют собой достаточно универсальный инструмент теоретического исследования методов решения оптимизационных и вариационных задач, и в связи с этим они обладают большой теоретической ценностью. Данные абстрактные схемы не просто являются средством получения результатов о локальной сходимости различных алгоритмов: они дают общее представление о необходимых ингредиентах такого анализа в различных предположениях, а также о факторах влияющих на скорость сходимости методов. В этой связи данные абстрактные схемы не просто являются удобным инструментом анализа конкретных алгоритмов, но также имеют несомненное методологическое значение.

В диссертации получены первые результаты о локальной сходимости метода модифицированных функций Лагранжа и метода множителей с линеаризованными ограничениями применительно к задачам с липшицевыми производными. При этом некоторые из указанных результатов являются новыми и в контексте задач, целевая функция и ограничения которых являются дважды непрерывно дифференцируемыми. Тем самым, диссертация вносит значимый вклад в теорию локальной сходимости метода модифицированных функций и метода множителей с линеаризованными ограничениями. С точки зрения практики, знание свойств и понимание механизмов и принципов локальной сходимости этих алгоритмов может быть полезным при их реализации в программных пакетах.

Доказанные в диссертации необходимые и достаточные условия прямой сверхлинейной сходимости полугладкого метода последовательного квадратичного программирования являются существенным дополнением теории его сходимости. Их практическая значимость объясняется тем, что они могут быть использованы при конструировании квазиньютоновских методов решения задач оптимизации с липшицевыми производными.

Полученная в работе полная картина взаимоотношений между важнейшими условиями регулярности для смешанных комплементарных задач позволяет сравнивать между собой теоретические результаты о сходимости различных численных методов, чем обусловлена ее теоретическая ценность. С практической точки зрения знание того, как соотносятся различные условия регулярности между собой, также имеет значение, поскольку помогает при выборе численных методов для решения задач, возникающих на практике.

Установленная в диссертации оценка расстояния до множества решений системы Каруша–Куна–Таккера задачи оптимизации с липшицевыми производными используется при теоретическом исследовании оптимизационных алгоритмов. С другой стороны, указанная оценка оказывается полезной при глобализации сходимости численных методов решения задач оптимизации с липшицевыми производными.

Наконец, доказанные в работе соотношения между рядом обобщенных дифференциальных объектов имеют теоретическое значение, поскольку позволяют сравнивать между собой различные теоретические результаты о сходимости оптимизационных алгоритмов, использующих эти объекты.

Соответствие диссертации паспорту научной специальности. В диссертации проведено исследование локальной сходимости ряда методов минимизации функций. Поэтому в ней решена задача, относящаяся к направлению «математическое программирование», которое входит в содержание специальности 01.01.09.

Апробация результатов. Результаты, полученные в диссертации, были представлены на XXI международном симпозиуме по математическому программированию «ISMP2012» (Берлин, Германия), на международной конференции по непрерывной оптимизации «ICCOPT2013» (Лиссабон, Португалия), на X всемирном конгрессе по структурной и междисциплинарной оптимизации «WCSMO-10» (Орландо, США, 2013), на ежегодных международных научных конференциях студентов и молодых ученых «Ломоносов-2011», «Ломоносов-2012» (Москва), на ежегодной научной конференции «Тихоновские чтения» (Москва, 2012), на ежегодной

научной конференции «Ломоносовские чтения» (Москва, 2011), а также на VII Московской международной конференции по исследованию операций «ORM2013».

Публикации. Полученные в диссертации результаты опубликованы в 12 работах, список которых приведен в конце автореферата. В их число входит 5 статей, опубликованных в журналах из списка ВАК РФ [4, 7, 9, 10, 11].

Структура и объем диссертации. Диссертация состоит из введения, трех глав, заключения и списка литературы из 91 источника. Общий объем диссертации — 134 страницы.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении дана постановка задачи оптимизации с липшицевыми производными, приведены примеры возникновения таких задач, обоснована актуальность исследования, описана его методика, отражена научная новизна диссертации, сформулированы основные результаты, выносимые на защиту, а также приведена информация об апробации результатов.

Первая глава посвящена вспомогательным вопросам вариационного и негладкого анализа.

В разделе 1.1 рассматривается ряд широко используемых концепций регулярности решения *смешанной комплементарной задачи* (СКЗ)

$$z \in [\ell, u], \quad \langle \Phi(z), y - z \rangle \geq 0 \quad \forall y \in [\ell, u],$$

где $\Phi: \mathbb{R}^s \mapsto \mathbb{R}^s$ — заданное отображение, а $[\ell, u]$ есть (обобщенный) параллелепипед,

$$[\ell, u] = \{z \in \mathbb{R}^s \mid \ell_i \leq z_i \leq u_i, i = 1, \dots, s\},$$

задаваемый числами $\ell_i \in \mathbb{R} \cup \{-\infty\}$ и $u_i \in \mathbb{R} \cup \{+\infty\}$, $\ell_i < u_i$, $i = 1, \dots, s$. В число рассматриваемых условий регулярности входят сильная регулярность, полуустойчивость, *BD* и *CD*-регулярности переформулировок СКЗ с использованием функции естественной невязки и функции Фишера–Бурмейстера и др. Особое внимание уделяется частному случаю

СКЗ — системе Каруша–Куна–Таккера (ККТ). Основным результатом раздела является полная картина взаимоотношений между рассматриваемыми условиями регулярности.

