

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМ. М. В. ЛОМОНОСОВА
ФАКУЛЬТЕТ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И КИБЕРНЕТИКИ
КАФЕДРА ИССЛЕДОВАНИЯ ОПЕРАЦИЙ

На правах рукописи



Усков Евгений Иванович

НЬЮТОНОВСКИЕ МЕТОДЫ РЕШЕНИЯ ЗАДАЧ ОПТИМИЗАЦИИ С НЕРЕГУЛЯРНЫМИ ОГРАНИЧЕНИЯМИ

Специальность 01.01.09 — дискретная математика и математическая кибернетика

ДИССЕРТАЦИЯ НА СОИСКАНИЕ УЧЁНОЙ СТЕПЕНИ
КАНДИДАТА ФИЗИКО-МАТЕМАТИЧЕСКИХ НАУК

Научный руководитель:
профессор, д.ф.-м.н.
Измаилов Алексей Феридович

Москва

2014

Оглавление

Введение	4
Список основных обозначений	12
Глава 1. Эффект притяжения к критическим множителям Лагранжа	13
1.1. Критические множители Лагранжа	14
1.2. Обоснование эффекта притяжения для метода Ньютона–Лагранжа	26
1.2.1. Одномерные задачи	27
1.2.2. Чисто квадратичные задачи	36
1.3. Локальная двойственная стабилизация	62
1.3.1. Стабилизированный метод последовательного квадратичного программирования	62
1.3.2. Метод модифицированных функций Лагранжа	66
Глава 2. Метод модифицированных функций Лагранжа	70
2.1. Результаты о глобальной сходимости	70
2.1.1. Общая теория глобальной сходимости	71
2.1.2. Глобальная сходимость для задач оптимизации с комплементарными ограничениями	79
2.2. Ускорение финальной фазы	92
Глава 3. Глобализация сходимости стабилизированного метода последовательного квадратичного программирования	97
3.1. Гибридные подходы к глобализации сходимости	98
3.2. Глобализация сходимости с помощью модифицированных функций Лагранжа	103
3.2.1. Алгоритм и его теоретические свойства	104
3.2.2. Связь с прямо-двойственным алгоритмом последовательного квадратичного программирования	115
3.3. Глобализация сходимости с помощью точных гладких штрафных функций . .	119
3.3.1. Глобализованный алгоритм	120

3.3.2. Анализ скорости сходимости	126
Глава 4. Метод последовательного квадратичного программирования, стабилизированный вдоль подпространства	139
4.1. Описание метода и результаты о локальной сходимости	140
4.1.1. Асимптотически исчезающая стабилизация	141
4.1.2. Неисчезающая стабилизация	147
4.2. Аппроксимация подпространства вырожденности	149
4.3. Глобализованный алгоритм	156
Приложение А. Численные результаты для метода модифицированных функций Лагранжа	162
A.1. Сравнение с другими алгоритмами	164
A.2. Эксперимент с правилами для параметра штрафа: критические множители и скорость сходимости	170
A.3. Эксперимент с правилами для параметра штрафа: общая эффективность . . .	172
A.4. Эксперимент с ускорителями	175
Приложение Б. Численные результаты для стабилизированного метода последовательного квадратичного программирования	180
B.1. Гибридная глобализация сходимости	180
B.2. Глобализация сходимости с помощью модифицированных функций Лагранжа	184
B.3. Глобализация сходимости с помощью точных гладких штрафных функций . .	188
Приложение В. Численные результаты для метода последовательного квадратичного программирования, стабилизированного вдоль подпространства	195
Заключение	202
Список литературы	204

Введение

В существующей литературе по теории и численным методам оптимизации основное внимание традиционно уделяется задачам, в решениях которых выполнены так называемые условия регулярности ограничений. Такие задачи достаточно подробно изучены, в частности, разработаны эффективные численные методы отыскания их решений, для которых обоснована глобальная сходимость и сверхлинейная скорость сходимости.

Вместе с тем, большой интерес представляют задачи, в решениях которых условия регулярности ограничений могут нарушаться, и их изучению посвящена данная работа. С такими задачами связан ряд серьезных теоретических трудностей (например, не гарантируется устойчивость решения по отношению к малым возмущениям входных данных). Более того, численное отыскание решений в таких задачах весьма проблематично и требует специальных подходов. В частности, многие традиционные методы оптимизации обычно демонстрируют невысокую эффективность на таких задачах.

Недавно было показано, что основной причиной низкой эффективности ньютоновских методов на нерегулярных задачах является эффект притяжения двойственных траекторий к так называемым критическим множителям Лагранжа. Эти множители обычно образуют тощее подмножество множества всех множителей Лагранжа, однако именно к ним обычно сходятся двойственные траектории ньютоновских методов, причем скорость сходимости прямой траектории в таких случаях оказывается в лучшем случае линейной. На сегодняшний день описанный эффект изучен весьма мало, и поэтому ему будет уделено значительное внимание в данной работе.

Отметим, что в недавнем прошлом активно разрабатывались эффективные и устойчивые методы поиска особых решений общих систем нелинейных уравнений (см. [14], а также более поздние работы [3, 5–7]). К сожалению, эффективное использование таких методов для решения нерегулярных задач оптимизации (например, путем применения их к системе Лагранжа или системе Ф. Джона) возможно только в очень обременительных предположениях, в связи с чем такой подход имеет весьма ограниченную область применения [15].

Еще одна причина интереса к нерегулярным задачам, помимо связанных с ними теоретических и практических трудностей, заключается в наличии важных практических при-

ложений. Одним из таких приложений являются *задачи оптимизации с распадающимися ограничениями*. Такое название связано с тем, что допустимое множество задач этого класса состоит из частей (кусков или ветвей), каждая из которых задается «обычными» ограничениями. В большинстве случаев такие задачи являются нерегулярными, несмотря на то, что ограничения, задающие каждую ветвь, обычно являются регулярными.

Среди задач оптимизации с распадающимися ограничениями наиболее важны *задачи оптимизации с комплементарными ограничениями* (МРСС, от англ. Mathematical program with complementarity constraints). Задачи этого класса в общей форме имеют следующий вид:

$$\begin{aligned} f(x) \rightarrow \min, \quad h(x) = 0, \quad g(x) \leq 0, \\ G(x) \geq 0, \quad H(x) \geq 0, \quad \langle G(x), H(x) \rangle \leq 0. \end{aligned}$$

Одним из приложений МРСС являются так называемые *двухуровневые задачи оптимизации* или *задачи двухуровневого программирования*, частный случай которых был впервые рассмотрен Штакельбергом в 1934 г. при исследовании иерархических рыночных структур. В данный класс входят задачи оптимизации, в которых одно из ограничений описывается с помощью другой оптимизационной задачи, называемой задачей нижнего уровня. Таким образом, задачи данного класса имеют двухуровневую иерархическую структуру. Один из способов сведения таких задач к одноуровневым задачам оптимизации заключается в замене задачи нижнего уровня на ее систему Каруша–Куна–Таккера (ККТ). В этом случае получим МРСС, которая в некоторых предположениях будет эквивалентна исходной задаче. Более подробную информацию о двухуровневых задачах можно найти в [43, 82].

Другим важным приложением МРСС являются задачи нахождения оптимальной формы упругопластичных балочных (стержневых) конструкций [49], относящиеся к так называемой оптимизации форм. В качестве примера рассмотрим следующую задачу. Необходимо спроектировать стержневую структуру таким образом, чтобы она была устойчивой к прилагаемым силам, и чтобы смещения и пластические деформации (если имеются) были в допустимых пределах. При этом геометрическая форма конструкции (т. е. взаимное расположение стержней) задана, а переменными являются площади поперечных сечений стержней. Чем больше площади поперечных сечений, тем прочнее конструкция, но тем больше ее размеры и масса. Возможная постановка задачи звучит следующим образом: необходимо минимизировать некоторую целевую функцию (отражающую размеры или массу) при условии, что конструкция будет удовлетворять указанным выше ограничениям.

При формализации данной задачи оптимизации могут естественным образом возникать комплементарные ограничения. Пластические деформации в таких моделях можно описывать в терминах так называемых пластических множителей $z \in \mathbb{R}^d$ и значений функции текучести $\omega \in \mathbb{R}^d$, где d — количество функций текучести (см., например, [22, 23]). Если

наблюдается текучесть (т. е. $\omega_i = 0$ для некоторого i), то может произойти пластическая деформация (т. е. $z_i \geq 0$). Если же выполнено условие упругости (т. е. $\omega_i > 0$), то пластические деформации невозможны (т. е. $z_i = 0$). Таким образом, описанные ограничения можно записать в виде следующего условия комплементарности:

$$\omega \geq 0, \quad z \geq 0, \quad \langle \omega, z \rangle = 0.$$

Более подробное исследование задач данного вида можно найти в работе [49].

Существует множество других приложений МРСС. Некоторые из них представлены в [82, 89].

Другой важный подкласс задач оптимизации с распадающимися ограничениями образуют *задачи оптимизации с исчезающими ограничениями* (MPVC, от англ. Mathematical program with vanishing constraints). Задачи этого класса имеют следующий вид:

$$\begin{aligned} f(x) \rightarrow \min, \quad h(x) = 0, \quad g(x) \leq 0, \\ H_i(x) \geq 0, \quad G_i(x)H_i(x) \leq 0, \quad i = 1, \dots, s. \end{aligned}$$

Наиболее известным применением MPVC являются задачи нахождения оптимального дизайна топологий механических конструкций [28, 29]. Целью такого дизайна является нахождение оптимальной геометрической формы структуры, которая включает в себя как взаимное расположение стержней, так и площади их поперечных сечений. Таким образом, задачи данного класса существенно отличаются от задач оптимизации форм, в которых взаимное расположение элементов предполагается заданным, а оптимизируются только площади поперечных сечений стержней.

Рассмотрим пример нахождения оптимального дизайна жесткой (упругой) конструкции с использованием так называемого подхода «базовой конструкции» (см. [29]). Введем множество M потенциальных стержней, причем для каждого потенциального стержня параметры материалов будем предполагать заданными. Для обозначения площади поперечного сечения стержня с номером i введем переменную $a_i \geq 0$. Положительное значение a_i означает существование стержня, а значение $a_i = 0$ — отсутствие. Как и в задачах оптимизации форм, обычно оптимизируется вес или размеры конструкции, а ограничения состоят в том, что структура должна быть устойчива к заданному внешнему воздействию, а деформации должны укладываться в допустимые пределы.

Одним из возможных и естественных способов формализации данной задачи является MPVC. Обозначим через d количество незакрепленных узлов стержневой структуры и введем вектор $u \in \mathbb{R}^d$ узловых смещений структуры под действием приложенной силы. Ограничения на напряжения и на продольный изгиб потенциального стержня с номером i можно представить в виде $f_i(a, u) \leq 0$, где f_i — некоторая функция. Очевидно, такие ограничения

имеют смысл только в случае, если $a_i > 0$. Если же $a_i = 0$, то такие ограничения должны исчезать из задачи. Если функции, задающие эти ограничения, могут быть непрерывно продолжены на случай $a_i = 0$ (в этом случае они могут не иметь физического смысла), то ограничения для стержня с номером i можно представить в виде

$$a_i \geq 0, \quad a_i f_i(a, u) \leq 0.$$

В результате получаем задачу оптимизации с исчезающими ограничениями.

Существует другой подход к нахождению оптимального дизайна топологий, который также приводит к MPVC (см. [28]). Разобьем область Ω , в которой должна располагаться структура, на конечное число частей. В каждую из получившихся частей можно распределить некоторое количество материала $\rho_i \geq 0$. Чем больше материала будет выделено для некоторой области, тем прочнее будет конструкция в этом месте. Однако, выделение дополнительного материала будет приводить к утяжелению конструкции. Так же, как и в предыдущем подходе, ограничения для каждой области имеют смысл только если $\rho_i > 0$, в противном случае они должны исчезать. Таким образом, вновь приходим к MPVC. Более подробно данные подходы рассмотрены в работах [28, 29].

Помимо задач оптимизации с распадающимися ограничениями существуют и другие примеры нерегулярных задач. Например, существуют классы задач большой размерности, ограничения которых имеют тенденцию оказываться нерегулярными или «почти» нерегулярными в решении.

Итак, **объектом исследования** в диссертационной работе являются задачи оптимизации с нерегулярными ограничениями.

Основной целью исследования является построение эффективных численных алгоритмов решения задач оптимизации с нерегулярными ограничениями.

Актуальность диссертационной работы в первую очередь обусловлена наличием важных приложений задач оптимизации с нерегулярными ограничениями, некоторые из которых были приведены выше. Кроме того, построение эффективных методов решения таких задач связано с существенными теоретическими трудностями. Поэтому такие задачи представляют как теоретический, так и практический интерес.

Методика исследования включает в себя 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.

Структура диссертации. Работа состоит из введения, четырех глав, трех приложений, заключения и списка литературы из 94 источников.

Первая глава посвящена эффекту притяжения двойственных траекторий ньютоновских методов к так называемым критическим множителям Лагранжа. Данный эффект игра-

ет ключевую роль в контексте задач оптимизации с нерегулярными ограничениями, и именно он является причиной медленной сходимости многих традиционных алгоритмов для таких задач. 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. В разделе 1.1 обсуждается понятие критического множителя Лагранжа и приводятся примеры влияния таких множителей на поведение метода последовательного квадратичного программирования. В разделе 1.2 впервые доказываются априорные результаты о сходимости двойственных траекторий метода последовательного квадратичного программирования к критическим множителям Лагранжа. Доказательство проводится для двух важных частных случаев, один из которых имеет ключевое значение для понимания эффекта притяжения, поскольку он отражает основные последствия нерегулярности 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 и метод модифицированных функций Лагранжа.

Вторая глава посвящена исследованию метода модифицированных функций Лагранжа в контексте задач оптимизации с нерегулярными ограничениями. В разделе 2.1 исследуются свойства глобальной сходимости метода. Для задач оптимизации общего вида доказывается результат о глобальной сходимости, не требующий традиционных условий регулярности ограничений. Затем для задач оптимизации с комплементарными ограничениями доказывается так называемая S -стационарность предельных точек в естественных для таких задач предположениях. В разделе 2.2 исследуется возможность ускорения финальной фазы метода модифицированных функций с помощью стабилизированного метода Ньютона–Лагранжа. В приложении А приводятся результаты вычислительного эксперимента, показывающие, что метод модифицированных функций 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 стабилизированного метода последовательного квадратичного программирования. Построение глобально сходящихся алгоритмов, которые вблизи решения превращаются в стабилизированный метод последовательного квадратичного программирования, является давно стоящей проблемой. В частности, на сегодняшний день отсутствуют алгоритмы такого рода, способные конкурировать с обычным методом последовательного квадратичного программирования даже на задачах с нерегулярными ограничениями. В разделе 3.1 исследуются гибридные подходы к глобализации

сходимости на основе некоторого метода внешней фазы. В разделе 3.2 предлагается комбинировать стабилизированный метод последовательного квадратичного программирования с методами модифицированных функций Лагранжа. В разделе 3.3 рассматривается подход к глобализации сходимости с помощью точных гладких штрафных функций. Результаты вычислительного эксперимента, приведенные в приложении Б, показывают, что первые два подхода уступают по общей эффективности обычному методу последовательного квадратичного программирования, однако последний подход имеет более высокую эффективность для некоторых классов задач. Проведенное исследование также показало, что стабилизированный метод последовательного квадратичного программирования обладает серьезным недостатком: для широкого класса задач метод имеет тенденцию генерировать вдали от решения длинные последовательности коротких шагов, существенно замедляющих сходимость.

В **четвертой главе** разрабатывается метод последовательного квадратичного программирования, стабилизированный вдоль подпространства. Он представляет собой модификацию стабилизированного метода последовательного квадратичного программирования, целью которой является устранение описанной выше проблемы коротких шагов. В разделе 4.1 обосновывается локальная сверхлинейная сходимость метода вблизи квалифицированного решения. В разделе 4.2 предлагается простая с вычислительной точки зрения процедура идентификации подпространства вырожденности, необходимая для достижения высокой общей эффективности метода. Раздел 4.3 посвящен глобализации сходимости метода с помощью точных гладких штрафных функций. В приложении В приводятся численные результаты, показывающие высокую эффективность метода (в частности, по сравнению с обычным методом последовательного квадратичного программирования).

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

Научная новизна исследования состоит в следующем.

1. Получены принципиально новые априорные результаты о сходимости двойственных траекторий метода последовательного квадратичного программирования к критическим множителям Лагранжа.
2. Получены новые результаты о глобальной сходимости метода модифицированных функций Лагранжа как для задач оптимизации общего вида, так и для задач оптимизации с комплементарными ограничениями.
3. Предложено несколько новых подходов к глобализации сходимости стабилизированного метода последовательного квадратичного программирования.
4. Разработан принципиально новый метод последовательного квадратичного программирования, стабилизированный вдоль подпространства, а также подход к глобализации

его сходимости с помощью точных гладких штрафных функций. Разработана теория локальной и глобальной сходимости метода.

5. Разработана экономичная процедура аппроксимации подпространства вырожденности, которая существенно опережает все известные альтернативы по вычислительной эффективности.

Перечислим **научные положения, 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 значения 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.

Достоверность научных положений обусловлена строгостью математических доказательств и использованных научных методов.

Список публикаций. Полученные в работе результаты опубликованы в [10, 12, 13, 15–21, 24–27, 73–77], в том числе 5 статей опубликовано в журналах из списка ВАК [15, 17, 27, 73, 77].

Апробация результатов. Результаты, полученные в диссертации, были представлены на XXI международном симпозиуме по математическому программированию «ISMP 2012» (Берлин, Германия), на международной конференции по непрерывной оптимизации «ICCOPT 2013» (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».

В работе используется следующая система нумерации ее частей. Номер раздела состоит из двух цифр, первая из которых обозначает номер главы, в которой расположен раздел. Аналогично, номер пункта состоит из трех цифр, первые две из которых составляют номер раздела, в котором находится этот пункт. Нумерация объектов (формул, теорем и т. д.) в каждой главе независимая. При ссылке на объект извне главы, в которой он находится, используется номер, состоящий из двух цифр, первая из которых является номером главы, а вторая номером объекта в главе.

Автор выражает огромную благодарность своему научному руководителю Алексею Феридовичу Измаилову за постановку задачи, постоянное внимание к работе, ценные советы и поддержку.

Список основных обозначений

\mathbb{R} — множество вещественных чисел;

\mathbb{R}_+ — множество неотрицательных вещественных чисел;

\mathbb{R}^n — n -мерное арифметическое пространство, снабженное евклидовым скалярным произведением и соответствующей нормой;

$\langle \cdot, \cdot \rangle$ — евклидово скалярное произведение;

$\|z\|$ — евклидова норма вектора z (если не оговорено иное);

$\|A\|$ — спектральная норма матрицы A , т. е. порожденная евклидовой нормой (если не оговорено иное);

$\mathcal{B}(\bar{x}, \rho) = \{x \in \mathbb{R}^s \mid \|x - \bar{x}\| \leq \rho\}$ — замкнутый шар радиуса ρ с центром в точке \bar{x} ;

$\text{dist}(x, S) = \inf_{\xi \in S} \|x - \xi\|$ — расстояние от точки x до множества S ;

$\text{im } A$ — образ (множество значений) матрицы (линейного оператора) A ;

$\text{ker } A$ — ядро (множество нулей) матрицы (линейного оператора) A ;

A^T — матрица, транспонированная к матрице A ;

$\det A$ — определитель матрицы A ;

$\text{rank } A = \dim \text{im } A$ — ранг матрицы (линейного оператора) A ;

$z_{\mathcal{I}}$ — подвектор вектора z с компонентами $z_i, i \in \mathcal{I}$;

$A_{\mathcal{I}}$ — подматрица матрицы A , состоящая из строк с индексами $i \in \mathcal{I}$;

$A_{\mathcal{I}, \mathcal{J}}$ — подматрица матрицы A , образуемая пересечением строк с индексами $i \in \mathcal{I}$ и столбцов с индексами $j \in \mathcal{J}$;

$|\mathcal{I}|$ — количество элементов в конечном множестве \mathcal{I} ;

■ — знак окончания доказательства.

Глава 1

Эффект притяжения к критическим множителям Лагранжа

В данной главе исследуется эффект притяжения двойственных траекторий методов ньютоновского типа к так называемым критическим множителям Лагранжа. Раздел 1.1 посвящен обсуждению необходимой терминологии, возможных сценариев двойственного поведения и демонстрации данного эффекта на примерах. Основное внимание уделяется методу последовательного квадратичного программирования, для которого этот эффект является крайне типичным и приводит к потере сверхлинейной скорости сходимости.

В разделе 1.2 для метода последовательного квадратичного программирования доказываются результаты о том, что множество критических множителей действительно является аттрактором для двойственных траекторий, а также характеризуется влияние данного эффекта на скорость прямой и двойственной сходимости. Доказательство проводится для двух частных случаев: задач оптимизации с одной переменной и одним ограничением и для чисто квадратичных задач. Второй частный случай имеет ключевое значение для понимания эффекта притяжения, поскольку он отражает наиболее существенные последствия вырожденности ограничений.

В разделе 1.3 кратко рассматриваются стабилизированный метод последовательного квадратичного программирования и метод модифицированных функций Лагранжа. Эти методы обладают свойством двойственной стабилизации, т. е. способностью подавлять эффект притяжения к критическим множителям. Однако, данное свойство является локальным, и для обоих методов эффект притяжения 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.

1.1. Критические множители Лагранжа

Будем рассматривать задачу математического программирования

$$f(x) \rightarrow \min, \quad h(x) = 0, \quad g(x) \leq 0, \quad (1)$$

где $f: \mathbb{R}^n \rightarrow \mathbb{R}$ — дважды дифференцируемая функция, а $h: \mathbb{R}^n \rightarrow \mathbb{R}^l$ и $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$ — дважды дифференцируемые отображения.

Напомним, что точка $\bar{x} \in \mathbb{R}^n$ называется стационарной точкой задачи (1), если существуют такие $\bar{\lambda} \in \mathbb{R}^l$ и $\bar{\mu} \in \mathbb{R}^m$, что тройка $(\bar{x}, \bar{\lambda}, \bar{\mu})$ удовлетворяет системе Каруша–Куна–Таккера (ККТ)

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

этой задачи. Здесь $L: \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m \rightarrow \mathbb{R}$ — функция Лагранжа задачи (1):

$$L(x, \lambda, \mu) = f(x) + \langle \lambda, h(x) \rangle + \langle \mu, g(x) \rangle. \quad (3)$$

При этом пара $(\bar{\lambda}, \bar{\mu})$ называется множителем Лагранжа, отвечающим стационарной точке \bar{x} . Множество таких пар будем обозначать $\mathcal{M}(\bar{x})$.

Определим $A(\bar{x}) = \{i = 1, \dots, m \mid g_i(\bar{x}) = 0\}$ — множество номеров ограничений-неравенств задачи (1), активных в допустимой точке \bar{x} этой задачи. Введем также разбиение множества $A(\bar{x})$ на подмножества

$$A_+(\bar{x}, \bar{\mu}) = \{i \in A(\bar{x}) \mid \bar{\mu}_i > 0\}, \quad A_0(\bar{x}, \bar{\mu}) = \{i \in A(\bar{x}) \mid \bar{\mu}_i = 0\}.$$

Кроме того, положим $N(\bar{x}) = \{1, \dots, m\} \setminus A(\bar{x})$.

Как хорошо известно, стационарность локального решения \bar{x} задачи (1) можно гарантировать при выполнении в точке \bar{x} тех или иных условий регулярности ограничений. Важнейшим условием регулярности является условие Мангасариана–Фромовица (MFCQ, от англ. Mangasarian–Fromovitz constraint qualification), которое состоит в следующем:

$$\text{rank } h'(\bar{x}) = l, \quad \exists \xi \in \ker h'(\bar{x}): g'_{A(\bar{x})} \xi < 0. \quad (4)$$

Выполнение MFCQ в стационарной точке \bar{x} задачи (1) равносильно ограниченности множества $\mathcal{M}(\bar{x})$. Требование единственности отвечающего \bar{x} множителя $(\bar{\lambda}, \bar{\mu})$ называется строгим условием регулярности Мангасариана–Фромовица (SMFCQ, от англ. Strict Mangasarian–Fromovitz constraint qualification). Еще более сильным условием, чем SMFCQ, является условие линейной независимости (LICQ, от англ. Linear independence constraint qualification):

$$\text{rank} \begin{pmatrix} h'(\bar{x}) \\ g'_{A(\bar{x})}(\bar{x}) \end{pmatrix} = l + |A(\bar{x})|. \quad (5)$$

Как было сказано во введении, основной интерес в данной работе представляют случаи, когда множество $\mathcal{M}(\bar{x})$ содержит более одного элемента (т. е. нарушается SMFCQ) или даже является неограниченным (т. е. нарушается MFCQ).

Ключевую роль в данной работе играет следующее понятие критического множителя Лагранжа, введенное в [71]: множитель $(\bar{\lambda}, \bar{\mu}) \in \mathcal{M}(\bar{x})$ 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

$$\begin{aligned} \frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda}, \bar{\mu})\xi + (h'(\bar{x}))^T \eta + (g'(\bar{x}))^T \zeta &= 0, & h'(\bar{x})\xi &= 0, & g'_{A_+(\bar{x}, \bar{\mu})}(\bar{x})\xi &= 0, \\ \zeta_{A_0(\bar{x}, \bar{\mu})} &\geq 0, & g'_{A_0(\bar{x}, \bar{\mu})}(\bar{x})\xi &\leq 0, & \zeta_i \langle g'_i(\bar{x}), \xi \rangle &= 0, \quad i \in A_0(\bar{x}, \bar{\mu}), & \zeta_{N(\bar{x})} &= 0, \end{aligned} \quad (6)$$

и *некритическим* иначе. Множество критических множителей обычно является тощим по отношению к множеству всех множителей. Например, при отсутствии ограничений-неравенств, если множество некритических множителей непусто, то оно является относительно открытым и всюду плотным в множестве всех множителей. Вместе с тем, критические множители обладают рядом особых свойств и оказывают значительное влияние на поведение ряда важнейших численных методов оптимизации.

Одно из таких свойств состоит в том, что для критических множителей неизбежно нарушается *достаточное условие второго порядка оптимальности* (SOSC, от англ. Second-order sufficient condition), которое в стационарной точке \bar{x} задачи (1) для отвечающего ей множителя Лагранжа $(\bar{\lambda}, \bar{\mu}) \in \mathcal{M}(\bar{x})$ имеет следующий вид:

$$\left\langle \frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda}, \bar{\mu})\xi, \xi \right\rangle > 0 \quad \forall \xi \in C(\bar{x}) \setminus \{0\}, \quad (7)$$

где

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

есть критический конус задачи (1) в точке \bar{x} . Напомним, что в случае выполнения (7) точка \bar{x} является строгим локальным решением задачи (1). Более того, если в стационарной точке \bar{x} для множителя $(\bar{\lambda}, \bar{\mu}) \in \mathcal{M}(\bar{x})$ выполнено *необходимое условие второго порядка оптимальности* (SONC, от англ. Second-order necessary condition)

$$\left\langle \frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda}, \bar{\mu})\xi, \xi \right\rangle \geq 0 \quad \forall \xi \in C(\bar{x}), \quad (8)$$

то множитель $(\bar{\lambda}, \bar{\mu})$ является некритическим тогда и только тогда, когда он удовлетворяет SOSC.

Другое важное свойство критических множителей устанавливается в следующем предложении, взятом из [72]:

Предложение 1. Пусть функция $f: \mathbb{R}^n \rightarrow \mathbb{R}$ и отображения $h: \mathbb{R}^n \rightarrow \mathbb{R}^l$, $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$ дважды дифференцируемы в точке \bar{x} . Пусть \bar{x} — стационарная точка задачи (1), а $(\bar{\lambda}, \bar{\mu})$ — отвечающий ей множитель Лагранжа.

Тогда следующие свойства являются эквивалентными:

- а) множитель $(\bar{\lambda}, \bar{\mu})$ является некритическим;
- б) выполнена оценка расстояния

$$\|x - \bar{x}\| + \text{dist}((\lambda, \mu), \mathcal{M}(\bar{x})) = O\left(\left\|\begin{pmatrix} \frac{\partial L}{\partial x}(x, \lambda, \mu) \\ h(x) \\ \min\{\mu, -g(x)\} \end{pmatrix}\right\|\right) \quad (9)$$

при $(x, \lambda, \mu) \rightarrow (\bar{x}, \bar{\lambda}, \bar{\mu})$;

- в) для любого $\sigma = (a, b, c) \in \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m$, достаточно близкого к нулю, любое решение $(x(\sigma), \lambda(\sigma), \mu(\sigma))$ канонически возмущенной системы ККТ

$$\begin{aligned} f'(x) + (h'(x))^T \lambda + (g'(x))^T \mu &= a, & h(x) &= b, \\ \mu &\geq 0, & g(x) &\leq c, & \langle \mu, g(x) - c \rangle &= 0, \end{aligned}$$

достаточно близкое к $(\bar{x}, \bar{\lambda}, \bar{\mu})$, удовлетворяет оценке

$$\|x(\sigma) - \bar{x}\| + \text{dist}((\lambda(\sigma), \mu(\sigma)), \mathcal{M}(\bar{x})) = O(\|\sigma\|). \quad (10)$$

Таким образом, оценки расстояния (9) и (10) не могут выполняться вблизи критического множителя.

При отсутствии ограничений-неравенств задача (1) принимает вид

$$f(x) \rightarrow \min, \quad h(x) = 0. \quad (11)$$

Стационарные точки и отвечающие им множители Лагранжа задачи (11) характеризуются системой Лагранжа

$$\frac{\partial L}{\partial x}(x, \lambda) = 0, \quad h(x) = 0, \quad (12)$$

где $L: \mathbb{R}^n \times \mathbb{R}^l \rightarrow \mathbb{R}$ — функция Лагранжа задачи (11):

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

Для задачи (11) условия регулярности MFCQ, SMFCQ и LICQ эквивалентны и имеют следующий вид:

$$\text{rank } h'(\bar{x}) = l. \quad (13)$$

Как следует из первых двух равенств в (6), для задачи с ограничениями-равенствами критичность множителя $\bar{\lambda} \in \mathcal{M}(\bar{x})$ означает существование такого элемента $\xi \in \ker h'(\bar{x}) \setminus \{0\}$, что

$$\frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda})\xi \in \text{im}(h'(\bar{x}))^T. \quad (14)$$

Для таких задач понятие критического множителя было введено в [8]. В случае полного вырождения (т.е. при $h'(\bar{x}) = 0$) критичность $\bar{\lambda}$ означает вырожденность матрицы Гессе $\frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda})$. В общем случае, поскольку $\text{im}(h'(\bar{x}))^T = (\ker h'(\bar{x}))^\perp$, условие (14) эквивалентно условию

$$\pi_{\ker h'(\bar{x})} \frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda})\xi = 0,$$

где через $\pi_{\ker h'(\bar{x})}$ обозначен оператор ортогонального проектирования на $\ker h'(\bar{x})$.

Симметричную матрицу линейного оператора $\xi \rightarrow \pi_{\ker h'(\bar{x})} \frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda})\xi: \ker h'(\bar{x}) \rightarrow \ker h'(\bar{x})$ называют суженной матрицей Гессе функции Лагранжа по x . Критичность множителя $\bar{\lambda}$ означает вырожденность этой матрицы, а условия SONC и SOSC эквивалентны ее неотрицательной и положительной определенности соответственно. Отсюда, в частности, очевидно, что SOSC является гораздо более сильным условием, чем некритичность.

Как уже упоминалось во введении, скорость сходимости ньютоновских методов к нерегулярным решениям обычно является лишь линейной. Однако, как было замечено в [94, раздел 6] и весьма убедительно продемонстрировано в [8, 67, 68, 70], причина медленной прямой сходимости состоит не в вырожденности ограничений как таковой, а в определенном нежелательном поведении двойственной траектории. А именно, анализ и вычислительный опыт свидетельствуют о том, что двойственные траектории ньютоновских методов имеют сильнейшую тенденцию притяжения к критическим множителям, и именно это обстоятельство является причиной медленной сходимости.

Рассмотрим данный эффект на примере метода последовательного квадратичного программирования (SQP, от англ. Sequential quadratic programming). Итерация метода заключается в следующем. 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

$$\begin{aligned} \langle f'(x), \xi \rangle + \frac{1}{2} \left\langle \frac{\partial^2 L}{\partial x^2}(x, \lambda, \mu)\xi, \xi \right\rangle &\rightarrow \min \\ h(x) + h'(x)\xi &= 0, \quad g(x) + g'(x)\xi \leq 0, \end{aligned} \quad (15)$$

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 (15). При отсутствии ограничений-неравенств итерацию можно записать в более простом виде. Для текущего приближения $(x, \lambda) \in \mathbb{R}^n \times \mathbb{R}^l$ вычисляется решение $(\xi, \eta) \in \mathbb{R}^n \times \mathbb{R}^l$ системы

$$\frac{\partial^2 L}{\partial x^2}(x, \lambda)\xi + (h'(x))^T \eta = -\frac{\partial L}{\partial x}(x, \lambda), \quad h'(x)\xi = -h(x), \quad (16)$$

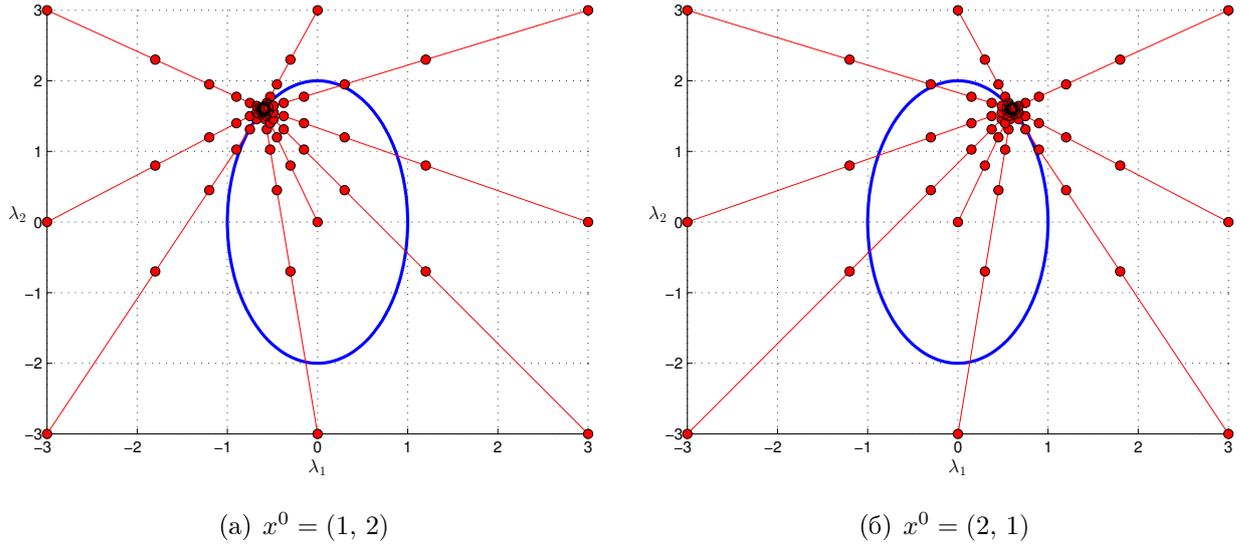


Рис. 1. Двойственные траектории МНЛ для примера 1.

и следующее приближение полагается равным $(x + \xi, \lambda + \eta)$. В этом случае метод эквивалентен применению метода Ньютона к системе Лагранжа задачи (11), и поэтому называется методом Ньютона–Лагранжа (МНЛ).

Сначала рассмотрим возможные сценарии поведения МНЛ для задач вида (11), в которых есть критические множители. Первый сценарий, а именно притяжение двойственной траектории к критическим множителям и линейная скорость сходимости прямой траектории, является весьма типичным и наблюдается в большинстве случаев [67, 68]. Такое поведение демонстрируется примерами 1–3, взятыми из [42, 67].

Пример 1. (Пример 2.1 из [67], тест 20203 из коллекции DEGEN [42].) Рассмотрим задачу

$$-x_1^2 - x_2^2 \rightarrow \min, \quad x_1^2 - x_2^2 = 0, \quad x_1 x_2 = 0.$$

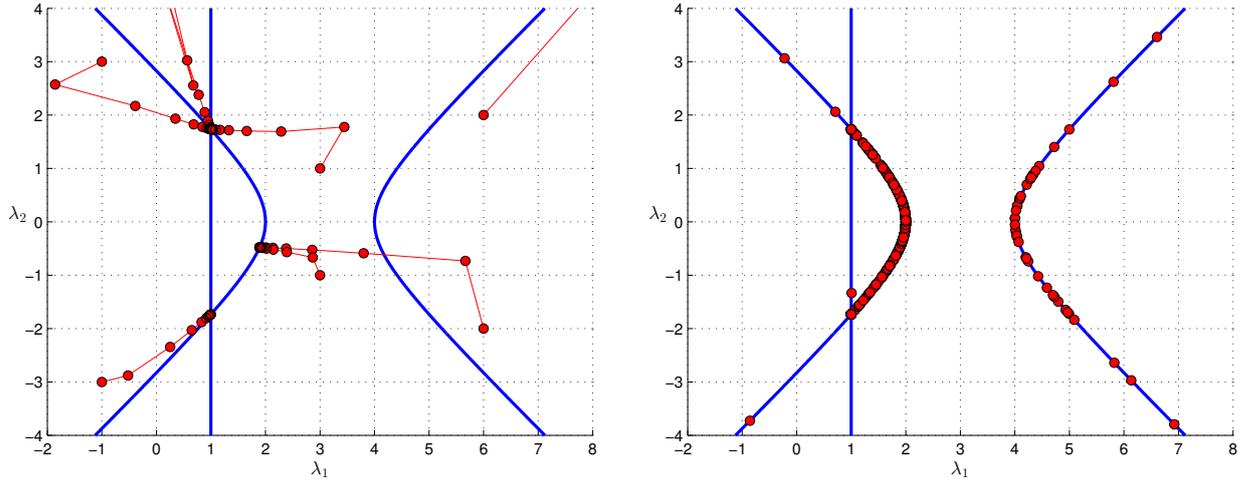
Задача имеет единственное решение $\bar{x} = 0$, $\mathcal{M}(\bar{x}) = \mathbb{R}^2$. Критическими являются множители $\lambda \in \mathbb{R}^2$, для которых $4\lambda_1^2 + \lambda_2^2 = 4$ (эллипс на рис. 1).

На рис. 1 показаны некоторые двойственные траектории МНЛ для двух различных значений x^0 . Как легко проверить аналитически, в данном примере при фиксированном x^0 двойственные траектории всегда сходятся к одному критическому множителю. Однако, как видно из рисунка, при различных x^0 предельные критические множители различаются.

Пример 2. (Пример 2.3 из [67], тест 20302 из коллекции DEGEN [42].) Рассмотрим задачу

$$x_1^2 - x_2^2 + 2x_3^2 \rightarrow \min, \quad -\frac{1}{2}x_1^2 + x_2^2 - \frac{1}{2}x_3^2 = 0, \quad x_1 x_3 = 0.$$

Она имеет единственное решение $\bar{x} = 0$, $\mathcal{M}(\bar{x}) = \mathbb{R}^2$. Критическими являются множители $\lambda \in \mathbb{R}^2$, для которых $\lambda_1 = 1$ или $(\lambda_1 - 3)^2 - \lambda_2^2 = 1$ (вертикальная линия и гипербола на



(а) Двойственные траектории МНЛ

(б) Распределение двойственных приближений на момент остановки

Рис. 2. Пример 2; $x^0 = (1, 2, 3)$.

рис. 2).

На рис. 2(а) показаны некоторые двойственные траектории МНЛ для начальной точки $x^0 = (1, 2, 3)$. На рис. 2(б) показано распределение двойственных приближений на момент остановки МНЛ, который запускался из точек λ^0 сетки $[-2, 8] \times [-2, 4]$ с шагом $1/4$. Как видно из рисунка, во всех случаях предельный множитель был критическим.

Следующий простой пример, взятый из [8, 68], позволяет аналитически продемонстрировать эффект притяжения к критическим множителям.

Пример 3. (Тест 20101 из коллекции DEGEN [42].) Пусть $n = l = 1$, $f(x) = x^2$, $h(x) = x^2$. Задача (11) с такими данными имеет единственную допустимую точку (а значит, единственное решение) $\bar{x} = 0$, причем $\mathcal{M}(0) = \mathbb{R}$ (поскольку $h'(0) = 0$). Это решение удовлетворяет SOS (7) при $\bar{\lambda} > -1$, а единственным критическим множителем Лагранжа является $\bar{\lambda} = -1$.

Согласно (16), любая прямо-двойственная траектория $\{(x_k, \lambda_k)\}$ МНЛ для данной задачи удовлетворяет равенствам

$$(1 + \lambda_k)(x_{k+1} - x_k) + x_k(\lambda_{k+1} - \lambda_k) = -(1 + \lambda_k)x_k, \quad 2x_k(x_{k+1} - x_k) = -(x_k)^2$$

для каждого k . Предполагая, что $x_k \neq 0$, из второго уравнения получаем

$$x_{k+1} = \frac{1}{2}x_k,$$

и тогда из первого уравнения

$$\lambda_{k+1} = \frac{1}{2}(\lambda_k - 1).$$

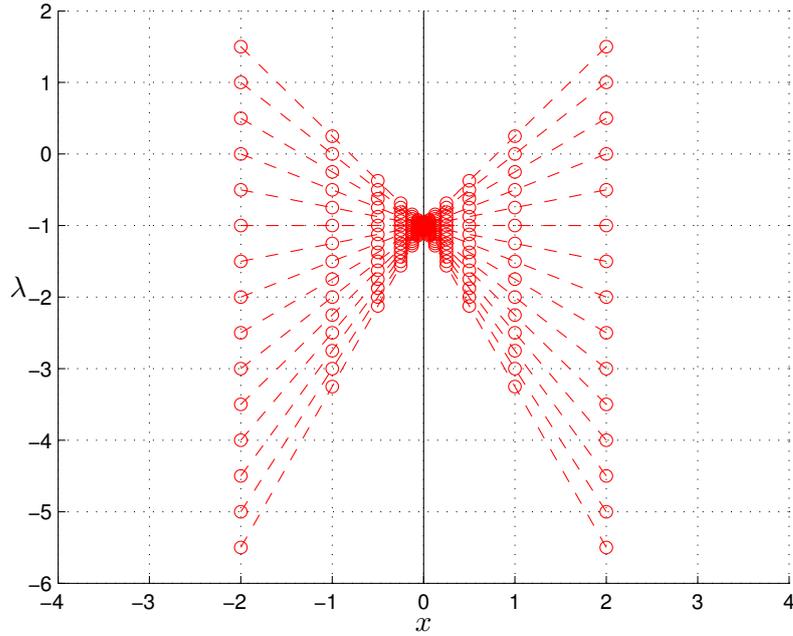


Рис. 3. Траектории МНЛ для примера 3.

В частности,

$$\lambda_{k+1} + 1 = \frac{1}{2}(\lambda_k + 1).$$

Таким образом, если $x^0 \neq 0$, то $x_k \neq 0$ для всех k , и последовательность $\{(x_k, \lambda_k)\}$ корректно определена и сходится к $(0, -1)$ с линейной скоростью. В частности, $\{\lambda_k\}$ сходится к критическому множителю $\bar{\lambda} = -1$. Более того, если $\lambda_0 \neq -1$, то каждая из последовательностей $\{x_k\}$ и $\{\lambda_k\}$ сходится с линейной скоростью.

На рис. 3 показаны прямо-двойственные траектории МНЛ. Вертикальная сплошная линия соответствует множеству $\{\bar{x}\} \times \mathcal{M}(\bar{x})$. Запуски выполнялись из точек (x^0, λ^0) , где x^0 равнялось -2 или 2 , а λ^0 принимало различные значения из сегмента $[-5.5, 0.5]$. Как видно из рисунка, все траектории сходятся к $(\bar{x}, \bar{\lambda})$ с линейной скоростью.

Другой возможный сценарий состоит в отсутствии сходимости двойственной траектории. В отличие от предыдущего сценария, такое поведение крайне нетипично, но возможно, как показывает следующий (значительно более сложный) пример.

Пример 4. Рассмотрим задачу

$$2x_1^2 + 2x_2^2 + 5x_2x_3 \rightarrow \min, \quad x_1x_2 + x_1x_3 + x_2x_3 = 0. \quad (17)$$

Систему Лагранжа задачи (17) можно записать в виде

$$H(\lambda)x = 0, \quad \frac{1}{2}\langle Bx, x \rangle = 0, \quad (18)$$

где

$$B = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \quad H(\lambda) = \begin{pmatrix} 4 & \lambda & \lambda \\ \lambda & 4 & 5 + \lambda \\ \lambda & 5 + \lambda & 0 \end{pmatrix}.$$

Здесь через $H(\lambda)$ обозначена матрица Гессе функции Лагранжа задачи (17).

Согласно (18), задача (17) имеет стационарную точку $\bar{x} = 0$, а множество множителей Лагранжа, отвечающих \bar{x} , есть все \mathbb{R} . Далее, $\det H(\lambda) = 2(\lambda - 5)(\lambda^2 + 6\lambda + 10)$, и поэтому существует единственный критический множитель $\bar{\lambda} = 5$.

Итерация МНЛ для задачи (17) выглядит следующим образом:

$$\begin{pmatrix} H(\lambda^k) & Bx^k \\ (Bx^k)^\top & 0 \end{pmatrix} \begin{pmatrix} x^{k+1} - x^k \\ \lambda^{k+1} - \lambda^k \end{pmatrix} = - \begin{pmatrix} H(\lambda^k)x^k \\ \frac{1}{2}\langle Bx^k, x^k \rangle \end{pmatrix}. \quad (19)$$

Несложно убедиться в том, что если $\lambda^k \neq 5$, то матрица в левой части (19) является невырожденной тогда и только тогда, когда

$$\langle Bx^k, \hat{H}(\lambda^k)Bx^k \rangle \neq 0. \quad (20)$$

Здесь и далее через \hat{H} обозначается присоединенная матрица для матрицы H (в частности, $\hat{H}H = H\hat{H} = I \det H$). Далее, если $\lambda_k \neq 5$, и выполнено (20) то из (19) получаем

$$x^{k+1} = \frac{1}{2} \frac{\langle Bx^k, x^k \rangle}{\langle Bx^k, \hat{H}(\lambda_k)Bx^k \rangle} \hat{H}(\lambda_k)Bx^k, \quad \lambda^{k+1} = \lambda^k - \frac{\det H(\lambda^k)}{2} \frac{\langle Bx^k, x^k \rangle}{\langle Bx^k, \hat{H}(\lambda^k)Bx^k \rangle}. \quad (21)$$

Рассмотрим линейное подпространство $\mathcal{L} = \ker \hat{H}(\bar{\lambda})B$ в \mathbb{R}^n . Легко проверить, что \mathcal{L} является линейной оболочкой векторов $a^1 = (1, -1, 3)$ и $a^2 = (3, 2, 2)$, и, в частности, $\dim \mathcal{L} = 2$. Далее, можно проверить, что при любом $\lambda \neq 5$ выполнено равенство

$$\hat{H}(\bar{\lambda})B\hat{H}(\lambda)B = 2(\lambda^2 + 6\lambda + 10)\hat{H}(\bar{\lambda})B,$$

из которого следуют равенства $\hat{H}(\bar{\lambda})B\hat{H}(\lambda)Ba^1 = 0$ и $\hat{H}(\bar{\lambda})B\hat{H}(\lambda)Ba^2 = 0$. Используя последние два равенства, получаем $\hat{H}(\lambda)B\mathcal{L} \subset \mathcal{L}$. Наконец, заметим, что если $x^k \in \mathcal{L}$ и $\lambda^k = 5$, то матрица в левой части (58) обязательно является вырожденной.

Приведенные рассуждения позволяют заключить, что если $x^k \in \mathcal{L}$, $\lambda^k \neq 5$, и выполнено (20), то равенства в (21) корректно определяют x^{k+1} и λ^{k+1} , причем $x^{k+1} \in \mathcal{L}$. Таким образом, если $x^0 \in \mathcal{L}$, $\lambda^0 \neq 5$, и (20) выполнено при $k = 0$, то либо для некоторого k будет выполнено $\lambda^k = 5$ или $\langle Bx^k, \hat{H}(\lambda^k)Bx^k \rangle = 0$ (и следующая итерация будет не определена), либо для всех k справедливо $x^k \in \mathcal{L}$, $\lambda^k \neq 5$, и выполнено (20) (откуда, в частности, следует $x^k \neq 0$).

Предположим, что имеет место последний случай, и покажем, что сходимость к критическому множителю $\bar{\lambda}$ при этом невозможна. Определим матрицу $U \in \mathbb{R}^{3 \times 2}$ со столбцами a^1 , a^2 , и рассмотрим (единственную) последовательность $\{z^k\} \subset \mathbb{R}^2 \setminus \{0\}$ такую, что $x^k = Uz^k$

для всех k . Непосредственными вычислениями можно убедиться, что выполнены соотношения

$$\hat{H}(\lambda)BU = 2(\lambda - 5)UM_1(\lambda), \quad U^T B \hat{H}(\lambda)BU = 4(\lambda - 5)M_2(\lambda),$$

где

$$M_1(\lambda) = \begin{pmatrix} \lambda + 2 & -2 \\ 1 & \lambda + 4 \end{pmatrix}, \quad M_2(\lambda) = \begin{pmatrix} 5 - \lambda & 7\lambda + 30 \\ 7\lambda + 30 & 16\lambda + 50 \end{pmatrix}.$$

Из данных соотношений и из (21) легко получить, что

$$\begin{aligned} x^{k+1} &= \frac{\langle Mz^k, z^k \rangle}{2\langle M_2(\lambda_k)z^k, z^k \rangle} UM_1(\lambda_k)z^k, \\ \lambda^k - \lambda^{k+1} &= ((\lambda^k)^2 + 6\lambda^k + 10) \frac{\langle Mz^k, z^k \rangle}{2\langle M_2(\lambda_k)z^k, z^k \rangle}, \end{aligned} \quad (22)$$

где

$$M = \frac{1}{2}U^T BU = \begin{pmatrix} -1 & 7 \\ 7 & 16 \end{pmatrix},$$

откуда следует

$$z^{k+1} = \frac{\langle Mz^k, z^k \rangle}{2\langle M_2(\lambda_k)z^k, z^k \rangle} M_1(\lambda_k)z^k. \quad (23)$$

Введем вспомогательную последовательность $\{\zeta^k\}$ такую, что $\zeta^0 = z^0/\|z^0\|$ и для каждого номера k вектор ζ^{k+1} определяется соотношением

$$\zeta^{k+1} = \frac{M_1(\lambda^k)\zeta^k}{\|M_1(\lambda^k)\zeta^k\|}. \quad (24)$$

Заметим, что из соотношений (23), (24) и условия $z^k \neq 0$, выполненного для всех k , вытекает, что последовательность $\{\zeta^k\}$ корректно определена, и для каждого k вектор ζ^k совпадает с $z^k/\|z^k\|$ с точностью до знака. Поэтому из (22) получаем

$$\lambda^k - \lambda^{k+1} = ((\lambda^k)^2 + 6\lambda^k + 10) \frac{\langle M\zeta^k, \zeta^k \rangle}{2\langle M_2(\lambda^k)\zeta^k, \zeta^k \rangle}. \quad (25)$$

Предположим, что последовательность $\{\lambda^k\}$ сходится к некоторому $\lambda^* \in \mathbb{R}$. Как следует из (25), сходимость $\{\lambda^k\}$ возможна только при условии, что $\langle M\zeta^k, \zeta^k \rangle$ сходится к 0. Следовательно, все предельные точки $\bar{\zeta}$ последовательности $\{\zeta^k\}$ должны удовлетворять $\langle M\bar{\zeta}, \bar{\zeta} \rangle = 0$. Однако, как следует из (24), $\hat{\zeta} = M_1(\lambda^*)\bar{\zeta}/\|M_1(\lambda^*)\bar{\zeta}\|$ также является предельной точкой $\{\zeta^k\}$, и поэтому $\langle M\hat{\zeta}, \hat{\zeta} \rangle = 0$, откуда следует $\langle M_1(\lambda^*)^T M M_1(\lambda^*)\bar{\zeta}, \bar{\zeta} \rangle = 0$. Далее, несложно проверить, что $M_1(\lambda)^T M M_1(\lambda) = 130(\lambda + 3)\hat{M} + (\lambda^2 - 10\lambda - 40)M$, где

$$\hat{M} = \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix}.$$

Следовательно, если $\lambda^* \neq -3$, то вектор $\bar{\zeta}$ также должен удовлетворять $\langle \hat{M}\bar{\zeta}, \bar{\zeta} \rangle = 0$. Однако, данное равенство не может выполняться одновременно с $\langle M\bar{\zeta}, \bar{\zeta} \rangle = 0$. Таким образом, последовательность $\{\lambda^k\}$ не может сходиться к точке, отличной от -3 .

Предположим теперь, что $\lambda^* = -3$. Из (25) с учетом равенства $M_2(\lambda) = MM_1(\lambda)$ получаем

$$\lambda^{k+1} - \lambda^{k+2} = ((\lambda^{k+1})^2 + 6\lambda^{k+1} + 10) \frac{\langle MM_1(\lambda^k)\zeta^k, M_1(\lambda^k)\zeta^k \rangle}{2\langle MM_1(\lambda^k)\zeta^k, M_1(\lambda^{k+1})M_1(\lambda^k)\zeta^k \rangle}.$$

Кроме того, второе равенство в (25) можно переписать в виде

$$\lambda^{k+1} + 3 = \frac{\langle MM_1(\lambda^k)\zeta^k, M_1(\lambda^k)\zeta^k \rangle}{2\langle M\zeta^k, M_1(\lambda^k)\zeta^k \rangle}.$$

Разделив предпоследнее равенство на последнее, получаем

$$1 - \frac{\lambda^{k+2} + 3}{\lambda^{k+1} + 3} = ((\lambda^{k+1})^2 + 6\lambda^{k+1} + 10) \frac{\langle M\zeta^k, M_1(\lambda^k)\zeta^k \rangle}{\langle MM_1(\lambda^k)\zeta^k, M_1(\lambda^{k+1})M_1(\lambda^k)\zeta^k \rangle}.$$

Заметим, что последовательность $\{\zeta^k\}$ может иметь не более четырех различных предельных точек, поскольку такие точки должны удовлетворять равенству $\langle M\bar{\zeta}, \bar{\zeta} \rangle = 0$. Более того, прямой подстановкой можно убедиться, что для каждой такой точки $\bar{\zeta}$ величина $\langle M\bar{\zeta}, M_1(\lambda^*)\bar{\zeta} \rangle / \langle MM_1(\lambda^*)\bar{\zeta}, M_1(\lambda^*)M_1(\lambda^*)\bar{\zeta} \rangle$ является отрицательной, откуда следует, что $(\lambda^{k+2} + 3) / (\lambda^{k+1} + 3) > 1 + \varepsilon$ для некоторого $\varepsilon > 0$ и для всех достаточно больших k . Очевидно, отсюда следует, что λ^{k+2} будет находиться дальше от -3 , чем λ^{k+1} , и поэтому последовательность $\{\lambda^k\}$ не может сходиться к -3 .

Таким образом, если $x^k \in \mathcal{L}$ для некоторого k , то последовательность $\{\lambda^k\}$ всегда расходится (и, в частности, не может сходиться к критическому множителю $\bar{\lambda} = 5$). Более того, как показывает вычислительный эксперимент, у данной последовательности всегда есть некритические предельные точки, и при этом последовательность $\{x^k\}$ всегда сходится к 0.

Отметим также, что описанное поведение наблюдается в данном примере достаточно часто, и является в некотором смысле устойчивым: оно не разрушается малым шевелением начальной точки. Даже если $x^0 \notin \mathcal{L}$, направления $x^k / \|x^k\|$ во многих случаях стремятся к подпространству \mathcal{L} . Пример такого поведения показан на рис. 4, где $x^0 = (1, -1, 2)$, $\lambda^0 = 1$.

Тем не менее, стоит отметить, что сходимость к критическому множителю по-прежнему является типичным сценарием в данном примере, особенно если направление $x^0 / \|x^0\|$ расположено достаточно далеко от \mathcal{L} , или если λ^0 достаточно близко к $\bar{\lambda}$. Конечно, в этом случае для всех k выполнено $x^k \notin \mathcal{L}$.

Наконец, возможен «смешанный» сценарий поведения, который также является достаточно типичным: двойственная траектория сначала ведет себя «хаотически», но потом попадает в область притяжения одного из критических множителей и далее сходится к этому множителю. Для демонстрации такого поведения вновь рассмотрим пример 2. На

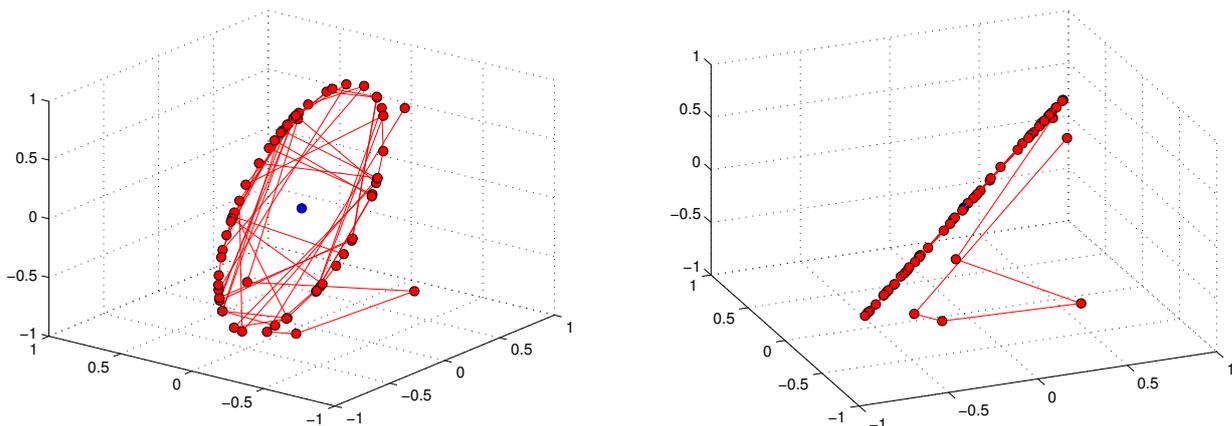


Рис. 4. Последовательность $\{x^k / \|x^k\|\}$ для примера 4 при $x^0 = (1, -1, 2)$, $\lambda^0 = 1$.

рис. 5 показана двойственная траектория МНЛ для начального приближения $x^0 = (3, 2, 1)$, $\lambda^0 = (3.75, -0.25)$, которая ведет себя описанным образом. (В [67, пример 2.3], сходимость двойственной траектории не была обнаружена в связи со слишком ранним завершением итерационного процесса.)

Отметим, что если в задаче (11) с нерегулярными ограничениями отсутствуют критические множители, то наиболее типичным является сценарий, когда двойственная траектория МНЛ расходится. Это связано с наличием другого эффекта, возникающего при сходимости МНЛ к решениям, которым отвечает более одного множителя Лагранжа — эффекта отталкивания двойственных траекторий от не критических множителей. Более того, данный эффект может возникать даже если в задаче присутствуют критические множители (в частности, именно он лежит в основе поведения, рассмотренного в примере 4), однако в таких случаях он является крайне нетипичным. Исследование данного эффекта можно найти в работах [68] и [77].

Теперь кратко обсудим, как меняются возможные сценарии поведения SQP при добавлении ограничений-неравенств. Для этого потребуются дополнительные определения. Из (2) немедленно следует, что если для множителя $(\bar{\lambda}, \bar{\mu}) \in \mathcal{M}(\bar{x})$ выполняется $\bar{\mu}_{A(\bar{x}) \setminus A} = 0$ для некоторого множества индексов $A \subset A(\bar{x})$, то \bar{x} является стационарной точкой задачи с ограничениями-равенствами

$$f(x) \rightarrow \min, \quad h(x) = 0, \quad g_A(x) = 0, \quad (26)$$

в то время как $(\bar{\lambda}, \bar{\mu}_A)$ является отвечающим этой стационарной точке множителем Лагранжа. Понятие критического множителя Лагранжа дополняется следующим, введенным в [67]: множитель $(\bar{\lambda}, \bar{\mu}) \in \mathcal{M}(\bar{x})$ называется *критическим относительно множества индексов* $A \subset A(\bar{x})$, если $\bar{\mu}_{A(\bar{x}) \setminus A} = 0$ и $(\bar{\lambda}, \bar{\mu}_A)$ — критический множитель Лагранжа, отвечающий

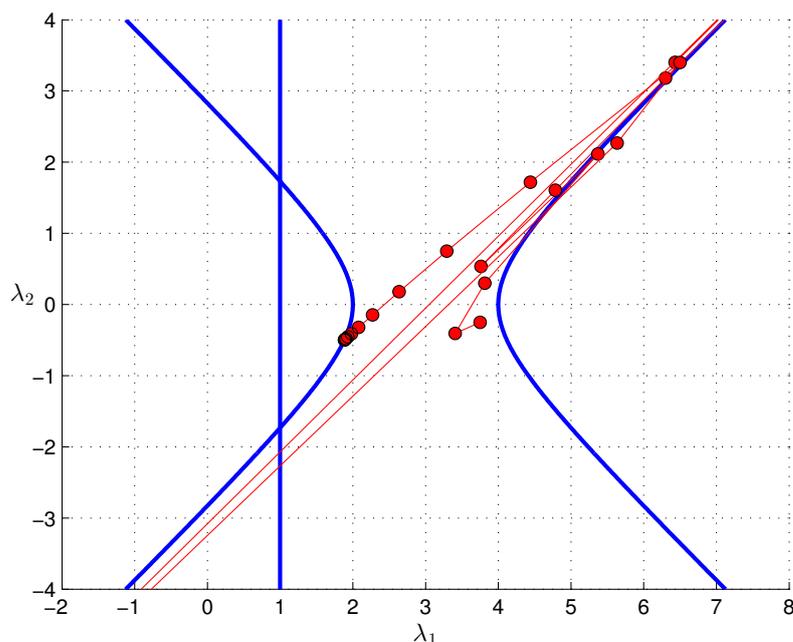


Рис. 5. Двойственная траектория для примера 2.

стационарной точке \bar{x} задачи (26). В противном случае $(\bar{\lambda}, \bar{\mu})$ называется *некритическим относительно A* .

Как легко видеть из приведенных определений, если множитель $(\bar{\lambda}, \bar{\mu})$ критический, то существует $I_1 \subset A_0(\bar{x}, \bar{\mu})$ такое, что этот множитель является критическим относительно множества индексов $A = A_+(\bar{x}, \bar{\mu}) \cup I_1$. В частности, если выполнено *условие строгой дополнителности* $\bar{\mu}_{A(\bar{x})} > 0$ (т. е. $A_0(\bar{x}, \bar{\mu}) = \emptyset$), то некритичность эквивалентна некритичности относительно $A(\bar{x})$.

Очевидно, что из выполнения SOS (7) вытекает выполнение соответствующего достаточного условия второго порядка оптимальности в стационарной точке \bar{x} задачи (26) при $A = A(\bar{x})$ для отвечающего этой стационарной точке множителя Лагранжа $(\bar{\lambda}, \bar{\mu}_{A(\bar{x})})$, и поэтому $(\bar{\lambda}, \bar{\mu})$ не может быть критическим множителем Лагранжа для задачи (1) относительно $A(\bar{x})$. Вместе с тем, критичность такого множителя относительно более узких множеств индексов $A \subset A(\bar{x})$ может иметь место (см. [67, пример 3.5]).

Рассмотрим произвольную траекторию $\{(x^k, \lambda^k, \mu^k)\}$ метода SQP для задачи (1) и для каждого k введем множество

$$A_k = \{i = 1, \dots, m \mid g_i(x^k) + \langle g'_i(x^k), x^{k+1} - x^k \rangle = 0\}$$

индексов ограничений-неравенств соответствующей подзадачи SQP, активных в точке x^k . Весьма естественным является предположение о том, что множества A_k стабилизируются на поздних итерациях, т. е. $A_k = A$ для всех достаточно больших k . В частности, данное предположение выполнено автоматически, если траектория сходится к точке $(\bar{x}, \bar{\lambda}, \bar{\mu})$, где \bar{x} — стационарная точка, а множитель $(\bar{\lambda}, \bar{\mu}) \in \mathcal{M}(\bar{x})$ удовлетворяет условию строгой до-

полнительности. Если $A_k = A$ для всех достаточно больших k , то метод SQP можно интерпретировать как МНЛ для задачи (26). В связи с этим, возможны следующие сценарии поведения.

Если ограничения задачи (26) регулярны в точке \bar{x} , т. е.

$$\text{rank} \begin{pmatrix} h'(\bar{x}) \\ g'_A(\bar{x}) \end{pmatrix} = l + |A|, \quad (27)$$

то точке \bar{x} отвечает единственный множитель $(\bar{\lambda}, \bar{\mu}_A)$, который обычно является не критическим. В таких случаях двойственная траектория метода сходится к этому множителю, а скорость прямой и прямо-двойственной сходимости является сверхлинейной.

Если же условие (27) нарушается в точке \bar{x} , то этой точке отвечает бесконечное множество множителей. В этом случае поведение метода зависит от наличия критических множителей относительно A . Если такие множители существуют, то двойственная траектория метода в большинстве случаев сходится к одному из них, и при этом скорость прямой сходимости является лишь линейной. Если же все множители являются не критическими относительно A , то двойственная траектория обычно расходится.

В завершение данного раздела отметим, что эффект притяжения двойственных траекторий к критическим множителям Лагранжа имеет место и для многих других ньютоновских методов. В частности, в [70] продемонстрировано, что данный эффект имеет место для методов модифицированных функций Лагранжа с линеаризованными ограничениями и квазиньютоновских методов последовательного квадратичного программирования.

В следующем разделе речь пойдет о теоретическом обосновании эффекта притяжения к критическим множителям для МНЛ.

1.2. Обоснование эффекта притяжения для метода Ньютона–Лагранжа

Существующие теоретические объяснения эффекта притяжения двойственных траекторий ньютоновских методов к критическим множителям Лагранжа [8, 68, 70] носят в основном «негативный» характер: они показывают, что было бы, если бы двойственная траектория сходилась к не критическому множителю, а также содержат аргументы в пользу того, что такое поведение крайне маловероятно.

В данном разделе разрабатываются первые «позитивные» результаты, показывающие, что множество критических множителей Лагранжа действительно является аттрактором для двойственных траекторий МНЛ. Обоснование проводится для двух частных случаев: задач с одной переменной и одним ограничением и чисто квадратичных задач. Второй частный

случай имеет ключевое значение для понимания эффекта притяжения, поскольку он отражает наиболее существенные последствия вырожденности ограничений. Кроме того, возможно, полученные результаты могут быть расширены на общий случай с применением процедуры Ляпунова–Шмидта (см., например, [58, гл. VII]). Вместе с тем, даже в случае чисто квадратичной задачи анализ оказывается крайне нетривиальным.

1.2.1. Одномерные задачи

Будем рассматривать задачу оптимизации с одной скалярной переменной и одним ограничением-равенством

$$f(x) \rightarrow \min, \quad h(x) = 0, \quad (28)$$

где функции $f: \mathbb{R} \rightarrow \mathbb{R}$ и $h: \mathbb{R} \rightarrow \mathbb{R}$ предполагаются дважды дифференцируемыми. Функция Лагранжа $L: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ задачи (28) имеет вид

$$L(x, \lambda) = f(x) + \lambda h(x). \quad (29)$$

Приводимые ниже предложения 2 и 3 полностью объясняют происходящее в примере 3, а также демонстрирует сложности такого анализа даже в таком простейшем специальном случае.

Предложение 2. Пусть функция $f: \mathbb{R} \rightarrow \mathbb{R}$ и отображение $h: \mathbb{R} \rightarrow \mathbb{R}$ дважды дифференцируемы в окрестности точки $\bar{x} \in \mathbb{R}$, причем их вторые производные непрерывны в этой точке. Предположим, что $f'(\bar{x}) = h(\bar{x}) = h'(\bar{x}) = 0$, $h''(\bar{x}) \neq 0$.

Тогда для любого $x_0 \in \mathbb{R} \setminus \{\bar{x}\}$, достаточно близкого к \bar{x} , и для любого $\lambda_0 \in \mathbb{R}$ существует единственная последовательность $\{(x_k, \lambda_k)\} \subset \mathbb{R} \times \mathbb{R}$ такая, что $(x_{k+1} - x_k, \lambda_{k+1} - \lambda_k)$ удовлетворяет системе (16) при $(x, \lambda) = (x_k, \lambda_k)$ для всех k ; эта последовательность сходится к $(\bar{x}, \bar{\lambda})$, где $\bar{\lambda} = -f''(\bar{x})/h''(\bar{x})$, причем $x_k \neq \bar{x}$ для всех k и

$$\lim_{k \rightarrow \infty} \frac{x_{k+1} - \bar{x}}{x_k - \bar{x}} = \frac{1}{2}. \quad (30)$$

Заметим, что из предположения $f'(\bar{x}) = h(\bar{x}) = h'(\bar{x}) = 0$ вытекает, что \bar{x} является стационарной точкой задачи (28), причем $\mathcal{M}(\bar{x}) = \mathbb{R}$. В то же время, из $h''(\bar{x}) \neq 0$ следует, что $\bar{\lambda} = -f''(\bar{x})/h''(\bar{x})$ является единственным критическим множителем Лагранжа.

Доказательство. Начнем с некоторых технических соотношений. Поскольку $h(\bar{x}) = h'(\bar{x}) = 0$, для $x \in \mathbb{R}$ выполняются оценки

$$\begin{aligned} h(x) &= \frac{1}{2}h''(\bar{x})(x - \bar{x})^2 + o((x - \bar{x})^2), \\ h'(x) &= h''(\bar{x})(x - \bar{x}) + o(x - \bar{x}), \end{aligned}$$

из которых для $x \neq \bar{x}$ вытекают соотношения

$$\lim_{x \rightarrow \bar{x}} \frac{h(x)}{(x - \bar{x})^2} = \frac{1}{2}h''(\bar{x}), \quad \lim_{x \rightarrow \bar{x}} \frac{h'(x)}{x - \bar{x}} = h''(\bar{x}). \quad (31)$$

Поскольку $h''(\bar{x}) \neq 0$, из этих соотношений следует, в частности, что $h'(x) \neq 0$ для всех x достаточно близких к \bar{x} , и

$$\lim_{x \rightarrow \bar{x}} \frac{h(x)}{h'(x)(x - \bar{x})} = \frac{1}{2}, \quad \lim_{x \rightarrow \bar{x}} \frac{h(x)}{(h'(x))^2} h''(x) = \frac{1}{2}. \quad (32)$$

Более того, по правилу Лопиталья

$$\lim_{x \rightarrow \bar{x}} \frac{f'(x)}{h'(x)} = \lim_{x \rightarrow \bar{x}} \frac{f''(x)}{h''(x)} = -\bar{\lambda}. \quad (33)$$

Пусть $(x_k, \lambda_k) \in \mathbb{R} \times \mathbb{R}$ — текущее приближение; тогда, согласно (16), следующее приближение (x_{k+1}, λ_{k+1}) должно быть решением системы

$$\begin{aligned} (f''(x_k) + \lambda_k h''(x_k))(x - x_k) + h'(x_k)(\lambda - \lambda_k) &= -f'(x_k) - \lambda_k h'(x_k), \\ h'(x_k)(x - x_k) &= -h(x_k). \end{aligned} \quad (34)$$

Если x_k достаточно близко к \bar{x} , второе равенство в (34) однозначно определяет

$$x_{k+1} = x_k - \frac{h(x_k)}{h'(x_k)}, \quad (35)$$

и, следовательно,

$$\frac{x_{k+1} - \bar{x}}{x_k - \bar{x}} = 1 - \frac{h(x_k)}{h'(x_k)(x_k - \bar{x})}. \quad (36)$$

Согласно первому соотношению в (32), отношение в правой части может быть сделано сколь угодно близким к $1/2$, если взять x_k достаточно близким к \bar{x} . Отсюда следует, что для всякого $x_0 \in \mathbb{R}$, достаточно близкого к \bar{x} , существует единственная последовательность $\{x_k\}$ такая, что для всякого k точка x_{k+1} удовлетворяет второму равенству в (34), эта последовательность сходится к \bar{x} с линейной скоростью, и, более того, выполняется (30).

Далее, из первого уравнения в (34) и из (35) следует, что λ_{k+1} должно удовлетворять уравнению

$$f'(x_k) + \lambda h'(x_k) = (f''(x_k) + \lambda_k h''(x_k)) \frac{h(x_k)}{h'(x_k)},$$

однозначно определяющему

$$\lambda_{k+1} = -\frac{f'(x_k)}{h'(x_k)} + \frac{h(x_k)}{(h'(x_k))^2} (f''(x_k) + \lambda_k h''(x_k)).$$

Следовательно,

$$\lambda_{k+1} - \bar{\lambda} = -\left(\frac{f'(x_k)}{h'(x_k)} + \bar{\lambda}\right) + \frac{h(x_k)}{(h'(x_k))^2} h''(x_k) \left(\frac{f''(x_k)}{h''(x_k)} + \bar{\lambda}\right) + \frac{h(x_k)}{(h'(x_k))^2} h''(x_k) (\lambda_k - \bar{\lambda}). \quad (37)$$

Используя второе соотношение в (32), а также (33), последнее равенство можно переписать в виде

$$\lambda_{k+1} - \bar{\lambda} = \left(\frac{1}{2} + \delta_k \right) (\lambda_k - \bar{\lambda}) + \varepsilon_k,$$

где последовательности $\{\varepsilon_k\}$ и $\{\delta_k\}$ сходятся к нулю. Применяя, например, [1, лемма 2.6.6], отсюда получаем сходимость $\{\lambda_k\}$ к $\bar{\lambda}$. ■

В предложении 2 установлена скорость сходимости лишь прямой траектории $\{x_k\}$. Предлагаемый ниже анализ скорости сходимости двойственной траектории $\{\lambda_k\}$ существенно сложнее. Для этого анализа потребуются следующие три технические леммы о числовых последовательностях.

Лемма 1. Пусть числовая последовательность $\{\alpha_k\}$ удовлетворяет для каждого k условию

$$\alpha_{k+1} = \theta_k \alpha_k + \varepsilon_k, \quad (38)$$

где числовая последовательность $\{\theta_k\}$ сходится к некоторому $\theta > 1$, а числовая последовательность $\{\varepsilon_k\}$ сходится к нулю.

Тогда либо $\alpha_k \rightarrow \infty$, либо $\alpha_k \rightarrow 0$.

Доказательство. Из (38) следует, что при всех k выполнено

$$|\alpha_{k+1}| - |\alpha_k| \geq (\theta_k - 1)|\alpha_k| - |\varepsilon_k|. \quad (39)$$

Предположим, что последовательность $\{\alpha_k\}$ не сходится к нулю, т.е. существует такое $\delta > 0$, что найдется сколь угодно большой номер k , для которого $|\alpha_k| > \delta$. Не ограничивая общности, будем считать, что $|\varepsilon_k| < \sigma\delta/2$, и что $\theta_k > 1 + \sigma$ при некотором $\sigma > 0$. Тогда из (39) следует выполнение

$$|\alpha_{k+1}| - |\alpha_k| > \sigma|\alpha_k| - \frac{\sigma\delta}{2}$$

для всех номеров k . Поэтому для всех k , для которых $|\alpha_k| > \delta$, будет выполнено

$$|\alpha_{k+1}| - |\alpha_k| > \frac{\sigma\delta}{2}$$

и, в частности, $|\alpha_{k+1}| > \delta$. Отсюда, очевидно, следует $\alpha_k \rightarrow \infty$. ■

Отметим, что в предположениях леммы 1 сходимость последовательности $\{\alpha_k\}$ к нулю возможна, но крайне нетипична. Действительно, для произвольной последовательности $\{\alpha_k\}$ с этими свойствами рассмотрим последовательность $\{\tilde{\alpha}_k\}$, также удовлетворяющую (38) при тех же $\{\theta_k\}$ и $\{\varepsilon_k\}$, но для которой $\tilde{\alpha}_0 = \alpha_0 + \delta$, где $\delta \neq 0$. Тогда для всякого k

$$\tilde{\alpha}_{k+1} = \alpha_{k+1} + \theta_k \cdots \theta_0 \delta,$$

и если $\theta_k \neq 0$ для всех k , то $\tilde{\alpha}_k \rightarrow \infty$. Таким образом, любое сколь угодно малое возмущение начального элемента последовательности $\{\alpha_k\}$ разрушает ее сходимость.

Лемма 2. Пусть числовая последовательность $\{\alpha_k\}$ удовлетворяет для каждого k условию

$$\alpha_{k+1} = \theta_k \alpha_k + 1, \quad (40)$$

где числовая последовательность $\{\theta_k\}$ сходится к некоторому $\theta > 0$.

Тогда справедливы следующие утверждения:

а) если $\theta < 1$, то

$$\alpha_k \rightarrow \frac{1}{1 - \theta}; \quad (41)$$

б) если $\theta = 1$, то $\alpha_k \rightarrow \infty$;

в) если $\theta > 1$, то либо $\alpha_k \rightarrow \infty$, либо выполнено (41).

Доказательство. Пусть сначала $\theta \neq 1$. Перепишем (40) в виде

$$\alpha_{k+1} - \frac{1}{1 - \theta} = \theta_k \left(\alpha_k - \frac{1}{1 - \theta} \right) + \frac{\theta_k - \theta}{1 - \theta}.$$

Последнее слагаемое в этом соотношении стремится к нулю. Поэтому, если $\theta < 1$, то для обоснования (41) достаточно снова применить [1, лемма 2.6.6]. Если же $\theta > 1$, то требуемое утверждение следует из леммы 1.

Пусть теперь $\theta = 1$. Предположим, что $\{\alpha_k\}$ не стремится к бесконечности, т. е. найдется бесконечный набор номеров $K \subset \{0, 1, \dots\}$ такой, что подпоследовательность $\{\alpha_k \mid k \in K\}$ сходится к некоторому числу α . Тогда для произвольного фиксированного номера m из (40) вытекает

$$\alpha_{k+m} = (\theta_{k+m-1} \cdots \theta_k) \alpha_k + (\theta_{k+m-1} \cdots \theta_{k+1}) + \dots + \theta_{k+m-1} + 1,$$

и поэтому последовательность $\{\alpha_{k+m} \mid k \in K\}$ сходится к $\alpha + m$. Последнее означает, что найдется неотрицательный элемент α_k со сколь угодно большим индексом k . Поскольку, не ограничивая общности, можно считать, что $\theta_k > 0$ для всех k , последнее означает, что начиная с некоторого номера k_1 все элементы последовательности $\{\alpha_k\}$ будут положительны.

Рассмотрим произвольное число $0 < \bar{\theta} < 1$. Поскольку $\theta_k \rightarrow 1$, то найдется номер $k_2 \geq k_1$ такой, что для всех $k \geq k_2$ выполнено $\theta_k > \bar{\theta}$. Учитывая, что при $k \geq k_2$ также выполнено $\alpha_k > 0$, из (40) получаем

$$\alpha_{k+1} > \bar{\theta} \alpha_k + 1$$

для таких k . Поэтому для произвольного фиксированного $k \geq k_2$ и любого достаточно большого номера m будет справедливо

$$\alpha_{k+m} > \bar{\theta}^m \alpha_k + \bar{\theta}^{m-1} + \dots + \bar{\theta} + 1 = \bar{\theta}^m \alpha_k + \frac{1 - \bar{\theta}^m}{1 - \bar{\theta}} > \frac{1}{2(1 - \bar{\theta})}.$$

Таким образом, начиная с некоторого номера k , все элементы последовательности $\{\alpha_k\}$ будут больше $1/(2(1 - \bar{\theta}))$. Учитывая произвольность выбора $\bar{\theta}$, получаем противоречие с тем, что $\{\alpha_k\}$ не стремится к бесконечности. ■

Как следует из доказательства леммы 2 и комментариев к лемме 1, если $\theta > 1$, то типичным является случай $\alpha_k \rightarrow \infty$.

Лемма 3. Пусть числовая последовательность $\{\alpha_k\}$ удовлетворяет для каждого k условию (38), где числовая последовательность $\{\theta_k\}$ сходится к некоторому $\theta \in (0, 1)$, а числовая последовательность $\{\varepsilon_k\}$ сходится к нулю, причем все элементы последовательности $\{\varepsilon_k\}$ с достаточно большими номерами имеют одинаковый знак, и

$$\lim_{k \rightarrow \infty} \frac{\varepsilon_{k+1}}{\varepsilon_k} = \gamma$$

при некотором $0 < \gamma < 1$.

Тогда все элементы последовательности $\{\alpha_k\}$ с достаточно большими номерами также имеют одинаковый знак, причем

$$\lim_{k \rightarrow \infty} \frac{\alpha_{k+1}}{\alpha_k} = \bar{\alpha}, \quad (42)$$

где $\bar{\alpha} \in \{\theta, \gamma\}$ и $\bar{\alpha} \geq \gamma$.

