МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМ. М. В. ЛОМОНОСОВА

На правах рукописи

УСКОВ ЕВГЕНИЙ ИВАНОВИЧ

НЬЮТОНОВСКИЕ МЕТОДЫ РЕШЕНИЯ ЗАДАЧ ОПТИМИЗАЦИИ С НЕРЕГУЛЯРНЫМИ ОГРАНИЧЕНИЯМИ

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, **Березнев Валентин Александрович**; кандидат физико-математических наук, **Павлова Наталья Геннадиевна**.
Ведущая организация:	**Институт динамики систем и теории управления Сибирского отделения Российской академии наук**.

Защита диссертации состоится 23 декабря 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. 939-30-10 (для оформления заявки на пропуск).

С диссертацией можно ознакомиться в библиотеке факультета ВМК МГУ. С текстом автореферата можно ознакомиться на официальном сайте ВМК МГУ `http://cs.msu.ru` в разделе «Наука» – «Работа диссертационных советов» – «Д 501.001.44».

Автореферат разослан 17 октября 2014 г.

Ученый секретарь
диссертационного совета
к. т. н., в. н. с. Костенко В. А.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

В работе рассматриваются задачи оптимизации, в решениях которых могут нарушаться так называемые условия регулярности ограничений. Работа посвящена исследованию причин низкой эффективности существующих методов на таких задачах, а также построению эффективных численных алгоритмов их решения.

Актуальность темы. Задачи оптимизации с нерегулярными ограничениями имеют многочисленные приложения. К ним относятся задачи оптимизации с комплементарными ограничениями, которые возникают, например, при решении задач двухуровнего программирования. Другим приложением являются задачи оптимизации с исчезающими ограничениями, которые находят применение в задачах нахождения оптимального дизайна топологий механических конструкций. Кроме того, существуют классы задач большой размерности, ограничения которых имеют тенденцию оказываться нерегулярными или почти нерегулярными в решении. В то же время, задачи оптимизации с нерегулярными ограничениями являются малоизученными. В частности, многие традиционные методы оптимизации обычно демонстрируют невысокую эффективность на таких задачах. Таким образом, задачи оптимизации с нерегулярными ограничениями представляют как теоретический, так и практический интерес.

Целью диссертационного исследования является построение эффективных численных алгоритмов решения задач оптимизации с нерегулярными ограничениями.

Предмет и объект исследования. Объектом исследования в диссертации являются задачи оптимизации с нерегулярными ограничениями. Предметом исследования является анализ причин низкой скорости сходимости существующих методов применительно к задачам этого класса, а также построение конкурентоспособных практических алгоритмов их решения.

Методы исследования. При выполнении диссертационного исследования использовались средства нелинейного анализа, линейной алгебры, теории оптимизации, теории чувствительности для оптимизационных задач, а также современные подходы к численной оптимизации, в том числе и к поиску особых решений.

Научная новизна. В диссертационной работе получены принципиально новые априорные результаты о сходимости двойственных траекторий метода последовательного квадратичного программирования к критическим множителям Лагранжа. Также в работе доказаны новые результаты о глобальной сходимости метода модифицированных функций Лагранжа, и предложено несколько новых подходов к глобализации сходимости стабилизированного метода последовательного квадратичного программирования. Кроме того, разработан принципиально новый метод последовательного квадратичного программирования, стабилизированный вдоль подпространства, а также подход к глобализации его сходимости с помощью точных гладких штрафных функций, и построена теория локальной и глобальной сходимости метода. Наконец, в качестве важного ингредиента этого метода разработана экономичная процедура аппроксимации подпространства вырожденности, которая существенно опережает все известные альтернативы по вычислительной эффективности.

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

Теоретическая и практическая значимость работы. В диссертации получены результаты, доказывающие эффект притяжения двойственных траекторий ньютоновских методов к критическим множителям Лагранжа в важных частных случаях, которые имеют ключевое значение для понимания данного эффекта. Полученные результаты крайне важны, поскольку именно притяжение к критическим множителям является причиной низкой эффективности существующих методов на задачах оптимизации с нерегулярными ограничениями.

В работе получены теоретические и численные результаты, свидетельствующие о хороших свойствах глобальной сходимости метода модифицированных функций Лагранжа на нерегулярных задачах. Также предложен ряд новых подходов к глобализации сходимости стабилизированного метода последовательного квадратичного программирования, и для каждого из них получены теоретические результаты о глобальной сходимости и сверхлинейной скорости сходимости. Кроме того, обнаружено важное с практической точки зрения негативное свойство стабилизированного метода последовательного квадратичного программирования: вдали от решения метод имеет тенденцию генерировать длинные последовательности коротких шагов, существенно замедляющие сходимость.

Также в работе предложен новый метод последовательного квадратичного программирования, стабилизированный вдоль подпространства, и для него доказаны результаты о глобальной сходимости и сверхлинейной скорости сходимости. Вычислительный эксперимент показал, что на нерегулярных задачах данный метод существенно опережает по эффективности метод последовательного квадратичного программирования.

Наконец, в диссертации разработан простой и экономичный способ аппроксимации подпространства вырожденности нелинейного отображения в особом решении соответствующего уравнения, который намного эффективнее всех существующих аналогов с вычислительной точки зрения. Данный способ имеет большое теоретические и практическое значение: он может быть использован в качестве ингредиента различных методов оптимизации (в частности, он используется в методе последовательного квадратичного программирования, стабилизированном вдоль подпространства), а также методов решения нерегулярных задач других классов.