В разделе 1.2 исследуются соотношения между рядом обобщенных дифференциальных объектов, в число которых входит частный дифференциал Кларка и проекция полного дифференциала Кларка. Дифференциалом Кларка отображения $\Phi: \mathbb{R}^s \mapsto \mathbb{R}^r$ в точке $\bar{z} \in \mathbb{R}^s$ называется множество

$$\partial\Phi(\bar{z}) = \text{conv}\{J \in \mathbb{R}^{r \times s} \mid \exists \{z^k\} \subset \mathcal{D}_\Phi: \{z^k\} \rightarrow \bar{z}, \{\Phi'(z^k)\} \rightarrow J\},$$

где через \mathcal{D}_Φ обозначено множество точек дифференцируемости отображения Φ . Пусть $s = s_1 + s_2$, $z = (z_1, z_2) \in \mathbb{R}^{s_1} \times \mathbb{R}^{s_2}$. Частным дифференциалом Кларка отображения Φ по переменной z_1 в точке $\bar{z} = (\bar{z}_1, \bar{z}_2) \in \mathbb{R}^{s_1} \times \mathbb{R}^{s_2}$ называется множество

$$\partial_{z_1}\Phi(\bar{z}_1, \bar{z}_2) = \partial\Phi(\cdot, \bar{z}_2) \Big|_{z_1=\bar{z}_1}.$$

Проекцией полного дифференциала Кларка отображения Φ в точке $\bar{z} = (\bar{z}_1, \bar{z}_2) \in \mathbb{R}^{s_1} \times \mathbb{R}^{s_2}$ называется множество

$$\Pi_{z_1}\partial\Phi(\bar{z}_1, \bar{z}_2) = \{J_1 \in \mathbb{R}^{r \times s_1} \mid \exists J_2 \in \mathbb{R}^{r \times s_2}: [J_1 \ J_2] \in \partial\Phi(\bar{z}_1, \bar{z}_2)\}.$$

Основной результат раздела состоит в том, что если отображение Φ имеет вид

$$\Phi(z_1, z_2) = K(z_1)z_2 + b(z_1), \quad (1)$$

где $K: \mathbb{R}^{s_1} \mapsto \mathbb{R}^{r \times s_2}$ и $b: \mathbb{R}^{s_1} \mapsto \mathbb{R}^r$ — заданные отображения, то в любой точке $\bar{z} = (\bar{z}_1, \bar{z}_2) \in \mathbb{R}^{s_1} \times \mathbb{R}^{s_2}$ такой, что отображения K и b локально липшицевы в \bar{z}_1 , выполнено равенство

$$\partial_{z_1}\Phi(\bar{z}_1, \bar{z}_2) = \Pi_{z_1}\partial\Phi(\bar{z}_1, \bar{z}_2).$$

Рассмотрение отображений вида (1) вызвано тем, что в таком виде представляется градиент функции Лагранжа задачи оптимизации по прямой переменной.

В разделе 1.3 доказывается оценка расстояния до множества решений системы ККТ задачи оптимизации

$$f(x) \rightarrow \min, \quad h(x) = 0, \quad g(x) \leq 0, \quad (2)$$

с липшицевыми производными. Напомним, что последнее означает, что функция $f: \mathbb{R}^n \mapsto \mathbb{R}$ и отображения $h: \mathbb{R}^n \mapsto \mathbb{R}^l$ и $g: \mathbb{R}^n \mapsto \mathbb{R}^m$ дифференцируемы, и что их производные являются локально липшицевыми. Отображение $\Phi: \mathbb{R}^s \mapsto \mathbb{R}^r$ называется локально липшицевым в точке $z \in \mathbb{R}^s$, если оно удовлетворяет условию Липшица в некоторой окрестности этой точки, т. е. если существует число $\ell > 0$ и окрестность $U \subset \mathbb{R}^s$ точки z такие, что

$$\|\Phi(z_1) - \Phi(z_2)\| \leq \ell \|z_1 - z_2\| \quad \forall z_1, z_2 \in U.$$

При этом отображение называется локально липшицевым, если оно локально липшицево в каждой точке своей области определения.

Система ККТ задачи (2) имеет вид

$$\begin{aligned} \frac{\partial L}{\partial x}(x, \lambda, \mu) &= 0, \\ h(x) &= 0, \quad \mu \geq 0, \quad g(x) \leq 0, \quad \langle \mu, g(x) \rangle = 0, \end{aligned} \quad (3)$$

где $L: \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m \rightarrow \mathbb{R}$ — функция Лагранжа задачи (2),

$$L(x, \lambda, \mu) = f(x) + \langle \lambda, h(x) \rangle + \langle \mu, g(x) \rangle.$$