Доказательство. Без ограничения общности будем считать, что все элементы последовательности $\{\varepsilon_k\}$ имеют один знак. Перепишем (38) в виде

$$\frac{\alpha_{k+1}}{\varepsilon_k} = \beta_k \frac{\alpha_k}{\varepsilon_{k-1}} + 1,$$

где $\beta_k = \theta_k \varepsilon_{k-1} / \varepsilon_k$, и, следовательно, $\beta_k \rightarrow \theta / \gamma$. Таким образом, начиная с некоторого номера k будет выполнено $\beta_k > 0$, и поэтому последовательность $\{\alpha_{k+1} / \varepsilon_k\}$, а значит и $\{\alpha_k\}$, может менять знак лишь конечное число раз. Применяя к последовательности $\{\alpha_{k+1} / \varepsilon_k\}$ лемму 2, получаем, что либо $\alpha_{k+1} / \varepsilon_k \rightarrow \infty$, либо $\alpha_{k+1} / \varepsilon_k \rightarrow \gamma / (\gamma - \theta)$, причем если $\theta < \gamma$, то выполнено последнее. Для доказательства (42) остается переписать (38) в виде

$$\frac{\alpha_{k+1}}{\alpha_k} = \theta_k + \frac{\varepsilon_{k-1}}{\alpha_k} \frac{\varepsilon_k}{\varepsilon_{k-1}}$$

и перейти к пределу при $k \rightarrow \infty$. ■

Заметим, что в условиях леммы 3 типичным является выполнение (42) при $\bar{\alpha} = \max\{\theta, \gamma\}$. Если $\gamma \geq \theta$, то это следует из утверждения леммы 3. Если же $\theta > \gamma$, то это следует из доказательства леммы 3 и комментариев к лемме 2, поскольку при этом типичным является случай $\alpha_{k+1}/\varepsilon_k \rightarrow \infty$.

Перейдем к анализу скорости сходимости последовательности $\{\lambda_k\}$. Перепишем равенство (37) следующим образом:

$$\lambda_{k+1} - \bar{\lambda} = Q(x_k)(\lambda_k - \bar{\lambda}) + \frac{R(x_k)}{(h'(x_k))^2}, \quad (43)$$

где функции $Q, R: \mathbb{R} \rightarrow \mathbb{R}$ имеют вид

$$Q(x) = h''(x) \frac{h(x)}{(h'(x))^2},$$

$$R(x) = h(x) \frac{\partial^2 L}{\partial x^2}(x, \bar{\lambda}) - h'(x) \frac{\partial L}{\partial x}(x, \bar{\lambda}).$$

Как показано в доказательстве предложения 2,

$$\lim_{k \rightarrow \infty} Q(x_k) = \frac{1}{2}, \quad \lim_{k \rightarrow \infty} \frac{R(x_k)}{(h'(x_k))^2} = 0. \quad (44)$$

Если

$$R(\cdot) \equiv 0, \quad (45)$$

то $\lambda_{k+1} - \bar{\lambda} = Q(x_k)(\lambda_k - \bar{\lambda})$. Отсюда следует, что если $\lambda_0 \neq \bar{\lambda}$, то $\lambda_k \neq \bar{\lambda}$ для всех k , и скорость сходимости последовательности $\{\lambda_k\}$ является линейной:

$$\lim_{k \rightarrow \infty} \frac{\lambda_{k+1} - \bar{\lambda}}{\lambda_k - \bar{\lambda}} = \frac{1}{2}. \quad (46)$$

Условие (45), в частности, выполнено, если $\frac{\partial L}{\partial x}(\cdot, \bar{\lambda}) \equiv 0$ (как, например, в примере 3). Если же (45) не выполнено, то линейную скорость сходимости $\{\lambda_k\}$ можно установить при дополнительных предположениях о гладкости функции R .

Заметим предварительно, что из соотношений (31) и равенств $\frac{\partial L}{\partial x}(\bar{x}, \bar{\lambda}) = 0$, $\frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda}) = 0$ следует предельное соотношение

$$\lim_{x \rightarrow \bar{x}} \frac{R(x)}{(x - \bar{x})^2} = \lim_{x \rightarrow \bar{x}} \frac{h(x) \frac{\partial^2 L}{\partial x^2}(x, \bar{\lambda}) - h'(x) \frac{\partial L}{\partial x}(x, \bar{\lambda})}{(x - \bar{x})^2} = 0,$$

что означает, что функция R дважды дифференцируема в точке \bar{x} , причем ее значение и первые две производные в этой точке равны нулю.

Предложение 3. Пусть выполнены предположения предложения 2. Пусть, кроме того, для некоторых $\alpha_1, \alpha_2 > 0$ и $\gamma_1, \gamma_2 \neq 0$ выполнены условия

$$\lim_{x \rightarrow \bar{x}-0} \frac{h(x) \frac{\partial^2 L}{\partial x^2}(x, \bar{\lambda}) - h'(x) \frac{\partial L}{\partial x}(x, \bar{\lambda})}{(\bar{x} - x)^{2+\alpha_1}} = \gamma_1, \quad (47)$$

$$\lim_{x \rightarrow \bar{x}+0} \frac{h(x) \frac{\partial^2 L}{\partial x^2}(x, \bar{\lambda}) - h'(x) \frac{\partial L}{\partial x}(x, \bar{\lambda})}{(x - \bar{x})^{2+\alpha_2}} = \gamma_2, \quad (48)$$

где L определяется согласно (29).

Тогда для любого $x_0 \in \mathbb{R} \setminus \{\bar{x}\}$, достаточно близкого к \bar{x} , и для любого $\lambda_0 \in \mathbb{R}$, для единственной последовательности $\{(x_k, \lambda_k)\} \subset \mathbb{R} \times \mathbb{R}$ такой, что $(x_{k+1} - x_k, \lambda_{k+1} - \lambda_k)$ удовлетворяет системе (16) при $(x, \lambda) = (x_k, \lambda_k)$ для всех k , выполняется следующее: начиная с некоторого номера k , имеет место $\lambda_k \neq \bar{\lambda}$, где $\bar{\lambda} = -f''(\bar{x})/h''(\bar{x})$, причем

$$\lim_{k \rightarrow \infty} \frac{\lambda_{k+1} - \bar{\lambda}}{\lambda_k - \bar{\lambda}} = \frac{1}{2^\sigma}, \quad (49)$$

где $\sigma \in \{1, \alpha_1, \alpha_2\}$.

Заметим, что условия (47), (48) выполнены, в частности, если для некоторых $p, q \geq 3$ существуют отличные от нуля левая производная p -го порядка и правая производная q -го порядка функции R в точке \bar{x} . Действительно, для минимальных таких p и q будут справедливы оценки

$$\begin{aligned} R(x) &= \frac{1}{p!} R^{(p)}(\bar{x} - 0)(x - \bar{x})^p + o((x - \bar{x})^p), \quad x < \bar{x}, \\ R(x) &= \frac{1}{q!} R^{(q)}(\bar{x} + 0)(x - \bar{x})^q + o((x - \bar{x})^q), \quad x > \bar{x}, \end{aligned}$$

откуда следуют условия (47), (48) при $\alpha_1 = p - 2$, $\alpha_2 = q - 2$ и $\gamma_1 = (-1)^p R^{(p)}(\bar{x} - 0)/p!$, $\gamma_2 = R^{(q)}(\bar{x} + 0)/q!$.

Кроме того, выполнение (47), (48) гарантировано, если

$$\lim_{x \rightarrow \bar{x}-0} \frac{\frac{\partial^2 L}{\partial x^2}(x, \bar{\lambda})}{(\bar{x} - x)^{\alpha_1}} = \gamma_1, \quad \lim_{x \rightarrow \bar{x}+0} \frac{\frac{\partial^2 L}{\partial x^2}(x, \bar{\lambda})}{(x - \bar{x})^{\alpha_2}} = \gamma_2 \quad (50)$$

при некоторых $\alpha_i > 0$, $\alpha_i \neq 1$ и $\gamma_i \neq 0$, $i = 1, 2$. Действительно, в этом случае, согласно правилу Лопиталья,

$$\lim_{x \rightarrow \bar{x}-0} \frac{\frac{\partial L}{\partial x}(x, \bar{\lambda})}{(\bar{x} - x)^{1+\alpha_1}} = \lim_{x \rightarrow \bar{x}-0} \frac{-\frac{\partial^2 L}{\partial x^2}(x, \bar{\lambda})}{(1 + \alpha_1)(\bar{x} - x)^{\alpha_1}} = -\frac{\gamma_1}{1 + \alpha_1},$$

и поэтому, с учетом (31),

$$\lim_{x \rightarrow \bar{x}-0} \frac{R(x)}{(\bar{x} - x)^{2+\alpha_1}} = \gamma_1 h''(\bar{x}) \left(\frac{1}{2} - \frac{1}{1 + \alpha_1} \right).$$

Аналогично проверяется (48).

Доказательство. Как следует из предложения 2, $(x_k, \lambda_k) \rightarrow (\bar{x}, \bar{\lambda})$ при $k \rightarrow \infty$, и последовательность $\{\lambda_k\}$ удовлетворяет (43). Обозначим

$$\varepsilon_k = \frac{R(x_k)}{(h'(x_k))^2}.$$

Согласно второму соотношению в (44), $\varepsilon_k \rightarrow 0$.

Из (36) следует, что если x_0 достаточно близко к \bar{x} , то все элементы последовательности $\{x_k - \bar{x}\}$ имеют одинаковый знак. Для определенности будем считать, что $x_k < \bar{x}$ для всех k . Тогда из (47) следует, что элементы последовательности $\{\varepsilon_k\}$ с достаточно большими номерами имеют один знак. Вновь привлекая второе равенство из (31) и (47), получаем

$$\lim_{k \rightarrow \infty} \frac{\varepsilon_k}{(\bar{x} - x_k)^{\alpha_1}} = \lim_{k \rightarrow \infty} \frac{R(x_k)}{(\bar{x} - x_k)^{2+\alpha_1}} \left(\frac{x_k - \bar{x}}{h'(x_k)} \right)^2 = \frac{\gamma_1}{(h''(\bar{x}))^2}.$$

Отсюда и из (30) следует

$$\lim_{k \rightarrow \infty} \frac{\varepsilon_{k+1}}{\varepsilon_k} = \lim_{k \rightarrow \infty} \left(\frac{x_{k+1} - \bar{x}}{x_k - \bar{x}} \right)^{\alpha_1} = \frac{1}{2^{\alpha_1}}.$$

С учетом (43) и (44), применяя лемму 3 к последовательности $\{\lambda_k - \bar{\lambda}\}$, получаем утверждение теоремы при $\sigma \in \{1, \alpha_1\}$.

Если же $x_k > \bar{x}$ при всех k , то, повторяя аналогичные рассуждения, также получаем (49), в котором $\sigma \in \{1, \alpha_2\}$. ■

Заметим, что если $\lambda_0 \neq \bar{\lambda}$, то утверждение предложения 3 остается справедливым, если условие (47) заменить на $R(x) = 0$ при $x < \bar{\lambda}$, или если заменить (48) на $R(x) = 0$ при $x > \bar{\lambda}$. В этом случае доказательство предложения модифицируется очевидным образом, с учетом того, что вся последовательность $\{x_k\}$ лежит либо слева, либо справа от \bar{x} .

Отметим также, что значение σ в предельном соотношении (49) зависит от того, с какой стороны $\{x_k\}$ приближается к \bar{x} . Как следует из комментариев к лемме 3, если $x_k < \bar{x}$ при всех k , то типичным является значение $\sigma = \min\{1, \alpha_1\}$, а если $x_k > \bar{x}$, то $\sigma = \min\{1, \alpha_2\}$.

Приводимый ниже пример, являющийся более общей версией примера 3, демонстрирует возможные сценарии поведения МНЛ, в том числе в ситуациях, когда какие-то из предположений предложения 2 нарушаются.

Пример 5. Пусть $f(x) = ax^p$, $h(x) = x^q$, где $p \geq 2$ и $q \geq 2$, причем p и q таковы, что в степени p и q можно возводить отрицательные числа. А именно, будем считать, что оба числа либо являются целыми, либо представляют собой рациональные дроби, в которых знаменатель является нечетным.

Тогда $\bar{x} = 0$ — единственное решение задачи (28), причем $\mathcal{M}(0) = \mathbb{R}$. Далее, если $q = 2$, то $h''(0) \neq 0$, и, следовательно, существует единственный критический множитель Лагранжа, равный $-a$ при $p = 2$ и нулю при $p > 2$. Если же $q > 2$, то при $p = 2$ критических множителей нет, а при $p > 2$, напротив, все множители Лагранжа являются критическими.

Заметим, что при $q = 2$ выполнены условия предложения 2. Дополнительно, если $p = 2$, то $\frac{\partial L}{\partial x}(\cdot, -a) \equiv 0$, а если $p = 3$, то $R(\cdot) \equiv 0$, и поэтому в обоих случаях должно выполняться

(46). Если же $p > 2$, $p \neq 3$, то выполнены условия предположения 3 при $\alpha_1 = \alpha_2 = p - 2$, $\gamma_2 = ap(p - 3)$ и $\gamma_1 = (-1)^p \gamma_2$, вследствие чего должно выполняться (49) при $\sigma \in \{1, p - 2\}$.

Согласно (16), имеем:

$$\begin{aligned} (ap(p-1)x_k^{p-2} + q(q-1)\lambda_k x_k^{q-2})(x_{k+1} - x_k) + qx_k^{q-1}(\lambda_{k+1} - \lambda_k) &= -apx_k^{p-1} - qx_k^{q-1}, \\ qx_k^{q-1}(x_{k+1} - x_k) &= -x_k^q \end{aligned}$$

для каждого k . Предполагая, что $x_k \neq 0$, из второго уравнения получаем

$$x_{k+1} = \left(1 - \frac{1}{q}\right) x_k, \quad (51)$$

и тогда из первого уравнения, после несложных преобразований,

$$\lambda_{k+1} = \left(1 - \frac{1}{q}\right) \lambda_k - \frac{ap(q-p+1)}{q^2} x_k^{p-q}. \quad (52)$$

Таким образом, последовательность $\{x_k\}$ сходится к 0 с линейной скоростью, но $1/2$ в правой части (30) нужно заменить на $1 - 1/q \geq 1/2$.

Поведение двойственной траектории зависит от значений обоих параметров p и q . Если $p < q$, то $\lambda_k \rightarrow \infty$. Действительно, в этом случае уравнения (51) и (52) имеют вид

$$x_{k+1} = \theta x_k, \quad \lambda_{k+1} = \theta \lambda_k + \gamma x_k^{-\alpha}, \quad (53)$$

где $0 < \theta < 1$, $\gamma \neq 0$ и $\alpha > 0$. Отсюда следует, что

$$\lambda_{k+1} x_{k+1}^\alpha = \theta^{1+\alpha} \lambda_k x_k^\alpha + \gamma \theta^\alpha. \quad (54)$$

Поскольку $0 < \theta^{1+\alpha} < 1$, то из последнего условия легко получить $\lambda_k x_k^\alpha \rightarrow \gamma \theta^\alpha / (1 - \theta^{1+\alpha})$, откуда, в свою очередь, следует $\lambda_k \rightarrow \infty$.

Если $p = q$, то равенство (52) можно переписать в виде

$$\lambda_{k+1} + a = \left(1 - \frac{1}{q}\right) (\lambda_k + a).$$

Отсюда следует, что $\{\lambda_k\}$ сходится к $\bar{\lambda} = -a$, причем, если $\lambda_0 \neq \bar{\lambda}$, то $\lambda_k \neq -a$ для всех k , и

$$\lim_{k \rightarrow \infty} \frac{\lambda_{k+1} - \bar{\lambda}}{\lambda_k - \bar{\lambda}} = 1 - \frac{1}{q}. \quad (55)$$

Если же $p > q$, то, применяя [1, лемма 2.6.6], получаем, что $\{\lambda_k\}$ сходится к 0. Очевидно, что если при этом $p = q + 1$, то (52) принимает вид

$$\lambda_{k+1} = \left(1 - \frac{1}{q}\right) \lambda_k,$$

и поэтому, если $\lambda_0 \neq 0$, то $\lambda_k \neq 0$ для всех k , и выполняется (55) при $\bar{\lambda} = 0$. Для того, чтобы установить скорость сходимости при $p \neq q + 1$, вновь перепишем (51) и (52) в виде (53), где

по-прежнему $0 < \theta < 1$ и $\gamma \neq 0$, но теперь $\alpha < 0$ и $\alpha \neq -1$. Отсюда вновь следует равенство (54), которое можно переписать следующим образом:

$$\lambda_{k+1}x_{k+1}^\alpha - \omega = \theta^{1+\alpha}(\lambda_k x_k^\alpha - \omega),$$

где $\omega = \gamma\theta^\alpha/(1 - \theta^{1+\alpha})$. Предполагая $\lambda_0 x_0^\alpha \neq \omega$, из последнего равенства получаем, что при $\alpha < -1$ выполняется $\lambda_k x_k^\alpha \rightarrow \infty$, а при $-1 < \alpha < 0$ выполняется $\lambda_k x_k^\alpha \rightarrow \omega$. Очевидно, что в обоих случаях $\lambda_k \neq 0$ для любого достаточно большого k . Для таких k из (53) получаем

$$\frac{\lambda_{k+1}}{\lambda_k} = \theta + \frac{\gamma}{\lambda_k x_k^\alpha},$$

откуда следует, что $\lambda_{k+1}/\lambda_k \rightarrow \theta$ при $\alpha < -1$, и $\lambda_{k+1}/\lambda_k \rightarrow \theta + \gamma/\omega = 1/\theta^\alpha$ при $-1 < \alpha < 0$. Таким образом, при $p > q$ выполняется

$$\lim_{k \rightarrow \infty} \frac{\lambda_{k+1}}{\lambda_k} = \left(1 - \frac{1}{q}\right)^{\min\{1, p-q\}}.$$

1.2.2. Чисто квадратичные задачи

Будем рассматривать следующую чисто квадратичную задачу оптимизации с ограничениями равенствами:

$$\frac{1}{2}\langle Ax, x \rangle \rightarrow \min, \quad \frac{1}{2}B[x, x] = 0, \quad (56)$$

где $A \in \mathbb{R}^{n \times n}$ — симметричная матрица, а $B: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^l$ — симметричное билинейное отображение (т.е. $B[x, x] = (\langle B_1 x, x \rangle, \dots, \langle B_l x, x \rangle)$, где $B_1, \dots, B_l \in \mathbb{R}^{n \times n}$ — симметричные матрицы).

Для всех $\lambda \in \mathbb{R}^l$ определим симметричную матрицу $H(\lambda) = A + \lambda B$, где $\lambda B = \sum_{i=1}^l \lambda_i B_i$. Стационарные точки задачи (56) и отвечающие им множители Лагранжа характеризуются системой Лагранжа

$$H(\lambda)x = 0, \quad \frac{1}{2}B[x, x] = 0 \quad (57)$$

относительно $(x, \lambda) \in \mathbb{R}^n \times \mathbb{R}^l$. В данном разделе будем исследовать поведение МНЛ в окрестности точки $\bar{x} = 0$. Данная точка является стационарной, причем множество отвечающих ей множителей Лагранжа есть \mathbb{R}^l . Как следует из определения, критическими являются множители, удовлетворяющие $\det H(\lambda) = 0$.

Итерация метода Ньютона для системы (57) определяется следующим образом: для текущего прямо-двойственного приближения $(x^k, \lambda^k) \in \mathbb{R}^n \times \mathbb{R}^l$ следующие приближение (x^{k+1}, λ^{k+1}) вычисляется как решение линейной системы

$$\begin{pmatrix} H(\lambda^k) & (B[x^k])^T \\ B[x^k] & 0 \end{pmatrix} \begin{pmatrix} x^{k+1} - x^k \\ \lambda^{k+1} - \lambda^k \end{pmatrix} = - \begin{pmatrix} H(\lambda^k)x^k \\ \frac{1}{2}B[x^k, x^k] \end{pmatrix}, \quad (58)$$

где для произвольного $\xi \in \mathbb{R}^n$ матрица $B[\xi] \in \mathbb{R}^{l \times n}$ определяется соотношением $B[\xi]x = B[\xi, x]$.

Далее будем использовать соглашение, что определитель пустой матрицы равен 1. Отсюда, в частности, следует, что определитель присоединенной матрицы для любой матрицы размера 1×1 также равен 1. Данное соглашение необходимо, чтобы избежать отдельного рассмотрения случая $n = 1$.

Равенство (58) можно переписать в виде

$$H(\lambda^k)x^{k+1} + (B[x^k])^T(\lambda^{k+1} - \lambda^k) = 0, \quad B[x^k, x^{k+1}] = \frac{1}{2}B[x^k, x^k]. \quad (59)$$

Предположим, что матрица $H(\lambda^k)$ невырождена. В этом случае из первого равенства получаем

$$x^{k+1} = -(H(\lambda^k))^{-1}(B[x^k])^T(\lambda^{k+1} - \lambda^k). \quad (60)$$

Подставляя полученное выражение во второе равенство в (59), находим

$$-B[x^k](H(\lambda^k))^{-1}(B[x^k])^T(\lambda^{k+1} - \lambda^k) = \frac{1}{2}B[x^k, x^k].$$

Если матрица $B[x^k](H(\lambda^k))^{-1}(B[x^k])^T$ невырождена (и, в частности, $x^k \neq 0$), то из последнего равенства получаем

$$\lambda^{k+1} - \lambda^k = -\frac{1}{2}(B[x^k](H(\lambda^k))^{-1}(B[x^k])^T)^{-1}B[x^k, x^k]. \quad (61)$$

Введем дополнительные обозначения. Для произвольных $\xi \in \mathbb{R}^n$ и $\lambda \in \mathbb{R}^l$ положим

$$\delta(\xi, \lambda) = (B[\xi](H(\lambda))^{-1}(B[\xi])^T)^{-1}B[\xi, \xi], \quad P(\xi, \lambda) = (H(\lambda))^{-1}(\delta(\xi, \lambda)B). \quad (62)$$

В этом случае условия (60) и (61) можно переписать в следующем виде:

$$x^{k+1} = \frac{1}{2}P(\xi^k, \lambda^k)x^k, \quad \lambda^{k+1} = \lambda^k - \frac{\delta(\xi^k, \lambda^k)}{2}, \quad (63)$$

где $\xi^k = x^k/\|x^k\|$.

Обозначим через \mathcal{M}_0 множество всех критических множителей:

$$\mathcal{M}_0 = \{\lambda \in \mathbb{R}^l \mid \Delta(\lambda) = 0\},$$

где $\Delta: \mathbb{R}^l \rightarrow \mathbb{R}$, $\Delta(\lambda) = \det H(\lambda)$. В данном разделе будет исследоваться поведение МНЛ (58) в окрестности некоторого критического множителя $\bar{\lambda} \in \mathcal{M}_0$, для которого выполнено

$$\dim \ker H(\bar{\lambda}) = 1. \quad (64)$$

Основную идею проводимого ниже анализа можно неформально описать следующим образом. Будет показано, что если λ^k лежит достаточно близко к некоторому критическому

множителю $\bar{\lambda} \in \mathcal{M}_0$, и при этом ξ^k достаточно близко к $\bar{\xi} \in \ker H(\bar{\lambda})$, то величина $P(\xi^k, \lambda^k)\xi^k$ близка к ξ^k , а величина $\delta(\xi^k, \lambda^k)$ близка к $\lambda^k - \tilde{\lambda}$ для некоторого множителя $\tilde{\lambda} \in \mathcal{M}_0$. Таким образом, условия (63) означают выполнение следующих приближительных соотношений:

$$x^{k+1} \approx \frac{1}{2}x^k, \quad \lambda^{k+1} - \tilde{\lambda} \approx \frac{1}{2}(\lambda^k - \tilde{\lambda}).$$

Неформально говоря, данные соотношения означают, что x^{k+1} приблизительно в два раза «ближе» к 0, чем x^k , а λ^{k+1} приблизительно в два раза «ближе» к множеству \mathcal{M}_0 , чем λ^k .

Для начала докажем следующую простую лемму.

Лемма 4. Пусть для некоторого критического множителя $\bar{\lambda} \in \mathcal{M}_0$ выполнено (64). Тогда для произвольного вектора $\bar{\xi} \in \ker H(\bar{\lambda})$ единичной нормы найдется константа $\rho > 0$ и аналитическое (в частности, дифференцируемое бесконечное число раз) отображение $\xi: \mathcal{B}(\bar{\lambda}, \rho) \rightarrow \mathbb{R}^n$, обладающее следующим свойством:

- а) $\xi(\bar{\lambda}) = \bar{\xi}$;
- б) $\|\xi(\lambda)\| = 1$ для всех $\lambda \in \mathcal{B}(\bar{\lambda}, \rho)$;
- в) множество $\ker H(\lambda)$ совпадает с линейной оболочкой вектора $\xi(\lambda)$ для всех $\lambda \in \mathcal{M}_0 \cap \mathcal{B}(\bar{\lambda}, \rho)$.

Доказательство. Зафиксируем вектор $\bar{\xi} \in \ker H(\bar{\lambda})$ единичной нормы и рассмотрим параметрическую линейную систему

$$H(\lambda)\xi - \tau\bar{\xi} = 0, \quad \langle \bar{\xi}, \xi \rangle = 1 \quad (65)$$

относительно $(\xi, \tau) \in \mathbb{R}^n \times \mathbb{R}$, в которой $\lambda \in \mathbb{R}^l$ является параметром. Очевидно, что при $\lambda = \bar{\lambda}$ система имеет решение $(\bar{\xi}, 0)$, и при этом матрица системы невырождена. Действительно, если равенства

$$H(\bar{\lambda})\xi - \tau\bar{\xi} = 0, \quad \langle \bar{\xi}, \xi \rangle = 0 \quad (66)$$

выполнены для некоторого (ξ, τ) , то, умножая первое из них скалярно на $\bar{\xi}$, получаем

$$0 = \langle H(\bar{\lambda})\xi, \bar{\xi} \rangle - \tau\|\bar{\xi}\|^2 = \langle \xi, H(\bar{\lambda})\bar{\xi} \rangle - \tau = -\tau,$$

откуда следует $\tau = 0$. Таким образом, из первого равенства в (66) следует $\xi \in \ker H(\bar{\lambda})$, что, с учетом (64), означает, что второе равенство в (66) может выполняться только в случае $\xi = 0$.

Очевидно, для любого $\lambda \in \mathbb{R}^l$, достаточно близкого к $\bar{\lambda}$, матрица системы (65) также будет невырождена, откуда следует, что система имеет единственное решение $(\xi(\lambda), \tau(\lambda))$.

Далее, из определения множества \mathcal{M}_0 следует, что для любого $\lambda \in \mathcal{M}_0$, первое уравнение в (65) имеет решение вида $(\xi, 0)$ для некоторого $\xi \in \mathbb{R}^n \setminus \{0\}$. Заметим, что если множитель λ достаточно близок к $\bar{\lambda}$, то $\langle \bar{\xi}, \xi \rangle \neq 0$. Действительно, предположим, что это не так. В этом случае существуют последовательности $\{\lambda^k\} \subset \mathcal{M}_0$ и $\{\xi^k\} \subset \mathbb{R}^n \setminus \{0\}$, для которых выполнено $\{\lambda^k\} \rightarrow \bar{\lambda}$, а также $\xi^k \in \ker H(\lambda^k)$ и $\langle \bar{\xi}, \xi^k \rangle = 0$ для всех k . Тогда для любой предельной точки $\tilde{\xi} \in \mathbb{R}^n$ последовательности $\{\xi^k / \|\xi^k\|\}$ выполнено $\tilde{\xi} \in \ker H(\bar{\lambda})$, $\|\tilde{\xi}\| = 1$ и $\langle \bar{\xi}, \tilde{\xi} \rangle = 0$, что противоречит (64). Таким образом, вектор $(\xi / \langle \bar{\xi}, \xi \rangle, 0)$ удовлетворяет обоим условиям в (65).

Таким образом, для любого $\lambda \in \mathcal{M}_0$, достаточно близкого к $\bar{\lambda}$, для единственного решения $(\xi(\lambda), \tau(\lambda))$ системы (65) обязательно выполнено $\tau(\lambda) = 0$, откуда следует $\xi(\lambda) \in \ker H(\lambda)$.

Наконец, применяя классическую теорему о неявной функции к системе (65) в точке $(\xi, \tau) = (\bar{\xi}, 0)$ и при значении параметра, равном $\bar{\lambda}$, получаем, что отображение $(\xi(\cdot), \tau(\cdot))$ является аналитическим в некоторой окрестности $\bar{\lambda}$. В частности, $\|\xi(\lambda)\| \rightarrow \|\bar{\xi}\| = 1$ при $\lambda \rightarrow \bar{\lambda}$. Таким образом, заменив $\xi(\cdot)$ на $\xi(\cdot) / \|\xi(\cdot)\|$, получаем корректно определенное отображение, которое является аналитическим в окрестности $\bar{\lambda}$ и удовлетворяет $H(\lambda)\xi(\lambda) = 0$ для всех $\lambda \in \mathcal{M}_0$, достаточно близких к $\bar{\lambda}$, и при этом $\|\xi(\lambda)\| = 1$ для всех λ в окрестности $\bar{\lambda}$. Остается заметить, что $\dim \ker H(\lambda) = 1$ для всех таких $\lambda \in \mathcal{M}_0$ (что следует, например, из непрерывности собственных значений симметричных матриц). Таким образом, построенное отображение $\xi(\cdot)$ обладает всеми необходимыми свойствами. ■

Очевидно, что из условия (64) следует, что ядро матрицы $H(\bar{\lambda})$ содержит ровно два различных вектора единичной нормы. В дальнейшем анализе зафиксируем один из этих векторов, обозначим его через $\bar{\xi}$ и определим отображение $\xi(\cdot)$ согласно предыдущей лемме. Подчеркнем, что проводимый анализ не зависит от выбора $\bar{\xi}$.

Для произвольной квадратной матрицы H обозначим через \hat{H} ее присоединенную матрицу. Напомним, что для любой квадратной матрицы H справедливо тождество

$$\hat{H}H = H\hat{H} = I \det H. \quad (67)$$

Лемма 5. Пусть условие (64) выполнено для некоторого $\bar{\lambda} \in \mathcal{M}_0$. Тогда найдется $\rho > 0$ и аналитическая функция $t: \mathcal{B}(\bar{\lambda}, \rho) \rightarrow \mathbb{R} \setminus \{0\}$, удовлетворяющая соотношению

$$\hat{H}(\lambda) = t(\lambda)\xi(\lambda)(\xi(\lambda))^T \quad \forall \lambda \in \mathcal{M}_0 \cap \mathcal{B}(\bar{\lambda}, \rho), \quad (68)$$

где отображение $\xi(\cdot)$ определено согласно лемме 4, и $\xi(\lambda)$ рассматривается как вектор-столбец.

Доказательство. Как следует из тождества (67), для всех критических множителей $\lambda \in$

\mathcal{M}_0 выполнено $H(\lambda)\hat{H}(\lambda) = 0$, и поэтому

$$\text{im } \hat{H}(\lambda) \subset \ker H(\lambda). \quad (69)$$

Отсюда, в частности, следует, что $\text{rank } \hat{H}(\lambda) \leq 1$ при условии, что λ достаточно близко к $\bar{\lambda}$, поскольку $\dim \ker H(\lambda) = 1$ для таких λ . Более того, из последнего условия также следует существование ненулевого минора размера $(n-1) \times (n-1)$ у матрицы $H(\lambda)$. Поэтому у матрицы $\hat{H}(\lambda)$ имеется ненулевой элемент, откуда следует $\text{rank } \hat{H}(\lambda) = 1$. С учетом данного условия, из (69) следует, что для всех $\lambda \in \mathcal{M}_0$, достаточно близких к $\bar{\lambda}$, выполнено

$$\text{im } \hat{H}(\lambda) = \ker H(\lambda). \quad (70)$$

С учетом определения $\xi(\cdot)$, отсюда вытекает существование разложения $\hat{H}(\lambda) = \xi(\lambda)(\zeta(\lambda))^T$ при некотором $\zeta(\lambda) \in \mathbb{R}^n$. Из последнего равенства, с учетом симметричности матрицы $\hat{H}(\lambda)$, получаем также равенство $\hat{H}(\lambda) = \zeta(\lambda)(\xi(\lambda))^T$, из которого, согласно (70), следует $\zeta(\lambda) \in \ker H(\lambda)$. Таким образом, $\zeta(\lambda) = t(\lambda)\xi(\lambda)$ для некоторого действительного числа $t(\lambda)$, что, в свою очередь, означает, что $\hat{H}(\lambda) = t(\lambda)\xi(\lambda)(\xi(\lambda))^T$. В частности,

$$t(\lambda) = t(\lambda)\|\xi(\lambda)\|^2 = \langle \hat{H}(\lambda)\xi(\lambda), \xi(\lambda) \rangle,$$

что означает выполнение (68) для функции $t: \mathcal{B}(\bar{\lambda}, \rho) \rightarrow \mathbb{R}$, определенной согласно последнему соотношению при достаточно малом $\rho > 0$.

Для завершения доказательства остается заметить, что функция t является аналитической в окрестности $\mathcal{B}(\bar{\lambda}, \rho)$ (поскольку отображение $\xi(\cdot)$ также аналитично в этой окрестности, а $\hat{H}(\cdot)$ содержит полиномиальные элементы), а также что $t(\bar{\lambda}) \neq 0$ (поскольку в противном случае было бы выполнено $\hat{H}(\bar{\lambda}) = 0$, что противоречит (64)). ■

Лемма 6. Пусть условие (64) выполнено для некоторого $\bar{\lambda} \in \mathcal{M}_0$. Тогда для некоторого $\rho > 0$ справедливо разложение

$$\Delta'(\lambda) = t(\lambda)B[\xi(\lambda), \xi(\lambda)] \quad \forall \lambda \in \mathcal{M}_0 \cap \mathcal{B}(\bar{\lambda}, \rho), \quad (71)$$

где отображение $\xi(\cdot)$ определено согласно лемме 4, а функция $t(\cdot)$ определена согласно лемме 5.

Доказательство. Вновь применяя тождество (67), находим, что для любого $\lambda \in \mathcal{M}_0 \cap \mathcal{B}(\bar{\lambda}, \rho)$ при достаточно малом $\rho > 0$ и для любого $\delta \in \mathbb{R}^l$ справедлива следующая цепочка равенств:

$$\begin{aligned} \Delta(\lambda + \delta) &= \|\xi(\lambda)\|^2 \Delta(\lambda + \delta) = \left\langle \hat{H}(\lambda + \delta)\xi(\lambda), H(\lambda + \delta)\xi(\lambda) \right\rangle = \\ &= \left\langle \hat{H}(\lambda + \delta)\xi(\lambda), H(\lambda)\xi(\lambda) + (H(\lambda + \delta) - H(\lambda))\xi(\lambda) \right\rangle = \left\langle \hat{H}(\lambda + \delta)\xi(\lambda), (B[\xi(\lambda)])^T \delta \right\rangle \end{aligned}$$

(здесь также использовались определения отображений $H(\cdot)$ и $\xi(\cdot)$). Таким образом,

$$\begin{aligned} \Delta(\lambda + \delta) - \Delta(\lambda) - \langle t(\lambda)B[\xi(\lambda), \xi(\lambda)], \delta \rangle = \\ \langle \hat{H}(\lambda + \delta)\xi(\lambda), (B[\xi(\lambda)])^T \delta \rangle - \langle t(\lambda)B[\xi(\lambda), \xi(\lambda)], \delta \rangle = \\ = \langle B[\xi(\lambda), \hat{H}(\lambda + \delta)\xi(\lambda) - t(\lambda)\xi(\lambda)], \delta \rangle. \end{aligned} \quad (72)$$

Далее, согласно (68),

$$\hat{H}(\lambda + \delta)\xi(\lambda) - t(\lambda)\xi(\lambda) \rightarrow t(\lambda)(\xi(\lambda)\|\xi(\lambda)\|^2 - \xi(\lambda)) = 0 \text{ as } \delta \rightarrow 0,$$

и поэтому правая часть в (72) есть $o(\|\delta\|)$, откуда следует (71). \blacksquare

Лемма 7. Для любых матриц $P, Q \in \mathbb{R}^{n \times s}$, где $s \leq n$, и любой невырожденной матрицы $H \in \mathbb{R}^{n \times n}$ выполнено соотношение

$$\det(P^T H^{-1} Q) \det H = \sum_{\mathcal{I} \in \mathcal{S}_s} \det S(P, Q, H; \mathcal{I}), \quad (73)$$

где \mathcal{S}_s представляет собой совокупность всех подмножеств длины s множества $\{1, \dots, n\}$, и для любого $\mathcal{I} \in \mathcal{S}_s$ матрица $S(P, Q, H; \mathcal{I}) \in \mathbb{R}^{n \times n}$ состоит из строк следующего вида:

$$(S(P, Q, H; \mathcal{I}))_i = \begin{cases} (PQ^T)_i & \text{если } i \in \mathcal{I}, \\ H_i & \text{если } i \in \{1, \dots, n\} \setminus \mathcal{I}. \end{cases}$$

Доказательство. Если $s = n$, то соотношение (73), очевидно, выполнено, поскольку

$$\det(P^T H^{-1} Q) \det H = \det P \det Q = \det(PQ^T),$$

и единственным элементом \mathcal{S}_s является множество $\{1, \dots, n\}$.

Пусть теперь $s < n$. Дважды применяя формулу Бине–Коши, находим

$$\det(P^T H^{-1} Q) = \sum_{\mathcal{I} \in \mathcal{S}_s} \det P_{\mathcal{I}} \det (H^{-1} Q^T)_{\mathcal{I}} = \sum_{\mathcal{I}, \mathcal{J} \in \mathcal{S}_s} \det P_{\mathcal{I}} \det (H^{-1})_{\mathcal{I}, \mathcal{J}} \det Q_{\mathcal{J}}.$$

Далее, применяя формулу для минора обратной матрицы (см., например, [2, пункт 4.70]), получаем

$$\det(H^{-1})_{\mathcal{I}, \mathcal{J}} = (-1)^{\sigma(\mathcal{I}) + \sigma(\mathcal{J})} \frac{\det H_{\hat{\mathcal{I}}, \hat{\mathcal{J}}}}{\det H},$$

где

$$\sigma(\mathcal{I}) = \sum_{i \in \mathcal{I}} i, \quad \hat{\mathcal{I}} = \{1, \dots, n\} \setminus \mathcal{I}, \quad \hat{\mathcal{J}} = \{1, \dots, n\} \setminus \mathcal{J}.$$

Таким образом,

$$\begin{aligned} \det(P^T H^{-1} Q) \det H &= \sum_{\mathcal{I}, \mathcal{J} \in \mathcal{S}_s} (-1)^{\sigma(\mathcal{I}) + \sigma(\mathcal{J})} \det P_{\mathcal{I}} \det H_{\hat{\mathcal{I}}, \hat{\mathcal{J}}} \det Q_{\mathcal{J}} = \\ &= \sum_{\mathcal{I}, \mathcal{J} \in \mathcal{S}_s} (-1)^{\sigma(\mathcal{I}) + \sigma(\mathcal{J})} \det(PQ^T)_{\mathcal{I}, \mathcal{J}} \det H_{\hat{\mathcal{I}}, \hat{\mathcal{J}}} = \sum_{\mathcal{I} \in \mathcal{S}_s} \det S(P, Q, H; \mathcal{I}), \end{aligned}$$

где последнее равенство следует из разложения Лапласа для определителя. \blacksquare

Используя обозначения из леммы 7, определим функцию $R: \mathbb{R}^{l \times n} \times \mathbb{R}^{l \times n} \times \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$,

$$R(M, N, H) = \sum_{\mathcal{I} \in \mathcal{S}_l} \det S(M^{\mathbb{T}}, N^{\mathbb{T}}, H; \mathcal{I}).$$

Очевидно, эта функция полиномиальна относительно элементов матриц M , N и H . Для следующей леммы понадобится новое предположение:

$$R(B[\bar{\xi}], B[\bar{\xi}], H(\bar{\lambda})) \neq 0. \quad (74)$$

Легко убедиться, что из данного свойства, в частности, следует

$$\text{rank } B[\bar{\xi}] = l. \quad (75)$$

Лемма 8. Пусть условие (64) выполнено для некоторого $\bar{\lambda} \in \mathcal{M}_0$. Пусть также выполнено (74) для некоторого вектора $\bar{\xi} \in \ker H(\bar{\lambda})$ единичной нормы.

Тогда найдется $\rho > 0$ такое, что матрица $B[\xi](H(\lambda))^{-1}(B[\xi])^{\mathbb{T}}$ невырождена для всех $\xi \in \mathcal{B}(\bar{\xi}, \rho)$ и всех $\lambda \in \mathcal{B}(\bar{\lambda}, \rho) \setminus \mathcal{M}_0$. Более того, существует аналитическое отображение $M: \mathcal{B}(\bar{\xi}, \rho) \times \mathcal{B}(\bar{\lambda}, \rho) \rightarrow \mathbb{R}^{l \times l}$, удовлетворяющее условию

$$M(\xi, \lambda) = (B[\xi](H(\lambda))^{-1}(B[\xi])^{\mathbb{T}})^{-1} \quad \forall \xi \in \mathcal{B}(\bar{\xi}, \rho), \forall \lambda \in \mathcal{B}(\bar{\lambda}, \rho) \setminus \mathcal{M}_0.$$

Доказательство. Если $\rho > 0$ достаточно мало, то для любых $\xi \in \mathcal{B}(\bar{\xi}, \rho)$ и $\lambda \in \mathcal{B}(\bar{\lambda}, \rho) \setminus \mathcal{M}_0$ невырожденность матрицы $B[\xi](H(\lambda))^{-1}(B[\xi])^{\mathbb{T}}$ мгновенно следует из леммы 7, предположения (74) и непрерывности функции R .

Предположим сначала, что $l \geq 2$. В дополнение к R , определим также функцию $\tilde{R}: \mathbb{R}^{(l-1) \times n} \times \mathbb{R}^{(l-1) \times n} \times \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$,

$$\tilde{R}(\tilde{M}, \tilde{N}, H) = \sum_{\mathcal{I} \in \mathcal{S}_{l-1}} \det S(\tilde{M}^{\mathbb{T}}, \tilde{N}^{\mathbb{T}}, H; \mathcal{I}).$$

Очевидно, эта функция также полиномиальна относительно элементов матриц \tilde{M} , \tilde{N} и H .

Обозначим $\mathcal{I} = \{1, \dots, l\}$. Применяя правило Крамера для вычисления обратной матрицы, получаем для произвольной пары индексов $i, j \in \mathcal{I}$ следующую цепочку равенств:

$$\begin{aligned} \left((B[\xi](H(\lambda))^{-1}(B[\xi])^{\mathbb{T}})^{-1} \right)_{ij} &= (-1)^{i+j} \frac{\det((B[\xi])_{\mathcal{I} \setminus \{j\}}(H(\lambda))^{-1}((B[\xi])_{\mathcal{I} \setminus \{i\}})^{\mathbb{T}})}{\det(B[\xi](H(\lambda))^{-1}(B[\xi])^{\mathbb{T}})} = \\ &= (-1)^{i+j} \frac{\det((B[\xi])_{\mathcal{I} \setminus \{j\}}(H(\lambda))^{-1}((B[\xi])_{\mathcal{I} \setminus \{i\}})^{\mathbb{T}}) \Delta(\lambda)}{\det(B[\xi](H(\lambda))^{-1}(B[\xi])^{\mathbb{T}}) \Delta(\lambda)} = \\ &= (-1)^{i+j} \frac{\tilde{R}((B[\xi])_{\mathcal{I} \setminus \{j\}}, (B[\xi])_{\mathcal{I} \setminus \{i\}}, H(\lambda))}{R(B[\xi], B[\xi], H(\lambda))}, \end{aligned}$$

где последнее равенство следует из леммы 7, если применить ее и к числителю, и к знаменателю.

Далее, из предположения (74) и непрерывности R следует, что для любого вектора $(\xi, \lambda) \in \mathbb{R}^n \times \mathbb{R}^l$, достаточно близкого к $(\bar{\xi}, \bar{\lambda})$, выполнено $R(B[\xi], B[\xi], H(\lambda)) \neq 0$, и поэтому можно определить матрицу $M(\xi, \lambda) \in \mathbb{R}^{l \times l}$ следующим образом:

$$(M(\xi, \lambda))_{ij} = (-1)^{i+j} \frac{\tilde{R}((B[\xi])_{\mathcal{I} \setminus \{j\}}, (B[\xi])_{\mathcal{I} \setminus \{i\}}, H(\lambda))}{R(B[\xi], B[\xi], H(\lambda))}, \quad i, j \in \mathcal{I}.$$

С учетом полиномиальности функций R и \tilde{R} можно заключить, что построенное отображение M является аналитическим в окрестности $(\bar{\xi}, \bar{\lambda})$, и поэтому обладает всеми необходимыми свойствами.

Предположим теперь, что $l = 1$. В этом случае

$$\begin{aligned} (B[\xi](H(\lambda))^{-1}(B[\xi])^T)^{-1} &= \frac{1}{B[\xi](H(\lambda))^{-1}(B[\xi])^T} = \frac{\Delta(\lambda)}{\det(B[\xi](H(\lambda))^{-1}(B[\xi])^T) \Delta(\lambda)} = \\ &= \frac{\Delta(\lambda)}{R(B[\xi], B[\xi], H(\lambda))}, \end{aligned} \quad (76)$$

и поэтому, определив $M: \mathbb{R}^n \times \mathbb{R}^l \rightarrow \mathbb{R}$,

$$M(\xi, \lambda) = \frac{\Delta(\lambda)}{R(B[\xi], B[\xi], H(\lambda))}$$

и проведя аналогичные рассуждения, легко убедиться, что функция M является аналитической и обладает всеми необходимыми свойствами. \blacksquare

Лемма 9. Пусть условие (64) выполнено для некоторого $\bar{\lambda} \in \mathcal{M}_0$. Пусть также выполнено (74) для вектора $\bar{\xi} = \xi(\bar{\lambda})$, где отображение $\xi(\cdot)$ определено согласно лемме 4.

Тогда найдется $\rho > 0$, для которого справедливо

$$\delta(\xi(\tilde{\lambda}), \lambda) = \lambda - \tilde{\lambda} \quad \forall \lambda \in \mathcal{B}(\bar{\lambda}, \rho) \setminus \mathcal{M}_0, \quad \forall \tilde{\lambda} \in \mathcal{M}_0 \cap \mathcal{B}(\bar{\lambda}, \rho),$$

где выражение в левой части корректно определено первым соотношением в (62).

Доказательство. Предполагая, что $\rho > 0$ достаточно мало, рассмотрим произвольный множитель $\tilde{\lambda} \in \mathcal{M}_0 \cap \mathcal{B}(\bar{\lambda}, \rho)$ и введем обозначение $\tilde{\xi} = \xi(\tilde{\lambda})$. Из непрерывности отображения $\xi(\cdot)$ и из леммы 8 следует, что выражение $\delta(\tilde{\xi}, \lambda)$ корректно определено первым соотношением в (62) и является единственным решением линейной системы

$$B[\tilde{\xi}](H(\lambda))^{-1}(\delta B)\tilde{\xi} = B[\tilde{\xi}, \tilde{\xi}] \quad (77)$$

относительно $\delta \in \mathbb{R}^l$. Заметим, что

$$\tilde{\xi} = (H(\lambda))^{-1}H(\lambda)\tilde{\xi} = (H(\lambda))^{-1}(H(\tilde{\lambda}) + (\lambda - \tilde{\lambda})B)\tilde{\xi} = (H(\lambda))^{-1}((\lambda - \tilde{\lambda})B)\tilde{\xi},$$

откуда немедленно следует, что $\delta = \lambda - \tilde{\lambda}$ удовлетворяет системе (77). Это и означает справедливость $\delta(\tilde{\xi}, \lambda) = \lambda - \tilde{\lambda}$. \blacksquare

Следствие 1. Пусть условие (64) выполнено для некоторого $\bar{\lambda} \in \mathcal{M}_0$. Пусть также выполнено (74) для вектора $\bar{\xi} = \xi(\bar{\lambda})$, где отображение $\xi(\cdot)$ определено согласно лемме 4.

Тогда найдутся $\rho > 0$ и $\ell > 0$, для которых справедливо

$$\|\delta(\xi, \lambda) - (\lambda - \tilde{\lambda})\| \leq \ell \|\xi - \xi(\tilde{\lambda})\| \quad \forall \xi \in \mathcal{B}(\bar{\xi}, \rho), \forall \lambda \in \mathcal{B}(\bar{\lambda}, \rho) \setminus \mathcal{M}_0, \forall \tilde{\lambda} \in \mathcal{M}_0 \cap \mathcal{B}(\bar{\lambda}, \rho), \quad (78)$$

где $\delta(\xi, \lambda)$ корректно определено первым соотношением в (62).

Доказательство. Как следует из леммы 8, для достаточно малой величины $\rho > 0$ можно корректно определить отображение $\hat{\delta}: \mathcal{B}(\bar{\xi}, \rho) \times \mathcal{B}(\bar{\lambda}, \rho) \rightarrow \mathbb{R}^l$, $\hat{\delta}(\xi, \lambda) = M(\xi, \lambda)B[\xi, \xi]$, удовлетворяющее условию Липшица с некоторой константой $\ell > 0$. При этом из (62) следует, что $\hat{\delta}(\xi, \lambda) = \delta(\xi, \lambda)$ для всех $(\xi, \lambda) \in \mathcal{B}(\bar{\xi}, \rho) \times (\mathcal{B}(\bar{\lambda}, \rho) \setminus \mathcal{M}_0)$. Уменьшая ρ в случае необходимости и применяя лемму 9 для любых таких ξ и λ и для любого $\tilde{\lambda} \in \mathcal{M}_0 \cap \mathcal{B}(\bar{\lambda}, \rho)$, получаем соотношение

$$\|\delta(\xi, \lambda) - (\lambda - \tilde{\lambda})\| = \|\hat{\delta}(\xi, \lambda) - \hat{\delta}(\xi(\tilde{\lambda}), \lambda)\| \leq \ell \|\xi - \xi(\tilde{\lambda})\|.$$

\blacksquare

Лемма 10. Пусть условие (64) выполнено для некоторого $\bar{\lambda} \in \mathcal{M}_0$. Тогда найдутся $\rho > 0$ и $\gamma > 0$, для которых справедливо

$$\|H(\tilde{\lambda})\xi\| \geq \gamma \|\xi - \xi(\tilde{\lambda})\| \quad \forall \xi \in \mathcal{B}(\bar{\xi}, \rho): \|\xi\| = 1, \forall \tilde{\lambda} \in \mathcal{M}_0 \cap \mathcal{B}(\bar{\lambda}, \rho),$$

где отображение $\xi(\cdot)$ определено согласно лемме 4.

Доказательство. Предположим, что утверждение неверно: существует последовательность $\{\varepsilon_k\}$ действительных чисел и последовательности $\{\xi^k\} \subset \mathbb{R}^n$ и $\{\tilde{\lambda}^k\} \subset \mathcal{M}_0$ такие, что $\varepsilon_k \rightarrow 0$, $\{\xi^k\} \rightarrow \bar{\xi}$, $\{\tilde{\lambda}^k\} \rightarrow \bar{\lambda}$, и при этом для всех k вектор $\xi(\tilde{\lambda}^k)$ корректно определен, $\|\xi^k\| = 1$, и выполнено $\|H(\tilde{\lambda}^k)\xi^k\| < \varepsilon_k \|\xi^k - \xi(\tilde{\lambda}^k)\|$. Таким образом, $\xi^k \neq \xi(\tilde{\lambda}^k)$ для всех k , и справедливо $\lim_{k \rightarrow \infty} \|H(\tilde{\lambda}^k)\xi^k\| / \|\xi^k - \xi(\tilde{\lambda}^k)\| = 0$, откуда следует

$$\left\{ H(\tilde{\lambda}^k) \frac{\xi^k - \xi(\tilde{\lambda}^k)}{\|\xi^k - \xi(\tilde{\lambda}^k)\|} \right\} \rightarrow 0. \quad (79)$$

Для каждого k введем обозначение $\zeta^k = (\xi^k - \xi(\tilde{\lambda}^k)) / \|\xi^k - \xi(\tilde{\lambda}^k)\|$. Заметим, что справедлива следующая цепочка равенств:

$$\langle \zeta^k, \xi(\tilde{\lambda}^k) \rangle^2 = \frac{(\langle \xi^k, \xi(\tilde{\lambda}^k) \rangle - 1)^2}{\|\xi^k - \xi(\tilde{\lambda}^k)\|^2} = \frac{(\langle \xi^k, \xi(\tilde{\lambda}^k) \rangle - 1)^2}{2(1 - \langle \xi^k, \xi(\tilde{\lambda}^k) \rangle)} = \frac{1}{2}(1 - \langle \xi^k, \xi(\tilde{\lambda}^k) \rangle).$$

С учетом предельных соотношений $\{\xi^k\} \rightarrow \bar{\xi}$ и $\{\xi(\tilde{\lambda}^k)\} \rightarrow \xi(\bar{\lambda}) = \bar{\xi}$, а также условия $\|\bar{\xi}\| = 1$, получаем

$$\lim_{k \rightarrow \infty} \langle \zeta^k, \xi(\tilde{\lambda}^k) \rangle = 0.$$

Таким образом, для любой предельной точки $\bar{\zeta}$ ограниченной последовательности $\{\zeta^k\}$ выполнено $\langle \bar{\zeta}, \xi(\bar{\lambda}) \rangle = 0$, что, в свою очередь, эквивалентно $\bar{\zeta} \in (\ker H(\bar{\lambda}))^\perp$. С другой стороны, из (79) следует $\bar{\zeta} \in \ker H(\bar{\lambda})$, что возможно только при $\bar{\zeta} = 0$, и поэтому противоречит $\|\bar{\zeta}\| = 1$. ■

Для следующего результата об оценке расстояния потребуется еще одно предположение

$$B[\bar{\xi}, \bar{\xi}] \neq 0. \quad (80)$$

Лемма 11. Пусть условие (64) выполнено для некоторого $\bar{\lambda} \in \mathcal{M}_0$. Пусть также выполнено (80) для некоторого вектора $\bar{\xi} \in \ker H(\bar{\lambda})$ единичной нормы.

Тогда найдутся $\rho > 0$ и $C_1, C_2 > 0$, для которых справедлива следующая оценка расстояния:

$$C_1 |\Delta(\lambda)| \leq \text{dist}(\lambda, \mathcal{M}_0) \leq C_2 |\Delta(\lambda)| \quad \forall \lambda \in \mathcal{B}(\bar{\lambda}, \rho). \quad (81)$$

Доказательство. Первая оценка в (81) немедленно следует из липшицевости полиномиальной функции Δ на ограниченных множествах. Вторая оценка в (81) следует из свойства

$$\Delta'(\bar{\lambda}) \neq 0, \quad (82)$$

которое, в свою очередь, следует из предположения (80) и леммы 6. ■

Определим множество

$$\mathcal{M}_{pd} = \{(\xi, \lambda) \in \mathbb{R}^n \times \mathbb{R}^l \mid H(\lambda)\xi = 0\}.$$

Очевидно,

$$\mathcal{M}_0 = \{\lambda \in \mathbb{R}^l \mid \exists \xi \in \mathbb{R}^n \setminus \{0\}: (\xi, \lambda) \in \mathcal{M}_{pd}\}.$$

Лемма 12. Пусть условие (64) выполнено для некоторого $\bar{\lambda} \in \mathcal{M}_0$. Пусть также выполнены условия (74), (80) для вектора $\bar{\xi} = \xi(\bar{\lambda})$, где отображение $\xi(\cdot)$ определено согласно лемме 4.

Тогда найдется такое $\rho > 0$, что для любого $\ell_1 > 0$ найдется $\ell_2 > 0$ со следующими свойствами: для любого вектора $(\xi, \lambda) \in \mathcal{B}(\bar{\xi}, \rho) \times (\mathcal{B}(\bar{\lambda}, \rho) \setminus \mathcal{M}_0)$, удовлетворяющего неравенству

$$\frac{\|\delta(\xi, \lambda)\|}{|\Delta(\lambda)|} \leq \ell_1, \quad (83)$$

где $\delta(\xi, \lambda)$ корректно определено первым соотношением в (62), и для любого $\theta \in [0, 1]$ справедливы соотношения

$$\left\| \xi \cdot (\Delta'(\lambda - \theta\delta(\xi, \lambda)))^T - \hat{H}(\lambda)(B[\xi])^T \right\| \leq \ell_2 \text{dist}((\xi, \lambda), \mathcal{M}_{pd}), \quad (84)$$

$$\left| \left\langle \Delta'(\lambda - \theta\delta(\xi, \lambda)), \frac{\delta(\xi, \lambda)}{\Delta(\lambda)} \right\rangle - 1 \right| \leq \ell_2 \text{dist}((\xi, \lambda), \mathcal{M}_{pd}). \quad (85)$$

Доказательство. Зафиксируем $\rho > 0$ таким образом, чтобы выполнялись утверждения лемм 4–6, 8 и 11, следствия 1, а также чтобы условие

$$\|B[\xi, \xi]\| \geq \varepsilon \quad \forall \xi \in \mathcal{B}(\bar{\xi}, \rho) \quad (86)$$

было выполнено для некоторой константы $\varepsilon > 0$ (возможность выбрать ρ , удовлетворяющее последнему свойству, следует из (80)). Далее, выберем константу $\tilde{\rho} \in (0, \rho)$ таким образом, чтобы оценка $\|\delta(\xi, \lambda)\| \leq \rho - \tilde{\rho}$ была справедлива для всех $(\xi, \lambda) \in \mathcal{B}(\bar{\xi}, \tilde{\rho}) \times (\mathcal{B}(\bar{\lambda}, \tilde{\rho}) \setminus \mathcal{M}_0)$ (возможность выбрать такое $\tilde{\rho}$ вытекает из следствия 1). Рассмотрим произвольный вектор $(\xi, \lambda) \in \mathcal{B}(\bar{\xi}, \tilde{\rho}) \times (\mathcal{B}(\bar{\lambda}, \tilde{\rho}) \setminus \mathcal{M}_0)$ со свойствами, заданными в условии леммы, и произвольную константу $\theta \in [0, 1]$. Обозначим через $(\tilde{\xi}, \tilde{\lambda}) \in \mathcal{M}_{pd}$ проекцию вектора (ξ, λ) на множество \mathcal{M}_{pd} . Очевидно, что, уменьшая $\tilde{\rho}$ в случае необходимости, можно обеспечить выполнение условий $(\tilde{\xi}, \tilde{\lambda}) \in \mathcal{B}(\bar{\xi}, \rho) \times \mathcal{B}(\bar{\lambda}, \rho)$ и

$$\xi(\tilde{\lambda}) = \frac{\tilde{\xi}}{\|\tilde{\xi}\|} \quad (87)$$

для любого $(\xi, \lambda) \in \mathcal{B}(\bar{\xi}, \tilde{\rho}) \times (\mathcal{B}(\bar{\lambda}, \tilde{\rho}) \setminus \mathcal{M}_0)$.

Обозначим для краткости $\delta = \delta(\xi, \lambda)$, $\eta = \delta/\Delta(\lambda)$ и $\hat{\lambda} = \lambda - \theta\delta$. Условие (83) означает, что $\|\eta\| \leq \ell_1$. Из леммы 11 следует

$$|\Delta(\lambda)| \leq \frac{1}{C_1} \text{dist}(\lambda, \mathcal{M}_0) \leq \frac{1}{C_1} \|\lambda - \tilde{\lambda}\|,$$

и поэтому, согласно (83),

$$\|\hat{\lambda} - \tilde{\lambda}\| \leq \|\lambda - \tilde{\lambda}\| + \|\delta\| \leq \ell_3 \|\lambda - \tilde{\lambda}\| \quad (88)$$

для $\ell_3 = 1 + \ell_1/C_1$. Заметим, что из выбора $\tilde{\rho}$ следует $\hat{\lambda} \in \mathcal{B}(\bar{\lambda}, \rho)$.

Очевидно, градиент функции Δ удовлетворяет условию Липшица с некоторой константой $\ell_4 > 0$ на множестве $\mathcal{B}(\bar{\lambda}, \rho)$, и поэтому, используя (88), получаем

$$\|\Delta'(\hat{\lambda}) - \Delta'(\tilde{\lambda})\| \leq \ell_4 \|\hat{\lambda} - \tilde{\lambda}\| \leq \ell_3 \ell_4 \|\lambda - \tilde{\lambda}\|. \quad (89)$$

Аналогично, функция $\hat{H}(\cdot)$ липшицева с некоторой константой $\ell_5 > 0$ на множестве $\mathcal{B}(\bar{\lambda}, \rho)$, и поэтому

$$\|\hat{H}(\lambda) - \hat{H}(\tilde{\lambda})\| \leq \ell_5 \|\lambda - \tilde{\lambda}\|. \quad (90)$$

Наконец, для некоторой константы $\ell_6 > 0$ выполнены условия

$$\|\Delta'(\hat{\lambda})\| \leq \ell_6, \quad \|\hat{H}(\lambda)\| \leq \ell_6, \quad \|\tilde{\xi}\| \leq \ell_6, \quad (91)$$

$$\|B[\xi]\| \leq \ell_6, \quad \|B[\tilde{\xi}]\| \leq \ell_6, \quad \|B[\xi - \tilde{\xi}]\| \leq \ell_6 \|\xi - \tilde{\xi}\|. \quad (92)$$

Подчеркнем, что для любого $\rho > 0$ константы ℓ_4 , ℓ_5 и ℓ_6 могут быть выбраны независимо от (ξ, λ) и $(\tilde{\xi}, \tilde{\lambda})$.

Далее, используя леммы 5, 6 и соотношение (87), получаем следующую цепочку равенств:

$$\begin{aligned} \tilde{\xi} \cdot (\Delta'(\tilde{\lambda}))^T &= t(\tilde{\lambda}) \tilde{\xi} \cdot (\xi(\tilde{\lambda}))^T (B[\xi(\tilde{\lambda})])^T = t(\tilde{\lambda}) \frac{\tilde{\xi}}{\|\tilde{\xi}\|} (\xi(\tilde{\lambda}))^T (B[\|\tilde{\xi}\| \xi(\tilde{\lambda})])^T = \\ &= t(\tilde{\lambda}) \xi(\tilde{\lambda}) (\xi(\tilde{\lambda}))^T (B[\tilde{\xi}])^T = \hat{H}(\tilde{\lambda}) (B[\tilde{\xi}])^T. \end{aligned}$$

С учетом оценок (89)–(92), отсюда находим

$$\begin{aligned} \|\xi \cdot (\Delta'(\hat{\lambda}))^T - \hat{H}(\lambda) (B[\xi])^T\| &\leq \|\xi - \tilde{\xi}\| \|\Delta'(\hat{\lambda})\| + \|\tilde{\xi}\| \|\Delta'(\hat{\lambda}) - \Delta'(\tilde{\lambda})\| + \\ &+ \|\hat{H}(\tilde{\lambda}) - \hat{H}(\lambda)\| \|B[\tilde{\xi}]\|^T + \|\hat{H}(\lambda)\| \|B[\tilde{\xi} - \xi]\|^T \leq \\ &\leq (\ell_6 + \ell_6^2) \|\xi - \tilde{\xi}\| + (\ell_3 \ell_4 + \ell_5) \ell_6 \|\lambda - \tilde{\lambda}\|. \end{aligned}$$

(Напомним, что спектральная норма матрицы совпадает с евклидовой нормой в случае, когда матрица состоит из одного столбца или одной строки; также напомним, что спектральная норма транспонированной матрицы совпадает со спектральной нормой самой матрицы.) Чтобы получить неравенство (84), остается положить $\ell_2 = \sqrt{2} \max\{\ell_6 + \ell_6^2, (\ell_3 \ell_4 + \ell_5) \ell_6\}$. Заметим, что константа ℓ_2 зависит только от ρ и ℓ_1 , и не зависит от векторов (ξ, λ) и $(\tilde{\xi}, \tilde{\lambda})$.

Теперь докажем неравенство (85). Из первого соотношения в (62) и из леммы 8 следует, что δ корректно определяется равенством $B[\xi] (H(\lambda))^{-1} (B[\xi])^T \delta = B[\xi, \xi]$, из которого, в свою очередь, следует $B[\xi] \hat{H}(\lambda) (B[\xi])^T \eta = B[\xi, \xi]$, и поэтому

$$B[\xi, \xi] (\langle \Delta'(\hat{\lambda}), \eta \rangle - 1) = B[\xi] (\xi \cdot (\Delta'(\hat{\lambda}))^T - \hat{H}(\lambda) (B[\xi])^T) \eta.$$

Применяя оценки (84), (86), а также первое неравенство в (92), получаем цепочку неравенств

$$\begin{aligned} |\langle \Delta'(\hat{\lambda}), \eta \rangle - 1| &= \frac{\|B[\xi] (\xi \cdot (\Delta'(\hat{\lambda}))^T - \hat{H}(\lambda) (B[\xi])^T) \eta\|}{\|B[\xi, \xi]\|} \leq \frac{\ell_2 \ell_6}{\varepsilon} \|\eta\| \operatorname{dist}((\xi, \lambda), \mathcal{M}_{pd}) \leq \\ &\leq \frac{\ell_1 \ell_2 \ell_6}{\varepsilon} \operatorname{dist}((\xi, \lambda), \mathcal{M}_{pd}), \end{aligned}$$

последнее из которых следует из определения η и из оценки (83).

Для завершения доказательства остается переопределить ρ как $\tilde{\rho}$, а также заменить ℓ_2 на $\ell_1 \ell_2 \ell_6 / \varepsilon$, если это произведение больше, чем ℓ_2 . \blacksquare

Лемма 13. Пусть условие (64) выполнено для некоторого $\bar{\lambda} \in \mathcal{M}_0$. Пусть также выполнены условия (74), (80) для вектора $\bar{\xi} = \xi(\bar{\lambda})$, где отображение $\xi(\cdot)$ определено согласно лемме 4.

Тогда найдется такое $\rho > 0$, что для любой константы $\ell_1 > 0$ и произвольных последовательностей $\{\xi^k\} \subset \mathcal{B}(\bar{\xi}, \rho)$ и $\{\lambda^k\} \subset \mathcal{B}(\bar{\lambda}, \rho) \setminus \mathcal{M}_0$, удовлетворяющих предельным соотношениям $\{\xi^k\} \rightarrow \xi(\tilde{\lambda})$ и $\{\lambda^k\} \rightarrow \tilde{\lambda}$ для некоторого $\tilde{\lambda} \in \mathcal{M}_0$, а также для всех k неравенству

$$\frac{\|\delta^k\|}{|\Delta(\lambda^k)|} \leq \ell_1, \quad (93)$$

где $\delta^k = \delta(\xi^k, \lambda^k)$ корректно определено первым соотношением в (62), выполнены следующие соотношения:

$$\lim_{k \rightarrow \infty} \frac{\Delta(\hat{\lambda}^k)}{\Delta(\lambda^k)} = \frac{1}{2}, \quad (94)$$

$$\{P(\xi^k, \lambda^k)\xi^k\} \rightarrow \xi(\tilde{\lambda}), \quad (95)$$

где $\hat{\lambda}^k = \lambda^k - \delta^k/2$. Более того, существует константа $\ell_2 > 0$ и последовательность $\{\tilde{\lambda}^k\} \subset \mathcal{M}_0$, для которой справедливо предельное соотношение $\{\tilde{\lambda}^k\} \rightarrow \tilde{\lambda}$, и для всех достаточно больших k имеют место следующие оценки:

$$\frac{\|\lambda^k - \delta^k - \tilde{\lambda}^k\|}{|\Delta(\lambda^k)|} \leq \ell_2 \text{dist}((\xi^k, \lambda^k), \mathcal{M}_{pd}), \quad (96)$$

$$\frac{\|\hat{\xi}^k - \xi(\tilde{\lambda}^k)\|}{|\Delta(\lambda^k)|} \leq \ell_2 \text{dist}((\xi^k, \lambda^k), \mathcal{M}_{pd}), \quad (97)$$

$$\|\eta^k - \hat{\eta}^k\| \leq \ell_2 \text{dist}((\xi^k, \lambda^k), \mathcal{M}_{pd}), \quad (98)$$

где

$$\hat{\xi}^k = \frac{P(\xi^k, \lambda^k)\xi^k}{\|P(\xi^k, \lambda^k)\xi^k\|}, \quad \eta^k = \frac{\delta^k}{\Delta(\lambda^k)}, \quad \hat{\eta}^k = \frac{\delta(\hat{\xi}^k, \hat{\lambda}^k)}{\Delta(\hat{\lambda}^k)},$$

причем справедливы неравенства $P(\xi^k, \lambda^k)\xi^k \neq 0$ и $\Delta(\hat{\lambda}^k) \neq 0$, первое из которых корректно определено вторым соотношением в (62).

Доказательство. Зафиксируем $\rho > 0$ таким образом, чтобы были выполнены утверждения лемм 10–12 и следствия 1. Выберем произвольную константу $\tilde{\rho} \in (0, \rho)$ и рассмотрим произвольные последовательности $\{\xi^k\} \subset \mathcal{B}(\bar{\xi}, \tilde{\rho})$ и $\{\lambda^k\} \subset \mathcal{B}(\bar{\lambda}, \tilde{\rho}) \setminus \mathcal{M}_0$ со свойствами, указанными в условии леммы. Условие (93) означает, что последовательность $\{\eta^k\}$ ограничена, откуда следует $\{\delta^k\} \rightarrow 0$ и $\{\hat{\lambda}^k\} \rightarrow \tilde{\lambda}$.

Сначала докажем существование константы $\ell_3 > 0$, удовлетворяющей для всех k неравенству

$$\left| \frac{\Delta(\hat{\lambda}^k)}{\Delta(\lambda^k)} - \frac{1}{2} \right| \leq \ell_3 \text{dist}((\xi^k, \lambda^k), \mathcal{M}_{pd}). \quad (99)$$

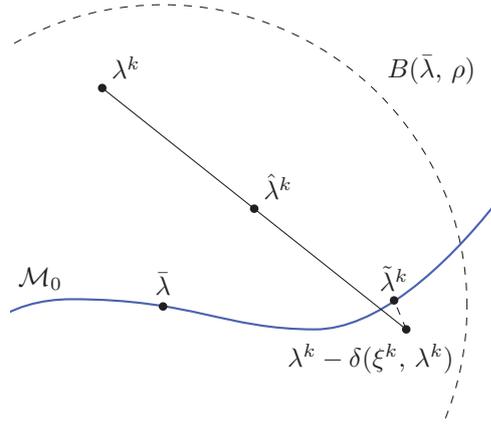


Рис. 6. Иллюстрация к лемме 13.

Действительно, из теоремы о среднем вытекает существование последовательности $\{\theta_k\} \subset [0, 1]$, для которой справедлива цепочка равенств

$$1 - \frac{\Delta(\hat{\lambda}^k)}{\Delta(\lambda^k)} = \frac{\Delta(\lambda^k) - \Delta(\hat{\lambda}^k)}{\Delta(\lambda^k)} = \frac{\left\langle \Delta' \left(\lambda^k - \theta_k \frac{\delta^k}{2} \right), \frac{\delta^k}{2} \right\rangle}{\Delta(\lambda^k)} = \frac{1}{2} \left\langle \Delta' \left(\lambda^k - \theta_k \frac{\delta^k}{2} \right), \eta^k \right\rangle.$$

Применяя соотношение (85) в лемме 12, получаем требуемое свойство (99) при $\ell_3 = \ell_2/2$. Заметим, что из (99) сразу же следует (94), и, в частности, $\Delta(\hat{\lambda}^k) \neq 0$ для всех достаточно больших k .

Далее, из теоремы о среднем также следует существование последовательности $\{\theta_k\} \subset [0, 1]$, для которой выполнено

$$1 - \frac{\Delta(\lambda^k - \delta^k)}{\Delta(\lambda^k)} = \langle \Delta'(\lambda^k - \theta_k \delta^k), \eta^k \rangle.$$

Вновь применяя соотношение (85) в лемме 12, получаем

$$\left| \frac{\Delta(\lambda^k - \delta^k)}{\Delta(\lambda^k)} \right| \leq \ell_2 \text{dist}((\xi^k, \lambda^k), \mathcal{M}_{pd}).$$

Таким образом, из второй оценки в (81) из леммы 11 следует, что для всех достаточно больших k выполнено

$$\left| \frac{\text{dist}(\lambda^k - \delta^k, \mathcal{M}_0)}{\Delta(\lambda^k)} \right| \leq C_2 \ell_2 \text{dist}((\xi^k, \lambda^k), \mathcal{M}_{pd})$$

(поскольку $\lambda^k - \delta^k \in \mathcal{B}(\bar{\lambda}, \rho)$ для таких k), откуда немедленно следует существование последовательности $\{\tilde{\lambda}^k\} \subset \mathcal{M}_0$, удовлетворяющей (96), если переопределить ℓ_2 как $C_2 \ell_2$ в случае, когда последнее больше чем ℓ_2 . Поскольку $\{\lambda^k\} \rightarrow \tilde{\lambda}$ и $\{\delta^k\} \rightarrow 0$, из (96) следует $\{\tilde{\lambda}^k\} \rightarrow \tilde{\lambda}$.

Для краткости введем обозначения $H_k = H(\lambda^k)$, $\hat{H}_k = \hat{H}(\lambda^k)$, $P_k = P(\xi^k, \lambda^k)$ для всех k . Покажем, что существует константа $\ell_4 > 0$, для которой выполнено неравенство

$$\|P_k \xi^k - \xi^k\| \leq \ell_4 \text{dist}((\xi^k, \lambda^k), \mathcal{M}_{pd}). \quad (100)$$

Действительно, из второго соотношения в (62), с учетом определения η^k , получаем

$$P_k \xi^k = H_k^{-1}(\delta^k B) \xi^k = \hat{H}_k(\eta^k B) \xi^k = \hat{H}_k(B[\xi^k])^T \eta^k,$$

и поэтому

$$P_k \xi^k - \xi^k = (\hat{H}_k(B[\xi^k])^T - \xi^k (\Delta'(\lambda^k))^T) \eta^k + \xi^k (\langle \Delta'(\lambda^k), \eta^k \rangle - 1).$$

Поскольку последовательности $\{\xi^k\}$ и $\{\eta^k\}$ ограничены, из леммы 12 сразу же следует оценка (100) при некотором $\ell_4 > 0$. Отметим, что из (100) также следует (95), и, в частности, $P_k \xi^k \neq 0$ для всех достаточно больших k .

Теперь покажем, что оценка (97) выполняется для указанной выше последовательности $\{\tilde{\lambda}^k\}$. Заметим, что $H_k P_k = \delta^k B$, и поэтому

$$\frac{(H_k - \delta^k B) P_k \xi^k}{\Delta_k} = \frac{(\delta^k B)(\xi^k - P_k \xi^k)}{\Delta_k} = (\eta^k B)(\xi^k - P_k \xi^k),$$

где $\Delta_k = \Delta(\lambda^k) (= \det H_k)$. Введем еще одно обозначение $\tilde{H}_k = H(\tilde{\lambda}^k)$. Тогда, по определению $\hat{\xi}^k$, получаем

$$\begin{aligned} \frac{\tilde{H}_k \hat{\xi}^k}{\Delta_k} &= \frac{\tilde{H}_k P_k \xi^k}{\Delta_k \|P_k \xi^k\|} = \frac{(H_k - (\lambda^k - \tilde{\lambda}^k) B) P_k \xi^k}{\Delta_k \|P_k \xi^k\|} = \\ &= \frac{1}{\|P_k \xi^k\|} \left(\frac{(H_k - \delta^k B) P_k \xi^k}{\Delta_k} - \frac{((\lambda^k - \delta^k - \tilde{\lambda}^k) B) P_k \xi^k}{\Delta_k} \right) = \\ &= \frac{\eta^k B}{\|P_k \xi^k\|} (\xi^k - P_k \xi^k) - \frac{(B[P_k \xi^k])^T (\lambda^k - \delta^k - \tilde{\lambda}^k)}{\|P_k \xi^k\| \Delta_k}. \end{aligned} \quad (101)$$

Поскольку последовательность $\{\eta^k\}$ ограничена, из (95) следует существование константы $\ell_5 > 0$, для которой для всех достаточно больших k выполнены неравенства

$$\frac{\|\eta^k B\|}{\|P_k \xi^k\|} \leq \ell_5, \quad \frac{\|(B[P_k \xi^k])^T\|}{\|P_k \xi^k\|} \leq \ell_5.$$

Объединяя (101) с (96) и (100), получаем

$$\frac{\|\tilde{H}_k \hat{\xi}^k\|}{\Delta_k} \leq \ell_6 \text{dist}((\xi^k, \lambda^k), \mathcal{M}_{pd}) \quad (102)$$

при $\ell_6 = (\ell_2 + \ell_4) \ell_5$. Наконец, из (95) и сходимости $\{\tilde{\lambda}^k\}$ к $\tilde{\lambda}$ следует, что для достаточно больших k выполнено $\hat{\xi}^k \in \mathcal{B}(\bar{\xi}, \rho)$ и $\tilde{\lambda}^k \in \mathcal{B}(\bar{\lambda}, \rho)$, откуда по лемме 10 получаем оценку $\|\tilde{H}_k \hat{\xi}^k\| \geq \gamma \|\hat{\xi}^k - \xi(\tilde{\lambda}^k)\|$. Комбинируя ее с неравенством (102), получаем цепочку неравенств

$$\frac{\|\hat{\xi}^k - \xi(\tilde{\lambda}^k)\|}{\Delta_k} \leq \frac{1}{\gamma} \frac{\|\tilde{H}_k \hat{\xi}^k\|}{\Delta_k} \leq \frac{\ell_6}{\gamma} \text{dist}((\xi^k, \lambda^k), \mathcal{M}_{pd}),$$

которая представляет собой оценку (97), если переопределить ℓ_2 как ℓ_6/γ .

Остается доказать оценку (98). Применяя (96) к тождеству

$$\frac{\hat{\lambda}^k - \tilde{\lambda}^k}{\Delta_k} = \frac{\eta^k}{2} + \frac{\lambda^k - \delta^k - \tilde{\lambda}^k}{\Delta_k}, \quad (103)$$

находим

$$\left\| \eta^k - 2 \frac{\hat{\lambda}^k - \tilde{\lambda}^k}{\Delta_k} \right\| \leq 2\ell_2 \operatorname{dist}((\xi^k, \lambda^k), \mathcal{M}_{pd}). \quad (104)$$

Далее, из тождества (103) и условий (93), (96) также следует, что для некоторого $\varepsilon > 0$ и для всех достаточно больших k справедливо $\|\hat{\lambda}^k - \tilde{\lambda}^k\|/|\Delta_k| \leq \ell_1/2 + \varepsilon$. Кроме того, из (94) легко получить, что для всех достаточно больших k выполнено

$$\left| \frac{\Delta_k}{\hat{\Delta}_k} \right| \leq 2 + \varepsilon, \quad (105)$$

где введено обозначение $\hat{\Delta}_k = \Delta(\hat{\lambda}^k) (= \det H(\hat{\lambda}^k))$. Комбинируя две последние оценки с (99), получаем

$$\left\| \frac{\hat{\lambda}^k - \tilde{\lambda}^k}{\Delta_k} - \frac{\hat{\lambda}^k - \tilde{\lambda}^k}{2\hat{\Delta}_k} \right\| = \left| \frac{\hat{\Delta}_k}{\Delta_k} - \frac{1}{2} \right| \left| \frac{\Delta_k}{\hat{\Delta}_k} \right| \frac{\|\hat{\lambda}^k - \tilde{\lambda}^k\|}{|\Delta_k|} \leq \ell_7 \operatorname{dist}((\xi^k, \lambda^k), \mathcal{M}_{pd}), \quad (106)$$

где $\ell_7 = (\ell_1/2 + \varepsilon)(2 + \varepsilon)\ell_3$. Наконец, согласно следствию 1, для всех достаточно больших k выполнено

$$\|\delta(\hat{\xi}^k, \hat{\lambda}^k) - (\hat{\lambda}^k - \tilde{\lambda}^k)\| \leq \ell \|\hat{\xi}^k - \xi(\tilde{\lambda}^k)\|,$$

откуда, с учетом определения $\hat{\eta}^k$ и свойств (97), (105), легко получить

$$\left\| \hat{\eta}^k - \frac{\hat{\lambda}^k - \tilde{\lambda}^k}{\hat{\Delta}_k} \right\| \leq \ell \frac{\|\hat{\xi}^k - \xi(\tilde{\lambda}^k)\|}{|\hat{\Delta}_k|} = \ell \left| \frac{\Delta_k}{\hat{\Delta}_k} \right| \frac{\|\hat{\xi}^k - \xi(\tilde{\lambda}^k)\|}{|\Delta_k|} \leq \ell_8 \operatorname{dist}((\xi^k, \lambda^k), \mathcal{M}_{pd}), \quad (107)$$

где $\ell_8 = \ell\ell_2(2 + \varepsilon)$. Из (104), (106) и (107) теперь выводим

$$\begin{aligned} \|\eta^k - \hat{\eta}^k\| &\leq \left\| \eta^k - 2 \frac{\hat{\lambda}^k - \tilde{\lambda}^k}{\Delta_k} \right\| + \left\| 2 \frac{\hat{\lambda}^k - \tilde{\lambda}^k}{\Delta_k} - \frac{\hat{\lambda}^k - \tilde{\lambda}^k}{\hat{\Delta}_k} \right\| + \left\| \frac{\hat{\lambda}^k - \tilde{\lambda}^k}{\hat{\Delta}_k} - \hat{\eta}^k \right\| \\ &\leq (2\ell_2 + 2\ell_7 + \ell_8) \operatorname{dist}((\xi^k, \lambda^k), \mathcal{M}_{pd}) \end{aligned}$$

для всех достаточно больших k . Переопределяя ℓ_2 как $2\ell_2 + 2\ell_7 + \ell_8$, получаем оценку (98).