Соответствие диссертации паспорту научной специальности. В диссертации проведено исследование сложных оптимизационных задач и разработаны эффективные численные методы их решения. Это соответствует паспорту 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, 2011 и 2012), на ежегодной научной конференции «Ломоносовские чтения» (Москва, 2011 и 2012), а также на VII Московской международной конференции по исследованию операций «ORM2013».

Публикации. Полученные в диссертации результаты опубликованы в 19 работах, список которых приведен в конце автореферата. В их число входит 5 статей, опубликованных в журналах из списка ВАК РФ [4, 6, 14, 15, 19].

Структура и объем диссертации. Диссертация состоит из введения, четырех глав, трех приложений, заключения и списка литературы из 94 источников. Общий объем диссертации 211 страниц.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во **введении** описана постановка задачи, приведены примеры возникновения задач оптимизации с нерегулярными ограничениями, обоснована актуальность исследования, описана его методика, отражена научная новизна диссертации, сформулированы основные положения, выносимые на защиту, а также приведена информация об апробации результатов.

Первая глава посвящена эффекту притяжения двойственных траекторий ньютоновских методов к так называемым критическим множителям Лагранжа. Данный эффект играет ключевую роль в контексте задач оптимизации с нерегулярными ограничениями, и именно он является причиной медленной сходимости многих традиционных алгоритмов для таких задач.

Раздел 1.1 посвящен обсуждению возможных сценариев двойственного поведения и демонстрации эффекта притяжения к критическим множителям на примерах. В данном разделе рассматривается задача математического программирования

$$f(x) \to \min, \quad h(x) = 0,\ g(x) \leqslant 0, \tag{1}$$

где $f\colon \mathbb{R}^n \to \mathbb{R}$ — гладкая функция, а $h\colon \mathbb{R}^n \to \mathbb{R}^l$ и $g\colon \mathbb{R}^n \to \mathbb{R}^m$ — гладкие отображения. Система Каруша–Куна–Таккера (ККТ) задачи (1) имеет вид

$$\frac{\partial L}{\partial x}(x,\, \lambda,\, \mu) = 0, \quad h(x) = 0, \quad \mu \geqslant 0, \quad g(x) \leqslant 0, \quad \langle \mu,\, g(x)\rangle = 0, \tag{2}$$

где $L\colon \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m \to \mathbb{R}$ — функция Лагранжа задачи (1). Если тройка $(\bar{x},\, \bar{\lambda},\, \bar{\mu}) \in \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m$ удовлетворяет системе (2), то $\bar{x}$ называется стационарной точкой задачи (1), а $(\bar{\lambda},\, \bar{\mu})$ — отвечающим ей множителем Лагранжа. Обозначим через $\mathcal{M}(\bar{x})$ множество всех множителей Лагранжа, отвечающих $\bar{x}$. Введем также множества индексов

$$A(\bar{x}) = \{i = 1,\, \ldots,\, m \mid g_i(\bar{x}) = 0\}, \quad N(\bar{x}) = \{1,\, \ldots,\, m\} \setminus A(\bar{x}),$$
$$A_+(\bar{x},\, \bar{\mu}) = \{i \in A(\bar{x}) \mid \bar{\mu}_i > 0\}, \quad A_0(\bar{x},\, \bar{\mu}) = \{i \in A(\bar{x}) \mid \bar{\mu}_i = 0\}.$$

Множитель Лагранжа $(\bar{\lambda},\, \bar{\mu}) \in \mathcal{M}(\bar{x})$ называется *критическим*, если существует набор $(\xi,\, \eta,\, \zeta) \in \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m$ такой, что $\xi \neq 0$, и выполнены соотношения

$$\frac{\partial^2 L}{\partial x^2}(\bar{x},\, \bar{\lambda},\, \bar{\mu})\xi + (h'(\bar{x}))^{\mathrm{T}}\eta + (g'(\bar{x}))^{\mathrm{T}}\zeta = 0, \quad h'(\bar{x})\xi = 0,$$
$$\zeta_{A_0(\bar{x},\bar{\mu})} \geqslant 0, \quad g'_{A_0(\bar{x},\bar{\mu})}(\bar{x})\xi \leqslant 0, \quad \zeta_i\langle g'_i(\bar{x}),\, \xi\rangle = 0,\ i \in A_0(\bar{x},\, \bar{\mu}),$$
$$g'_{A_+(\bar{x},\bar{\mu})}(\bar{x})\xi = 0, \quad \zeta_{N(\bar{x})} = 0,$$

и *некритическим* иначе. Множество критических множителей обычно является тощим по отношению к множеству всех множителей, однако эти множители обладают рядом особых свойств и оказывают значительное влияние на поведение ряда важнейших численных методов оптимизации. В разделе 1.1 приводится ряд численных примеров, убедительно демонстрирующих, что двойственные траектории метода последовательного квадратичного программирования (SQP) имеют сильнейшую тенденцию притяжения к критическим множителям, причем данный эффект является причиной медленной прямой сходимости. Напомним, что на каждой итерации метода SQP для текущего приближения $(x,\, \lambda,\, \mu) \in \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m$ вычисляется направление $\xi \in \mathbb{R}^n$ как стационарная точка задачи квадратичного программирования

$$\begin{gathered} \langle f'(x),\, \xi\rangle + \frac{1}{2}\left\langle \frac{\partial^2 L}{\partial x^2}(x,\, \lambda,\, \mu)\xi,\, \xi\right\rangle \to \min \\ h(x) + h'(x)\xi = 0, \quad g(x) + g'(x)\xi \leqslant 0, \end{gathered} \tag{3}$$

и следующее приближение полагается равным $(x + \xi,\, y,\, z)$, где $y \in \mathbb{R}^l$ и $z \in \mathbb{R}^m_+$ — множители Лагранжа, отвечающие стационарной точке ξ подзадачи (3).