Пусть $\mathcal{M}(\bar{x}) \subset \mathbb{R}^l \times \mathbb{R}^m$ — совокупность множителей Лагранжа, отвечающих точке $\bar{x} \in \mathbb{R}^n$ (т. е. множество пар $(\lambda, \mu) \in \mathbb{R}^l \times \mathbb{R}^m$ таких, что тройка (\bar{x}, λ, μ) удовлетворяет системе (3)). Точка $\bar{x} \in \mathbb{R}^n$ называется стационарной точкой задачи (2), если $\mathcal{M}(\bar{x}) \neq \emptyset$. Будем обозначать через

$$A = A(\bar{x}) = \{i = 1, \dots, m \mid g_i(\bar{x}) = 0\}$$

множество индексов активных ограничений в точке $\bar{x} \in \mathbb{R}^n$. Кроме того, в том случае, если \bar{x} — стационарная точка задачи (2), для всякого множителя Лагранжа $(\bar{\lambda}, \bar{\mu}) \in \mathcal{M}(\bar{x})$ через

$$\begin{aligned} A_+ &= A_+(\bar{x}, \bar{\mu}) = \{i \in A(\bar{x}) \mid \bar{\mu}_i > 0\}, \\ A_0 &= A_0(\bar{x}, \bar{\mu}) = \{i \in A(\bar{x}) \mid \bar{\mu}_i = 0\} \end{aligned}$$

будут обозначаться множества индексов сильно и слабо активных ограничений соответственно.

В данном разделе устанавливается эквивалентность следующих трех свойств.

Свойство 1 (верхняя липшицева устойчивость решений системы ККТ при канонических возмущениях). Существует число $\ell > 0$ и окрестность \mathcal{U} точки $(\bar{x}, \bar{\lambda}, \bar{\mu})$ такие, что для любого $\sigma = (a, b, c) \in \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m$, достаточно близкого к точке $(0, 0, 0)$, всякое решение $(x(\sigma), \lambda(\sigma), \mu(\sigma)) \in \mathcal{U}$ возмущенной системы ККТ

$$\frac{\partial L}{\partial x}(x, \lambda, \mu) = a, \quad h(x) = b, \quad \mu \geq 0, \quad g(x) \leq c, \quad \langle \mu, g(x) - c \rangle = 0$$

удовлетворяет оценке

$$\|x(\sigma) - \bar{x}\| + \text{dist}((\lambda(\sigma), \mu(\sigma)), \mathcal{M}(\bar{x})) \leq \ell \|\sigma\|.$$

Свойство 2 (оценка расстояния для системы ККТ). Существует число $\ell > 0$ и окрестность \mathcal{U} точки $(\bar{x}, \bar{\lambda}, \bar{\mu})$ такие, что для любого вектора $(x, \lambda, \mu) \in \mathcal{U}$ выполняется неравенство

$$\|x - \bar{x}\| + \text{dist}((\lambda, \mu), \mathcal{M}(\bar{x})) \leq \ell \left\| \begin{pmatrix} \frac{\partial L}{\partial x}(x, \lambda, \mu) \\ h(x) \\ \min\{\mu, -g(x)\} \end{pmatrix} \right\|.$$

Свойство 3 (некритичность множителя $(\bar{\lambda}, \bar{\mu})$). Не существует тройки $(\xi, \eta, \zeta) \in \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m$ с $\xi \neq 0$, удовлетворяющей системе

$$d + (h'(\bar{x}))^T \eta + (g'(\bar{x}))^T \zeta = 0, \quad h'(\bar{x})\xi = 0, \quad g'_{A_+}(\bar{x})\xi = 0,$$

$$\zeta_{A_0} \geq 0, \quad g'_{A_0}(\bar{x})\xi \leq 0, \quad \zeta_i \langle g'_i(\bar{x}), \xi \rangle = 0, \quad i \in A_0, \quad \zeta_{\{1, \dots, m\} \setminus A} = 0$$

с некоторым $d \in C_x \frac{\partial L}{\partial x}(\bar{x}, \bar{\lambda}, \bar{\mu})(\xi)$, где через $C_x \frac{\partial L}{\partial x}(\bar{x}, \bar{\lambda}, \bar{\mu})(\xi)$ обозначена контингентная производная отображения $\frac{\partial L}{\partial x}(\cdot, \bar{\lambda}, \bar{\mu})$ в точке \bar{x} по направлению ξ , то есть множество

$$C_x \frac{\partial L}{\partial x}(\bar{x}, \bar{\lambda}, \bar{\mu})(\xi) = \left\{ w \in \mathbb{R}^r \mid \exists \{t_k\} \subset \mathbb{R}_+, \{t_k\} \rightarrow 0+ : \right. \\ \left. \left\{ \left(\frac{\partial L}{\partial x}(\bar{x} + t_k \xi, \bar{\lambda}, \bar{\mu}) - \frac{\partial L}{\partial x}(\bar{x}, \bar{\lambda}, \bar{\mu}) \right) / t_k \right\} \rightarrow w \right\}.$$

Результаты главы 1 опубликованы в статьях [1, 2, 5, 9, 11].