Для завершения доказательства остается переопределить ρ как $\tilde{\rho}$. \blacksquare

Вообще говоря, константа ℓ_2 в данной лемме не обязательно должна быть одинаковой для различных последовательностей $\{\xi^k\}$, $\{\lambda^k\}$. Однако, приводимое ниже следствие показывает, что эту константу можно выбрать независимо от последовательностей.

Следствие 2. В условиях леммы 12, для любых $\ell_1 > 0$ и $\varepsilon > 0$ найдутся такие $\rho > 0$ и $\ell_2 > 0$, что для всех $(\xi, \lambda) \in \mathcal{B}(\bar{\xi}, \rho) \times (\mathcal{B}(\bar{\lambda}, \rho) \setminus \mathcal{M}_0)$, удовлетворяющих (83), где $\delta(\xi, \lambda)$

корректно определено первым соотношением в (62), выполнены следующие соотношения:

$$\left| \frac{\Delta(\hat{\lambda})}{\Delta(\lambda)} - \frac{1}{2} \right| \leq \varepsilon, \quad (108)$$

$$\left\| \frac{\delta(\xi, \lambda)}{\Delta(\lambda)} - \frac{\delta(\hat{\lambda}, \hat{\xi})}{\Delta(\hat{\lambda})} \right\| \leq \ell_2 \text{dist}((\xi, \lambda), \mathcal{M}_{pd}), \quad (109)$$

$$\text{dist}((\hat{\xi}, \hat{\lambda}), \mathcal{M}_{pd}) \leq \ell_2 |\Delta(\lambda)|, \quad (110)$$

$$\|\hat{\xi} - \bar{\xi}\| \leq \ell_2 \|\lambda - \bar{\lambda}\|, \quad (111)$$

где

$$\hat{\xi} = \frac{P(\xi, \lambda)\xi}{\|P(\xi, \lambda)\xi\|}, \quad \hat{\lambda} = \lambda - \frac{\delta(\xi, \lambda)}{2}, \quad (112)$$

причем $P(\xi, \lambda)\xi \neq 0$, где $P(\xi, \lambda)$ корректно определено вторым соотношением в (62).

Доказательство. Зафиксируем произвольные константы $\ell_1 > 0$ и $\varepsilon > 0$. Рассуждая от противного, из соотношений (94), (95) и (98) в лемме 13 легко получить существование таких $\rho > 0$ и $\ell_2 > 0$, что условия (108), (109) и неравенство $P(\xi, \lambda)\xi \neq 0$ будут выполнены для всех $\xi \in \mathcal{B}(\bar{\xi}, \rho)$ и $\lambda \in \mathcal{B}(\bar{\lambda}, \rho) \setminus \mathcal{M}_0$.

Зафиксируем константы $\varepsilon_1 > 0$, $\varepsilon_2 > 0$. Из соотношений (96) и (97) в лемме 13 следует, что, уменьшая $\rho > 0$ при необходимости, можно для любых $\xi \in \mathcal{B}(\bar{\xi}, \rho)$ и $\lambda \in \mathcal{B}(\bar{\lambda}, \rho) \setminus \mathcal{M}_0$, удовлетворяющих (83), гарантировать существование множителя $\tilde{\lambda} \in \mathcal{M}_0$, для которого выполнено

$$\frac{\|(\lambda - \tilde{\lambda}) - \delta(\xi, \lambda)\|}{|\Delta(\lambda)|} \leq \varepsilon_1, \quad \frac{\|\hat{\xi} - \xi(\tilde{\lambda})\|}{|\Delta(\lambda)|} \leq \varepsilon_2. \quad (113)$$

Используя определение $\hat{\lambda}$ и условие (83), отсюда получаем

$$\|\hat{\lambda} - \tilde{\lambda}\| \leq \|(\lambda - \tilde{\lambda}) - \delta(\xi, \lambda)\| + \frac{\|\delta(\xi, \lambda)\|}{2} \leq \left(\varepsilon_1 + \frac{\ell_1}{2} \right) |\Delta(\lambda)|.$$

Поэтому из второго неравенства в (113) вытекает

$$\text{dist}((\hat{\xi}, \hat{\lambda}), \mathcal{M}_{pd}) \leq \|\hat{\xi} - \xi(\tilde{\lambda})\| + \|\hat{\lambda} - \tilde{\lambda}\| \leq \left(\varepsilon_1 + \varepsilon_2 + \frac{\ell_1}{2} \right) |\Delta(\lambda)|,$$

которое совпадает с (110), если положить $\ell_2 = \varepsilon_1 + \varepsilon_2 + \ell_1/2$.

Наконец, из (83) и первого неравенства в (113) имеем

$$\frac{\|\lambda - \tilde{\lambda}\|}{|\Delta(\lambda)|} \leq \frac{\|(\lambda - \tilde{\lambda}) - \delta(\xi, \lambda)\|}{|\Delta(\lambda)|} + \frac{\|\delta(\xi, \lambda)\|}{|\Delta(\lambda)|} \leq \varepsilon_1 + \ell_1,$$

и поэтому $\|\lambda - \tilde{\lambda}\| \leq (\varepsilon_1 + \ell_1)|\Delta(\lambda)|$, откуда, в свою очередь, следует

$$\|\tilde{\lambda} - \bar{\lambda}\| \leq \|\lambda - \tilde{\lambda}\| + \|\lambda - \bar{\lambda}\| \leq (\varepsilon_1 + \ell_1)|\Delta(\lambda)| + \|\lambda - \bar{\lambda}\|. \quad (114)$$

Данная оценка позволяет заключить, что $\tilde{\lambda} \rightarrow \bar{\lambda}$ при $\lambda \rightarrow \bar{\lambda}$, и поэтому для любого заданного $\tilde{\rho} > 0$ с помощью выбора ρ можно гарантировать, что $\tilde{\lambda} \in \mathcal{B}(\bar{\lambda}, \tilde{\rho})$. Поскольку отображение $\xi(\cdot)$ является аналитическим, существует константа $\ell_3 > 0$, для которой выполнено

$$\|\xi(\tilde{\lambda}) - \bar{\xi}\| \leq \ell_3 \|\tilde{\lambda} - \bar{\lambda}\| \quad \forall \tilde{\lambda} \in \mathcal{M}_0 \cap \mathcal{B}(\bar{\lambda}, \tilde{\rho}),$$

при условии, что константа $\tilde{\rho} > 0$ достаточно мала для того, чтобы отображение $\xi(\tilde{\lambda})$ было корректно определено. Комбинируя последнее неравенство со вторым неравенством в (113), а также с (114), находим

$$\begin{aligned} \|\hat{\xi} - \bar{\xi}\| &\leq \|\hat{\xi} - \xi(\tilde{\lambda})\| + \|\xi(\tilde{\lambda}) - \bar{\xi}\| \leq \varepsilon_2 |\Delta(\lambda)| + \ell_3 \|\tilde{\lambda} - \bar{\lambda}\| \leq \\ &\leq ((\varepsilon_1 + \ell_1)\ell_3 + \varepsilon_2) |\Delta(\lambda)| + \ell_3 \|\lambda - \bar{\lambda}\|. \end{aligned} \quad (115)$$

Согласно лемме 11, если $\rho > 0$ достаточно мало, то

$$|\Delta(\lambda)| \leq \frac{1}{C_1} \text{dist}(\lambda, \mathcal{M}_0) \leq \frac{1}{C_1} \|\lambda - \bar{\lambda}\|,$$

и поэтому из (115) следует (111) при $\ell_2 = ((\varepsilon_1 + \ell_1)\ell_3 + \varepsilon_2)/C_1 + \ell_3$. \blacksquare

Предложение 4. В условиях леммы 12, для любых $\ell_0 > 0$ и $\ell_1 > 0$, $\ell_0 < \ell_1$, и для любого $\rho > 0$ найдется такая константа $\rho_0 \in (0, \rho)$, что для любой начальной точки $(\xi^0, \lambda^0) \in \mathcal{B}(\bar{\xi}, \rho_0) \times (\mathcal{B}(\bar{\lambda}, \rho_0) \setminus \mathcal{M}_0)$, удовлетворяющей

$$\frac{\|\delta(\xi^0, \lambda^0)\|}{|\Delta(\lambda^0)|} \leq \ell_0, \quad (116)$$

где $\delta(\xi^0, \lambda^0)$ корректно определено первым соотношением в (62), итерационная схема

$$\xi^{k+1} = \frac{P(\xi^k, \lambda^k)\xi^k}{\|P(\xi^k, \lambda^k)\xi^k\|}, \quad \lambda^{k+1} = \lambda^k - \frac{\delta(\xi^k, \lambda^k)}{2}, \quad (117)$$

где $\delta(\xi^k, \lambda^k)$ и $P(\xi^k, \lambda^k)$ вводятся согласно (62), генерирует корректно определенную траекторию $\{(\xi^k, \lambda^k)\} \subset \mathcal{B}(\bar{\xi}, \rho) \times (\mathcal{B}(\bar{\lambda}, \rho) \setminus \mathcal{M}_0)$; данная траектория сходится к $(\xi(\lambda^*), \lambda^*)$ для некоторого множителя $\lambda^* \in \mathcal{M}_0$ и для всех k удовлетворяет соотношению

$$\frac{\|\delta(\xi^k, \lambda^k)\|}{|\Delta(\lambda^k)|} \leq \ell_1, \quad (118)$$

и выполнено

$$\lim_{k \rightarrow \infty} \frac{\|\lambda^{k+1} - \lambda^*\|}{\|\lambda^k - \lambda^*\|} = \frac{1}{2}, \quad (119)$$

$$\left\{ \frac{\lambda^k - \lambda^*}{\|\lambda^k - \lambda^*\|} \right\} \rightarrow d \quad (120)$$

для некоторого $d \in \mathbb{R}^l$, удовлетворяющего

$$\langle \Delta'(\lambda^*), d \rangle \geq \frac{1}{\ell_1}. \quad (121)$$

Доказательство. Зафиксируем произвольные константы $\ell_0 > 0$, $\ell_1 > 0$, $\ell_0 < \ell_1$, а также $\rho > 0$. Зафиксируем $\varepsilon \in (0, 1/2)$ и положим $q_1 = 1/2 - \varepsilon$, $q_2 = 1/2 + \varepsilon$. Согласно следствию 2, уменьшая при необходимости $\rho > 0$, можно выбрать $\ell_2 > 0$ таким образом, что для всех $\xi \in \mathcal{B}(\bar{\xi}, \rho)$ и $\lambda \in \mathcal{B}(\bar{\lambda}, \rho) \setminus \mathcal{M}_0$, удовлетворяющих (83), будут справедливы соотношения (108)–(111), где $\hat{\xi}$ и $\hat{\lambda}$ корректно определяются условием (112), и, в частности, $P(\xi, \lambda)\xi \neq 0$. Более того, уменьшая при необходимости $\rho > 0$ и увеличивая $\ell_2 > 0$, можно гарантировать, что $\rho < 1$, выполнено утверждение леммы 12, и для всех $\tilde{\lambda} \in \mathcal{M}_0 \cap \mathcal{B}(\bar{\lambda}, \rho)$ справедливо $\dim \ker H(\tilde{\lambda}) = 1$, $\xi(\tilde{\lambda})$ корректно определено леммой 4 и $\|\xi(\tilde{\lambda}) - \bar{\xi}\| < 1$. Наконец, выберем константу $\rho_0 \in (0, \rho)$ таким образом, чтобы следующие неравенства были выполнены для любого множителя $\lambda^0 \in \mathcal{B}(\bar{\lambda}, \rho_0) \setminus \mathcal{M}_0$:

$$\rho_0 + \frac{\ell_1 |\Delta(\lambda^0)|}{2(1 - q_2)} \leq \frac{\rho}{\max\{1, \ell_2\}}, \quad (122)$$

$$\ell_0 + \sqrt{2}\ell_2\rho_0 + \frac{\ell_2^2 |\Delta(\lambda^0)|}{1 - q_2} \leq \ell_1. \quad (123)$$

Доказательство проведем по индукции. Предположим, что вектор $(\xi^0, \lambda^0) \in \mathcal{B}(\bar{\xi}, \rho_0) \times (\mathcal{B}(\bar{\lambda}, \rho_0) \setminus \mathcal{M}_0) \subset \mathcal{B}(\bar{\xi}, \rho) \times (\mathcal{B}(\bar{\lambda}, \rho) \setminus \mathcal{M}_0)$ удовлетворяет (116), и для некоторого неотрицательного числа k векторы (ξ^j, λ^j) корректно определены соотношениями

$$\xi^{j+1} = \frac{P(\xi^j, \lambda^j)\xi^j}{\|P(\xi^j, \lambda^j)\xi^j\|}, \quad \lambda^{j+1} = \lambda^j - \frac{\delta(\xi^j, \lambda^j)}{2} \quad (124)$$

для всех $j = 0, 1, \dots, k-1$, причем для всех $j = 0, 1, \dots, k$ выполнено $(\xi^j, \lambda^j) \in \mathcal{B}(\bar{\xi}, \rho) \times (\mathcal{B}(\bar{\lambda}, \rho) \setminus \mathcal{M}_0)$ и

$$\frac{\|\delta(\xi^j, \lambda^j)\|}{|\Delta(\lambda^j)|} \leq \ell_1. \quad (125)$$

Тогда, согласно следствию 2, вектор $(\xi^{k+1}, \lambda^{k+1})$ (совпадающий с $(\hat{\xi}, \hat{\lambda})$ для $\xi = \xi^k$ и $\lambda = \lambda^k$) корректно определен соотношением (117). Более того, с учетом (108)–(111), получаем

$$\left| \frac{\Delta(\lambda^{j+1})}{\Delta(\lambda^j)} - \frac{1}{2} \right| \leq \varepsilon, \quad (126)$$

$$\|\eta^j - \eta^{j+1}\| \leq \ell_2 \operatorname{dist}((\xi^j, \lambda^j), \mathcal{M}_{pd}), \quad (127)$$

$$\operatorname{dist}((\xi^{j+1}, \lambda^{j+1}), \mathcal{M}_{pd}) \leq \ell_2 |\Delta(\lambda^j)|, \quad (128)$$

$$\|\xi^{j+1} - \bar{\xi}\| \leq \ell_2 \|\lambda^j - \bar{\lambda}\| \quad (129)$$

для всех $j = 0, 1, \dots, k$, где введено обозначение $\eta^j = \delta(\xi^j, \lambda^j)/\Delta(\lambda^j)$.

Очевидно, из (126) следует оценка

$$q_1 \leq \left| \frac{\Delta(\lambda^{j+1})}{\Delta(\lambda^j)} \right| \leq q_2. \quad (130)$$

В частности, $\Delta(\lambda^{k+1}) \neq 0$, откуда следует $\lambda^{k+1} \notin \mathcal{M}_0$. Комбинируя (130) со вторым соотношением в (124) и с неравенством (125) и принимая во внимание (122), получаем

$$\begin{aligned} \|\lambda^j - \bar{\lambda}\| &= \left\| \lambda^0 - \frac{1}{2} \sum_{i=0}^{j-1} \delta(\xi^i, \lambda^i) - \bar{\lambda} \right\| \leq \|\lambda^0 - \bar{\lambda}\| + \frac{1}{2} \sum_{i=0}^{j-1} \|\delta(\xi^i, \lambda^i)\| \leq \\ &\leq \|\lambda^0 - \bar{\lambda}\| + \frac{\ell_1}{2} \sum_{i=0}^{j-1} |\Delta(\lambda^i)| \leq \|\lambda^0 - \bar{\lambda}\| + \frac{\ell_1 |\Delta(\lambda^0)|}{2} \sum_{i=0}^{j-1} q_2^i \leq \rho_0 + \frac{\ell_1 |\Delta(\lambda^0)|}{2(1-q_2)} \leq \rho \end{aligned} \quad (131)$$

для всех $j = 0, 1, \dots, k+1$, и, в частности,

$$\|\lambda^{k+1} - \bar{\lambda}\| \leq \rho. \quad (132)$$

Из предпоследнего неравенства в (131) и из (122), (129) далее находим

$$\|\xi^{k+1} - \bar{\xi}\| \leq \ell_2 \|\lambda^k - \bar{\lambda}\| \leq \ell_2 \left(\rho_0 + \frac{\ell_1 |\Delta(\lambda^0)|}{2(1-q_2)} \right) \leq \rho. \quad (133)$$

Наконец, из соотношений (127), (128) и (130) выводим

$$\begin{aligned} \|\eta^{k+1}\| &\leq \|\eta^0\| + \sum_{j=0}^k \|\eta^{j+1} - \eta^j\| \leq \|\eta^0\| + \ell_2 \sum_{j=0}^k \text{dist}((\xi^j, \lambda^j), \mathcal{M}_{pd}) \leq \\ &\leq \|\eta^0\| + \ell_2 \text{dist}((\xi^0, \lambda^0), \mathcal{M}_{pd}) + \ell_2^2 \sum_{j=0}^{k-1} |\Delta(\lambda^j)| \leq \\ &\leq \|\eta^0\| + \ell_2 \text{dist}((\xi^0, \lambda^0), \mathcal{M}_{pd}) + \ell_2^2 |\Delta(\lambda^0)| \sum_{j=0}^{k-1} q_2^j. \end{aligned}$$

Поскольку $\text{dist}((\xi^0, \lambda^0), \mathcal{M}_{pd}) \leq \|(\xi^0 - \bar{\xi}, \lambda^0 - \bar{\lambda})\| \leq \sqrt{2}\rho_0$, из полученного неравенства и из (116), (123) получаем

$$\|\eta^{k+1}\| \leq \ell_0 + \sqrt{2}\ell_2\rho_0 + \frac{\ell_2^2 |\Delta(\lambda^0)|}{1-q_2} \leq \ell_1. \quad (134)$$

Комбинируя (132)–(134) и используя определение η^{k+1} , можно заключить, что $(\xi^{k+1}, \lambda^{k+1}) \in \mathcal{B}(\bar{\xi}, \rho) \times (\mathcal{B}(\bar{\lambda}, \rho) \setminus \mathcal{M}_0)$, и соотношение (125) выполнено для $j = k+1$.

Приводимое выше рассуждение показывает, что итерационная схема (117) генерирует корректно определенную траекторию $\{(\xi^k, \lambda^k)\} \subset \mathcal{B}(\bar{\xi}, \rho) \times (\mathcal{B}(\bar{\lambda}, \rho) \setminus \mathcal{M}_0)$, удовлетворяющую неравенству (118) (для всех k) и неравенствам (126)–(129) (и, следовательно, (130), для всех j). Докажем теперь, что эта траектория сходится к точке $(\xi(\lambda^*), \lambda^*)$ для некоторого множителя $\lambda^* \in \mathcal{M}_0$. Действительно, для любых неотрицательных чисел k и j , используя (122), второе соотношение в (124), а также (125) и (130), получаем

$$\|\lambda^{k+j} - \lambda^k\| \leq \frac{1}{2} \sum_{i=k}^{k+j-1} \|\delta(\xi^i, \lambda^i)\| \leq \frac{\ell_1}{2} \sum_{i=k}^{k+j-1} |\Delta(\lambda^i)| \leq \frac{\ell_1 |\Delta(\lambda^0)|}{2} \sum_{i=k}^{k+j-1} q_2^i \leq \frac{\ell_1 |\Delta(\lambda^0)|}{2(1-q_2)} q_2^k \leq \rho q_2^k.$$

Таким образом, последовательность $\{\lambda^k\} \subset \mathcal{B}(\bar{\lambda}, \rho)$ является фундаментальной, и поэтому она сходится к некоторому множителю $\lambda^* \in \mathcal{B}(\bar{\lambda}, \rho)$.

Далее, согласно первому соотношению в (117), для всех $k = 1, \dots$ выполнено $\|\xi^k\| = 1$, и поэтому у последовательности $\{\xi^k\} \subset \mathcal{B}(\bar{\xi}, \rho)$ есть предельная точка $\tilde{\xi} \in \mathcal{B}(\bar{\xi}, \rho)$, $\|\tilde{\xi}\| = 1$. Из свойства (130) следует предельное соотношение $\lim_{k \rightarrow \infty} \Delta(\lambda^k) = 0$, и поэтому, переходя в неравенстве (128) к пределу вдоль соответствующей подпоследовательности, получаем $(\tilde{\xi}, \lambda^*) \in \mathcal{M}_{pd}$, откуда следует $\lambda^* \in \mathcal{M}_0$ и $\tilde{\xi} \in \ker H(\lambda^*)$. Равенство $\dim \ker H(\lambda^*) = 1$ позволяет заключить, что $\tilde{\xi}$ совпадает либо с $\xi(\lambda^*)$, либо с $-\xi(\lambda^*)$. Однако, из неравенства $\|\xi(\lambda^*) - \bar{\xi}\| < 1$ следует

$$\|\bar{\xi} - (-\xi(\lambda^*))\| = \|2\bar{\xi} - (\bar{\xi} - \xi(\lambda^*))\| \geq 2\|\bar{\xi}\| - \|\bar{\xi} - \xi(\lambda^*)\| > 1,$$

и поскольку $\rho < 1$, то $\tilde{\xi} \in \mathcal{B}(\bar{\xi}, \rho)$ не может совпадать с $-\xi(\lambda^*)$. Таким образом, $\tilde{\xi} = \xi(\lambda^*)$. В частности, $\xi(\lambda^*)$ является единственной предельной точкой последовательности $\{\xi^k\}$, т. е., $\{\xi^k\} \rightarrow \xi(\lambda^*)$.

Покажем теперь, что последовательность $\{\eta^k\}$ сходится. Действительно, из соотношений (127), (128), (130) легко получить, что для любых положительных чисел k и j справедлива цепочка неравенств

$$\begin{aligned} \|\eta^{k+j} - \eta^k\| &\leq \ell_2 \sum_{i=k}^{k+j-1} \text{dist}((\xi^i, \lambda^i), \mathcal{M}_{pd}) \leq \ell_2^2 \sum_{i=k-1}^{k+j-2} |\Delta(\lambda^i)| \leq \ell_2^2 |\Delta(\lambda^0)| \sum_{i=k-1}^{k+j-2} q_2^i \leq \\ &\leq \frac{\ell_2^2 |\Delta(\lambda^0)|}{1 - q_2} q_2^{k-1} \leq \ell_1 q_2^{k-1}. \end{aligned}$$

где последнее неравенство следует из (123). Таким образом, последовательность $\{\eta^k\}$ является фундаментальной, и поэтому она сходится к некоторому вектору $\tilde{\eta} \in \mathbb{R}^l$. Очевидно, из (118) следует $\|\tilde{\eta}\| \leq \ell_1$. Более того, используя оценку (85) в лемме 12, получаем

$$\langle \Delta'(\lambda^*), \tilde{\eta} \rangle = 1, \quad (135)$$

и, в частности, $\tilde{\eta} \neq 0$.

Теперь докажем соотношения (119) и (120). Как следует из условия (94) в лемме 13, для любого $\varepsilon \in (0, 1/2)$ справедливо

$$\left| \frac{\Delta(\lambda^{k+1})}{\Delta(\lambda^k)} - \frac{1}{2} \right| \leq \varepsilon \quad (136)$$

для достаточно больших k . Заново обозначим $q_1 = 1/2 - \varepsilon$, $q_2 = 1/2 + \varepsilon$. Тогда из (136) следует, что для любых таких k и любого неотрицательного числа j выполнено

$$q_1^j \leq \frac{\Delta(\lambda^{k+j})}{\Delta(\lambda^k)} \leq q_2^j. \quad (137)$$

Отсюда далее получаем

$$\frac{1 - q_1^j}{1 - q_1} \leq \sum_{i=k}^{k+j-1} \frac{\Delta(\lambda^i)}{\Delta(\lambda^k)} \leq \frac{1 - q_2^j}{1 - q_2}. \quad (138)$$

Заметим, что из второго соотношения в (117) следует

$$\lambda^k - \lambda^{k+j} = \frac{1}{2} \sum_{i=k}^{k+j-1} \delta(\xi^i, \lambda^i) = \frac{1}{2} \sum_{i=k}^{k+j-1} \Delta(\lambda^i) \eta^i,$$

и поэтому,

$$\frac{\lambda^k - \lambda^{k+j}}{\Delta(\lambda^k)} - \tilde{\eta} = \frac{1}{2} \sum_{i=k}^{k+j-1} \frac{\Delta(\lambda^i)}{\Delta(\lambda^k)} (\eta^i - \tilde{\eta}) + \frac{1}{2} \left(\sum_{i=k}^{k+j-1} \frac{\Delta(\lambda^i)}{\Delta(\lambda^k)} - 2 \right) \tilde{\eta}.$$

Таким образом, поскольку $\|\eta^k - \tilde{\eta}\| \leq \varepsilon$ для всех достаточно больших k , то из (137) и (138) получаем

$$\begin{aligned} \left\| \frac{\lambda^k - \lambda^{k+j}}{\Delta(\lambda^k)} - \tilde{\eta} \right\| &\leq \frac{1}{2} \sum_{i=k}^{k+j-1} \left| \frac{\Delta(\lambda^i)}{\Delta(\lambda^k)} \right| \|\eta^i - \tilde{\eta}\| + \frac{\|\tilde{\eta}\|}{2} \left| \sum_{i=k}^{k+j-1} \frac{\Delta(\lambda^i)}{\Delta(\lambda^k)} - 2 \right| \leq \\ &\leq \frac{\varepsilon}{2} \sum_{i=k}^{k+j-1} q_2^{i-k} + \frac{\|\tilde{\eta}\|}{2} \max \left\{ \left| \frac{1 - q_1^j}{1 - q_1} - 2 \right|, \left| \frac{1 - q_2^j}{1 - q_2} - 2 \right| \right\} \leq \\ &\leq \frac{\varepsilon}{2(1 - q_2)} + \frac{\|\tilde{\eta}\|}{2} \max \left\{ \left| \frac{1 - q_1^j}{1 - q_1} - 2 \right|, \left| \frac{1 - q_2^j}{1 - q_2} - 2 \right| \right\}. \end{aligned}$$

Переходя к пределу при $j \rightarrow \infty$, получаем, что для всех достаточно больших k выполнено

$$\left\| \frac{\lambda^k - \lambda^*}{\Delta(\lambda^k)} - \tilde{\eta} \right\| \leq \frac{\varepsilon}{2(1 - q_2)} + \frac{\|\tilde{\eta}\|}{2} \max \left\{ \left| \frac{1}{1 - q_1} - 2 \right|, \left| \frac{1}{1 - q_2} - 2 \right| \right\} \leq \frac{\varepsilon}{1 - 2\varepsilon} (1 + 2\|\tilde{\eta}\|),$$

где также были приняты во внимание определения q_1 и q_2 . Учитывая произвольность выбора ε , заключаем, что

$$\left\{ \frac{\lambda^k - \lambda^*}{\Delta(\lambda^k)} \right\} \rightarrow \tilde{\eta}. \quad (139)$$

Из этого соотношения, очевидно, следует

$$\lim_{k \rightarrow \infty} \frac{\|\lambda^{k+1} - \lambda^*\|}{\|\lambda^k - \lambda^*\|} \left| \frac{\Delta(\lambda^k)}{\Delta(\lambda^{k+1})} \right| = 1,$$

и поэтому (119) сразу же следует из (136) (напомним, что последнее соотношение справедливо для произвольного $\varepsilon > 0$ для всех достаточно больших k).

Наконец, из соотношения (139) также следует (120) при $d = \tilde{\eta}/\|\tilde{\eta}\|$. Поскольку $\|\tilde{\eta}\| \leq \ell_1$, то из (135) получаем $\langle \Delta'(\lambda^*), d \rangle = 1/\|\tilde{\eta}\| \geq 1/\ell_1$, доказывая (121). ■

Сформулируем теперь основной результат данного раздела, который легко следует из предложения 4, равенства

$$\frac{x^{k+1}}{\|x^k\|} = \frac{1}{2} P(\xi^k, \lambda^k) \xi^k,$$

(которое легко получить из первого соотношения в (63)) и из условия $\|P(\xi^k, \lambda^k) \xi^k\| \rightarrow 1$ (которое следует из (95) в лемме 13).

Теорема 1. Пусть выполнено условие (64) для некоторого $\bar{\lambda} \in \mathcal{M}_0$, а также условия (74), (80) для некоторого $\bar{\xi} \in \ker(A + \bar{\lambda}B)$.

Тогда для любых $\ell_0 > 0$ и $\ell_1 > 0$, $\ell_0 < \ell_1$, и для любого $\rho > 0$ найдется такая константа $\rho_0 \in (0, \rho)$, что для любой начальной точки $(x^0, \lambda^0) \in (\mathbb{R}^n \setminus \{0\}) \times (\mathcal{B}(\bar{\lambda}, \rho_0) \setminus \mathcal{M}_0)$, удовлетворяющей $x^0/\|x^0\| \in \mathcal{B}(\bar{\xi}, \rho_0)$ и

$$\frac{\|\delta(x^0/\|x^0\|, \lambda^0)\|}{|\det(A + \lambda^0 B)|} \leq \ell_0, \quad (140)$$

где $\delta(x^0/\|x^0\|, \lambda^0)$ корректно определено первым соотношением в (62), существует единственная траектория $\{(x^k, \lambda^k)\} \subset (\mathbb{R}^n \setminus \{0\}) \times (\mathcal{B}(\bar{\lambda}, \rho) \setminus \mathcal{M}_0)$, удовлетворяющая (58) для всех $k = 0, 1, \dots$; эта траектория сходится к $(0, \lambda^*)$ для некоторого множителя $\lambda^* \in \mathcal{M}_0$, последовательность $\{x^k/\|x^k\|\} \subset \mathcal{B}(\bar{\xi}, \rho)$ сходится к некоторому $\xi^* \in \ker(A + \lambda^* B)$,

$$\lim_{k \rightarrow \infty} \frac{\|x^{k+1}\|}{\|x^k\|} = \frac{1}{2},$$

и свойства (119), (120) выполнены для некоторого $d \in \mathbb{R}^l$, удовлетворяющего неравенству (121) (из которого, в частности, следует, что вектор d не является касательным к \mathcal{M}_0 в точке λ^*).

Утверждение данной теоремы, в частности, означает, что $(\xi^*, \lambda^*) \rightarrow (\bar{\xi}, \bar{\lambda})$, если $(x^0/\|x^0\|, \lambda^0) \rightarrow (\bar{\xi}, \bar{\lambda})$.

Заметим, что поведение, наблюдаемое в примере 4 (например, продемонстрированное на рис. 4) не противоречит теореме 1. Дело в том, что последовательность $\{x^k/\|x^k\|\}$ остается отделенной от $\bar{\xi} = \xi(\bar{\lambda}) = \pm(10, -5, -3)/\sqrt{134}$, что, в свою очередь, не дает двойственной траектории сходиться к единственному критическому множителю $\bar{\lambda} = 5$, в котором выполнены предположения (64), (74) и (80). Такое поведение возможно, если $x^0/\|x^0\|$ не достаточно близко к $\bar{\xi}$.

Начальная часть траектории, показанной на рис. 5, также свидетельствует о том, что $x^0/\|x^0\|$ должно быть достаточно близко к $\bar{\xi}$ для того, чтобы имело место утверждение теоремы 1. Более того, согласно полученному вычислительному опыту, близость $x^0/\|x^0\|$ к $\bar{\xi}$ играет наиболее важную роль для сходимости двойственной траектории к критическому множителю в окрестности $\bar{\lambda}$.

Следующий пример показывает, что предположение (140), ограничивающее множество начальных точек, также нельзя опустить.

Пример 6. (Пример 2.2 из [67], тест 20301 из коллекции DEGEN [42].) Рассмотрим задачу оптимизации

$$x_1^2 - x_2^2 + x_3^2 \rightarrow \min, \quad x_1^2 + x_2^2 - x_3^2 = 0, \quad x_1 x_3 = 0.$$

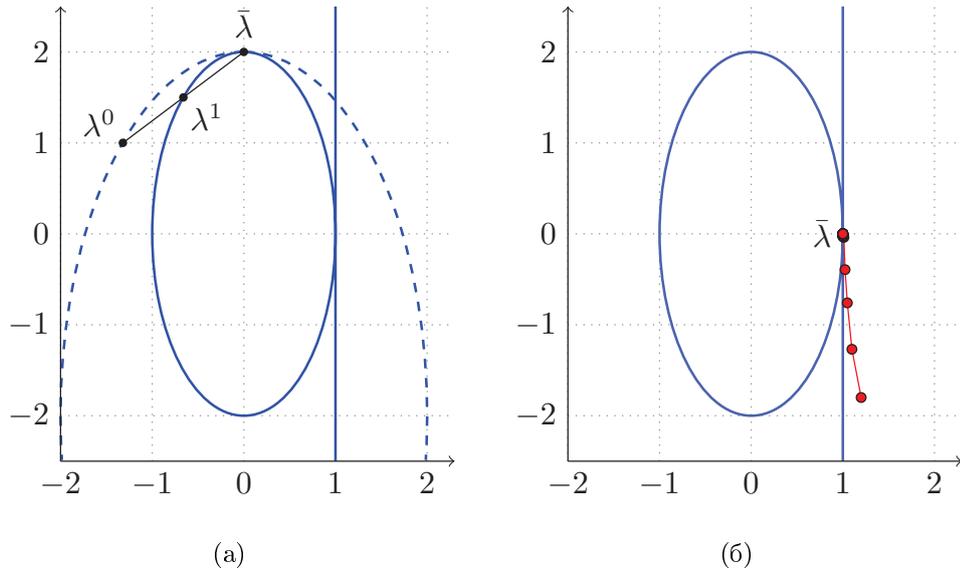


Рис. 7. Иллюстрация к примеру 6.

Здесь $\Delta(\lambda) = -2(1 - \lambda_1)(4 - 4\lambda_1^2 - \lambda_2^2)$, и поэтому критическими являются множители $\lambda \in \mathbb{R}^2$, для которых $\lambda_1 = 1$ или $4\lambda_1^2 + \lambda_2^2 = 4$ (вертикальная линия и эллипс на рис. 7).

Несложно убедиться в том, что предположения (64), (74) и (80) выполняются для всех критических множителей $\bar{\lambda} \in \mathbb{R}^2$, кроме тех, для которых выполнено $\bar{\lambda}_1 = 1$, и для всех векторов $\bar{\xi} \in \ker H(\bar{\lambda})$ единичной нормы. (Если $\bar{\lambda} = (1, 0)$, то $\dim \ker H(\bar{\lambda}) = 2$. Если же $\bar{\lambda}_1 = 1$, $\bar{\lambda}_2 \neq 0$, то $\ker H(\bar{\lambda})$ натянута на вектор $\bar{\xi} = (0, 1, 0)$, но при этом $\text{rank } B[\bar{\xi}] = 1$, а значит (74) не выполняется.) Рассмотрим, например, множитель $\bar{\lambda} = (0, 2)$ и вектор $\bar{\xi} = (1/\sqrt{2}, 0, -1/\sqrt{2}) \in \ker H(\bar{\lambda})$ и возьмем произвольный множитель $\lambda^0 \in \mathbb{R}^2$ из эллипса $4(\lambda_1^0)^2 + (\lambda_2^0 + 2)^2 = 16$. Заметим, что среди таких λ^0 существуют сколь угодно близкие к $\bar{\lambda}$, а также что для любого $\ell_0 > 0$ неравенство (116) не выполняется для $\xi^0 = \bar{\xi}$ при условии, что множитель λ^0 достаточно близок к $\bar{\lambda}$. Далее, легко проверить, что $\delta(\bar{\xi}, \lambda^0) = (\lambda_1^0, \lambda_2^0 - 2)$, и поэтому $\lambda^1 = \lambda^0 - \delta(\bar{\xi}, \lambda^0)/2 = (\lambda_1^0/2, \lambda_2^0/2 + 1)$. В частности, $4(\lambda_1^1)^2 + (\lambda_2^1)^2 = 4$, т. е. множитель λ^1 является критическим, а значит утверждение теоремы не выполнено (см. рис. 7(a)). В частности, выражение $\delta(x^1/\|x^1\|, \lambda^1)$ не является корректно определенным, и представленный выше анализ неприменим.

В то же время, предположение (140) на самом деле не является слишком ограничивающим. Действительно, следующее утверждение показывает, что если ρ_0 достаточно мало, то условие (116) при некотором $\ell_0 > 0$ выполнено для всех $\xi^0 \in \mathcal{B}(\bar{\xi}, \rho_0)$ и для всех λ^0 из $\mathcal{B}(\bar{\lambda}, \rho_0) \setminus \mathcal{M}_0$, кроме «узкого» множества вокруг \mathcal{M}_0 .

Предложение 5. *В условиях леммы 12 найдется такое $\rho > 0$, что для любого $\varepsilon \in (0, 1)$ существует константа $\ell_0 > 0$, обладающая следующими свойствами: для любого $\rho_0 \in (0, \rho]$ условие (116) выполнено для всех $(\xi^0, \lambda^0) \in \mathcal{B}(\bar{\xi}, \rho_0) \times \mathcal{B}(\bar{\lambda}, \rho_0)$, для которых*

$$\text{dist}(\lambda^0, \mathcal{M}_0) \geq \varepsilon \rho_0.$$

Доказательство. Зафиксируем произвольную константу $\rho > 0$, для которой выполнены утверждения леммы 11 и следствия 1. Как следует из (78), для всех $(\xi, \lambda) \in \mathcal{B}(\bar{\xi}, \rho) \times (\mathcal{B}(\bar{\lambda}, \rho) \setminus \mathcal{M}_0)$ выполнено

$$\|\delta(\xi, \lambda) - (\lambda - \bar{\lambda})\| \leq \ell \|\xi - \bar{\xi}\|,$$

и поэтому

$$\|\delta(\xi, \lambda)\| \leq \|\lambda - \bar{\lambda}\| + \ell \|\xi - \bar{\xi}\| \leq (1 + \ell) \max\{\|\lambda - \bar{\lambda}\|, \|\xi - \bar{\xi}\|\}.$$

Применяя второе неравенство в (81), далее получаем

$$\left\| \frac{\delta(\xi, \lambda)}{\Delta(\lambda)} \right\| \leq C_2(1 + \ell) \frac{\max\{\|\lambda - \bar{\lambda}\|, \|\xi - \bar{\xi}\|\}}{\text{dist}(\lambda, \mathcal{M}_0)}.$$

Таким образом, для любого $\varepsilon \in (0, 1)$, если $(\xi^0, \lambda^0) \in \mathcal{B}(\bar{\xi}, \rho_0) \times \mathcal{B}(\bar{\lambda}, \rho_0)$ и $\text{dist}(\lambda^0, \mathcal{M}_0) \geq \varepsilon \rho_0$ для некоторого $\rho_0 \in (0, \rho]$, то

$$\left\| \frac{\delta(\xi^0, \lambda^0)}{\Delta(\lambda^0)} \right\| \leq C_2(1 + \ell) \frac{\rho_0}{\varepsilon \rho_0} \leq \ell_0,$$

где константа $\ell_0 = C_2(1 + \ell)/\varepsilon$ не зависит от ρ_0 и (ξ^0, λ^0) . ■

Из теоремы 1 и предложения 5 следует, что для любого $\varepsilon \in (0, 1)$ найдется такая константа $\rho_0 > 0$, что для любой начальной точки $(x^0, \lambda^0) \subset (\mathbb{R}^n \setminus \{0\}) \times \mathcal{B}(\bar{\lambda}, \rho_0)$, удовлетворяющей $x^0/\|x^0\| \in \mathcal{B}(\bar{\xi}, \rho_0)$ и $\text{dist}(\lambda^0, \mathcal{M}_0) \geq \varepsilon \rho_0$, рассматриваемый метод сходится к критическому множителю.

Отметим также, что условие (140) в теореме 1 можно опустить, если в рассматриваемой задаче есть только одно ограничение, т. е. если $l = 1$. Действительно, в этом случае из первого равенства в (62) и из (76) следует, что

$$\frac{\delta(\xi, \lambda)}{\Delta(\lambda)} = \frac{B[\xi, \xi]}{R(B[\xi], B[\xi], H(\lambda))}$$

для всех $\xi \in \mathcal{B}(\bar{\xi}, \rho)$ и $\lambda \in (\mathcal{B}(\bar{\lambda}, \rho) \setminus \mathcal{M}_0)$ при условии, что $\rho > 0$ достаточно мало. Таким образом, из предположения (74) следует существование таких констант $\rho_0 > 0$ и $\ell_0 > 0$, что условие (116) выполнено для всех $(\xi^0, \lambda^0) \in \mathcal{B}(\bar{\xi}, \rho_0) \times (\mathcal{B}(\bar{\lambda}, \rho_0) \setminus \mathcal{M}_0)$.

Другое интересное наблюдение состоит в том, что предположение (74) также можно опустить в случае $l = 1$. Действительно, в этом случае для любых $\xi \in \mathbb{R}^n$ и $\lambda \in \mathbb{R}^l \setminus \{0\}$ выполнено

$$\det(B[\xi](H(\lambda))^{-1}(B[\xi])^T) \det H(\lambda) = B[\xi] \hat{H}(\lambda) (B[\xi])^T,$$

и поэтому свойством, необходимым для доказательства леммы 8, является не (74), а лишь $B[\bar{\xi}]\hat{H}(\bar{\lambda})(B[\bar{\xi}])^T \neq 0$. С учетом равенства $\hat{H}(\bar{\lambda}) = t(\bar{\lambda})\bar{\xi}\bar{\xi}^T$ (которое следует из леммы 5), требуемое свойство эквивалентно (80).

Однако, предположение (74) нельзя опустить в общем случае: без него нельзя гарантировать невырожденность матрицы в левой части системы (58). Действительно, рассмотрим произвольную задачу оптимизации, в которой одно из ограничений тождественно равно нулю. Ясно, что предположения (64) и (80) вполне могут быть выполнены, однако $\text{rank } B[\bar{\xi}] < l$, и поэтому (74) нарушается. В этом примере матрица в левой части системы (58) всегда содержит нулевую строку. Следовательно, решение данной системы определяется не единственным образом, а значит двойственная траектория вовсе не обязана сходиться. В частности, утверждение теоремы 1 в этом случае не выполнено.

Заметим далее, что два предположения (64), (80) на самом деле эквивалентны условию (82). Действительно, если выполнено (64), то из леммы 6 следует, что (82) эквивалентно (80). Таким образом, достаточно показать, что из (82) следует (64). Для этого, в свою очередь, достаточно показать, что для любых матриц $H, B \in \mathbb{R}^{n \times n}$, где $\dim \ker H \geq 2$, выполнено

$$\left. \frac{d}{dt} \det(H + tB) \right|_{t=0} = 0. \quad (141)$$

Если матрица B невырождена, то $\dim \ker B^{-1}H \geq 2$, а значит нулевое собственное значение матрицы $B^{-1}H$ имеет кратность не меньше 2. Таким образом, $\det(B^{-1}H - tI) = O(t^2)$, откуда следует

$$\left. \frac{d}{dt} \det(H + tB) \right|_{t=0} = -\det B \left. \frac{d}{dt} \det(B^{-1}H - tI) \right|_{t=0} = 0.$$

Если же матрица B вырождена, рассмотрим произвольную последовательность $\{B_k\} \subset \mathbb{R}^{n \times n}$ невырожденных матриц, сходящуюся к B (очевидно, такая последовательность всегда существует). Тогда, согласно сказанному выше,

$$\left. \frac{d}{dt} \det(H + tB_k) \right|_{t=0} = 0$$

для всех k . Переходя к пределу, вновь получаем (141).

Возвращаясь к примеру 6, рассмотрим критические множители, для которых $\lambda_1 = 1$. Предположения (64) и (80) (или, что то же самое, (82)) выполнены для всех таких множителей за исключением $(1, 0)$. Однако для любого множителя $\bar{\lambda} = (1, \bar{\lambda}_2)$, где $\bar{\lambda}_2 \neq 0$, выполнено $\text{rank } B[\bar{\xi}] \leq 1$ для $\bar{\xi} = (0, 1, 0) \in \ker H(\bar{\lambda})$, и поэтому предположение (74) нарушается для таких $\bar{\lambda}$. Более того, предположение (64) (и, следовательно, (82)) нарушается для $\bar{\lambda} = (1, 0)$, поскольку $\dim \ker H(\bar{\lambda}) = 2$. При этом, если рассматривать начальные точки (x^0, λ^0) , для которых $(x^0/\|x^0\|, \lambda^0)$ достаточно близко к $\{(0, 1, 0)\} \times (\{1\} \times \mathbb{R})$, то двойственные траектории во многих случаях будут сходиться к $\bar{\lambda} = (1, 0)$, а не к критическому множителю, близкому к λ^0 . Пример такого поведения показан на рис. 7(б) (где $x^0 = (0.1, 1.1, 0.1)$ и $\lambda^0 = (1.2, -1.8)$).

Таким образом, критические множители большей «кратности» (т. е. те, для которых нарушается (64) и, следовательно, (82)) притягивают двойственные траектории даже сильнее, чем другие критические множители.

В завершение отметим, что предположения (64), (75), (80) обычно выполнены «практически для всех» критических множителей. При этом не удалось найти ни одного примера, в котором выполнены предположения (64), (75), (80), но при этом нарушается (74). Поэтому гипотеза состоит в том, что в случае выполнения (64) и (80) условие (74) на самом деле эквивалентно (75).

1.3. Локальная двойственная стабилизация

В предыдущих разделах было показано, что эффект притяжения двойственных траекторий к критическим множителям является крайне типичным для метода последовательного квадратичного программирования: двойственные траектории метода двигаются практически параллельно множеству множителей, постепенно приближаясь к критическим множителям.

Данный раздел посвящен обсуждению двух методов, обладающих свойством локальной двойственной стабилизации, т. е. способностью подавлять эффект притяжения к критическим множителям: стабилизированного метода последовательного квадратичного программирования и метода модифицированных функций Лагранжа. Для этих методов эффект притяжения значительно менее выражен: существуют широкие области, из которых метод сходится сверхлинейно, и при этом двойственная траектория сходится к некритическому множителю.

1.3.1. Стабилизированный метод последовательного квадратичного программирования

Стабилизированный метод последовательного квадратичного программирования (sSQP, от англ. Stabilized sequential quadratic programming) был изначально введен в [93] для задач с ограничениями-неравенствами и состоял в решении последовательности минимаксных подзадач. Позднее было показано, что такие минимаксные подзадачи эквивалентны задачам квадратичного программирования в прямо-двойственном пространстве (см. [81]), и что метод применим также к задачам с ограничениями-равенствами.

Итерация метода заключается в следующем. 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

$$\begin{aligned} \langle f'(x), \xi \rangle + \frac{1}{2} \left\langle \frac{\partial^2 L}{\partial x^2}(x, \lambda, \mu) \xi, \xi \right\rangle + \frac{\sigma}{2} (\|\lambda + \eta\|^2 + \|\mu + \zeta\|^2) \rightarrow \min_{(\xi, \eta, \zeta)} \quad (142) \\ h(x) + h'(x)\xi - \sigma\eta = 0, \quad g(x) + g'(x)\xi - \sigma\zeta \leq 0, \end{aligned}$$

где σ — параметр стабилизации, и следующее приближение полагается равным $(x + \xi, \lambda + \eta, \mu + \zeta)$. При отсутствии ограничений-неравенств итерацию можно записать в более простом виде: для текущего приближения $(x, \lambda) \in \mathbb{R}^n \times \mathbb{R}^l$ вычисляется решение $(\xi, \eta) \in \mathbb{R}^n \times \mathbb{R}^l$ системы

$$\frac{\partial^2 L}{\partial x^2}(x, \lambda)\xi + (h'(x))^T \eta = -\frac{\partial L}{\partial x}(x, \lambda), \quad h'(x)\xi - \sigma\eta = -h(x), \quad (143)$$

и следующее приближение полагается равным $(x + \xi, \lambda + \eta)$. В этом случае метод также называют стабилизированным методом Ньютона–Лагранжа (СМНЛ), поскольку система (143) отличается от системы (16) метода Ньютона–Лагранжа только наличием стабилизирующего члена $-\sigma\eta$ во втором уравнении.

Одно из сильных свойств данного метода состоит в том, что для сверхлинейной локальной сходимости не требуется выполнение условий регулярности ограничений. Данное свойство устанавливается приводимой ниже теоремой 2, которая представляет собой результат из [48, теорема 5] с некоторыми улучшениями, которые могут быть получены из доказательства в [50, теорема 1] или [64, теорема 4.1]; см. также [72, глава 7].

Теорема 2. Пусть $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $h: \mathbb{R}^n \rightarrow \mathbb{R}^l$ и $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$ дважды дифференцируемы в окрестности точки $\bar{x} \in \mathbb{R}^n$, а их вторые производные непрерывны в этой точке. Пусть \bar{x} — стационарная точка задачи (1), и пусть для отвечающего ей множителя Лагранжа $(\bar{\lambda}, \bar{\mu}) \in \mathbb{R}^l \times \mathbb{R}^m$ выполняется SOSC (7). Пусть $\sigma: \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m \rightarrow \mathbb{R}_+$ — произвольная функция, для которой справедливы оценки расстояния

$$\beta_1(\|x - \bar{x}\| + \text{dist}((\lambda, \mu), \mathcal{M}(\bar{x}))) \leq \sigma(x, \lambda, \mu) \leq \beta_2(\|x - \bar{x}\| + \text{dist}((\lambda, \mu), \mathcal{M}(\bar{x})))$$

для некоторых $\beta_1, \beta_2 > 0$ и для всех (x, λ, μ) , достаточно близких к $(\bar{x}, \bar{\lambda}, \bar{\mu})$.

Тогда для любого достаточно большого $\ell > 0$ найдется окрестность \mathcal{V} точки $(\bar{x}, \bar{\lambda}, \bar{\mu})$ и непрерывная в нуле функция $\chi: \mathbb{R}_+ \rightarrow \mathbb{R}_+$, удовлетворяющая оценке $\chi(t) = o(t)$ при $t \rightarrow 0$, для которых справедливы следующие утверждения: для любого начального приближения $(x^0, \lambda^0, \mu^0) \in \mathcal{V} \cap (\mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}_+^m)$ существует такая траектория $\{(x^k, \lambda^k, \mu^k)\} \subset \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m$, что для всех k направление $(\xi^k, \eta^k, \zeta^k) = (x^{k+1} - x^k, \lambda^{k+1} - \lambda^k, \mu^{k+1} - \mu^k)$ удовлетворяет системе ККТ задачи (142) при $(x, \lambda, \mu) = (x^k, \lambda^k, \mu^k)$, $\sigma = \sigma(x^k, \lambda^k, \mu^k)$ и условию локализации

$$\|(\xi^k, \eta^k, \zeta^k)\| \leq \ell(\|x^k - \bar{x}\| + \text{dist}((\lambda^k, \mu^k), \mathcal{M}(\bar{x}))); \quad (144)$$

любая такая траектория сходится к некоторой точке $(\bar{x}, \lambda^*, \mu^*)$, где $(\lambda^*, \mu^*) \in \mathcal{M}(\bar{x})$, и для каждого k справедливо неравенство

$$\|x^{k+1} - \bar{x}\| + \|\lambda^{k+1} - \lambda^*\| + \|\mu^{k+1} - \mu^*\| \leq \chi(\|x^k - \bar{x}\| + \text{dist}((\lambda^k, \mu^k), \mathcal{M}(\bar{x}))). \quad (145)$$

Заметим, что из (145) вытекают следующие оценки для всех k :

$$\|x^{k+1} - \bar{x}\| + \text{dist}((\lambda^{k+1}, \mu^{k+1}), \mathcal{M}(\bar{x})) \leq \chi(\|x^k - \bar{x}\| + \text{dist}((\lambda^k, \mu^k), \mathcal{M}(\bar{x}))), \quad (146)$$

$$\|x^{k+1} - \bar{x}\| + \|\lambda^{k+1} - \lambda^*\| + \|\mu^{k+1} - \mu^*\| \leq \chi(\|x^k - \bar{x}\| + \|\lambda^k - \lambda^*\| + \|\mu^k - \mu^*\|). \quad (147)$$

В частности, скорость сходимости последовательности $\{(x^k, \lambda^k, \mu^k)\}$ к $(\bar{x}, \lambda^*, \mu^*)$ и последовательности $\{\|x^k - \bar{x}\| + \text{dist}((\lambda^k, \mu^k), \mathcal{M}(\bar{x}))\}$ к нулю является сверхлинейной.

Согласно предложению 1, требованиям теоремы удовлетворяет, например, функция естественной невязки $\sigma: \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m \rightarrow \mathbb{R}_+$ системы ККТ (2):

$$\sigma(x, \lambda, \mu) = \left\| \left(\frac{\partial L}{\partial x}(x, \lambda, \mu), h(x), \min\{\mu, -g(x)\} \right) \right\|.$$

При отсутствии ограничений-неравенств теорему 2 можно уточнить. Во-первых, как показано в [71], SOS можно заменить на некритичность множителя $\bar{\lambda}$. Во-вторых, для любой точки (x, λ) , достаточно близкой к решению, но не удовлетворяющей системе Лагранжа (12), система (143) имеет единственное решение. Таким образом, имеет место следующий факт, который представляет собой результат из [71, теорема 1] с некоторыми улучшениями, которые могут быть получены из доказательства в [50, теорема 1].

Теорема 3. Пусть $f: \mathbb{R}^n \rightarrow \mathbb{R}$ и $h: \mathbb{R}^n \rightarrow \mathbb{R}^l$ дважды дифференцируемы в окрестности точки $\bar{x} \in \mathbb{R}^n$, а их вторые производные непрерывны в этой точке. Пусть \bar{x} — стационарная точка задачи (11), и пусть множитель $\bar{\lambda} \in \mathcal{M}(\bar{x})$ является некритическим. Пусть $\sigma: \mathbb{R}^n \times \mathbb{R}^l \rightarrow \mathbb{R}$ — любая функция, непрерывная и равная нулю во всех точках множества $\{\bar{x}\} \times \mathcal{M}(\bar{x})$, достаточно близких к $(\bar{x}, \bar{\lambda})$, и такая, что для некоторого $C > 0$ и всех $(x, \lambda) \notin \{\bar{x}\} \times \mathcal{M}(\bar{x})$, достаточно близких к $(\bar{x}, \bar{\lambda})$, справедливо

$$\sigma(x, \lambda) \neq 0, \quad \|x - \bar{x}\| \leq C|\sigma(x, \lambda)|.$$

Тогда справедливы следующие утверждения:

- а) в некоторой окрестности \mathcal{V} точки $(\bar{x}, \bar{\lambda})$ существует единственное отображение $d(\cdot) = (\xi(\cdot), \eta(\cdot)): \mathcal{V} \rightarrow \mathbb{R}^n \times \mathbb{R}^l$ со следующими свойствами: $(\xi(x, \lambda), \eta(x, \lambda))$ удовлетворяет системе (143) при $\sigma = \sigma(x, \lambda)$ для всех $(x, \lambda) \in \mathcal{V}$, и $d(\bar{x}, \lambda) = 0$, если $\lambda \in \mathcal{M}(\bar{x})$;

б) окрестность \mathcal{V} можно выбрать таким образом, что найдется константа $\ell > 0$ и непрерывная в нуле функция $\chi: \mathbb{R}_+ \rightarrow \mathbb{R}_+$, удовлетворяющая оценке $\chi(t) = o(t)$ при $t \rightarrow 0$, для которых будут справедливы оценки

$$\|d(x, \lambda)\| \leq \ell(\|x - \bar{x}\| + \text{dist}(\lambda, \mathcal{M}(\bar{x}))), \quad (148)$$

$$\|x + \xi(x, \lambda) - \bar{x}\| + \text{dist}(\lambda + \eta(x, \lambda), \mathcal{M}(\bar{x})) \leq \chi(\|x - \bar{x}\| + \text{dist}(\lambda, \mathcal{M}(\bar{x}))) \quad (149)$$

для всех $(x, \lambda) \in \mathcal{V}$;

в) существует $M > 0$, такое что для любого $\varepsilon > 0$ найдется такое $\varepsilon_0 > 0$, что для любого $(x^0, \lambda^0) \in \mathcal{B}((\bar{x}, \bar{\lambda}), \varepsilon_0)$ последовательность $\{(x^k, \lambda^k)\}$ корректно определяется равенствами $(x^{k+1}, \lambda^{k+1}) = (x^k, \lambda^k) + d(x^k, \lambda^k)$ для всех k , и для нее справедливо $\{(x^k, \lambda^k)\} \subset \mathcal{B}((\bar{x}, \bar{\lambda}), \varepsilon)$; эта последовательность сходится к (\bar{x}, λ^*) при некотором $\lambda^* = \lambda^*(x^0, \lambda^0) \in \mathcal{M}(\bar{x})$, и для всех k выполнено

$$\|x^{k+1} - \bar{x}\| + \|\lambda^{k+1} - \lambda^*\| \leq M\chi(\|x^k - \bar{x}\| + \text{dist}(\lambda^k, \mathcal{M}(\bar{x}))). \quad (150)$$

Как следует из предложения 1, требованиям теоремы удовлетворяет, например, функция $\sigma: \mathbb{R}^n \times \mathbb{R}^l \rightarrow \mathbb{R}$,

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

для любого фиксированного $\beta \in (0, 1]$.

Из теоремы 2 следует, что если sSQP инициализируется в окрестности множителя $(\bar{\lambda}, \bar{\mu})$, удовлетворяющего SOSC, то траектории метода будут сходиться некоторому множителю (λ^*, μ^*) , близкому к $(\bar{\lambda}, \bar{\mu})$, причем скорость прямо-двойственной сходимости будет сверхлинейной. Если эта окрестность достаточно мала, то множитель (λ^*, μ^*) также будет удовлетворять SOSC и, в частности, будет не критическим. Более того, согласно теореме 3, при отсутствии ограничений-неравенств траектории sSQP будут сходиться к не критическим множителям при условии, что метод инициализируется в достаточно малой окрестности не критического множителя.

Тем не менее, эффект притяжения к критическим множителям присутствует и для sSQP: во многих задачах существуют широкие прямо-двойственные области, при запуске из которых метод будет линейно сходиться к некоторому критическому множителю. Отметим, однако, что для sSQP данный эффект проявляется значительно реже, чем для обычного SQP. Такое поведение подтверждается, в частности, результатами вычислительного эксперимента, приведенными в работе [70].

Чтобы продемонстрировать сравнительное поведение обычного и стабилизированного SQP при наличии критических множителей Лагранжа, вновь рассмотрим задачу из примера 3. На рис. 8 красным и синим цветом показаны прямо-двойственные траектории обычного и стабилизированного МНЛ; вертикальная сплошная линия соответствует множеству

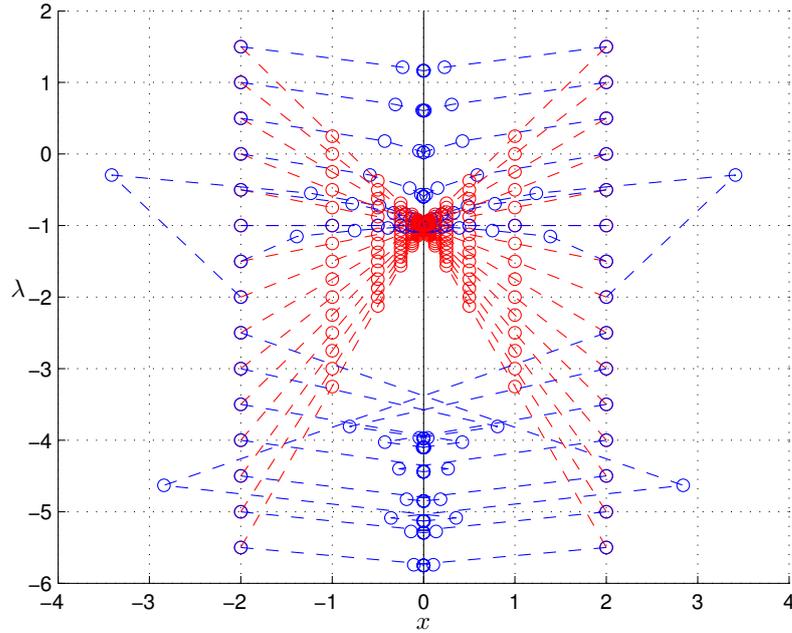


Рис. 8. Траектории МНЛ и СМНЛ для примера 3.

$\{\bar{x}\} \times \mathcal{M}(\bar{x})$. Запуски выполнялись из точек (x^0, λ^0) , где x^0 равнялось -2 или 2 , а λ^0 принимало различные значения из сегмента $[-5.5, 0.5]$. Рис. 8 демонстрирует эффект двойственной стабилизации: траектории МНЛ всегда сходятся к $(\bar{x}, \bar{\lambda})$ с линейной скоростью, в то время как многие траектории СМНЛ сходятся сверхлинейно, причем соответствующая двойственная компонента сходит к некритическому множителю. Отметим, однако, что некоторые траектории СМНЛ по-прежнему сходятся к $(\bar{x}, \bar{\lambda})$ с линейной скоростью.

На рис. 9 серым цветом показана область начальных приближений (x^0, λ^0) , из которых СМНЛ сходит к критическому множителю Лагранжа. Как видно из рисунка, это происходит, если двойственное начальное приближение лежит недостаточно далеко от критического множителя. Подчеркнем, что данная картина не противоречит теореме 3: вокруг каждого некритического множителя существует окрестность, при запуске из которой метод будет сходить сверхлинейно, но размеры этих окрестностей уменьшаются при приближении к критическому множителю.

1.3.2. Метод модифицированных функций Лагранжа

Метод модифицированных функций Лагранжа (МФЛ) или просто метод множителей является одним из традиционных подходов к решению задачи (1). Данный метод восходит к работам [60,90], но продолжает активно развиваться и по сей день; см. также [4,11]. Он также является основой весьма успешных солверов, таких как LANCELOT [79] и ALGENCAN [30].

Определим семейство модифицированных функций Лагранжа задачи (1): $L_c: \mathbb{R}^n \times \mathbb{R}^l \times$

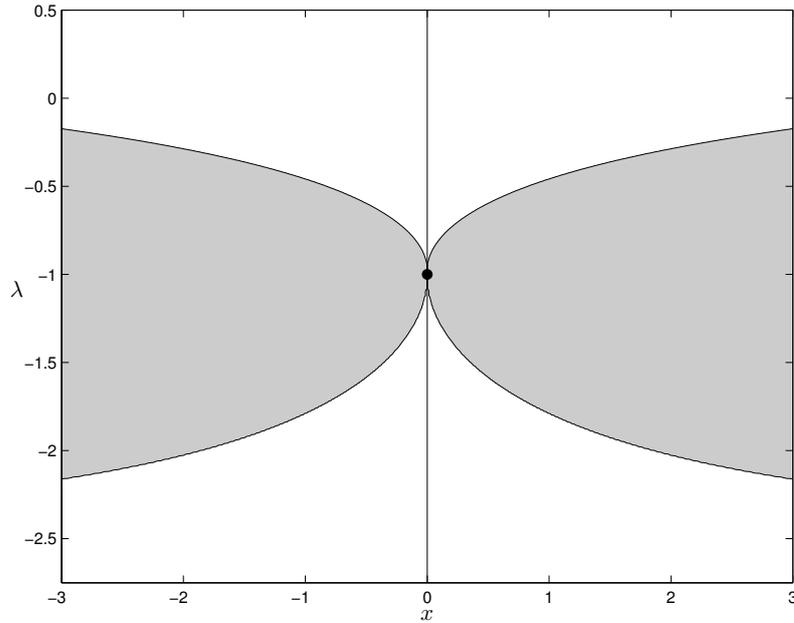


Рис. 9. Область притяжения СМНЛ к критическому множителю для примера 3.

$\mathbb{R}^m \rightarrow \mathbb{R}$,

$$L_c(x, \lambda, \mu) = f(x) + \frac{1}{2c} (\|\lambda + ch(x)\|^2 + \|\max\{0, \mu + cg(x)\}\|^2), \quad (151)$$

где $c > 0$ — параметр штрафа, а максимум берется покомпонентно. Базовый вариант метода модифицированных функций состоит в следующем.

Алгоритм 1. Выбираем $(x^0, \lambda^0, \mu^0) \in \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m$, фиксируем $c_0 > 0$ и полагаем $k = 0$.

1. Вычисляем $x^{k+1} \in \mathbb{R}^n$ как стационарную точку задачи безусловной оптимизации

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

т. е.

$$\frac{\partial L_{c_k}}{\partial x}(x^{k+1}, \lambda^k, \mu^k) = 0.$$

2. Полагаем

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

где максимум берется покомпонентно.

3. Выбираем $c_{k+1} \geq c_k$, увеличиваем номер шага k на 1 и переходим к шагу 1.

Отметим, что в практических реализациях МФЛ обычно используются более сложные варианты данного алгоритма. Некоторые из них будут рассмотрены последующих главах данной работы.

Метод модифицированных функций обладает рядом привлекательных свойств [37], одно из которых заключается в сильной теории глобальной сходимости: стационарность предельных точек траекторий метода обосновываются при весьма слабых требованиях регулярности ограничений [31, 33] и, в частности, при возможной неограниченности множества $\mathcal{M}(\bar{x})$, что делает МФЛ перспективной глобальной стратегией для задач оптимизации с нерегулярными ограничениями.

Другой привлекательной особенностью МФЛ является локальная линейная и даже сверхлинейная сходимость без каких-либо условий регулярности. А именно, справедлива следующая теорема, более общий вариант которой доказан в [48].

Теорема 4. Пусть функция $f: \mathbb{R}^n \rightarrow \mathbb{R}$ и отображения $h: \mathbb{R}^n \rightarrow \mathbb{R}^l$ и $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$ дважды дифференцируемы в окрестности точки $\bar{x} \in \mathbb{R}^n$, и их вторые производные непрерывны в этой точке. Пусть \bar{x} — стационарная точка задачи (1), удовлетворяющая достаточному условию второго порядка (7) при некотором $(\bar{\lambda}, \bar{\mu}) \in \mathcal{M}(\bar{x})$.

Тогда для всякого $M > 0$ найдется $\bar{c} > 0$ такое, что для любого $c_0 \geq \bar{c}$ и любого начального приближения $(x^0, \lambda^0, \mu^0) \in \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}_+^m$, достаточно близкого к $(\bar{x}, \bar{\lambda}, \bar{\mu})$, существует траектория $\{(x^k, \lambda^k, \mu^k)\}$ алгоритма 1, удовлетворяющая для каждого $k = 0, 1, \dots$ условию

$$\|(x^{k+1} - x^k, \lambda^{k+1} - \lambda^k, \mu^{k+1} - \mu^k)\| \leq M\sigma(x^k, \lambda^k, \mu^k).$$

Всякая такая траектория сходится к $(\bar{x}, \lambda^*, \mu^*)$ при некотором $(\lambda^*, \mu^*) \in \mathcal{M}(\bar{x})$, причем скорость сходимости линейная, и сверхлинейная, если $c_k \rightarrow +\infty$.

Как было недавно показано в [65], при отсутствии ограничений-неравенств и некоторых дополнительных ограничениях на выбор параметра штрафа для локальной сверхлинейной сходимости достаточно не критичности множителя $\bar{\lambda}$.

Результат о локальной сходимости для МФЛ очень похож на соответствующий результат для sSQP. Более того, как было показано в работе [9], МФЛ также подвержен эффекту притяжения к критическим множителям: для него тоже существуют прямо-двойственные области, при запуске из которых метод сходится к некоторому критическому множителю, причем во многих случаях с линейной скоростью (см. примеры 2–7 из [9]).

Отметим, что для МФЛ эффект притяжения к критическим множителям проявляется несколько по-другому, чем для sSQP. Если в некоторый момент двойственное приближение оказывается достаточно близким к области множества множителей Лагранжа, во всех точках которой выполнено SOSC (7) (а значит и не критичность), то типичным сценарием для МФЛ является сходимость к точке этой области. Если же нужная близость не достигается, то метод обычно сходится к точке относительной границы области выполнения SOSC, а всякая

такая точка по необходимости является критическим множителем. Например, для примера 3 область сходимости к критическим множителям будет иметь вид $\{(x, \lambda) \in \mathbb{R}^n \times \mathbb{R}^l \mid \lambda \leq -1\}$ (т. е. существенно отличается от аналогичной области для sSQP, показанной на рис. 9).

Наличие эффекта притяжения к критическим множителям также подтверждается результатами вычислительного эксперимента, которые приведены в приложении А.2: сходимость к критическим множителям наблюдалась для большей части задач. Вместе с тем, в отличие от SQP и его стабилизированной версии, в большинстве таких случаев имела место сверхлинейная сходимость, если параметр штрафа стремился к бесконечности.

Глава 2

Метод модифицированных функций Лагранжа

Данная глава посвящена исследованию поведения метода модифицированных функций Лагранжа (МФЛ) для задач с нерегулярными ограничениями.

В разделе 2.1 для задач оптимизации общего вида и для задач оптимизации с комплементарными ограничениями доказываются результаты о глобальной сходимости МФЛ в весьма слабых предположениях. В частности, не требуется выполнение традиционных условий регулярности ограничений, что делает МФЛ перспективной глобальной стратегией для нерегулярных задач.

Раздел 2.2 посвящен анализу возможности ускорения финальной фазы МФЛ с помощью ньютоновских методов. Обосновывается локальная сверхлинейная сходимость ускорителя на основе стабилизированного метода Ньютона–Лагранжа. Однако, несмотря на неплохие теоретические свойства, высокая стоимость итераций и притяжение к критическим множителям приводят к тому, что МФЛ уступает по эффективности другим существующим методам.

2.1. Результаты о глобальной сходимости

В данном разделе исследуется глобальная сходимость МФЛ. Сначала рассматриваются задачи общего вида, для которых доказывается глобальная сходимость метода без требования регулярности ограничений (требуется более слабая оценка расстояния и некоторые ограничения на решения подзадач). Далее отдельно рассматриваются задачи оптимизации с комплементарными ограничениями, поскольку для них неприменима стандартная теория глобальной сходимости. Для таких задач доказывается так называемая S -стационарность предельных точек метода в предположении так называемого МРСС-условия линейной независимости.

Теоретические результаты, полученные в данном разделе, а также результаты вычислительного эксперимента, приведенные в приложении А.1, свидетельствуют о том, что МФЛ обладает очень хорошими свойствами глобальной сходимости. В частности, этот метод может быть хорошим выбором для потенциально вырожденных задач в случаях, когда главную роль играют робастность и качество получаемых решений.

2.1.1. Общая теория глобальной сходимости

Будем рассматривать задачу оптимизации со смешанными ограничениями

$$f(x) \rightarrow \min, \quad h(x) = 0, \quad g(x) \leq 0, \quad (1)$$

где $f: \mathbb{R}^n \rightarrow \mathbb{R}$ — дважды дифференцируемая функция, а $h: \mathbb{R}^n \rightarrow \mathbb{R}^l$ и $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$ — дважды дифференцируемые отображения.

Напомним, что стационарные точки и отвечающие им множители Лагранжа задачи (1) характеризуются системой Каруша–Куна–Таккера (ККТ)

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

где $L: \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m \rightarrow \mathbb{R}$ — функция Лагранжа задачи (1):

$$L(x, \lambda, \mu) = f(x) + \langle \lambda, h(x) \rangle + \langle \mu, g(x) \rangle. \quad (3)$$

Определим семейство модифицированных функций Лагранжа задачи (1): $L_c: \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m \rightarrow \mathbb{R}$,

$$L_c(x, \lambda, \mu) = f(x) + \frac{1}{2c} (\|\lambda + ch(x)\|^2 + \|\max\{0, \mu + cg(x)\}\|^2), \quad (4)$$

где $c > 0$ — параметр штрафа, а максимум берется покомпонентно. Отметим, что если среди ограничений задачи (1) присутствуют простые ограничения (или в более общем случае, линейные ограничения), то на практике они часто не заносятся в модифицированную функцию Лагранжа, а трактуются напрямую, т. е. на каждой итерации алгоритма текущая модифицированная функция минимизируется при этих ограничениях (см., например, [31, 41]). Для простоты изложения в данном разделе не будут учитываться эти обобщения. По этой же причине не будут рассматриваться важные с практической точки зрения модификации алгоритма, в которой подзадачи решаются приближенно, а также модификации, в которой различным ограничениям соответствуют различные параметры штрафа. Легко убедиться, что эти модификации практически не сказываются на проводимом далее анализе, поскольку они никак не связаны с нерегулярностью ограничений.

Сформулируем алгоритм, который будет использоваться в дальнейшем анализе. С учетом упомянутых выше упрощений, этот алгоритм соответствует тому, который лежит в основе солвера ALGENCAN [30] (см. [31, алгоритм 3.1], а также некоторые совсем недавние модификации и улучшения в [33, 36]).

Алгоритм 1. Выбираем скалярные параметры $\bar{\lambda}_{\min}, \bar{\lambda}_{\max}, \bar{\lambda}_{\min} \leq \bar{\lambda}_{\max}, \bar{\mu}_{\max} \geq 0, c_0 > 0, \theta \in [0, 1)$ и $\rho > 1$. Полагаем $k = 0$.

1. Выбираем вектор $(\bar{\lambda}^k, \bar{\mu}^k) \in \mathbb{R}^l \times \mathbb{R}^m$, удовлетворяющий $\bar{\lambda}_{\min} \leq \bar{\lambda}_j^k \leq \bar{\lambda}_{\max}, j = 1, \dots, l$, и $0 \leq \bar{\mu}_j^k \leq \bar{\mu}_{\max}, j = 1, \dots, m$. (Для $k \geq 1$ естественным является выбор $(\bar{\lambda}^k, \bar{\mu}^k)$ как проекции вектора (λ^k, μ^k) , получаемого в п. 2 алгоритма, на множество $\bigotimes_{j=1}^l [\bar{\lambda}_{\min}, \bar{\lambda}_{\max}] \times \bigotimes_{j=1}^m [0, \bar{\mu}_{\max}]$.)

Вычисляем $x^{k+1} \in \mathbb{R}^n$ как стационарную точку задачи

$$L_{c_k}(x, \bar{\lambda}^k, \bar{\mu}^k) \rightarrow \min, \quad x \in \mathbb{R}^n. \quad (5)$$

2. Полагаем

$$\lambda^{k+1} = \bar{\lambda}^k + c_k h(x^{k+1}), \quad \mu^{k+1} = \max\{0, \bar{\mu}^k + c_k g(x^{k+1})\}, \quad (6)$$

$$\tau^{k+1} = \min\{\mu^{k+1}, -g(x^{k+1})\}. \quad (7)$$

3. Если $k = 0$ или выполнено условие

$$\max\{\|h(x^{k+1})\|_{\infty}, \|\tau^{k+1}\|_{\infty}\} \leq \theta \max\{\|h(x^k)\|_{\infty}, \|\tau^k\|_{\infty}\},$$

выбираем любое значение $c_{k+1} \geq c_k$. В противном случае выбираем $c_{k+1} \geq \rho c_k$. Увеличиваем номер шага k на 1 и переходим к п. 1.

В ALGENCAN используются следующие значения параметров: $\bar{\lambda}_{\min} = -10^{20}$, $\bar{\lambda}_{\max} = \bar{\mu}_{\max} = 10^{20}$, $\theta = 0.5$ и $\rho = 10$. Начальное значение параметра штрафа c_0 определяется с помощью специальной процедуры инициализации и зависит от значений f, h и g в имеющейся начальной точке (см. [38]). Последующие значения параметра штрафа на шаге 3 алгоритма выбираются следующим образом: $c_{k+1} = c_k$, если условие (7) выполнено, и $c_{k+1} = \rho c_k$ в противном случае.

Отметим, что помимо процедуры, описанной в алгоритме 1, ALGENCAN использует ряд эвристик, направленных на повышение его эффективности (см. [38]). В частности, используются «рестарты» для параметра штрафа: если текущее приближение оказывается «почти допустимым», но при этом невязка остальных условий в системе ККТ (2) не является достаточно малой, то очередное значение параметра штрафа снова вычисляется с помощью процедуры инициализации. Кроме того, в ALGENCAN присутствует процедура масштабирования данных задачи: алгоритм 1 применяется не к исходной задаче (1), а к ее масштабированному варианту

$$\tilde{f}(x) \rightarrow \min, \quad \tilde{h}(x) = 0, \quad \tilde{g}(x) \leq 0, \quad (8)$$

где $\tilde{f}(x) = f(x)/\max\{1, \|f'(x^0)\|_\infty\}$, $\tilde{h}_i(x) = h_i(x)/\max\{1, \|h'_i(x^0)\|_\infty\}$, $i \in 1, \dots, l$, и $\tilde{g}_i(x) = g_i(x)/\max\{1, \|g'_i(x^0)\|_\infty\}$, $i \in 1, \dots, m$, а $x^0 \in \mathbb{R}^n$ — начальное приближение. Разумеется, переход к задаче (8) приводит к изменению как прямой, так и двойственной траектории метода. Вместе с тем, очевидно, что точка \bar{x} является стационарной в исходной задаче (1) тогда и только тогда, когда она является стационарной в масштабированной задаче (8), хотя соответствующие множители Лагранжа могут меняться.

При анализе глобальной сходимости алгоритма 1 необходимо отдельно рассматривать случаи допустимых и недопустимых предельных точек. Возможность сходимости к недопустимым точкам является одним из слабых мест МФЛ. С другой стороны, подобные случаи никак не связаны с нерегулярностью ограничений в решениях или в допустимых точках. В связи с этим, в данном разделе основное внимание уделяется допустимым предельным точкам. Возможность существования недопустимых предельных точек кратко обсуждается в конце раздела.

Приводимое ниже рассуждение в значительной степени основано на результатах из [31, теоремы 4.1, 4.2], а также на некоторых недавних улучшениях, полученных в [33].

Из (3), (4) и (6) легко следует, что траектория, генерируемая алгоритмом 1, удовлетворяет для всех $k = 1, 2, \dots$ условиям

$$\frac{\partial L}{\partial x}(x^k, \lambda^k, \mu^k) = \frac{\partial L_{c_{k-1}}}{\partial x}(x^k, \bar{\lambda}^{k-1}, \bar{\mu}^{k-1}) = 0, \quad \mu^k \geq 0. \quad (9)$$

Предположим, что последовательность $\{x^k\}$ имеет допустимую предельную точку. Тогда из (6) и (7), с учетом ограниченности $\{\bar{\mu}^k\}$ и правила обновления параметра штрафа на шаге 3 алгоритма 1, получаем, что если для некоторого множества индексов $K \subset \{0, 1, \dots\}$ подпоследовательность $\{x^k \mid k \in K\}$ сходится к допустимой точке \bar{x} , то

$$\{\mu_{\{1, \dots, m\} \setminus A(\bar{x})}^k \mid k \in K\} \rightarrow 0 \quad \text{при} \quad k \rightarrow \infty, \quad (10)$$

где через $A(\bar{x}) = \{i = 1, \dots, m \mid g_i(\bar{x}) = 0\}$ обозначено множество индексов активных ограничений-неравенств.

Таким образом, если последовательность $\{(\lambda^k, \mu^k) \mid k \in K\}$ имеет предельную точку $(\bar{\lambda}, \bar{\mu})$ (в частности, если эта последовательность ограничена), то переходя к пределу в (9) вдоль соответствующей подпоследовательности и применяя (10), сразу же получаем, что \bar{x} является предельной точкой задачи (1), а $(\bar{\lambda}, \bar{\mu})$ — отвечающим ей множителем Лагранжа.

Вновь заметим, что последовательность $\{(\bar{\lambda}^k, \bar{\mu}^k)\}$ ограничена, и поэтому, согласно (6), последовательность $\{(\lambda^k, \mu^k)\}$ также будет ограничена в случае, когда ограничена $\{c_k\}$. Таким образом, основные трудности при анализе глобальной сходимости связаны именно со случаем, когда $c_k \rightarrow +\infty$ и последовательность $\{(\lambda^k, \mu^k)\}$ не ограничена. Для того, чтобы обойти эти трудности, можно потребовать выполнения условий регулярности.