В разделе 1.2 впервые доказываются априорные результаты о сходимости двойственных траекторий метода SQP к критическим множителям Лагранжа. Обоснование данного эффекта представляет собой трудную теоретическую проблему. В частности, все известные до последнего времени результаты были «негативными»: они показывали, что двойственные траектории не могут сходиться к некритическому множителю, либо такая сходимость маловероятна.

Сначала рассматриваются задачи оптимизации с одной скалярной переменной и одним ограничением-равенством:

$$f(x) \to \min, \quad h(x) = 0,$$

где $f,\, h\colon \mathbb{R} \to \mathbb{R}$ — гладкие функции. Для таких задач доказывается результат о том, что если $\bar{x}$ — стационарная точка, а $\bar{\lambda}$ — отвечающий $\bar{x}$ критический множитель Лагранжа, причем выполнено $h''(\bar{x}) \neq 0$, то для любого x^0, достаточно близкого к $\bar{x}$, и для любого $\bar{\lambda}^0$ существует единственная траектория $\{(x_k,\, \lambda_k)\} \subset \mathbb{R} \times \mathbb{R}$ метода SQP, и эта траектория сходится к $(0,\, \bar{\lambda})$, причем выполнено

$$\lim_{k\to\infty} \frac{x_{k+1} - \bar{x}}{x_k - \bar{x}} = \frac{1}{2},$$

т. е. скорость прямой сходимости является линейной.

Далее эффект притяжения к критическим множителям полностью обосновывается для чисто квадратичных задач, имеющих вид

$$\frac{1}{2}\langle Ax,\, x\rangle \to \min, \quad \frac{1}{2}B[x,\, x] = 0,$$

где $A \in \mathbb{R}^{n\times n}$ — симметричная матрица, а $B\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^l$ — симметричное билинейное отображение. Полученный результат имеет ключевое значение для понимания эффекта притяжения, поскольку чисто квадратичные задачи несут в себе все основные проблемы связанные с нерегулярностью ограничений.

Наконец, в разделе 1.3 рассматриваются существующие методы, обладающие свойством локальной двойственной стабилизации, т. е. способностью подавлять эффект притяжения: стабилизированный метод последовательного квадратичного программирования (sSQP) и метод модифицированных функций Лагранжа (МФЛ). На каждой итерации метода sSQP для текущего приближения $(x,\, \lambda,\, \mu) \in \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m$ вычисляется направление $(\xi,\, \eta,\, \zeta) \in \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m$ как стационарная точка задачи квадратичного программирования

$$\langle f'(x),\, \xi\rangle + \frac{1}{2}\left\langle \frac{\partial^2 L}{\partial x^2}(x,\, \lambda,\, \mu)\xi,\, \xi \right\rangle + \frac{\sigma}{2}(\|\lambda+\eta\|^2 + \|\mu+\zeta\|^2) \to \min$$
$$h(x) + h'(x)\xi - \sigma\eta = 0, \quad g(x) + g'(x)\xi - \sigma\zeta \leqslant 0,$$

где σ — параметр стабилизации, и следующее приближение полагается равным $(x+\xi,\, \lambda+\eta,\, \mu+\zeta)$. В базовом варианте МФЛ на каждой итерации для текущего двойственного приближения $(\lambda^k,\, \mu^k) \in \mathbb{R}^l \times \mathbb{R}^m$ и текущего значения параметра штрафа $c_k > 0$ вычисляется $x^{k+1} \in \mathbb{R}^n$ как стационарная точка задачи безусловной оптимизации

$$L_{c_k}(x,\, \lambda^k,\, \mu^k) \to \min, \quad x \in \mathbb{R}^n,$$

где $L_c\colon \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m \to \mathbb{R}$ — модифицированная функция Лагранжа для задачи (1):

$$L_c(x,\, \lambda,\, \mu) = f(x) + \frac{1}{2c}\left(\|\lambda + ch(x)\|^2 + \|\max\{0,\, \mu + cg(x)\}\|^2\right),$$

после чего двойственное приближение пересчитывается по формулам

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

где максимум берется покомпонентно, и выбирается $c_{k+1} \geqslant c_k$.