Вторая глава посвящена разработке и анализу итерационных схем решения *обобщенных уравнений*, т. е. включений вида

$$\Phi(u) + N(u) \ni 0, \quad (4)$$

где $\Phi: \mathbb{R}^\nu \mapsto \mathbb{R}^\nu$ — однозначное отображение (именуемое *базовым* или *базой*), а $N: \mathbb{R}^\nu \mapsto 2^{\mathbb{R}^\nu}$ — многозначное (иногда называемое *полем*). Теоремы о сходимости этих итерационных схем являются основным средством анализа численных методов решения задач оптимизации с липшицевыми производными в главе 3.

Итерационные схемы, рассматриваемые в разделе 2.1, предназначены для решения обобщенных уравнений с произвольной непрерывной базой и имеют абстрактный характер. В первых двух пунктах раздела исследуются свойства локальной сходимости схемы, итерация которой состоит в следующем. По текущей точке $u^k \in \mathbb{R}^\nu$ и текущему значению параметра π^k , принадлежащему множеству Π произвольной природы, очередная точка u^{k+1} генерируется как решение подзадачи

$$\mathcal{A}(\pi^k, u^k, u) + N(u) \ni 0, \quad \|u^{k+1} - u^k\| \leq \delta, \quad (5)$$

где для любых $\pi \in \Pi$ и $\tilde{u} \in \mathbb{R}^\nu$ точечно-множественное отображение $\mathcal{A}(\pi, \tilde{u}, \cdot)$ из \mathbb{R}^ν в $2^{\mathbb{R}^\nu}$ представляет собой аппроксимацию отображения Φ вблизи \tilde{u} , а $\delta > 0$ — фиксированное число. Второе условие в формуле (5) называется *условием локализации* и служит для исключения из рассмотрения тех решений подзадачи, которые далеки от искомого решения обобщенного уравнения.

В пункте 2.1.1 устанавливается локальная сходимость итерационной схемы (5) в предположении о *сильной метрической регулярности* искомого решения обобщенного уравнения (4). Решение \bar{u} обобщенного уравнения (4) называется *сильно метрически регулярным*, если существует константа $\ell > 0$ такая, что для всякого $r \in \mathbb{R}^\nu$, достаточно близкого к 0, возмущенное обобщенное уравнение

$$\Phi(u) + N(u) \ni r \quad (6)$$

имеет вблизи точки \bar{u} единственное решение $u(r)$, причем отображение $u(\cdot)$ является локально липшицевым в точке 0 с константой ℓ . В соответствующем результате о локальной сходимости предполагается, что отображение \mathcal{A} является однозначным и аппроксимирует отображение Φ в достаточно сильном смысле: разность $\Phi(\cdot) - \mathcal{A}(\pi, \tilde{u}, \cdot)$ должна быть локально липшицевой в искомом решении с малой константой липшица равномерно по параметру π и точкам $\tilde{u} \in \mathbb{R}^\nu$, близким к этому решению. В указанных предположениях гарантируется существование числа $\delta > 0$ такого, что из любого начального приближения, достаточно близкого к искомому решению обобщенного уравнения (4), итерационная схема (5) генерирует единственную траекторию, которая сходится к этому решению с линейной скоростью. При этом скорость сходимости является сверхлинейной, если точность аппроксимации отображения Φ возрастает с ростом номера итерации (см. теорему 2.1 диссертации).

В пункте 2.1.2 сильная метрическая регулярность заменяется более слабым условием — *полуустойчивостью*. Решение \bar{u} обобщенного уравнения (4) называется полуустойчивым, если существует константа $\ell > 0$ такая, что для любого $r \in \mathbb{R}^\nu$ всякое решение $u(r)$ возмущенного обобщенного уравнения (6), достаточно близкое к \bar{u} , удовлетворяет оценке

$$\|u(r) - \bar{u}\| \leq \ell \|r\|.$$

Требование на качество аппроксимации отображения Φ также ослабляется. В частности, отображение \mathcal{A} может быть многозначным. В отличие от предположений пункта 2.1.1, эти ослабленные предположения не гарантируют разрешимость подзадач итерационной схемы (5) вблизи искомого решения. Выполнение этого свойства приходится требовать явно. Кроме того, доказываемая в пункте 2.1.2 теорема о локальной сходимости итерационной схемы (5) не гарантирует единственность генерируемой ей траектории.

Итерационная схема, рассматриваемая в пункте 2.1.3, предназначена для поиска неизолированных решений обобщенного уравнения (4). Такие решения не могут обладать ни полуустойчивостью, ни, тем более, сильной метрической регулярностью. По текущей точке $u^k \in \mathbb{R}^\nu$ и текущему

значению параметра $\pi^k \in \Pi$ очередная точка u^{k+1} генерируется в этой схеме как решение подзадачи

$$\mathcal{A}(\pi^k, u^k, u) + N(u) \ni 0, \quad \|u^{k+1} - u^k\| \leq c \operatorname{dist}(u^k, \bar{U}), \quad (7)$$