Легко проверить, что условие MFSCQ (1.4) допускает следующую двойственную формулировку: оно выполняется в допустимой точке \bar{x} задачи (1), если система

$$(h'(\bar{x}))^T \eta + (g'_{A(\bar{x})}(\bar{x}))^T \tilde{\zeta} = 0, \quad \tilde{\zeta} \geq 0,$$

относительно переменных $(\eta, \tilde{\zeta}) \in \mathbb{R}^l \times \mathbb{R}^{|A(\bar{x})|}$ имеет только нулевое решение. В контексте алгоритма 1 выполнение MFSCQ исключает случаи неограниченности $\{(\lambda^k, \mu^k) \mid k \in K\}$. Действительно, если разделить условия из (9) на $\|(\lambda^k, \mu^k)\|$, то, переходя к пределу и учитывая (10), получаем противоречие с MFSCQ.

Существенно более сложные рассуждения в [33] показывают, что анализ глобальной сходимости алгоритма 1 может быть проведен при следующем (более слабом) условии регулярности.

Для заданной допустимой точки \bar{x} задачи (1) обозначим через $J \subset \{1, \dots, l\}$ такое множество индексов, что $\{h'_j(\bar{x}) \mid j \in J\}$ образует базис подпространства, натянутого на векторы $\{h'_j(\bar{x}) \mid j = 1, \dots, l\}$. Тогда в точке \bar{x} задачи (1) выполняется условие RCPLD (от англ. Relaxed constant positive linear dependence) [33], если существует окрестность U точки \bar{x} , для которой:

- 1) матрица $h'(x)$ имеет одинаковый ранг для всех $x \in U$;
- 2) для любого $I \subset A(\bar{x})$, если существуют $\tilde{\eta} \in \mathbb{R}^{|J|}$ и $\tilde{\zeta} \in \mathbb{R}^{|I|}$, не равные нулю одновременно и удовлетворяющие

$$(h'_J(\bar{x}))^T \tilde{\eta} + (g'_I(\bar{x}))^T \tilde{\zeta} = 0, \quad \tilde{\zeta} \geq 0, \quad (11)$$

то

$$\text{rank} \begin{pmatrix} h'_J(x) \\ g'_I(x) \end{pmatrix} < |J| + |I| \quad \forall x \in U. \quad (12)$$

Данное условие действительно является ослабленной версией условия CPLD (от англ. Constant positive linear dependence constraint qualification) в допустимой точке \bar{x} задачи (1), которое состоит в том, что существует такая окрестность U точки \bar{x} , что для любых $J \subset \{1, \dots, l\}$ и $I \subset A(\bar{x})$, если существуют $\tilde{\eta} \in \mathbb{R}^{|J|}$ и $\tilde{\zeta} \in \mathbb{R}^{|I|}$, не равные нулю одновременно и удовлетворяющие (11), то выполнено (12). Легко видеть, что из MFSCQ следует CPLD, а значит и RCPLD. Отметим, что CPLD также следует из другого известного условия регулярности — условия постоянного ранга (CRCQ, от англ. Constant rank constraint qualification).

Следующее утверждение устанавливает свойства глобальной сходимости алгоритма 1 в случае допустимых предельных точек [31, 33].

Теорема 1. Пусть $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $h: \mathbb{R}^n \rightarrow \mathbb{R}^l$ и $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$ непрерывно дифференцируемы на \mathbb{R}^n . Пусть последовательность $\{x^k\} \subset \mathbb{R}^n$ сгенерирована алгоритмом 1, а $\bar{x} \in \mathbb{R}^n$ — ее предельная точка.

Тогда если точка \bar{x} является допустимой в задаче (1), и в ней выполнено условие RCPLD, то она является стационарной точкой задачи (1).

Более того, если в точке \bar{x} выполнено MFCQ, то для любого бесконечного множества индексов $K \subset \{0, 1, \dots\}$, для которого подпоследовательность $\{x^k \mid k \in K\}$ сходится к \bar{x} , соответствующая подпоследовательность $\{(\lambda^k, \mu^k) \mid k \in K\}$ является ограниченной, и любая ее предельная точка является множителем Лагранжа, отвечающим \bar{x} .

Если условие RCPLD нарушается, то утверждение теоремы не имеет места, как показывает следующий простой пример.

Пример 1. Пусть $n = l = 1$, $m = 0$, $f(x) = x$, $h(x) = x^2$. В этом случае задача (1) имеет единственную допустимую точку (а значит и единственное решение) $\bar{x} = 0$, которая не является стационарной, и в которой нарушается RCPLD. Имеем:

$$L_c(x, \lambda) = x + \frac{1}{2c}\lambda^2 + \lambda x^2 + \frac{c}{2}x^4, \quad \frac{\partial L_c}{\partial x}(x, \lambda) = 1 + 2\lambda x + 2cx^3.$$

Полагая, например, $\bar{\lambda}^k = 0$, для любого $c_k > 0$ получаем, что точка $x^{k+1} = -1/(2c_k)^{1/3}$ является единственной стационарной точкой и глобальным решением подзадачи (5). Если при этом $c_k \rightarrow +\infty$, то последовательность $\{x^k\}$ сходится к \bar{x} .

Приведенные выше результаты о глобальной сходимости применимы и при наличии вырожденных решений: такие решения вполне могут быть предельными точками траекторий, генерируемых методом множителей, и это никак не противоречит теории. Проблемы возникают в случаях, когда существуют допустимые точки, которые не являются стационарными, но в которых нарушаются даже такие слабые условия регулярности как RCPLD. Ситуация становится еще хуже, когда все допустимые точки задачи могут быть нерегулярными. В таких случаях теорема 1 разрешает сходимость к любым допустимым точкам, что, конечно, не является удовлетворительным результатом. Например, в случае МРСС, условие MFCQ нарушается во всех допустимых точках. Что касается RCPLD, то данное условие может выполняться для МРСС, однако в разделе 2.1.2 показано, что его выполнение крайне нетипично. В связи с этим, для МРСС будет представлен отдельный анализ, использующий специальную структуру этих задач.

Чтобы дополнить картину, докажем новый результат о глобальной сходимости, не использующий условия регулярности ограничений. Вместо этого будет сделано два предположения, которые не являются стандартными, но при этом весьма обоснованы и должны выполняться во многих случаях. Первое состоит в липшицевой оценке расстояния (см. условие

(14) ниже) до допустимого множества задачи (1) в терминах невязки ограничений. Заметим, что такая оценка слабее, чем условия регулярности в точке \bar{x} : согласно [33], оценка расстояния следует из RCPLD (а значит и из MFCQ), но в то же время, согласно [33, контрпример 3], RCPLD из нее не следует. Второе предположение состоит в том, что на каждой итерации алгоритма 1 следующее приближение ищется не просто как стационарная точка подзадачи (5), но как точка, обладающая некоторыми свойствами (приближенного) минимума, в смысле соотношения (13) ниже. Данное предположение 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} . Последнее, в свою очередь, выполняется автоматически при некоторых дополнительных условиях, обеспечивающих локальную выпуклость модифицированной функции в некоторой фиксированной (не зависящей от k) окрестности предельной точки.

Теорема 2. Пусть D — допустимое множество задачи (1). В предположениях теоремы 1 пусть существует такое бесконечное множество индексов $K \subset \{1, 2, \dots\}$, что подпоследовательность $\{x^k \mid k \in K\}$ сходится к $\bar{x} \in D$, и 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

$$L_{c_{k-1}}(x^k, \bar{\lambda}^{k-1}, \bar{\mu}^{k-1}) \leq L_{c_{k-1}}(\bar{x}^k, \bar{\lambda}^{k-1}, \bar{\mu}^{k-1}) + \beta(\|h(x^k)\| + \|\max\{0, g(x^k)\}\|) + \frac{\theta\|(\lambda^k, \mu^k)\|^2}{c_{k-1}}. \quad (13)$$

Пусть также справедлива оценка расстояния

$$\text{dist}(x^k, D) = O(\|h(x^k)\| + \|\max\{0, g(x^k)\}\|) \quad (14)$$

для $k \in K$ при $k \rightarrow \infty$.

Тогда 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} .

(Отметим, что допустимое множество D не обязательно выпукло, и поэтому проекция точки x^k на D может быть не единственной. В теореме 2 подходит любая проекция \bar{x}^k .)

Доказательство. Применяя (4) и (6), получаем

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

С другой стороны, точка \bar{x}^k является допустимой в задаче (1), и поэтому $h(\bar{x}^k) = 0$ и $\max\{0, \bar{\mu}^{k-1} + c_{k-1}g(\bar{x}^k)\} \leq \bar{\mu}^{k-1}$, откуда следует

$$\begin{aligned} L_{c_{k-1}}(\bar{x}^k, \bar{\lambda}^{k-1}, \bar{\mu}^{k-1}) &= f(\bar{x}^k) + \frac{1}{2c_{k-1}}(\|\bar{\lambda}^{k-1} + c_{k-1}h(\bar{x}^k)\|^2 + \\ &+ \|\max\{0, \bar{\mu}^{k-1} + c_{k-1}g(\bar{x}^k)\}\|^2) \leq f(\bar{x}^k) + \frac{\|(\bar{\lambda}^{k-1}, \bar{\mu}^{k-1})\|^2}{2c_{k-1}}. \end{aligned}$$

Комбинируя последние два соотношения с (13), находим

$$\begin{aligned} \frac{\|(\lambda^k, \mu^k)\|^2}{2c_{k-1}} &\leq f(\bar{x}^k) - f(x^k) + \frac{\|(\bar{\lambda}^{k-1}, \bar{\mu}^{k-1})\|^2}{2c_{k-1}} + \beta(\|h(x^k)\| + \|\max\{0, g(x^k)\}\|) + \\ &+ \frac{\theta\|(\lambda^k, \mu^k)\|^2}{c_{k-1}}. \end{aligned} \quad (15)$$

Заметим, что функция f локально липшицева с некоторой константой $\ell > 0$. Используя правило пересчета множителей (6), оценку расстояния (14), неравенство (15), а также ограниченность $\{(\bar{\lambda}^k, \bar{\mu}^k)\}$ и неравенство $\bar{\mu}^k \geq 0$ для всех k , легко получить существование константы $M > 0$, для которой выполнена следующая цепочка неравенств:

$$\begin{aligned} (1 - 2\theta)\|(\lambda^k, \mu^k)\|^2 &\leq 2\ell c_{k-1}\|\bar{x}^k - x^k\| + \|(\bar{\lambda}^{k-1}, \bar{\mu}^{k-1})\|^2 + \\ &+ 2\beta c_{k-1}(\|h(x^k)\| + \|\max\{0, g(x^k)\}\|) \leq M(\|c_{k-1}h(x^k)\| + \|\max\{0, c_{k-1}g(x^k)\}\| + 1) \leq \\ &\leq M(\|\lambda^k - \bar{\lambda}^{k-1}\| + \|\mu^k\| + 1) \leq M(\|\lambda^k\| + \|\mu^k\| + \|\bar{\lambda}^{k-1}\| + 1). \end{aligned}$$

Очевидно, из полученного соотношения следует, что последовательность $\{(\lambda^k, \mu^k)\}$ не может быть неограниченной. Переходя к пределу в (9) вдоль любой сходящейся подпоследовательности и принимая во внимание (10), отсюда немедленно получаем утверждение теоремы. ■

В частности, если выполнено (13), то любое условие, из которого следует липшицева оценка расстояния (14) (например, RCPLD [33]), обеспечивает ограниченность подпоследовательности $\{(\lambda^k, \mu^k) \mid k \in K\}$ и стационарность предельной точки \bar{x} .

Наконец, как хорошо известно, в общем случае невозможно гарантировать отсутствие сходимости МФЛ к недопустимым точкам. В частности, справедлива следующая теорема, которая представляет собой результат из [31, теорема 4.1] для задачи (1).

Теорема 3. *В предположениях теоремы 1, если последовательность параметров штрафа $\{c_k\}$ ограничена, то точка \bar{x} является допустимой в задаче (1); в противном случае она является стационарной точкой задачи*

$$\|h(x)\|^2 + \|\max\{0, g(x)\}\|^2 \rightarrow \min, \quad x \in \mathbb{R}^n. \quad (16)$$

Стационарность точки \bar{x} в задаче (16) означает, что выполнено соотношение

$$(h'(\bar{x}))^T h(\bar{x}) + (g'(\bar{x}))^T \max\{0, g(\bar{x})\} = 0. \quad (17)$$

Очевидно, из (17) не следует допустимость \bar{x} в задаче (1). Заметим, однако, что возможная недопустимость \bar{x} связана с редкой ситуацией. А именно, определим множества индексов $J(\bar{x}) = \{j = 1, \dots, l \mid h_j(\bar{x}) \neq 0\}$ и $I(\bar{x}) = \{i = 1, \dots, m \mid g_i(\bar{x}) > 0\}$ и введем следующее «условие регулярности»:

$$\begin{aligned} &\text{либо } I(\bar{x}) = \emptyset \text{ и } \text{rank } h'_{J(\bar{x})}(\bar{x}) = |J(\bar{x})|, \\ &\text{либо } I(\bar{x}) \neq \emptyset \text{ и } \exists \xi \in \ker h'_{J(\bar{x})}(\bar{x}): g'_{I(\bar{x})}(\bar{x})\xi < 0. \end{aligned} \quad (18)$$

Если в недопустимой точке \bar{x} выполняется это условие, то \bar{x} не может удовлетворять (17), и поэтому она не может быть предельной точкой алгоритма 1. В частности, при отсутствии ограничений-неравенств выполнение (17) в недопустимой точке \bar{x} возможно только при $\text{rank } h'(\bar{x}) < l$. Существуют и другие дополнительные условия, обеспечивающие допустимость предельных точек, но использующие информацию о структуре ограничений (см., например, [34]).

Следующий пример показывает, что сходимость траекторий алгоритма 1 к недопустимым точкам, удовлетворяющим (17), является возможным сценарием.

Пример 2. Пусть $n = l = 1$, $m = 0$, $f \equiv 0$, $h(x) = x^2 + 1$. Задача (1) с такими данными не имеет допустимых точек. Тем не менее,

$$\begin{aligned} L_c(x, \lambda) &= \frac{1}{2c}\lambda^2 + \lambda(x^2 + 1) + \frac{c}{2}(x^2 + 1)^2, \\ \frac{\partial L_c}{\partial x}(x, \lambda) &= 2\lambda x + 2c(x^2 + 1)x = 2x(\lambda + c(x^2 + 1)), \end{aligned}$$

и поэтому для любых $\bar{\lambda}^k$ и c_k , для которых $\bar{\lambda}^k + c_k > 0$, точка $\bar{x} = 0$ является единственной стационарной точкой и глобальным решением подзадачи (5). Таким образом, недопустимая точка \bar{x} может быть предельной точкой траектории алгоритма 1. Причина состоит в том, что $h'(\bar{x}) = 0$, и, в частности, условие (18) не выполняется.

Отметим, однако, что подобное поведение МФЛ совершенно нетипично. В частности, результаты вычислительного эксперимента, приведенные в приложении А.1, показывают, что сходимость к недопустимым точкам для ALGENCAN происходит крайне редко (хотя и не может быть исключена теоретически). Напротив, в плане свойств глобальной сходимости ALGENCAN превосходит многие существующие альтернативы, в том числе и для вырожденных задач.

Замечание 1. Тест на шаге 3 алгоритма 1, определяющий, нужно ли увеличивать параметр штрафа на данной итерации, представляет собой тест на линейное убывание естественной невязки системы ККТ (2) для задачи (1). Данный тест был введен в [36]. В предыдущих версиях ALGENCAN (см., например, [31]) для вычисления τ^{k+1} вместо (7) использовалось следующее выражение:

$$\tau^{k+1} = \max \left\{ g(x^{k+1}), -\frac{\bar{\mu}^k}{c_k} \right\}.$$

Несложно показать, что все утверждения и комментарии, приведенные выше, а также в следующем разделе, посвященном МРСС, остаются справедливыми и для такой версии алгоритма. Более того, анализ даже несколько упрощается, поскольку в этом случае свойство (10) можно заменить на более сильное:

$$\mu_{\{1, \dots, m\} \setminus A(\bar{x})}^k = 0$$

для всех достаточно больших $k \in K$.

2.1.2. Глобальная сходимость для задач оптимизации с комплементарными ограничениями

В данном разделе будем рассматривать МРСС следующего вида:

$$f(x) \rightarrow \min, \quad G(x) \geq 0, \quad H(x) \geq 0, \quad \langle G(x), H(x) \rangle \leq 0. \quad (19)$$

Конечно, «обычные» ограничения-равенства и ограничения-неравенства также могут присутствовать в МРСС. Однако все сложности проводимого ниже анализа связаны именно с комплементарными ограничениями, и поэтому для простоты изложения будем считать, что других ограничений нет. Легко показать, что все полученные результаты распространяются и на общий случай.

Сначала покажем, что в контексте МРСС теоремы 1 оказывается недостаточно, и необходим дальнейший анализ. Напомним, что для МРСС условие МФСQ нарушается в каждой допустимой точке. Что касается RCPLD, то можно построить примеры, в которых выполнено даже более сильное условие CPLD. С другой стороны, приведенное ниже рассуждение показывает, что подобные примеры являются в значительной степени искусственными, и выполнение RCPLD является крайне нетипичным.

Покажем сначала, что CPLD может выполняться для МРСС. Рассмотрим задачу (19), в которой отображения $G, H: \mathbb{R}^n \rightarrow \mathbb{R}^s$ тождественно равны нулю. В этом случае каждая точка $\bar{x} \in \mathbb{R}^n$ является допустимой и удовлетворяет CPLD: для всех $x \in \mathbb{R}^n$ матрица Якоби ограничений тождественно равна нулю, и поэтому любое подмножество ее строк является линейно зависимым.

Теперь покажем, что выполнение (R)CPLD на практике является крайне нетипичным. Заметим, что в каждой допустимой точке выполнено s равенств

$$G_i(x)H_i(x) = 0, \quad i = 1, \dots, s, \quad (20)$$

и поэтому на практике s обычно не превосходит $n - 1$. В противном случае не остается «степеней свободы» для оптимизации, и МРСС по сути становится задачей допустимости. Если $s \geq n$, то (R)CPLD вполне может выполняться, но в контексте МРСС типичным является именно случай $s \leq n - 1$. Рассматривая ограничения МРСС (19) как частный случай ограничений-неравенств (1), заметим, что матрица Якоби для всех $x \in \mathbb{R}^n$ имеет вид

$$\begin{pmatrix} -G'(x) \\ -H'(x) \\ (G'(x))^T H(x) + (H'(x))^T G(x) \end{pmatrix}. \quad (21)$$

Согласно (20), для произвольной допустимой точки $\bar{x} \in \mathbb{R}^n$ задачи (19) без ограничения общности можно полагать, что $H(\bar{x}) = 0$ (этого всегда можно добиться путем перемещения положительных элементов из H в G , а соответствующих нулевых элементов из G в H). В этом случае последняя строка матрицы Якоби (21) в точке \bar{x} имеет вид $\sum_{i=1}^s G_i(\bar{x})H'_i(\bar{x})$ с неотрицательными коэффициентами $G_i(\bar{x})$, $i = 1, \dots, s$. Следовательно, строки $-H'_i(\bar{x})$, $i = 1, \dots, s$, матрицы (21) в совокупности с последней строкой являются положительно линейно зависимыми. Поэтому если (R)CPLD выполнено в точке \bar{x} , то векторы $-H'_i(x)$, $i = 1, \dots, s$, и $\sum_{i=1}^s (G_i(x)H'_i(x) + H_i(x)G'_i(x))$ должны быть линейно зависимы для всех $x \in \mathbb{R}^n$, достаточно близких к \bar{x} . Если $s \leq n - 1$, такая система содержит не более чем n строк. Таким образом, не более чем n векторов в пространстве \mathbb{R}^n остаются линейно зависимыми при произвольных малых возмущениях точки \bar{x} , что является крайне нетипичным свойством (в большинстве случаев малое возмущение линейно зависимой системы делает ее линейно независимой).

Данное рассуждение позволяет заключить, что (R)CPLD обычно не выполняется для МРСС, и поэтому теорема 1 неприменима к таким задачам. Таким образом, необходим дополнительный анализ. Сначала введем некоторую стандартную терминологию.

Определим обычную функцию Лагранжа $L: \mathbb{R}^n \times (\mathbb{R}^s \times \mathbb{R}^s \times \mathbb{R}) \rightarrow \mathbb{R}$ задачи (19):

$$L(x, \mu) = f(x) - \langle \mu_G, G(x) \rangle - \langle \mu_H, H(x) \rangle + \mu_0 \langle G(x), H(x) \rangle \quad (22)$$

(которая полностью согласуется с (3)), а также семейство модифицированных функций Лагранжа $L_c: \mathbb{R}^n \times (\mathbb{R}^s \times \mathbb{R}^s \times \mathbb{R}) \rightarrow \mathbb{R}$:

$$L_c(x, \mu) = f(x) + \frac{1}{2c} (\| \max\{0, \mu_G - cG(x)\} \|^2 + \| \max\{0, \mu_H - cH(x)\} \|^2 + (\max\{0, \mu_0 + c\langle G(x), H(x) \rangle\})^2)$$

(которое полностью согласуется с (4)), где $\mu = (\mu_G, \mu_H, \mu_0)$. Кроме того, определим так называемую *МРСС-функцию Лагранжа* $\mathcal{L}: \mathbb{R}^n \times (\mathbb{R}^s \times \mathbb{R}^s) \rightarrow \mathbb{R}$:

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

где $\lambda = (\lambda_G, \lambda_H)$.

Для произвольной точки $\bar{x} \in \mathbb{R}^n$, допустимой в задаче (19), определим множества индексов

$$\begin{aligned} I_G(\bar{x}) &= \{i = 1, \dots, s \mid G_i(\bar{x}) = 0\}, & I_H(\bar{x}) &= \{i = 1, \dots, s \mid H_i(\bar{x}) = 0\}, \\ I_0(\bar{x}) &= I_G(\bar{x}) \cap I_H(\bar{x}). \end{aligned}$$

Заметим, что обязательно выполнено $I_G(\bar{x}) \cup I_H(\bar{x}) = \{1, \dots, s\}$. *МРСС-условие линейной независимости* (МРСС-LICQ) состоит в том, что векторы

$$G'_i(\bar{x}), \quad i \in I_G(\bar{x}), \quad H'_i(\bar{x}), \quad i \in I_H(\bar{x})$$

линейно независимы. Допустимая точка \bar{x} задачи (19) называется *слабо стационарной*, если существует вектор $\bar{\lambda} = (\bar{\lambda}_G, \bar{\lambda}_H) \in \mathbb{R}^s \times \mathbb{R}^s$, удовлетворяющий

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

Если дополнительно

$$(\bar{\lambda}_G)_i (\bar{\lambda}_H)_i \geq 0 \quad \forall i \in I_0(\bar{x}),$$

то \bar{x} называется *C-стационарной* точкой. Если

$$\forall i \in I_0(\bar{x}) \text{ либо } (\bar{\lambda}_G)_i (\bar{\lambda}_H)_i = 0, \text{ либо } (\bar{\lambda}_G)_i > 0, (\bar{\lambda}_H)_i > 0,$$

то \bar{x} называется *M-стационарной* точкой. Если же

$$(\bar{\lambda}_G)_{I_0(\bar{x})} \geq 0, \quad (\bar{\lambda}_H)_{I_0(\bar{x})} \geq 0,$$

то \bar{x} называется *сильно стационарной* точкой. Очевидно, все эти понятия стационарности совпадают, если выполнено *условие строгой дополнителности нижнего уровня*, т.е. если $I_0(\bar{x}) = \emptyset$.

В случае МРСС (19) равенства из (6) записываются следующим образом:

$$\mu_G^{k+1} = \max\{0, \bar{\mu}_G^k - c_k G(x^{k+1})\}, \quad \mu_H^{k+1} = \max\{0, \bar{\mu}_H^k - c_k H(x^{k+1})\}, \quad (23)$$

$$\mu_0^{k+1} = \max\{0, \bar{\mu}_0^k + c_k \langle G(x^{k+1}), H(x^{k+1}) \rangle\}. \quad (24)$$

Основные соотношения (9) и (10) принимают следующий вид:

$$\frac{\partial L}{\partial x}(x^k, \mu^k) = 0, \quad \mu^k = (\mu_G^k, \mu_H^k, \mu_0^k) \geq 0 \quad (25)$$

для всех $k = 1, 2, \dots$, и если для некоторого $K \subset \{0, 1, \dots\}$ подпоследовательность $\{x^k \mid k \in K\}$ сходится к допустимой точке \bar{x} , то

$$\{(\mu_G^k)_{I_H(\bar{x}) \setminus I_G(\bar{x})} \mid k \in K\} \rightarrow 0, \quad \{(\mu_H^k)_{I_G(\bar{x}) \setminus I_H(\bar{x})} \mid k \in K\} \rightarrow 0 \quad (26)$$

для $k \in K$ при $k \rightarrow \infty$.

Сначала докажем следующую техническую лемму, которая потребуется в дальнейшем.

Лемма 1. Пусть функция $f: \mathbb{R}^n \rightarrow \mathbb{R}$ и отображение $h: \mathbb{R}^n \rightarrow \mathbb{R}^l$ дифференцируемы в некоторой окрестности $\bar{x} \in \mathbb{R}^n$, причем их производные непрерывны в точке \bar{x} . Предположим, что $\text{rank } h'(\bar{x}) = l$. Пусть задана последовательность $\{x^k\} \subset \mathbb{R}^n$, сходящаяся к \bar{x} , и для всех k выполнено равенство

$$\alpha_k f'(x^k) - (h'(x^k))^T \lambda^k = \omega^k \quad (27)$$

при некоторых $\alpha_k \in \mathbb{R}$, $\lambda^k \in \mathbb{R}^l$ и $\omega^k \in \mathbb{R}^n$.

Если $\alpha_k \rightarrow \bar{\alpha}$ и $\{\omega^k\} \rightarrow 0$ при $k \rightarrow \infty$, то существует единственный вектор $\bar{\lambda} \in \mathbb{R}^l$, для которого $\{\lambda^k\} \rightarrow \bar{\lambda}$ при $k \rightarrow \infty$ и

$$\bar{\alpha} f'(\bar{x}) - (h'(\bar{x}))^T \bar{\lambda} = 0. \quad (28)$$

В частности, если $\bar{\alpha} = 0$, то $\{\lambda^k\} \rightarrow 0$.

Доказательство. Предположим сначала, что последовательность $\{\lambda^k\}$ неограничена. Тогда существует бесконечное множество индексов $K \subset \{0, 1, \dots\}$, для которого $\{\|\lambda^k\| \mid k \in K\} \rightarrow +\infty$. Разделив (27) на $\|\lambda^k\|$ и перейдя к пределу вдоль любой сходящейся подпоследовательности ограниченной последовательности $\{\lambda^k / \|\lambda^k\| \mid k \in K\}$, получаем противоречие с условием $\text{rank } h'(\bar{x}) = l$.

Поэтому последовательность $\{\lambda^k\}$ ограничена. Переходя к пределу вдоль любой сходящейся подпоследовательности, получаем равенство (28) для некоторого $\bar{\lambda} \in \mathbb{R}^l$. Единственность $\bar{\lambda}$ следует из условия $\text{rank } h'(\bar{x}) = l$. ■

Следующая теорема устанавливает, что допустимые предельные точки алгоритма 1, применяемого к МРСС, являются по крайней мере S -стационарными при условии, что выполнено МРСС-LICQ. Более того, они являются сильно стационарными, если определенная двойственная последовательность ограничена. В совокупности с примерами ниже, которые показывают, что в общем случае M -стационарность (а значит и сильная стационарность) может не иметь места, данная теорема дает полную картину глобальной сходимости для МФЛ, применяемого к МРСС. Несмотря на то, что свойства глобальной сходимости не идеальны,

МФЛ не уступает никаким известным практическим алгоритмам для МРСС. Более того, в приложении А.1 показано, что в большинстве случаев предельные точки являются именно сильно стационарными.

Теорема 4. Пусть функция $f: \mathbb{R}^n \rightarrow \mathbb{R}$ и отображения $G, H: \mathbb{R}^n \rightarrow \mathbb{R}^s$ непрерывно дифференцируемы на \mathbb{R}^n . Пусть траектория $\{x^k\} \subset \mathbb{R}^n$ сгенерирована алгоритмом 1 для задачи (19), и пусть $\bar{x} \in \mathbb{R}^n$ является предельной точкой $\{x^k\}$. Предположим, что \bar{x} является допустимой в задаче (19), и в \bar{x} выполнено условие МРСС-LICQ.

Если последовательность $\{\mu_0^k\}$, генерируемая алгоритмом 1, имеет ограниченную подпоследовательность, причем соответствующая подпоследовательность $\{x^k\}$ сходится к \bar{x} (в частности, если последовательность $\{c_k\}$ ограничена), то \bar{x} является сильно стационарной точкой задачи (19). В противном случае \bar{x} является по крайней мере C -стационарной точкой задачи (19).

Доказательство. Без ограничения общности будем считать, что $G(\bar{x}) = 0$ (как уже упоминалось выше, этого всегда можно достичь путем перемещения положительных компонент из G в H , а соответствующих нулевых компонент из H в G). Определим множества индексов

$$I_0 = \{i = 1, \dots, s \mid H_i(\bar{x}) = 0\}, \quad I_+ = \{i = 1, \dots, s \mid H_i(\bar{x}) > 0\}. \quad (29)$$

Ясно, что из $G(\bar{x}) = 0$ следует $I_0 = I_0(\bar{x})$ и $I_+ = I_G(\bar{x}) \setminus I_H(\bar{x})$, где множества $I_0(\bar{x})$, $I_G(\bar{x})$ и $I_H(\bar{x})$ были введены выше. Кроме того, условие МРСС-LICQ в точке \bar{x} принимает вид

$$\text{rank} \begin{pmatrix} G'(\bar{x}) \\ H'_{I_0}(\bar{x}) \end{pmatrix} = s + |I_0|.$$

Из (22) и (25) получаем

$$\begin{aligned} 0 &= f'(x^k) - (G'(x^k))^T(\mu_G^k - \mu_0^k H(x^k)) - (H'(x^k))^T(\mu_H^k - \mu_0^k G(x^k)) = \\ &= f'(x^k) - (G'_{I_0}(x^k))^T(\lambda_G^k)_{I_0} - (G'_{I_+}(x^k))^T(\lambda_G^k)_{I_+} - \\ &\quad - (H'_{I_0}(x^k))^T(\lambda_H^k)_{I_0} - (H'_{I_+}(x^k))^T(\lambda_H^k)_{I_+}, \end{aligned} \quad (30)$$

где

$$\lambda_G^k = \mu_G^k - \mu_0^k H(x^k), \quad \lambda_H^k = \mu_H^k - \mu_0^k G(x^k). \quad (31)$$

Рассмотрим произвольное бесконечное множество индексов $K \subset \{0, 1, \dots\}$, для которого подпоследовательность $\{x^k \mid k \in K\}$ сходится к \bar{x} . Как упоминалось выше, если последовательность $\{c_k\}$ ограничена, то \bar{x} является стационарной точкой в обычном смысле

условий ККТ. Для МРСС это эквивалентно сильной стационарности \bar{x} . Поэтому ниже будем рассматривать случай $c_k \rightarrow +\infty$. В этом случае из (23) и (29) находим, что

$$(\mu_H^k)_{I_+} = 0 \quad (32)$$

для всех достаточно больших k .

Из второго соотношения в (31) и из (32) вытекает, что если для некоторого бесконечного множества индексов $K_0 \subset K$ подпоследовательность $\{\mu_0^k \mid k \in K_0\}$ ограничена, то последовательность $\{(\lambda_H^k)_{I_+} \mid k \in K_0\}$ сходится к нулю. Применяя лемму 1 к (30) при $\alpha_k = 1$, получаем, что последовательности $\{\lambda_G^k \mid k \in K_0\}$ и $\{(\lambda_H^k)_{I_0} \mid k \in K_0\}$ сходятся к некоторым векторам $\bar{\lambda}_G$ и $(\bar{\lambda}_H)_{I_0}$ соответственно, и при этом

$$f'(\bar{x}) - (G'(\bar{x}))^T \bar{\lambda}_G - (H'_{I_0}(\bar{x}))^T (\bar{\lambda}_H)_{I_0} = 0. \quad (33)$$

Переходя к пределу в (31) вдоль K_0 , получаем $(\bar{\lambda}_G)_{I_0} \geq 0$ и $(\bar{\lambda}_H)_{I_0} \geq 0$, и поэтому \bar{x} является сильно стационарной точкой. Заметим, что приведенное выше рассуждение также применимо в случаях $I_0 = \emptyset$ или $I_+ = \emptyset$.

Остается рассмотреть случай $\{\mu_0^k \mid k \in K\} \rightarrow +\infty$. Сначала докажем, что в этом случае множество I_0 непусто, и выполнено

$$\frac{c_{k-1}}{\mu_0^k} \langle G_{I_0}(x^k), H_{I_0}(x^k) \rangle \rightarrow M > 0, \quad (34)$$

если $k \in K$ стремится к бесконечности.

Поскольку $\{\mu_0^k \mid k \in K\} \rightarrow +\infty$, то из (24) следует

$$\mu_0^k = \bar{\mu}_0^{k-1} + c_{k-1} \langle G(x^k), H(x^k) \rangle.$$

Поэтому для всех достаточно больших $k \in K$ выполнено

$$\langle G(x^k), H(x^k) \rangle > 0, \quad (35)$$

и

$$\frac{c_{k-1}}{\mu_0^k} \langle G(x^k), H(x^k) \rangle = 1 - \frac{\bar{\mu}_0^{k-1}}{\mu_0^k} \rightarrow 1 \quad (36)$$

при $k \rightarrow \infty$. Очевидно также, что в этом случае

$$\frac{\mu_0^k}{c_{k-1}} = \langle G(x^k), H(x^k) \rangle + \frac{\bar{\mu}_0^{k-1}}{c_{k-1}},$$

и поэтому

$$\frac{\mu_0^k}{c_{k-1}} \rightarrow 0. \quad (37)$$

Ясно, что если $I_+ = \emptyset$ (и, следовательно, $I_0 \neq \emptyset$), то (34) сразу же следует из (36). Предположим, что $I_+ \neq \emptyset$. Из второго равенства в (31) и из (32) получаем, что для всех достаточно больших $k \in K$ выполнено

$$\frac{(\lambda_H^k)_{I_+}}{\mu_0^k} = -G_{I_+}(x^k) \quad (38)$$

Следовательно, $\{(\lambda_H^k)_{I_+}/\mu_0^k \mid k \in K\} \rightarrow 0$ при $k \rightarrow \infty$. Разделив (30) на μ_0^k и применяя лемму 1 при $\alpha_k = 1/\mu_0^k$, получаем соотношения

$$\left\{ \frac{\lambda_G^k}{\mu_0^k} \mid k \in K \right\} \rightarrow 0, \quad \left\{ \frac{(\lambda_H^k)_{I_0}}{\mu_0^k} \mid k \in K \right\} \rightarrow 0. \quad (39)$$

Далее, из (31) имеем

$$\frac{(\lambda_G^k)_{I_+}}{\mu_0^k} = \frac{(\mu_G^k)_{I_+}}{\mu_0^k} - H_{I_+}(x^k),$$

и поскольку $H_{I_+}(\bar{x}) > 0$, то из первого соотношения в (39) вытекает $\{(\mu_G^k)_i \mid k \in K\} \rightarrow +\infty$ для всех $i \in I_+$. Поэтому из (23) следует, что $G_{I_+}(x^k) < 0$ для всех достаточно больших $k \in K$. Из первого равенства в (31) для таких k имеем

$$\begin{aligned} (\lambda_G^k)_{I_+} &= \max\{0, (\bar{\mu}_G^{k-1})_{I_+} - c_{k-1}G_{I_+}(x^k)\} - \mu_0^k H_{I_+}(x^k) = \\ &= (\bar{\mu}_G^{k-1})_{I_+} - c_{k-1}G_{I_+}(x^k) - \mu_0^k H_{I_+}(x^k). \end{aligned} \quad (40)$$

Следовательно,

$$\frac{c_{k-1}}{\mu_0^k} G_{I_+}(x^k) + H_{I_+}(x^k) = \frac{(\bar{\mu}_G^{k-1})_{I_+}}{\mu_0^k} - \frac{(\lambda_G^k)_{I_+}}{\mu_0^k}.$$

Тогда из первого соотношения в (39) находим

$$\left\{ \frac{c_{k-1}}{\mu_0^k} G_{I_+}(x^k) + H_{I_+}(x^k) \mid k \in K \right\} \rightarrow 0,$$

откуда следует

$$\left\{ -\frac{c_{k-1}}{\mu_0^k} G_{I_+}(x^k) \mid k \in K \right\} \rightarrow H_{I_+}(\bar{x}), \quad (41)$$

а значит

$$\frac{c_{k-1}}{\mu_0^k} \langle G_{I_+}(x^k), H_{I_+}(x^k) \rangle \rightarrow -\|H_{I_+}(\bar{x})\|^2 \quad (42)$$

для $k \in K$ при $k \rightarrow \infty$.

Далее, поскольку множество I_+ непустое, то из (35) и (42) можно заключить, что I_0 также должно быть непустым. Комбинируя (42) и (36), получаем соотношение

$$\frac{c_{k-1}}{\mu_0^k} \langle G_{I_0}(x^k), H_{I_0}(x^k) \rangle \rightarrow 1 + \|H_{I_+}(\bar{x})\|^2, \quad (43)$$

которое совпадает с (34) при $M = 1 + \|H_{I_+}(\bar{x})\|^2$. Из (38) и (41) также получаем предельное соотношение

$$\left\{ \frac{c_{k-1}}{(\mu_0^k)^2} (\lambda_H^k)_{I_+} \mid k \in K \right\} \rightarrow H_{I_+}(\bar{x}). \quad (44)$$

Покажем теперь, что существует бесконечное множество индексов $K_0 \subset K$ и векторы $\bar{\lambda}_G, \bar{\lambda}_H \in \mathbb{R}^s$, для которых справедливо

$$\{\lambda_G^k \mid k \in K_0\} \rightarrow \bar{\lambda}_G, \quad \{\lambda_H^k \mid k \in K_0\} \rightarrow \bar{\lambda}_H, \quad (45)$$

$(\bar{\lambda}_H)_{I_+} = 0$, и выполнено соотношение (33). Отсюда, в частности, следует слабая стационарность точки \bar{x} .

Согласно (31), имеем

$$\begin{aligned}(\lambda_G^k)_{I_0} &= \max\{0, (\bar{\mu}_G^{k-1})_{I_0} - c_{k-1}G_{I_0}(x^k)\} - \mu_0^k H_{I_0}(x^k), \\(\lambda_H^k)_{I_0} &= \max\{0, (\bar{\mu}_H^{k-1})_{I_0} - c_{k-1}H_{I_0}(x^k)\} - \mu_0^k G_{I_0}(x^k).\end{aligned}$$

Предположим, что для некоторого $i \in I_0$ существует такое бесконечное множество индексов $K_i \subset K$, что неравенство

$$(\bar{\mu}_G^{k-1})_i - c_{k-1}G_i(x^k) \geq 0 \quad (46)$$

выполнено для всех $k \in K_i$. В этом случае выполнено

$$(\lambda_G^k)_i = (\bar{\mu}_G^{k-1})_i - c_{k-1}G_i(x^k) - \mu_0^k H_i(x^k),$$

откуда следует

$$\frac{c_{k-1}}{\mu_0^k} G_i(x^k) = \frac{(\bar{\mu}_G^{k-1})_i}{\mu_0^k} - \frac{(\lambda_G^k)_i}{\mu_0^k} - H_i(x^k) \rightarrow 0,$$

и поэтому

$$\frac{c_{k-1}}{\mu_0^k} G_i(x^k) H_i(x^k) \rightarrow 0 \quad (47)$$

для всех $k \in K_i$. Очевидно, что (47) также справедливо, если выполнено неравенство

$$(\bar{\mu}_H^{k-1})_i - c_{k-1}H_i(x^k) \geq 0 \quad (48)$$

для всех k из некоторого бесконечного множества индексов $K_i \subset K$.

Построим теперь множество индексов J_0 следующим образом. Сначала возьмем $J_0 = \emptyset$ и $K_1 = K$. Если для некоторого $i \in I_0 \setminus J_0$ существует такое бесконечное множество индексов $K_2 \subset K_1$, что (46) или (48) выполнено для всех $k \in K_2$, то добавим индекс i к множеству J_0 и повторим данный процесс, используя K_2 вместо K_1 . Будем повторять описанный процесс до тех пор, пока не получим такие множества J_0 и K_q , что для всех $i \in I_0 \setminus J_0$ неравенства (46) и (48) выполняются лишь для конечного числа индексов из K_q .

Введем обозначение $J_+ = I_0 \setminus J_0$. Очевидно, что условие (47) выполнено для всех $i \in J_0$ при $k \in K_q$. С учетом (34), отсюда получаем, что $J_+ \neq \emptyset$, и предельное соотношение

$$\frac{c_{k-1}}{\mu_0^k} \langle G_{J_+}(x^k), H_{J_+}(x^k) \rangle \rightarrow M \quad (49)$$

выполнено для $k \in K_q$ при $k \rightarrow \infty$. Поскольку для $i \in J_+$ неравенства (46) и (48) выполнены лишь для конечного числа элементов последовательности $\{x^k \mid k \in K_q\}$, то из (23) и (31) следует, что для всех достаточно больших $k \in K_q$ выполнено

$$(\lambda_G^k)_{J_+} = -\mu_0^k H_{J_+}(x^k), \quad (\lambda_H^k)_{J_+} = -\mu_0^k G_{J_+}(x^k)$$

С учетом (49), последнее означает справедливость

$$\frac{c_{k-1}}{(\mu_0^k)^3} \langle (\lambda_G^k)_{J_+}, (\lambda_H^k)_{J_+} \rangle \rightarrow M \quad (50)$$

для $k \in K_q$ при $k \rightarrow \infty$.

Покажем теперь, что либо $I_+ = \emptyset$, либо

$$\{(\lambda_H^k)_{I_+} \mid k \in K_q\} \rightarrow 0. \quad (51)$$

Действительно, предположим, что $I_+ \neq \emptyset$, и для некоторого $\varepsilon > 0$ и некоторого бесконечного множества индексов $K_q^0 \subset K_q$ неравенство

$$\|(\lambda_H^k)_{I_+}\| \geq \varepsilon \quad (52)$$

выполнено для всех $k \in K_q^0$. Из соотношения (44) следует

$$\frac{c_{k-1}^2}{(\mu_0^k)^4} \|(\lambda_H^k)_{I_+}\|^2 \rightarrow \|H_{I_+}(\bar{x})\|^2,$$

для $k \in K$, и поэтому, используя (50), получаем

$$\frac{\mu_0^k}{c_{k-1}} \frac{\langle (\lambda_G^k)_{J_+}, (\lambda_H^k)_{J_+} \rangle}{\|(\lambda_H^k)_{I_+}\|^2} \rightarrow \frac{M}{\|H_{I_+}(\bar{x})\|^2} > 0$$

для $k \in K_q$ при $k \rightarrow \infty$. Отсюда и из (37) следует

$$\frac{\langle (\lambda_G^k)_{J_+}, (\lambda_H^k)_{J_+} \rangle}{\|(\lambda_H^k)_{I_+}\|^2} \rightarrow +\infty$$

для $k \in K_q^0$ при $k \rightarrow \infty$. Используя очевидную цепочку неравенств

$$\langle (\lambda_G^k)_{J_+}, (\lambda_H^k)_{J_+} \rangle \leq \|(\lambda_G^k)_{J_+}\| \|(\lambda_H^k)_{J_+}\| \leq \|(\lambda_G^k, (\lambda_H^k)_{I_0})\|^2,$$

получаем

$$\frac{\|(\lambda_G^k, (\lambda_H^k)_{I_0})\|}{\|(\lambda_H^k)_{I_+}\|} \rightarrow \infty$$

для $k \in K_q^0$. Комбинируя последнее соотношение с (52), приходим к соотношению

$$\{\|(\lambda_G^k, (\lambda_H^k)_{I_0})\| \mid k \in K_q^0\} \rightarrow \infty.$$

Наконец, разделив (30) на $\|(\lambda_G^k, (\lambda_H^k)_{I_0})\|$ и применяя лемму 1 для $k \in K_q^0$ при

$$\alpha_k = \frac{1}{\|(\lambda_G^k, (\lambda_H^k)_{I_0})\|}, \quad \omega_k = \frac{(\lambda_H^k)_{I_+}}{\|(\lambda_G^k, (\lambda_H^k)_{I_0})\|},$$

(очевидно, $\{\alpha_k \mid k \in K_q^0\} \rightarrow 0$ и $\{\omega_k \mid k \in K_q^0\} \rightarrow 0$), получаем

$$\left\{ \frac{(\lambda_G^k, (\lambda_H^k)_{I_0})}{\|(\lambda_G^k, (\lambda_H^k)_{I_0})\|} \mid k \in K_q^0 \right\} \rightarrow 0,$$

что, очевидно, является противоречием.

Таким образом, справедливо (51) (в случае $I_+ = \emptyset$ оно становится тривиальным). Применяя лемму 1 при $\alpha_k = 1$ к равенству (30) для $k \in K_q$, получаем, что справедливо (33) для некоторых $\bar{\lambda}_G$ и $(\bar{\lambda}_H)_{I_0}$, удовлетворяющих соотношениям

$$\{\lambda_G^k \mid k \in K_q\} \rightarrow \bar{\lambda}_G, \quad \{(\lambda_H^k)_{I_0} \mid k \in K_q\} \rightarrow (\bar{\lambda}_H)_{I_0}.$$

С учетом (51), для получения (45) остается положить $(\bar{\lambda}_H)_{I_+} = 0$ и $K_0 = K_q$.

Наконец, докажем, что точка \bar{x} является С-стационарной. Поскольку соотношения (33), (45) выполнены для некоторого множества индексов $K_0 \subset K$ и некоторых векторов $\bar{\lambda}_G, \bar{\lambda}_H \in \mathbb{R}^s$, где $(\bar{\lambda}_H)_{I_+} = 0$, достаточно показать, что $(\bar{\lambda}_G)_i (\bar{\lambda}_H)_i \geq 0$ для всех $i \in I_0$.

Построим множества индексов J_0, J_+ и множество K_q тем же способом, что и выше, но используя K_0 вместо K . Ясно, что соотношение (50) выполнено и для новых множеств J_0, J_+ и K_q . Заметим, что подпоследовательность $\{(\lambda_G^k, \lambda_H^k) \mid k \in K_0\}$ ограничена, а значит

$$\left\{ \frac{\langle (\lambda_G^k)_{J_+}, (\lambda_H^k)_{J_+} \rangle}{\mu_0^k} \mid k \in K_0 \right\} \rightarrow 0.$$

Следовательно, из (50) вытекает

$$\frac{c_{k-1}}{(\mu_0^k)^2} \rightarrow +\infty \quad (53)$$

для $k \in K_q$ при $k \rightarrow \infty$.

Поскольку $I_0 = J_0 \cup J_+$, остается показать, что $(\bar{\lambda}_G)_i (\bar{\lambda}_H)_i \geq 0$ для $i \in J_0$ и для $i \in J_+$. Из определения J_0 следует, что для всех $i \in J_0$ выполнено либо (46) для всех $k \in K_q$, либо (48) для всех $k \in K_q$. Ясно, что для любого $i \in J_0$ возможны два случая: либо существует такое бесконечное множество индексов $K_q^1 \subset K_q$, что для всех $k \in K_q^1$ выполнены оба неравенства (46), (48), либо существует бесконечное множество индексов $K_q^2 \subset K_q$, для которого одно из неравенств выполнено для всех $k \in K_q^2$, а второе нарушается для всех $k \in K_q^2$.

Сначала рассмотрим случай, когда для некоторого $i \in J_0$ и некоторого бесконечного множества индексов $K_q^1 \subset K_q$ неравенства (46) и (48) выполнены для всех $k \in K_q^1$. В этом случае из (23) и (31) имеем

$$(\lambda_G^k)_i = (\bar{\mu}_G^{k-1})_i - c_{k-1} G_i(x^k) - \mu_0^k H_i(x^k), \quad (54)$$

$$(\lambda_H^k)_i = (\bar{\mu}_H^{k-1})_i - c_{k-1} H_i(x^k) - \mu_0^k G_i(x^k) \quad (55)$$

для $k \in K_q^1$. Выражая $G_i(x^k)$ из второго уравнения и подставляя в первое, получаем

$$(\lambda_G^k)_i = (\bar{\mu}_G^{k-1})_i - \mu_0^k H_i(x^k) - \frac{c_{k-1}}{\mu_0^k} ((\bar{\mu}_H^{k-1})_i - c_{k-1} H_i(x^k) - (\lambda_H^k)_i),$$

и поэтому

$$H_i(x^k) = \frac{\mu_0^k ((\lambda_G^k)_i - (\bar{\mu}_G^{k-1})_i) - c_{k-1} ((\lambda_H^k)_i - (\bar{\mu}_H^{k-1})_i)}{c_{k-1}^2 - (\mu_0^k)^2},$$

$$c_{k-1} H_i(x^k) = \frac{c_{k-1}^2}{c_{k-1}^2 - (\mu_0^k)^2} \left(\frac{\mu_0^k}{c_{k-1}} ((\lambda_G^k)_i - (\bar{\mu}_G^{k-1})_i) - ((\lambda_H^k)_i - (\bar{\mu}_H^{k-1})_i) \right).$$

Из соотношений (37) и (45) следует, что для $k \in K_q^1$ первый член произведения в правой части последнего уравнения стремится к 1, а второй ограничен. Поэтому последовательность $\{c_{k-1}H_i(x^k) \mid k \in K_q^1\}$ является ограниченной. Повторяя аналогичное рассуждение для $G_i(x^k)$, получаем также ограниченность $\{c_{k-1}G_i(x^k) \mid k \in K_q^1\}$. С учетом (37), из полученных условий следует, что $\mu_0^k G_i(x^k)$ и $\mu_0^k H_i(x^k)$ стремятся к нулю для $k \in K_q^1$. Комбинируя эти соотношения с (46), (48) и применяя к (54), (55), получаем $(\bar{\lambda}_G)_i \geq 0$ и $(\bar{\lambda}_H)_i \geq 0$.

Рассмотрим теперь случай, когда для некоторого $i \in J_0$ и бесконечного множества индексов $K_q^2 \subset K_q$ одно из неравенств (46), (48) выполнено для всех $k \in K_q^2$, а второе нарушается для всех $k \in K_q^2$. Будем предполагать, что выполнено первое неравенство. В этом случае из (23) и (31) следует, что для всех $k \in K_q^2$ выполнено (54), а также равенство

$$(\lambda_H^k)_i = -\mu_0^k G_i(x^k) \quad (56)$$

вместо (55). Если $(\bar{\lambda}_H)_i = 0$, то равенство $(\bar{\lambda}_G)_i(\bar{\lambda}_H)_i = 0$ выполнено автоматически. Предположим, что $(\bar{\lambda}_H)_i \neq 0$. Тогда из (54) и (56) находим

$$\begin{aligned} (\lambda_G^k)_i &= (\bar{\mu}_G^{k-1})_i + \frac{c_{k-1}}{\mu_0^k} (\lambda_H^k)_i - \mu_0^k H_i(x^k), \\ \frac{c_{k-1}}{(\mu_0^k)^2} (\lambda_H^k)_i &= H_i(x^k) - \frac{(\lambda_G^k)_i - (\bar{\mu}_G^{k-1})_i}{\mu_0^k}, \end{aligned}$$

откуда следует

$$\frac{c_{k-1}}{(\mu_0^k)^2} (\lambda_H^k)_i \rightarrow 0$$

при $k \rightarrow \infty$. Поскольку $(\bar{\lambda}_H)_i \neq 0$, полученное соотношение противоречит (53).

Остается показать, что $(\bar{\lambda}_G)_i(\bar{\lambda}_H)_i \geq 0$ для всех $i \in J_+$. Из определения J_+ вытекает существование бесконечного множества индексов $K_q^3 \subset K_q$, для которого неравенства (46), (48) нарушаются для всех $k \in K_q^3$. Поэтому из (23), (31) получаем, что для всех $i \in J_+$ и $k \in K_q^3$ справедливо

$$(\lambda_G^k)_i = -\mu_0^k H_i(x^k), \quad (\lambda_H^k)_i = -\mu_0^k G_i(x^k). \quad (57)$$

Если $(\bar{\lambda}_G)_i = 0$ или $(\bar{\lambda}_H)_i = 0$ для некоторого i то условие $(\bar{\lambda}_G)_i(\bar{\lambda}_H)_i \geq 0$, очевидно, выполнено. Предположим, что $(\bar{\lambda}_G)_i \neq 0$ и $(\bar{\lambda}_H)_i \neq 0$. В этом случае из (37) и (57) вытекает, что $|c_{k-1}G_i(x^k)|$ и $|c_{k-1}H_i(x^k)|$ стремятся к бесконечности для $k \in K_q^3$ при $k \rightarrow \infty$. Тогда из равенства (23), а также из нарушения (46) и (48) следует, что $G_i(x^k) > 0$ и $H_i(x^k) > 0$ для всех достаточно больших k . Применяя эти неравенства к (57), получаем $(\bar{\lambda}_G)_i < 0$ и $(\bar{\lambda}_H)_i < 0$ (поскольку они не равны нулю). Таким образом, $(\bar{\lambda}_G)_i(\bar{\lambda}_H)_i > 0$. ■

Следующие два примера показывают, что в предположениях теоремы 4 предельные точки траекторий, генерируемых алгоритмом 1, не обязательно являются сильно стационарными или даже М-стационарными. В первом примере все индексы лежат в I_0 , а во втором как I_0 , так и его дополнение являются непустыми.

Пример 3. Рассмотрим пример `scholtes3` из MacMPES [83] (похожие эффекты также наблюдаются в примерах `scale4` и `scale5` из этой коллекции). Пусть $n = 2$, $s = 1$, $f(x) = ((x_1 - 1)^2 + (x_2 - 1)^2)/2$, $G(x) = x_1$, $H(x) = x_2$. Задача (19) с такими данными имеет два решения $(1, 0)$ и $(0, 1)$, оба из которых являются сильно стационарными, а также одну неоптимальную С-стационарную (но не М-стационарную) точку $\bar{x} = (0, 0)$, удовлетворяющую условию МРСС-LICQ.

Если запустить solver ALGENCAN (с отключенным ньютоновским ускорителем, а также с отключенными средствами решения подзадач «до второго порядка оптимальности», т. е. чистый алгоритм 1) из начальной точки по умолчанию ($x^0 \in \mathbb{R}^2$ вблизи 0 и $\mu^0 = (0, 0, 0)$), то он будет сходиться к указанной С-стационарной точке \bar{x} , причем $\mu_G^k = 0$ и $\mu_H^k = 0$ для всех k и $c_k \rightarrow +\infty$. Это происходит потому, что метод безусловной минимизации, используемый для решения подзадач вида (5) в алгоритме 1, выбирает седловую точку $x^{k+1} = (t_k, t_k)$ функции $L_{c_k}(\cdot, \bar{\mu}^k)$, где $t_k \approx 1/c_k^{1/3}$ определяется уравнением

$$t - 1 + (\bar{\mu}_0^k + c_k t^2)t = 0.$$

(Конечно, такой исход возможен только для очень специальных начальных точек, используемых методом безусловной минимизации, в данном случае, точек на прямой $x_1 = x_2$.) Заметим, что согласно (24), $\mu_0^{k+1} = \bar{\mu}_0^k + c_k t_k^2 \approx \bar{\mu}_0^k + c_k^{1/3} \rightarrow +\infty$.

Отметим также, что ALGENCAN трактует ограничения неотрицательности как простые ограничения, не включая их в модифицированную функцию Лагранжа. С другой стороны, поскольку $\mu_G^k = 0$ и $\mu_H^k = 0$, такая частичная модифицированная функция совпадает с полной на \mathbb{R}_+^2 , и поэтому полученные результаты адекватно описывают поведение ALGENCAN: алгоритм, использующий полную модифицированную функцию, показывал бы аналогичное поведение.

Поведение, похожее на описанное в примере 3, отмечалось в [32, пример 2] для задачи с такими же ограничениями, но с целевой функцией $f(x) = -x_1 - x_2$. В такой задаче ALGENCAN сходится к С-стационарной точке $\bar{x} = 0$, которая является глобальным максимумом. Однако следует подчеркнуть, что небольшое возмущение начальной точки приводит к сходимости к сильно стационарным точкам. Последнее означает, что нежелательная сходимость к неоптимальным С-стационарным точкам является крайне нетипичной.

Пример 4. Рассмотрим задачу (19) со следующими данными: $n = 4$, $s = 2$, $f(x) = -\sqrt{2}(x_1 + x_3) - x_2 + (x_4 - 1)^2/2$, $G(x) = (x_1, x_2)$, $H(x) = (x_3, x_4)$. Рассмотрим допустимую точку $\bar{x} = (0, 0, 0, 1)$, которая является неоптимальной и удовлетворяет МРСС-LICQ. Легко видеть, что эта точка является С-стационарной, но не М-стационарной.

Для такой задачи равенство в (25) для каждого k эквивалентно следующей системе:

$$\begin{aligned} -\sqrt{2} - (\mu_G^k)_1 + \mu_0^k x_3^k &= 0, & -1 - (\mu_G^k)_2 + \mu_0^k x_4^k &= 0, \\ -\sqrt{2} - (\mu_H^k)_1 + \mu_0^k x_1^k &= 0, & x_4^k - 1 - (\mu_H^k)_2 + \mu_0^k x_2^k &= 0, \end{aligned}$$

где μ_G^k , μ_H^k и μ_0^k определяются согласно (23), (24). Для каждого $k = 1, 2, \dots$ будем выбирать $\bar{\mu}_G^{k-1} = \bar{\mu}_H^{k-1} = 0$ и $\bar{\mu}_0^{k-1} = 1$. Тогда точки $x_1^k = x_3^k = \sqrt{2}/t_k$, $x_2^k = -1/t_k^2$, $x_4^k = 1 + 1/t_k$ и $\mu_0^k = t_k$, $\mu_G^k = (0, t_k)$, $\mu_H^k = 0$, где $t_k = c_{k-1}^{1/3}$, удовлетворяют всех необходимым соотношениям. Ясно, что если $c_k \rightarrow +\infty$, то последовательность $\{x^k\}$ сходится к \bar{x} .

Замечание 2. Легко показать, что первая часть утверждения теоремы 4 (т.е. сильная стационарность точки \bar{x} в случае, когда $\{\mu_0^k\}$ имеет ограниченную подпоследовательность, и соответствующая подпоследовательность $\{x^k\}$ сходится к \bar{x}) остается справедливым, если заменить условие МРСС-LICQ в точке \bar{x} на MFCQ для так называемой расширенной задачи нелинейного программирования (RNLP, от англ. Relaxed nonlinear programming problem), соответствующей МРСС (19):

$$f(x) \rightarrow \min, \quad G_{I_G(\bar{x}) \setminus I_H(\bar{x})}(x) = 0, \quad H_{I_H(\bar{x}) \setminus I_G(\bar{x})}(x) = 0, \quad G_{I_0(\bar{x})}(x) \geq 0, \quad H_{I_0(\bar{x})}(x) \geq 0.$$

Очевидно, что MFCQ для RNLP слабее условия МРСС-LICQ.

Однако, условие МРСС-LICQ является существенным для второй части утверждения теоремы 4. Следующий пример показывает, что если $\{\mu_0^k\}$ стремится к бесконечности, то выполнение MFCQ для RNLP не обеспечивает даже слабой стационарности предельных точек.

Пример 5. Рассмотрим задачу (19) со следующими данными: $n = 6$, $s = 3$, $f(x) = -3x_1/2 + x_3(-1 - x_4^2 + x_5^2) - x_6$, $G(x) = (|x_1|^{3/2} + x_2, -|x_1|^{3/2} + x_2, x_3)$, $H(x) = (x_4, x_5, x_6)$. Легко видеть, что допустимая точка $\bar{x} = 0$ удовлетворяет условию MFCQ для RNLP, но при этом она не является ни оптимальной, ни слабо стационарной.

Пусть $x_1 \geq 0$. Тогда первое равенство в (25) для каждого k эквивалентно следующей системе:

$$\begin{aligned} -1 - x_1^{1/2}((\mu_G^k)_1 - (\mu_G^k)_2 - \mu_0^k(x_4 - x_5)) &= 0, & -(\mu_G^k)_1 - (\mu_G^k)_2 + \mu_0^k(x_4 + x_5) &= 0, \\ -1 - x_4^2 + x_5^2 - (\mu_G^k)_3 + \mu_0^k x_6 &= 0, & -2x_3 x_4 - (\mu_H^k)_1 + \mu_0^k(x_1^{3/2} + x_2) &= 0, \\ 2x_3 x_5 - (\mu_H^k)_2 + \mu_0^k(-x_1^{3/2} + x_2) &= 0, & -1 - (\mu_H^k)_3 + \mu_0^k x_3 &= 0, \end{aligned}$$

где μ_G^k , μ_H^k и μ_0^k определяются согласно (23), (24).

Для каждого $k = 1, 2, \dots$ будем выбирать $\bar{\mu}_G^{k-1} = \bar{\mu}_H^{k-1} = 0$ и $\bar{\mu}_0^{k-1} = 0$. Тогда точки $x^k = (t_k^6, 0, t_k^4, t_k/2, t_k/2, t_k^4)$ и $\mu_0^k = 1/t_k^4$, $\mu_G^k = (0, 1/t_k^3, 0)$, $\mu_H^k = 0$, где $t_k = 1/c_{k-1}^{1/12}$, удовлетворяют всем необходимым соотношениям. Ясно, что если $c_k \rightarrow +\infty$, то последовательность $\{x^k\}$ сходится к \bar{x} .

В завершение данного раздела отметим пример в [53, раздел 7.2], который показывает, что метод SQP, применяемый к MPC, может сходиться к произвольной допустимой точке, которая может не быть слабо стационарной даже при выполненном условии MPC-LICQ. Как было показано выше, такие ситуации исключены для алгоритма 1.

2.2. Ускорение финальной фазы

Как уже обсуждалось в разделе 1.3.2, МФЛ также обладает сильной теорией локальной сходимости. Более того, результаты вычислительного эксперимента, приведенные в приложении А.2, свидетельствуют о том, что если параметр штрафа стремится к бесконечности, то сверхлинейная сходимость часто сохраняется даже в случае сходимости двойственной траектории к критическим множителям Лагранжа. В этом состоит существенное преимущество МФЛ по сравнению с SQP и его стабилизированной версией, для которых сходимость к критическим множителям всегда приводит к потере сверхлинейной сходимости.

Однако, к сожалению, МФЛ обладает следующим серьезным недостатком: увеличение параметра штрафа, необходимое для достижения сверхлинейной скорости сходимости, приводит к ухудшению обусловленности подзадач метода, что повышает вычислительные затраты на их решение. Как показывают результаты вычислительного эксперимента, приведенные в приложении А.3, высокая скорость сходимости обычно не компенсирует эти затраты, и метод со сверхлинейной скоростью сходимости часто уступает по общей эффективности методу, в котором последовательность параметров штрафа остается ограниченной, а скорость сходимости — линейной. Более того, согласно результатам, приведенным в приложении А.1, солвер ALGENCAN, реализующий МФЛ, обладает меньшей эффективностью, чем другие алгоритмы.

Возможное решение описанной проблемы состоит в том, чтобы в финальной фазе метода (когда есть основания считать, что текущее приближение уже достаточно близко к решению) переключаться на некоторый локальный быстрый алгоритм. Данный подход к ускорению методов множителей, предложенный в [37], реализован в текущей версии солвера ALGENCAN: при достаточной малости невязки системы ККТ (2) в текущей точке осуществляется оценка множества активных в искомом решении ограничений-неравенств, формируется соответствующая задача с ограничениями-равенствами, и к ее системе Лагранжа применяется метод Ньютона. Однако, как отмечено выше, в случае нерегулярных ограничений сверхлинейная скорость сходимости такого метода Ньютона–Лагранжа обычно все равно теряется из-за притяжения его двойственной траектории к критическим множителям Лагранжа.

Напрашивающейся альтернативой является использование в качестве ускорителя не метода Ньютона–Лагранжа, а его стабилизированной версии (СМНЛ). Как было показано в

разделе 1.3.1, тенденция притяжения к критическим множителям Лагранжа для этого метода не столь выражена: существуют широкие прямо-двойственные области, при старте из которых такое притяжение для данного метода не имеет места. Вопрос состоит в том, будут ли эти привлекательные свойства СМНЛ достаточными для его использования в качестве ускорителя МФЛ: если двойственные траектории последнего притягиваются к критическим множителям, то точки этих траекторий могут никогда не попадать в область начальных приближений, подходящих для успешного применения такого ускорителя. Таким образом, и здесь роль критических множителей Лагранжа оказывается определяющей.

В данном разделе будем рассматривать алгоритм 1 со следующей модификацией: на шаге 1 будем вычислять $x^{k+1} \in \mathbb{R}^n$ как приближенное решение подзадачи (5), а именно, как точку, удовлетворяющую условию

$$\left\| \frac{\partial L_{ck}}{\partial x}(x^{k+1}, \bar{\lambda}^k, \bar{\mu}^k) \right\|_{\infty} \leq \varepsilon_k,$$

где $\{\varepsilon_k\} \subset \mathbb{R}_+$ — некоторая последовательность.

Обсудим предложенную в [37] конструкцию ускорителя алгоритма 1, которая с незначительными изменениями реализована в ALGENCAN. Определим функцию естественной невязки $\sigma: \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m \rightarrow \mathbb{R}_+$ системы ККТ (2):

$$\sigma(x, \lambda, \mu) = \left\| \left(\frac{\partial L}{\partial x}(x, \lambda, \mu), h(x), \min\{\mu, -g(x)\} \right) \right\| \quad (58)$$

(в ALGENCAN в (58) используется ∞ -норма). При заданной точности остановки $\varepsilon > 0$, ускоритель запускается после того, как алгоритмом 1 достигается заданная точность $\hat{\varepsilon} > \varepsilon$, а именно, когда для текущего приближения (x^k, λ^k, μ^k) оказывается выполненным условие

$$\sigma(x^k, \lambda^k, \mu^k) \leq \hat{\varepsilon}, \quad (59)$$

но при этом еще не достигнута точность остановки, т. е. (59) не выполняется с $\hat{\varepsilon}$ замененным на ε ; в противном случае алгоритм останавливается, поскольку приближение (x^k, λ^k, μ^k) нужного качества найдено самим методом множителей.

Если запуск ускорителя оканчивается неудачей, то все его итерации отбрасываются, $\hat{\varepsilon}$ уменьшается (оставаясь больше ε), и продолжается выполнение алгоритма 1 с того места, где оно было прервано, и до тех пор, пока не будет достигнута новая точность $\hat{\varepsilon}$. После этого вновь запускается ускоритель, и т. д. При этом под неудачей ускорителя понимается ситуация, когда после заданного количества K его шагов не удается получить приближение $(\tilde{x}, \tilde{\lambda}, \tilde{\mu}) \in \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m$ удовлетворяющее

$$\sigma(\tilde{x}, \tilde{\lambda}, \tilde{\mu}) \leq \varepsilon. \quad (60)$$

Запуск ускорителя из точки $(\hat{x}, \hat{\lambda}, \hat{\mu}) = (x^k, \lambda^k, \mu^k)$ осуществляется следующим образом. Сначала формируется рабочее множество индексов $\hat{A} \subset \{1, \dots, m\}$, оценивающее множество

номеров ограничений-неравенств, активных в искомом решении:

$$\hat{A} = \{i = 1, \dots, m \mid g_i(\hat{x}) \geq -\hat{\varepsilon}\}. \quad (61)$$

Ограничения-неравенства с номерами $i \in \{1, \dots, m\} \setminus \hat{A}$ отбрасываются, а остальные ограничения-неравенства заменяются на равенства

$$g_i(x) + \frac{1}{2}s_i^2 = 0, \quad i \in \hat{A},$$

где s_i , $i \in \hat{A}$, — вспомогательные переменные. Согласно [37], введение вспомогательных переменных значительно повышает робастность ускорителя, снимая негативный эффект от возможного включения в \hat{A} индексов ограничений, не являющихся активными в решении.

Далее, не ограничивая общности, предположим, что $\hat{A} = \{1, \dots, \hat{m}\}$ при некотором $\hat{m} \in \{0, 1, \dots, m\}$. Вводится отображение $\hat{h}: \mathbb{R}^n \times \mathbb{R}^{\hat{m}} \rightarrow \mathbb{R}^l \times \mathbb{R}^{\hat{m}}$,

$$\hat{h}(x, s) = \left(h(x), g_1(x) + \frac{1}{2}s_1^2, \dots, g_{\hat{m}}(x) + \frac{1}{2}s_{\hat{m}}^2 \right), \quad (62)$$

и к задаче

$$f(x) \rightarrow \min, \quad \hat{h}(x, s) = 0, \quad (63)$$

применяется МНЛ, т.е. метод Ньютона для системы Лагранжа

$$\frac{\partial \hat{L}}{\partial(x, s)}((x, s), (\lambda, \nu)) = 0, \quad \hat{h}(x, s) = 0, \quad (64)$$

этой задачи, где $\hat{L}: (\mathbb{R}^n \times \mathbb{R}^{\hat{m}}) \times (\mathbb{R}^l \times \mathbb{R}^{\hat{m}}) \rightarrow \mathbb{R}$,

$$\hat{L}((x, s), (\lambda, \nu)) = f(x) + \langle (\lambda, \nu), \hat{h}(x, s) \rangle, \quad (65)$$

есть функция Лагранжа задачи (63). А именно, для текущего приближения $((x^k, s^k), (\lambda^k, \nu^k)) \in (\mathbb{R}^n \times \mathbb{R}^{\hat{m}}) \times (\mathbb{R}^l \times \mathbb{R}^{\hat{m}})$ к решению системы (64) следующее приближение $((x^{k+1}, s^{k+1}), (\lambda^{k+1}, \nu^{k+1}))$ ищется как решение линейной системы

$$\begin{aligned} \begin{pmatrix} \frac{\partial^2 \hat{L}}{\partial(x, s)^2}((x^k, s^k), (\lambda^k, \nu^k)) & (\hat{h}'(x^k, s^k))^T \\ \hat{h}'(x^k, s^k) & 0 \end{pmatrix} \begin{pmatrix} (x, s) - (x^k, s^k) \\ (\lambda, \nu) - (\lambda^k, \nu^k) \end{pmatrix} = \\ = - \begin{pmatrix} \frac{\partial \hat{L}}{\partial(x, s)}((x^k, s^k), (\lambda^k, \nu^k)) \\ \hat{h}(x^k, s^k) \end{pmatrix}. \end{aligned} \quad (66)$$

В роли начального приближения для МНЛ используется точка $((\hat{x}, s^0), (\hat{\lambda}, \nu^0))$, где $s_i^0 = \sqrt{2 \max\{0, -g_i(\hat{x})\}}$, $\nu_i^0 = \hat{\mu}_i$, $i \in \hat{A}$. Для каждого $k = 1, \dots, K$ проверяется выполнение условия

$$\hat{\sigma}((x^k, s^k), (\lambda^k, \nu^k)) \leq \varepsilon, \quad (67)$$

где функция $\hat{\sigma}: (\mathbb{R}^n \times \mathbb{R}^{\hat{m}}) \times (\mathbb{R}^l \times \mathbb{R}^{\hat{m}}) \rightarrow \mathbb{R}_+$ представляет собой невязку системы Лагранжа (64) и определяется как

$$\hat{\sigma}((x, s), (\lambda, \nu)) = \left\| \left(\frac{\partial \hat{L}}{\partial(x, s)}((x, s), (\lambda, \nu)), \hat{h}(x, s) \right) \right\|. \quad (68)$$

Если условие (67) выполнено для некоторого k , то далее проверяется выполнение условия (60) для $\tilde{x} = x^k$, $\tilde{\lambda} = \lambda^k$, $\tilde{\mu}_i = \max\{0, \nu_i^k\}$, $i \in \hat{A}$, $\tilde{\mu}_i = 0$, $i \in \{1, \dots, m\} \setminus \hat{A}$. Если это условие выполняется, то алгоритм завершает работу, поскольку найдено приближение $(\tilde{x}, \tilde{\lambda}, \tilde{\mu})$ нужного качества. В противном случае, если $k < K$, то ищется следующее приближение согласно (66); если же $k = K$, то запуск ускорителя объявляется неудачным.

В ALGENCAN по умолчанию используются $K = 10$, $\varepsilon = 10^{-8}$ и начальное значение $\hat{\varepsilon} = 10^{-4}$. Возможное правило уменьшения $\hat{\varepsilon}$ см. в [37]. Кроме того, ускоритель в ALGENCAN всегда запускается для исходной задачи, даже если алгоритм 1 запускается для масштабированного варианта (8). В последнем случае двойственное начальное приближение ускорителя $(\hat{\lambda}, \hat{\mu})$ модифицируется соответствующим образом.

Имея в виду возможное улучшение поведения описанной схемы применительно к задачам с нерегулярными ограничениями, рассмотрим также вариант ускорителя, в котором вместо итераций МНЛ для задачи (63) осуществляются итерации СМНЛ, т. е. (66) заменяется линейной системой

$$\begin{pmatrix} \frac{\partial^2 \hat{L}}{\partial(x, s)^2}((x^k, s^k), (\lambda^k, \nu^k)) & (\hat{h}'(x^k, s^k))^T \\ \hat{h}'(x^k, s^k) & -\sigma_k I \end{pmatrix} \begin{pmatrix} (x, s) - (x^k, s^k) \\ (\lambda, \nu) - (\lambda^k, \nu^k) \end{pmatrix} = \\ = - \begin{pmatrix} \frac{\partial \hat{L}}{\partial(x, s)}((x^k, s^k), (\lambda^k, \nu^k)) \\ \hat{h}(x^k, s^k) \end{pmatrix}, \quad (69)$$

где $\sigma_k \geq 0$ — параметр стабилизации, выбор которого основан на оценке расстояния от $((x^k, s^k), (\lambda^k, \nu^k))$ до множества решений системы Лагранжа (64) задачи (63). А именно, будем полагать, что $\sigma_k = \hat{\sigma}((x^k, s^k), (\lambda^k, \nu^k))$.

Для СМНЛ, применяемого к задаче (63), справедлива следующая теорема о локальной сходимости, вытекающая из результатов в [71]. Эта теорема показывает, что если СМНЛ запускается в качестве ускорителя алгоритма 1 вблизи стационарной точки при двойственном приближении, достаточно близком к некритическому множителю Лагранжа, удовлетворяющему условию строгой дополненности, то имеет место сверхлинейная сходимость итераций ускорителя.

Теорема 5. Пусть функция $f: \mathbb{R}^n \rightarrow \mathbb{R}$ и отображения $h: \mathbb{R}^n \rightarrow \mathbb{R}^l$ и $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$ дважды дифференцируемы в окрестности точки $\bar{x} \in \mathbb{R}^n$, и их вторые производные непрерывны в этой точке. Пусть \bar{x} — стационарная точка задачи (1), а $(\bar{\lambda}, \bar{\mu}) \in \mathcal{M}(\bar{x})$ — некри-

тический множитель Лагранжа, удовлетворяющий условию строгой дополнителности $\bar{\mu}_A(\bar{x}) > 0$.

Тогда для любого $\hat{\varepsilon} > 0$ и любого начального приближения $(x^0, \lambda^0, \mu^0) \in \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m$, достаточно близкого к $(\bar{x}, \bar{\lambda}, \bar{\mu})$, если \hat{A} определяется согласно (61) при $\hat{x} = x^0$, и если $s^0 = \sqrt{2 \max\{0, -g_{\hat{A}}(x^0)\}}$, $\nu^0 = \mu_{\hat{A}}^0$, то существует единственная траектория $\{((x^k, s^k), (\lambda^k, \nu^k))\}$ такая, что для каждого $k = 0, 1, \dots$ точка $\{((x^{k+1}, s^{k+1}), (\lambda^{k+1}, \nu^{k+1}))\}$ удовлетворяет системе (69), в которой \hat{h} определяется согласно (62), \hat{L} согласно (65), а $\sigma_k = \hat{\sigma}((x^k, s^k), (\lambda^k, \nu^k))$ согласно (68); траектория сверхлинейно сходится к $((\bar{x}, \bar{s}), (\lambda^*, \nu^*))$, где $\bar{s} = \sqrt{-2g_{\hat{A}}(\bar{x})}$, $\nu^* = \mu_{\hat{A}}^*$, при некотором $(\lambda^*, \mu^*) \in \mathcal{M}(\bar{x})$, причем $\mu_i^* = 0$, $i \in \{1, \dots, m\} \setminus \hat{A}$.

Доказательство. Достаточная близость x^0 к \bar{x} гарантирует включение $A(\bar{x}) \subset \hat{A}$. Отсюда следует, что \bar{x} является стационарной точкой задачи

$$f(x) \rightarrow \min, \quad h(x) = 0, \quad g_{\hat{A}}(x) \leq 0, \quad (70)$$

а $(\bar{\lambda}, \bar{\mu}_{\hat{A}})$ — отвечающий этой стационарной точке не критический множитель Лагранжа, для которого выполнено условие строгой дополнителности. Применяя к задаче (70) теорему 2 из [71], получаем требуемое утверждение. ■

Отметим, что при нарушении условия строгой дополнителности на сверхлинейную сходимость СМНЛ, применяемого к задаче (63), рассчитывать не приходится. Действительно, в этом случае к задаче (70) можно применить предложение 5 из [71], из которого следует, что в задаче (63) множитель $(\bar{\lambda}, \bar{\mu}_{\hat{A}})$, отвечающий стационарной точке (\bar{x}, \bar{s}) , неизбежно является критическим.

К сожалению, несмотря на неплохие теоретические свойства, описанная выше конструкция ускорителя не позволяет добиться существенного повышения эффективности. Результаты вычислительного эксперимента, приведенные в приложении А.4, показывают, что замена используемого в ALGENCAN стандартного ускорителя МНЛ на СМНЛ не приводит к сколько-нибудь существенному повышению эффективности или робастности солвера. 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 СМНЛ обычно все равно сходится к критическому множителю.

Глава 3

Глобализация сходимости стабилизированного метода последовательного квадратичного программирования

Как было продемонстрировано в разделе 1.3.1, стабилизированный метод последовательного квадратичного программирования (sSQP) обладает сильной теорией локальной сходимости: согласно теоремам 1.2 и 1.3, для сверхлинейной сходимости метода не требуется выполнение условий регулярности ограничений.

Вместе с тем, вопрос о глобализации сходимости метода sSQP, по сути, является открытым. К известным попыткам решения этой проблемы относятся комбинирование sSQP с методами неточного восстановления допустимости [47], а также прямо-двойственные методы множителей [55, 56]. Однако, убедительных свидетельств хорошего численного поведения этих подходов на задачах с нерегулярными ограничениями на сегодняшний день неизвестно.

Отметим, что для метода SQP известны эффективные способы глобализации сходимости [11, § 5.4], [72, гл. 6], которые используются во многих успешных оптимизационных пакетах. Однако, данные способы неприменимы к sSQP. В частности, решения подзадач sSQP не обязательно являются направлениями убывания для штрафных функций, которые обычно используются для глобализации сходимости (например, для ℓ_1 -штрафной функции).

В данной главе предлагается ряд новых подходов к глобализации сходимости sSQP. В разделе 3.1 рассматриваются гибридные подходы к глобализации сходимости, в которых sSQP объединяется с некоторым методом внешней фазы, обладающим разумными свойствами глобальной сходимости. В разделе 3.2 предлагается подход, основанный на комбинировании sSQP и метода модифицированных функций Лагранжа. Наконец, раздел 3.3 посвящен

глобализации сходимости sSQP с помощью прямо-двойственных точных гладких штрафных функций.

3.1. Гибридные подходы к глобализации сходимости

Будем рассматривать задачу оптимизации со смешанными ограничениями

$$f(x) \rightarrow \min, \quad h(x) = 0, \quad g(x) \leq 0, \quad (1)$$

где $f: \mathbb{R}^n \rightarrow \mathbb{R}$ — дважды дифференцируемая функция, а $h: \mathbb{R}^n \rightarrow \mathbb{R}^l$ и $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$ — дважды дифференцируемые отображения.

В этом разделе предлагаются два способа гибридной глобализации сходимости sSQP для задачи (1): алгоритм с возвратами и алгоритм с рекордами. Идея алгоритма с возвратами заключается в том, что на каждой итерации из текущей точки (x^k, λ^k, μ^k) 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.

Определим функцию естественной невязки $\sigma: \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m \rightarrow \mathbb{R}_+$ системы ККТ задачи (1):

$$\sigma(x, \lambda, \mu) = \left\| \left(\frac{\partial L}{\partial x}(x, \lambda, \mu), h(x), \min\{\mu, -g(x)\} \right) \right\|,$$

где $L: \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m \rightarrow \mathbb{R}$ — функция Лагранжа для задачи (11):

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

Алгоритм 1. Фиксируем параметр $q \in (0, 1)$. Выбираем $(x^0, \lambda^0, \mu^0) \in \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m$ и полагаем $k = 0$.