Оба метода обладают сильной теорией локальной сходимости: если любой их них инициализируется в достаточно малой окрестности точки $(\bar{x},\, \bar{\lambda},\, \bar{\mu})$, где $\bar{x}$ — стационарная точка задачи (1), а множитель $(\bar{\lambda},\, \bar{\mu}) \in \mathcal{M}(\bar{x})$, удовлетворяет *достаточному условию второго порядка оптимальности* (SOSC)

$$\left\langle \frac{\partial^2 L}{\partial x^2}(\bar{x},\, \bar{\lambda},\, \bar{\mu})\xi,\, \xi \right\rangle > 0 \quad \forall\, \xi \in C(\bar{x}) \setminus \{0\}, \tag{4}$$

где

$$C(\bar{x}) = \{\xi \in \mathbb{R}^n \mid h'(\bar{x})\xi = 0,\ g'_{A(\bar{x})}(\bar{x})\xi \leqslant 0,\ \langle f'(\bar{x}),\, \xi\rangle \leqslant 0\},$$

то траектории метода будут сверхлинейно сходиться к некоторой точке $(\bar{x},\, \lambda^*,\, \mu^*)$, где множитель $(\lambda^*,\, \mu^*) \in \mathcal{M}(\bar{x})$ также удовлетворяет SOSC (для МФЛ сверхлинейная сходимость имеет место только при $c_k \to \infty$). Однако, свойство двойственной стабилизации является лишь локальным, и для обоих методов эффект притяжения к критическим множителям присутствует, хотя и значительно менее выражен. А именно, во многих задачах есть широкие прямо-двойственные области, при старте из которых двойственные траектории методов будут сходиться к критическим множителям.

Вторая глава посвящена исследованию поведения МФЛ для задач оптимизации с нерегулярными ограничениями. В данной главе рассматривается алгоритм, на каждой итерации которого выполняются следующие действия. Выбирается вектор $(\bar{\lambda}^k, \bar{\mu}^k) \in \mathbb{R}^l \times \mathbb{R}^m$ из некоторого фиксированного параллелепипеда, и для текущего значения параметра штрафа c_k вычисляется x^{k+1} как стационарная точка задачи оптимизации

$$L_{c_k}(x,\, \bar{\lambda}^k,\, \bar{\mu}^k) \to \min, \quad x \in \mathbb{R}^n.$$

Новое двойственное приближение вычисляется по формулам $\lambda^{k+1} = \bar{\lambda}^k + c_k h(x^{k+1})$ и $\mu^{k+1} = \max\{0,\, \bar{\mu}^k + c_k g(x^{k+1})\}$. Если

$$\|(h(x^{k+1}),\, \min\{\mu^{k+1},\, -g(x^{k+1})\})\| \leqslant \theta \|(h(x^k),\, \min\{\mu^k,\, -g(x^k)\})\|,$$

где $\theta \in (0,\, 1)$, то выбирается любое значение $c_{k+1} \geqslant c_k$; в противном случае выбирается $c_{k+1} \geqslant \rho c_k$, где $\rho > 1$.

В разделе 2.1 исследуются свойства глобальной сходимости МФЛ. Раздел 2.1.1 посвящен глобальной сходимости МФЛ для задачи (1). Сначала кратко обсуждаются существующие результаты о глобальной сходимости метода, после чего доказывается результат о глобальной сходимости, не требующий традиционных условий регулярности ограничений. Вместо этого делается два предположения, которые не являются стандартными, но при этом весьма обоснованы и должны выполняться во многих случаях. Пусть траектория $\{(x^k,\, \lambda^k,\, \mu^k)\} \subset \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m$ сгенерирована описанным выше алгоритмом, и для некоторого бесконечного множества индексов $K \subset \{1,\, 2,\, \ldots\}$ подпоследовательность $\{x^k \mid k \in K\}$ сходится к $\bar{x} \in D$, где D — допустимое множество задачи (1). Первое предположение состоит в том, что справедлива липшицева оценка расстояния

$$\operatorname{dist}(x^k,\, D) = O(\|h(x^k)\| + \|\max\{0,\, g(x^k)\}\|)$$

для $k \in K$ при $k \to \infty$. Данная оценка слабее многих традиционных условий регулярности ограничений (например, RCPLD или MFCQ). Второе предположение состоит в том, что для всех $k \in K$, некоторого $\beta > 0$, $\theta \in (0,\, 1/2)$ и некоторой проекции $\bar{x}^k$ точки x^k на множество D выполнено

$$L_{c_{k-1}}(x^k,\, \bar{\lambda}^{k-1},\, \bar{\mu}^{k-1}) \leqslant L_{c_{k-1}}(\bar{x}^k,\, \bar{\lambda}^{k-1},\, \bar{\mu}^{k-1}) + \\ + \beta(\|h(x^k)\| + \|\max\{0,\, g(x^k)\}\|) + \frac{\theta\|(\lambda^k,\, \mu^k)\|^2}{c_{k-1}}.$$

Данное требование не выглядит обременительным с учетом того, что приближения x^k ищутся посредством минимизации $L_{c_{k-1}}(\cdot,\, \bar{\lambda}^{k-1},\, \bar{\mu}^{k-1})$. В частности, оно выполнено автоматически, если x^k является глобальным решением подзадачи в некоторой окрестности предельной точки $\bar{x}$. Если выполнены оба предположения, то последовательность $\{(\lambda^k,\, \mu^k) \mid k \in K\}$ ограничена, $\bar{x}$ является стационарной точкой задачи (1), и любая предельная точка подпоследовательности $\{(\lambda^k,\, \mu^k) \mid k \in K\}$ является множителем Лагранжа, отвечающим $\bar{x}$.

В разделе 2.1.2 свойства глобальной сходимости МФЛ исследуются для *задач оптимизации с комплементарными ограничениями* (MPCC) следующего вида:

$$f(x) \to \min, \quad G(x) \geqslant 0,\ H(x) \geqslant 0,\ \langle G(x),\, H(x)\rangle \leqslant 0. \tag{5}$$

В MPCC могут также присутствовать «обычные» ограничения-равенства и ограничения-неравенства, однако все сложности проводимого анализа связаны именно с комплементарными ограничениями. Легко показать, что все полученные результаты распространяются и на общий случай.