где \bar{U} — множество решений обобщенного уравнения (4), а $c > 0$ — фиксированное число. От схемы (5) схема (7) отличается более сильным условием локализации. Для схемы (7) доказывается теорема о локальной сходимости и скорости сходимости. Ее предположения состоят в требовании разрешимости подзадач схемы, определенных условиях на качество аппроксимации отображения Φ , а также в выполнении свойства, которое носит название *верхне-липшицевого поведения решения при канонических возмущениях*. Решение $\bar{u} \in \bar{U}$ обладает этим свойством, если существует константа $\ell > 0$ такая, что для любого $r \in \mathbb{R}^r$ всякое решение $u(r)$ возмущенного обобщенного уравнения (6), достаточно близкое к \bar{u} , удовлетворяет оценке

$$\operatorname{dist}(u(r), \bar{U}) \leq \ell \|r\|.$$

Если решение \bar{u} изолировано, то верхне-липшицево поведение при канонических возмущениях эквивалентно полуустойчивости. Заметим, что траектория, генерируемая схемой (7), сходится не к \bar{u} , а к некоторому решению u^* обобщенного уравнения (4), которое можно сделать сколь угодно близким к \bar{u} за счет выбора начального приближения.

Метод решения обобщенных уравнений, исследуемый в разделе 2.2, является менее абстрактным по сравнению с итерационными схемами, рассматриваемыми в разделе 2.1. При этом он накладывает на базу обобщенного уравнения более сильные ограничения гладкости: предполагается, что отображение Φ является *полугладким* в искомом решении \bar{u} обобщенного уравнения (4). Отображение $\Phi: \mathbb{R}^r \mapsto \mathbb{R}^r$ называется *полугладким* в точке $u \in \mathbb{R}^r$, если оно является локально липшицевым в u , дифференцируемым в u по любому направлению и удовлетворяет условию

$$\sup_{J \in \partial \Phi(u+v)} \|\Phi(u+v) - \Phi(u) - Jv\| = o(\|v\|).$$

Отображение называется полугладким, если оно полугладко в каждой точке своей области определения. Метод решения обобщенных уравне-

ний, о котором идет речь, называется полугладким методом Джозефи–Ньютона. Его итерация состоит в решении (частично) линеаризованного обобщенного уравнения

$$\Phi(u^k) + J_k(u - u^k) + N(u) \ni 0,$$

где $u^k \in \mathbb{R}^\nu$ — текущее приближение к искомому решению уравнения (4), а J_k — произвольная матрица из множества $\partial\Phi(u^k)$. Основным результатом раздела является теорема о локальной сверхлинейной сходимости этого метода в предположении о *CD-регулярности* искомого решения \bar{u} . Решение \bar{u} обобщенного уравнения (4) называется *CD-регулярным*, если для любой матрицы $J \in \partial\Phi(\bar{u})$ и любого $r \in \mathbb{R}^\nu$, достаточно близкого 0, возмущенное частично линеаризованное уравнение

$$\Phi(\bar{u}) + J(u - \bar{u}) + N(u) \ni r$$

имеет вблизи \bar{u} единственное решение $u_J(r)$, и отображение $u_J(\cdot)$ является локально липшицевым в точке 0.

Результаты главы 2 опубликованы в работах [3, 6, 7, 10].

В третьей главе осуществляется анализ ряда эффективных численных методов применительно к задачам оптимизации с липшицевыми производными. Для этих методов устанавливаются результаты о локальной сходимости и скорости сходимости в различных предположениях. В качестве инструментов анализа при этом выступают итерационные схемы, представленные в главе 2. А именно, исследуемые методы представляются в виде частных случаев этих итерационных схем, и соответствующие результаты о локальной сходимости выводятся из теорем, доказанных в главе 2.

Раздел 3.1 посвящен методу модифицированных функций Лагранжа. В пункте 3.1.1 исследуется этот метод для задачи (2). Модифицированной функцией Лагранжа $L_\sigma: \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m \mapsto \mathbb{R}$ задачи (2) называется функция

$$L_\sigma(x, \lambda, \mu) = f(x) + 2\sigma(\|\lambda + h(x)/\sigma\|_2^2 + \|\max\{0, \mu + g(x)/\sigma\}\|_2^2),$$

где максимум берется покомпонентно. Параметр σ в определении модифицированной функции Лагранжа называется *обратным параметром*

штрафа. Итерация метода модифицированных функций Лагранжа для задачи (2) состоит в следующем. По текущему двойственному приближению $(\lambda^k, \mu^k) \in \mathbb{R}^l \times \mathbb{R}^m$ и текущему значению обратного параметра штрафа $\sigma_k > 0$ очередное прямое приближение x^{k+1} ищется как стационарная точка задачи

$$L_{\sigma_k}(x, \lambda^k, \mu^k) \rightarrow \min, \quad x \in \mathbb{R}^n, \quad (8)$$

а очередное двойственное приближение $(\lambda^{k+1}, \mu^{k+1})$ вычисляется по формулам

$$\lambda^{k+1} = \lambda^k + h(x^{k+1})/\sigma_k, \quad \mu^{k+1} = \max\{0, \mu^k + g(x^{k+1})/\sigma_k\}. \quad (9)$$