1. Запоминаем $\hat{k} = k$ и $(\hat{x}, \hat{\lambda}, \hat{\mu}) = (x^k, \lambda^k, \mu^k)$.
2. Осуществляем шаг sSQP, т. е. вычисляем $(x^{k+1}, \lambda^{k+1}, \mu^{k+1}) \in \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m$ как стационарную точку задачи

$$\begin{aligned} \langle f'(x^k), x - x^k \rangle + \frac{1}{2} \left\langle \frac{\partial^2 L}{\partial x^2}(x^k, \lambda^k, \mu^k)(x - x^k), x - x^k \right\rangle + \frac{\sigma_k}{2} (\|\lambda\|^2 + \|\mu\|^2) \rightarrow \min \\ h(x^k) + h'(x^k)(x - x^k) - \sigma_k(\lambda - \lambda^k) = 0, \\ g(x^k) + g'(x^k)(x - x^k) - \sigma_k(\mu - \mu^k) \leq 0, \end{aligned} \quad (2)$$

где $\sigma_k = \sigma(x^k, \lambda^k, \mu^k)$. Если задача (2) не имеет стационарных точек, переходим к п. 4.

3. Если

$$\sigma(x^{k+1}, \lambda^{k+1}, \mu^{k+1}) \leq q\sigma(x^k, \lambda^k, \mu^k), \quad (3)$$

то увеличиваем номер шага k на 1 и переходим к п. 2.

4. Если $k > \hat{k}$ (т.е. текущая точка (x^k, λ^k, μ^k) сгенерирована с помощью шага sSQP), то восстанавливаем сохраненные значения, полагая $k = \hat{k}$ и $(x^k, \lambda^k, \mu^k) = (\hat{x}, \hat{\lambda}, \hat{\mu})$.
5. Вычисляем $(x^{k+1}, \lambda^{k+1}, \mu^{k+1}) \in \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m$ с помощью шага метода внешней фазы, увеличиваем номер шага k на 1 и переходим к п. 1.

Свойства глобальной сходимости приведенного концептуального алгоритма весьма прозрачны. Действительно, возможны только два сценария: либо все итерации алгоритма 1, начиная с некоторого номера k , являются итерациями метода sSQP, либо все итерации сгенерированы с помощью метода внешней фазы, и в последнем случае алгоритм 1 наследует свойства глобальной сходимости этого метода. В первом же случае из (3) вытекает следующая теорема.

Теорема 1. Пусть функция $f: \mathbb{R}^n \rightarrow \mathbb{R}$ и отображения $h: \mathbb{R}^n \rightarrow \mathbb{R}^l$ и $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$ дважды дифференцируемы на \mathbb{R}^n . Пусть последовательность $\{(x^k, \lambda^k, \mu^k)\}$ сгенерирована алгоритмом 1, и пусть все члены этой последовательности, начиная с некоторого, получены шагами sSQP.

Тогда

$$\sigma(x^k, \lambda^k, \mu^k) \rightarrow 0 \text{ при } k \rightarrow \infty,$$

и, в частности, если $(\bar{x}, \bar{\lambda}, \bar{\mu}) \in \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m$ является предельной точкой последовательности $\{(x^k, \lambda^k, \mu^k)\}$, то \bar{x} — стационарная точка задачи (1), а $(\bar{\lambda}, \bar{\mu})$ — отвечающий ей множитель Лагранжа.

Напомним, что достаточное условие второго порядка оптимальности (SOSC) в стационарной точке \bar{x} задачи (1) для отвечающего ей множителя Лагранжа $(\bar{\lambda}, \bar{\mu})$ имеет вид

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

где

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

есть критический конус задачи (1) в точке \bar{x} .

Для обоснования сверхлинейной скорости сходимости алгоритма 1 потребуется следствие из теоремы 1.2.

Следствие 1. Пусть $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $h: \mathbb{R}^n \rightarrow \mathbb{R}^l$ и $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$ дважды дифференцируемы в окрестности точки $\bar{x} \in \mathbb{R}^n$, а их вторые производные непрерывны в этой точке. Пусть \bar{x} — стационарная точка задачи (1), и пусть для отвечающего ей множителя Лагранжа $(\bar{\lambda}, \bar{\mu}) \in \mathbb{R}^l \times \mathbb{R}^m$ выполняется SOS (4).

Тогда найдется функция $\chi: \mathbb{R}_+ \rightarrow \mathbb{R}_+$ такая, что $\chi(t) = o(t)$, и для любой начальной точки $(x^0, \lambda^0, \mu^0) \in \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m$, достаточно близкой к $(\bar{x}, \bar{\lambda}, \bar{\mu})$, метод sSQP, в котором в качестве $(x^{k+1}, \lambda^{k+1}, \mu^{k+1})$ берется ближайшая к (x^k, λ^k, μ^k) стационарная точка задачи (2), определяет траекторию $\{(x^k, \lambda^k, \mu^k)\}$, причем

$$\sigma(x^{k+1}, \lambda^{k+1}, \mu^{k+1}) \leq \chi(\sigma(x^k, \lambda^k, \mu^k)). \quad (5)$$

Доказательство. Из теоремы 1.2 вытекает оценка

$$\|x^{k+1} - \bar{x}\| + \text{dist}((\lambda^{k+1}, \mu^{k+1}), \mathcal{M}(\bar{x})) \leq \chi(\|x^k - \bar{x}\| + \text{dist}((\lambda^k, \mu^k), \mathcal{M}(\bar{x})))$$

для некоторой функции $\chi: \mathbb{R}_+ \rightarrow \mathbb{R}_+$, удовлетворяющей $\chi(t) = o(t)$, и для всех достаточно больших k . Поэтому для доказательства следствия достаточно показать, что для некоторых констант $\ell_1, \ell_2 > 0$ выполнены неравенства

$$\ell_1 \sigma(x, \lambda, \mu) \leq \|x - \bar{x}\| + \text{dist}((\lambda, \mu), \mathcal{M}(\bar{x})) \leq \ell_2 \sigma(x, \lambda, \mu). \quad (6)$$

для всех $(x, \lambda, \mu) \in \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m$, достаточно близких к $(\bar{x}, \bar{\lambda}, \bar{\mu})$. Первое неравенство в (6) следует из липшицевости функции σ в окрестности точки $(\bar{x}, \bar{\lambda}, \bar{\mu})$ и равенства $\sigma(\bar{x}, \bar{\lambda}, \bar{\mu}) = 0$. Второе неравенство в (6), согласно предложению 1.1, эквивалентно не критичности множителя $(\bar{\lambda}, \bar{\mu})$ (а значит, следует из SOS). ■

Отметим, что в следствии 1 можно ослабить предположение о том, что на каждом шаге метода sSQP в качестве $(x^{k+1}, \lambda^{k+1}, \mu^{k+1})$ берется ближайшая к (x^k, λ^k, μ^k) стационарная точка соответствующей подзадачи. Данное предположение позволяет гарантировать, что для всех (x^k, λ^k, μ^k) , достаточно близких к $(\bar{x}, \bar{\lambda}, \bar{\mu})$, и для некоторого $\ell > 0$ будет выполнено неравенство

$$\|(x^{k+1} - x^k, \lambda^{k+1} - \lambda^k, \mu^{k+1} - \mu^k)\| \leq \ell(\|x^k - \bar{x}\| + \text{dist}((\lambda^k, \mu^k), \mathcal{M}(\bar{x}))), \quad (7)$$

а значит для генерируемой траектории будет справедливо утверждение теоремы 1.2. Поэтому следствие 1 (а также последующие утверждения о скорости сходимости) останется справедливым, если для некоторого $\ell > 0$ и для всех k в качестве $(x^{k+1}, \lambda^{k+1}, \mu^{k+1})$ выбирать любую стационарную точку подзадачи (2), удовлетворяющую (7). Впрочем, в практических реализациях sSQP условия такого рода обычно игнорируются.

Сверхлинейная скорость сходимости алгоритма 1 устанавливается следующей теоремой.

Теорема 2. Пусть в предположениях следствия 1 точка $(\bar{x}, \bar{\lambda}, \bar{\mu})$ является предельной точкой траектории $\{(x^k, \lambda^k, \mu^k)\}$, сгенерированной алгоритмом 1, в котором на шаге sSQP в качестве $(x^{k+1}, \lambda^{k+1}, \mu^{k+1})$ берется ближайшая к (x^k, λ^k, μ^k) стационарная точка задачи (2).

Тогда вся траектория $\{(x^k, \lambda^k, \mu^k)\}$ сходится к $(\bar{x}, \bar{\lambda}, \bar{\mu})$ сверхлинейно.

Доказательство. Согласно следствию 1, найдутся окрестность \mathcal{U}_1 точки $(\bar{x}, \bar{\lambda}, \bar{\mu})$ и функция $\chi: \mathbb{R}_+ \rightarrow \mathbb{R}_+$ такие, что $\chi(t) = o(t)$, и для любой начальной точки из \mathcal{U}_1 метод sSQP определяет траекторию, вдоль которой выполняется неравенство (5). При этом из равенства $\sigma(\bar{x}, \bar{\lambda}, \bar{\mu}) = 0$ и непрерывности функции σ следует, что найдется такая окрестность \mathcal{U}_2 точки $(\bar{x}, \bar{\lambda}, \bar{\mu})$, что для произвольной точки $(x, \lambda, \mu) \in \mathcal{U}_2$ справедливо неравенство

$$\chi(\sigma(x, \lambda, \mu)) \leq q\sigma(x, \lambda, \mu). \quad (8)$$

Поскольку $(\bar{x}, \bar{\lambda}, \bar{\mu})$ является предельной точкой последовательности $\{(x^k, \lambda^k, \mu^k)\}$, то найдется номер k , для которого $(x^k, \lambda^k, \mu^k) \in \mathcal{U}_1 \cap \mathcal{U}_2$. Тогда метод sSQP, запущенный из точки (x^k, λ^k, μ^k) как из начальной точки, определяет траекторию, причем из оценок (5) и (8) следует, что вдоль этой траектории будет выполняться условие (3). Поэтому все итерации алгоритма 1, начиная с номера k , будут итерациями sSQP, и сверхлинейная сходимость алгоритма 1 следует из теоремы 1.2. ■

Очевидный (по крайней мере, теоретический) недостаток алгоритма с возвратами заключается в том, что ранние попытки переключиться на sSQP могут приводить к напрасным вычислительным затратам на итерации sSQP, которые впоследствии отбрасываются.

Альтернативная стратегия гибридной глобализации сходимости состоит в следующем. Вместо того, чтобы сравнивать невязку системы ККТ в пробной точке с невязкой, полученной на предыдущей итерации, можно 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.

Алгоритм 2. Фиксируем параметр $q \in (0, 1)$. Выбираем $(x^0, \lambda^0, \mu^0) \in \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m$ и полагаем $k = 0$. Вычисляем $f_{\text{rec}} = \sigma(x^0, \lambda^0, \mu^0)$.

1. Осуществляем шаг sSQP, т.е. вычисляем $(x^{k+1}, \lambda^{k+1}, \mu^{k+1}) \in \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m$ как стационарную точку задачи (2) при $\sigma_k = \sigma(x^k, \lambda^k, \mu^k)$. Если задача (2) не имеет стационарных точек, переходим к п. 3.

2. Если

$$\sigma(x^{k+1}, \lambda^{k+1}, \mu^{k+1}) \leq qf_{\text{rec}}, \quad (9)$$

то полагаем $f_{\text{rec}} = \sigma(x^{k+1}, \lambda^{k+1}, \mu^{k+1})$, увеличиваем номер шага k на 1 и переходим к п. 1.

3. Вычисляем $(x^{k+1}, \lambda^{k+1}, \mu^{k+1}) \in \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m$ с помощью шага метода внешней фазы. Если

$$\sigma(x^{k+1}, \lambda^{k+1}, \mu^{k+1}) \leq f_{\text{rec}}, \quad (10)$$

то полагаем $f_{\text{rec}} = \sigma(x^{k+1}, \lambda^{k+1}, \mu^{k+1})$. Увеличиваем номер шага k на 1 и переходим к п. 1.

Как и в случае алгоритма с возвратами, здесь возможны лишь два сценария: либо все итерации алгоритма 2, начиная с некоторого номера k , сгенерированы с помощью метода внешней фазы, либо бесконечное число итераций алгоритма 2 являются итерациями sSQP. В первом случае свойства глобальной сходимости алгоритма 2 будут теми же, что и у метода внешней фазы. Во втором случае из условия (9) легко выводится следующий результат.

Теорема 3. Пусть функция $f: \mathbb{R}^n \rightarrow \mathbb{R}$ и отображения $h: \mathbb{R}^n \rightarrow \mathbb{R}^l$ и $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$ дважды дифференцируемы на \mathbb{R}^n . Пусть последовательность $\{(x^k, \lambda^k, \mu^k)\}$ сгенерирована алгоритмом 2, и пусть члены $(x^{k_j}, \lambda^{k_j}, \mu^{k_j})$ этой последовательности получены шагами sSQP для бесконечного числа номеров j .

Тогда

$$\sigma(x^{k_j}, \lambda^{k_j}, \mu^{k_j}) \rightarrow 0 \text{ при } j \rightarrow \infty,$$

и, в частности, если $(\bar{x}, \bar{\lambda}, \bar{\mu}) \in \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m$ является предельной точкой последовательности $\{(x^{k_j}, \lambda^{k_j}, \mu^{k_j})\}$, то \bar{x} — стационарная точка задачи (1), а $(\bar{\lambda}, \bar{\mu})$ — отвечающий ей множитель Лагранжа.

Заметим, что свойства глобальной сходимости алгоритма 2, установленные в теореме 3, несколько слабее чем соответствующие свойства алгоритма 1, установленные в теореме 1. Тем не менее, и результат теоремы 3 вполне разумен: он говорит о том, что в случае бесконечного числа принятых шагов sSQP запуск алгоритма 2 будет успешным, если критерий остановки состоит в достаточной малости невязки системы ККТ.

Сверхлинейная скорость сходимости алгоритма 2 устанавливается в несколько более сильных предположениях, чем для алгоритма 1.

Теорема 4. Пусть в предположениях следствия 1 траектория $\{(x^k, \lambda^k, \mu^k)\}$, сгенерированная алгоритмом 2, в котором на шаге sSQP в качестве $(x^{k+1}, \lambda^{k+1}, \mu^{k+1})$ берется ближайшая к (x^k, λ^k, μ^k) стационарная точка задачи (2), сходится к точке $(\bar{x}, \bar{\lambda}, \bar{\mu})$.

Тогда скорость сходимости сверхлинейная.

Доказательство. Если рекордное значение $f_{\text{гес}}$ меняется бесконечное число раз, то можно указать точку (x^k, λ^k, μ^k) , сколь угодно близкую к $(\bar{x}, \bar{\lambda}, \bar{\mu})$, и для которой выполнено $f_{\text{гес}} = \sigma(x^k, \lambda^k, \mu^k)$. Тогда требуемый результат получается аналогично доказательству теоремы 2.

Теперь допустим, что, начиная с некоторой итерации, рекордное значение $f_{\text{гес}} > 0$ не меняется (в частности, начиная с этой итерации, шаги sSQP не принимаются алгоритмом 1). Но $\sigma(x^k, \lambda^k, \mu^k) \rightarrow 0$ при $k \rightarrow \infty$, поскольку $\{(x^k, \lambda^k, \mu^k)\}$ сходится к точке $(\bar{x}, \bar{\lambda}, \bar{\mu})$, функция σ непрерывна в этой точке, и $\sigma(\bar{x}, \bar{\lambda}, \bar{\mu}) = 0$. Поэтому поэтому для достаточно большого k условие (9) должно выполняться, а значит, рекордное значение $f_{\text{гес}}$ должно измениться, что дает противоречие. ■

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 близко к решению. Основной причиной этого является эффект притяжения ньютоновских методов условной оптимизации к критическим множителям Лагранжа.

3.2. Глобализация сходимости с помощью модифицированных функций Лагранжа

В данном разделе предлагается подход к глобализации сходимости sSQP, основанный на комбинировании с методами модифицированных функций Лагранжа. Основная идея состоит в том, чтобы использовать направления sSQP для приближенного решения подзадач МФЛ. Отметим, что такой способ комбинирования методов sSQP и МФЛ существенно отличается от рассмотренного в разделе 2.2, где предлагалось переключаться на sSQP лишь на финальной фазе работы МФЛ. Заметим также, что общая идея комбинирования стабилизированных ньютоновских методов с методами модифицированных функций Лагранжа далеко не новая; она восходит по крайней мере к [57] (см. также [4, с. 249]).

В разделе 3.2.1 предлагается алгоритм, комбинирующий методы sSQP и МФЛ, и для него доказываются результаты о глобальной сходимости и сверхлинейной скорости сходимости. В разделе 3.2.2 показано, что данный алгоритм тесно связан с прямо-двойственным алгоритмом SQP из [56], но в то же время обладает рядом преимуществ. Наконец, 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 из [56]), однако проигрывает по эффективности обычному SQP.

3.2.1. Алгоритм и его теоретические свойства

Будем рассматривать задачу оптимизации со смешанными ограничениями

$$f(x) \rightarrow \min, \quad h(x) = 0, \quad g(x) \leq 0, \quad (11)$$

где $f: \mathbb{R}^n \rightarrow \mathbb{R}$ — дважды дифференцируемая функция, а $h: \mathbb{R}^n \rightarrow \mathbb{R}^l$ и $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$ — дважды дифференцируемые отображения.

Функция Лагранжа $L: \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m \rightarrow \mathbb{R}$ для задачи (11) имеет вид

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

Определим также семейство модифицированных функций Лагранжа задачи (11): $L_\sigma: \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m \rightarrow \mathbb{R}$,

$$L_\sigma(x, \lambda, \mu) = f(x) + \frac{\sigma}{2} \left(\left\| \lambda + \frac{1}{\sigma} h(x) \right\|^2 + \left\| \max \left\{ 0, \mu + \frac{1}{\sigma} g(x) \right\} \right\|^2 \right), \quad (12)$$

где $\sigma > 0$ — обратный параметр штрафа. Введем функцию естественной невязки $\sigma: \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m \rightarrow \mathbb{R}_+$ системы ККТ задачи (11):

$$\sigma(x, \lambda, \mu) = \left\| \left(\frac{\partial L}{\partial x}(x, \lambda, \mu), h(x), \min\{\mu, -g(x)\} \right) \right\|.$$

Сформулируем алгоритм, комбинирующий sSQP и МФЛ.

Алгоритм 3. Фиксируем $r_0, \varepsilon_0, \sigma_0, \gamma > 0$ и $q, \theta, \tau, \varepsilon, \kappa, \delta \in (0, 1)$. Фиксируем $\bar{\lambda}_{\min}, \bar{\lambda}_{\max} \in \mathbb{R}^l$, $\bar{\lambda}_{\min} < \bar{\lambda}_{\max}$, $\bar{\mu}_{\max} \in \mathbb{R}_+^m$. Выбираем $(x^0, \lambda^0, \mu^0) \in \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m$, $\hat{x}^0 \in \mathbb{R}^n$, $\bar{\lambda}^0 \in [\bar{\lambda}_{\min}, \bar{\lambda}_{\max}]$ и $\bar{\mu}^0 \in [0, \bar{\mu}_{\max}]$. Полагаем $\hat{x}^0 = x^0$, $k = 0$ и $j = 0$.

1. Выбираем симметричную матрицу $H_j \in \mathbb{R}^{n \times n}$ таким образом, чтобы система

$$\frac{\partial L}{\partial x}(\hat{x}^j, \bar{\lambda}^k, \bar{\mu}^k) + H_j \xi + (h'(\hat{x}^j))^T \eta + (g'(\hat{x}^j))^T \zeta = 0, \quad (13)$$

$$h(\hat{x}^j) + h'(\hat{x}^j) \xi - \sigma_k \eta = 0, \quad (14)$$

$$\bar{\mu}^k + \zeta \geq 0, \quad g(\hat{x}^j) + g'(\hat{x}^j) \xi - \sigma_k \zeta \leq 0, \quad \langle \bar{\mu}^k + \zeta, g(\hat{x}^j) + g'(\hat{x}^j) \xi - \sigma_k \zeta \rangle = 0. \quad (15)$$

имела вычислимое решение $(\xi^j, \eta^j, \zeta^j) \in \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m$, удовлетворяющее неравенству

$$\left\langle \frac{\partial L_{\sigma_k}}{\partial x}(\hat{x}^j, \bar{\lambda}^k, \bar{\mu}^k), \xi^j \right\rangle \leq -\gamma \|\xi^j\|^2. \quad (16)$$

2. Если $\bar{\lambda}^k + \eta^j \in [\bar{\lambda}_{\min}, \bar{\lambda}_{\max}]$, $\bar{\mu}^k + \zeta^j \in [0, \bar{\mu}_{\max}]$ и

$$\sigma(\hat{x}^j + \xi^j, \bar{\lambda}^k + \eta^j, \bar{\mu}^k + \zeta^j) \leq r_k, \quad (17)$$

полагаем

$$x^{k+1} = \hat{x}^j + \xi^j, \quad \lambda^{k+1} = \bar{\lambda}^{k+1} = \bar{\lambda}^k + \eta^j, \quad \mu^{k+1} = \bar{\mu}^{k+1} = \bar{\mu}^k + \zeta^j, \quad (18)$$

$$\sigma_{k+1} = \sigma(x^{k+1}, \lambda^{k+1}, \mu^{k+1}), \quad \varepsilon_{k+1} = \varepsilon_k, \quad r_{k+1} = qr_k. \quad (19)$$

Полагаем $j = 0$, $\hat{x}^0 = x^{k+1}$, увеличиваем k на 1 и переходим к шагу 1 (*итерация sSQP*).

3. Полагаем $\hat{x}^{j+1} = \hat{x}^j + \tau^i \xi^j$, где i — наименьшее неотрицательное целое число, для которого справедливо

$$L_{\sigma_k}(\hat{x}^j + \tau^i \xi^j, \bar{\lambda}^k, \bar{\mu}^k) \leq L_{\sigma_k}(\hat{x}^j, \bar{\lambda}^k, \bar{\mu}^k) + \varepsilon \tau^i \left\langle \frac{\partial L_{\sigma_k}}{\partial x}(\hat{x}^j, \bar{\lambda}^k, \bar{\mu}^k), \xi^j \right\rangle. \quad (20)$$

4. Если

$$\left\| \frac{\partial L_{\sigma_k}}{\partial x}(\hat{x}^{j+1}, \bar{\lambda}^k, \bar{\mu}^k) \right\| > \varepsilon_k, \quad (21)$$

увеличиваем j на 1 и переходим к шагу 1 (*внутренняя итерация*).

5. Полагаем $\varepsilon_{k+1} = \theta \varepsilon_k$, $x^{k+1} = \hat{x}^{j+1}$ и

$$\lambda^{k+1} = \bar{\lambda}^k + \frac{1}{\sigma_k} h(\hat{x}^{j+1}), \quad \mu^{k+1} = \max \left\{ 0, \bar{\mu}^k + \frac{1}{\sigma_k} g(\hat{x}^{j+1}) \right\}.$$

Если $\lambda^{k+1} \in [\bar{\lambda}_{\min}, \bar{\lambda}_{\max}]$ и $\mu^{k+1} \in [0, \bar{\mu}_{\max}]$, полагаем $\bar{\lambda}^{k+1} = \lambda^{k+1}$, $\bar{\mu}^{k+1} = \mu^{k+1}$; в противном случае выбираем $\bar{\lambda}^{k+1} \in [\bar{\lambda}_{\min}, \bar{\lambda}_{\max}]$ и $\bar{\mu}^{k+1} \in [0, \bar{\mu}_{\max}]$. Если выполнено

$$\sigma(x^{k+1}, \lambda^{k+1}, \mu^{k+1}) \leq r_k, \quad (22)$$

полагаем $\sigma_{k+1} = \sigma(x^{k+1}, \lambda^{k+1}, \mu^{k+1})$ и $r_{k+1} = qr_k$. В противном случае полагаем

$$\sigma_{k+1} = \begin{cases} \sigma_k, & \text{если } \psi(x^{k+1}, \mu^{k+1}) \leq \delta \psi(x^k, \mu^k), \\ \kappa \sigma_k, & \text{иначе,} \end{cases} \quad (23)$$

и $r_{k+1} = r_k$. Полагаем $j = 0$, $\hat{x}^0 = x^{k+1}$, увеличиваем k на 1 и переходим к шагу 1 (*итерация МФЛ*).

В приведенном алгоритме индекс k используется для нумерации объектов, которые могут меняться только на внешних итерациях (т.е. итерациях sSQP или МФЛ), но остаются неизменными для любой последовательности внутренних итераций.

Неравенство (16) на шаге 1 алгоритма 3 гарантирует, что условие Армихо (20) на шаге 3 выполняется для некоторого неотрицательного целого i .

Для того, чтобы удовлетворить требованиям шага 1 алгоритма 3, матрицу H_j достаточно выбрать так, чтобы выполнялось

$$\left\langle \left(H_j + \frac{1}{\sigma_k} (h'(\hat{x}^j))^T h'(\hat{x}^j) \right) \xi, \xi \right\rangle \geq \gamma \|\xi\|^2 \quad \forall \xi \in \mathbb{R}^n. \quad (24)$$

Действительно, система (13)–(15) эквивалентна системе Каруша–Куна–Таккера подзадачи метода sSQP, а именно, задачи квадратичного программирования

$$\begin{aligned} \langle f'(\hat{x}^j), \xi \rangle + \frac{1}{2} \langle H_j \xi, \xi \rangle + \frac{\sigma_k}{2} (\|\bar{\lambda}^k + \eta\|^2 + \|\bar{\mu}^k + \zeta\|^2) &\rightarrow \min, \\ h(\hat{x}^j) + h'(\hat{x}^j)\xi - \sigma_k \eta = 0, \quad g(\hat{x}^j) + g'(\hat{x}^j)\xi - \sigma_k \zeta &\leq 0 \end{aligned} \quad (25)$$

относительно переменных $(\xi, \eta, \zeta) \in \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m$ (поскольку, как легко убедиться, из системы ККТ задачи (25) вытекает, что прямые переменные η и ζ равны соответствующим множителям Лагранжа для ограничений-равенств и ограничений-неравенств из (25)). Заметим, что для всякой такой ненулевой тройки, удовлетворяющей равенству

$$h'(\hat{x}^j)\xi - \sigma_k \eta = 0, \quad (26)$$

для квадратичной формы в целевой функции задачи (25) имеет место

$$\begin{aligned} \langle H_j \xi, \xi \rangle + \sigma_k (\|\eta\|^2 + \|\zeta\|^2) &= \langle H_j \xi, \xi \rangle + \sigma_k \left(\left\| \frac{1}{\sigma_k} h'(\hat{x}^j)\xi \right\|^2 + \|\zeta\|^2 \right) = \\ &= \left\langle \left(H_j + \frac{1}{\sigma_k} (h'(\hat{x}^j))^T h'(\hat{x}^j) \right) \xi, \xi \right\rangle + \|\zeta\|^2 > 0, \end{aligned}$$

откуда следует сильная выпуклость целевой функции задачи (25) на линейном подпространстве, задаваемом равенством (26). Отсюда, в свою очередь, следует сильная выпуклость этой функции на допустимом множестве задачи (25), поскольку, как легко проверить, для квадратичных функций из сильной выпуклости на линейном подпространстве следует сильная выпуклость на любом параллельном ему аффинном многообразии. Наконец, допустимое множество задачи (25) непусто при любом $\sigma_k > 0$, и поэтому, согласно сказанному выше, задача (25) всегда имеет единственное решение, которому, в силу линейности ограничений этой задачи, соответствует решение системы (13)–(15). Кроме того, в силу выпуклости задачи (25), ее условия ККТ являются достаточными для оптимальности, и поэтому, с учетом сказанного выше, решение системы (13)–(15) единственно (и совпадает с решением задачи (25)). Следующая лемма устанавливает, что для этого решения выполняется (16).

Лемма 1. Пусть $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $h: \mathbb{R}^n \rightarrow \mathbb{R}^l$ и $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$ дифференцируемы в точке $\hat{x}^j \in \mathbb{R}^n$. Пусть (ξ^j, η^j, ζ^j) – решение системы (13)–(15) при некоторых $\bar{\lambda}^k \in \mathbb{R}^l$, $\bar{\mu}^k \in \mathbb{R}^m$, симметричной матрице $H_j \in \mathbb{R}^{n \times n}$ и $\sigma_k > 0$, удовлетворяющих (24) при некотором $\gamma > 0$.

Тогда выполняется (16), и, в частности, условие Армихо (20) на шаге 3 алгоритма 3 выполняется для некоторого неотрицательного целого i .

Доказательство. Из (12) для любых $(x, \lambda, \mu) \in \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m$ и $\sigma > 0$ имеем

$$\begin{aligned} \frac{\partial L_\sigma}{\partial x}(x, \lambda, \mu) &= f'(x) + (h'(x))^T \left(\lambda + \frac{1}{\sigma} h(x) \right) + (g'(x))^T \max \left\{ 0, \mu + \frac{1}{\sigma} g(x) \right\} = \\ &= \frac{\partial L}{\partial x}(x, \lambda, \mu) + \frac{1}{\sigma} (h'(x))^T h(x) + (g'(x))^T \max \left\{ -\mu, \frac{1}{\sigma} g(x) \right\}, \end{aligned} \quad (27)$$

и поэтому для всякого решения (ξ^j, η^j, ζ^j) системы (13)–(15) справедливо

$$\begin{aligned} \left\langle \frac{\partial L_{\sigma_k}}{\partial x}(\hat{x}^j, \bar{\lambda}^k, \bar{\mu}^k), \xi^j \right\rangle &= \left\langle \frac{\partial L}{\partial x}(\hat{x}^j, \bar{\lambda}^k, \bar{\mu}^k), \xi^j \right\rangle + \frac{1}{\sigma_k} \langle h(\hat{x}^j), h'(\hat{x}^j) \xi^j \rangle + \\ &+ \left\langle \max \left\{ -\bar{\mu}^k, \frac{1}{\sigma_k} g(\hat{x}^j) \right\}, g'(\hat{x}^j) \xi^j \right\rangle = -\langle H_j \xi^j, \xi^j \rangle - \langle \eta^j, h'(\hat{x}^j) \xi^j \rangle - \\ &- \langle \zeta^j, g'(\hat{x}^j) \xi^j \rangle + \frac{1}{\sigma_k} \langle h(\hat{x}^j), h'(\hat{x}^j) \xi^j \rangle + \left\langle \max \left\{ -\bar{\mu}^k, \frac{1}{\sigma_k} g(\hat{x}^j) \right\}, g'(\hat{x}^j) \xi^j \right\rangle = \\ &= -\left\langle \left(H_j + \frac{1}{\sigma_k} (h'(\hat{x}^j))^T h'(\hat{x}^j) \right) \xi^j, \xi^j \right\rangle - \left\langle \zeta^j - \max \left\{ -\bar{\mu}^k, \frac{1}{\sigma_k} g(\hat{x}^j) \right\}, g'(\hat{x}^j) \xi^j \right\rangle. \end{aligned}$$

Полагая

$$\begin{aligned} R_i &= \left(\zeta_i^j - \max \left\{ -\bar{\mu}_i^k, \frac{1}{\sigma_k} g_i(\hat{x}^j) \right\} \right) \langle g'_i(\hat{x}^j), \xi^j \rangle = \\ &= \min \left\{ \bar{\mu}_i^k + \zeta_i^j, \zeta_i^j - \frac{1}{\sigma_k} g_i(\hat{x}^j) \right\} \langle g'_i(\hat{x}^j), \xi^j \rangle, \quad i = 1, \dots, m, \end{aligned}$$

отсюда и из (24) получаем

$$\left\langle \frac{\partial L_{\sigma_k}}{\partial x}(\hat{x}^j, \bar{\lambda}^k, \bar{\mu}^k), \xi^j \right\rangle \leq -\gamma \|\xi^j\|^2 - \sum_{i=1}^m R_i. \quad (28)$$

Для всякого $i = 1, \dots, m$, если $\bar{\mu}_i^k \leq -g_i(\hat{x}^j)/\sigma_k$, то

$$R_i = (\bar{\mu}_i^k + \zeta_i^j) \langle g'_i(\hat{x}^j), \xi^j \rangle.$$

Если $\bar{\mu}_i^k + \zeta_i^j = 0$, то $R_i = 0$. Кроме того, согласно (15), в оставшемся случае, когда $\bar{\mu}_i^k + \zeta_i^j > 0$, имеет место равенство

$$g_i(\hat{x}^j) + \langle g'_i(\hat{x}^j), \xi^j \rangle - \sigma_k \zeta_i^j = 0, \quad (29)$$

и поэтому

$$\langle g'_i(\hat{x}^j), \xi^j \rangle = -g_i(\hat{x}^j) + \sigma_k \zeta_i^j = -\sigma_k \left(\bar{\mu}_i^k + \frac{1}{\sigma_k} g_i(\hat{x}^j) \right) + \sigma_k (\bar{\mu}_i^k + \zeta_i^j) > 0,$$

т. е. в этом случае $R_i > 0$.

Если же $\bar{\mu}_i^k > -g_i(\hat{x}^j)/\sigma_k$, то

$$R_i = \left(\zeta_i^j - \frac{1}{\sigma_k} g_i(\hat{x}^j) \right) \langle g'_i(\hat{x}^j), \xi^j \rangle.$$

Если $\bar{\mu}_i^k + \zeta_i^j = 0$, то

$$\zeta_i^j - \frac{1}{\sigma_k} g_i(\hat{x}^j) = - \left(\bar{\mu}_i^k + \frac{1}{\sigma_k} g_i(\hat{x}^j) \right) < 0,$$

откуда и из (15) вытекает, что

$$\langle g'_i(\hat{x}^j), \xi^j \rangle \leq \sigma_k \left(\zeta_i^j - \frac{1}{\sigma_k} g_i(\hat{x}^j) \right) < 0,$$

и поэтому $R_i > 0$. Кроме того, согласно (15), в оставшемся случае, когда $\bar{\mu}_i^k + \zeta_i^j > 0$, опять выполнено (29), и поэтому

$$R_i = \frac{1}{\sigma_k} (\langle g'_i(\hat{x}^j), \xi^j \rangle)^2 \geq 0.$$

Таким образом, в любом случае $R_i \geq 0$ для всех $i = 1, \dots, m$, откуда и из (28) вытекает (16). ■

Свойство (24) является автоматическим, если сама матрица H_j достаточно положительно определена, чего всегда можно добиться, например, прибавляя βI к имеющейся матрице при достаточно большом $\beta > 0$. Вместе с тем, с точки зрения высокой скорости сходимости (см. теорему 6 ниже), матрицы H_0 на поздних внешних итерациях должны совпадать с $\frac{\partial^2 L}{\partial x^2}(x^k, \bar{\lambda}^k, \bar{\mu}^k)$, что может не согласовываться с требованием положительной определенности H_0 . В будущем здесь может помочь использование квазиньютоновских аппроксимаций $\frac{\partial^2 L}{\partial x^2}(x^k, \bar{\lambda}^k, \bar{\mu}^k)$, но существующая на данный момент теория локальной сходимости квазиньютоновских методов sSQP [46] также предполагает фактическую (по норме) близость этих аппроксимаций к $\frac{\partial^2 L}{\partial x^2}(x^k, \bar{\lambda}^k, \bar{\mu}^k)$, что может быть препятствием для их положительной определенности.

При отсутствии ограничений-неравенств требование (24) не слишком обременительно, поскольку оно выполняется автоматически для всех достаточно больших k при $j = 0$ и $H_0 = \frac{\partial^2 L}{\partial x^2}(x^k, \bar{\lambda}^k)$, если $\sigma_k \rightarrow 0$ при $k \rightarrow \infty$, и последовательность $\{(x^k, \bar{\lambda}^k)\}$ сходится к $(\bar{x}, \bar{\lambda})$, где $\bar{x} \in \mathbb{R}^n$ — стационарная точка задачи

$$f(x) \rightarrow \min, \quad h(x) = 0, \tag{30}$$

а $\bar{\lambda} \in \mathbb{R}^l$ — отвечающий ей множитель Лагранжа, для которых выполняется достаточное условие второго порядка оптимальности

$$\left\langle \frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda}) \xi, \xi \right\rangle \geq \hat{\gamma} \|\xi\|^2 \quad \forall \xi \in \ker h'(\bar{x})$$

при некотором $\hat{\gamma} > \gamma$ (требуемое свойство легко получить с помощью леммы Финслера–Дебре; см., например, [4, лемма 1.25]).

Более того, (14) эквивалентно равенству

$$\eta = \frac{1}{\sigma_k} (h(\hat{x}^j) + h'(\hat{x}^j) \xi). \tag{31}$$

Подставляя данное выражение в (13), при отсутствии ограничений-неравенств приходим к уравнению

$$\left(H_j + \frac{1}{\sigma_k} (h'(\hat{x}^j))^T h'(\hat{x}^j) \right) \xi = -\frac{\partial L_{\sigma_k}}{\partial x}(\hat{x}^j, \bar{\lambda}^k). \quad (32)$$

В случае невырожденности матрицы в левой части, соотношения (31), (32) корректно определяют единственное решение (ξ^j, η^j) итерационной системы. Достаточную положительную определенность этой матрицы (т.е. выполнение (24)) можно обеспечить в процессе ее факторизации по Холецкому [88, раздел 3.4], т.е. одновременно с отысканием ξ^j как решения линейной системы (32) (с последующим определением η^j по явной формуле (31)).

Вместе с тем, можно предложить и другие способы реализации шага 1 алгоритма 3, не связанные с обеспечением выполнения (24). А именно, для всякого j пробуем решить систему (13)–(15) для некоторой начальной матрицы H_j (например, для $H_j = \frac{\partial^2 L}{\partial x^2}(\hat{x}^j, \bar{\lambda}^k, \bar{\mu}^k)$). Если решение системы не может быть найдено, или если найденное решение не удовлетворяет (16), выбираем другую матрицу H_j (например, заменяем H_j на $H_j + \beta I$ при некотором фиксированном $\beta > 0$), и повторяем процедуру.

Следующее утверждение показывает, что алгоритм 3 либо генерирует бесконечную последовательность $\{(x^k, \lambda^k, \mu^k)\}$, либо для некоторого k генерирует расходящуюся на бесконечность последовательность внутренних итераций $\{\hat{x}^j\}$ (последний случай можно трактовать как неудачный запуск алгоритма).

Предложение 1. Пусть $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $h: \mathbb{R}^n \rightarrow \mathbb{R}^l$ и $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$ непрерывно дифференцируемы на \mathbb{R}^n . Пусть для некоторого номера k шаги (1)–(4) алгоритма 3 генерируют бесконечную последовательность $\{\hat{x}^j\}$, причем последовательность $\{H_j\}$ ограничена.

Тогда $\|\hat{x}^j\| \rightarrow \infty$ при $j \rightarrow \infty$.

Доказательство. Последовательность $\{\hat{x}^j\}$ генерируется при фиксированных значениях $\bar{\lambda}^k$, $\bar{\mu}^k$ и σ_k , и поэтому одномерный поиск на шаге 3 выполняется для фиксированной непрерывно дифференцируемой функции $L_{\sigma_k}(\cdot, \bar{\lambda}^k, \bar{\mu}^k)$. Покажем, что для этой функции последовательность $\{\xi^j\}$ является равномерно градиентной относительно $\{\hat{x}^j\}$ (в терминологии [4]).

Рассмотрим произвольное бесконечное множество индексов $J \subset \{0, 1, \dots\}$, такое что подпоследовательность $\{\hat{x}^j \mid j \in J\}$ сходится к некоторой точке \hat{x} , удовлетворяющей неравенству

$$\frac{\partial L_{\sigma_k}}{\partial x}(\hat{x}, \bar{\lambda}^k, \bar{\mu}^k) \neq 0. \quad (33)$$

Отсюда и из соотношения (16) следует ограниченность подпоследовательности $\{\xi^j \mid j \in J\}$.

Предположим, что последовательность $\{\xi^j \mid j \in J\}$ имеет нулевую предельную точку. Без ограничения общности можем считать, что вся эта последовательность сходится к нулю.

При этом, согласно (14), для любого $j \in J$ справедливо

$$\eta^j = \frac{1}{\sigma_k}(h(\hat{x}^j) + h'(\hat{x}^j)\xi^j),$$

и поэтому $\{\eta^j \mid j \in J\} \rightarrow h(\hat{x})/\sigma_k$. С другой стороны, в силу (15) и конечности числа различных подмножеств множества $\{1, \dots, m\}$ можно считать, что существует множество $I \subset \{1, \dots, m\}$ такое, что для любого $j \in J$

$$\begin{aligned} \bar{\mu}_i^k + \zeta_i^j &> 0, & g_i(\hat{x}^j) + \langle g'_i(\hat{x}^j), \xi^j \rangle - \sigma_k \zeta_i^j &= 0 \quad \forall i \in I, \\ \bar{\mu}_i^k + \zeta_i^j &= 0, & g_i(\hat{x}^j) + \langle g'_i(\hat{x}^j), \xi^j \rangle - \sigma_k \zeta_i^j &\leq 0 \quad \forall i \in \{1, \dots, m\} \setminus I. \end{aligned}$$

В частности,

$$\begin{aligned} -\bar{\mu}_i^k &< \zeta_i^j = \frac{1}{\sigma_k}(g_i(\hat{x}^j) + \langle g'_i(\hat{x}^j), \xi^j \rangle) \quad \forall i \in I, \\ -\bar{\mu}_i^k &= \zeta_i^j \geq \frac{1}{\sigma_k}(g_i(\hat{x}^j) + \langle g'_i(\hat{x}^j), \xi^j \rangle) \quad \forall i \in \{1, \dots, m\} \setminus I, \end{aligned}$$

откуда следует, что $\{\zeta_i^j \mid j \in J\} \rightarrow g_i(\hat{x})/\sigma_k \geq -\bar{\mu}_i^k$ для всех $i \in I$, $\zeta_i^j = -\bar{\mu}_i^k \geq g_i(\hat{x})/\sigma_k$ для всех $i \in \{1, \dots, m\} \setminus I$, а значит, объединяя эти два случая, $\{\zeta^j \mid j \in J\} \rightarrow \max\{-\bar{\mu}^k, g(\hat{x})/\sigma_k\}$. В силу полученных предельных соотношений, переходя к пределу в (13) и используя (27), имеем

$$0 = \frac{\partial L}{\partial x}(\hat{x}, \bar{\lambda}^k, \bar{\mu}^k) + \frac{1}{\sigma_k}(h'(\hat{x}))^T h(\hat{x}) + (g'(\hat{x}))^T \max\left\{-\bar{\mu}^k, \frac{1}{\sigma_k}g(\hat{x})\right\} = \frac{\partial L_{\sigma_k}}{\partial x}(\hat{x}, \bar{\lambda}^k, \bar{\mu}^k),$$

что противоречит (33). Таким образом, последовательность $\{\xi^j \mid j \in J\}$ не может иметь нулевую предельную точку, а значит, в силу (16), ее не может иметь и последовательность $\{\langle \frac{\partial L_{\sigma_k}}{\partial x}(\hat{x}^j, \bar{\lambda}^k, \bar{\mu}^k), \xi^j \rangle \mid j \in J\}$.

Полученные свойства как раз и означают, что последовательность $\{\xi^j\}$ является равномерно градиентной. Применяя [4, теорема 1.8], получаем, что для любой предельной точки \hat{x} последовательности $\{\hat{x}^j\}$ выполнено

$$\frac{\partial L_{\sigma_k}}{\partial x}(\hat{x}, \bar{\lambda}^k, \bar{\mu}^k) = 0.$$

Но тогда, с учетом теста (21) на шаге 4 алгоритма 3 последовательность $\{\hat{x}^j\}$ не может иметь предельных точек, что и доказывает предложение. \blacksquare

Следующая теорема устанавливает свойства глобальной сходимости алгоритма 3 в случае, когда он генерирует бесконечную последовательность $\{(x^k, \lambda^k, \mu^k)\}$.

Теорема 5. Пусть $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $h: \mathbb{R}^n \rightarrow \mathbb{R}^l$ и $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$ непрерывно дифференцируемы на \mathbb{R}^n . Пусть последовательность $\{(x^k, \lambda^k, \mu^k)\}$ сгенерирована алгоритмом 3.

Тогда для любой предельной точки $(\bar{x}, \bar{\lambda}, \bar{\mu})$ последовательности $\{(x^k, \lambda^k, \mu^k)\}$ выполняется следующее:

$$\frac{\partial L}{\partial x}(\bar{x}, \bar{\lambda}, \bar{\mu}) = 0, \quad \bar{\mu} \geq 0, \quad (h'(\bar{x}))^T h(\bar{x}) + (g'(\bar{x}))^T \max\{0, g(\bar{x})\} = 0, \quad (34)$$

причем либо \bar{x} — стационарная точка задачи (11), а $(\bar{\lambda}, \bar{\mu})$ — отвечающий ей множитель Лагранжа, либо для любой сходящейся к $(\bar{x}, \bar{\lambda}, \bar{\mu})$ подпоследовательности последовательности $\{(x^k, \lambda^k, \mu^k)\}$ все ее элементы, начиная с некоторого, получены итерациями МФЛ алгоритма 3. При этом, если последовательность $\{\sigma_k\}$, сгенерированная алгоритмом 3, отделена от нуля, то всегда реализуется первый случай.

Доказательство. Рассмотрим произвольное множество индексов $K \subset \{1, 2, \dots\}$, для которого подпоследовательность $\{(x^k, \lambda^k, \mu^k) \mid k \in K\}$ сходится к некоторой точке $(\bar{x}, \bar{\lambda}, \bar{\mu})$. Заметим, что для любого $k \in K$ либо выполнено

$$\sigma(x^k, \lambda^k, \mu^k) \leq r_{k-1}, \quad (35)$$

либо точка (x^k, λ^k, μ^k) получена в результате итерации МФЛ, а величина σ_k вычислена согласно (23). Если для бесконечного набора индексов $k \in K$ выполнено (35), то $r_k \rightarrow 0$, поскольку $r_k = qr_{k-1}$ для всех таких k , и последовательность $\{r_k\}$ является монотонно невозрастающей. Но тогда из (35) вытекает равенство $\sigma(\bar{x}, \bar{\lambda}, \bar{\mu}) = 0$, которое означает, что \bar{x} — стационарная точка задачи (11), а $(\bar{\lambda}, \bar{\mu})$ — отвечающий ей множитель Лагранжа (и, в частности, выполнено (34)).

Предположим теперь, что (35) нарушается для всех $k \in K$, начиная с некоторого номера. Тогда, согласно сказанному выше, для всех достаточно больших $k \in K$ выполнено $\varepsilon_k = \theta\varepsilon_{k-1}$,

$$\left\| \frac{\partial L_{\sigma_{k-1}}}{\partial x}(x^k, \bar{\lambda}^{k-1}, \bar{\mu}^{k-1}) \right\| \leq \varepsilon_{k-1}, \quad (36)$$

и

$$\lambda^k = \bar{\lambda}^{k-1} + \frac{1}{\sigma_{k-1}} h(x^k), \quad \mu^k = \max \left\{ 0, \bar{\mu}^{k-1} + \frac{1}{\sigma_{k-1}} g(x^k) \right\}. \quad (37)$$

Условие неотрицательности $\bar{\mu}$ в (34) следует немедленно из второго равенства в (37). Кроме того, из монотонного невозрастания последовательности $\{\varepsilon_k\}$ следует, что $\varepsilon_k \rightarrow 0$. С другой стороны, из (27) и (37) получаем цепочку равенств

$$\begin{aligned} \frac{\partial L_{\sigma_{k-1}}}{\partial x}(x^k, \bar{\lambda}^{k-1}, \bar{\mu}^{k-1}) &= f'(x^k) + (h'(x^k))^T \left(\bar{\lambda}^{k-1} + \frac{1}{\sigma_{k-1}} h(x^k) \right) + \\ &+ (g'(x^k))^T \max \left\{ 0, \bar{\mu}^{k-1} + \frac{1}{\sigma_{k-1}} g(x^k) \right\} = \frac{\partial L}{\partial x}(x^k, \lambda^k, \mu^k), \end{aligned}$$

откуда и из (36), переходя к пределу вдоль соответствующей подпоследовательности, получаем первое равенство в (34). Если при этом $\psi(\bar{x}, \bar{\mu}) = 0$, то снова имеет место $\sigma(\bar{x}, \bar{\lambda}, \bar{\mu}) = 0$.

Заметим, что для всякого k значение σ_{k+1} может быть больше σ_k только если выполняется (22) (выполнение (17) также влечет выполнение (22) для соответствующей пары (x^{k+1}, λ^{k+1})), причем в этом случае $\sigma_{k+1} \leq r_k$ и $r_{k+1} = qr_k$. Таким образом, если для бесконечного числа номеров k выполняется (22), то $\sigma_k \rightarrow 0$. Если же (22) выполняется лишь конечное число раз, то, начиная с некоторого номера \hat{k} , все внешние итерации алгоритма являются итерациями МФЛ, последовательность $\{\sigma_k\}$ является монотонно невозрастающей, причем для всех $k \geq \hat{k}$ значение σ_{k+1} выбирается согласно (23). В частности, последовательность $\{\sigma_k\}$ может быть отделена от нуля только если $\psi(x^{k+1}, \mu^{k+1}) \leq \delta\psi(x^k, \mu^k)$ для всех достаточно больших k , откуда следует, что $\psi(x^k, \mu^k) \rightarrow 0$, а значит $\sigma(\bar{x}, \bar{\lambda}, \bar{\mu}) = 0$.

Из сказанного выше вытекает, в частности, что если $\psi(\bar{x}, \bar{\mu}) \neq 0$, то $\sigma_k \rightarrow 0$. Из (27) и (36) следует неравенство

$$\begin{aligned} & \|\sigma_{k-1} (f'(x^k) + (h'(x^k))^T \bar{\lambda}^{k-1} + (g'(x^k))^T \bar{\mu}^{k-1}) + \\ & \quad + (h'(x^k))^T h(x^k) + (g'(x^k))^T \max\{-\sigma_{k-1} \bar{\mu}^{k-1}, g(x^k)\}\| \leq \varepsilon_{k-1} \sigma_{k-1} \end{aligned}$$

для всех $k \in K$. Используя ограниченность $\{(\bar{\lambda}^k, \bar{\mu}^k) \mid k \in K\}$ и переходя к пределу в полученном неравенстве вдоль соответствующей подпоследовательности, получаем второе равенство в (34). Теорема доказана. \blacksquare

Доказанная теорема не исключает того, что в случаях, когда обратный параметр штрафа не отделен от нуля, траектории алгоритма 3 могут иметь удовлетворяющие (34) предельные точки с недопустимой прямой частью. Однако, любая подпоследовательность траектории алгоритма, сходящаяся к такой предельной точке, обладает следующим свойством: все ее элементы с достаточно большими номерами генерируются итерациями МФЛ, т.е. на этих итерациях алгоритм ведет себя как метод модифицированных функций Лагранжа, а точнее, как вариант этого метода, используемый в солвере ALGENCAN [30] (см. алгоритм 2.1 и [31, алгоритм 3.1]). Возможная сходимость к недопустимым точкам является известным теоретическим недостатком МФЛ [4, 31]. Однако, как показано в приложении А, сходимость к недопустимым точкам для ALGENCAN совершенно нетипична. Более того, в плане свойств глобальной сходимости ALGENCAN превосходит многие существующие альтернативы (в том числе и для вырожденных задач).

Скорость сходимости алгоритма 3 устанавливается следующей теоремой. Напомним, что, согласно предложению 1.1, если для отвечающего стационарной точке \bar{x} задачи (11) множителя Лагранжа $(\bar{\lambda}, \bar{\mu})$ выполнено достаточное условие второго порядка оптимальности

$$\left\langle \frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda}, \bar{\mu})\xi, \xi \right\rangle > 0 \quad \forall \xi \in C(\bar{x}) \setminus \{0\}, \quad (38)$$

где

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

то имеет место оценка расстояния до множества решений

$$\|x - \bar{x}\| + \text{dist}((\lambda, \mu), \mathcal{M}(\bar{x})) = O(\sigma(x, \lambda, \mu)) \quad (39)$$

при $(x, \lambda, \mu) \rightarrow (\bar{x}, \bar{\lambda}, \bar{\mu})$.

Теорема 6. Пусть $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $h: \mathbb{R}^n \rightarrow \mathbb{R}^l$ и $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$ дважды дифференцируемы в окрестности точки $\bar{x} \in \mathbb{R}^n$, причем их вторые производные непрерывны в этой точке. Пусть \bar{x} — стационарная точка задачи (11), и пусть для отвечающего ей множителя Лагранжа $(\bar{\lambda}, \bar{\mu}) \in \mathbb{R}^l \times \mathbb{R}^m$ выполняется SOSC (38), причем $\bar{\lambda} \in (\bar{\lambda}_{\min}, \bar{\lambda}_{\max})$, $\bar{\mu} < \bar{\mu}_{\max}$. Пусть траектория $\{(x^k, \lambda^k, \mu^k)\}$ сгенерирована алгоритмом 3, и пусть для некоторого множества индексов $K \subset \{1, 2, \dots\}$ справедливо $\{(x^k, \lambda^k, \mu^k) \mid k \in K\} \rightarrow (\bar{x}, \bar{\lambda}, \bar{\mu})$, причем $r_k < r_{k-1}$ для бесконечного числа индексов $k \in K$. Предположим, что для всех достаточно больших индексов k , если точка $(\hat{x}^0, \bar{\lambda}^k, \bar{\mu}^k)$, используемая на шаге 1 алгоритма 3, достаточно близка к $(\bar{x}, \bar{\lambda}, \bar{\mu})$, то матрица H_0 совпадает с $\frac{\partial^2 L}{\partial x^2}(\hat{x}^0, \bar{\lambda}^k, \bar{\mu}^k)$, а направление (ξ^0, η^0, ζ^0) удовлетворяет условию локализации

$$\|(\xi^0, \eta^0, \zeta^0)\| \leq c \text{dist}((\hat{x}^0, \bar{\lambda}^k, \bar{\mu}^k), \{\bar{x}\} \times \mathcal{M}(\bar{x})) \quad (40)$$

для некоторого $c > 0$, не зависящего от k .

Тогда вся траектория $\{(x^k, \lambda^k, \mu^k)\}$ сходится к $(\bar{x}, \bar{\lambda}, \bar{\mu})$, все итерации, начиная с некоторого номера, являются итерациями sSQP, и скорость сходимости $\{(x^k, \lambda^k, \mu^k)\}$ является сверхлинейной.

Доказательство. По построению алгоритма, для каждого номера k неравенство $r_{k+1} < r_k$ имеет место тогда и только тогда, когда выполнено (22) (в результате итерации sSQP или МФЛ). Поэтому из предположений теоремы следует, что соотношение (22) должно быть выполнено для бесконечного числа индексов k , для которых $k+1 \in K$. Заметим, что для каждого такого индекса также выполнено $\sigma_{k+1} = \sigma(x^{k+1}, \lambda^{k+1}, \mu^{k+1})$ и $r_{k+1} \geq q\sigma(x^{k+1}, \lambda^{k+1}, \mu^{k+1})$.

Обозначим через \mathcal{U} такую окрестность точки $(\bar{x}, \bar{\lambda}, \bar{\mu})$, что если для достаточно большого k выполнено $(\hat{x}^0, \bar{\lambda}^k, \bar{\mu}^k) \in \mathcal{U}$, то матрица H_0 совпадает с $\frac{\partial^2 L}{\partial x^2}(\hat{x}^0, \bar{\lambda}^k, \bar{\mu}^k)$, а направление (ξ^0, η^0, ζ^0) удовлетворяет условию (40) (такая окрестность существует согласно предположениям теоремы). Из сходимости $\{(\lambda^k, \mu^k) \mid k \in K\}$ к $(\bar{\lambda}, \bar{\mu})$ следует, что для всех достаточно больших $k \in K$ выполнено $\lambda^k \in [\bar{\lambda}_{\min}, \bar{\lambda}_{\max}]$, $\mu^k \in [0, \bar{\mu}_{\max}]$. Таким образом, для всех таких k также выполнено $\bar{\lambda}^k = \lambda^k$, $\bar{\mu}^k = \mu^k$, и, в частности, $(\hat{x}^0, \bar{\lambda}^k, \bar{\mu}^k) \in \mathcal{U}$.

Обозначим через $K_0 \subset \{0, 1, \dots\}$ множество всех индексов k , для которых на шаге 1 алгоритма 3 направление (ξ^0, η^0, ζ^0) вычисляется как решение системы (13)–(15) при $\hat{x}^0 = x^k$, $\bar{\lambda}^k = \lambda^k$, $\bar{\mu}^k = \mu^k$, $\sigma_k = \sigma(x^k, \lambda^k, \mu^k)$ и

$$H_0 = \frac{\partial^2 L}{\partial x^2}(x^k, \lambda^k, \mu^k), \quad (41)$$

т. е. оно совпадает с направлением sSQP в точке (x^k, λ^k, μ^k) , и при этом выполнено неравенство $r_k \geq q\sigma(x^k, \lambda^k, \mu^k)$ и условие локализации (40). Как следует из сказанного выше, множество индексов $K \cap K_0$ является бесконечным. В частности, с учетом сходимости $\{(x^k, \lambda^k, \mu^k) \mid k \in K\}$ к $(\bar{x}, \bar{\lambda}, \bar{\mu})$, можно заключить, что для любой окрестности \mathcal{V} точки $(\bar{x}, \bar{\lambda}, \bar{\mu})$ найдется индекс $k \in K_0$, для которого $(x^k, \lambda^k, \mu^k) \in \mathcal{V}$.

Рассмотрим произвольную последовательность $\{(\tilde{x}^k, \tilde{\lambda}^k, \tilde{\mu}^k)\}$, сходящуюся к $(\bar{x}, \bar{\lambda}, \bar{\mu})$, и любую последовательность $\{(\tilde{\xi}^k, \tilde{\eta}^k, \tilde{\zeta}^k)\}$, такую что для всех k направление $(\tilde{\xi}^k, \tilde{\eta}^k, \tilde{\zeta}^k)$ удовлетворяет следующим условиям: оно является решением системы (13)–(15) при $\hat{x}^0 = \tilde{x}^k$, $\bar{\lambda}^k = \tilde{\lambda}^k$, $\bar{\mu}^k = \tilde{\mu}^k$, $\sigma_k = \sigma(\tilde{x}^k, \tilde{\lambda}^k, \tilde{\mu}^k)$, $H_0 = \frac{\partial^2 L}{\partial x^2}(\tilde{x}^k, \tilde{\lambda}^k, \tilde{\mu}^k)$, и выполнено

$$\|(\tilde{\xi}^k, \tilde{\eta}^k, \tilde{\zeta}^k)\| \leq c \operatorname{dist}((\tilde{x}^k, \tilde{\lambda}^k, \tilde{\mu}^k), \{\bar{x}\} \times \mathcal{M}(\bar{x})) \quad (42)$$

для некоторого $c > 0$, не зависящего от k . Применяя теорему 1.2 и оценку расстояния (39), получаем, что для достаточно больших номеров k выполнено $\tilde{\lambda}^k + \tilde{\eta}^k \in [\bar{\lambda}_{\min}, \bar{\lambda}_{\max}]$, $\tilde{\mu}^k + \tilde{\zeta}^k \in [0, \bar{\mu}_{\max}]$, последовательность $\{(\tilde{x}^k + \tilde{\xi}^k, \tilde{\lambda}^k + \tilde{\eta}^k, \tilde{\mu}^k + \tilde{\zeta}^k)\}$ сходится к $(\bar{x}, \bar{\lambda}, \bar{\mu})$, и выполнено

$$\begin{aligned} \sigma(\tilde{x}^k + \tilde{\xi}^k, \tilde{\lambda}^k + \tilde{\eta}^k, \tilde{\mu}^k + \tilde{\zeta}^k) &= O(\|\tilde{x}^k + \tilde{\xi}^k - \bar{x}\| + \operatorname{dist}((\tilde{\lambda}^k + \tilde{\eta}^k, \tilde{\mu}^k + \tilde{\zeta}^k), \mathcal{M}(\bar{x}))) = \\ &= o(\|\tilde{x}^k - \bar{x}\| + \operatorname{dist}((\tilde{\lambda}^k, \tilde{\mu}^k), \mathcal{M}(\bar{x}))) = o(\sigma(\tilde{x}^k, \tilde{\lambda}^k, \tilde{\mu}^k)) \leq q\sigma(\tilde{x}^k, \tilde{\lambda}^k, \tilde{\mu}^k). \end{aligned}$$

Возвращаясь к последовательности $\{(x^k, \lambda^k, \mu^k)\}$, из сказанного выше несложно получить, что если для достаточно большого $k \in K_0$ точка (x^k, λ^k, μ^k) достаточно близка к $(\bar{x}, \bar{\lambda}, \bar{\mu})$, то $\bar{\lambda}^k + \eta^0 \in [\bar{\lambda}_{\min}, \bar{\lambda}_{\max}]$, $\bar{\mu}^k + \zeta^0 \in [0, \bar{\mu}_{\max}]$, $(\hat{x}^0 + \xi^0, \bar{\lambda}^k + \eta^0, \bar{\mu}^k + \zeta^0) \in \mathcal{U}$, и выполнено

$$\sigma(\hat{x}^0 + \xi^0, \bar{\lambda}^k + \eta^0, \bar{\mu}^k + \zeta^0) \leq q\sigma(\hat{x}^0, \bar{\lambda}^k, \bar{\mu}^k) = q\sigma(x^k, \lambda^k, \mu^k) \leq r_k,$$

т. е. тест (17) выполнен при $j = 0$. В частности, для таких k соответствующая итерация является итерацией sSQP, и из выполнения (18), (19) при $j = 0$ вытекает, что $k + 1 \in K_0$.

Таким образом, существует такая окрестность \mathcal{V} точки $(\bar{x}, \bar{\lambda}, \bar{\mu})$, что для любого $k \in K_0$, если $(x^k, \lambda^k, \mu^k) \in \mathcal{V}$, то итерация является итерацией sSQP, и выполнено $k + 1 \in K_0$. Последнее означает, что если $(x^k, \lambda^k, \mu^k) \in \mathcal{V}$ для некоторого $k \in K_0$, то все последующие итерации являются итерациями sSQP до тех пор, пока траектория остается в \mathcal{V} .

Наконец, согласно теореме 1.2, существует окрестность $\mathcal{V}_0 \subset \mathcal{V}$ точки $(\bar{x}, \bar{\lambda}, \bar{\mu})$ со следующими свойствами: если для некоторого k выполнено $(x^k, \lambda^k, \mu^k) \in \mathcal{V}_0$, то любая траектория

$\{(\tilde{x}^k, \tilde{\lambda}^k, \tilde{\mu}^k)\}$, сгенерированная чистым методом sSQP, запущенным из точки (x^k, λ^k, μ^k) , и для всех k удовлетворяющая (42) при $(\tilde{\xi}^k, \tilde{\eta}^k, \tilde{\zeta}^k) = (\tilde{x}^{k+1} - \tilde{x}^k, \tilde{\lambda}^{k+1} - \tilde{\lambda}^k, \tilde{\mu}^{k+1} - \tilde{\mu}^k)$, полностью лежит в \mathcal{V} и сверхлинейно сходится к некоторому решению. Для завершения доказательства остается заметить, что для некоторого $k \in K_0$ имеет место $(x^k, \lambda^k, \mu^k) \in \mathcal{V}_0$. Поэтому, согласно сказанному выше, все итерации, начиная с номера k , являются итерациями sSQP, и вся последующая траектория остается в \mathcal{V} , откуда следует, что последовательность $\{(x^k, \lambda^k, \mu^k)\}$ сверхлинейно сходится к $(\bar{x}, \bar{\lambda}, \bar{\mu})$. ■

Отметим, что если последовательность $\{(x^k, \lambda^k, \mu^k)\}$ сходится к точке $(\bar{x}, \bar{\lambda}, \bar{\mu})$, то неравенство $r_k < r_{k-1}$ будет выполнено для бесконечного числа индексов k . Действительно, последовательность $\{r_k\}$ является монотонно невозрастающей, причем для каждого номера k неравенство $r_{k+1} < r_k$ имеет место тогда и только тогда, когда выполнено (22) (в результате итерации sSQP или МФЛ). Поэтому требуемый факт следует из соотношения $\{\sigma(x^k, \lambda^k, \mu^k)\} \rightarrow 0$.

Подчеркнем, что условие локализации (40) в теореме 5 не является слишком ограничивающим. Согласно теореме 1.2, для произвольного (x^k, λ^k, μ^k) , достаточно близкого к точке $(\bar{x}, \bar{\lambda}, \bar{\mu})$, удовлетворяющей SOS, задача (25), в которой $\hat{x}^0 = x^k, \bar{\lambda}^k = \lambda^k, \bar{\mu}^k = \mu^k, \sigma_k = \sigma(x^k, \lambda^k, \mu^k)$ и H_0 выбирается согласно (41), всегда имеет стационарную точку (ξ^0, η^0, ζ^0) , удовлетворяющую (40) для некоторого фиксированного $c > 0$, не зависящего от (x^k, λ^k, μ^k) . В частности, если H_0 удовлетворяет (24), то условие (40) выполняется автоматически для единственного решения задачи (25).

При отсутствии ограничений-неравенств теорему 6 можно уточнить, заменив SOS (38) на более слабое требование некритичности множителя $\bar{\lambda}$. Для этого при доказательстве нужно воспользоваться теоремой 1.3 вместо теоремы 1.2.

3.2.2. Связь с прямо-двойственным алгоритмом последовательного квадратичного программирования

Алгоритм 3, рассмотренный в предыдущем разделе, имеет много общего с алгоритмом 2.1 из [56], который авторы называют прямо-двойственным алгоритмом последовательного квадратичного программирования (pdSQP). Сначала сравним эти алгоритмы для задачи (30).

Следуя [55, 56], введем семейство прямо-двойственных модифицированных функций Лагранжа: $M_\sigma: \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^l \rightarrow \mathbb{R}$,

$$M_\sigma(x, \lambda; \lambda_E) = L_\sigma(x, \lambda_E) + \frac{\nu\sigma}{2} \left\| \lambda - \left(\lambda_E + \frac{1}{\sigma} h(x) \right) \right\|^2, \quad (43)$$

где λ_E играет роль оценки множителя Лагранжа $\bar{\lambda}$, $\sigma > 0$ — обратный параметр штрафа, а $\nu > 0$ — дополнительный фиксированный параметр. Алгоритм pdSQP состоит в следующем. На каждой итерации с номером k для заданных $x^k \in \mathbb{R}^n, \lambda^k \in \mathbb{R}^l, \lambda_E^k \in \mathbb{R}^l$ и $\sigma_k > 0$

прямо-двойственное направление поиска $d^k = (\xi^k, \eta^k)$ вычисляется как стационарная точка подзадачи

$$\begin{aligned} \langle f'(x^k), \xi \rangle + \frac{1}{2} \langle H_k \xi, \xi \rangle + \frac{\sigma_k}{2} \|\lambda^k + \eta\|^2 &\rightarrow \min, \\ h(x^k) + h'(x^k)\xi - \sigma_k(\lambda^k + \eta - \lambda_E^k) &= 0, \end{aligned} \quad (44)$$

где матрица H_k выбирается таким образом, чтобы матрица $H_k + (h'(x^k))^T h'(x^k) / \sigma_k$ была положительно определенной. Далее вычисляется $(x^{k+1}, \lambda^{k+1}) = (x^k, \lambda^k) + \alpha_k d^k$, где α_k определяется с помощью процедуры гибкого одномерного поиска для семейства функций $M_\sigma(\cdot; \lambda_E^k)$. После этого обновляются λ_E^k и σ_k (а также некоторые другие параметры) и осуществляется переход к следующей итерации.

Процедура обновления λ_E^k устроена следующим образом. Введем функции $\varphi_V, \varphi_O: \mathbb{R}^n \times \mathbb{R}^l \rightarrow \mathbb{R}$,

$$\varphi_V(x, \lambda) = \beta \left\| \frac{\partial L}{\partial x}(x, \lambda) \right\| + \|h(x)\|, \quad \varphi_O(x, \lambda) = \left\| \frac{\partial L}{\partial x}(x, \lambda) \right\| + \beta \|h(x)\|,$$

где $\beta > 0$ — фиксированное число. Если $\varphi_V(x^{k+1}, \lambda^{k+1}) \leq \varphi_V^{\max}/2$ или $\varphi_O(x^{k+1}, \lambda^{k+1}) \leq \varphi_O^{\max}/2$, где φ_V^{\max} и φ_O^{\max} — текущие рекордные значения функций φ_V и φ_O соответственно, то полагается $\lambda_E^{k+1} = \lambda^{k+1}$ (при этом говорят, что осуществляется V- или O-итерация соответственно). В противном случае, если выполнено

$$\left\| \frac{\partial M_{\sigma_k}}{\partial (x, \lambda)}(x^{k+1}, \lambda^{k+1}; \lambda_E^k) \right\| \leq \varepsilon_k, \quad (45)$$

где $\varepsilon_k > 0$ — еще один параметр (роль которого та же, что и в алгоритме 3), то в качестве λ_E^{k+1} выбирается проекция λ^{k+1} на множество $[\bar{\lambda}_{\min}, \bar{\lambda}_{\max}]$ (M-итерация). Если же (45) нарушается, то полагается $\lambda_E^{k+1} = \lambda_E^k$ (F-итерация).

Легко убедиться, что стационарные точки подзадачи (44) описываются линейной системой

$$H_k \xi + (h'(x^k))^T \eta = -\frac{\partial L}{\partial x}(x^k, \lambda^k), \quad h'(x^k)\xi - \sigma_k \eta = -h(x^k) + \sigma_k(\lambda^k - \lambda_E^k).$$

Из второго уравнения системы получаем

$$\eta = \frac{1}{\sigma_k} (h(x^k) + h'(x^k)\xi) - (\lambda^k - \lambda_E^k). \quad (46)$$

Подставляя полученное равенство в первое уравнение системы, находим

$$\left(H_k + \frac{1}{\sigma_k} (h'(x^k))^T h'(x^k) \right) \xi = -\frac{\partial L_{\sigma_k}}{\partial x}(x^k, \lambda_E^k). \quad (47)$$

Поскольку на каждой итерации матрица в левой части (47) положительно определена, из (46), (47) следует, что подзадача (44) имеет единственную стационарную точку. Далее, сравнивая (46), (47) с соотношениями (31), (32), получаем следующую связь между направлениями поиска алгоритма 3 и алгоритма pdSQP.

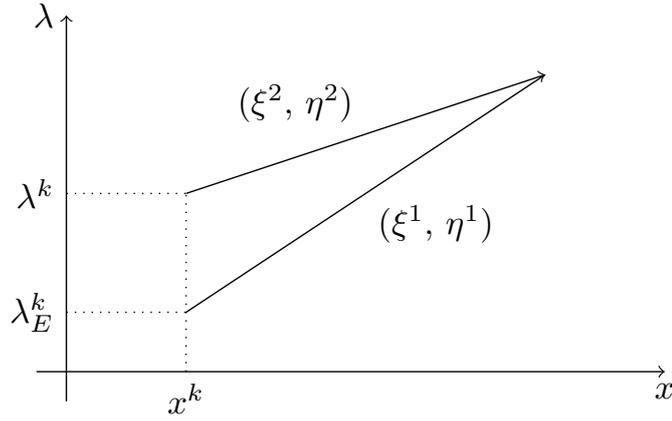


Рис. 1. Иллюстрация к предложению 2.

Предложение 2. Пусть (ξ^1, η^1) — единственное решение системы (31), (32) при $\hat{x}^j = x^k$, $\bar{\lambda}^k = \lambda_E^k$ и $H_j = H_k$, где матрица H_k такова, что матрица $H_k + (h'(x^k))^T h'(x^k) / \sigma_k$ положительно определена. Пусть (ξ^2, η^2) — единственное решение системы (46), (47).

Тогда $\xi^1 = \xi^2$ и $\lambda_E^k + \eta^1 = \lambda^k + \eta^2$.

Из данного предложения, в частности, следует, что при фиксированных x^k , λ_E^k и H_k значения ξ и $\lambda^k + \eta$, где (ξ, η) определяется из соотношений (46), (47), не зависят от λ^k . Таким образом, если на некоторой итерации (внешней или внутренней) $\hat{x}^j = x^k$, $\bar{\lambda}^k = \lambda_E^k$ и $H_j = H_k$, то алгоритм 3 «попытается сделать шаг» в ту же точку, что и pdSQP, вне зависимости от текущего значения λ^k . Связь между направлениями поиска показана на рис. 1.

Сказанное свидетельствует о том, что алгоритм 3 и алгоритм pdSQP тесно связаны. Двойственные аппроксимации λ_E^k играют ту же роль, что и $\bar{\lambda}^k$ в алгоритме 3. Поскольку

$$\frac{\partial M_\sigma}{\partial(x, \lambda)}(x, \lambda; \lambda_E) = \begin{pmatrix} \frac{\partial L_\sigma}{\partial x}(x, \lambda_E) - \nu (h'(x))^T \left(\lambda - \left(\lambda_E + \frac{1}{\sigma} h(x) \right) \right) \\ \nu \sigma \left(\lambda - \left(\lambda_E + \frac{1}{\sigma} h(x) \right) \right) \end{pmatrix},$$

то условие

$$\frac{\partial M_{\sigma_k}}{\partial(x, \lambda)}(x^{k+1}, \lambda^{k+1}; \lambda_E^k) = 0$$

эквивалентно равенствам

$$\frac{\partial L_{\sigma_k}}{\partial x}(x^{k+1}, \lambda_E^k) = 0, \quad \lambda^{k+1} = \lambda_E^k + \frac{1}{\sigma_k} h(x^{k+1})$$

(что соответствует точному шагу МФЛ для текущей двойственной аппроксимации λ_E^k). Таким образом, в обоих алгоритмах техника глобализации сходимости основана на безусловной минимизации модифицированной функции Лагранжа при фиксированном значении двойственной переменной. Действительно, двойственное приближение ($\bar{\lambda}^k$ в случае алгоритма 3, и λ_E^k в случае алгоритма pdSQP) может меняться только в двух случаях:

— если достигнуто существенное улучшение рекордного значения r_k в алгоритме 3, или значений φ_V^{\max} или φ_O^{\max} в алгоритме pdSQP (осуществляется соответственно итерация sSQP, или V- или O-итерация);

— если найдено достаточно хорошее приближение к решению задачи безусловной минимизации модифицированной функции Лагранжа при соответствующем значении двойственной переменной (осуществляется итерация МФЛ, или M-итерация).

В противном случае двойственное приближение не меняется, т. е. F-итерация алгоритма pdSQP является по сути внутренней итерацией.

Правила обновления текущего множителя $\bar{\lambda}^k$ или λ_E^k также во многих случаях работают схожим образом. Предположим, что $\hat{x}^j = x^k$, $\bar{\lambda}^k = \lambda_E^k$ и $H_j = H_k$, причем λ_E^k не слишком велико по норме. Если алгоритмы осуществляют итерацию sSQP и V- или O-итерацию соответственно, причем $\alpha_k = 1$, то справедливо $\bar{\lambda}^{k+1} = \lambda_E^{k+1}$. Если же $\alpha_k < 1$, то значения $\bar{\lambda}^{k+1}$ и λ_E^{k+1} могут различаться, однако, как показывает вычислительный эксперимент, в большинстве случаев различия незначительны. Далее, если алгоритмы осуществляют итерацию МФЛ и M-итерацию соответственно, то $\lambda^{k+1} = \bar{\lambda}^k + h(x^{k+1})/\sigma_k$ для алгоритма 3, и $\lambda^{k+1} \approx \lambda_E^k + h(x^{k+1})/\sigma_k$ для алгоритма pdSQP, т. е. алгоритмы генерируют близкие значения λ^{k+1} . При этом в качестве $\bar{\lambda}^{k+1}$ и λ_E^{k+1} берется либо само λ^{k+1} , либо его проекция на заданный параллелепипед.

Наконец, отметим, что при одинаковых начальных значениях соответствующих переменных процедура прямо-двойственного гибкого одномерного поиска из [56] для семейства функций $M_\sigma(\cdot; \lambda_E^k)$ во многих случаях выдает такое же значение α_k параметра длины шага, что и обычная процедура прямого одномерного поиска для функции $L_{\sigma_k}(\cdot, \bar{\lambda}^k)$.

Таким образом, несмотря на то, что в основе алгоритма 3 лежит обычная модифицированная функция Лагранжа, а в основе алгоритма pdSQP — более сложная прямо-двойственная модифицированная функция Лагранжа, на большинстве шагов эти алгоритмы выполняют эквивалентные или очень близкие действия. Вместе с тем, у алгоритмов имеются некоторые различия, связанные со стремлением получить наиболее сильные теоретические результаты для алгоритма 3 и, в частности, сохранить сверхлинейную скорость сходимости метода sSQP. А именно, различаются правила обновления обратных параметров штрафа: в алгоритме pdSQP вблизи решения в большинстве случаев принимается $\sigma_{k+1} = \|\sigma(x^{k+1}, \lambda^{k+1})\|^{3/2}$. Такой выбор приводит к тому, что обратный параметр штрафа стремится к нулю слишком быстро, что может разрушать стабилизирующий эффект метода sSQP: в этом случае отсутствуют теоретические гарантии сверхлинейной сходимости последнего (см. комментарии к теореме 1.3, а также [72, глава 7]). Более подробно данный вопрос обсуждается в приложении Б.2.

В завершение данного раздела отметим, что алгоритм 3 допускает наличие произволь-

ных ограничений-неравенств, в то время как алгоритм pdSQP применим лишь к задачам вида

$$f(x) \rightarrow \min, \quad h(x) = 0, \quad l \leq x \leq u, \quad (48)$$

где $l_i \in \mathbb{R} \cup \{-\infty\}$, $u_i \in \mathbb{R} \cup \{+\infty\}$, $i = 1, \dots, n$. Разумеется, произвольную задачу вида (11) легко привести к форме (48) с помощью введения дополнительных переменных. Однако, помимо неизбежного увеличения размерности, такие переформулировки могут ухудшать свойства задачи. В частности, вычислительный эксперимент показывает, что поведение алгоритма 3 существенно ухудшается, если его применять к таким переформулировкам.

Отметим также, что когда алгоритм pdSQP применяется к задаче (48), его подзадачи имеют вид

$$\begin{aligned} \langle f'(x^k), \xi \rangle + \frac{1}{2} \langle H_k \xi, \xi \rangle + \frac{\sigma_k}{2} \|\lambda^k + \eta\|^2 &\rightarrow \min, \\ h(x^k) + h'(x^k)\xi - \sigma_k(\lambda^k + \eta - \lambda_E^k) &= 0, \quad l \leq x \leq u, \end{aligned}$$

и, вообще говоря, не являются подзадачами метода sSQP. В частности, отсутствуют какие-либо теоретические результаты о локальной сверхлинейной сходимости метода с такими подзадачами. В то же время, подзадачи, возникающие на шаге 1 алгоритма 3, являются именно подзадачами sSQP.

3.3. Глобализация сходимости с помощью точных гладких штрафных функций

В данном разделе предлагается подход к глобализации сходимости sSQP с помощью прямо-двойственных точных гладких штрафных функций. Основное отличие данного подхода от других способов глобализации сходимости, предложенных в этой главе, состоит в том, что он не является гибридным: на каждой итерации вычисляется прямо-двойственное направление поиска sSQP, вдоль которого затем осуществляется одномерный поиск для точной гладкой штрафной функции. Подходы, предложенные в разделах 3.1 и 3.2, напротив, являются гибридными в следующем смысле: в них либо вообще не вычисляется направление sSQP на некоторых итерациях, либо это направление или его часть может не использоваться (даже будучи вычисленным) для значительной доли итераций.

В разделе 3.3.1 приводится глобализованный алгоритм, и доказывается результат о его глобальной сходимости. Раздел 3.3.2 посвящен анализу скорости сходимости данного алгоритма. В приложении Б.3 приводятся результаты вычислительного эксперимента, которые показывают, что предлагаемый алгоритм имеет хорошую эффективность, и, в частности, существенно опережает обычный SQP на задачах с полным вырождением.

3.3.1. Глобализованный алгоритм

Будем рассматривать задачу с ограничениями-равенствами

$$f(x) \rightarrow \min, \quad h(x) = 0, \quad (49)$$

где $f: \mathbb{R}^n \rightarrow \mathbb{R}$ — дважды дифференцируемая функция, а $h: \mathbb{R}^n \rightarrow \mathbb{R}^l$ — дважды дифференцируемое отображение. Напомним, что стационарные точки и отвечающие им множители Лагранжа задачи (49) характеризуются системой Лагранжа

$$\frac{\partial L}{\partial x}(x, \lambda) = 0, \quad h(x) = 0, \quad (50)$$

где $L: \mathbb{R}^n \times \mathbb{R}^l \rightarrow \mathbb{R}$ — функция Лагранжа задачи (49):

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

Определим семейство прямо-двойственных гладких штрафных функций $\varphi_{c_1, c_2}: \mathbb{R}^n \times \mathbb{R}^l \rightarrow \mathbb{R}$:

$$\varphi_{c_1, c_2}(x, \lambda) = L(x, \lambda) + \frac{c_1}{2} \|h(x)\|^2 + \frac{c_2}{2} \left\| \frac{\partial L}{\partial x}(x, \lambda) \right\|^2, \quad (51)$$

где $c_1 > 0$ и $c_2 > 0$ — параметры штрафа. Впервые такие штрафные функции были предложены в [44]. Согласно теории в [4, раздел 4.3] и [35], функция φ_{c_1, c_2} является точной в следующем смысле. Если $c_2 > 0$ достаточно мало, а $c_1 > 0$ достаточно велико, то любая стационарная точка функции φ_{c_1, c_2} удовлетворяет системе Лагранжа (50); если прямо-двойственное решение (50) удовлетворяет условию регулярности $\text{rank } h'(\bar{x}) = l$ и достаточному условию второго порядка

$$\left\langle \frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda})\xi, \xi \right\rangle > 0 \quad \forall \xi \in \ker h'(\bar{x}) \setminus \{0\}, \quad (52)$$

то оно является строгим локальным минимумом функции φ_{c_1, c_2} . Таким образом, при использовании φ_{c_1, c_2} в роли функции качества для глобализации сходимости следует по возможности держать параметр c_2 малым, увеличивая его только в случае необходимости.