Определим так называемую *MPCC-функцию Лагранжа* $\mathcal{L}\colon \mathbb{R}^n \times (\mathbb{R}^s \times \mathbb{R}^s) \to \mathbb{R}$:

$$\mathcal{L}(x,\, \lambda) = f(x) - \langle \lambda_G,\, G(x)\rangle - \langle \lambda_H,\, H(x)\rangle,$$

где $\lambda = (\lambda_G,\, \lambda_H)$. Для произвольной точки $\bar{x} \in \mathbb{R}^n$, допустимой в задаче (5), определим множества индексов

$$I_G(\bar{x}) = \{i = 1,\, \ldots,\, s \mid G_i(\bar{x}) = 0\}, \quad I_H(\bar{x}) = \{i = 1,\, \ldots,\, s \mid H_i(\bar{x}) = 0\},$$
$$I_0(\bar{x}) = I_G(\bar{x}) \cap I_H(\bar{x}).$$

MPCC-условие линейной независимости (MPCC-LICQ) состоит в том, что векторы

$$G_i'(\bar{x}),\ i \in I_G(\bar{x}), \quad H_i'(\bar{x}),\ i \in I_H(\bar{x})$$

линейно независимы. Допустимая точка $\bar{x}$ задачи (5) называется *слабо стационарной*, если существует вектор $\bar{\lambda} = (\bar{\lambda}_G,\, \bar{\lambda}_H) \in \mathbb{R}^s \times \mathbb{R}^s$, удовлетворяющий

$$\frac{\partial \mathcal{L}}{\partial x}(\bar{x},\, \bar{\lambda}) = 0, \quad (\bar{\lambda}_G)_{I_H(\bar{x}) \setminus I_G(\bar{x})} = 0,\ (\bar{\lambda}_H)_{I_G(\bar{x}) \setminus I_H(\bar{x})} = 0.$$

Если дополнительно $(\bar{\lambda}_G)_i(\bar{\lambda}_H)_i \geqslant 0$ для всех $i \in I_0(\bar{x})$, то $\bar{x}$ называется *C-стационарной* точкой. Если же $(\bar{\lambda}_G)_{I_0(\bar{x})} \geqslant 0$, $(\bar{\lambda}_H)_{I_0(\bar{x})} \geqslant 0$, то $\bar{x}$ называется *сильно стационарной* точкой.

Для задачи (5) доказывается результат о том, что допустимые предельные точки МФЛ являются по крайней мере C-стационарными при условии, что в них выполнено MPCC-LICQ. Более того, они являются сильно стационарными, если определенная двойственная последовательность ограничена. В совокупности с примерами, показывающими, что в общем случае сильная стационарность может не иметь места, доказанный результат дает полную картину глобальной сходимости для МФЛ, применяемого к MPCC.

В разделе 2.2 исследуется возможность ускорения финальной фазы МФЛ с помощью sSQP. Основной результат состоит в том, что если ускоритель инициализируется в окрестности точки $(\bar{x}, \bar{\lambda}, \bar{\mu})$, где $\bar{x}$ — стационарная точка задачи (1), а $(\bar{\lambda}, \bar{\mu})$ — некритический множитель Лагранжа, удовлетворяющий условию строгой дополнительности $\bar{\mu}_{A(\bar{x})} > 0$, то траектория ускорителя будет сверхлинейно сходиться к некоторой точке $(\bar{x}, \lambda^*, \mu^*)$, где $(\lambda^*, \mu^*) \in \mathcal{M}(\bar{x})$.

К сожалению, несмотря на неплохие теоретические свойства, данная конструкция ускорителя не позволяет добиться существенного повышения эффективности. Основная причина состоит в притяжении двойственных траекторий МФЛ к критическим множителям. Из-за данного эффекта ускоритель запускается из точек, двойственная компонента которых близка к некоторому критическому множителю, и в таких случаях sSQP обычно все равно сходится к критическому множителю.

В приложении А приводятся результаты вычислительного эксперимента, показывающие, что МФЛ обладает хорошими свойствами глобальной сходимости, однако несколько уступает по эффективности другим алгоритмам; ускорители финальной фазы не позволяют существенно повысить эффективность из-за притяжения к критическим множителям.

Третья глава посвящена глобализации сходимости метода sSQP. Построение глобально сходящихся алгоритмов, которые вблизи решения превращаются в sSQP, является давно стоящей проблемой. В частности, на сегодняшний день отсутствуют алгоритмы такого рода, способные конкурировать с обычным SQP даже на задачах с нерегулярными ограничениями.

В разделе 3.1 исследуются гибридные подходы к глобализации сходимости на основе некоторого метода внешней фазы. Для задачи (1) предлагаются два способа гибридной глобализации сходимости sSQP: алгоритм с возвратами и алгоритм с рекордами.

Идея алгоритма с возвратами заключается в том, что на каждой итерации из текущей точки делается шаг sSQP, который принимается, если он приводит к линейному убыванию невязки системы ККТ. Если на какой-то из последующих итераций шаг sSQP не принимается, то происходит возврат к той точке, откуда был сделан первый в данной серии шаг sSQP, и из этой точки делается шаг метода внешней фазы.

Стратегия гибридной глобализации сходимости с рекордами состоит в следующем. Вместо того, чтобы сравнивать невязку системы ККТ в пробной точке с невязкой, полученной на предыдущей итерации, будем сравнивать ее с наименьшим достигнутым значением невязки по всем предыдущим итерациям (т. е. с рекордом), и принимать шаг sSQP только в том случае, если наблюдается линейное убывание рекорда.