Для данного метода доказываются теоремы о локальной сходимости и скорости сходимости в различных наборах предположений. В первой из этих теорем предполагается, что в искомой стационарной точке \bar{x} задачи (2) выполнено *условие линейной независимости*, и что единственный отвечающий этой точке множитель Лагранжа $(\bar{\lambda}, \bar{\mu})$ удовлетворяет *сильному достаточному условию второго порядка оптимальности*. Первое из этих условий по определению означает, что градиенты ограничений-равенств и активных ограничений-неравенств в точке \bar{x} линейно независимы. Второе условие заключается в том, что

$$\forall H \in \partial_x \frac{\partial L}{\partial x}(\bar{x}, \bar{\lambda}, \bar{\mu}) \quad \langle H\xi, \xi \rangle > 0 \quad \forall \xi \in C_+(\bar{x}, \bar{\mu}) \setminus \{0\},$$

где

$$C_+(\bar{x}, \bar{\mu}) = \{\xi \in \mathbb{R}^n \mid h'(\bar{x})\xi = 0, g'_{A_+(\bar{x}, \bar{\mu})}(\bar{x})\xi = 0\}.$$

В указанных предположениях найдутся числа $\bar{\sigma} > 0$ и $\delta > 0$ такие, что для любой последовательности $\{\sigma_k\} \subset (0, \bar{\sigma}]$ из любого начального приближения (x^0, λ^0, μ^0) , достаточно близкого к тройке $(\bar{x}, \bar{\lambda}, \bar{\mu})$, метод модифицированных функций Лагранжа (8), (9) генерирует единственную траекторию, удовлетворяющую условию

$$\|(x^{k+1}, \lambda^{k+1}, \mu^{k+1}) - (x^k, \lambda^k, \mu^k)\| \leq \delta, \quad k = 0, 1, \dots,$$

и эта траектория сходится к $(\bar{x}, \bar{\lambda}, \bar{\mu})$ с линейной скоростью. При этом скорость сходимости является сверхлинейной, если $\{\sigma_k\} \rightarrow 0$.

Во второй теореме условие линейной независимости ослабляется до *строгого условия Мангасариана–Фромовица*, которое заключается в единственности множителя $(\bar{\lambda}, \bar{\mu})$, отвечающего стационарной точке \bar{x} задачи (2). От этого множителя требуется, чтобы он был не критическим. Кроме того, предполагается, что для задачи (2) в точке \bar{x} выполнено *условие квадратичного роста*: существуют числа $\gamma > 0$ и $\rho > 0$ такие, что

$$f(x) \geq f(\bar{x}) + \gamma \|x - \bar{x}\|^2 \quad \forall x \in \mathbb{R}^n : \|x - \bar{x}\| \leq \rho, \quad h(x) = 0, \quad g(x) \leq 0.$$

Отметим, что при выполнении строгого условия Мангасариана–Фромовица условие квадратичного роста следует из *достаточного условия второго порядка оптимальности*, состоящего в том, что

$$\forall H \in \partial_x \frac{\partial L}{\partial x}(\bar{x}, \bar{\lambda}, \bar{\mu}) \quad \langle H\xi, \xi \rangle > 0 \quad \forall \xi \in C(\bar{x}) \setminus \{0\},$$

где $C(\bar{x})$ — критический конус задачи (2),

$$\begin{aligned} C(\bar{x}) &= \{\xi \in \mathbb{R}^n \mid h'(\bar{x})\xi = 0, \quad g'_{A(\bar{x})}(\bar{x})\xi \leq 0, \quad \langle f'(\bar{x}), \xi \rangle \leq 0\} = \\ &= \{\xi \in \mathbb{R}^n \mid h'(\bar{x})\xi = 0, \quad g'_{A_+(\bar{x}, \bar{\mu})}(\bar{x})\xi = 0, \quad g'_{A_0(\bar{x}, \bar{\mu})}(\bar{x})\xi \leq 0\}. \end{aligned}$$

Поскольку последнее условие в свою очередь следует из сильного достаточного условия второго порядка оптимальности, предположения второй теоремы пункта 3.1.1 слабее предположений первой теоремы этого пункта. В части утверждений эти теоремы отличаются тем, что вторая не гарантирует единственность траектории метода.

В третьей теореме пункта 3.1.1 устанавливается локальная сходимость метода модифицированных функций Лагранжа без каких-либо предположений о регулярности ограничений в искомой стационарной точке $\bar{x} \in \mathbb{R}^n$ задачи (2). В ней предполагается только то, что множитель Лагранжа $(\bar{\lambda}, \bar{\mu}) \in \mathcal{M}(\bar{x})$, вблизи которого стартует двойственная траектория, удовлетворяет достаточному условию второго порядка оптимальности. Используемое при этом условие локализации является более сильным, чем в первой и второй теоремах данного пункта. Заметим, что двойственная траектория, генерируемая алгоритмом, сходится не к $(\bar{\lambda}, \bar{\mu})$, а к множителю $(\lambda^*, \mu^*) \in \mathcal{M}(\bar{x})$, который можно сделать сколь угодно близким к $(\bar{x}, \bar{\lambda})$ за счет выбора начального приближения.