Напомним, что итерация sSQP для задачи (49) состоит в следующем: для текущего приближения $(x, \lambda) \in \mathbb{R}^n \times \mathbb{R}^l$ вычисляется решение $(\xi, \eta) \in \mathbb{R}^n \times \mathbb{R}^l$ системы

$$\frac{\partial^2 L}{\partial x^2}(x, \lambda)\xi + (h'(x))^T \eta = -\frac{\partial L}{\partial x}(x, \lambda), \quad h'(x)\xi - \sigma \eta = -h(x), \quad (53)$$

где σ — параметр стабилизации, и следующее приближение полагается равным $(x + \xi, \lambda + \eta)$.

Для всех $x \in \mathbb{R}^n$, $\lambda \in \mathbb{R}^l$ и $\sigma > 0$ определим матрицу

$$H_\sigma(x, \lambda) = \frac{\partial^2 L}{\partial x^2}(x, \lambda) + \frac{1}{\sigma} (h'(x))^T h'(x). \quad (54)$$

Напомним, что при достаточно малых значениях $\sigma > 0$ эта матрица является невырожденной для всех (x, λ) , не удовлетворяющих системе (50), но при этом близких к некоторому решению $(\bar{x}, \bar{\lambda})$, где $\bar{\lambda}$ — некритический множитель [71].

Следующая лемма показывает, что параметры c_1 и c_2 можно выбрать таким образом, что направление sSQP будет направлением убывания для функции φ_{c_1, c_2} .

Лемма 2. Пусть $f: \mathbb{R}^n \rightarrow \mathbb{R}$ и $h: \mathbb{R}^n \rightarrow \mathbb{R}^l$ дважды дифференцируемы в точке $x \in \mathbb{R}^n$, и пусть (ξ, η) — решение системы (53) для данного x при некоторых $\lambda \in \mathbb{R}^l$ и $\sigma > 0$.

Тогда справедливы следующие утверждения:

а) если $\langle h(x), h'(x)\xi \rangle < 0$, то

$$\langle \varphi'_{c_1, c_2}(x, \lambda), (\xi, \eta) \rangle \leq -\omega \quad (55)$$

для любого $\omega > 0$ и для всех $c_1 \geq \bar{c}_1(\omega; x, \lambda; \xi, \eta)$ и c_2 , где

$$\bar{c}_1(\omega; x, \lambda; \xi, \eta) = -\frac{\left\langle \frac{\partial L}{\partial x}(x, \lambda), \xi \right\rangle + \langle h(x), \eta \rangle + \omega}{\langle h(x), h'(x)\xi \rangle}; \quad (56)$$

б) если $\frac{\partial L}{\partial x}(x, \lambda) \neq 0$, то (55) выполнено для любого $\omega > 0$ и для всех c_1 и $c_2 \geq \bar{c}_2(\omega; x, \lambda; \xi, \eta; c_1)$, где

$$\bar{c}_2(\omega; x, \lambda; \xi, \eta; c_1) = \frac{\left\langle \frac{\partial L}{\partial x}(x, \lambda), \xi \right\rangle + \langle h(x), \eta \rangle + c_1 \langle h(x), h'(x)\xi \rangle + \omega}{\left\| \frac{\partial L}{\partial x}(x, \lambda) \right\|^2}; \quad (57)$$

в) если матрица $H_\sigma(x, \lambda)$ невырождена, то вектор (ξ, η) единственным образом определяется равенствами

$$\xi = -(H_\sigma(x, \lambda))^{-1} \left(\frac{\partial L}{\partial x}(x, \lambda) + \frac{1}{\sigma} (h'(x))^T h(x) \right), \quad (58)$$

$$\eta = \frac{1}{\sigma} \left(h(x) - h'(x) (H_\sigma(x, \lambda))^{-1} \left(\frac{\partial L}{\partial x}(x, \lambda) + \frac{1}{\sigma} (h'(x))^T h(x) \right) \right). \quad (59)$$

Доказательство. Дифференцируя равенство (51), для всех c_1 и c_2 получаем

$$\varphi'_{c_1, c_2}(x, \lambda) = \begin{pmatrix} I + c_2 \frac{\partial^2 L}{\partial x^2}(x, \lambda) & c_1 (h'(x))^T \\ c_2 h'(x) & I \end{pmatrix} \begin{pmatrix} \frac{\partial L}{\partial x}(x, \lambda) \\ h(x) \end{pmatrix}. \quad (60)$$

В частности,

$$\begin{aligned} \langle \varphi'_{c_1, c_2}(x, \lambda), (\xi, \eta) \rangle &= \left\langle \frac{\partial L}{\partial x}(x, \lambda), \xi \right\rangle + \langle h(x), \eta \rangle + c_1 \langle h(x), h'(x)\xi \rangle + \\ &+ c_2 \left\langle \frac{\partial L}{\partial x}(x, \lambda), \frac{\partial^2 L}{\partial x^2}(x, \lambda)\xi + (h'(x))^T \eta \right\rangle = \\ &= \left\langle \frac{\partial L}{\partial x}(x, \lambda), \xi \right\rangle + \langle h(x), \eta \rangle + c_1 \langle h(x), h'(x)\xi \rangle - c_2 \left\| \frac{\partial L}{\partial x}(x, \lambda) \right\|^2, \end{aligned} \quad (61)$$

где последнее равенство следует из первого уравнения в (53). Отсюда очевидным образом следуют утверждения а) и б).

Далее, второе равенство в (53) может быть записано в виде

$$\eta = \frac{1}{\sigma}(h(x) + h'(x)\xi). \quad (62)$$

Подставляя это соотношение в первое равенство в (53), с учетом (54) получаем

$$H_\sigma(x, \lambda)\xi = -\frac{\partial L}{\partial x}(x, \lambda) - \frac{1}{\sigma}(h'(x))^T h(x).$$

Если матрица $H_\sigma(x, \lambda)$ невырождена, последнее равенство единственным образом определяет ξ , и тогда (62) единственным образом определяет η . Из полученных соотношений следуют равенства (58) и (59), что доказывает утверждение в). ■

Из утверждений а) и б) леммы 2 несложно вывести правила выбора c_1 и c_2 таким образом, чтобы направление sSQP было направлением убывания для функции φ_{c_1, c_2} .

Введем отображение $\Phi: \mathbb{R}^n \times \mathbb{R}^l \rightarrow \mathbb{R}^n \times \mathbb{R}^l$ системы Лагранжа (50):

$$\Phi(x, \lambda) = \left(\frac{\partial L}{\partial x}(x, \lambda), h(x) \right). \quad (63)$$

Алгоритм 4. Выбираем параметры $\bar{\sigma} > 0$, $c_1 > 0$, $c_2 > 0$, $\delta > 0$, $\rho > 0$, $q > 1$, $\varepsilon, \theta \in (0, 1)$. Фиксируем непрерывные функции $\psi_1, \psi_2, \psi_3: \mathbb{R}_+ \rightarrow \mathbb{R}_+$, положительные везде кроме нуля. Выбираем $(x^0, \lambda^0) \in \mathbb{R}^n \times \mathbb{R}^l$ и полагаем $k = 0$.

1. Если $\Phi(x^k, \lambda^k) = 0$, где Φ определяется согласно (63), завершаем работу алгоритма.
2. Полагаем $\sigma_k = \min\{\bar{\sigma}, \|\Phi(x^k, \lambda^k)\|\}$ и вычисляем $d^k = (\xi^k, \eta^k)$ как решение системы (53) при $\sigma = \sigma_k$ и $(x, \lambda) = (x^k, \lambda^k)$. Если система (53) не имеет решений, переходим к шагу 4.
3. Если

$$\langle \varphi'_{c_1, c_2}(x^k, \lambda^k), d^k \rangle \leq -\rho \|d^k\|^q, \quad (64)$$

переходим к шагу 5.

Если

$$\|h(x^k)\| \geq \psi_1(\sigma_k) \quad (65)$$

и

$$\langle h(x^k), h'(x^k)\xi^k \rangle \leq -\psi_2(\|h(x^k)\|), \quad (66)$$

полагаем $c_1 = \bar{c}_{1, k} + \delta$, где $\bar{c}_{1, k} = \bar{c}_1(\rho \|d^k\|^q; x^k, \lambda^k; \xi^k, \eta^k)$, и переходим к шагу 5.

Если

$$\left\| \frac{\partial L}{\partial x}(x^k, \lambda^k) \right\| \geq \psi_3(\sigma_k), \quad (67)$$

полагаем $c_2 = \bar{c}_{2, k} + \delta$, где $\bar{c}_{2, k} = \bar{c}_2(\rho \|d^k\|^q; x^k, \lambda^k; \xi^k, \eta^k; c_1)$, и переходим к шагу 5.

4. Полагаем $d^k = -\varphi'_{c_1, c_2}(x^k, \lambda^k)$.

5. Вычисляем $\alpha_k = \theta^j$, где j — наименьшее неотрицательное целое число, удовлетворяющее неравенству Армихо

$$\varphi_{c_1, c_2}((x^k, \lambda^k) + \theta^j d^k) \leq \varphi_{c_1, c_2}(x^k, \lambda^k) + \varepsilon \theta^j \langle \varphi'_{c_1, c_2}(x^k, \lambda^k), d^k \rangle. \quad (68)$$

6. Полагаем $(x^{k+1}, \lambda^{k+1}) = (x^k, \lambda^k) + \alpha_k d^k$, увеличиваем k на 1 и переходим к шагу 1.

Замечание 1. На каждой итерации алгоритма либо оба параметра c_1, c_2 остаются неизменными, либо ровно один из них увеличивается на величину, не меньшую чем δ .

Действительно, параметры c_1 или c_2 могут измениться только если условие (64) не выполнено для текущих значений. Если при этом выполнены тесты (65) и (66), то, согласно утверждению а) леммы 2, до обновления имело место неравенство $c_1 < \bar{c}_1(\rho \|d^k\|^q; x^k, \lambda^k; \xi^k, \eta^k)$, поскольку в противном случае (64) было бы выполнено. Таким образом, на шаге 3 значение c_1 увеличивается на величину, не меньшую чем δ . Если же выполнен тест (67), то легко провести аналогичное рассуждение для параметра c_2 , используя утверждение б) леммы 2. Отметим, что в обоих случаях неравенство (64) будет справедливо для новых значений c_1 и c_2 .

Нетрудно видеть, что алгоритм корректно определен: либо он завершается на шаге 1 в точке, удовлетворяющей системе Лагранжа (50), либо d^k является направлением убывания для штрафной функции (возможно, после обновления c_1 или c_2), и это направление принимается на шаге 5 для некоторого j .

Отметим также, что на шаге 4 вместо антиградиента можно использовать любые другие направления поиска, обеспечивающие глобальную сходимость соответствующего метода спуска.

Следующая теорема показывает, что алгоритм 4 обладает разумными свойствами глобальной сходимости.

Теорема 7. Пусть $f: \mathbb{R}^n \rightarrow \mathbb{R}$ и $h: \mathbb{R}^n \rightarrow \mathbb{R}^l$ дважды непрерывно дифференцируемы на \mathbb{R}^n .

Тогда справедливы следующие утверждения:

а) если значения c_1 и c_2 не меняются для всех достаточно больших номеров k , то любая предельная точка $(\bar{x}, \bar{\lambda})$ последовательности $\{(x^k, \lambda^k)\}$ удовлетворяет равенству

$$\varphi'_{c_1, c_2}(\bar{x}, \bar{\lambda}) = 0; \quad (69)$$

б) если существует такое бесконечное множество K , что c_1 или c_2 увеличивается на всех итерациях с номерами $k \in K$, то любая предельная точка $(\bar{x}, \bar{\lambda})$ последовательности $\{(x^k, \lambda^k) \mid k \in K\}$ либо удовлетворяет системе Лагранжа (50), либо такова, что последовательность $\{d^k \mid k \in K\}$ неограничена, а матрица $H_{\tilde{\sigma}}(\bar{x}, \bar{\lambda})$ вырождена, где $\tilde{\sigma} = \tilde{\sigma}(\bar{x}, \bar{\lambda}) = \min\{\bar{\sigma}, \|\Phi(\bar{x}, \bar{\lambda})\|\} > 0$.

Доказательство. Предположим сначала, что параметры c_1 и c_2 не меняются для всех достаточно больших k . В этом случае все элементы последовательности $\{(x^k, \lambda^k)\}$, начиная с некоторого номера, генерируются методом спуска, глобализованным с помощью одномерного поиска с использованием правила Армихо для фиксированной гладкой функции φ_{c_1, c_2} . Покажем, что последовательность $\{d^k\}$ направлений поиска является равномерно градиентной (в терминологии [4]). Последнее означает, что если для некоторого бесконечного множества индексов K подпоследовательность $\{(x^k, \lambda^k) \mid k \in K\}$ сходится к точке $(\bar{x}, \bar{\lambda})$, для которой $\varphi'_{c_1, c_2}(\bar{x}, \bar{\lambda}) \neq 0$, то подпоследовательность $\{d^k \mid k \in K\}$ ограничена, и $\limsup_{K \ni k \rightarrow \infty} \langle \varphi'_{c_1, c_2}(x^k, \lambda^k), d^k \rangle < 0$.

Обозначим через $K_1 \subset K$ множество номеров $k \in K$, для которых d^k генерируется как решение системы (53). Тогда неравенство (64) выполнено для всех $k \in K_1$, и, в частности, $\|\varphi'_{c_1, c_2}(x^k, \lambda^k)\| \|d^k\| \geq \rho \|d^k\|^q$. Поэтому из условия $q > 1$ и непрерывности φ'_{c_1, c_2} в точке $(\bar{x}, \bar{\lambda})$, очевидно, следует ограниченность $\{d^k \mid k \in K_1\}$. Далее, из условий $\varphi'_{c_1, c_2}(\bar{x}, \bar{\lambda}) \neq 0$ и (60) получаем, что $\Phi(\bar{x}, \bar{\lambda}) \neq 0$. Отсюда и из (53) следует, что последовательность $\{d^k \mid k \in K_1\}$ не может иметь нулевую предельную точку. Таким образом, из (64) следует, что подпоследовательность $\{\langle \varphi'_{c_1, c_2}(x^k, \lambda^k), d^k \rangle \mid k \in K_1\}$ отделена от нуля некоторой отрицательной константой. Остается заметить, что для всех $k \in K \setminus K_1$ направление d^k получено на шаге 4 алгоритма 4, и требуемые свойства данного направления очевидны.

Таким образом, последовательность $\{d^k\}$ является равномерно градиентной. Поэтому, согласно [4, теорема 1.8], любая предельная точка $(\bar{x}, \bar{\lambda})$ последовательности $\{(x^k, \lambda^k)\}$ удовлетворяет (69).

Предположим теперь, что существует такое бесконечное множество индексов K , что один из параметров c_1, c_2 увеличивается на всех итерациях с номерами $k \in K$, причем подпоследовательность $\{(x^k, \lambda^k) \mid k \in K\}$ имеет предельную точку $(\bar{x}, \bar{\lambda})$, которая не удовлетворяет системе Лагранжа (50). Предположим далее, что подпоследовательность $\{d^k \mid k \in K\}$ ограничена. Без ограничения общности будем считать, что $\{(x^k, \lambda^k) \mid k \in K\} \rightarrow (\bar{x}, \bar{\lambda})$ и $\{d^k \mid k \in K\} \rightarrow \bar{d}$.

Сначала предположим, что параметр c_1 увеличивается на всех итерациях с номерами $k \in K$. Последнее означает, что тесты (65) и (66) выполнены для всех $k \in K$. Поскольку $\Phi(\bar{x}, \bar{\lambda}) \neq 0$, подпоследовательность $\{\sigma_k \mid k \in K\}$ имеет положительный предел. Переходя к

пределу в (65) вдоль соответствующей подпоследовательности, получаем неравенство $h(\bar{x}) \neq 0$. Таким образом, $\psi_2(\|h(\bar{x})\|) > 0$, и поэтому выражения в левой части (66) отделены от нуля для $k \in K$. Отсюда и из (56), с учетом сходимости подпоследовательностей $\{(x^k, \lambda^k) \mid k \in K\}$ и $\{d^k \mid k \in K\}$, получаем сходимость $\{\bar{c}_{1,k} \mid k \in K\}$. С другой стороны (см. замечание 1), для всех $k \in K$ параметр c_1 увеличивается по крайней мере на $\delta > 0$, а значит этот параметр стремится к бесконечности. Последнее, очевидно, противоречит ограниченности $\{\bar{c}_{1,k} \mid k \in K\}$, поскольку для всех $k \in K$ значение c_1 становится равным $\bar{c}_{1,k} + \delta$.

Проводя аналогичное рассуждение с использованием соотношений (57) и (67), несложно показать что c_2 также не может увеличиваться на всех итерациях с номерами $k \in K$. Таким образом, получено противоречие, а значит, в сделанных предположениях последовательность $\{d^k \mid k \in K\}$ не может быть ограниченной.

Остается показать, что если подпоследовательность $\{d^k \mid k \in K\}$ является неограниченной, причем $\{(x^k, \lambda^k) \mid k \in K\} \rightarrow (\bar{x}, \bar{\lambda})$ и $\Phi(\bar{x}, \bar{\lambda}) \neq 0$, то матрица $H_{\bar{\sigma}}(\bar{x}, \bar{\lambda})$, где $\bar{\sigma} = \min\{\bar{\sigma}, \|\Phi(\bar{x}, \bar{\lambda})\|\} > 0$, обязательно вырождена. Действительно, предположим, что это не так. Тогда, с учетом сходимости $\{\sigma_k \mid k \in K\} \rightarrow \bar{\sigma}$, получаем, что матрицы $H_{\sigma_k}(x^k, \lambda^k)$ невырождены для всех достаточно больших $k \in K$. Поэтому из утверждения в) леммы 2 следует, что для всех $k \in K$ направление $d^k = (\xi^k, \eta^k)$ единственным образом определяется формулами (58) и (59). Отсюда легко получить сходимость подпоследовательности $\{d^k \mid k \in K\}$, что, очевидно, противоречит ее неограниченности. ■

Отметим, что если предельная точка $(\bar{x}, \bar{\lambda})$ траектории алгоритма удовлетворяет (69), и при этом матрица в правой части (60) невырождена при $(x, \lambda) = (\bar{x}, \bar{\lambda})$ (что справедливо «почти для всех» значений c_1 и c_2), то $(\bar{x}, \bar{\lambda})$ удовлетворяет системе Лагранжа (50).

В случае, когда реализуется исход б), теорема 7 допускает существование у подпоследовательности $\{(x^k, \lambda^k) \mid k \in K\}$ предельных точек, не удовлетворяющих системе Лагранжа (50). Тем не менее, как следует из теоремы, подобная ситуация является очень специальной, поскольку в этом случае последовательность направлений sSQP должна быть неограниченной, а матрица вида (54) в пределе (вдоль соответствующей подпоследовательности) должна быть вырожденной. Результаты вычислительного эксперимента, приведенные в приложении Б.3, также показывают, что такие исходы крайне маловероятны.

Несмотря на то, что алгоритм 4 обладает хорошими свойствами глобальной сходимости, приведенное ниже замечание показывает, что эти свойства можно улучшить, если ввести в алгоритме верхние границы $C_1 > 0$ и $C_2 > 0$ для параметров штрафа c_1 и c_2 .

Замечание 2. Рассмотрим алгоритм 4 со следующей модификацией: для обновления c_1 и c_2 на шаге 3 будем использовать правила $c_1 = \min\{C_1, \bar{c}_{1,k} + \delta\}$ и $c_2 = \min\{C_2, \bar{c}_{2,k} + \delta\}$ при некоторых фиксированных $C_1 > 0$ и $C_2 > 0$. Несложно убедиться, что в этом случае,

начиная с некоторой итерации, значения c_1 и c_2 не будут меняться.

Действительно, рассуждая так же, как в замечании 1, можно показать, что последовательности параметров c_1 и c_2 остаются неубывающими, причем если один из этих параметров увеличивается, то он либо становится равным соответствующей верхней границе (и впоследствии не меняется), либо увеличивается на величину, не меньшую δ . Поэтому число итераций, на которых увеличивается параметр c_1 или c_2 не может превосходить C_1/δ и C_2/δ соответственно.

Таким образом, согласно утверждению а) теоремы 7, любая предельная точка траектории алгоритма будет удовлетворять (69).

Отметим, что введение верхних границ C_1 и C_2 имеет и практический смысл: использование слишком больших значений c_1 или c_2 неразумно с вычислительной точки зрения. Напомним также, что большие значения c_2 могут ухудшать теоретические свойства выбранной функции качества, и поэтому в случаях, когда нежелательна сходимость к решениям, в которых нарушается необходимое условие оптимальности второго порядка, следует выбирать малое значение C_2 . Однако, введение верхних границ потенциально может привести к тому, что направление sSQP будет использоваться реже.

3.3.2. Анализ скорости сходимости

Данный раздел посвящен анализу скорости сходимости алгоритма 4. Сначала доказывается несколько результатов о принятии единичного параметра длины шага в общих методах спуска. Эти результаты имеют отношение к известной теореме Денниса-Море, однако покрывают случаи отсутствующих вторых производных и неизолированных решений (последнее имеет особое значение в контексте sSQP). Затем с помощью полученных результатов обосновывается локальная сверхлинейная сходимость алгоритма 4.

Принятие единичного параметра длины шага в методах спуска

Будем рассматривать задачу безусловной оптимизации

$$\varphi(u) \rightarrow \min, \quad u \in \mathbb{R}^{\nu} \quad (70)$$

с гладкой целевой функцией $\varphi: \mathbb{R}^{\nu} \rightarrow \mathbb{R}$. Рассмотрим общий метод спуска для задачи (70). А именно, возьмем произвольный вектор $u^0 \in \mathbb{R}^{\nu}$, и для каждого k будем полагать

$$u^{k+1} = u^k + \alpha_k d^k,$$

где $d^k \in \mathbb{R}^{\nu}$ — некоторое направление убывания функции φ в точке u^k , а параметр длины шага α_k вычисляется с помощью одномерного поиска с использованием правила Армихо: для

некоторых фиксированных параметров $\varepsilon, \theta \in (0, 1)$ положим $\alpha_k = \theta^j$, где j — наименьшее неотрицательное целое число, удовлетворяющее неравенству

$$\varphi(u^k + \theta^j d^k) \leq \varphi(u^k) + \varepsilon \theta^j \langle \varphi'(u^k), d^k \rangle.$$

В данном разделе будет получено несколько результатов о достаточных условиях принятия единичного шага в описанном выше методе спуска. Эти результаты имеют отношение к соответствующей части теоремы Денниса-Море (см., например, [4, предложение 1.15]), которая состоит в следующем.

Предложение 3. Пусть функция $\varphi: \mathbb{R}^{\nu} \rightarrow \mathbb{R}$ дважды дифференцируема в окрестности точки $\bar{u} \in \mathbb{R}^{\nu}$, а ее вторая производная непрерывна в этой точке. Пусть \bar{u} — стационарная точка задачи (70), т. е. $\varphi'(\bar{u}) = 0$, и пусть в этой точке выполнено достаточное условие оптимальности второго порядка, т. е. матрица $\varphi''(\bar{u})$ положительно определена. Предположим, что последовательность $\{u^k\}$ сгенерирована описанным выше методом спуска, в котором полагалось $\varepsilon \in (0, 1/2)$. Предположим также, что $\{u^k\}$ сходится к \bar{u} , и

$$\|d^k + (\varphi''(\bar{u}))^{-1} \varphi'(u^k)\| = o(\|\varphi'(u^k)\|) \quad (71)$$

при $k \rightarrow \infty$.

Тогда $\alpha_k = 1$ для всех достаточно больших k .

Следующее предложение можно рассматривать как теорему Денниса-Море «первого порядка».

Предложение 4. Пусть функция $\varphi: \mathbb{R}^{\nu} \rightarrow \mathbb{R}$ дифференцируема в окрестности точки $\bar{u} \in \mathbb{R}^{\nu}$. Пусть последовательность $\{u^k\} \subset \mathbb{R}^{\nu}$ сходится к \bar{u} , а последовательность $\{d^k\} \subset \mathbb{R}^{\nu}$ сходится к нулю. Предположим, что существует $\beta > 0$, для которого выполнено

$$\varphi(u^k + d^k) - \varphi(u^k) \leq \beta \langle \varphi'(u^k), d^k \rangle + o(\|d^k\|^2) \quad (72)$$

при $k \rightarrow \infty$, и существует такое $\gamma > 0$, что

$$\langle \varphi'(u^k), d^k \rangle \leq -\gamma \|d^k\|^2 \quad (73)$$

для всех достаточно больших k .

Тогда для любого $\varepsilon \in (0, \beta)$ неравенство

$$\varphi(u^k + d^k) \leq \varphi(u^k) + \varepsilon \langle \varphi'(u^k), d^k \rangle \quad (74)$$

выполнено для всех достаточно больших k .

Доказательство. Объединяя неравенства (72) и (73), получаем

$$\begin{aligned} \varphi(u^k + d^k) - \varphi(u^k) - \varepsilon \langle \varphi'(u^k), d^k \rangle &\leq (\beta - \varepsilon) \langle \varphi'(u^k), d^k \rangle + o(\|d^k\|^2) \leq \\ &\leq -\gamma(\beta - \varepsilon) \|d^k\|^2 + o(\|d^k\|^2) \leq 0 \end{aligned}$$

для всех достаточно больших k , что и доказывает предложение. \blacksquare

Приведенный выше простой результат имеет следующие особенности по сравнению с существующими аналогами. Во-первых, не требуется никакая связь направления d^k с градиентом $\varphi'(u^k)$ помимо соотношений (72) и (73). Во-вторых, не предполагается существование второй производной функции φ , и поэтому нельзя использовать ни условие Денниса-Море (71), ни положительную определенность $\varphi''(\bar{u})$. Более того, рассматриваемый метод не обязательно должен быть ньютоновским или квазиньютоновским (вторая производная может даже не существовать в окрестности \bar{u}), и, в частности, полные шаги не обязательно должны обеспечивать сверхлинейное убывание расстояния до решения.

Предположения (72) и (73) могут показаться несколько неожиданными (особенно (72)), однако их осмысленность подтверждается по крайней мере приводимым ниже следствием 2, которое уже значительно ближе к утверждению теоремы Денниса-Море, но по-прежнему покрывает более широкий набор методов спуска.

Следствие 2. Пусть функция $\varphi: \mathbb{R}^{\nu} \rightarrow \mathbb{R}$ дважды дифференцируема в окрестности точки $\bar{u} \in \mathbb{R}^{\nu}$, а ее вторая производная непрерывна в этой точке. Пусть $\bar{u} \in \bar{U} = \{u \in \mathbb{R}^{\nu} \mid \varphi'(u) = 0\}$. Предположим, что последовательность $\{u^k\} \subset \mathbb{R}^{\nu}$ сходится к \bar{u} , последовательность $\{d^k\} \subset \mathbb{R}^{\nu}$ сходится к нулю,

$$\text{dist}(u^k + d^k, \bar{U}) = o(\text{dist}(u^k, \bar{U})) \quad (75)$$

при $k \rightarrow \infty$, и существует такое $\gamma > 0$, что

$$\langle \varphi''(\bar{u})d^k, d^k \rangle \geq \gamma \|d^k\|^2 \quad (76)$$

для всех достаточно больших k .

Тогда для любого $\varepsilon \in (0, 1/2)$ неравенство (74) выполнено для всех достаточно больших k .

Доказательство. Покажем, что предположения предложения 4 выполнены для $\beta = 1/2$.

Для каждого k обозначим через \hat{u}^k произвольную проекцию точки $u^k + d^k$ на множество \bar{U} (отметим, что множество \bar{U} не обязательно выпуклое, и поэтому проекция не обязательно

единственна). Поскольку $\{u^k\} \rightarrow \bar{u}$ и $\{d^k\} \rightarrow 0$, то $\{\hat{u}^k\} \rightarrow \bar{u}$. Тогда из (75) получаем

$$\|u^k + d^k - \hat{u}^k\| = o(\text{dist}(u^k, \bar{U})) = o(\|u^k - \hat{u}^k\|), \quad (77)$$

откуда, очевидно, следует

$$u^k - \hat{u}^k = -d^k + o(\|d^k\|) \quad (78)$$

при $k \rightarrow \infty$.

Из теоремы о среднем и непрерывности φ'' в точке \bar{u} получаем

$$\varphi'(u^k + d^k) = \varphi'(u^k + d^k) - \varphi'(\hat{u}^k) = O(\|u^k + d^k - \hat{u}^k\|) = o(\|u^k - \hat{u}^k\|) = o(\|d^k\|)$$

при $k \rightarrow \infty$, где последние две оценки следуют из соотношений (77) и (78) соответственно.

С другой стороны,

$$\varphi'(u^k + d^k) = \varphi'(u^k) + \varphi''(u^k)d^k + o(\|d^k\|),$$

и поэтому, с учетом предыдущей оценки, получаем

$$\varphi'(u^k) + \varphi''(u^k)d^k = o(\|d^k\|) \quad (79)$$

при $k \rightarrow \infty$. Далее,

$$\varphi(u^k + d^k) - \varphi(u^k) = \langle \varphi'(u^k), d^k \rangle + \frac{1}{2} \langle \varphi''(u^k)d^k, d^k \rangle + o(\|d^k\|^2) = \frac{1}{2} \langle \varphi'(u^k), d^k \rangle + o(\|d^k\|^2) \quad (80)$$

при $k \rightarrow \infty$, где последнее равенство следует из (79). Отсюда получаем (72) при $\beta = 1/2$.

Наконец, вновь применяя (79), выводим

$$\langle \varphi'(u^k), d^k \rangle = -\langle \varphi''(\bar{u})d^k, d^k \rangle + o(\|d^k\|^2)$$

при $k \rightarrow \infty$, и поэтому из (76) следует (73) при некотором $\gamma > 0$. ■

Основные отличия следствия 2 от предложения 3 (части теоремы Денниса-Море) состоят в следующем. Вместо положительной определенности матрицы $\varphi''(\bar{u})$ и условия (71) здесь явно предполагается, что полный шаг метода обеспечивает сверхлинейное убывание расстояние до множества решений, и что квадратичная форма, задаваемая матрицей Гессе целевой функции, равномерно положительно определена на направлениях, генерируемых методом. Важное отличие состоит в том, что следствие 2 может быть применимо даже если матрица $\varphi''(\bar{u})$ не является положительно определенной (и, в частности, если решение не является изолированным). Данная особенность крайне важна для анализа сходимости глобализованного метода sSQP, поскольку область применения данного метода включает в себя вырожденные задачи с неизолрованными решениями.

Легко проверить, что из предположений предложения 3 следуют предположения следствия 2, а значит и предложения 4. В то же время, предложения 4 может быть применимо тогда, когда следствие 2 (а значит и предложение 3) применить нельзя. Последнее

возможно даже в случаях, когда функция φ дважды дифференцируема в окрестности \bar{u} , ее вторая производная непрерывна в точке \bar{u} , и матрица $\varphi''(\bar{u})$ положительно определена. Действительно, рассмотрим в данных предположениях масштабированный метод Ньютона $d^k = -\tau(\varphi''(u^k))^{-1}\varphi'(u^k)$ с множителем $\tau \in (0, 2)$. Тогда

$$\langle \varphi'(u^k), d^k \rangle = -\frac{1}{\tau} \langle \varphi''(u^k) d^k, d^k \rangle = -\frac{1}{\tau} \langle \varphi''(\bar{u}) d^k, d^k \rangle + o(\|d^k\|^2)$$

при $k \rightarrow \infty$, откуда следует (73) при некотором $\gamma > 0$. Далее, вновь применяя первое равенство в (80), получаем

$$\varphi(u^k + d^k) - \varphi(u^k) = \left(1 - \frac{\tau}{2}\right) \langle \varphi'(u^k), d^k \rangle + o(\|d^k\|^2)$$

при $k \rightarrow \infty$, а значит (72) выполнено для любого $\beta \in (0, 1 - \tau/2)$. Таким образом, согласно предложению 4, в методе спуска, использующем такие направления поиска, будет приниматься единичный параметр длины шага на всех итерациях с достаточно большими номерами. С другой стороны, полный шаг такого метода не обеспечивает сверхлинейную скорость сходимости, если $\tau \neq 1$: в этом случае не выполняется (79), а значит и (75), и поэтому следствие 2 неприменимо.

Скорость сходимости глобализованного алгоритма

Способы выбора подходящих начальных значений параметров c_1 и c_2 в алгоритме 4 описаны в [35]. Увеличение этих параметров, конечно, не является необходимым для доказательства глобальной сходимости (увеличение c_2 даже нежелательно). Однако, для сверхлинейной сходимости алгоритма 4 необходимо, чтобы полные шаги метода sSQP принимались в окрестности квалифицированного решения, т.е. точки $(\bar{x}, \bar{\lambda})$, удовлетворяющей системе Лагранжа (50), и такой, что $\bar{\lambda}$ является не критическим множителем Лагранжа, отвечающим \bar{x} . Достижение данной цели без увеличения c_2 часто возможно, но может привести к слишком быстрому росту c_1 .

Чтобы продемонстрировать данную ситуацию, рассмотрим задачу из примера 1.3. Пусть $n = l = 1$, $f(x) = x^2$, $h(x) = x^2$. Задача (49) с такими данными имеет единственную допустимую точку (а значит, единственное решение) $\bar{x} = 0$, причем $\mathcal{M}(0) = \mathbb{R}$, и единственным критическим множителем Лагранжа является $\lambda = -1$.

Если $1 + \lambda + 2x^2/\sigma \neq 0$, то система (53) для данной задачи имеет единственное решение

$$\xi = -S(\sigma; x, \lambda)x, \quad \eta = \frac{1}{\sigma}(1 - 2S(\sigma; x, \lambda))x^2, \quad (81)$$

где

$$S(\sigma; x, \lambda) = \frac{1 + \lambda + x^2/\sigma}{1 + \lambda + 2x^2/\sigma}.$$

Предполагая, что вектор $x \neq 0$ достаточно близок к 0, а λ близок к любому множителю $\bar{\lambda} \neq -1$, получаем, что величина $\sigma = \|\Phi(x, \lambda)\|$ имеет порядок $|x|$, и выполнено $h(x) = O(\sigma^2)$. Поэтому тест (65) может выполняться в точке $(x^k, \lambda^k) = (x, \lambda)$, если выбрать $\psi_1(t) = O(t^2)$. Далее, $S(\sigma; x, \lambda) = 1 + O(|x|)$, а значит,

$$h(x)h'(x)\xi = -2S(\sigma; x, \lambda)x^4 = -2x^4 + O(x^5).$$

Таким образом, тест (66) может выполняться, если выбрать $\psi_2(t) = O(t^2)$ (достаточно естественный выбор), поскольку в этом случае $\psi_2(\|h(x)\|) = O(x^4)$. В то же время, из (61) несложно получить цепочку равенств

$$\begin{aligned} \langle \varphi'_{c_1, c_2}(x, \lambda), (\xi, \eta) \rangle &= \\ &= -2S(\sigma; x, \lambda)x^2(1 + \lambda) + \frac{1}{\sigma}(1 - 2S(\sigma; x, \lambda))x^4 - 2c_1S(\sigma; x, \lambda)x^4 - 4c_2x^2(1 + \lambda)^2 = \\ &= -2(1 + O(|x|))x^2(1 + \lambda) - \frac{1}{\sigma}(1 + O(|x|))x^4 - 2c_1(1 + O(|x|))x^4 - 4c_2x^2(1 + \lambda)^2 = \\ &= -2x^2(1 + \lambda) + O(|x|^3) - 2c_1(x^4 + O(|x|^5)) - 4c_2x^2(1 + \lambda)^2. \end{aligned}$$

Если $\bar{\lambda} < -1$, что означает нарушение SOS (52) в точке $(\bar{x}, \bar{\lambda})$, то для того, чтобы сделать выражение в правой части отрицательным только посредством изменения c_1 (т.е. если $c_2 \geq 0$ остается фиксированным и небольшим), может потребоваться выбирать c_1 порядка $1/x^2$. Если $q \geq 2$, то при таком выборе c_1 также возможно выполнение (64); если же $q < 2$, то c_1 должно расти даже быстрее, поскольку величина $\|(\xi, \eta)\|$ имеет порядок $|x|$. С другой стороны, (ξ, η) может быть направлением убывания, если взять достаточно большое, но ограниченное значение c_2 . Для этого необходимо, чтобы $\psi_1(t)$ убывала медленнее, чем t^2 , при $t \rightarrow 0+$, и выбрать $\psi_3(t) = O(t)$: в этом случае тест (65) не будет выполнен для достаточно больших k , в то время как тест (67) может выполняться.

Начнем локальный анализ со следующего результата о принятии направлений sSQP и стабилизации параметров штрафа.

Лемма 3. Пусть $f: \mathbb{R}^n \rightarrow \mathbb{R}$ и $h: \mathbb{R}^n \rightarrow \mathbb{R}^l$ дважды дифференцируемы в окрестности точки $\bar{x} \in \mathbb{R}^n$, а их вторые производные непрерывны в этой точке. Пусть \bar{x} — стационарная точка задачи (49), а $\bar{\lambda} \in \mathbb{R}^l$ — отвечающий ей не критический множитель Лагранжа. Пусть функции ψ_1, ψ_2, ψ_3 в алгоритме 4 выбраны следующим образом: $\psi_1(t) = \rho_1 t$, $\psi_2(t) = \rho_2 t^2$, $\psi_3(t) = \rho_3 t$, где $\rho_1, \rho_3 > 0$, $\rho_1^2 + \rho_3^2 < 1$, $\rho_2 \in (0, 1)$.

Тогда для любой точки $(x^k, \lambda^k) \in \mathbb{R}^n \times \mathbb{R}^l$, не удовлетворяющей системе Лагранжа (50) и достаточно близкой к $(\bar{x}, \bar{\lambda})$, система (53) при $\sigma = \sigma_k = \|\Phi(x^k, \lambda^k)\|$ и $(x, \lambda) = (x^k, \lambda^k)$ имеет единственное решение $d^k = (\xi^k, \eta^k)$, причем выполняются тесты (65) и (66) или выполняется тест (67) (т.е. алгоритм 4 принимает направление d^k и устанавливает c_1 и c_2 таким образом, чтобы неравенство (64) было выполнено для этого направления).

Более того, если $q \geq 2$, то величина

$$\bar{c}_1(\rho \|d^k\|^q; x^k, \lambda^k; \xi^k, \eta^k)$$

равномерно ограничена для всех (x^k, λ^k) , достаточно близких к $(\bar{x}, \bar{\lambda})$ и удовлетворяющих (65) и (66), а величина

$$\bar{c}_2(\rho \|d^k\|^q; x^k, \lambda^k; \xi^k, \eta^k; c_1)$$

равномерно ограничена для всех (x^k, λ^k) , достаточно близких к $(\bar{x}, \bar{\lambda})$ и удовлетворяющих (67), и для ограниченных значений c_1 . В частности, если траектория алгоритма 4, начиная с некоторой итерации, лежит в достаточно малой окрестности точки $(\bar{x}, \bar{\lambda})$, то значения c_1 и c_2 , начиная с некоторой итерации, остаются постоянными.

Доказательство. Существование и единственность решения d^k системы (53) следует из утверждения а) теоремы 1.3.

Если оба теста (65), (67) не выполнены в точке (x^k, λ^k) , достаточно близкой к $(\bar{x}, \bar{\lambda})$, то

$$\sigma_k^2 = \|\Phi(x^k, \lambda^k)\|^2 \leq (\rho_1^2 + \rho_3^2)\sigma_k^2,$$

что противоречит предположению $\rho_1^2 + \rho_3^2 < 1$, поскольку $\sigma_k \neq 0$. Таким образом, один из этих тестов обязательно выполнен. Если тест (65) не выполняется, но выполнен тест (67), то алгоритм принимает направление sSQP (возможно, модифицируя c_2).

Предположим теперь, что выполнен тест (65). Применяя предложение 1.1 и теорему 1.3 б), получаем оценку

$$\|(\xi^k, \eta^k)\| \leq \ell \|\Phi(x^k, \lambda^k)\| \quad (82)$$

для некоторого $\ell > 0$ и для всех (x^k, λ^k) , достаточно близких к $(\bar{x}, \bar{\lambda})$. Поэтому, используя второе равенство в (53), получаем, что

$$\begin{aligned} \langle h(x^k), h'(x^k)\xi^k \rangle &= \langle h(x^k), -h(x^k) + \sigma_k \eta^k \rangle \leq -\|h(x^k)\|^2 + \sigma_k \|h(x^k)\| \|\eta^k\| \leq \\ &\leq -\|h(x^k)\|^2 + \ell \sigma_k^2 \|h(x^k)\| \leq -\|h(x^k)\|^2 + \frac{\ell}{\rho_1^2} \|h(x^k)\|^3 = -\left(1 - \frac{\ell}{\rho_1^2} \|h(x^k)\|\right) \|h(x^k)\|^2, \end{aligned}$$

откуда следует, что тест (66) с заданной функцией ψ_2 выполняется для всех x^k , достаточно близких к \bar{x} . Таким образом, в этом случае алгоритм также принимает направление sSQP (возможно, модифицируя c_1).

Далее, из (82) следует оценка

$$\left\langle \frac{\partial L}{\partial x}(x^k, \lambda^k), \xi^k \right\rangle + \langle h(x^k), \eta^k \rangle = O(\|\Phi(x^k, \lambda^k)\|^2)$$

при $(x^k, \lambda^k) \rightarrow (\bar{x}, \bar{\lambda})$, и поэтому последнее утверждение леммы легко следует из (56), (57), и (82). ■

Остается показать, что единичный параметр длины шага будет приниматься правилом Армихо на всех итерациях с достаточно большими номерами, если метод сходится к точке $(\bar{x}, \bar{\lambda})$ удовлетворяющей системе Лагранжа (50), и если множитель $\bar{\lambda}$ является некритическим.

Лемма 4. Пусть $f: \mathbb{R}^n \rightarrow \mathbb{R}$ и $h: \mathbb{R}^n \rightarrow \mathbb{R}^l$ дважды дифференцируемы в окрестности точки $\bar{x} \in \mathbb{R}^n$, а их вторые производные непрерывны в этой точке. Пусть \bar{x} — стационарная точка задачи (49), а $\bar{\lambda} \in \mathbb{R}^l$ — отвечающий ей множитель Лагранжа. Пусть $\sigma: \mathbb{R}^n \times \mathbb{R}^l \rightarrow \mathbb{R}_+$ — произвольная функция, для которой выполнено $\sigma(u) \rightarrow 0$ при $u \rightarrow \bar{u} = (\bar{x}, \bar{\lambda})$ и $\sigma(0) = 0$. Предположим, что для любого $u = (x, \lambda)$, достаточно близкого к \bar{u} , существует вектор $d(u) = (\xi, \eta)$, удовлетворяющий системе (53) при $\sigma = \sigma(u)$, и для которого выполнено $d(u) \rightarrow 0$ при $u \rightarrow \bar{u}$ и $d(\bar{u}) = 0$.

Тогда для любых $c_1 > 0$ и $c_2 > 0$ существует окрестность \mathcal{V} точки \bar{u} и непрерывная нуле функция $\chi: \mathbb{R}^n \times \mathbb{R}^l \rightarrow \mathbb{R}_+$, для которой справедлива оценка $\chi(u) = o(\|d(u)\|^2)$ при $u \rightarrow \bar{u}$ и неравенство

$$\varphi_{c_1, c_2}(u + d(u)) - \varphi_{c_1, c_2}(u) \leq \frac{1}{2} \langle \varphi'_{c_1, c_2}(u), d(u) \rangle + \chi(u) \quad (83)$$

для всех $u \in \mathcal{V}$.

Доказательство. Из (53) следует существование окрестности \mathcal{V}_1 точки \bar{u} и функции $\chi_1: \mathbb{R}^n \times \mathbb{R}^l \rightarrow \mathbb{R}_+$, таких что вектор $d(u) = (\xi, \eta)$ с нужными свойствами существует для всех $u \in \mathcal{V}_1$, $\chi_1(u) = o(\|\xi\|)$ при $u \rightarrow \bar{u}$, и выполнены неравенства

$$\left\| \frac{\partial L}{\partial x}(x + \xi, \lambda + \eta) \right\| \leq \left\| \frac{\partial L}{\partial x}(x, \lambda) + \frac{\partial^2 L}{\partial x^2}(x, \lambda)\xi + (h'(x))^T \eta \right\| + \chi_1(u) = \chi_1(u), \quad (84)$$

$$\|h(x + \xi)\| \leq \|h(x) + h'(x)\xi\| + \chi_1(u) = \sigma(u)\|\eta\| + \chi_1(u) \quad (85)$$

для всех $u \in \mathcal{V}_1$.

Далее, из (85) следует, что

$$L(x + \xi, \lambda + \eta) = L(x + \xi, \lambda) + \langle \eta, h(x + \xi) \rangle \leq L(x + \xi, \lambda) + \sigma(u)\|\eta\|^2 + \chi_1(u)\|\eta\|$$

для всех $u \in \mathcal{V}_1$, а значит существует окрестность $\mathcal{V} \subset \mathcal{V}_1$ точки \bar{u} и функция $\chi_2: \mathbb{R}^n \times \mathbb{R}^l \rightarrow \mathbb{R}_+$, для которой справедлива оценка $\chi_2(u) = o(\|d(u)\|^2)$ при $u \rightarrow \bar{u}$ и неравенство

$$L(x + \xi, \lambda + \eta) - L(x, \lambda) \leq \left\langle \frac{\partial L}{\partial x}(x, \lambda), \xi \right\rangle + \frac{1}{2} \left\langle \frac{\partial^2 L}{\partial x^2}(x, \lambda)\xi, \xi \right\rangle + \chi_2(u) \quad (86)$$

для всех $u \in \mathcal{V}$. С другой стороны, вновь используя (53), получаем

$$\left\langle \frac{\partial L}{\partial x}(x, \lambda) + \frac{\partial^2 L}{\partial x^2}(x, \lambda)\xi, \xi \right\rangle = -\langle (h'(x))^T \eta, \xi \rangle = \langle \eta, h(x) - \sigma(u)\eta \rangle \leq \langle \eta, h(x) \rangle + \sigma(u)\|\eta\|^2,$$

откуда, в свою очередь, следует выполнение

$$\begin{aligned} \langle L'(x, \lambda), d(u) \rangle &= \left\langle \frac{\partial L}{\partial x}(x, \lambda), \xi \right\rangle + \langle h(x), \eta \rangle \geq \\ &\geq 2 \left\langle \frac{\partial L}{\partial x}(x, \lambda), \xi \right\rangle + \left\langle \frac{\partial^2 L}{\partial x^2}(x, \lambda) \xi, \xi \right\rangle - \sigma(u) \|\eta\|^2 \end{aligned}$$

для всех $u \in \mathcal{V}_1$. Комбинируя полученное неравенство с (86), выводим оценку

$$L(x + \xi, \lambda + \eta) - L(x, \lambda) \leq \frac{1}{2} \langle L'(x, \lambda), d(u) \rangle + \frac{1}{2} \sigma(u) \|\eta\|^2 + \chi_2(u) \quad (87)$$

для всех $u \in \mathcal{V}$.

Из соотношений (51), (84), (85) и (87) следует неравенство

$$\begin{aligned} \varphi_{c_1, c_2}(u + d(u)) - \varphi_{c_1, c_2}(u) &\leq \frac{1}{2} \langle L'(x, \lambda), d(u) \rangle + \frac{1}{2} \sigma(u) \|\eta\|^2 + \chi_2(u) + \\ &+ \frac{c_1}{2} (\sigma(u) \|\eta\| + \chi_1(u))^2 + \frac{c_2}{2} (\chi_1(u))^2 - \frac{c_1}{2} \|h(x)\|^2 - \frac{c_2}{2} \left\| \frac{\partial L}{\partial x}(x, \lambda) \right\|^2 \end{aligned}$$

для всех $u \in \mathcal{V}$.

Заметим, что функция $\tilde{\chi}: \mathbb{R}^n \times \mathbb{R}^l \rightarrow \mathbb{R}_+$, определяемая равенством

$$\tilde{\chi}(u) = \frac{c_1}{2} (\sigma(u) \|\eta\| + \chi_1(u))^2 + \frac{c_2}{2} (\chi_1(u))^2 + \frac{1}{2} \sigma(u) \|\eta\|^2 + \chi_2(u),$$

удовлетворяет оценке $\tilde{\chi}(u) = o(\|d(u)\|^2)$ при $u \rightarrow \bar{u}$. Далее, согласно (61),

$$\langle \varphi'_{c_1, c_2}(x, \lambda), d(u) \rangle = \langle L'(x, \lambda), d(u) \rangle + c_1 \langle h(x), h'(x) \xi \rangle - c_2 \left\| \frac{\partial L}{\partial x}(x, \lambda) \right\|^2,$$

и поэтому, вновь применяя (53), получаем соотношения

$$\begin{aligned} \varphi_{c_1, c_2}(u + d(u)) - \varphi_{c_1, c_2}(u) - \frac{1}{2} \langle \varphi'_{c_1, c_2}(u), d(u) \rangle &\leq -\frac{c_1}{2} (\|h(x)\|^2 + \langle h(x), h'(x) \xi \rangle) + \tilde{\chi}(u) = \\ &= -\frac{c_1}{2} \langle h(x), h(x) + h'(x) \xi \rangle + \tilde{\chi}(u) = -\frac{c_1}{2} \langle h(x), \sigma(u) \eta \rangle + \tilde{\chi}(u) = \\ &= \sigma(u) \frac{c_1}{2} \langle h'(x) \xi - \sigma(u) \eta, \eta \rangle + \tilde{\chi}(u) \leq \frac{c_1}{2} \sigma(u) (\|h'(x) \xi\| + \sigma(u) \|\eta\|) \|\eta\| + \tilde{\chi}(u) \end{aligned}$$

для всех $u \in \mathcal{V}$. Для того чтобы получить (83), остается определить $\chi: \mathbb{R}^n \times \mathbb{R}^l \rightarrow \mathbb{R}_+$ как

$$\chi(u) = \frac{c_1}{2} \sigma(u) (\|h'(x) \xi\| + \sigma(u) \|\eta\|) \|\eta\| + \tilde{\chi}(u)$$

и заметить, что $\chi(u) = o(\|d(u)\|^2)$ при $u \rightarrow \bar{u}$. ■

Предполагая теперь выполнение $q = 2$ вместо $q \geq 2$ и применяя леммы 3 и 4, а также предложение 4, приходим к следующей теореме.

Теорема 8. *В дополнение к предположениям леммы 3, пусть в алгоритме 4 используются значения $q = 2$ и $\varepsilon \in (0, 1/2)$. Предположим, что траектория $\{(x^k, \lambda^k)\}$, сгенерированная алгоритмом 4, сходится к точке $(\bar{x}, \bar{\lambda})$.*

Тогда $\alpha_k = 1$ для всех достаточно больших k , и скорость сходимости последовательности $\{(x^k, \lambda^k)\}$ является сверхлинейной.

Доказательство. Согласно лемме 3, на всех итерациях с достаточно большими номерами алгоритм принимает направление $d^k = d(u^k)$ метода sSQP, где $u^k = (x^k, \lambda^k)$, параметры штрафа c_1 и c_2 остаются постоянными, и выполняется неравенство (64) при $q = 2$.

Более того, согласно оценке (82) (которая следует из утверждения б) теоремы 1.3), $d(u^k) \rightarrow 0$ при $u^k \rightarrow \bar{u}$. Применяя лемму 4, получаем оценку

$$\varphi_{c_1, c_2}(u^k + d^k) - \varphi_{c_1, c_2}(u^k) \leq \frac{1}{2} \langle \varphi'_{c_1, c_2}(u^k), d^k \rangle + o(\|d^k\|^2)$$

при $k \rightarrow \infty$.

Таким образом, применяя предложение 4 (при $\varphi = \varphi_{c_1, c_2}$ и $\beta = 1/2$), получаем, что $\alpha_k = 1$ для всех достаточно больших k . Сверхлинейная сходимость следует из утверждения в) теоремы 1.3. \blacksquare

Оставшаяся часть данного раздела посвящена получению несколько иного результата о принятии единичного параметра длины шага. Проводимый ниже анализ не использует никаких дополнительных предположений на степень q . Более того, он применим не только к алгоритму 4, но и к более общим итерационным схемам. В частности, полученные ниже результаты справедливы для любого метода спуска, использующего направления поиска sSQP и правило Армихо для функции φ_{c_1, c_2} при любых фиксированных параметрах $c_1 > 0$ и $c_2 > 0$, если значение c_1 достаточно велико. Однако, для достижения данной общности требуются более сильные предположения: SOSC (52) вместо более слабого требования не критичности предельного множителя $\bar{\lambda}$, а также существование третьих производных у функции f и отображения h в окрестности точки \bar{x} . Отметим, что эти результаты не могут быть получены с использованием теоремы Денниса-Море, поскольку матрица Гессе функции $\varphi''_{c_1, c_2}(\bar{x}, \bar{\lambda})$ (которая существует, если f и h трижды дифференцируемы) обязательно вырождена, если $\text{rank } h'(\bar{x}) < l$ (см. предложение 5 ниже). Вместо этого используется следствие 2, которое допускает вырожденность этой матрицы.

Следующее предложение представляет собой вариант [4, теорема 4.16 (а)], в котором не требуется регулярность ограничений.

Предложение 5. Пусть $f: \mathbb{R}^n \rightarrow \mathbb{R}$ и $h: \mathbb{R}^n \rightarrow \mathbb{R}^l$ трижды дифференцируемы в точке $\bar{x} \in \mathbb{R}^n$. Пусть \bar{x} — стационарная точка задачи (49), а отвечающий ей множитель Лагранжа $\bar{\lambda} \in \mathbb{R}^l$ удовлетворяет SOSC (52).

Тогда для любого $\bar{c}_2 > 0$ существует такое $\bar{c}_1 \geq 0$, что для любых $c_1 \geq \bar{c}_1$ и $c_2 \geq \bar{c}_2$ матрица Гессе $\varphi''_{c_1, c_2}(\bar{x}, \bar{\lambda})$ неотрицательно определена, причем $\langle \varphi''_{c_1, c_2}(\bar{x}, \bar{\lambda})(\xi, \eta), (\xi, \eta) \rangle = 0$ тогда и только тогда, когда $\xi = 0$ и $\eta \in \ker(h'(\bar{x}))^T$.

Доказательство. Дифференцируя (60) и используя предположение о том, что $(\bar{x}, \bar{\lambda})$ удовлетворяет системе Лагранжа (50), получаем для всех c_1 и c_2 равенство

$$\varphi''_{c_1, c_2}(\bar{x}, \bar{\lambda}) = \begin{pmatrix} \frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda}) + c_1(h'(\bar{x}))^T h'(\bar{x}) & (h'(\bar{x}))^T \\ h'(\bar{x}) & 0 \end{pmatrix} + c_2 \begin{pmatrix} \left(\frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda})\right)^2 & \frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda})(h'(\bar{x}))^T \\ h'(\bar{x}) \frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda}) & h'(\bar{x})(h'(\bar{x}))^T \end{pmatrix}.$$

Вводя обозначение $x(\eta) = (h'(\bar{x}))^T \eta$, получаем

$$\langle \varphi''_{c_1, c_2}(\bar{x}, \bar{\lambda})(\xi, \eta), (\xi, \eta) \rangle = Q_{c_2}(\xi, x(\eta)) + c_1 R(\xi, x(\eta)), \quad (88)$$

где квадратичные формы $Q_{c_2}: \mathbb{R}^n \times \text{im}(h'(\bar{x}))^T \rightarrow \mathbb{R}$ и $R: \mathbb{R}^n \times \text{im}(h'(\bar{x}))^T \rightarrow \mathbb{R}$ определяются равенствами

$$Q_{c_2}(\xi, x) = \left\langle \frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda})\xi, \xi \right\rangle + 2\langle \xi, x \rangle + c_2 \left\| \frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda})\xi + x \right\|^2, \\ R(\xi, x) = \|h'(\bar{x})\xi\|^2.$$

Если $R(\xi, x) = 0$ для некоторого $(\xi, x) \in (\mathbb{R}^n \times \text{im}(h'(\bar{x}))^T) \setminus \{(0, 0)\}$, то $\xi \in \ker h'(\bar{x}) = (\text{im}(h'(\bar{x}))^T)^\perp$, и поэтому, $\langle \xi, x \rangle = 0$. Таким образом,

$$Q_{c_2}(\xi, x) = \left\langle \frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda})\xi, \xi \right\rangle + c_2 \left\| \frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda})\xi + x \right\|^2.$$

Если $\xi \neq 0$, то неравенство $Q_{c_2}(\xi, x) > 0$ следует из SOS (52). Если $\xi = 0$, но $x \neq 0$, то $Q_{c_2}(\xi, x) = c_2 \|x\|^2 > 0$ для любого $c_2 > 0$. Учитывая очевидную неотрицательную определенность R и применяя лемму Финслера–Дебре (см., например, [4, лемма 1.25]), получаем, что для любого $c_2 > 0$ найдется такое $c_1 \geq 0$, что квадратичная форма $Q_{c_2} + c_1 R$ будет положительно определена.

Таким образом, согласно (88), значение $\langle \varphi''_{c_1, c_2}(\bar{x}, \bar{\lambda})(\xi, \eta), (\xi, \eta) \rangle$ всегда положительно за исключением случая $\xi = 0$ и $x(\eta) = 0$. Для завершения доказательства остается заметить, что увеличение c_1 или c_2 может только увеличить это значение. ■

Теорема 9. Пусть $f: \mathbb{R}^n \rightarrow \mathbb{R}$ и $h: \mathbb{R}^n \rightarrow \mathbb{R}^l$ трижды дифференцируемы в окрестности точки $\bar{x} \in \mathbb{R}^n$, причем их третьи производные непрерывны в этой точке. Пусть \bar{x} — стационарная точка задачи (49), а отвечающий ей множитель Лагранжа $\bar{\lambda} \in \mathbb{R}^l$ удовлетворяет SOS (52). Положим $\varepsilon \in (0, 1/2)$, $\theta \in (0, 1)$ и рассмотрим итерационную схему $(x^{k+1}, \lambda^{k+1}) = (x^k, \lambda^k) + \alpha_k d^k$, где для всех достаточно больших номеров k направление поиска $d^k = (\xi^k, \eta^k)$ является решением системы (53) при $\sigma = \|\Phi(x^k, \lambda^k)\|$ и $(x, \lambda) = (x^k, \lambda^k)$,

а $\alpha_k = \theta^j$, где j — наименьшее неотрицательное целое число, для которого выполнено неравенство Армихо (68) для некоторых фиксированных $c_1 \geq 0$ и $c_2 \geq 0$.

Тогда для любого $\bar{c}_2 > 0$ найдется константа $\bar{c}_1 \geq 0$, обладающая следующим свойством: для любых $c_1 \geq \bar{c}_1$ и $c_2 \geq \bar{c}_2$ и для любой последовательности $\{(x^k, \lambda^k)\} \subset (\mathbb{R}^n \times \mathbb{R}^l) \setminus (\{\bar{x}\} \times \mathcal{M}(\bar{x}))$, генерируемой описанной выше итерационной схемой и сходящейся к $(\bar{x}, \bar{\lambda})$, выполнено $\alpha_k = 1$ для всех достаточно больших k , и скорость сходимости является сверхлинейной.

Доказательство. Для произвольного $\bar{c}_2 > 0$ выберем $\bar{c}_1 \geq 0$ таким образом, чтобы выполнялось утверждение предложения 5, и для любых $c_1 \geq \bar{c}_1$ и $c_2 \geq \bar{c}_2$ рассмотрим произвольную траекторию $\{(x^k, \lambda^k)\}$ с указанными выше свойствами.

Согласно теореме 1.3 б), полный шаг $d^k = (\xi^k, \eta^k)$ метода sSQP удовлетворяет оценке

$$\text{dist}((x^k, \lambda^k) + d^k, \{\bar{x}\} \times \mathcal{M}(\bar{x})) = o(\text{dist}((x^k, \lambda^k), \{\bar{x}\} \times \mathcal{M}(\bar{x}))) \quad (89)$$

при $k \rightarrow \infty$, а значит выполнена оценка (75) при $u^k = (x^k, \lambda^k)$ и $\bar{u} = (\bar{x}, \bar{\lambda})$.

Покажем, что для траектории $\{(x^k, \lambda^k)\}$ будет выполнено неравенство

$$\langle \varphi''_{c_1, c_2}(\bar{x}, \bar{\lambda}) d^k, d^k \rangle \geq \gamma \|d^k\|^2 \quad (90)$$

для некоторого $\gamma > 0$ и для всех достаточно больших k . Согласно предложению 5, достаточно показать, что последовательность $\{(\xi^k, \eta^k)/\|(\xi^k, \eta^k)\|\}$ не имеет предельных точек вида $(0, \eta)$, где $\eta \in \ker(h'(\bar{x}))^T$. Предположим, напротив, что такие предельные точки существуют. Без ограничения общности будем полагать, что последовательность $\{(\xi^k, \eta^k)/\|(\xi^k, \eta^k)\|\}$ сходится к $(0, \eta)$, откуда автоматически следует $\|\eta\| = 1$.

Обозначим через \bar{P} ортогональный проектор на подпространство $(\text{im } h'(\bar{x}))^\perp$. Тогда $\bar{P}h'(\bar{x}) = 0$, и поэтому из (62) и гладкости отображения h следует оценка

$$\begin{aligned} \sigma_k \bar{P}\eta^k &= \bar{P}(h(x^k) + h'(x^k)\xi^k) = \\ &= \bar{P}(h'(\bar{x})(x^k - \bar{x}) + h'(\bar{x})\xi^k + (h'(x^k) - h'(\bar{x}))\xi^k) + O(\|x^k - \bar{x}\|^2) = \\ &= O(\|\xi^k\| \|x^k - \bar{x}\|) + O(\|x^k - \bar{x}\|^2) \end{aligned} \quad (91)$$

при $k \rightarrow \infty$.

Покажем теперь, что

$$x^k - \bar{x} = o(\|\eta^k\|) \quad (92)$$

при $k \rightarrow \infty$. Действительно, если это не так, то без ограничения общности можно считать, что выполнено $\eta^k = O(\|x^k - \bar{x}\|)$, и поэтому из соотношения $\xi^k = o(\|\eta^k\|)$ получаем, что $\xi^k = o(\|x^k - \bar{x}\|)$. Значит, существует такая константа $\mu > 0$, что неравенство

$$\|x^k + \xi^k - \bar{x}\| \geq \mu \|x^k - \bar{x}\|$$

выполнено для всех достаточно больших k . Таким образом, из (89) следует соотношение

$$\|x^k - \bar{x}\| + \text{dist}(\lambda^k + \eta^k, \mathcal{M}(\bar{x})) = o(\|x^k - \bar{x}\| + \text{dist}(\lambda^k, \mathcal{M}(\bar{x}))),$$

из которого, очевидно, вытекает оценка

$$\|x^k - \bar{x}\| + \text{dist}(\lambda^k + \eta^k, \mathcal{M}(\bar{x})) = o(\text{dist}(\lambda^k, \mathcal{M}(\bar{x}))) \quad (93)$$

при $k \rightarrow \infty$. Поскольку

$$\text{dist}(\lambda^k, \mathcal{M}(\bar{x})) \leq \text{dist}(\lambda^k + \eta^k, \mathcal{M}(\bar{x})) + \|\eta^k\|,$$

оценка (93) означает, что $\text{dist}(\lambda^k, \mathcal{M}(\bar{x})) = O(\|\eta^k\|)$, откуда следует выполнение (92), что противоречит предположению.

Комбинируя (91), (92) и соотношение $\xi^k = o(\|\eta^k\|)$, получаем оценку

$$\sigma_k \bar{P} \eta^k = o(\|\eta^k\|^2) = o(\sigma_k \|\eta^k\|)$$

при $k \rightarrow \infty$, где последнее равенство следует из (82). Отсюда и из условия $\sigma_k \neq 0$ получаем $\bar{P} \eta = 0$ или, что то же самое, $\eta \in \ker \bar{P} = \text{im } h'(\bar{x})$. С другой стороны $\eta \in \ker(h'(\bar{x}))^T = (\text{im } h'(\bar{x}))^\perp$, что возможно только если $\eta = 0$. Таким образом, получено противоречие, которое и устанавливает неравенство (90).

Для завершения доказательства остается применить следствие 2. ■

Напомним, что при доказательстве следствия 2 было попутно доказано выполнение (73) для некоторого $\gamma > 0$ и для всех достаточно больших k . Отсюда следует, что направления sSQP действительно могут использоваться в методе спуска из теоремы 9 для достаточно больших номеров k .

Глава 4

Метод последовательного квадратичного программирования, стабилизированный вдоль подпространства

В предыдущей главе было предложено несколько различных подходов к глобализации сходимости стабилизированного метода последовательного квадратичного программирования, но ни один из них не позволяет достичь уверенного преимущества по сравнению с обычным SQP на задачах с нерегулярными ограничениями. Однако, как показано в приложении Б.3, основная причина этого связана не с недостатками конкретных подходов к глобализации сходимости, а с определенным нежелательным поведением самого sSQP. А именно, на задачах с неполным вырождением sSQP часто генерирует длинные последовательности коротких шагов, что приводит к медленному движению вдали от решения.

В настоящей главе предлагается принципиально новый метод последовательного квадратичного программирования, стабилизированный вдоль подпространства (s-sSQP, от англ. Subspace-stabilized SQP). Такое название связано с тем, что в s-sSQP двойственная траектория стабилизируется только вдоль подпространства вырожденности матрицы Якоби ограничений, что позволяет преодолеть описанный выше негативный эффект.

В разделе 4.1 приводится описание итерации метода s-sSQP и устанавливаются результаты о локальной сверхлинейной сходимости. В разделе 4.2 предлагается простая и эффективная процедура аппроксимации подпространства вырожденности. Раздел 4.3 посвящен глобализации сходимости метода на основе подхода, предложенного в разделе 3.3. Наконец, в приложении В 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.

4.1. Описание метода и результаты о локальной сходимости

Будем рассматривать задачу с ограничениями-равенствами

$$f(x) \rightarrow \min, \quad h(x) = 0, \quad (1)$$

где $f: \mathbb{R}^n \rightarrow \mathbb{R}$ — дважды дифференцируемая функция, а $h: \mathbb{R}^n \rightarrow \mathbb{R}^l$ — дважды дифференцируемое отображение. Напомним, что стационарные точки и отвечающие им множители Лагранжа задачи (1) характеризуются системой Лагранжа

$$\frac{\partial L}{\partial x}(x, \lambda) = 0, \quad h(x) = 0, \quad (2)$$

где $L: \mathbb{R}^n \times \mathbb{R}^l \rightarrow \mathbb{R}$ — функция Лагранжа задачи (1):

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

Нарушение условия регулярности в допустимой точке \bar{x} задачи (1) означает, что множество $(\text{im } h'(\bar{x}))^\perp = \ker(h'(\bar{x}))^\text{T}$ является нетривиальным подпространством в \mathbb{R}^l . Назовем его *подпространством вырожденности* и обозначим через \bar{P} ортогональный проектор на данное подпространство (для фиксированного \bar{x}).

Опишем итерацию метода s-sSQP. Пусть заданы отображение $P: \mathbb{R}^n \times \mathbb{R}^l \rightarrow \mathbb{R}^{l \times l}$ и функция $\sigma: \mathbb{R}^n \times \mathbb{R}^l \rightarrow \mathbb{R}$. Линейный оператор $P(x, \lambda)$ представляет собой аппроксимацию проектора на подпространство вырожденности (в некотором смысле, см. ниже) при $(x, \lambda) \rightarrow (\bar{x}, \bar{\lambda})$. 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

$$\frac{\partial^2 L}{\partial x^2}(x, \lambda)\xi + (h'(x))^\text{T}\eta = -\frac{\partial L}{\partial x}(x, \lambda), \quad h'(x)\xi - \sigma(x, \lambda)P(x, \lambda)\eta = -h(x) \quad (3)$$

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

Возьмем произвольную стационарную точку \bar{x} задачи (1) и любой отвечающий ей множитель Лагранжа $\bar{\lambda}$. Для доказательства локальной сходимости описанного выше метода воспользуемся результатами для абстрактной итерационной схемы из [50] и применим их к системе Лагранжа (2) и итеративной процедуре, шаг которой определяется системой (3).

Таким образом, необходимо проверить предположения 1–3 из [50, теорема 1]. Согласно предположению 1.1, предположение 1 эквивалентно оценке расстояния

$$\|x - \bar{x}\| + \text{dist}(\lambda, \mathcal{M}(\bar{x})) = O\left(\left\|\left(\frac{\partial L}{\partial x}(x, \lambda), h(x)\right)\right\|\right) \quad (4)$$

при $(x, \lambda) \rightarrow (\bar{x}, \bar{\lambda})$, которая, в свою очередь, эквивалентна некритичности множителя $\bar{\lambda}$ (см. также [71, замечание 1], [71, предложение 1]).

Проверка предположений 2 и 3 из [50] потребует введения дополнительных предположений на P и σ . Ниже рассматриваются два различных набора таких предположений.

4.1.1. Асимптотически исчезающая стабилизация

Пусть P и σ удовлетворяют следующим предположениям:

(P1) отображение P непрерывно в точке $(\bar{x}, \bar{\lambda})$, и подпространство $(\text{im } h'(\bar{x}))^\perp$ инвариантно относительно $P(\bar{x}, \bar{\lambda})$, т. е.,

$$P(\bar{x}, \bar{\lambda})\eta = \eta \quad \forall \eta \in (\text{im } h'(\bar{x}))^\perp;$$

(S1) функция σ непрерывна в каждой точке из $\{\bar{x}\} \times \mathcal{M}(\bar{x})$, достаточно близкой к $(\bar{x}, \bar{\lambda})$, $\sigma(\bar{x}, \lambda) = 0$ для всех $\lambda \in \mathcal{M}(\bar{x})$, достаточно близких к $\bar{\lambda}$, $\sigma(x, \lambda) \neq 0$ для всех $(x, \lambda) \in (\mathbb{R}^n \times \mathbb{R}^l) \setminus (\{\bar{x}\} \times \mathcal{M}(\bar{x}))$, достаточно близких к $(\bar{x}, \bar{\lambda})$, и

$$\|x - \bar{x}\| = O(|\sigma(x, \lambda)|) \quad (5)$$

при $(x, \lambda) \rightarrow (\bar{x}, \bar{\lambda})$.

Заметим, что предположение (P1) позволяет выбрать, например, $P(\cdot) \equiv I$, т. е. оно выполнено для обычного sSQP. Другие способы выбора P будут обсуждаться далее. Заметим также, что предположение (S1) выполнено, если множитель $\bar{\lambda}$ является некритическим, а σ определяется равенством

$$\sigma(x, \lambda) = \|\Phi(x, \lambda)\|^\beta \quad (6)$$

при любом фиксированном $\beta \in (0, 1]$, где $\Phi: \mathbb{R}^n \times \mathbb{R}^l \rightarrow \mathbb{R}^n \times \mathbb{R}^l$ — оператор системы Лагранжа:

$$\Phi(x, \lambda) = \left(\frac{\partial L}{\partial x}(x, \lambda), h(x)\right). \quad (7)$$

Предположение 2 в [50] характеризует требуемое качество аппроксимации исходной задачи итерационными подзадачами. Определим отображение $\mathcal{A}: (\mathbb{R}^n \times \mathbb{R}^l) \times (\mathbb{R}^n \times \mathbb{R}^l) \rightarrow \mathbb{R}^n \times \mathbb{R}^l$,

$$\mathcal{A}((x, \lambda), (\xi, \eta)) = \left(\frac{\partial L}{\partial x}(x, \lambda) + \frac{\partial^2 L}{\partial x^2}(x, \lambda)\xi + (h'(x))^T \eta, h(x) + h'(x)\xi - \sigma(x, \lambda)P(x, \lambda)\eta\right). \quad (8)$$

Заметим, что подзадача (3) может быть записана в виде

$$\mathcal{A}((x, \lambda), (\xi, \eta)) = 0.$$

Требуемое качество аппроксимации Φ (отображением \mathcal{A}) устанавливается следующим предложением.

Предложение 1. Пусть $f: \mathbb{R}^n \rightarrow \mathbb{R}$ и $h: \mathbb{R}^n \rightarrow \mathbb{R}^l$ дважды дифференцируемы в окрестности точки $\bar{x} \in \mathbb{R}^n$, а их вторые производные непрерывны в этой точке. Пусть \bar{x} — стационарная точка задачи (1), и пусть $\bar{\lambda} \in \mathcal{M}(\bar{x})$. Предположим, что отображение $P: \mathbb{R}^n \times \mathbb{R}^l \rightarrow \mathbb{R}^{l \times l}$ ограничено в окрестности $(\bar{x}, \bar{\lambda})$, а функция $\sigma: \mathbb{R}^n \times \mathbb{R}^l \rightarrow \mathbb{R}$ непрерывна в каждой точке множества $\{\bar{x}\} \times \mathcal{M}(\bar{x})$, достаточно близкой к $(\bar{x}, \bar{\lambda})$, и $\sigma(\bar{x}, \lambda) = 0$ для всех $\lambda \in \mathcal{M}(\bar{x})$, достаточно близких к $\bar{\lambda}$.

Тогда найдется такое $\varepsilon > 0$, что для любого $C > 0$ существует такая функция $\omega: \mathbb{R}_+ \rightarrow \mathbb{R}_+$, что $\omega(t) = o(t)$ при $t \rightarrow 0$, $\omega(0) = 0$, и для отображений Φ и \mathcal{A} , определяемых равенствами (7) и (8), выполнено

$$\|\Phi(x + \xi, \lambda + \eta) - \mathcal{A}((x, \lambda), (\xi, \eta))\| \leq \omega(\|x - \bar{x}\| + \text{dist}(\lambda, \mathcal{M}(\bar{x}))) \quad (9)$$

для всех $(x, \lambda) \in \mathcal{B}((\bar{x}, \bar{\lambda}), \varepsilon)$ и всех $(\xi, \eta) \in \mathbb{R}^n \times \mathbb{R}^l$, удовлетворяющих неравенству

$$\|(\xi, \eta)\| \leq C(\|x - \bar{x}\| + \text{dist}(\lambda, \mathcal{M}(\bar{x}))). \quad (10)$$

Доказательство. Зафиксируем такую константу $\varepsilon > 0$, что f и h дважды дифференцируемы на множестве $\mathcal{B}((\bar{x}, \bar{\lambda}), \varepsilon)$, величина $\|P(x, \lambda)\|$ ограничена некоторой фиксированной константой для всех $(x, \lambda) \in \mathcal{B}((\bar{x}, \bar{\lambda}), \varepsilon)$, функция σ непрерывна в каждой точке $(\{\bar{x}\} \times \mathcal{M}(\bar{x})) \cap \mathcal{B}((\bar{x}, \bar{\lambda}), \varepsilon)$, и выполнено $\sigma(\bar{x}, \lambda) = 0$ для всех $\lambda \in \mathcal{M}(\bar{x}) \cap \mathcal{B}(\bar{\lambda}, \varepsilon)$.

Очевидно,

$$\begin{aligned} \|\Phi(x + \xi, \lambda + \eta) - \mathcal{A}((x, \lambda), (\xi, \eta))\| &\leq \\ &\leq \left\| \frac{\partial L}{\partial x}(x + \xi, \lambda + \eta) - \frac{\partial L}{\partial x}(x, \lambda) - \frac{\partial^2 L}{\partial x^2}(x, \lambda)\xi - (h'(x))^T \eta \right\| + \\ &\quad + \|h(x + \xi) - h(x) - h'(x)\xi\| + |\sigma(x, \lambda)| \|P(x, \lambda)\eta\|. \end{aligned}$$

Таким образом, достаточно найти функции $\omega_i: \mathbb{R}_+ \rightarrow \mathbb{R}_+$, удовлетворяющие $\omega_i(t) = o(t)$ для $i = 1, 2, 3$, и для которых выполнено

$$\left\| \frac{\partial L}{\partial x}(x + \xi, \lambda + \eta) - \frac{\partial L}{\partial x}(x, \lambda) - \frac{\partial^2 L}{\partial x^2}(x, \lambda)\xi - (h'(x))^T \eta \right\| \leq \omega_1(\|x - \bar{x}\| + \text{dist}(\lambda, \mathcal{M}(\bar{x}))), \quad (11)$$

$$\|h(x + \xi) - h(x) - h'(x)\xi\| \leq \omega_2(\|x - \bar{x}\| + \text{dist}(\lambda, \mathcal{M}(\bar{x}))), \quad (12)$$

$$|\sigma(x, \lambda)| \|P(x, \lambda)\eta\| \leq \omega_3(\|x - \bar{x}\| + \text{dist}(\lambda, \mathcal{M}(\bar{x}))) \quad (13)$$

для всех $(x, \lambda) \in \mathcal{B}((\bar{x}, \bar{\lambda}), \varepsilon)$ и всех (ξ, η) , удовлетворяющих (10).

Положим

$$R((x, \lambda), (\xi, \eta)) = \frac{\partial L}{\partial x}(x + \xi, \lambda + \eta) - \frac{\partial L}{\partial x}(x, \lambda) - \frac{\partial^2 L}{\partial x^2}(x, \lambda)\xi - (h'(x))^T \eta$$

и определим функцию ω_1 следующим образом:

$$\omega_1(t) = \sup \left\{ \|R((x, \lambda), (\xi, \eta))\| \left| \begin{array}{l} (x, \lambda) \in \mathcal{B}((\bar{x}, \bar{\lambda}), \varepsilon), (\xi, \eta) \in \mathbb{R}^n \times \mathbb{R}^l, \\ \|x - \bar{x}\| + \text{dist}(\lambda, \mathcal{M}(\bar{x})) \leq t, \\ \|(\xi, \eta)\| \leq C(\|x - \bar{x}\| + \text{dist}(\lambda, \mathcal{M}(\bar{x}))) \end{array} \right. \right\}. \quad (14)$$

Заметим, что множество, по которому берется точная верхняя грань, непусто и компактно для любого фиксированного $t \geq 0$. Таким образом, функция ω_1 корректно определена, и для любого $t \geq 0$ в этом множестве существуют векторы (x, λ) и (ξ, η) , для которых выполнено $\omega_1(t) = \|R((x, \lambda), (\xi, \eta))\|$.

Очевидно, функция ω_1 удовлетворяет (11) для всех $(x, \lambda) \in \mathcal{B}((\bar{x}, \bar{\lambda}), \varepsilon)$ и всех (ξ, η) , удовлетворяющих (10). Остается показать, что $\omega_1(t) = o(t)$ при $t \rightarrow 0$. Предположим, что это не так, т. е. существует $\gamma > 0$ и сходящаяся к нулю последовательность положительных чисел $\{t_k\}$, для которых выполнено $\omega_1(t_k) \geq \gamma t_k$ для всех k . Последнее означает существование таких последовательностей $\{(x^k, \lambda^k)\} \subset \mathcal{B}((\bar{x}, \bar{\lambda}), \varepsilon)$ и $\{(\xi^k, \eta^k)\} \subset \mathbb{R}^n \times \mathbb{R}^l$, что для всех k имеет место

$$\|R((x^k, \lambda^k), (\xi^k, \eta^k))\| \geq \gamma t_k, \quad (15)$$

$$\|x^k - \bar{x}\| + \text{dist}(\lambda^k, \mathcal{M}(\bar{x})) \leq t_k, \quad (16)$$

$$\|(\xi^k, \eta^k)\| \leq C(\|x^k - \bar{x}\| + \text{dist}(\lambda^k, \mathcal{M}(\bar{x}))). \quad (17)$$

Из этих соотношений и теоремы о среднем выводим цепочку неравенств

$$\begin{aligned} \gamma t_k &\leq \left\| \frac{\partial L}{\partial x}(x^k + \xi^k, \lambda^k + \eta^k) - \frac{\partial L}{\partial x}(x^k, \lambda^k) - \frac{\partial^2 L}{\partial x^2}(x^k, \lambda^k)\xi^k - (h'(x^k))^T \eta^k \right\| \leq \\ &\leq \left\| \frac{\partial L}{\partial x}(x^k + \xi^k, \lambda^k) - \frac{\partial L}{\partial x}(x^k, \lambda^k) - \frac{\partial^2 L}{\partial x^2}(x^k, \lambda^k)\xi^k \right\| + \|(h'(x^k + \xi^k) - h'(x^k))^T \eta^k\| \leq \\ &\leq \sup_{\tau \in [0, 1]} \left\| \frac{\partial^2 L}{\partial x^2}(x^k + \tau \xi^k, \lambda^k) - \frac{\partial^2 L}{\partial x^2}(x^k, \lambda^k) \right\| \|\xi^k\| + \sup_{\tau \in [0, 1]} \|h''(x^k + \tau \xi^k)\| \|\xi^k\| \|\eta^k\| = \\ &= o(\|\xi^k\|) + O(\|\xi^k\| \|\eta^k\|) = o(t_k) \end{aligned}$$

при $k \rightarrow \infty$, получая противоречие.