Для обоих алгоритмов доказываются теоретические результаты о глобальной сходимости и о сверхлинейной скорости сходимости.

В приложении Б.1 приводятся результаты вычислительного эксперимента, в котором sSQP, глобализованный с использованием предложенных подходов, сравнивается с обычным квазиньютоновским SQP. К сожалению, предложенные способы гибридной глобализации сходимости, будучи полностью теоретически обоснованными, на практике не сохраняют локальный выигрыш от стабилизации, если начальное приближение недостаточно близкó к решению. Основной причиной этого является эффект притяжения ньютоновских методов к критическим множителям Лагранжа.

В разделе 3.2 для задачи (1) предлагается алгоритм, комбинирующий свойства sSQP и МФЛ. Основная идея состоит в том, что направления sSQP можно использовать для приближенного решения подзадач МФЛ. На каждой итерации алгоритма сначала вычисляется направление (ξ^k, η^k, ζ^k) метода sSQP. Если шаг sSQP приводит к существенному улучшению невязки системы ККТ по сравнению с рекордным значением, то этот шаг принимается (осуществляется итерация sSQP). В противном случае осуществляется одномерный поиск вдоль направления ξ^k для модифицированной функции Лагранжа при текущих значениях двойственных переменных. Если в результате одномерного поиска получен приближенный минимум этой функции, то осуществляется итерация МФЛ. В противном случае осуществляется внутренняя итерация (двойственное приближение и параметр штрафа не меняются).

Для данного алгоритма доказывается следующий результат о глобальной сходимости: любая предельная точка ($\bar{x}$, $\bar{\lambda}$, $\bar{\mu}$) траектории алгоритма либо является решением системы ККТ (2), либо удовлетворяет соотношениям

$$\frac{\partial L}{\partial x}(\bar{x},\, \bar{\lambda},\, \bar{\mu}) = 0, \quad \bar{\mu} \geqslant 0, \quad (h'(\bar{x}))^{\mathrm{T}} h(\bar{x}) + (g'(\bar{x}))^{\mathrm{T}} \max\{0,\, g(\bar{x})\} = 0,$$

причем в последнем случае все итерации, начиная с некоторого номера, являются итерациями МФЛ. Также доказывается результат о сверхлинейной скорости сходимости. В частности, если траектория алгоритма сходится к точке, удовлетворяющей SOSC (4), то в естественных предположениях скорость сходимости будет сверхлинейной.

В разделе 3.2.2 показано, что данный алгоритм тесно связан с прямо-двойственным алгоритмом SQP, но в то же время обладает рядом преимуществ.

В приложении Б.2 приводятся результаты вычислительного эксперимента, которые показывают, что предложенный алгоритм оказывается более эффективным, чем некоторые другие методы (включая прямо-двойственный алгоритм SQP), однако проигрывает по эффективности обычному SQP.

В разделе 3.3 для задачи оптимизации с ограничениями-равенствами

$$f(x) \to \min, \quad h(x) = 0, \tag{6}$$

где $f\colon \mathbb{R}^n \to \mathbb{R}$ — гладкая функция, а $h\colon \mathbb{R}^n \to \mathbb{R}^l$ — гладкое отображение, разрабатывается подход к глобализации сходимости с помощью точных гладких штрафных функций. Основное отличие данного подхода от других способов глобализации сходимости, предложенных в главе 3, состоит в том, что он не является гибридным. А именно, на каждой итерации вычисляется прямо-двойственное направление поиска sSQP, вдоль которого затем осуществляется одномерный поиск для точной гладкой штрафной функции $\varphi_{c_1,\,c_2}\colon \mathbb{R}^n \times \mathbb{R}^l \to \mathbb{R}$:

$$\varphi_{c_1,\,c_2}(x,\,\lambda) = L(x,\,\lambda) + \frac{c_1}{2}\|h(x)\|^2 + \frac{c_2}{2}\left\|\frac{\partial L}{\partial x}(x,\,\lambda)\right\|^2,$$

где $c_1 > 0$ и $c_2 > 0$ — параметры штрафа.

В разделе 3.3.1 приводится глобализованный алгоритм, и для него доказывается следующий результат о глобальной сходимости. Если последовательность параметров штрафа ограничена, то любая предельная точка $(\bar{x},\,\bar{\lambda})$ траектории алгоритма $\{(x^k,\,\lambda^k)\}$ удовлетворяет $\varphi'_{c_1,\,c_2}(\bar{x},\,\bar{\lambda}) = 0$, откуда в естественных предположениях следует, что $(\bar{x},\,\bar{\lambda})$ является решением системы Лагранжа задачи (6). Если же любой из параметров увеличивается бесконечное число раз, то любая предельная точка соответствующей подпоследовательности траектории $\{(x^k,\,\lambda^k)\}$ в естественных предположениях является решением системы Лагранжа задачи (6).

В разделе 3.3.2 доказывается результат о сверхлинейной скорости сходимости: если траектория алгоритма сходится к точке $(\bar{x},\,\bar{\lambda})$, где $\bar{x}$ — стационарная точка задачи (6), а $\bar{\lambda}$ — отвечающий ей некритический множитель Лагранжа, то скорость сходимости является сверхлинейной.