В пункте 3.1.2 рассматривается метод модифицированных функций для задачи оптимизации с липшицевыми производными, в которой отсутствуют ограничения-неравенства

$$f(x) \rightarrow \min, \quad h(x) = 0. \quad (10)$$

Итерация этого метода состоит в следующем. По текущему двойственному приближению $\lambda^k \in \mathbb{R}^l$ и текущему значению обратного параметра штрафа $\sigma^k > 0$ очередное прямое приближение $x^{k+1} \in \mathbb{R}^n$ находится как стационарная точка задачи

$$f(x) + \langle \lambda^k, h(x) \rangle + \frac{1}{2\sigma_k} \|h(x)\|_2^2 \rightarrow \min, \quad x \in \mathbb{R}^n,$$

а очередное двойственное приближение $\lambda^{k+1} \in \mathbb{R}^l$ полагается равным $\lambda^k + h(x^{k+1})/\sigma_k$. Для этого метода доказываются две теоремы о локальной сходимости и скорости сходимости, не требующие регулярности ограничений в искомой стационарной точке $\bar{x} \in \mathbb{R}^n$ задачи (10) и не предполагающие, что отвечающий точке \bar{x} множитель $\bar{\lambda}$, вблизи которого стартует двойственная траектория метода, удовлетворяет достаточному условию второго порядка оптимальности

$$\forall H \in \partial_x \frac{\partial L}{\partial x}(\bar{x}, \bar{\lambda}) \quad \langle H\xi, \xi \rangle > 0 \quad \forall \xi \in \ker h'(\bar{x}) \setminus \{0\},$$

где $L: \mathbb{R}^n \times \mathbb{R}^l \mapsto \mathbb{R}$ — функция Лагранжа задачи (10),

$$L(x, \lambda) = f(x) + \langle \lambda, h(x) \rangle.$$

Вместо достаточного условия второго порядка оптимальности на множитель $\bar{\lambda}$ в этих теоремах накладывается более слабое условие, состоящее в том, что

$$\forall H \in \partial_x \frac{\partial L}{\partial x}(\bar{x}, \bar{\lambda}) \quad H\xi \notin \text{im}(h'(\bar{x}))^T \quad \forall \xi \in \ker h'(\bar{x}) \setminus \{0\}.$$

В первой из этих теорем предполагается, что обратный параметр штрафа σ_k в методе модифицированных функций выбирается согласно правилу

$$\sigma_k = \left\| \left(\frac{\partial L}{\partial x}(x^k, \lambda^k), h(x^k) \right) \right\|,$$

и устанавливается сверхлинейная сходимость метода. Во второй теореме доказывается линейная сходимость в предположении, что обратный параметр штрафа фиксирован.

Раздел 3.2 посвящен методу множителей с линеаризованными ограничениями. Этот метод вводится для задачи вида (2), в которой $m = n$ и $g(x) = -x$, $x \in \mathbb{R}^n$, то есть для задачи с ограничениями-равенствами и условием неотрицательности переменных. Итерация метода множителей с линеаризованными ограничениями состоит в следующем. По текущему приближению $(x^k, \lambda^k, \mu^k) \in \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^n$ и значению параметра штрафа $s_k \geq 0$ очередное прямое приближение $x^{k+1} \in \mathbb{R}^n$ вычисляется как стационарная точка задачи оптимизации

$$\begin{aligned} f(x) + \langle \lambda^k, h(x) \rangle + \frac{s_k}{2} \|h(x)\|_2^2 &\rightarrow \min, \\ h(x^k) + h'(x^k)(x - x^k) &= 0, x \geq 0, \end{aligned}$$

а очередная точка $(\lambda^{k+1}, \mu^{k+1}) \in \mathbb{R}^l \times \mathbb{R}^n$ двойственной траектории выбирается таким образом, чтобы пара $(\lambda^{k+1} - \lambda^k, \mu^{k+1})$ была отвечающим x^{k+1} множителем Лагранжа. Для этого метода доказывается две теоремы о локальной квадратичной сходимости. Первая теорема требует выполнения условия линейной независимости и сильного достаточного условия второго порядка оптимальности. Во второй теореме аналогичный результат (но без единственности траектории) устанавливается при выполнении строгого условия Мангасариана–Фромовица и достаточного условия второго порядка оптимальности.

Раздел 3.3 посвящен получению необходимых и достаточных условий для прямой сверхлинейной сходимости метода последовательного квадратичного программирования для задачи (2) в случае, когда отображения f' , h' и g' являются полугладкими. Итерация метода последовательного квадратичного программирования состоит в следующем. По текущему приближению (x^k, λ^k, μ^k) к решению системы ККТ задачи (2) и текущей симметричной итерационной матрице $H_k \in \mathbb{R}^{n \times n}$ очередное прямое

приближение $x^{k+1} \in \mathbb{R}^n$ вычисляется как стационарная точка задачи

$$\begin{aligned} \langle f'(x^k), x - x^k \rangle + \frac{1}{2} \langle H_k(x - x^k), x - x^k \rangle &\rightarrow \min, \\ h(x^k) + h'(x^k)(x - x^k) = 0, \quad g(x^k) + g'(x^k)(x - x^k) &\leq 0, \end{aligned}$$