Непрерывность h'' в точке \bar{x} означает выполнение неравенства $h''(x) \leq M$ для некоторой константы $M > 0$ и всех x , достаточно близких к \bar{x} . Вновь применяя теорему о среднем, получаем, что

$$\begin{aligned} \|h(x + \xi) - h(x) - h'(x)\xi\| &\leq \sup_{\tau \in [0, 1]} \|h'(x + \tau \xi) - h'(x)\| \|\xi\| \leq \sup_{\tau \in [0, 1]} \|h''(x + \tau \xi)\| \|\xi\|^2 \leq \\ &\leq \omega_2(\|x - \bar{x}\| + \text{dist}(\lambda, \mathcal{M}(\bar{x}))) \end{aligned}$$

для всех (x, λ) , достаточно близких к $(\bar{x}, \bar{\lambda})$, и всех (ξ, η) , для которых выполнено (10), где функция ω_2 определяется равенством $\omega_2(t) = C^2 M t^2$. Очевидно, что при $t \rightarrow 0$ имеет место $\omega_2(t) = o(t)$, откуда следует справедливость (12) для всех $(x, \lambda) \in \mathcal{B}((\bar{x}, \bar{\lambda}), \varepsilon)$ и всех (ξ, η) , удовлетворяющих (10) (возможно, после уменьшения $\varepsilon > 0$).

Наконец, положим

$$R((x, \lambda), \eta) = \sigma(x, \lambda)P(x, \lambda)\eta \quad (18)$$

и рассмотрим функцию ω_3 , определяемую равенством (14), в котором $R((x, \lambda), (\xi, \eta))$ заменено на только что введенное выражение $R((x, \lambda), \eta)$. Очевидно, функция ω_3 удовлетворяет (13) для всех $(x, \lambda) \in \mathcal{B}((\bar{x}, \bar{\lambda}), \varepsilon)$ и всех (ξ, η) , для которых выполнено (10), и поэтому остается показать, что $\omega_3(t) = o(t)$ при $t \rightarrow 0$. Предполагая противное, получаем существование константы $\gamma > 0$, сходящейся к нулю последовательности положительных чисел $\{t_k\}$ и последовательностей $\{(x^k, \lambda^k)\} \subset \mathcal{B}((\bar{x}, \bar{\lambda}), \varepsilon)$ и $\{(\xi^k, \eta^k)\} \subset \mathbb{R}^n \times \mathbb{R}^l$, таких что неравенства (15)–(17) выполнены для всех k .

Согласно (16) и (17), без ограничения общности можно считать, что последовательность $\{(x^k, \lambda^k)\}$ сходится к (\bar{x}, λ^*) при некотором $\lambda^* \in \mathcal{M}(\bar{x}) \cap \mathcal{B}(\bar{\lambda}, \varepsilon)$, и что $\|\eta^k\| \leq C t_k$ для всех k . Выбор ε подразумевает, что величины $\|P(x^k, \lambda^k)\|$ образуют ограниченную последовательность, $\sigma(\bar{x}, \lambda^*) = 0$, а функция σ непрерывна в точке (\bar{x}, λ^*) . Отсюда легко получить оценку

$$\gamma t_k \leq |\sigma(x^k, \lambda^k)| \|P(x^k, \lambda^k)\| \|\eta^k\| = o(t_k),$$

при $k \rightarrow \infty$, образующую противоречие. ■

Очевидно, из предположений (P1) и (S1) следует выполнение требований для P и σ в только что доказанном предложении.

Наконец, предположение 3 из [50] выполнено, если для всех $(x, \lambda) \in \mathbb{R}^n \times \mathbb{R}^l$, достаточно близких к $(\bar{x}, \bar{\lambda})$, система (3) имеет решение (ξ, η) , удовлетворяющее оценке

$$\|(\xi, \eta)\| = O(\|x - \bar{x}\| + \text{dist}(\lambda, \mathcal{M}(\bar{x}))) \quad (19)$$

при $(x, \lambda) \rightarrow (\bar{x}, \bar{\lambda})$. Справедливость этого предположения устанавливается следующим утверждением.

Предложение 2. Пусть $f: \mathbb{R}^n \rightarrow \mathbb{R}$ и $h: \mathbb{R}^n \rightarrow \mathbb{R}^l$ дважды дифференцируемы в окрестности точки $\bar{x} \in \mathbb{R}^n$, а их вторые производные непрерывны в этой точке. Пусть \bar{x} — стационарная точка задачи (1), и пусть множитель $\bar{\lambda} \in \mathcal{M}(\bar{x})$ является некритическим. Предположим, что $P: \mathbb{R}^n \times \mathbb{R}^l \rightarrow \mathbb{R}^{l \times l}$ и $\sigma: \mathbb{R}^n \times \mathbb{R}^l \rightarrow \mathbb{R}$ удовлетворяют предположениям (P1) и (S1).

Тогда для всех $(x, \lambda) \in (\mathbb{R}^n \times \mathbb{R}^l) \setminus (\{\bar{x}\} \times \mathcal{M}(\bar{x}))$, достаточно близких к $(\bar{x}, \bar{\lambda})$, система (3) имеет единственное решение (ξ, η) , и это решение удовлетворяет оценке (19).

Доказательство. Покажем сначала, что для всех $(x, \lambda) \in (\mathbb{R}^n \times \mathbb{R}^l) \setminus (\{\bar{x}\} \times \mathcal{M}(\bar{x}))$, достаточно близких к $(\bar{x}, \bar{\lambda})$, матрица

$$\begin{pmatrix} \frac{\partial^2 L}{\partial x^2}(x, \lambda) & (h'(x))^T \\ h'(x) & -\sigma(x, \lambda)P(x, \lambda) \end{pmatrix}$$

линейной системы (3) является невырожденной.

Предположим, что это не так. Тогда существуют такие последовательности $\{(x^k, \lambda^k)\} \subset (\mathbb{R}^n \times \mathbb{R}^l) \setminus (\{\bar{x}\} \times \mathcal{M}(\bar{x}))$ и $\{(\xi^k, \eta^k)\} \subset \mathbb{R}^n \times \mathbb{R}^l$, что $\{(x^k, \lambda^k)\} \rightarrow (\bar{x}, \bar{\lambda})$, и для всех k выполнены соотношения $\|(\xi^k, \eta^k)\| = 1$,

$$\frac{\partial^2 L}{\partial x^2}(x^k, \lambda^k)\xi^k + (h'(x^k))^T \eta^k = 0, \quad h'(x^k)\xi^k - \sigma(x^k, \lambda^k)P(x^k, \lambda^k)\eta^k = 0. \quad (20)$$

Без ограничения общности будем полагать, что $\{(\xi^k, \eta^k)\} \rightarrow (\xi, \eta)$. Тогда, переходя к пределу в соотношениях выше, получаем равенства $\|(\xi, \eta)\| = 1$ и

$$\frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda})\xi + (h'(\bar{x}))^T \eta = 0, \quad h'(\bar{x})\xi = 0, \quad (21)$$

последнее из которых следует из непрерывности P и σ в точке $(\bar{x}, \bar{\lambda})$ и из равенства $\sigma(\bar{x}, \bar{\lambda}) = 0$ (предположения (P1) и (S1)).

Если $\xi \neq 0$, то (21) противоречит предположению о некритичности $\bar{\lambda}$. Поэтому $\xi = 0$, а значит $\|\eta\| = 1$ и $(h'(\bar{x}))^T \eta = 0$. Далее, согласно предположению (P1), подпространство $(\text{im } h'(\bar{x}))^\perp = \ker(h'(\bar{x}))^T$ инвариантно относительно $P(\bar{x}, \bar{\lambda})$, откуда следует $P(\bar{x}, \bar{\lambda})\eta = \eta$.

Из второго равенства в (20) вытекает цепочка соотношений

$$\begin{aligned} \sigma(x^k, \lambda^k)\bar{P}P(x^k, \lambda^k)\eta^k &= \bar{P}h'(x^k)\xi^k = \bar{P}(h'(x^k) - h'(\bar{x}))\xi^k = O(\|x^k - \bar{x}\|\|\xi^k\|) = \\ &= O(|\sigma(x^k, \lambda^k)|\|\xi^k\|) \end{aligned}$$

при $k \rightarrow \infty$, последнее из которых следует из (5). Отсюда, из неравенства $\sigma(x^k, \lambda^k) \neq 0$, выполненного для всех k (предположение (S1)), и соотношения $\{\xi^k\} \rightarrow \xi = 0$ выводим равенства

$$\eta = \bar{P}\eta = \bar{P}P(\bar{x}, \bar{\lambda})\eta = \lim_{k \rightarrow \infty} \bar{P}P(x^k, \lambda^k)\eta^k = 0,$$

которые противоречат условию $\|\eta\| = 1$.

Остается доказать, что для всех $(x, \lambda) \in (\mathbb{R}^n \times \mathbb{R}^l) \setminus (\{\bar{x}\} \times \mathcal{M}(\bar{x}))$, достаточно близких к $(\bar{x}, \bar{\lambda})$, единственное решение (ξ, η) системы (3) удовлетворяет оценке (19).

Предположим противное: существуют такие последовательности $\{(x^k, \lambda^k)\} \subset \mathbb{R}^n \times \mathbb{R}^l$ и $\{(\xi^k, \eta^k)\} \subset \mathbb{R}^n \times \mathbb{R}^l$, что $\{(x^k, \lambda^k)\} \rightarrow (\bar{x}, \bar{\lambda})$, для всех k выполнено $(x^k, \lambda^k) \notin \{\bar{x}\} \times \mathcal{M}(\bar{x})$, $(\xi^k, \eta^k) \neq 0$,

$$\frac{\partial^2 L}{\partial x^2}(x^k, \lambda^k)\xi^k + (h'(x^k))^T \eta^k = -\frac{\partial L}{\partial x}(x^k, \lambda^k), \quad h'(x^k)\xi^k - \sigma(x^k, \lambda^k)P(x^k, \lambda^k)\eta^k = -h(x^k) \quad (22)$$

и

$$\|x^k - \bar{x}\| + \text{dist}(\lambda^k, \mathcal{M}(\bar{x})) = o(\|(\xi^k, \eta^k)\|) \quad (23)$$

при $k \rightarrow \infty$. Заметим, что из последнего соотношения следует

$$\begin{aligned} \left\| \left(\frac{\partial L}{\partial x}(x^k, \lambda^k), h(x^k) \right) \right\| &= \left\| \left(\frac{\partial L}{\partial x}(x^k, \lambda^k) - \frac{\partial L}{\partial x}(\bar{x}, \bar{\lambda}^k), h(x^k) - h(\bar{x}) \right) \right\| = \\ &= O(\|x^k - \bar{x}\| + \text{dist}(\lambda^k, \mathcal{M}(\bar{x}))) = o(\|(\xi^k, \eta^k)\|), \end{aligned} \quad (24)$$

где через $\bar{\lambda}^k$ обозначена проекция λ^k на множество $\mathcal{M}(\bar{x})$.

Без ограничения общности будем считать, что $\{(\xi^k, \eta^k)/\|(\xi^k, \eta^k)\|\} \rightarrow (\xi, \eta)$, откуда следует $\|(\xi, \eta)\| = 1$. Разделив (22) на $\|(\xi^k, \eta^k)\|$, перейдя к пределу и воспользовавшись (24), получаем (21). Аналогично тому, как это делалось выше, из некритичности $\bar{\lambda}$ получаем соотношения $\xi = 0$, $\|\eta\| = 1$, $(h'(\bar{x}))^T \eta = 0$ и $P(\bar{x}, \bar{\lambda})\eta = \eta$.

Из второго равенства в (22) следует, что

$$\begin{aligned} \sigma(x^k, \lambda^k) \bar{P}P(x^k, \lambda^k)\eta^k &= \bar{P}(h(x^k) + h'(x^k)\xi^k) = \\ &= \bar{P}(h(x^k) - h(\bar{x}) - h'(\bar{x})(x^k - \bar{x}) + (h'(x^k) - h'(\bar{x}))\xi^k) = \\ &= O(\|x^k - \bar{x}\|^2) + O(\|x^k - \bar{x}\|\|\xi^k\|) = o(|\sigma(x^k, \lambda^k)|\|(\xi^k, \eta^k)\|) + O(|\sigma(x^k, \lambda^k)|\|\xi^k\|) \end{aligned}$$

при $k \rightarrow \infty$, где второе равенство вытекает из теоремы о среднем, а последнее следует из (5) и (23). Таким образом, поскольку $\sigma(x^k, \lambda^k) \neq 0$ для всех k , $\sigma(x^k, \lambda^k) \rightarrow 0$ при $k \rightarrow \infty$ (предположение (S1)), и $\{\xi^k/\|(\xi^k, \eta^k)\|\} \rightarrow \xi = 0$, вновь получаем цепочку равенств

$$\eta = \bar{P}\eta = \bar{P}P(\bar{x}, \bar{\lambda})\eta = \lim_{k \rightarrow \infty} \frac{\bar{P}P(x^k, \lambda^k)\eta^k}{\|(\xi^k, \eta^k)\|} = 0,$$

которая противоречит $\|\eta\| = 1$. ■

Отметим, что если $x = \bar{x}$ и множитель $\lambda \in \mathcal{M}(\bar{x})$ достаточно близок к $\bar{\lambda}$, то система (3) всегда имеет нулевое решение, которое, очевидно, удовлетворяет (19).

С учетом предложений 1 и 2 и сказанного выше, из [50, теорема 1] (а также некоторых соотношений из ее доказательства) вытекает следующий результат о локальной сходимости.

Теорема 1. Пусть выполнены предположения предложения 2. Тогда справедливы следующие утверждения:

- а) в некоторой окрестности \mathcal{V} точки $(\bar{x}, \bar{\lambda})$ существует единственное отображение $d(\cdot) = (\xi(\cdot), \eta(\cdot)): \mathcal{V} \rightarrow \mathbb{R}^n \times \mathbb{R}^l$ со следующими свойствами: $(\xi(x, \lambda), \eta(x, \lambda))$ удовлетворяет системе (3) для всех $(x, \lambda) \in \mathcal{V}$, и $d(\bar{x}, \bar{\lambda}) = 0$, если $\bar{\lambda} \in \mathcal{M}(\bar{x})$;

б) окрестность \mathcal{V} можно выбрать таким образом, что найдется константа $\ell > 0$ и непрерывная в нуле функция $\chi: \mathbb{R}_+ \rightarrow \mathbb{R}_+$, удовлетворяющая оценке $\chi(t) = o(t)$ при $t \rightarrow 0$, для которых будут справедливы оценки

$$\|d(x, \lambda)\| \leq \ell(\|x - \bar{x}\| + \text{dist}(\lambda, \mathcal{M}(\bar{x}))), \quad (25)$$

$$\|x + \xi(x, \lambda) - \bar{x}\| + \text{dist}(\lambda + \eta(x, \lambda), \mathcal{M}(\bar{x})) \leq \chi(\|x - \bar{x}\| + \text{dist}(\lambda, \mathcal{M}(\bar{x}))) \quad (26)$$

для всех $(x, \lambda) \in \mathcal{V}$;

в) существует $M > 0$, такое что для любого $\varepsilon > 0$ найдется такое $\varepsilon_0 > 0$, что для любого $(x^0, \lambda^0) \in \mathcal{B}((\bar{x}, \bar{\lambda}), \varepsilon_0)$ последовательность $\{(x^k, \lambda^k)\}$ корректно определяется равенствами $(x^{k+1}, \lambda^{k+1}) = (x^k, \lambda^k) + d(x^k, \lambda^k)$ для всех k , и для нее справедливо $\{(x^k, \lambda^k)\} \subset \mathcal{B}((\bar{x}, \bar{\lambda}), \varepsilon)$; эта последовательность сходится к (\bar{x}, λ^*) при некотором $\lambda^* = \lambda^*(x^0, \lambda^0) \in \mathcal{M}(\bar{x})$, и для всех k выполнено

$$\|x^{k+1} - \bar{x}\| + \|\lambda^{k+1} - \lambda^*\| \leq M\chi(\|x^k - \bar{x}\| + \text{dist}(\lambda^k, \mathcal{M}(\bar{x}))). \quad (27)$$

Заметим, что из (27) вытекают следующие оценки для всех k :

$$\|x^{k+1} - \bar{x}\| + \text{dist}(\lambda^{k+1}, \mathcal{M}(\bar{x})) \leq M\chi(\|x^k - \bar{x}\| + \text{dist}(\lambda^k, \mathcal{M}(\bar{x}))),$$

$$\|x^{k+1} - \bar{x}\| + \|\lambda^{k+1} - \lambda^*\| \leq M\chi(\|x^k - \bar{x}\| + \|\lambda^k - \lambda^*\|).$$

В частности, скорость сходимости последовательности $\{(x^k, \lambda^k)\}$ к (\bar{x}, λ^*) и последовательности $\{\|x^k - \bar{x}\| + \text{dist}(\lambda^k, \mathcal{M}(\bar{x}))\}$ к нулю является сверхлинейной.

4.1.2. Неисчезающая стабилизация

Пусть теперь P и σ удовлетворяют следующим предположениям:

(P2) отображение P непрерывно во всех точках множества $\{\bar{x}\} \times \mathcal{M}(\bar{x})$, достаточно близких к $(\bar{x}, \bar{\lambda})$; $\text{im } P(\bar{x}, \lambda) \cap \text{im } h'(\bar{x}) = \{0\}$ для всех $\lambda \in \mathcal{M}(\bar{x})$, достаточно близких к $\bar{\lambda}$, и $\ker P(\bar{x}, \bar{\lambda}) \cap (\text{im } h'(\bar{x}))^\perp = \{0\}$;

(S2) функция σ непрерывна в точке $(\bar{x}, \bar{\lambda})$, и $\sigma(\bar{x}, \bar{\lambda}) \neq 0$.

Предположения (P2) и (S2) отличаются от (P1) и (S1). Если P и σ удовлетворяют (P2) и (S2), то итерационный процесс, определенный в соответствии с (3), имеет много общего с методом из [66].

Согласно [50, замечание 2], предположение 2 из [50] достаточно проверить лишь для решений подзадач. Требуемый факт устанавливается следующим утверждением.

Предложение 3. Пусть $f: \mathbb{R}^n \rightarrow \mathbb{R}$ и $h: \mathbb{R}^n \rightarrow \mathbb{R}^l$ дважды дифференцируемы в окрестности точки $\bar{x} \in \mathbb{R}^n$, а их вторые производные непрерывны в этой точке. Пусть \bar{x} — стационарная точка задачи (1), и пусть $\bar{\lambda} \in \mathcal{M}(\bar{x})$. Предположим, что $P: \mathbb{R}^n \times \mathbb{R}^l \rightarrow \mathbb{R}^{l \times l}$ и $\sigma: \mathbb{R}^n \times \mathbb{R}^l \rightarrow \mathbb{R}$ удовлетворяют предположениям (P2) и (S2).

Тогда найдется такое $\varepsilon > 0$, что для любого $C > 0$ существует такая функция $\omega: \mathbb{R}_+ \rightarrow \mathbb{R}_+$, что $\omega(t) = o(t)$ при $t \rightarrow 0$, $\omega(0) = 0$, и для отображений Φ и \mathcal{A} , определяемых равенствами (7) и (8), выполнена оценка (9) для всех $(x, \lambda) \in \mathcal{B}((\bar{x}, \bar{\lambda}), \varepsilon)$ и всех $(\xi, \eta) \in \mathbb{R}^n \times \mathbb{R}^l$, удовлетворяющих (3) и (10).

Доказательство. Зафиксируем такую константу $\varepsilon > 0$, что f и h дважды дифференцируемы на множестве $\mathcal{B}((\bar{x}, \bar{\lambda}), \varepsilon)$, отображение P непрерывно в каждой точке множества $(\{\bar{x}\} \times \mathcal{M}(\bar{x})) \cap \mathcal{B}((\bar{x}, \bar{\lambda}), \varepsilon)$, и $\text{im } P(\bar{x}, \lambda) \cap \text{im } h'(\bar{x}) = \{0\}$ для всех $\lambda \in \mathcal{M}(\bar{x}) \cap \mathcal{B}(\bar{\lambda}, \varepsilon)$ (предположение (P2)).

Рассуждая аналогично тому, как это делалось в доказательстве предложения 1, легко убедиться, что достаточно доказать неравенство (13) для всех $(x, \lambda) \in \mathcal{B}((\bar{x}, \bar{\lambda}), \varepsilon)$ и всех (ξ, η) , удовлетворяющих (3) и (10), и для некоторой функции $\omega_3: \mathbb{R}_+ \rightarrow \mathbb{R}_+$, удовлетворяющей $\omega_3(t) = o(t)$ при $t \rightarrow 0$.

Для $R((x, \lambda), \eta)$, заданной равенством (18), положим

$$\omega_3(t) = \sup \left\{ \|R((x, \lambda), \xi)\| \left| \begin{array}{l} (x, \lambda) \in \mathcal{B}((\bar{x}, \bar{\lambda}), \varepsilon), \\ (\xi, \eta) \in \mathbb{R}^n \times \mathbb{R}^l \text{ удовлетворяет (3)}, \\ \|x - \bar{x}\| + \text{dist}(\lambda, \mathcal{M}(\bar{x})) \leq t, \\ \|(\xi, \eta)\| \leq C(\|x - \bar{x}\| + \text{dist}(\lambda, \mathcal{M}(\bar{x}))) \end{array} \right. \right\}.$$

Множество, по которому берется точная верхняя грань, компактно при любом фиксированном $t \geq 0$. Для тех $t \geq 0$, для которых это множество пусто, положим $\omega_3(t) = 0$, и в этом случае функция ω_3 будет корректно определена. Более того, для любого $t \geq 0$, для которого $\omega_3(t) \neq 0$, найдутся векторы (x, λ) и (ξ, η) в этом множестве, для которых выполнено $\omega_3(t) = \|R((x, \lambda), \eta)\|$.

Очевидно, что функция ω_3 удовлетворяет (13) для всех $(x, \lambda) \in \mathcal{B}((\bar{x}, \bar{\lambda}), \varepsilon)$ и всех (ξ, η) , для которых выполнено (3) и (10). Остается показать, что $\omega_3(t) = o(t)$ при $t \rightarrow 0$. Предположим, что это не так: существует $\gamma > 0$ и сходящаяся к нулю последовательность положительных чисел $\{t_k\}$, для которых $\omega_3(t_k) \geq \gamma t_k$ для всех k . Отсюда получаем существование таких последовательностей $\{(x^k, \lambda^k)\} \subset (\mathbb{R}^n \times \mathbb{R}^l) \cap \mathcal{B}((\bar{x}, \bar{\lambda}), \varepsilon)$ и $\{(\xi^k, \eta^k)\} \subset \mathbb{R}^n \times \mathbb{R}^l$, что неравенства (15)–(17) и равенства из (22) выполняются для всех k .

Согласно (16) и (17), без ограничения общности можно предположить, что $\{(x^k, \lambda^k)\}$ сходится к (\bar{x}, λ^*) при некотором $\lambda^* \in \mathcal{M}(\bar{x})$, и что для всех k справедливо $\|(\xi^k, \eta^k)\| \leq C t_k$,

откуда следует $\|x^k + \xi^k - \bar{x}\| \leq (1 + C)t_k$. Поэтому, также не ограничивая общности, можно считать, что

$$\left\{ \frac{1}{t_k} \eta^k \right\} \rightarrow \eta, \quad \left\{ \frac{1}{t_k} (x^k + \xi^k - \bar{x}) \right\} \rightarrow \xi \quad (28)$$

для некоторых $\xi \in \mathbb{R}^n$ и $\eta \in \mathbb{R}^l$. Наконец, согласно предположению (S2), можно считать (возможно, уменьшив $\varepsilon > 0$), что $\sigma(x^k, \lambda^k) \rightarrow \bar{\sigma} \neq 0$ при $k \rightarrow \infty$.

Из второго равенства в (3) можно вывести соотношения

$$\begin{aligned} \sigma(x^k, \lambda^k) P(x^k, \lambda^k) \eta^k &= h(x^k) + h'(x^k) \xi^k = \\ &= (h(x^k) - h(\bar{x}) - h'(\bar{x})(x^k - \bar{x})) + (h'(x^k) - h'(\bar{x})) \xi^k + h'(\bar{x})(x^k + \xi^k - \bar{x}) = \\ &= h'(\bar{x})(x^k + \xi^k - \bar{x}) + O(\|x^k - \bar{x}\|^2) + O(\|x^k - \bar{x}\| \|\xi^k\|) = h'(\bar{x})(x^k + \xi^k - \bar{x}) + O(t_k^2), \end{aligned} \quad (29)$$

предпоследнее из которых следует из теоремы о среднем, а последнее — из (16) и (17). Разделив (29) на t_k , перейдя к пределу и воспользовавшись соотношениями (28) и предположением (P2), получаем, что

$$\bar{\sigma} P(\bar{x}, \lambda^*) \eta = h'(\bar{x}) \xi.$$

Поскольку $\text{im } P(\bar{x}, \lambda^*) \cap \text{im } h'(\bar{x}) = \{0\}$ (что следует из предположения (P2)), полученное равенство может выполняться только если $\bar{\sigma} P(\bar{x}, \lambda^*) \eta = 0$. С другой стороны, из (15) следует $\|\bar{\sigma} P(\bar{x}, \lambda^*) \eta\| \geq \gamma$, образуя противоречие. \blacksquare

Проводя рассуждение, аналогичное доказательству в [66, теорема 2.4], несложно убедиться, что если выполнены предположения (P2) и (S2), и множитель $\bar{\lambda}$ является некритическим, то матрица

$$\begin{pmatrix} \frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda}) & (h'(\bar{x}))^T \\ h'(\bar{x}) & -\sigma(\bar{x}, \bar{\lambda}) P(\bar{x}, \bar{\lambda}) \end{pmatrix}$$

невырождена, откуда немедленно следует справедливость предположения 3 из [50]. Более того, система (3) имеет единственное решение для всех $(x, \lambda) \in \mathbb{R}^n \times \mathbb{R}^l$, достаточно близких к $(\bar{x}, \bar{\lambda})$.

Вновь применяя [50, теорема 1], получаем следующий результат о локальной сходимости.

Теорема 2. *В предположениях предложения 2, но с заменой (P1) и (S1) на (P2) и (S2) соответственно, утверждение теоремы 1 остается справедливым.*

4.2. Аппроксимация подпространства вырожденности

Будем рассматривать следующую задачу. Пусть $\bar{A} \in \mathbb{R}^{l \times n}$ — некоторая матрица, для которой известен ее ранг $r = \text{rank } \bar{A}$. Необходимо построить такое отображение $U: \mathbb{R}^{l \times n} \rightarrow \mathbb{R}^{l \times (l-r)}$,

что для любой предельной точки \bar{U} отображения $U(\cdot)$ в точке \bar{A} выполнены соотношения $\text{rank } \bar{U} = l - r$ и $\bar{A}^T \bar{U} = 0$.

Следующий алгоритм для заданной матрицы $A \in \mathbb{R}^{l \times n}$ вычисляет матрицу $U(A) \in \mathbb{R}^{l \times (l-r)}$ таким образом, что получаемое отображение U обладает требуемыми свойствами.

Алгоритм 1. Полагаем $\mathcal{I}_0 = \emptyset$, $U^{(0)} = I \in \mathbb{R}^{l \times l}$, $A^{(0)} = A$, $s = 0$.

1. Если $s = r$, переходим к шагу 4. В противном случае находим индексы $i_s \in \{1, \dots, l\} \setminus \mathcal{I}_s$ и $j_s \in \{1, \dots, n\}$, для которых справедливо

$$|A_{i_s j_s}^{(s)}| = \max_{\substack{i \in \{1, \dots, l\} \setminus \mathcal{I}_s, \\ j \in \{1, \dots, n\}}} |A_{ij}^{(s)}|. \quad (30)$$

Если $A_{i_s j_s}^{(s)} = 0$, останавливаем работу алгоритма, считая запуск неудачным. В противном случае полагаем $\mathcal{I}_{s+1} = \mathcal{I}_s \cup \{i_s\}$.

2. Определяем матрицу преобразования $T^{(s)} \in \mathbb{R}^{l \times l}$,

$$T_{ij}^{(s)} = \begin{cases} -\frac{A_{jj_s}^{(s)}}{A_{i_s j_s}^{(s)}}, & \text{если } i = i_s, j \notin \mathcal{I}_{s+1}, \\ 1, & \text{если } i = j, \\ 0, & \text{в противном случае,} \end{cases} \quad i, j = 1, \dots, l. \quad (31)$$

Полагаем $U^{(s+1)} = U^{(s)} T^{(s)}$ и $A^{(s+1)} = (T^{(s)})^T A^{(s)}$.

3. Увеличиваем s на 1 и переходим к шагу 1.
4. Определяем $U(A)$ как матрицу, составленную из столбцов $U^{(r)}$ с номерами из множества $\{1, \dots, l\} \setminus \mathcal{I}_r$.

В принципе, $U(A)$ может зависеть от выбора i_s и j_s на шаге 1 (как и другие объекты, получаемые в результате работы алгоритма 1 и алгоритма 2 ниже), если максимум в (30) достигается на нескольких наборах индексов. Тем не менее, устанавливаемые ниже свойства справедливы для любой возможной реализации $U(A)$. Отметим также, что любая конкретная реализация этих алгоритмов будет вычислять одно и то же $U(A)$ для заданной матрицы A .

По построению для всех s выполнены равенства

$$U^{(s)} = U^{(s-1)} T^{(s-1)} = U^{(0)} T^{(0)} \dots T^{(s-1)} = T^{(0)} \dots T^{(s-1)}, \quad (32)$$

$$A^{(s)} = (T^{(s-1)})^T A^{(s-1)} = (T^{(0)} \dots T^{(s-1)})^T A^{(0)} = (U^{(s)})^T A, \quad (33)$$

и $\det T^{(s)} = 1$, а значит $\det U^{(s)} = 1$ и $\text{rank } A^{(s)} = \text{rank } A$.

Предположим, что запуск алгоритма 1 объявляется неудачным на шаге 1. В этом случае все строки матрицы $A^{(s)}$ с номерами из $\{1, \dots, l\} \setminus \mathcal{I}_s$ содержат лишь нулевые элементы,

откуда следует, что $\text{rank } A^{(s)} \leq s < r$. Поэтому если $\text{rank } A \geq r$ (что выполнено для всех A , достаточно близких к \bar{A}), то запуск алгоритма 1 всегда является успешным.

Далее, для каждого s по построению выполнены неравенства $|T_{ij}^{(s)}| \leq 1$ для всех $i, j = 1, \dots, l$. Таким образом, из (32) следует, что величина $\|U^{(r)}\|$ (а значит и $\|U(A)\|$) ограничена некоторой константой, которая зависит от r , но не зависит от матрицы A .

Лемма 1. *Если ранг матрицы $A \in \mathbb{R}^{l \times n}$ равен r , то для матрицы $A^{(r)}$, получаемой в ходе работы алгоритма 1, справедливо $A_{ij}^{(r)} = 0$ для всех $i \in \{1, \dots, l\} \setminus \mathcal{I}_r$ и $j \in \{1, \dots, n\}$.*

Доказательство. Для упрощения изложения будем считать, что для всех $s = 0, \dots, r-1$ выбираются индексы $i_s = s+1$ и $j_s = s+1$ (легко убедиться, что такое предположение не ограничивает общности). В этом случае матрица $A^{(s)}$ имеет следующий вид для всех $s = 0, \dots, r$:

$$A^{(s)} = \begin{pmatrix} A_{11}^{(0)} & A_{12}^{(0)} & \cdots & A_{1s}^{(0)} & A_{1(s+1)}^{(0)} & \cdots & A_{1n}^{(0)} \\ 0 & A_{22}^{(1)} & \cdots & A_{2s}^{(1)} & A_{2(s+1)}^{(1)} & \cdots & A_{2n}^{(1)} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_{ss}^{(s-1)} & A_{s(s+1)}^{(s-1)} & \cdots & A_{sn}^{(s-1)} \\ 0 & 0 & \cdots & 0 & A_{(s+1)(s+1)}^{(s)} & \cdots & A_{(s+1)n}^{(s)} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & A_{l(s+1)}^{(s)} & \cdots & A_{ln}^{(s)} \end{pmatrix}.$$

Заметим, что $A_{i_s j_s}^{(s)} \neq 0$ для таких s (в противном случае запуск алгоритма был бы неудачным, что невозможно, так как $\text{rank } A = r$), и поэтому $A_{ii}^{(s)} \neq 0$ для $i \in \{1, \dots, s\}$. Отсюда следует, что левый верхний $s \times s$ -минор матрицы $A^{(s)}$ является ненулевым.

Поэтому, если хотя бы в одной из строк матрицы $A^{(r)}$ с номерами от $r+1$ до l присутствуют ненулевые элементы, то ее ранг превышает r , что невозможно, поскольку $\text{rank } A = r$.

■

Лемма 2. *Если матрица $A \in \mathbb{R}^{l \times n}$ такова, что $\text{rank } A \geq r$, то для матрицы $U(A)$, получаемой в результате работы алгоритма 1, выполнено $\text{rank } U(A) = l - r$. Более того, если $\text{rank } A = r$, то $A^T U(A) = 0$.*

Доказательство. Равенство $\text{rank } U(A) = l - r$ немедленно следует из определения $U(A)$ на шаге 4 алгоритма, поскольку $\det U^{(r)} = 1$. Равенство $A^T U(A) = 0$ следует из определения $U(A)$, соотношения (33) и леммы 1. ■

Еще раз подчеркнем, что полученные свойства не зависят от выбора i_s и j_s на шаге 1 алгоритма 1, хотя сама матрица $U(A)$ может зависеть от их выбора.

Предложение 4. Пусть ранг матрицы $\bar{A} \in \mathbb{R}^{l \times n}$ равен r , и пусть последовательность матриц $\{A_k\} \subset \mathbb{R}^{l \times n}$ сходится к \bar{A} . Пусть для всех достаточно больших k матрица $U(A_k)$ получена в результате работы алгоритма 1.

Тогда для любой предельной точки \bar{U} ограниченной последовательности $\{U(A_k)\}$ справедливы соотношения $\text{rank } \bar{U} = l - r$ и $\bar{A}^T \bar{U} = 0$.

Доказательство. Заметим, что $\text{rank } A_k \geq r$ для всех достаточно больших k , и поэтому запуск алгоритма 1 для матрицы $A = A_k$ или $A = \bar{A}$ будет успешным.

Для доказательства соотношений $\text{rank } \bar{U} = l - r$ и $\bar{A}^T \bar{U} = 0$ будем рассуждать от противного: предположим, что существует предельная точка \bar{U} последовательности $\{U(A_k)\}$, для которой выполнено $\text{rank } \bar{U} < l - r$ или $\bar{A}^T \bar{U} \neq 0$. Без ограничения общности будем считать, что $\{U(A_k)\}$ сходится к \bar{U} , и что для каждого номера $s = 0, \dots, r - 1$ индексы i_s и j_s , выбираемые на шаге 1 алгоритма 1, применяемого к матрице $A = A_k$, совпадают для всех k .

Легко убедиться, что для каждого $s = 0, \dots, r - 1$ алгоритм 1, применяемый к матрице $A = \bar{A}$, также может использовать указанные выше индексы i_s и j_s . Действительно, соотношение (30) при $A = A_k$ выполнено для всех k , и поскольку $\{A_k\}$ сходится к \bar{A} , то, переходя в (30) к пределу при $k \rightarrow \infty$, получаем равенство

$$|\bar{A}_{i_s j_s}^{(s)}| = \max_{\substack{i \in \{1, \dots, l\} \setminus \mathcal{I}_s, \\ j \in \{1, \dots, n\}}} |\bar{A}_{ij}^{(s)}|.$$

Следовательно, (30) выполнено и при $A = \bar{A}$.

Для всех k и всех $s = 0, \dots, r - 1$ обозначим через $T_k^{(s)}$, $U_k^{(r)}$ и $U(A_k)$ соответствующие матрицы, полученные в ходе работы алгоритма 1, примененного к $A = A_k$ (при указанных выше i_s и j_s). Обозначим через $T^{(s)}$, $U^{(r)}$ и $U(\bar{A})$ матрицы, полученные в ходе работы алгоритма 1, примененного к $A = \bar{A}$ (с теми же i_s и j_s).

Из соотношения (31) и сходимости $\{A_k\}$ к \bar{A} следует сходимость $\{T_k^{(s)}\}$ к $T^{(s)}$ для всех $s = 0, \dots, r - 1$. Отсюда, с учетом (32), получаем сходимость $\{U_k^{(r)}\}$ к $U^{(r)}$, а значит и сходимость $\{U(A_k)\}$ к $U(\bar{A})$. Таким образом, $U(\bar{A}) = \bar{U}$, и поэтому любое из свойств $\text{rank } \bar{U} < l - r$ и $\bar{A}^T \bar{U} \neq 0$ противоречит лемме 2. \blacksquare

Полученные конструкции и результаты можно распространить на случай, когда ранг r матрицы \bar{A} неизвестен. Пусть для заданной матрицы A и номера $s \in \{0, \dots, r\}$ объекты \mathcal{I}_s и $A^{(s)}$ получены в ходе работы алгоритма 1. Обозначим через $R^{(s)}(A) \in \mathbb{R}^{(l-s) \times n}$ матрицу, состоящую из строк $A^{(s)}$ с номерами $i \in \{1, \dots, l\} \setminus \mathcal{I}_s$. Из леммы 1 вытекает, что если $\text{rank } A = r$, то $R^{(r)}(A) = 0$. Докажем теперь более сильный результат.

Лемма 3. Пусть ранг матрицы $\bar{A} \in \mathbb{R}^{l \times n}$ равен r .

Тогда найдутся такие константы $\varepsilon > 0$ и $M > 0$, что для любой матрицы $A \in \mathbb{R}^{l \times n}$, достаточно близкой к \bar{A} , и для соответствующих матриц $R^{(s)}(A)$, сгенерированных алгоритмом 1 при $s \in \{0, \dots, r\}$, выполняются следующие соотношения:

$$\|R^{(s)}(A)\| \geq \varepsilon \quad \forall s = 0, \dots, r-1, \quad (34)$$

$$\|R^{(r)}(A)\| \leq M\|A - \bar{A}\|. \quad (35)$$

Доказательство. Докажем сначала, что существует $\varepsilon > 0$, для которого выполнено неравенство (34) для всех матриц A , достаточно близких к \bar{A} . Предположим, что это не так: существуют $q \in \{0, \dots, r-1\}$ и последовательность $\{A_k\} \subset \mathbb{R}^{l \times n}$, сходящаяся к \bar{A} , такие, что для матриц $R^{(q)}(A_k)$, сгенерированных алгоритмом 1, справедливо

$$\lim_{k \rightarrow \infty} R^{(q)}(A_k) = 0. \quad (36)$$

Без ограничения общности будем считать, что для каждого $s = 0, \dots, q-1$ индексы i_s и j_s , выбираемые на шаге 1 алгоритма 1, применяемого к матрице $A = A_k$, совпадают для всех k . Повторяя рассуждение из доказательства предложения 4, легко показать, что для каждого $s = 0, \dots, q-1$ алгоритм 1, применяемый к матрице $A = \bar{A}$, также может использовать указанные индексы i_s и j_s , и что полученные в ходе его работы матрицы $R^{(q)}(\bar{A})$ удовлетворяют соотношению

$$\lim_{k \rightarrow \infty} R^{(q)}(A_k) = R^{(q)}(\bar{A}).$$

С учетом (36), из полученного соотношения получаем равенство $R^{(q)}(\bar{A}) = 0$. Таким образом, для матрицы $\bar{A}^{(q)}$, полученной в ходе работы алгоритма 1, примененного к $A = \bar{A}$, справедлива цепочка соотношений $\text{rank } \bar{A} = \text{rank } \bar{A}^{(q)} \leq q < r$, которая противоречит предположениям.

Докажем теперь существование константы $M > 0$, удовлетворяющей (35) для всех матриц A , достаточно близких к \bar{A} . Заметим, что неравенство (35) выполнено для $A = \bar{A}$ при любом $M > 0$, поскольку обе части неравенства равны нулю. Предположим, что для некоторой последовательности $\{A_k\} \subset \mathbb{R}^{l \times n} \setminus \{\bar{A}\}$, сходящейся к \bar{A} , матрицы $R^{(r)}(A_k)$, сгенерированные алгоритмом 1, удовлетворяют соотношению

$$\lim_{k \rightarrow \infty} \frac{\|R^{(r)}(A_k)\|}{\|A_k - \bar{A}\|} = +\infty. \quad (37)$$

Вновь предположим без ограничения общности, что для каждого $s = 0, \dots, q-1$ индексы i_s и j_s , выбираемые на шаге 1 алгоритма 1, применяемого к матрице $A = A_k$, совпадают для всех k , откуда следует, что для каждого $s = 0, \dots, q-1$ алгоритм 1, применяемый к матрице $A = \bar{A}$, также может использовать указанные индексы i_s и j_s .

Заметим, что для каждого $s = 0, \dots, r-1$ матрица $T^{(s)}$, определяемая согласно (31) при любых фиксированных i_s и j_s , рассматриваемая как функция от A (отображающая $\mathbb{R}^{l \times n}$ в

$\mathbb{R}^{l \times l}$), является аналитической в окрестности \bar{A} . Поэтому матрица $A^{(r)} = A^{(r)}(A)$, определяемая согласно (32), (33), также является аналитической в окрестности \bar{A} (как функция от A), а значит, липшицева в окрестности \bar{A} . Поэтому из определения $R^{(r)}$ и равенства $R^{(r)}(\bar{A}) = 0$ следует, что для некоторой константы $M > 0$ выполнено

$$\|R^{(r)}(A_k)\| = \|R^{(r)}(A_k) - R^{(r)}(\bar{A})\| \leq \|A^{(r)}(A_k) - A^{(r)}(\bar{A})\| \leq M\|A_k - \bar{A}\|$$

для всех k , что, очевидно, противоречит (37). \blacksquare

Доказанная лемма позволяет модифицировать алгоритм 1 таким образом, чтобы он мог автоматически оценивать r . Предположим, что помимо матрицы $A \in \mathbb{R}^{l \times n}$ известна величина $t \geq 0$, некоторым образом оценивающая расстояние от A до \bar{A} .

Алгоритм 2. Полагаем $\mathcal{I}_0 = \emptyset$, $U^{(0)} = I \in \mathbb{R}^{l \times l}$, $A^{(0)} = A$, $R^{(0)}(A, t) = A$, $s = 0$.

1. Если $s = l$ или

$$\|R^{(s)}(A, t)\| \leq t, \quad (38)$$

полагаем $r = s$ и переходим к шагу 4. В противном случае находим индексы $i_s \in \{1, \dots, l\} \setminus \mathcal{I}_s$ и $j_s \in \{1, \dots, n\}$, для которых справедливо (30). Полагаем $\mathcal{I}_{s+1} = \mathcal{I}_s \cup \{i_s\}$.

2. Определяем матрицу преобразования $T^{(s)} \in \mathbb{R}^{l \times l}$ согласно (31). Полагаем $U^{(s+1)} = U^{(s)}T^{(s)}$, $A^{(s+1)} = (T^{(s)})^T A^{(s)}$ и определяем $R^{(s+1)}(A, t)$ как матрицу, составленную из столбцов $A^{(s+1)}$ с номерами из множества $\{1, \dots, l\} \setminus \mathcal{I}_{s+1}$.

3. Увеличиваем s на 1 и переходим к шагу 1.

4. Определяем $U(A, t)$ как матрицу, составленную из столбцов $U^{(r)}$ с номерами из множества $\{1, \dots, l\} \setminus \mathcal{I}_r$.

Заметим, что значение $A_{i_s j_s}^{(s)}$ на шаге 1 алгоритма 2 всегда отлично от нуля, поскольку в противном случае тест (38) был бы выполнен. Таким образом, запуск алгоритма 2 всегда является успешным.

Предложение 5. Пусть $\tau > 0$ и $\theta \in (0, 1)$ фиксированы. Пусть последовательность $\{A_k\} \subset \mathbb{R}^{l \times n}$ сходится к некоторой матрице $\bar{A} \in \mathbb{R}^{l \times n}$. Пусть числовая последовательность $\{\sigma_k\}$ сходится к нулю при $k \rightarrow \infty$, и существует такое $M > 0$, что для всех k справедлива оценка

$$\|A_k - \bar{A}\| \leq M|\sigma_k|.$$

Тогда для всех достаточно больших k алгоритм 2, применяемый к $A = A_k$ и $t = \tau|\sigma_k|^\theta$, завершается со значением $r = \text{rank } \bar{A}$, и для любой предельной точки \bar{U} ограниченной последовательности $\{U(A_k, \tau|\sigma_k|^\theta)\}$ справедливы соотношения $\text{rank } \bar{U} = l - \text{rank } \bar{A}$ и $\bar{A}^T \bar{U} = 0$.

Доказательство. С учетом предложения 4, достаточно показать, что для всех достаточно больших k алгоритм 2 завершается при $r = \text{rank } \bar{A}$, поскольку в этом случае матрица $U(A_k, \tau|\sigma_k|^\theta)$ совпадает с $U(A_k)$, полученной в результате работы алгоритма 1 (предполагая, что в обоих алгоритмах выбираются одинаковые индексы i_s и j_s для всех $s = 0, \dots, r-1$).

Предположим сначала, что существует бесконечное число номеров k , для которых алгоритм 2, применяемый к $A = A_k$ и $t = \tau|\sigma_k|^\theta$, завершается при $r < \text{rank } \bar{A}$. Поскольку $\text{rank } \bar{A} \leq l$, то $r < l$, а значит тест (38) выполняется для таких k при $s = r$. Таким образом, соответствующая подпоследовательность последовательности $\{R^{(r)}(A_k, \tau|\sigma_k|^\theta)\}$ сходится к нулю, что противоречит лемме 3, согласно которой найдется $\varepsilon > 0$, для которого $\|R^{(r)}(A_k)\| \geq \varepsilon$ для всех достаточно больших k , а значит и $\|R^{(r)}(A_k, \tau|\sigma_k|^\theta)\| \geq \varepsilon$ для всех достаточно больших k (матрицы $R^{(s)}(A_k)$, генерируемые алгоритмом 1, и $R^{(s)}(A_k, \tau|\sigma_k|^\theta)$, генерируемые алгоритмом 2, совпадают при $s \leq \text{rank } A_k$).

Предположим теперь, что существует бесконечное число номеров k , для которых алгоритм 2 завершается при $r > \text{rank } \bar{A}$. Очевидно, для всех таких k справедливо

$$\|R^{(\text{rank } \bar{A})}(A_k, \tau|\sigma_k|^\theta)\| > \tau|\sigma_k|^\theta \geq \frac{\tau\|A_k - \bar{A}\|^\theta}{M^\theta}.$$

С другой стороны, из леммы 3 получаем оценку

$$\|R^{(\text{rank } \bar{A})}(A_k)\| = O(\|A_k - \bar{A}\|)$$

при $k \rightarrow \infty$, что невозможно, поскольку $\theta \in (0, 1)$. ■

Возвращаясь к s-sSQP, определим отображение $P: \mathbb{R}^n \times \mathbb{R}^l \rightarrow \mathbb{R}^{l \times l}$ следующим образом:

$$P(x, \lambda) = U(U^T U)^{-1} U^T, \quad (39)$$

где матрица $U = U(h'(x), \tau|\sigma(x, \lambda)|^\theta)$ определяется с помощью алгоритма 2, применяемого к $A = h'(x)$ и $t = \tau|\sigma(x, \lambda)|^\theta$ при некоторых фиксированных значениях $\tau > 0$, $\theta \in (0, 1)$.

Нетрудно убедиться, что если матрица $\bar{U} \in \mathbb{R}^{l \times (l-r)}$ удовлетворяет соотношениям $\text{rank } \bar{U} = l-r$ и $(h'(\bar{x}))^T \bar{U} = 0$, то ортогональный проектор \bar{P} на подпространство вырожденности $(\text{im } h'(\bar{x}))^\perp$ имеет вид

$$\bar{P} = \bar{U}(\bar{U}^T \bar{U})^{-1} \bar{U}^T.$$

Поэтому из предложения 5 следует, что если функция σ удовлетворяет предположению (S1), то для любого множителя $\lambda^* \in \mathcal{M}(\bar{x})$, достаточно близкого к $\bar{\lambda}$, отображение P , определяемое согласно (39), удовлетворяет соотношению

$$\lim_{k \rightarrow \infty} P(x^k, \lambda^k) = \bar{P}$$

для любой последовательности $\{(x^k, \lambda^k)\}$, сходящейся к (\bar{x}, λ^*) . Последнее означает, что P удовлетворяет обоим предположениям (P1) и (P2).

4.3. Глобализованный алгоритм

В данном разделе будет показано, что стратегия глобализации сходимости, предложенная в разделе 3.3 для sSQP, может быть использована и для s-sSQP. При этом сохраняются все полученные результаты о глобальной сходимости и о сверхлинейной скорости сходимости.

Легко убедиться, что утверждения а) и б) леммы 3.2 остаются справедливыми, если в предположениях леммы заменить направление (ξ, η) на направление поиска метода s-sSQP, т. е. на произвольное решение системы (3). Доказательство этих утверждений леммы при этом практически не меняется, поскольку в нем используется только первое уравнение системы (3.53), которое совпадает с первым уравнением системы s-sSQP (3). Из леммы 3.2 получаем правила выбора c_1 и c_2 , при которых направление s-sSQP является направлением убывания для штрафной функции

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

Эти правила приводят к следующему алгоритму.

Алгоритм 3. Выбираем параметры $c_1 > 0$, $c_2 > 0$, $\delta > 0$, $\rho > 0$, $q > 1$, $\varepsilon, \nu \in (0, 1)$. Фиксируем отображение $P: \mathbb{R}^n \times \mathbb{R}^l \rightarrow \mathbb{R}^{l \times l}$ и функцию $\sigma: \mathbb{R}^n \times \mathbb{R}^l \rightarrow \mathbb{R}$. Фиксируем непрерывные функции $\psi_1, \psi_2, \psi_3: \mathbb{R}_+ \rightarrow \mathbb{R}_+$, положительные везде кроме нуля. Выбираем $(x^0, \lambda^0) \in \mathbb{R}^n \times \mathbb{R}^l$ и полагаем $k = 0$.

1. Если $\Phi(x^k, \lambda^k) = 0$, завершаем работу алгоритма.
2. Вычисляем $d^k = (\xi^k, \eta^k)$ как решение системы (3) при $(x, \lambda) = (x^k, \lambda^k)$. Если система (3) не имеет решений, переходим к шагу 4.
3. Если

$$\langle \varphi'_{c_1, c_2}(x^k, \lambda^k), d^k \rangle \leq -\rho \|d^k\|^q, \quad (40)$$

переходим к шагу 5.

Если

$$\|h(x^k)\| \geq \psi_1(\|\Phi(x^k, \lambda^k)\|) \quad (41)$$

и

$$\langle h(x^k), h'(x^k)\xi^k \rangle \leq -\psi_2(\|h(x^k)\|), \quad (42)$$

полагаем $c_1 = \bar{c}_{1, k} + \delta$, где $\bar{c}_{1, k} = \bar{c}_1(\rho \|d^k\|^q; x^k, \lambda^k; \xi^k, \eta^k)$ и \bar{c}_1 определяется согласно (3.56), и переходим к шагу 5.

Если

$$\left\| \frac{\partial L}{\partial x}(x^k, \lambda^k) \right\| \geq \psi_3(\|\Phi(x^k, \lambda^k)\|), \quad (43)$$

полагаем $c_2 = \bar{c}_{2,k} + \delta$, где $\bar{c}_{2,k} = \bar{c}_2(\rho \|d^k\|^q; x^k, \lambda^k; \xi^k, \eta^k; c_1)$ и \bar{c}_2 определяется согласно (3.57), и переходим к шагу 5.

4. Полагаем $d^k = -\varphi'_{c_1, c_2}(x^k, \lambda^k)$.

5. Вычисляем $\alpha_k = \nu^j$, где j — наименьшее неотрицательное целое число, удовлетворяющее неравенству Армихо

$$\varphi_{c_1, c_2}((x^k, \lambda^k) + \nu^j d^k) \leq \varphi_{c_1, c_2}(x^k, \lambda^k) + \varepsilon \nu^j \langle \varphi'_{c_1, c_2}(x^k, \lambda^k), d^k \rangle. \quad (44)$$

6. Полагаем $(x^{k+1}, \lambda^{k+1}) = (x^k, \lambda^k) + \alpha_k d^k$, увеличиваем k на 1 и переходим к шагу 1.

Легко убедиться, что замечание 3.1 остается справедливым для алгоритма 3. Отсюда следует, что алгоритм корректно определен: либо он завершается на шаге 1, либо направление d^k принимается на шаге 5 для некоторого j .

Глобальная сходимость алгоритма 3 устанавливается следующей теоремой, доказательство которой полностью аналогично доказательству теоремы 3.7.

Теорема 3. Пусть $f: \mathbb{R}^n \rightarrow \mathbb{R}$ и $h: \mathbb{R}^n \rightarrow \mathbb{R}^l$ дважды непрерывно дифференцируемы на \mathbb{R}^n . Пусть последовательность $\{(x^k, \lambda^k)\}$ сгенерирована алгоритмом 3.

Тогда справедливы следующие утверждения:

а) если значения c_1 и c_2 не меняются для всех достаточно больших номеров k , то любая предельная точка $(\bar{x}, \bar{\lambda})$ последовательности $\{(x^k, \lambda^k)\}$ удовлетворяет равенству

$$\varphi'_{c_1, c_2}(\bar{x}, \bar{\lambda}) = 0; \quad (45)$$

б) если существует такое бесконечное множество K , что c_1 или c_2 увеличивается на всех итерациях с номерами $k \in K$, то любая предельная точка $(\bar{x}, \bar{\lambda})$ последовательности $\{(x^k, \lambda^k) \mid k \in K\}$, такая что отображение P и функция σ непрерывны в этой точке, либо удовлетворяет системе Лагранжа (2), либо такова, что последовательность $\{d^k \mid k \in K\}$ неограничена, а матрица

$$\begin{pmatrix} \frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda}) & (h'(\bar{x}))^T \\ h'(\bar{x}) & -\sigma(\bar{x}, \bar{\lambda})P(\bar{x}, \bar{\lambda}) \end{pmatrix} \quad (46)$$

вырождена.

Напомним, что если выполнено (45) и при этом матрица в правой части (3.60) невырождена при $(x, \lambda) = (\bar{x}, \bar{\lambda})$ (что справедливо «почти для всех» значений c_1 и c_2), то $(\bar{x}, \bar{\lambda})$

удовлетворяет системе Лагранжа (2). Заметим также, что для алгоритма 3 остается справедливым замечание 3.2, согласно которому можно улучшить свойства глобальной сходимости алгоритма, если ввести верхние границы $C_1 > 0$ и $C_2 > 0$ на параметры c_1 и c_2 .

Перейдем к анализу локальной сходимости. Покажем сначала, что леммы 3.3 и 3.4 остаются справедливыми, если направление поиска метода sSQP в них заменить на направление s-sSQP.

Лемма 4. *Пусть предположения леммы 3.3 выполнены для алгоритма 3. Пусть отображение P и функция σ удовлетворяют предположениям (P1) и (S1) или (P2) и (S2) соответственно.*

Тогда утверждение леммы 3.3 также справедливо для алгоритма 3 и решения $d^k = (\xi^k, \eta^k)$ системы (3).

Доказательство. Существование и единственность решения d^k системы (3) следует из теоремы 1 а) и соответствующего утверждения теоремы 2.

Рассуждая так же, как при доказательстве леммы 3.3, легко убедиться, что один из тестов (41) и (43) обязательно должен быть выполнен, причем если тест (43) выполнен, а (41) нарушается, то алгоритм модифицирует c_2 и принимает направление s-sSQP.

Предположим, что тест (41) выполняется. Легко проверить, что второе и третье равенства в (29) справедливы при $(x^k, \lambda^k) \rightarrow (\bar{x}, \bar{\lambda})$. Поэтому, применяя (4) и оценки (25), (26) из теоремы 1 б) или соответствующие оценки из теоремы 2, получаем

$$h(x^k) + h'(x^k)\xi^k = o(\|x^k - \bar{x}\|) + O(\|x^k - \bar{x}\|\|\xi^k\|) = o(\|\Phi(x^k, \lambda^k)\|). \quad (47)$$

Таким образом, из (41) следует

$$\begin{aligned} \langle h(x^k), h'(x^k)\xi^k \rangle &= -\|h(x^k)\|^2 + \langle h(x^k), h(x^k) + h'(x^k)\xi^k \rangle \leq \\ &\leq -\|h(x^k)\|^2 + \|h(x^k)\|o(\|\Phi(x^k, \lambda^k)\|) = -\|h(x^k)\|^2 + o(\|h(x^k)\|^2) \end{aligned}$$

при $(x^k, \lambda^k) \rightarrow (\bar{x}, \bar{\lambda})$, откуда вытекает, что тест (42) выполнен для заданной функции ψ_2 и для всех (x^k, λ^k) , достаточно близких к $(\bar{x}, \bar{\lambda})$. Таким образом, алгоритм модифицирует c_1 и принимает направление s-sSQP.

Наконец, из (4) и (25) вытекают оценки

$$\begin{aligned} \left\langle \frac{\partial L}{\partial x}(x^k, \lambda^k), \xi^k \right\rangle + \langle h(x^k), \eta^k \rangle &= O(\|\Phi(x^k, \lambda^k)\|^2), \\ \langle h(x^k), h'(x^k)\xi^k \rangle &= O(\|\Phi(x^k, \lambda^k)\|^2) \end{aligned}$$

при $(x^k, \lambda^k) \rightarrow (\bar{x}, \bar{\lambda})$, и поэтому последнее утверждение леммы следует из (25), (3.56) и (3.57). ■

Лемма 5. Пусть $f: \mathbb{R}^n \rightarrow \mathbb{R}$ и $h: \mathbb{R}^n \rightarrow \mathbb{R}^l$ дважды дифференцируемы в окрестности точки $\bar{x} \in \mathbb{R}^n$, а их вторые производные непрерывны в этой точке. Пусть \bar{x} — стационарная точка задачи (1), а $\bar{\lambda} \in \mathbb{R}^l$ — отвечающий ей множитель Лагранжа. Пусть отображение $P: \mathbb{R}^n \times \mathbb{R}^l \rightarrow \mathbb{R}^{l \times l}$ и функция $\sigma: \mathbb{R}^n \times \mathbb{R}^l \rightarrow \mathbb{R}_+$ удовлетворяют предположениям (P1) и (S1) или (P2) и (S2) соответственно, причем в последнем случае множитель $\bar{\lambda}$ является некритическим. Предположим, что для любого $u = (x, \lambda)$, достаточно близкого к \bar{u} , существует вектор $d(u) = (\xi, \eta)$, удовлетворяющий системе (3) при $P = P(u)$ и $\sigma = \sigma(u)$, и для которого выполнено $d(u) \rightarrow 0$ при $u \rightarrow \bar{u}$ и $d(\bar{u}) = 0$.

Тогда для любых $c_1 > 0$ и $c_2 > 0$ существует окрестность \mathcal{V} точки \bar{u} и непрерывная в нуле функция $\chi: \mathbb{R}^n \times \mathbb{R}^l \rightarrow \mathbb{R}_+$, для которой справедлива оценка $\chi(u) = o(\|d(u)\|^2)$ при $u \rightarrow \bar{u}$ и неравенство

$$\varphi_{c_1, c_2}(u + d(u)) - \varphi_{c_1, c_2}(u) \leq \frac{1}{2} \langle \varphi'_{c_1, c_2}(u), d(u) \rangle + \chi(u) \quad (48)$$

для всех $u \in \mathcal{V}$.

Доказательство. Рассуждая точно так же, как при доказательстве леммы 3.4, но используя направление $d(u)$ метода s-sSQP, несложно показать, что неравенство (48) выполнено для всех $u \in \mathcal{V}$, если определить функцию $\chi: \mathbb{R}^n \times \mathbb{R}^l \rightarrow \mathbb{R}_+$,

$$\begin{aligned} \chi(u) = & \frac{c_1}{2} (\|h'(x)\xi\| + \|\sigma(u)P(u)\eta\|) \|\sigma(u)P(u)\eta\| + \frac{3}{2} \|\sigma(u)P(u)\eta\| \|\eta\| + \\ & + \frac{c_1}{2} (\|\sigma(u)P(u)\eta\| + \chi_1(u))^2 + \frac{c_2}{2} (\chi_1(u))^2 + \chi_2(u), \end{aligned}$$

где χ_1, χ_2 определяются аналогично тому, как это делалось в доказательстве леммы 3.4.

Из предположений (P1) и (S1), очевидно, следует оценка

$$\sigma(u)P(u)\eta = o(\|d(u)\|) \quad (49)$$

при $u \rightarrow \bar{u}$.

В противном случае, если P и σ ограничены в окрестности $(\bar{x}, \bar{\lambda})$, то любое решение (ξ, η) системы (3), очевидно, удовлетворяет оценке

$$\Phi(x, \lambda) = O(\|(\xi, \eta)\|) \quad (50)$$

при $(x, \lambda) \rightarrow (\bar{x}, \bar{\lambda})$. Более того, предполагая (P2), (S2) и некритичность $\bar{\lambda}$, легко проверить, что при $u \rightarrow \bar{u}$ остаются справедливыми равенства из (47) с заменой (x^k, λ^k) и ξ^k на (x, λ) и ξ соответственно. Поэтому из равенства $\sigma(u)P(u)\eta = h(x) + h'(x)\xi$ и оценки (50) вновь получаем (49).

В обоих случаях из (49) следует оценка $\chi(u) = o(\|d(u)\|^2)$ при $u \rightarrow \bar{u}$. ■

Предполагая теперь выполнение $q = 2$ вместо $q \geq 2$ и применяя леммы 4 и 5, а также предложение 3.4, приходим к следующей теореме. Ее доказательство полностью аналогично доказательству теоремы 3.8, но использует теоремы 1 и 2 вместо теоремы 1.3.

Теорема 4. Пусть $f: \mathbb{R}^n \rightarrow \mathbb{R}$ и $h: \mathbb{R}^n \rightarrow \mathbb{R}^l$ дважды дифференцируемы в окрестности точки $\bar{x} \in \mathbb{R}^n$, а их вторые производные непрерывны в этой точке. Пусть \bar{x} — стационарная точка задачи (1), а $\bar{\lambda} \in \mathbb{R}^l$ — отвечающий ей некритический множитель Лагранжа. Пусть отображение P и функция σ удовлетворяют предположениям (P1) и (S1) или (P2) и (S2) соответственно. Предположим, что в алгоритме 3 используются значения $q = 2$ и $\varepsilon \in (0, 1/2)$, а функции ψ_1, ψ_2, ψ_3 выбраны следующим образом: $\psi_1(t) = \rho_1 t$, $\psi_2(t) = \rho_2 t^2$, $\psi_3(t) = \rho_3 t$, где $\rho_1, \rho_3 > 0$, $\rho_1^2 + \rho_3^2 < 1$, $\rho_2 \in (0, 1)$. Предположим, что траектория $\{(x^k, \lambda^k)\}$, сгенерированная алгоритмом 3, сходится к точке $(\bar{x}, \bar{\lambda})$.

Тогда $\alpha_k = 1$ для всех достаточно больших k , и скорость сходимости последовательности $\{(x^k, \lambda^k)\}$ является сверхлинейной.

В завершение данного раздела покажем, что теорема 3.9 также остается справедливой, если направление d^k в ней заменить на направление поиска метода s-sQP.

Теорема 5. Пусть $f: \mathbb{R}^n \rightarrow \mathbb{R}$ и $h: \mathbb{R}^n \rightarrow \mathbb{R}^l$ трижды дифференцируемы в окрестности точки $\bar{x} \in \mathbb{R}^n$, причем их третьи производные непрерывны в этой точке. Пусть \bar{x} — стационарная точка задачи (1), а отвечающий ей множитель Лагранжа $\bar{\lambda} \in \mathbb{R}^l$ удовлетворяет SOSC

$$\left\langle \frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda})\xi, \xi \right\rangle > 0 \quad \forall \xi \in \ker h'(\bar{x}) \setminus \{0\}.$$

Пусть $P: \mathbb{R}^n \times \mathbb{R}^l \rightarrow \mathbb{R}^{l \times l}$ и $\sigma: \mathbb{R}^n \times \mathbb{R}^l \rightarrow \mathbb{R}$ удовлетворяют предположениям (P1) и (S1) или (P2) и (S2) соответственно. Положим $\varepsilon \in (0, 1/2)$, $\theta \in (0, 1)$ и рассмотрим итерационную схему $(x^{k+1}, \lambda^{k+1}) = (x^k, \lambda^k) + \alpha_k d^k$, где для всех достаточно больших номеров k направление поиска $d^k = (\xi^k, \eta^k)$ является решением системы (3) при $(x, \lambda) = (x^k, \lambda^k)$, а $\alpha_k = \nu^j$, где j — наименьшее неотрицательное целое число, для которого выполнено неравенство Армихо (44) для некоторых фиксированных $c_1 \geq 0$ и $c_2 \geq 0$.

Тогда для любого $\bar{c}_2 > 0$ найдется константа $\bar{c}_1 \geq 0$, обладающая следующим свойством: для любых $c_1 \geq \bar{c}_1$ и $c_2 \geq \bar{c}_2$ и для любой последовательности $\{(x^k, \lambda^k)\} \subset (\mathbb{R}^n \times \mathbb{R}^l) \setminus (\{\bar{x}\} \times \mathcal{M}(\bar{x}))$, генерируемой описанной выше итерационной схемой и сходящейся к $(\bar{x}, \bar{\lambda})$, выполнено $\alpha_k = 1$ для всех достаточно больших k , и скорость сходимости является сверхлинейной.

Доказательство. Для произвольного $\bar{c}_2 > 0$ выберем $\bar{c}_1 \geq 0$ таким образом, чтобы выполня-

лось утверждение предложения 3.5, и для любых $c_1 \geq \bar{c}_1$ и $c_2 \geq \bar{c}_2$ рассмотрим произвольную траекторию $\{(x^k, \lambda^k)\}$ с указанными выше свойствами.

Согласно теореме 1 б) и соответствующему утверждению теоремы 2, вблизи прямо-двойственного решения, в котором выполнено SOSC, полный шаг метода s-sSQP удовлетворяет оценке (26), а значит,

$$\text{dist}((x^k, \lambda^k) + d^k, \{\bar{x}\} \times \mathcal{M}(\bar{x})) = o(\text{dist}((x^k, \lambda^k), \{\bar{x}\} \times \mathcal{M}(\bar{x})))$$

при $k \rightarrow \infty$.

Покажем, что для траектории $\{(x^k, \lambda^k)\}$ будет выполнено неравенство

$$\langle \varphi''_{c_1, c_2}(\bar{x}, \bar{\lambda})d^k, d^k \rangle \geq \gamma \|d^k\|^2 \quad (51)$$

для некоторого $\gamma > 0$ и для всех достаточно больших k . Согласно предложению 3.5, достаточно показать, что последовательность $\{(\xi^k, \eta^k)/\|(\xi^k, \eta^k)\|\}$ не имеет предельных точек вида $(0, \eta)$, где $\eta \in \ker(h'(\bar{x}))^T$. Предположим, напротив, что такие предельные точки существуют, откуда автоматически следует $\|\eta\| = 1$.

Предполагая (P1) и (S1), противоречие несложно получить, практически дословно повторяя доказательство теоремы 3.9. Предполагая (P2) и (S2), из (49) выводим равенство $P(\bar{x}, \bar{\lambda})\eta = 0$, откуда получаем, что $\eta \in \ker P(\bar{x}, \bar{\lambda})$. С другой стороны, $\eta \in \ker(h'(\bar{x}))^T = (\text{im } h'(\bar{x}))^\perp$, что возможно только если $\eta = 0$, поскольку $\ker P(\bar{x}, \bar{\lambda}) \cap (\text{im } h'(\bar{x}))^\perp = \{0\}$. Отсюда вновь получаем противоречие.

Таким образом, в обоих случаях выполнено неравенство (51), и для завершения доказательства остается применить следствие 3.2. ■

Приложение А

Численные результаты для метода модифицированных функций Лагранжа

В данном приложении приводятся результаты вычислительного эксперимента, целью которого являлось исследование эффективности и робастности метода модифицированных функций Лагранжа на задачах с нерегулярными ограничениями. Раздел А.1 посвящен сравнению солвера ALGENCAN, реализующего МФЛ, с другими известными алгоритмами. В разделе А.2 исследуется эффект притяжения двойственных траекторий МФЛ к критическим множителям и его влияние на скорость сходимости. Раздел А.3 посвящен анализу общей эффективности метода в зависимости от способа управления параметром штрафа. Наконец, в разделе А.4 исследуется влияние критических множителей на поведение описанных в разделе 2.2 ускорителей финальной фазы этого метода.

В эксперименте использовалась версия 2.3.7 пакета ALGENCAN с AMPL-интерфейсом, скомпилированная с использованием библиотеки MA57 [62]. Данная библиотека предназначена для эффективного решения разреженных систем линейных уравнений, что позволяет существенно повысить общую эффективность солвера на задачах большой размерности.

Эксперимент проводился на двух тестовых коллекциях: DEGEN [42] и MacMPEC [83] (обе существуют в виде AMPL моделей).

MacMPEC представляет собой коллекцию из 180 задач оптимизации с комплементарными ограничениями (на момент проведения эксперимента), т. е. задач вида

$$\begin{aligned} f(x) \rightarrow \min, \quad h(x) = 0, \quad g(x) \leq 0, \\ G(x) \geq 0, \quad H(x) \geq 0, \quad \langle G(x), H(x) \rangle \leq 0. \end{aligned} \tag{1}$$

Данная коллекция содержит как небольшие академические примеры, так и прикладные задачи, в том числе большой размерности. В эксперименте использовалась 161 задача из MacMPEC: 13 задач были исключены, поскольку в них присутствуют так называемые смешанные комплементарные ограничения; одна задача содержит двоичную переменную;

остальные 5 исключенных задач не имеют допустимых точек, согласно [83].

Большинство результатов сравнения алгоритмов на коллекции MacMPEC в этом и последующих приложениях представлено в форме так называемых «performance profiles» (см. [45]). Для каждого из алгоритмов значение функции, график которой изображен на рисунке, в точке $\tau \in [1, \infty)$ есть доля задач в наборе, на которых результат данного алгоритма был не более чем в τ раз хуже наилучшего результата для данной задачи среди всех сравниваемых алгоритмов. При этом считается, что результат неудачного запуска в бесконечное число раз хуже любого результата. В частности, значение этой функции при $\tau = 1$ соответствует доле запусков, на которых результат данного алгоритма был наилучшим. Значение функции при больших τ характеризует робастность алгоритма, т. е. долю его успешных запусков.

Коллекция DEGEN содержит 111 небольших задач математического программирования с нерегулярными ограничениями, из которых в эксперименте использовались 98. Остальные 13 задач были исключены по следующим причинам: задачи 20205, 20223, 30204 и 40207 имеют неограниченные на допустимом множестве целевые функции (хотя в них существуют стационарные точки, в которых ограничения нерегулярны); задачи 2DD01-50h, 2DD01-50v, 2DD01-500h, 2DD01-500v, 2DD01-5000h и 2DD01-5000v слишком большие по сравнению с остальными; задачи 20201, 40210 и 40211 содержат только простые ограничения, вследствие чего ALGENCAN не делает для этих задач ни одной внешней итерации. Отметим, что в 52 из оставшихся 98 задач существуют критические относительно некоторого множества индексов $A \subset A(\bar{x})$ множители Лагранжа; в 33 задачах критических множителей нет; в 13 задачах решению вообще не отвечает ни один множитель Лагранжа, что вполне возможно в случае нерегулярных ограничений.

В задачах из коллекции DEGEN прямо-двойственная начальная точка не задана по умолчанию, а выбирается случайным образом из заданной окрестности. Для каждой задачи делалось некоторое количество запусков из случайных начальных точек, при этом использовался размер окрестности 100, заданный по умолчанию.

Для представления большинства результатов на коллекции DEGEN используется модификация подхода из [45], учитывающая многократность запусков алгоритмов для задач из этой коллекции. А именно, для каждого из алгоритмов на рисунках представлен график функции $\pi: [1, +\infty) \rightarrow [0, 1]$, определяемой следующим образом. Обозначим через k_p средний результат на один успешный запуск, а через s_p — долю успешных запусков данного алгоритма для p -й задачи в коллекции. Пусть r_p — наименьшее значение k_p по всем алгоритмам. Тогда для каждого $\tau \in [1, +\infty)$ полагаем

$$\pi(\tau) = \frac{1}{P} \sum_{p \in R(\tau)} s_p,$$

где P — общее количество используемых задач в коллекции, а $R(\tau)$ — множество задач, для которых результат данного алгоритма не более чем в τ раз хуже наилучшего:

$$R(\tau) = \{p = 1, \dots, P \mid k_p \leq \tau r_p\}.$$

В частности, значение $\pi(1)$ примерно соответствует доле запусков, на которых солвер показал наилучший результат, а значения $\pi(\tau)$ для больших τ характеризуют робастность солвера.

А.1. Сравнение с другими алгоритмами

В данном разделе приводятся результаты вычислительного эксперимента, в котором солвер ALGENCAN [30], реализующий МФЛ, сравнивался со следующими солверами:

- SNOPT [54] и filterSQP [51], две хорошо известные реализации метода SQP;
- MINOS [86], реализующий метод модифицированных функций Лагранжа с линейризованными ограничениями;
- IPOPT [63, 92] и KNITRO [40, 78], реализующие метод внутренней точки.

Кроме того, для МРСС использовался солвер IPOPT-C [91], который представляет собой модификацию IPOPT, использующую специальную структуру комплементарных ограничений. KNITRO также имеет некоторые особенности, использующие структуру МРСС [80]. Методы SQP участвуют в сравнении, поскольку известно [52], что они показывают сравнительно высокую эффективность и робастность для МРСС. Сравнение с методом модифицированных функций Лагранжа с линейризованными ограничениями, реализованном в MINOS, связано с тем, что данный метод в некотором смысле связан как с МФЛ, так и с SQP; см. [69, 70].

Что касается остальных солверов, использовались пакеты SNOPT 7.2-8 и MINOS 5.51, входящие в базовый комплект AMPL, и IPOPT 3.8.0. Солверы KNITRO 8.0.0 и filterSQP 20020316 запускались на сервере NEOS [87]. Для МРСС использовалась последняя доступная версия IPOPT-C 2.2.1e. Для всех солверов, включая ALGENCAN, использовались значения по умолчанию для всех параметров с единственным исключением: максимальное число внешних итераций для MINOS было увеличено с 50 до 1000, поскольку в этом случае существенно улучшается робастность солвера на вырожденных задачах.

В данном эксперименте основной интерес представляли робастность алгоритмов и качество результатов (т. е. значений целевой функции, если метод останавливается в допустимой точке). В этом состоит одна из причин того, что не уделялось существенного внимания сравнению эффективности и, в частности, процессорного времени. Другая очевидная причина состоит в том, что некоторые солверы запускались на сервере NEOS, и поэтому сравнение

процессорного времени просто не имеет смысла. Тем не менее, чтобы предоставить дополнительную информацию, было также проведено сравнение солверов по количеству вычислений целевой функции и ограничений, поскольку эти данные частично отражают эффективность и в то же время доступны для всех солверов (за исключением числа вычислений ограничений для KNITRO, которое отсутствует в выдаче солвера).

При сравнении числа вычислений есть следующая техническая сложность: солверы MINOS и SNOPT не учитывают линейные функции при подсчете числа вычислений, в то время как остальные солверы выдают число вычислений как линейных, так и нелинейных функций. Чтобы сделать сравнение относительно честным, в тех случаях, когда целевая функция или все ограничения являются линейными, соответствующее число вычислений полагалось равным 1 для всех солверов. Отметим, что МРСС всегда имеет по крайней мере одно нелинейное ограничение (последняя компонента в (1)), и поэтому такие действия не требуются для числа вычислений ограничений на коллекции MacMPEC.

Наконец, было проведено сравнение солверов с точки зрения качества результатов. Для каждого солвера приводится число неудачных запусков (т. е. случаев, когда солвер заканчивает работу с флагом, отличным от успешного), а также число случаев сходимости к неоптимальным значениям. Кроме того, был проведен анализ ограниченности двойственных траекторий, генерируемых солвером ALGENCAN. Напомним, что ограниченность двойственной траектории МФЛ означает стационарность допустимых предельных точек прямой траектории.

Сначала приведем результаты на коллекции MacMPEC. Отметим, что все задачи из этой коллекции написаны с использованием оператора `complements` языка AMPL, что создает некоторые технические сложности. Во-первых, среди рассматриваемых солверов только KNITRO и IPOPT-C распознают данный оператор. Остальные солверы просто игнорируют ограничения, записанные с помощью данного оператора. Для таких солверов необходимо было переформулировать комплементарные ограничения в виде обычных ограничений-равенств или ограничений-неравенств. Во-вторых, даже KNITRO может напрямую работать только с теми тестами из MacMPEC, в которых каждая компонента отображений G и H имеет вид $ax_j + b$ для некоторых $a, b \in \mathbb{R}$ и $j \in \{1, \dots, n\}$. В связи с этим, солверы KNITRO и IPOPT-C запускались на моделях, в которых комплементарные ограничения были переформулированы с использованием дополнительных переменных. А именно, если компонента отображения G с номером $i \in \{1, \dots, s\}$ имеет вид, отличный от $ax_j + b$, то вводилась дополнительная переменная s_i , и ограничение

$$G_i(x) \geq 0 \quad \text{complements} \quad H_i(x) \geq 0$$

заменялось на

$$G_i(x) - s_i = 0,$$

$$s_i \geq 0 \quad \text{complements} \quad H_i(x) \geq 0.$$

Затем аналогичным образом модифицировалось отображение H .

Для всех солверов, кроме KNITRO и IPOPT-C (которые корректно распознают оператор `complements`), для представления задачи (1) использовались 4 эквивалентные переформулировки, в которых последняя компонента в (1) задается одним из условий

$$\langle G(x), H(x) \rangle \leq 0; \tag{2}$$

$$\langle G(x), H(x) \rangle = 0; \tag{3}$$

$$G_i(x)H_i(x) \leq 0, \quad i = 1, \dots, s; \tag{4}$$

$$G_i(x)H_i(x) = 0, \quad i = 1, \dots, s. \tag{5}$$

Более того, каждая из форм (2)–(5) может быть далее переформулирована с использованием дополнительных переменных. Таким образом, получаем 8 различных переформулировок для комплементарных ограничений. Полученный вычислительный опыт свидетельствует о том, что для ALGENCAN чуть более предпочтительной является форма (2), хотя различия не слишком существенные. В то же время, поведение солверов MINOS и SNOPT значительно лучше на формах (2) и (3). Кроме того, введение дополнительных переменных несколько улучшает поведение всех солверов, что согласуется с существующими результатами [52, 53]. В связи с этим, для солверов ALGENCAN, SNOPT, filterSQP и MINOS далее приводятся результаты для переформулировки MPCC с использованием (2) и с введенными дополнительными переменными (в случае необходимости). Отметим, что солвер filterSQP, применяемый к такой переформулировке, по сути эквивалентен солверу filtermpcc [52], применяемому к исходной задаче.

Для каждой задачи из MacMPCC было проведено по одному запуску из начальной точки по умолчанию, заданной в модели. Диаграмма на рис. 1 показывает число неудачных запусков (столбцы черного цвета) и число случаев сходимости к неоптимальным значениям (столбцы серого цвета) для различных солверов. Значение целевой функции в момент завершения алгоритма считалось неоптимальным, если оно отличается от наилучшего известного значения более чем на 0.1. Оптимальные значения были взяты из [83], за исключением случаев, когда в ходе экспериментов были найдены допустимые точки с лучшим значением целевой функции. Диаграмма показывает, что ALGENCAN имеет наименьшее число случаев сходимости к неоптимальным значениям, а также что только IPOPT-C слегка опережает ALGENCAN по числу неудачных запусков. Отметим, впрочем, что для IPOPT-C отсутствуют какие-либо теоретические гарантии глобальной сходимости [91].