В приложении Б.3 приводятся результаты вычислительного эксперимента, которые показывают, что предлагаемый алгоритм имеет высокую эффективность, и, в частности, существенно опережает обычный SQP на задачах с полным вырождением. Проведенное исследование также показало, что метод sSQP обладает серьезным недостатком: для широкого класса задач он имеет тенденцию генерировать вдали от решения длинные последовательности коротких шагов, существенно замедляющие сходимость.

В **четвертой главе** для задачи (6) разрабатывается метод последовательного квадратичного программирования, стабилизированный вдоль подпространства (s-sSQP). Он представляет собой модификацию метода sSQP, в которой двойственная траектория стабилизируется только вдоль подпространства вырожденности матрицы Якоби ограничений. Целью модификации является устранение описанной выше проблемы коротких шагов. Пусть заданы отображение $P\colon \mathbb{R}^n \times \mathbb{R}^l \to \mathbb{R}^{l\times l}$ и функция $\sigma\colon \mathbb{R}^n \times \mathbb{R}^l \to \mathbb{R}$, где линейный оператор $P(x,\, \lambda)$ аппроксимирует проектор на подпространство вырожденности при $(x,\, \lambda) \to (\bar{x},\, \bar{\lambda})$. На каждой итерации s-sSQP для текущего приближения $(x,\, \lambda) \in \mathbb{R}^n \times \mathbb{R}^l$ вычисляется направление поиска $(\xi,\, \eta) \in \mathbb{R}^n \times \mathbb{R}^l$ как решение линейной системы

$$\begin{gathered}\frac{\partial^2 L}{\partial x^2}(x,\, \lambda)\xi + (h'(x))^{\mathrm{T}}\eta = -\frac{\partial L}{\partial x}(x,\, \lambda),\\ h'(x)\xi - \sigma(x,\, \lambda)P(x,\, \lambda)\eta = -h(x),\end{gathered}$$

и следующее приближение определяется как $(x+\xi,\, \lambda+\eta)$.

Раздел 4.1 посвящен обоснованию локальной сверхлинейной сходимости метода вблизи квалифицированного решения. Отдельно рассматриваются два различных набора предположений на P и σ. Ключевое различие состоит в том, что в первом случае $\sigma(x,\, \lambda) \to 0$ при $(x,\, \lambda) \to (\bar{x},\, \bar{\lambda})$ (асимптотически исчезающая стабилизация), а во втором случае $\sigma(x,\, \lambda) \to \bar{\sigma} \neq 0$ при $(x,\, \lambda) \to (\bar{x},\, \bar{\lambda})$ (асимптотически неисчезающая стабилизация). Для каждого набора предположений доказывается теорема о локальной сверхлинейной сходимости, полностью аналогичная соответствующей теореме для sSQP.

В разделе 4.2 предлагается простая с вычислительной точки зрения процедура идентификации подпространства вырожденности, необходимая для достижения высокой общей эффективности метода. Данная процедура предназначена для реализации метода s-sSQP, но может быть использована в качестве ингредиента других методов, а также методов решения нерегулярных задач других классов.

Раздел 4.3 посвящен глобализации сходимости метода с помощью точных гладких штрафных функций. В данном разделе приводится описание глобализованного алгоритма, и для каждого набора предположений доказываются результаты о глобальной сходимости и сверхлинейной скорости сходимости, аналогичные соответствующим результатам, полученным в разделе 3.3.

В приложении B приводятся результаты вычислительного эксперимента, свидетельствующие о том, что стабилизация вдоль нужного подпространства крайне важна для высокой эффективности методов с двойственной стабилизацией: метод s-sSQP существенно опережает по эффективности и робастности как обычный, так и стабилизированный SQP. Предложенная процедура идентификации подпространства вырожденности весьма эффективна с вычислительной точки зрения (в частности, она значительно экономичнее известных альтернатив), и при этом демонстрирует высокую надежность.

В **заключении** сформулированы основные результаты и выводы, полученные в диссертации.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ, ВЫНОСИМЫЕ НА ЗАЩИТУ

1. Доказан эффект притяжения двойственных траекторий метода последовательного квадратичного программирования к критическим множителям в ряде важных частных случаев. Полученные результаты имеют ключевое значение для понимания данного эффекта.

2. Полученные теоретические и численные результаты демонстрируют, что метод модифицированных функций Лагранжа обладает хорошими свойствами глобальной сходимости даже при нарушении условий регулярности ограничений, и поэтому является хорошим выбором для задач оптимизации с нерегулярными ограничениями, если приоритет имеет высокая робастность и качество получаемых решений.

3. Предложен подход к глобализации сходимости стабилизированного метода последовательного квадратичного программирования с использованием точных гладких штрафных функций. Глобализованный таким способом метод существенно опережает по эффективности обычный метод последовательного квадратичного программирования на некоторых классах задач.

4. Разработан новый метод последовательного квадратичного программирования, стабилизированный вдоль подпространства, и получены результаты о его локальной и глобальной сходимости. Разработанный метод существенно опережает по эффективности обычный метод последовательного квадратичного программирования на задачах с нерегулярными ограничениями.

5. Разработана процедура аппроксимации подпространства вырожденности отображения в особом решении, которая, в отличие от существующих альтернатив, целесообразна с практической точки зрения. Данная процедура может быть использована для построения самых разных методов, использующих информацию о структуре подпространства вырожденности.