а очередное двойственное приближение $(\lambda^{k+1}, \mu^{k+1}) \in \mathbb{R}^l \times \mathbb{R}_+^m$ — как отвечающий ей множитель Лагранжа. Основным результатом раздела является теорема, которая утверждает, что в естественных предположениях прямая сверхлинейная сходимость метода последовательного квадратичного программирования эквивалентна условию

$$\begin{aligned} \pi_{C(\bar{x})} \left(\frac{\partial L}{\partial x}(x^{k+1}, \lambda^k, \mu^k) - \frac{\partial L}{\partial x}(x^k, \lambda^k, \mu^k) - H_k(x^{k+1} - x^k) \right) = \\ = o(\|x^{k+1} - x^k\|), \end{aligned}$$

где \bar{x} — решение задачи, к которому сходится прямая траектория метода, а $\pi_{C(\bar{x})}(\cdot)$ — оператор взятия евклидовой проекции на критический конус задачи в точке \bar{x} . Это условие представляет собой полугладкую версию условия Дениса–Морэ.

Результаты главы 3 опубликованы в статьях [4, 7, 8, 10, 12].

В заключении перечислены основные результаты, полученные в диссертации.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ, ВЫНОСИМЫЕ НА ЗАЩИТУ

1. Разработаны абстрактные итерационные схемы решения обобщенных уравнений, которые составляют удобный набор средств для тонкого теоретического анализа локальной сходимости и скорости сходимости численных методов решения задач оптимизации и вариационного анализа в различных предположениях.
2. Развита теория локальной сходимости метода модифицированных функций Лагранжа и метода множителей с линеаризованными ограничениями применительно к задачам с липшицевыми производными при различных требованиях регулярности ограничений и при отсутствии таковых.

3. Теория локальной сходимости полугладкого метода последовательного квадратичного программирования дополнена необходимыми и достаточными условиями его прямой сверхлинейной сходимости.

ПУБЛИКАЦИИ АВТОРА ПО ТЕМЕ ДИССЕРТАЦИИ

1. Измаилов А. Ф., Куренной А. С. Частный дифференциал Кларка и другие обобщенные дифференциалы // Теоретические и прикладные задачи нелинейного анализа. — М. : ВЦ РАН, 2010. — С. 77–90.
2. Измаилов А. Ф., Куренной А. С. Частный дифференциал Кларка и проекция полного дифференциала Кларка // XVIII международная научная конференция студентов, аспирантов и молодых ученых «Ломоносов-2011», секция «Вычислительная математика и математическая кибернетика»: Сб. тезисов / МГУ им. М. В. Ломоносова. — М. : МАКС Пресс, 2011. — 11–15 апреля. — С. 38.
3. Измаилов А. Ф., Куренной А. С. Абстрактная ньютоновская схема для нахождения изолированных решений негладких обобщенных уравнений // Тихоновские чтения: Научная конференция / МГУ им. М. В. Ломоносова. — Т. 1. — М. : МАКС Пресс, 2012. — 23–31 октября. — С. 41.
4. Измаилов А. Ф., Куренной А. С. Методы множителей для задач оптимизации с липшицевыми производными // Журнал вычислительной математики и математической физики. — 2012. — Т. 52, № 12. — С. 2140–2148.
5. Измаилов А. Ф., Куренной А. С. Об условиях регулярности для комплементарных задач // XIX международная научная конференция студентов, аспирантов и молодых ученых «Ломоносов-2012», секция «Вычислительная математика и математическая кибернетика»: Сб. тезисов / МГУ им. М. В. Ломоносова. — М. : МАКС Пресс, 2012. — 9–13 апреля. — С. 61–62.

6. Измаилов А. Ф., Куренной А. С., Солодов М. В. Метод Джозефи–Ньютона для полугладких обобщенных уравнений // Ломоносовские чтения: Научная конференция, посвященная 300-летию со дня рождения М. В. Ломоносова / МГУ им. М. В. Ломоносова. — М. : МАКС Пресс, 2011. — С. 24.
7. Izmailov A. F., Kurennoy A. S. Abstract Newtonian frameworks and their applications // SIAM Journal on Optimization. — 2013. — V. 23, № 4. — P. 2369–2396.
8. Izmailov A. F., Kurennoy A. S. Multiplier methods for optimization problems with Lipschitzian derivatives // VII Московская международная конференция по исследованию операций (ORM2013): Труды. — Т. 1. — М. : МАКС Пресс, 2013. — С. 64–66.
9. Izmailov A. F., Kurennoy A. S. On regularity conditions for complementarity problems // Computational Optimization and Applications. — 2013. — DOI: 10.1007/s10589-013-9604-1.
10. Izmailov A. F., Kurennoy A. S., Solodov M. V. The Josephy–Newton method for semismooth generalized equations and semismooth SQP for optimization // Set-Valued and Variational Analysis. — 2013. — V. 21. — P. 17–45.
11. Izmailov A. F., Kurennoy A. S., Solodov M. V. A note on upper Lipschitz stability, error bounds, and critical multipliers for Lipschitz-continuous KKT systems // Mathematical Programming. — 2013. — V. 142. — P. 591–604.
12. Izmailov A. F., Kurennoy A. S., Solodov M. V. Local convergence of the method of multipliers for variational and optimization problems under the sole noncriticality assumption // Optimization Online. — 2013. http://www.optimization-online.org/DB_HTML/2013/08/3999.html