ПУБЛИКАЦИИ АВТОРА ПО ТЕМЕ ДИССЕРТАЦИИ

1. Измаилов А.Ф., Крылова А.М., Усков Е.И. Гибридная глобализация стабилизированного метода последовательного квадратичного программирования // Теоретические и прикладные задачи нелинейного анализа. — М.: ВЦ РАН, 2011. — С. 47–66.

2. Измаилов А.Ф., Солодов М.В., Усков Е.И. Метод модифицированных функций Лагранжа для вырожденных задач оптимизации // Ломоносовские чтения: Научная конференция, посвященная 300-летию со дня рождения М.В. Ломоносова: Москва, факультет ВМК МГУ име-

ни М.В. Ломоносова, 14–23 ноября 2011 г.: Тезисы докладов. — М.: МАКС Пресс, 2011. — С. 23–24.

3. Измаилов А.Ф., Солодов М.В., Усков Е.И. Метод множителей как средство глобализации сходимости стабилизированного метода последовательного квадратичного программирования для задач оптимизации с ограничениями-равенствами // Теоретические и прикладные задачи нелинейного анализа. — М.: ВЦ РАН, 2014. — С. 46–64.

4. Измаилов А.Ф., Усков Е.И. О применении ньютоновских методов к системе условий оптимальности Ф. Джона // Журнал вычислительной математики и математической физики. — 2011. — Т. 51, № 7. — С. 1194–1208.

5. Измаилов А.Ф., Усков Е.И. О применении ньютоновских методов к системе условий оптимальности Ф. Джона // Тихоновские чтения: Научная конференция, Москва, МГУ имени М.В. Ломоносова, 14 июня 2011 г.: Тезисы докладов. — М.: МАКС Пресс, 2011. — С. 40.

6. Измаилов А.Ф., Усков Е.И. О влиянии критических множителей Лагранжа на скорость сходимости метода множителей // Журнал вычислительной математики и математической физики. — 2012. — Т. 52, № 11. — С. 1959–1975.

7. Измаилов А.Ф., Усков Е.И. О притяжении метода Ньютона к критическим множителям Лагранжа // Тихоновские чтения: Научная конференция, Москва, МГУ имени М.В. Ломоносова, 29–31 октября 2012 г.: Тезисы докладов. — М.: МАКС Пресс, 2012. — С. 41–42.

8. Измаилов А.Ф., Усков Е.И. Эффект притяжения метода Ньютона–Лагранжа к критическим множителям Лагранжа: полный анализ в одномерном случае // Теоретические и прикладные задачи нелинейного анализа. — М.: ВЦ РАН, 2012. — С. 53–71.

9. Измаилов А.Ф., Усков Е.И. Стабилизированный метод Ньютона–Лагранжа для минимизации модифицированной функции Лагранжа // Теоретические и прикладные задачи нелинейного анализа. — М.: ВЦ РАН, 2013. — С. 39–54.

10. Измаилов А.Ф., Усков Е.И. Эффективная численная аппроксимация подпространства вырожденности нелинейного отображения // Теоретические и прикладные задачи нелинейного анализа. — М.: ВЦ РАН, 2014. — С. 74–87.

11. Усков Е.И. О применении ньютоновских методов к системе условий оптимальности Ф. Джона // «Ломоносов-2011»: XVIII Международная научная конференция студентов, аспирантов и молодых ученых; секция «Вычислительная математика и кибернетика»: Москва, МГУ имени М.В. Ломоносова, 11–15 апреля 2011 г.: Сб. тезисов. — М.: МАКС Пресс, 2011. — С. 50.

12. Усков Е.И. Метод модифицированных функций Лагранжа для вырожденных задач оптимизации // «Ломоносов-2012»: XIX Международная научная конференция студентов, аспирантов и молодых ученых; секция «Вычислительная математика и кибернетика»: Москва, МГУ имени М.В. Ломоносова, 9–13 апреля 2012 г.: Сб. тезисов. — М.: МАКС Пресс, 2012. — С. 69–70.

13. Усков Е.И. Численное сравнение оптимизационных алгоритмов // Теоретические и прикладные задачи нелинейного анализа. — М.: ВЦ РАН, 2012. — С. 118–131.

14. Усков Е.И. О притяжении метода Ньютона к критическим множителям Лагранжа // Журнал вычислительной математики и математической физики. — 2013. — Т. 53, № 8. — С. 1272–1286.

15. Izmailov A.F., Solodov M.V., Uskov E.I. Global convergence of augmented Lagrangian methods applied to optimization problems with

degenerate constraints, including problems with complementarity constraints // SIAM Journal on Optimization. — 2012. — Vol. 22, no. 4. — P. 1579–1606.

16. Izmailov A.F., Solodov M.V., Uskov E.I. Combining stabilized SQP with the augmented Lagrangian algorithm // IMPA Preprint A754. — 2014.

17. Izmailov A.F., Solodov M.V., Uskov E.I. Globalizing stabilized SQP by smooth primal-dual exact penalty function // IMPA Preprint A752. — 2014.

18. Izmailov A.F., Uskov E.I. Attraction of Newton method to critical Lagrange multipliers // VII Московская международная конференция по исследованию операций (ORM2013): Москва, 15–19 октября 2013 г.: Труды. — Т. 1. — М.: МАКС Пресс, 2013. — С. 67–69.

19. Izmailov A.F., Uskov E.I. Attraction of Newton method to critical Lagrange multipliers: fully quadratic case // Mathematical Programming. — 2014. — DOI: 10.1007/s10107-014-0777-x.