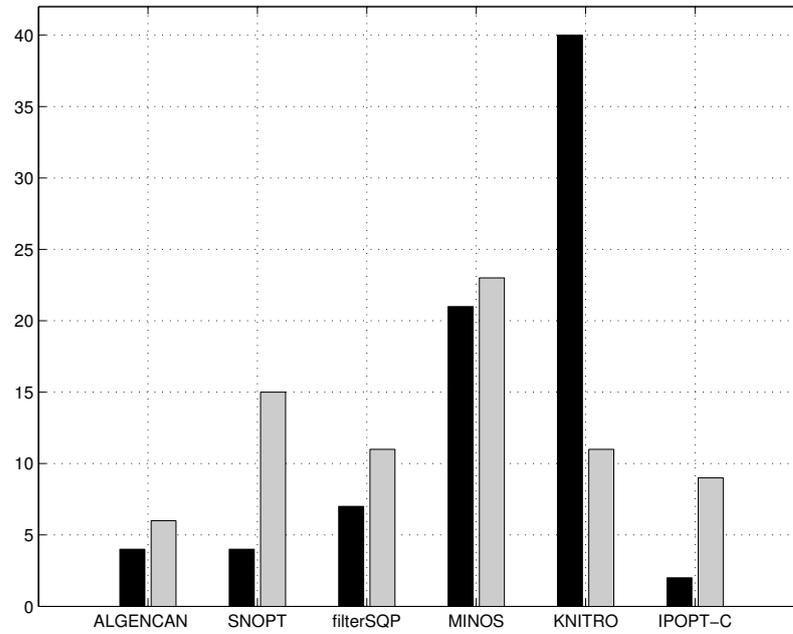


Рис. 1. Неудачные запуски и случаи сходимости к неоптимальным значениям на MacMPEC.

Таблица 1. Задачи из MacMPEC с «аномальным» поведением ALGENCAN.

Задача	НЗ	НО	НДТ	Задача	НЗ	НО	НДТ
bilevel1		•		qpсec-200-4		•	
bilevel3		•		qpсec2			•
bilin		•		ralphmod	•		•
design-cent-2	•		•	scale4			•
design-cent-4		•		scholtes4			•
ex9.2.2			•	siouxfls1	•		•
hakonsen	•		•	water-FL		•	•
pack-comp2p-8			•	water-net			•

На рисунках 2(а) и 2(б) показаны результаты сравнения алгоритмов по числу вычислений целевой функции и ограничений соответственно. Как видно из рисунков, ALGENCAN, SNOPT, filterSQP и IPOPT-C продемонстрировали практически одинаково высокую робастность (решено около 97% задач), в то время как MINOS и KNITRO оказались менее робастны на MacMPEC (решено около 85% и 75% задач соответственно). Рис. 2 также показывает, что результат ALGENCAN по количеству вычислений целевой функции и ограничений сравним с результатом MINOS, однако по этим показателям оба они значительно уступают остальным солверам.

Следует отметить, что ALGENCAN опережает остальные солверы по числу внешних итераций, показывая лучший результат примерно на 60% задач; filterSQP лидирует на менее

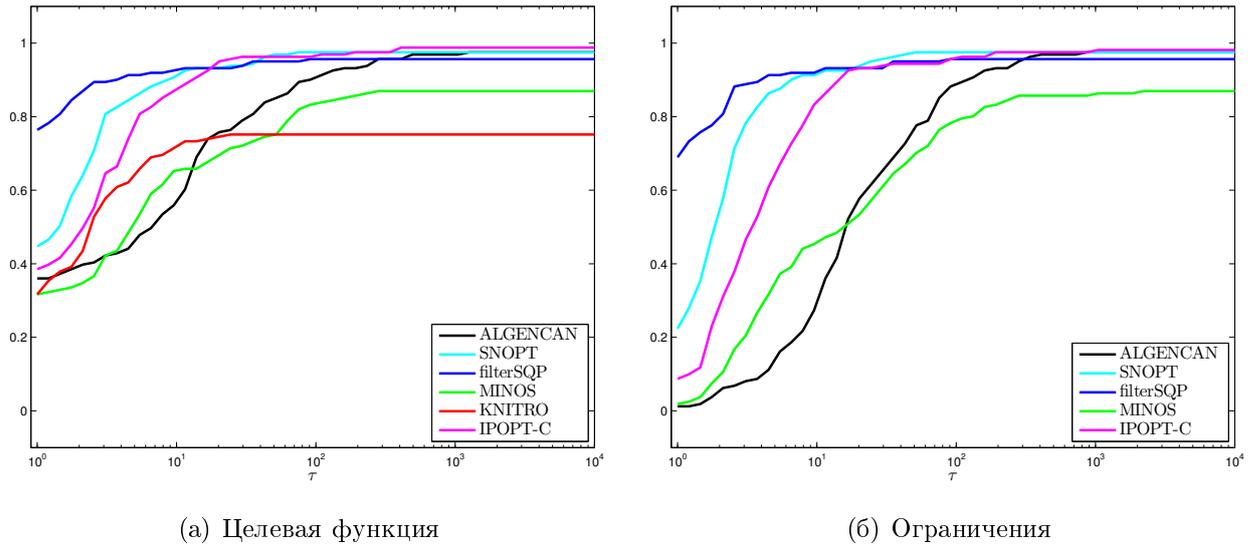


Рис. 2. Число вычислений нелинейных функций на МасМРЕС.

чем 40% задач, а для остальных солверов доля соответствующих задач не превосходит 15%.

Наконец, в таблице 1 собрана информация о задачах из МасМРЕС, на которых солвер ALGENCAN демонстрировал «плохое» поведение: неудачные запуски (столбцы «НЗ»), сходимость к неоптимальным значениям (столбцы «НО») или неограниченность двойственной траектории (столбцы «НДТ»). Двойственная траектория считалась неограниченной, если на момент завершения имело место неравенство $\|(\lambda^k, \mu^k)\|_\infty \geq 10^4$. Как видно из таблицы 1, сама по себе неограниченность двойственной траектории часто не является проблемой: в большинстве таких случаев алгоритм все равно сходился к оптимальному значению. Более того, можно заключить, что ограниченность двойственных траекторий является типичным сценарием для МФЛ даже в тех случаях, когда множество множителей Лагранжа является неограниченным, и в таких случаях для МРСС гарантирована сильная стационарность предельных точек при условии, что выполнено МРСС-LICQ. И даже если двойственная траектория является неограниченной, алгоритм вполне может прийти к оптимальному решению.

Перейдем к результатам на коллекции DEGEN. Для каждой задачи делалось по 100 запусков из случайных начальных точек. Для того, чтобы выделить случаи сходимости к неоптимальным значениям, использовалось пороговое значение 10^{-2} для расстояния до оптимального значения, указанного в модели. Число неудачных запусков, а также случаев сходимости к неоптимальным значениям для различных солверов показано на рис. 3 (цвета те же, что и на рис. 1). Эксперименты показали, что доля случаев сходимости к оптимальным значениям среди успешных запусков для ALGENCAN приблизительно равна 95%, в то время как для всех остальных солверов она не превышает 90%. Более того, доля случаев сходимости к оптимальным значениям среди всех запусков, включая неудачные, для ALGENCAN также выше, чем для всех остальных солверов.

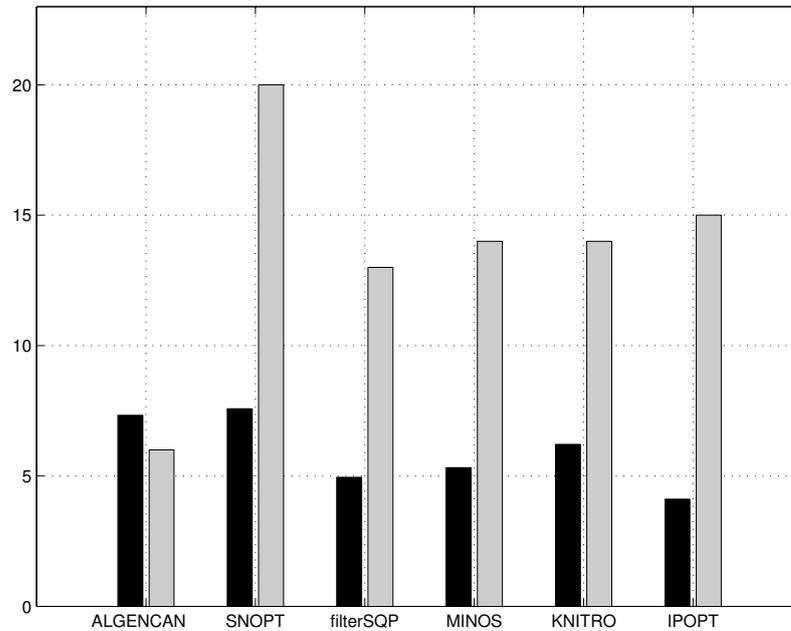


Рис. 3. Неудачные запуски и случаи сходимости к неоптимальным значениям на DEGEN.

Практически во всех случаях неудачных запусков солвера ALGENCAN причиной неудачи являлось слишком большое значение параметра штрафа. Последнее обычно связано со сходимостью к недопустимым точкам, а также в редких случаях с медленной скоростью сходимости.

Сравнение солверов по числу вычислений целевой функции и ограничений представлено на рисунках 4(а) и 4(б) соответственно. Согласно рисункам, все солверы имеют примерно одинаковую робастность на коллекции DEGEN: около 94% запусков являются успешными. С точки зрения эффективности картина во многом аналогична результатам для МасМРЕС: ALGENCAN несколько более эффективен, чем MINOS, однако менее эффективен, чем остальные солверы. Однако, различия не столь значительны, как на коллекции МасМРЕС.

По внешним итерациям ALGENCAN вновь существенно опережает остальные солверы, показывая лучший результат практически на всех задачах. Более того, примерно на половине задач результат ALGENCAN по крайней мере в 4 раза лучше, чем результаты остальных солверов.

Наконец, доля запусков, в которых двойственная траектория оказалась неограниченной, составила лишь 9.1%. Более того, лишь для 12 задач доля таких запусков оказалась выше 20%. Таким образом, несмотря на неограниченность множества множителей для большинства задач из DEGEN, двойственные траектории ALGENCAN обычно остаются ограниченными.

Таким образом, если наибольшее значение имеет робастность и качество получаемых результатов, то ALGENCAN является хорошим выбором. Если же более важна скорость, то

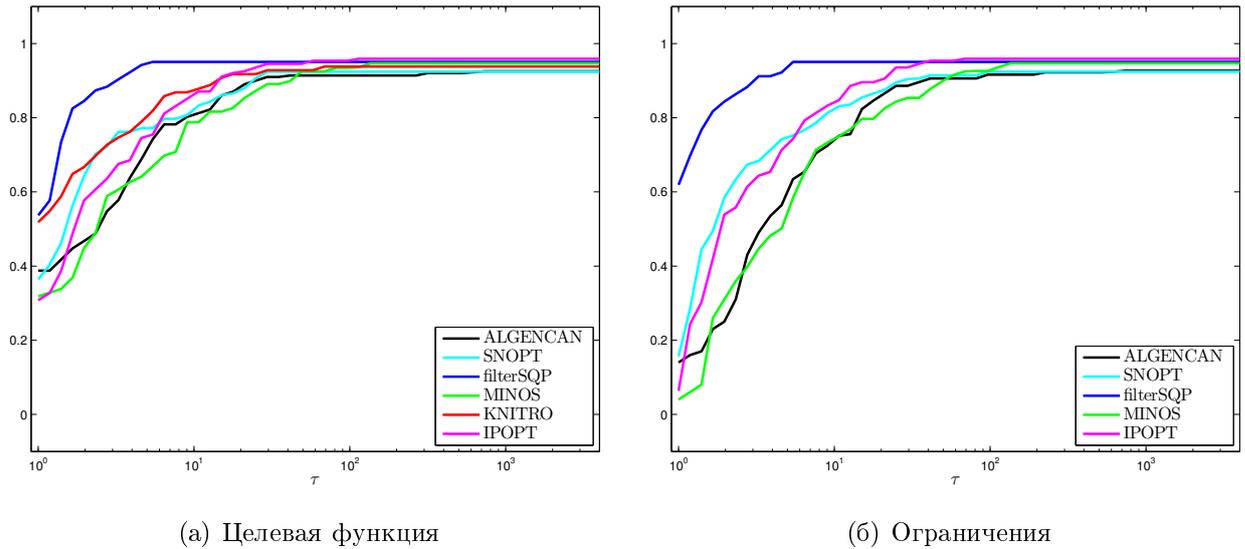


Рис. 4. Число вычислений нелинейных функций на DEGEN.

требуются дополнительные модификации для ускорения ALGENCAN.

А.2. Эксперимент с правилами для параметра штрафа: критические множители и скорость сходимости

В этом разделе приводятся результаты вычислительного эксперимента, в котором исследовалось влияние критических множителей на скорость сходимости солвера ALGENCAN с различными способами управления параметром штрафа. В данном эксперименте сравнивались солвер ALGENCAN (см. алгоритм 2.1 и сопровождающие его комментарии; в дальнейшем этот базовый вариант будем обозначать M1), а также его модификация (в дальнейшем обозначаемая M2), в которой параметр штрафа увеличивался в два раза на каждой итерации, причем на первой итерации он брался равным 1. Таким образом, согласно теореме 1.4, для M2 можно ожидать сверхлинейную скорость сходимости.

В обоих вариантах M1 и M2 алгоритма использовались следующие настройки. Во-первых, был отключен ньютоновский ускоритель, описанный в разделе 2.2, поскольку в данном эксперименте изучается поведение «чистого» метода модифицированных функций. Во-вторых, ограничение на число внутренних итераций (т. е. итераций метода, применяемого для решения подзадач (2.5)) на первой внешней итерации было увеличено до 100 (значение по умолчанию 10). Кроме того, в экспериментах на коллекции DEGEN было отключено масштабирование, поскольку оно приводит к изменению множества множителей Лагранжа, отвечающих решению.

Для каждой задачи из коллекции DEGEN осуществлялось по 10 запусков каждого из алгоритмов M1 и M2 из случайных начальных точек. Для каждого из запусков анализирова-

лась скорость убывания расстояния до прямого решения, а также притяжение двойственной траектории к критическим множителям (решения и критические множители указаны в моделях коллекции DEGEN).

На коллекции DEGEN результаты для M1 и M2 оказались практически одинаковыми. Это связано с тем, что для многих задач из DEGEN алгоритм 2.1, в котором подзадачи решаются точно (т. е. $\varepsilon_k = 0$ для всех k), находит точное решение за конечное число шагов. Для таких задач поведение алгоритма в значительной степени зависит от точности решения подзадач и мало зависит от правила изменения параметра штрафа. Поэтому ниже приводятся результаты только для M2.

Из 98 задач, использовавшихся в эксперименте, только для 80 наблюдалась сходимость к решениям, в которых нарушаются условия регулярности ограничений, и ниже обсуждаются результаты именно для этих задач. Для остальных задач либо запуски завершались неудачей, либо имела место сходимость к невырожденному решению.

Для всех задач, в которых множество множителей Лагранжа, отвечающих решению, пусто (таких задач 12 из 80), наблюдалась линейная сходимость прямой траектории, а норма двойственных приближений стремилась к бесконечности. Для тех задач, в которых решению отвечают только некритические множители (27 задач из 80), практически во всех случаях наблюдалась сверхлинейная сходимость прямой траектории, причем двойственная траектория сходилась к некоторому множителю Лагранжа.

Результаты для задач, в которых решению отвечает хотя бы один критический множитель (41 задача из 80), приведены в таблице 2. В столбце «СЛ» отмечены случаи сверхлинейной прямой сходимости, а в столбце «КМ» — случаи сходимости к критическому множителю. При этом черный кружок означает, что соответствующее поведение наблюдалось во всех запусках, а белый — что только для части запусков. Как видно из таблицы, сходимость к критическим множителям является типичной, но при этом сверхлинейная скорость прямой сходимости обычно сохраняется. Вместе с тем, отсутствие сверхлинейной сходимости всегда сопряжено именно со сходимостью к критическим множителям.

Для задач из МасМРЕС осуществлялось по одному запуску каждого варианта алгоритма из указанных в моделях коллекции начальных точек. Для большинства этих задач ни решение, ни тем более отвечающие ему множители Лагранжа не известны. В связи с этим, под скоростью сходимости здесь понималась скорость убывания выдаваемой ALGENCAN на каждой итерации невязки системы ККТ.

Для варианта M2 были получены следующие результаты. Для 28 задач из 161 запуск завершился неудачей. Из оставшихся 133 задач лишь для 16 наблюдалась линейная скорость сходимости, а для остальных — сверхлинейная.

Базовый вариант M1 оказался значительно более робастным: лишь для 10 задач его за-

Таблица 2. Результаты для задач из DEGEN с критическими множителями.

Задача	СЛ	КМ	Задача	СЛ	КМ	Задача	СЛ	КМ
10201	•	•	20209	○	○	20309		•
10203	•	•	20210	•	•	2DD01-5v	•	
10206	•	•	20211	•		30210	•	
10207	•	•	20213	○	•	30211	•	
10301		•	20214		•	30301	•	
20101	•		20216	•	•	30401	•	•
20103	•	•	20217	○	•	40201	•	
20105	○	○	20219	•	•	40202	•	
20107	•		20226	•	•	40205	•	•
20109	•	•	20227	•	•	40206	•	•
20202	•	•	20302	•	•	40208	•	•
20203		•	20303	•		40401	•	
20204	•		20304	•		40402	•	•
20207	•	•	20307		•			

пуск был неудачным. Однако, как и следовало ожидать, отсутствие сверхлинейной скорости сходимости наблюдалось значительно чаще, чем для M2, а именно, для 98 запусков из 151 успешных. Правда, при этом для 30 из этих задач линейная скорость была весьма быстрой: невязка системы ККТ убывала на каждой итерации более чем в 10 раз. Сверхлинейная скорость сходимости наблюдалась для 53 задач.

Из полученных результатов можно сделать следующие выводы. Если параметр штрафа стремится к бесконечности, то МФЛ во многих случаях сохраняет сверхлинейную скорость сходимости несмотря на нерегулярность ограничений решаемой задачи, и несмотря на притяжение двойственных траекторий к критическим множителям Лагранжа (при их наличии). Однако, как и следовало ожидать, бесконечное увеличение параметра штрафа негативно влияет на робастность метода.

А.3. Эксперимент с правилами для параметра штрафа: общая эффективность

Как отмечено в предыдущем разделе, высокая скорость сходимости внешних итераций МФЛ может не компенсировать расходов на ее обеспечение: быстрый рост стоимости решения подзадач может оказывать решающее негативное влияние на общую эффективность (и даже

на робастность) метода.

Для сравнения эффективности МФЛ с разными способами управления параметром штрафа варианты М1 и М2 алгоритма 2.1, описанные в предыдущем разделе, были дополнены следующими двумя. Вариант М3 аналогичен М2, но параметр штрафа увеличивается на каждой итерации в $\rho = 10$ раз. В варианте М4 параметр штрафа также увеличивается в 10 раз на каждой итерации, но, как и в базовом ALGENCAN, использовались «рестарты» для этого параметра, а значение c_0 определялось с помощью процедуры инициализации (см. раздел 2.1.1).

Похожие эксперименты проводились в работах [36, 38], где авторы исследовали влияние различных способов управления параметром штрафа на робастность и эффективность ALGENCAN. В [36] сравнивались базовый ALGENCAN и его модификации, в которых параметр штрафа растет на каждой итерации. В [38] исследовалось влияние на базовый ALGENCAN различных значений параметра ρ и «рестартов». Однако, в отличие от экспериментов, представленных в данном разделе, тестирование в [36, 38] проводилось на задачах с регулярными ограничениями.

В данном эксперименте по-прежнему был отключен ускоритель для всех вариантов алгоритма. Остальные настройки ALGENCAN были установлены по умолчанию. В частности, использовалось масштабирование, а ограничение на число внутренних итераций для первой внешней итерации было оставлено равным 10.

Сравнение вариантов ALGENCAN проводилось по следующим четырем показателям: количество внешних и внутренних итераций, а также количество вычислений целевой функции и ограничений. Количество внешних итераций отражает скорость сходимости алгоритма, а остальные характеристики могут рассматриваться как различные показатели его эффективности.

Результаты в данном разделе представлены в форме «performance profiles», описанной выше. На рис. 5 приведены результаты на коллекции MacMPEC. Как видно из рис. 5, для всех вариантов алгоритма, в которых параметр штрафа увеличивается на каждой итерации, робастность ниже, чем для базового ALGENCAN: в вариантах М2–М4 параметр штрафа чаще достигает величины 10^{20} , и в этих случаях запуск объявляется неудачным. При этом робастность варианта М4 выше, чем М2 и (особенно) М3, т. е. адаптивная инициализация параметра штрафа и его «рестарты» повышают вероятность успешности запуска.

Из рис. 5(а) также следует, что варианты М3 и (особенно) М4 имеют лучший результат по количеству внешних итераций по сравнению с базовым вариантом М1, что свидетельствует о более высокой скорости сходимости М3 и (особенно) М4. Результат варианта М2, напротив, хуже, чем результат М1: хотя параметр штрафа в М2 растет на каждой итерации, это происходит слишком медленно.

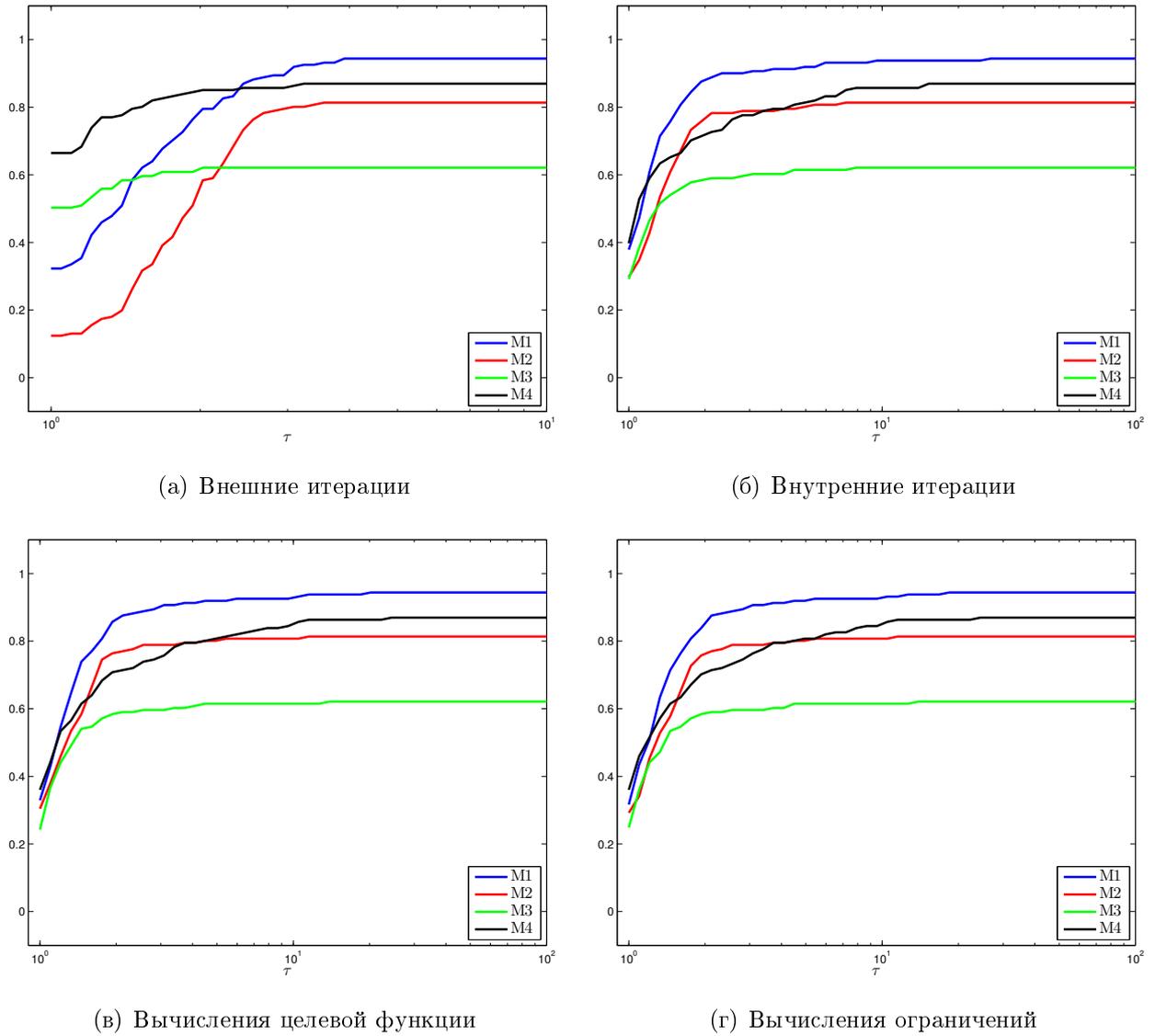


Рис. 5. Результаты на МасМРЕС.

С другой стороны, из рис. 5(б)–5(г) следует, что все варианты алгоритма, в которых параметр штрафа увеличивается на каждой итерации (и даже те, которые демонстрируют более высокую скорость сходимости), не обыгрывают базовый ALGENCAN по эффективности. Отметим также, что вариант М4 несколько более эффективен, чем М2 и М3, хотя и незначительно.

На рис. 6 представлены результаты на коллекции DEGEN. Как видно из рисунка, на DEGEN все варианты алгоритма имеют примерно одинаковую робастность. Результаты по числу внешних итераций, показанные на рис. 6(а), повторяют картину для МасМРЕС, хотя различие между вариантами М1 и М2, и между М3 и М4 гораздо менее значительно. Однако, как следует из рис. 6(б)–6(г), в отличие от результатов для МасМРЕС, варианты М3 и М4 обгоняют М1 и М2 также и по остальным характеристикам, хотя и незначительно.

Приведенные результаты подтверждают высказанные выше опасения: высокая скорость

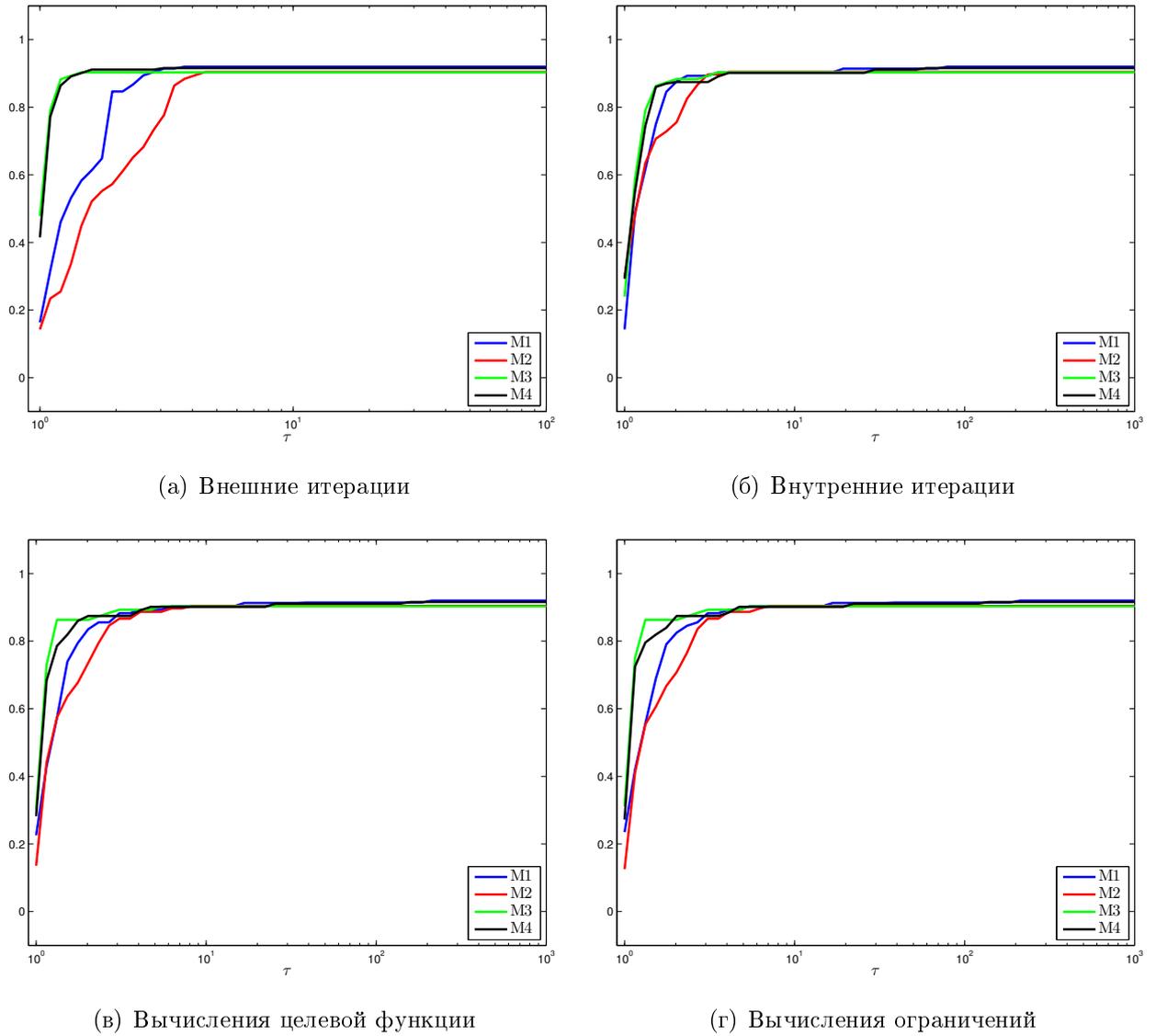


Рис. 6. Результаты на DEGEN.

сходимости МФЛ, достигаемая даже на задачах с нерегулярными ограничениями за счет бесконечного увеличения параметра штрафа, во многих случаях не приводит к высокой итоговой эффективности метода.

А.4. Эксперимент с ускорителями

В заключение приведем результаты сравнения двух вариантов ALGENCAN с включенным ускорителем: базового варианта, в котором ускорителем является МНЛ, и варианта, использующего в качестве ускорителя СМНЛ (см. описание в разделе 2.2). В данном эксперименте использовались настройки ALGENCAN по умолчанию.

Тестирование на коллекции DEGEN проводилось так же, как и в разделе А.2. Для 10 задач из 98 все запуски обоих тестируемых вариантов алгоритма заканчивались либо неуда-

чей, либо сходимостью к решению, в котором ограничения регулярны. Еще для 14 задач ни один из ускорителей вообще не запускался ни разу: ALGENCAN успевал набрать требуемую точность (т.е. найти приближение, удовлетворяющее (2.60)) шагами МФЛ до того, как оказывался выполнен тест (2.59) на включение ускорителя. Поэтому результаты приводятся для оставшихся 74 задач. Для некоторых из этих задач ускоритель запускался на нескольких внешних итерациях (т.е. на всех таких итерациях, кроме, возможно, последней, его запуски завершались неудачей). В таких случаях анализировалась только траектория, сгенерированная ускорителем на последней внешней итерации.

Для 11 из 12 задач, в которых решению не отвечает ни один множитель Лагранжа, запуски ускорителей всегда завершались неудачей. В оставшемся примере такого рода 2DD01-5h все запуски обоих ускорителей завершились успехом, и при этом во всех случаях скорость прямой сходимости была сверхлинейной.

Таблица 3. Результаты для задач из DEGEN с критическими множителями.

Задача	УУ	СЛ	КМ	Задача	УУ	СЛ	КМ
10201	• / •	/ •	• /	20214	• / •	/	• / •
10203	• / •	/	• / •	20216	• / •	/ ◦	• / ◦
10206	• / •	/ •	• /	20217	• / •	/	• / •
10207	• / •	/ •	• /	20219	• / •	/	• / •
10301	/	/	/	20226	• / •	/ ◦	• / •
10403	◦ / ◦	• / •	/	20227	• / •	/	• / •
20101	• / •	• / •	/	20302	• / •	/ ◦	• / ◦
20103	• / •	/ ◦	• / ◦	20303	• / •	◦ / •	◦ /
20105	• / •	• / •	/	20307	• /	/	• /
20107	• / •	◦ / •	◦ /	20309	• / •	/	• / •
20109	• / •	/	• / •	20312	◦ / ◦	/	• / •
20202	• / •	/	• / •	2DD01-5v	• / •	◦ / ◦	◦ / ◦
20203	◦ / ◦	/	• / •	30211	◦ / ◦	• / •	/
20204	• / •	◦ / •	◦ /	30301	• / •	• / •	◦ /
20207	• / •	/	• / •	30401	• / •	◦ / •	◦ /
20209	◦ / ◦	◦ / ◦	◦ / ◦	40201	◦ / ◦	• / •	/
20210	• / •	/	• / •	40206	/	/	/
20211	• / •	• / •	/	40208	/	/	/
20213	/ •	/	/ •	40402	• / •	/ •	• /

Результаты для задач, в которых есть критические множители (38 задач из 74), пред-

Таблица 4. Результаты для задач из DEGEN без критических множителей.

Задача	УУ	СЛ	Задача	УУ	СЛ	Задача	УУ	СЛ
00301	● / ●	● / ●	20106	/	/	30206	○ / ○	● / ●
00302	○ / ○	/	20110	● / ●	● / ●	30303	● / ●	● / ●
00501	/	/	20111	/	/	40203	/	/
10202	● / ●	● / ●	20221	● / ●	● / ●	40204	● / ●	● / ●
10204	/ ●	/ ●	20222	○ / ●	● / ●	01001	● / ●	● / ●
10205	○ / ○	/	20225	○ / ○	○ / ○	10401	○ / ○	○ / ○
20102	● / ●	● / ●	30202	● / ○	● / ●	10404	○ / ○	○ / ○
20104	○ / ●	/ ●	30205	● / ●	● / ●	10701	○ / ○	○ / ○

ставлены в таблице 3 (для МНЛ, и через косую черту для СМНЛ). В столбце «УУ» отмечены случаи успешных запусков ускорителей, т. е. те случаи, когда решение находилось именно шагами ускорителя. В столбце «СЛ» отмечены случаи сверхлинейной прямой сходимости, а в столбце «КМ» — случаи притяжения двойственной траектории к критическим множителям. Черные и белые кружки имеют тот же смысл, что и в таблице 2.

Как видно из таблицы, для обоих ускорителей практически все запуски завершились успехом. Однако скорость сходимости и поведение двойственной траектории ускорителей несколько различаются.

Для МНЛ практически во всех случаях имела место сходимость к критическим множителям; при этом было лишь несколько случаев сверхлинейной прямой сходимости, причем практически во всех из них двойственная траектория сходилась к некритическому множителю. Для СМНЛ сверхлинейная сходимость прямой траектории наблюдалась значительно чаще, а двойственная траектория реже сходилась к критическому множителю, хотя таких случаев, как и случаев линейной прямой сходимости, по-прежнему большинство. Как и для МНЛ, практически во всех случаях сверхлинейной сходимости двойственная траектория сходилась к некритическому множителю. Таким образом, сходимость к критическим множителям в большинстве случаев разрушает сверхлинейную скорость сходимости обоих тестируемых ускорителей.

Следует отметить, что при наличии ограничений-неравенств аналогичный эффект может наблюдаться даже в том случае, когда решению отвечают только некритические множители, поскольку критические множители могут присутствовать в соответствующей задаче (2.63) со вспомогательными переменными, к которой применяется МНЛ или СМНЛ. Более того, несложно убедиться, что наличие критических множителей в задаче (2.63) является в определенном смысле типичным. Действительно, как следует из [71, предложение 5], если

$A(\bar{x}) \subset \hat{A}$, то любому множителю Лагранжа $(\lambda, \mu) \in \mathcal{M}(\bar{x})$, для которого нарушается условие строгой дополнителности (т. е. $\mu_i = 0$ при некотором $i \in A(\bar{x})$), соответствует критический множитель Лагранжа $(\lambda, \mu_{\hat{A}})$ в задаче (2.63). Далее, из (2.2) очевидно, что множество

$$\mathcal{M}_A(\bar{x}) = \{\mu_{A(\bar{x})} \mid (\lambda, \mu) \in \mathcal{M}(\bar{x})\}$$

представляет собой пересечение аффинного множества в $\mathbb{R}^{|A(\bar{x})|}$ с неотрицательным ортантом, и поэтому выполнение $\mu_{A(\bar{x})} > 0$ для всех $\mu_{A(\bar{x})} \in \mathcal{M}_A(\bar{x})$ возможно только в том случае, когда $\mathcal{M}_A(\bar{x})$ состоит из единственного элемента. Последнее означает, что вектор μ одинаковый для всех множителей $(\lambda, \mu) \in \mathcal{M}(\bar{x})$, что нетипично для задач с нерегулярными ограничениями.

Результаты для задач, в которых решению отвечают только некритические множители (24 задачи из 74), приведены в таблице 4. Отметим, что практически для всех из них в соответствующих задачах со вспомогательными переменными присутствуют критические множители. Вместе с тем, картина несколько отличается от представленной в таблице 3. Результаты для двух используемых ускорителей почти не различаются. При этом доля неудачных запусков ускорителей весьма существенна, однако успешных запусков все же значительно больше. Кроме того, для успешных запусков скорость прямой сходимости практически всегда была сверхлинейной.

Поведение ускорителей, демонстрируемое таблицами 3 и 4, связано еще с одной эвристикой, реализованной в ALGENCAN, которая состоит в адаптивном возмущении диагональных элементов матриц, стоящих в левой части систем (2.66) и (2.69), с целью преодоления возможной вырожденности этих матриц (см. [37]). Возмущение диагональных элементов матрицы в (2.66) приближает поведение МНЛ к СМНЛ, и для некоторых задач приводит к сверхлинейной прямой сходимости МНЛ. Отключение данной эвристики приближает поведение ускорителей к ожидаемому: в частности, прямая сходимость МНЛ во многих случаях становится линейной, при этом двойственная траектория сходится к одному из критических множителей Лагранжа задачи (2.63). В некоторых случаях прямая сходимость МНЛ остается сверхлинейной, что связано с наличием вспомогательных переменных: прямая сходимость по всем переменным, включая вспомогательные, тем не менее, является линейной, как и сходимость невязки (2.68) системы Лагранжа (2.64).

Тестирование на MacMPEC проводилось так же, как в разделе А.2. Для МНЛ результаты следующие: для 22 задач из 161 не было сделано ни одного запуска ускорителя; еще для 29 задач все запуски ускорителя завершились неудачей; для 82 из оставшихся 110 имела место сверхлинейная сходимость невязки системы Лагранжа вдоль траектории ускорителя. Для СМНЛ ускоритель также не запускался ни разу для 22 задач; для 39 задач все запуски ускорителя были неудачными; из оставшихся 100 задач сверхлинейная сходимость

также наблюдалась в большинстве случаев (для 74 задач). Отметим, что сверхлинейная сходимость МНЛ во многих случаях связана с описанной выше эвристикой, реализованной в ALGENCAN: ее отключение практически всегда приводит к линейной скорости сходимости.

Сравнение ALGENCAN с двумя вариантами ускорителя по робастности и эффективности, аналогичное эксперименту в разделе А.3, также не выявляет существенных различий в их поведении.

Таким образом, замена используемого в ALGENCAN стандартного ускорителя МНЛ на СМНЛ не приводит к сколько-нибудь существенному повышению эффективности или робастности солвера.

Приложение Б

Численные результаты для стабилизированного метода последовательного квадратичного программирования

Данное приложение посвящено численному тестированию различных подходов к глобализации сходимости стабилизированного метода последовательного квадратичного программирования, предложенных в главе 3. В разделе Б.1 рассматриваются гибридные подходы к глобализации сходимости. Раздел Б.2 посвящен подходу, основанном на комбинировании sSQP с методами модифицированных функций Лагранжа. Наконец, в разделе Б.3 рассматривается подход к глобализации сходимости sSQP с помощью точных гладких штрафных функций.

Б.1. Гибридная глобализация сходимости

В разделе 3.1 была теоретически обоснована глобальная сходимость и сверхлинейная скорость сходимости алгоритмов 3.1 и 3.2. Вычислительный эксперимент, результаты которого представлены в данном разделе, имел целью оценку практической перспективности гибридной глобализации сходимости метода sSQP.

Эксперименты проводились в среде MATLAB. Все подзадачи квадратичного программирования решались алгоритмом активного множества посредством встроенной функции `quadprog` среды MATLAB. В эксперименте использовались тестовые коллекции DEGEN [42] и Hock–Schittkowski [61].

Коллекция Hock–Schittkowski — известное средство численного тестирования алгорит-

мов условной оптимизации общего назначения. Основная цель экспериментов на данной коллекции — выяснить объем дополнительных вычислительных затрат, связанных с использованием шага sSQP, на невырожденных задачах, когда каких-либо преимуществ от использования этого шага ожидать не приходится. Иными словами, результат можно считать положительным, если тестируемые гибридные алгоритмы будут не слишком сильно проигрывать используемому в них методу внешней фазы без попыток переключения на sSQP. В экспериментах использовалась версия коллекции Hock–Schittkowski, реализованная в среде MATLAB, которая включает в себя 101 тестовую задачу, каждая из которых содержит рекомендуемую начальную точку.

В экспериментах на коллекции DEGEN использовались все задачи, кроме 2DD01-50h, 2DD01-50v, 2DD01-500h, 2DD01-500v, 2DD01-5000h и 2DD01-5000v (остается 105 задач). На задачах из DEGEN можно ожидать выигрыша от переключения на шаги sSQP.

Ниже используются следующие обозначения для сравниваемых алгоритмов.

1. SQP — глобализованный посредством одномерного поиска квазиньютоновский метод SQP (см. [11, алгоритм 5.4.1]) с выбором матрицы H_k по правилу Бройдена–Флетчера–Голдфарба–Шанно (BFGS) с модификацией Пауэлла (см. [88, раздел 18.3]).
2. sSQP (B) — алгоритм 3.1, в котором в качестве метода внешней фазы используется метод из п. 1.
3. sSQP (R1) — алгоритм 3.2, в котором в качестве метода внешней фазы используется метод из п. 1, причем если после принятого шага sSQP осуществлялся шаг метода внешней фазы, то последний выполнялся с единичной матрицей H_k .
4. sSQP (R2) — то же, что и алгоритм из п. 3, но со следующим отличием: если после принятого шага sSQP осуществлялся шаг метода внешней фазы, то последний выполнялся с матрицей H_k , получаемой по формуле BFGS с модификацией Пауэлла из последней матрицы, использовавшейся на шаге метода внешней фазы.

Параметры в [11, алгоритм 5.4.1] были выбраны следующим образом: $\bar{c} = 1$, $\varepsilon = 0.1$, $\theta = 0.5$. Для алгоритмов 3.1 и 3.2 использовалось значение параметра $q = 0.9$. Процесс останавливался при выполнении неравенства $\sigma(x^k, \lambda^k, \mu^k) \leq 10^{-8}$. Запуск считался неудачным, если это условие остановки не было достигнуто за 500 итераций, или если алгоритм не мог осуществить очередную итерацию метода внешней фазы.

Сравнение алгоритмов проводилось по количеству итераций, причем для алгоритма с возвратами в это количество включаются отброшенные шаги sSQP.

Для коллекции Hock–Schittkowski осуществлялся запуск из рекомендуемой в коллекции точки x^0 и из $\lambda^0 = 0$, $\mu^0 = 0$. Для коллекции DEGEN осуществлялось 100 запусков из случай-

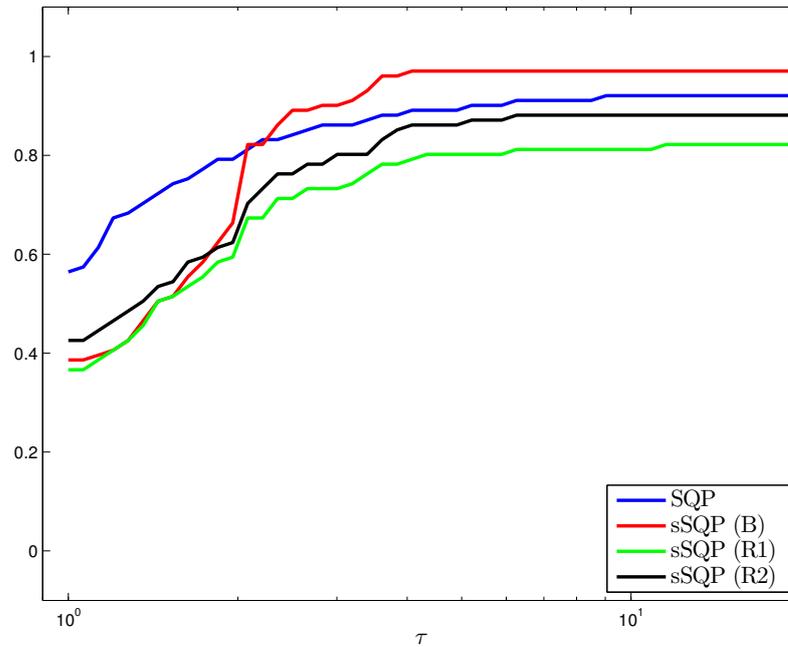


Рис. 1. Результаты на коллекции Hock–Schittkowski.

но сгенерированных начальных точек в кубической окрестности прямо-двойственного решения, размер которой определялся параметром $\delta > 0$. Точки выбирались удовлетворяющими простым ограничениям задачи вида $a \leq x \leq b$ (при их наличии) и условиям неотрицательности на множители, отвечающие ограничениями-неравенствам.

Результаты на коллекции Hock–Schittkowski представлены на рис. 1 в форме «performance profiles» из [45] (см. описание в приложении А). Рис. 1 демонстрирует, что ожидаемый проигрыш глобализованных методов sSQP на коллекции Hock–Schittkowski является вполне умеренным. Примерно для 40% задач каждый из этих алгоритмов показывает наилучший результат (заметим, что сразу несколько алгоритмов могут показывать одинаковый наилучший результат), и примерно для 80% задач результат sSQP с возвратами хуже наилучшего не более чем в 2 раза. Более того, алгоритм sSQP с возвратами демонстрирует наивысшую робастность.

Для представления результатов для коллекции DEGEN использовалась модификация подхода из [45], учитывающая многократность запусков алгоритмов для каждой задачи (см. описание в приложении А).

Чтобы дать представление о локальном эффекте от стабилизации, на рис. 2(а), приведены результаты на коллекции DEGEN при размере окрестности для начальных точек $\delta = 0.001$. Выигрыш sSQP перед квазиньютоновским SQP очевиден, как по эффективности, так и по робастности. Однако, уже при $\delta = 0.1$ выигрыш становится не столь очевидным (рис. 2(б)), при $\delta = 1$ выигрыш по эффективности исчезает вовсе (рис. 2(в)), а при $\delta = 10$ уже наблюдается существенный проигрыш по эффективности, хотя по-прежнему наблюда-

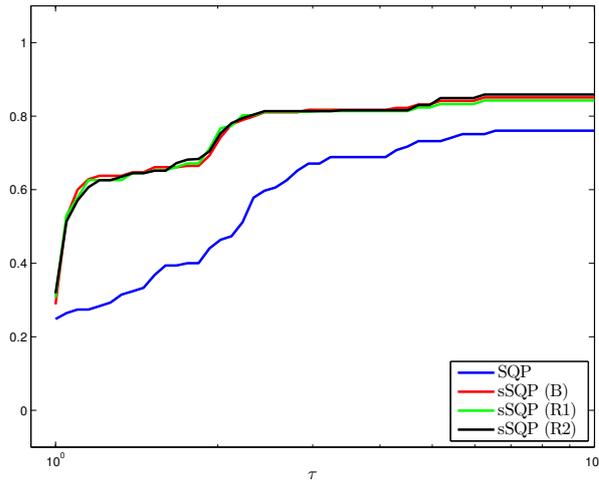
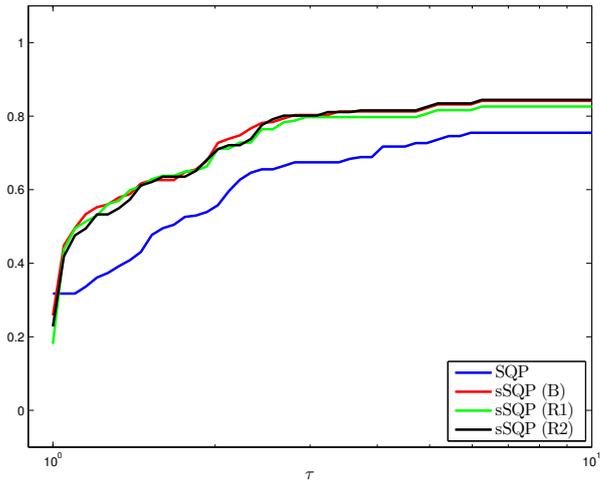
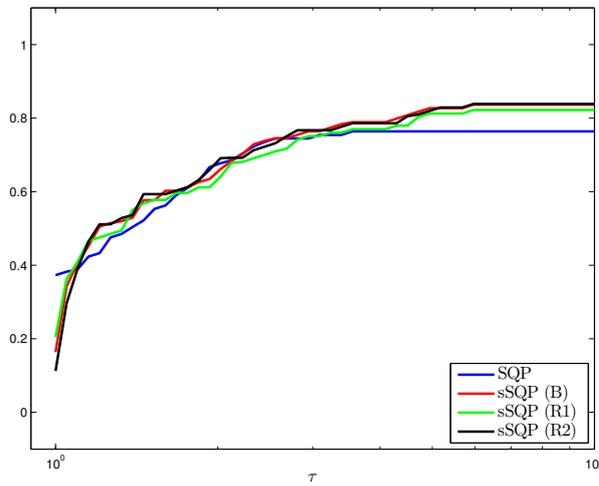
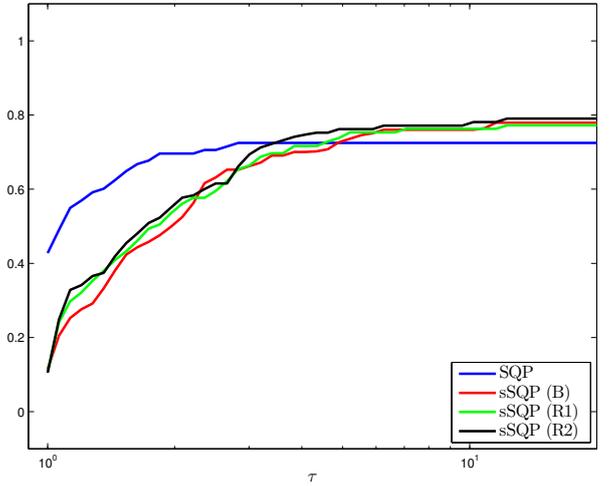
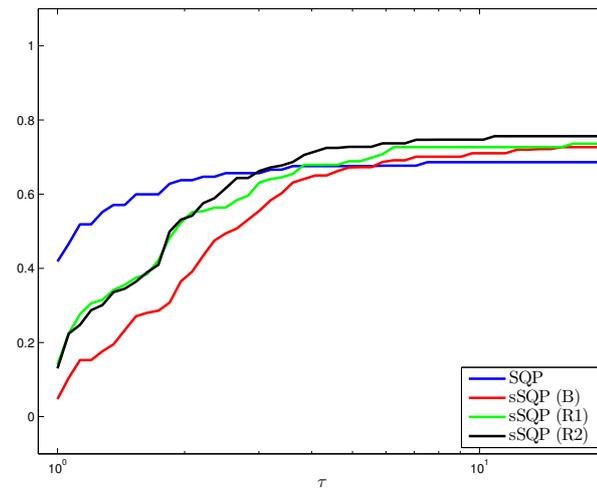
(a) $\delta = 0.001$ (б) $\delta = 0.1$ (в) $\delta = 1$ (г) $\delta = 10$ (д) $\delta = 1000$

Рис. 2. Результаты на коллекции DEGEN.

ется некоторый выигрыш по робастности (рис. 2(г)). Похожая ситуация сохраняется и при больших значениях δ (рис. 2(д)).

Б.2. Глобализация сходимости с помощью модифицированных функций Лагранжа

В данном разделе приводятся результаты вычислительного эксперимента с алгоритмом 3.3, который комбинирует свойства sSQP и МФЛ. Согласно теоретическим результатам, приведенным в разделе 3.2.1, можно ожидать, что вблизи решения, удовлетворяющего SOSC, алгоритм 3.3 будет иметь сверхлинейную скорость сходимости. Однако, как для sSQP, так и для МФЛ обычно существуют широкие (хотя и не такие широкие, как для SQP) области притяжения двойственных траекторий к критическим множителям (см. раздел 1.3, а также работы [9, 17, 70]), причем в случае сходимости к критическому множителю метод sSQP обычно имеет лишь линейную скорость сходимости. Таким образом, основной вопрос состоит в том, насколько типичной для алгоритма 3.3 является асимптотическая сверхлинейная сходимость, и влечет ли она более высокую общую эффективность алгоритма по сравнению с имеющимися альтернативами.

Эксперименты проводились на коллекции DEGEN [42], из которой были исключены задачи 2DD01-50h, 2DD01-50v, 2DD01-500h, 2DD01-500v, 2DD01-5000h и 2DD01-5000v, а также задача 2DD01-5v. В большинстве экспериментов использовались все 104 оставшиеся задачи этой коллекции, однако в некоторых случаях использовались только задачи с ограничениями-равенствами (всего 34 задачи). Для каждой задачи проводилось по 20 запусков из случайных начальных точек $(x^0, \lambda^0, \mu^0) \in \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}_+^m$, удовлетворяющих неравенству $\|(x^0, \lambda^0, \mu^0)\|_\infty \leq 100$. Запуск считался успешным, если для некоторого номера $k \leq 500$ выполнялся критерий остановки $\sigma(x^k, \lambda^k, \mu^k) \leq 10^{-8}$.

В алгоритме 3.3 значения параметров выбирались таким образом, чтобы по возможности уменьшить количество внутренних итераций при фиксированном k . В частности, использовались значения $r_0 = 10^4$, $\varepsilon_0 = 10^2$, $q = 0.5$ и $\theta = 0.5$. Согласно полученному вычислительному опыту, чем слабее используемые критерии принятия итераций sSQP или МФЛ, тем выше скорость сходимости. Причина этого заключается, видимо, в том, что на внутренних итерациях двойственное приближение не обновляется, и при этом используется только часть информации, полученной в ходе решения подзадачи sSQP (а именно, только прямая компонента ξ вычисляемого направления поиска).

В правиле Армихо (3.68) использовались параметры $\tau = 0.5$ и $\varepsilon = 0.1$, а на итерации МФЛ — параметры $\delta = 0.5$, $\kappa = 0.1$. Остальные параметры были выбраны следующим

образом: $\gamma = 1$, $\bar{\lambda}_{\min} = -10^{10}$, $\bar{\lambda}_{\max} = \bar{\mu}_{\max} = 10^{10}$. При этом в каждом из запусков полагалось $\hat{x}^0 = x^0$, $\bar{\lambda}^0 = \lambda^0$ и $\bar{\mu}^0 = \mu^0$. Выбор параметра σ_0 будет обсуждаться ниже.

Матрица H_j на шаге 1 выбиралась следующим образом:

$$H_j = \frac{\partial^2 L}{\partial x^2}(\hat{x}^j, \bar{\lambda}^k, \bar{\mu}^k) + \max\{\gamma - \gamma_{\min}, 0\}I,$$

где γ_{\min} — минимальное собственное значение матрицы $\frac{\partial^2 L}{\partial x^2}(\hat{x}^j, \bar{\lambda}^k, \bar{\mu}^k) + (h'(\hat{x}^j))^T h'(\hat{x}^j) / \sigma_k$. При таком выборе матрица H_j удовлетворяет соотношению (3.24) для всех номеров k и j , и поэтому, согласно обсуждению после описания алгоритма 3.3, система (3.13)–(3.15) имеет единственное решение, удовлетворяющее (3.16). Отметим, что такой выбор H_j является достаточно дорогим с вычислительной точки зрения и используется исключительно для упрощения реализации. Однако, использование других процедур выбора H_j , которые обеспечивают выполнение (3.24), но по возможности стараются приблизить H_j к настоящей матрице Гессе, не должно сильно влиять на поведение алгоритма 3.3.

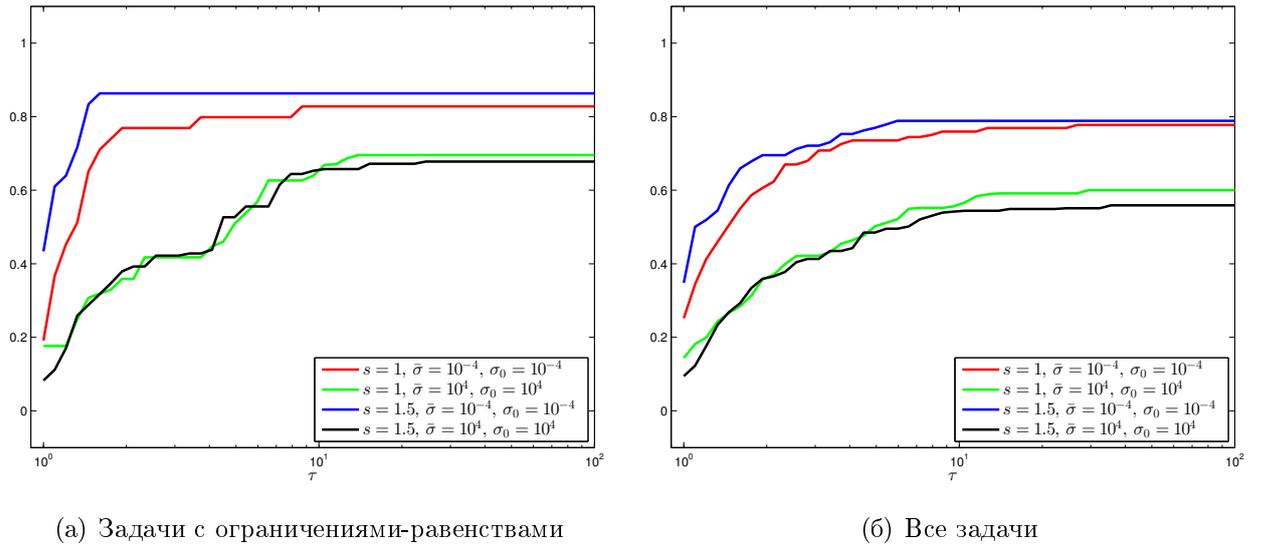
Наконец, в случае выполнения одного из тестов (3.17) или (3.22) использовалось несколько модифицированное правило обновления параметра σ_k :

$$\sigma_{k+1} = \min\{\sigma(x^{k+1}, \lambda^{k+1}, \mu^{k+1})^s, \bar{\sigma}\},$$

где $\bar{\sigma} > 0$ и $s > 0$ — фиксированные числа. Если $s = 1$, то все полученные в разделе 3.2.1 теоретические результаты остаются справедливыми. Если же $s > 1$, то сверхлинейная сходимость алгоритма 3.3 может теряться, поскольку параметр стабилизации стремится к нулю слишком быстро, и даже вблизи квалифицированного решения шаг метода sSQP может не обеспечивать сверхлинейного убывания расстояния до множества решений. Если параметр стабилизации становится слишком мал, то шаг метода sSQP мало отличается от шага обычного метода SQP. При этом, как было показано в главе 1, типичным поведением SQP на вырожденных задачах является притяжение двойственных траекторий к критическим множителям Лагранжа и линейная скорость сходимости. Отметим, что аналогичный эффект возникает также если значение $\bar{\sigma}$ слишком мало. В последнем случае на ранних итерациях шаги sSQP мало отличаются от шагов SQP, вследствие чего двойственная траектория успевает достаточно сильно приблизиться к множеству критических множителей, вблизи которых для sSQP также имеет место притяжение к таким множителям и линейная скорость сходимости.

Результаты в данном разделе представлены в форме «performance profiles». В качестве показателя эффективности алгоритма 3.3 использовалось суммарное количество всех итераций (sSQP, МФЛ и внутренних), т. е. количество решенных систем на шаге 1.

На рис. 3 представлено сравнение эффективности алгоритма 3.3 при различных значениях s , $\bar{\sigma}$ и σ_0 . Как уже упоминалось выше, малые значения $\bar{\sigma}$ или значения $s > 1$ практически

Рис. 3. Алгоритм 3.3 при различных значениях s , $\bar{\sigma}$ и σ_0 .

исключают возможность сверхлинейной сходимости алгоритма, фактически превращая его в SQP. Тем не менее, из рис. 3 видно, что увеличение параметра s приводит к повышению общей эффективности алгоритма, а при $\bar{\sigma} = 10^{-4}$ его эффективность намного выше, чем при $\bar{\sigma} = 10^4$. Одна из причин этого состоит в том, что при больших значениях $\bar{\sigma}$ обратный параметр штрафа на начальных итерациях часто оказывается слишком большим, и это приводит к тому, что решения подзадач МФЛ оказываются далекими от решений исходной задачи. Таким образом, возможная сверхлинейная сходимость при больших $\bar{\sigma}$ не компенсирует отсутствие серьезного прогресса на ранних итерациях. Отметим также, что даже при больших значениях $\bar{\sigma}$ сверхлинейная сходимость наблюдается не слишком часто. Во многих случаях это является следствием выбора матрицы H_j , отличной от настоящей матрицы Гессе функции Лагранжа в точке $(x^j, \bar{\lambda}^k, \bar{\mu}^k)$. Другие такие случаи связаны с притяжением двойственной траектории к критическим множителям Лагранжа.

С учетом полученных результатов, в дальнейших экспериментах в алгоритме 3.3 будем использовать значения $s = 1.5$, $\bar{\sigma} = 10^{-4}$, $\sigma_0 = 10^{-4}$. При таких значениях параметров направления поиска (ξ^j, η^j, ζ^j) , получаемые на шаге 1 алгоритма, оказываются близкими к направлениям поиска метода SQP, и поэтому алгоритм 3.3 сочетает в себе свойства SQP и МФЛ.

На рис. 4 показано сравнение алгоритма 3.3 с SQP и МФЛ. В сравнении использовалась квазиньютоновская версия SQP, глобализованная посредством одномерного поиска для ℓ_1 -штрафной функции; матрица H_k выбиралась по правилу BFGS с модификацией Пауэлла (см. описание алгоритма SQP в разделе Б.1). В реализации МФЛ для решения подзадач

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

также использовался глобализованный посредством одномерного поиска квазиньютоновский

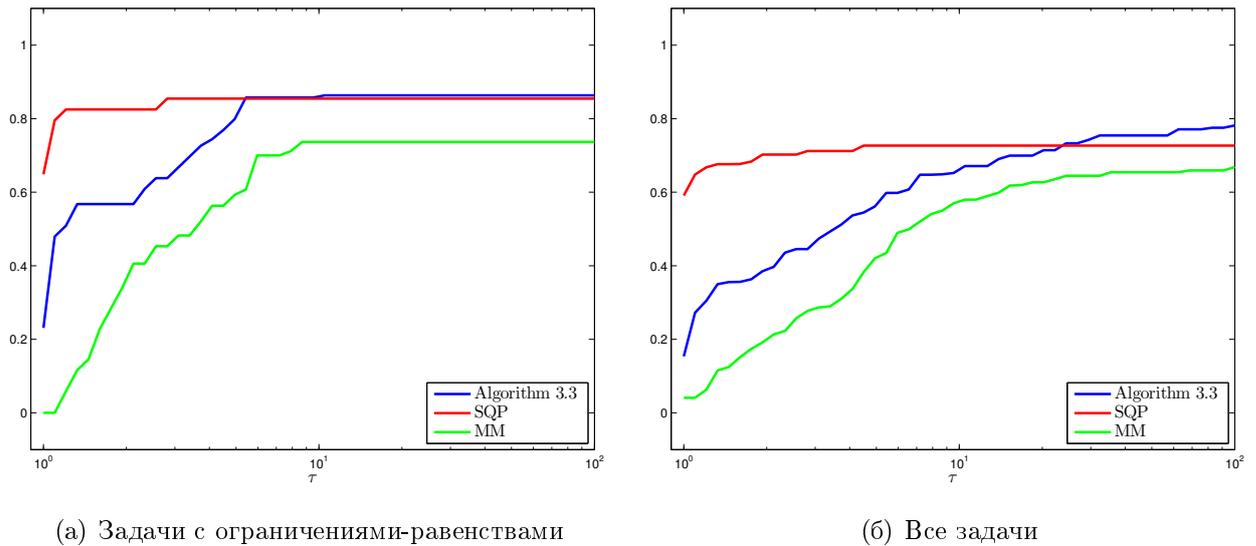


Рис. 4. Сравнение алгоритмов.

метод, в котором матрица H_k выбиралась по правилу BFGS. Управление точностью решения подзадач, а также обновление векторов λ^k , μ^k , $\bar{\lambda}^k$, $\bar{\mu}^k$ осуществлялось так же, как в случае итерации МФЛ в алгоритме 3.3. Для обновления σ_k использовалось правило (3.23) (с теми же параметрами); начальное значение σ_0 было равно 10^{-4} .

В качестве показателя эффективности для SQP использовалось число внешних итераций (т.е. количество решенных подзадач квадратичного программирования), а для МФЛ — число внутренних итераций (т.е. количество решенных ньютоновских линейных систем). При отсутствии ограничений-неравенств эти показатели для всех алгоритмов соответствуют числу решенных линейных систем, и поэтому достаточно согласованы. Отметим, однако, что метод SQP решает линейные системы размера $(n + l) \times (n + l)$, а алгоритм 3.3 и МФЛ — системы размера $n \times n$. В частности, в алгоритме 3.3 прямую компоненту направления поиска можно найти как решение системы (32), после чего двойственная компонента находится по явной формуле (31) (решение линейных систем меньшего размера можно рассматривать как практическое преимущество sSQP над SQP). Таким образом, для задач с ограничениями-равенствами итерации SQP несколько дороже, чем у других алгоритмов. Для задач со смешанными ограничениями решение подзадач квадратичного программирования существенно дороже, чем решение линейных систем. Сравнение с МФЛ приводится для того, чтобы убедиться, что более дорогие итерации алгоритма 3.3 позволяют увеличить скорость сходимости по сравнению с МФЛ.

Как видно из рис. 4, алгоритм 3.3 существенно опережает МФЛ по эффективности, причем в большинстве случаев выигрыш достигается за счет итераций sSQP. Однако, алгоритм 3.3 значительно уступает методу SQP, что в значительной степени связано с «отбрасыванием» двойственной компоненты направления поиска, если не принимается итерация

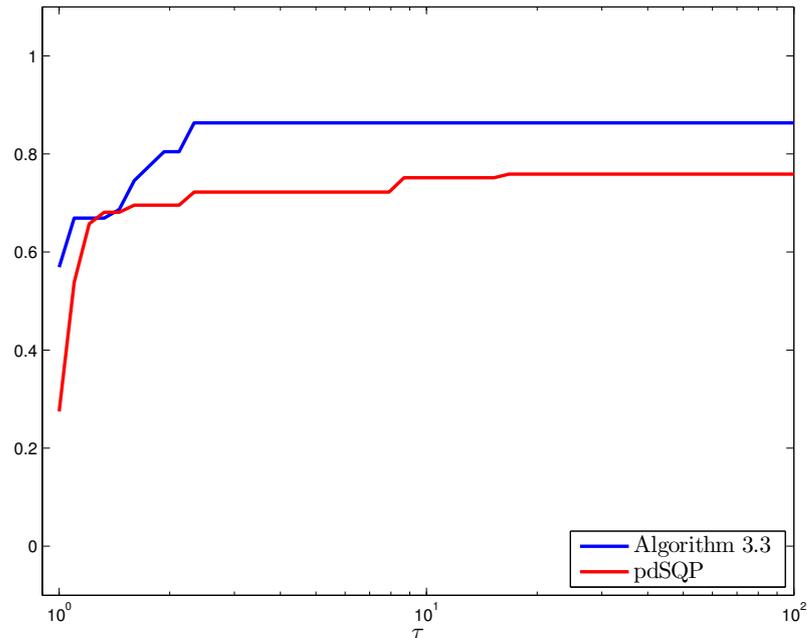


Рис. 5. Сравнение алгоритмов.

sSQP. Отметим, что по робастности алгоритм 3.3 несколько опережает остальные методы.

Наконец, на рис. 5 показано сравнение алгоритма 3.3 и алгоритма pdSQP на задачах с ограничениями-равенствами. В реализации алгоритма pdSQP использовались рекомендованные значения параметров из [56], за исключением точности в критерии останова, которая полагалась равной 10^{-8} . Как видно из рис. 5, алгоритм 3.3 опережает алгоритм pdSQP как по эффективности, так и по робастности. В ряде случаев выигрыш по эффективности объясняется тем, что в алгоритме 3.3 сначала делается попытка принять полный шаг метода sSQP, и только в случае ее неудачи осуществляется одномерный поиск.

Б.3. Глобализация сходимости с помощью точных гладких штрафных функций

В данном разделе представлены результаты вычислительного эксперимента с алгоритмом 3.4 на задачах с нерегулярными ограничениями. Как было показано в разделе 1.3.1, для sSQP эффект притяжения двойственных траекторий к критическим множителям Лагранжа присутствует, однако значительно менее выражен, чем для SQP. Основной вопрос состоит в том, сможет ли алгоритм 3.4 благодаря возможной сверхлинейной сходимости показать более высокую эффективность, чем обычный SQP. Для этого, в частности, необходимо, чтобы в случае сходимости к критическим множителям алгоритм 3.4 не слишком сильно проигрывал SQP по эффективности.

В данном эксперименте алгоритм 3.4 сравнивался с квазиньютоновской версией SQP,

глобализованной посредством одномерного поиска для ℓ_1 -штрафной функции; матрица Гессе функции Лагранжа выбиралась по правилу BFGS с модификацией Пауэлла (см. описание алгоритма SQP в разделе Б.1). В алгоритме 3.4 использовались параметры $\delta = 10$, $\varepsilon = 0.3$, $\theta = 0.5$. Начальные значения параметров штрафа были взяты из [35]: $c_1 = 100$, $c_2 = 0.01$. В тестах (3.64)–(3.67) использовались параметры $\rho = 0.1$, $q = 2$ и функции $\psi_1(t) = \psi_3(t) = 0.5t$ и $\psi_2(t) = 0.5t^2$. Выбор $\bar{\sigma}$ будет обсуждаться далее.

В экспериментах использовались все задачи с ограничениями-равенствами из коллекции DEGEN [42], кроме 2DD01-5v, 2DD01-50h, 2DD01-50v, 2DD01-500h, 2DD01-500v, 2DD01-5000h и 2DD01-5000v. Для каждой из оставшихся 34 задач проводилось по 20 запусков из случайных начальных точек $(x^0, \lambda^0) \in \mathbb{R}^n \times \mathbb{R}^l$, удовлетворяющих неравенству $\|(x^0, \lambda^0)\|_\infty \leq 100$. Запуск считался успешным, если для некоторого номера $k \leq 500$ выполнялся критерий остановки $\sigma(x^k, \lambda^k) \leq 10^{-8}$.

В первой части экспериментов использовалось большое значение $\bar{\sigma}$, а именно, $\bar{\sigma} = 10^4$. Относительная эффективность алгоритмов представлена на рис. 6(а) в форме, предложенной в работе [84]. Для двух заданных алгоритмов А и В и для каждой задачи p определим величину

$$r_p^{AB} = -\log_2(k_p^A/k_p^B),$$

где k_p^A и k_p^B обозначают среднее число итераций соответствующего алгоритма среди успешных запусков для задачи p . Если все запуски алгоритма А или В для задачи p являются неудачными, будем полагать значение r_p^{AB} равным соответственно -5 или 5 (если все запуски обоих алгоритмов являются неудачными, соответствующая задача не показана на диаграмме). В частности, значение r_p^{AB} положительно, если алгоритм А опережает В по числу итераций, и отрицательно в противном случае; при этом результат победителя лучше в $2^{|r_p^{AB}|}$ раз. Задачи на диаграмме расположены в порядке убывания величины $|r_p^{AB}|$. Верхняя часть рис. 6(а) соответствует алгоритму 3.4, а нижняя — методу SQP.

Эти же результаты представлены на рис. 6(б) в форме «performance profiles». Из рис. 6 видно, что алгоритм 3.4 более эффективен для большинства задач. С другой стороны, робастность алгоритма 3.4, а также его эффективность для некоторых задач ниже, чем у SQP.

Рассмотрим подробнее задачи, на которых алгоритм 3.4 показывает низкую эффективность. Во-первых, 9 из этих задач имеют неизолированные решения. Несколько неожиданно, квазиньютоновская версия SQP демонстрирует очень хорошее поведение на таких задачах: практически всегда наблюдается сверхлинейная сходимость к некоторому решению. Глобализованный sSQP, напротив, обычно сходится достаточно медленно. Еще в 4 задачах решению не отвечает ни один множитель Лагранжа. Для таких задач нет оснований ожидать, что двойственная стабилизация будет давать положительный эффект, и во многих случаях скорость сходимости sSQP оказывается лишь линейной.

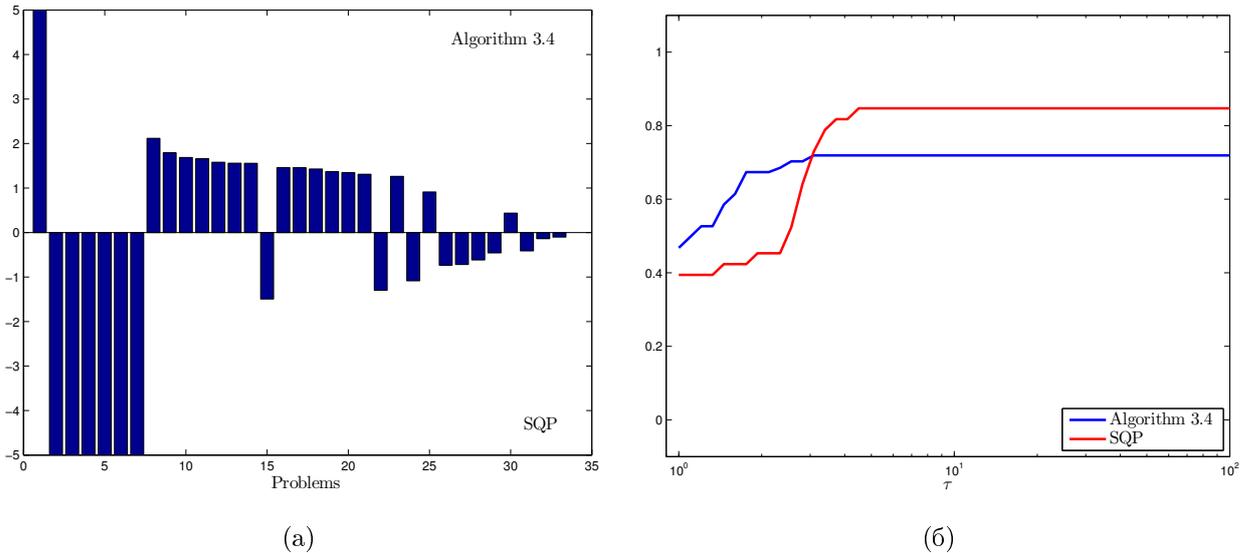


Рис. 6. Результаты для всех задач с ограничениями-равенствами из DEGEN.

Данные два класса вырожденных задач (задачи с неизолрованными решениями и задачи без множителей Лагранжа), конечно, являются очень специальными, и разумно исключить их из эксперимента, тем самым улучшив относительную эффективность алгоритма 3.4. Вместе с тем, как показали эксперименты, медленная сходимость sSQP может иметь место в гораздо более типичной ситуации, а именно, в задачах с неполным вырождением, т. е. когда в решении \bar{x} выполнено $h'(\bar{x}) \neq 0$. В коллекции DEGEN только 4 таких задачи с ограничениями-равенствами (хотя на практике такая ситуация намного более типична, чем полное вырождение). Если их также исключить эксперимента, получим картину, показанную на рис. 7. Алгоритм 3.4 теперь значительно опережает SQP по эффективности, и имеет почти такую же робастность. Более того, алгоритм 3.4 уступает SQP по эффективности только на задаче 20113, которая также является достаточно специальной: все множители Лагранжа, отвечающие решению, являются критическими.

Медленная сходимость алгоритма 3.4 на задачах с неполным вырождением заслуживает более подробного рассмотрения. Отметим, что такое поведение обычно наблюдается, когда принимаются полные шаги sSQP (т. е. $\alpha_k = 1$), и поэтому оно не может быть связано с какими-либо недостатками предложенного подхода к глобализации сходимости, а скорее является свойством самого метода sSQP вдали от решения. Для лучшего понимания происходящего рассмотрим следующий невырожденный пример.

Пример 1. Пусть $n = l = 1$, $f(x) = x^2/2$, $h(x) = x$. Тогда $\bar{x} = 0$ — единственное решение задачи (3.49). Это решение является невырожденным, и ему отвечает единственный множитель Лагранжа $\bar{\lambda} = 0$.

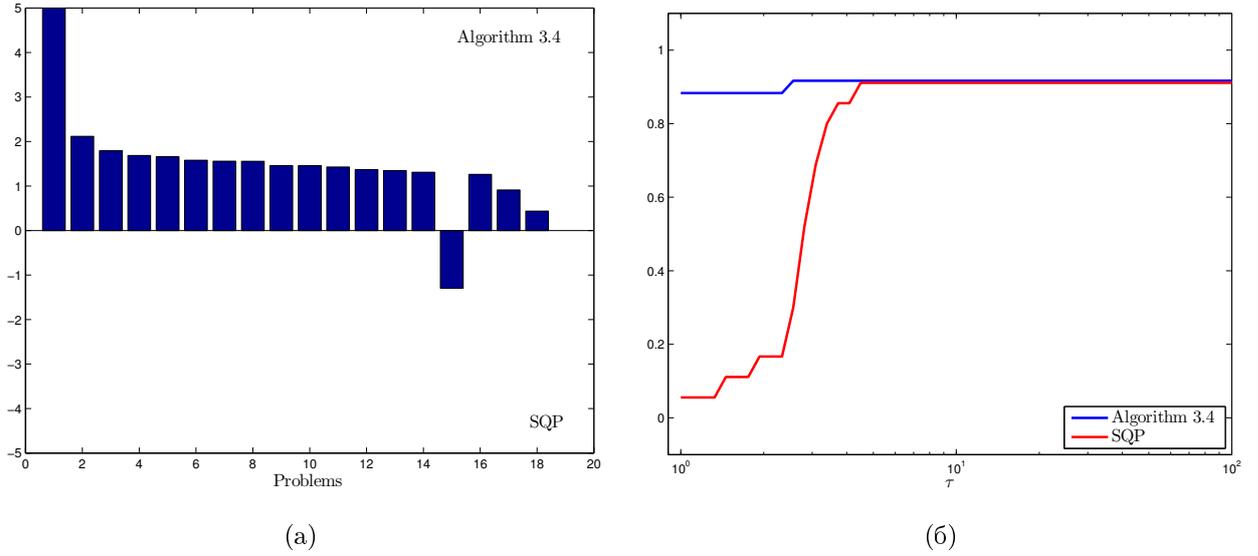


Рис. 7. Результаты для выбранных задач с ограничениями-равенствами из DEGEN.

Итерационная система sSQP (3.53) принимает вид

$$\begin{pmatrix} 1 & 1 \\ 1 & -\sigma \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix} = - \begin{pmatrix} x + \lambda \\ x \end{pmatrix},$$

и поэтому метод генерирует направление

$$\begin{pmatrix} \xi^k \\ \eta^k \end{pmatrix} = \frac{1}{\sigma_k + 1} \begin{pmatrix} -\sigma_k & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x^k + \lambda^k \\ x^k \end{pmatrix} = \frac{1}{\sigma_k + 1} \begin{pmatrix} -(\sigma_k + 1)x^k - \sigma_k \lambda^k \\ -\lambda^k \end{pmatrix}.$$

Таким образом, $x^{k+1} = -S_k \lambda^k$, $\lambda^{k+1} = S_k \lambda^k$, где $S_k = \sigma_k / (\sigma_k + 1)$. Если значение σ_k велико (например, $\sigma_k = \|\Phi(x^k, \lambda^k)\|$ для (x^k, λ^k) , далеких от решения), то S_k близко к 1, и поэтому приближения меняются очень медленно. С другой стороны, если σ_k мало, то S_k близко к 0, что означает быструю скорость сходимости.

Траектория чистого метода sSQP для этого примера показана на рис. 8(a). Критерий остановки был выполнен только после 38 итераций, даже несмотря на то, что асимптотическая скорость сходимости по-прежнему сверхлинейная.

Поведение sSQP, наблюдаемое в примере 1, является типичным для задач с неполным вырождением: вдали от решения метод часто генерирует длинную последовательность коротких шагов до того, как траектория попадет в область, где проявляется сверхлинейная сходимость. Такое поведение уже наблюдалось в некоторых примерах из работы [85] (в которой приводятся результаты эксперимента с sSQP без какой-либо глобализации сходимости). Ситуацию можно неформально описать следующим образом. Вокруг решения имеется «малая» область O_1 , в которой гарантируется сверхлинейная сходимость, и «большая» область O_2 , за пределами которой уже нельзя ожидать хороших свойств от направления sSQP. Короткие шаги возникают, если траектория попадает в область $O_2 \setminus O_1$. В задачах с полным

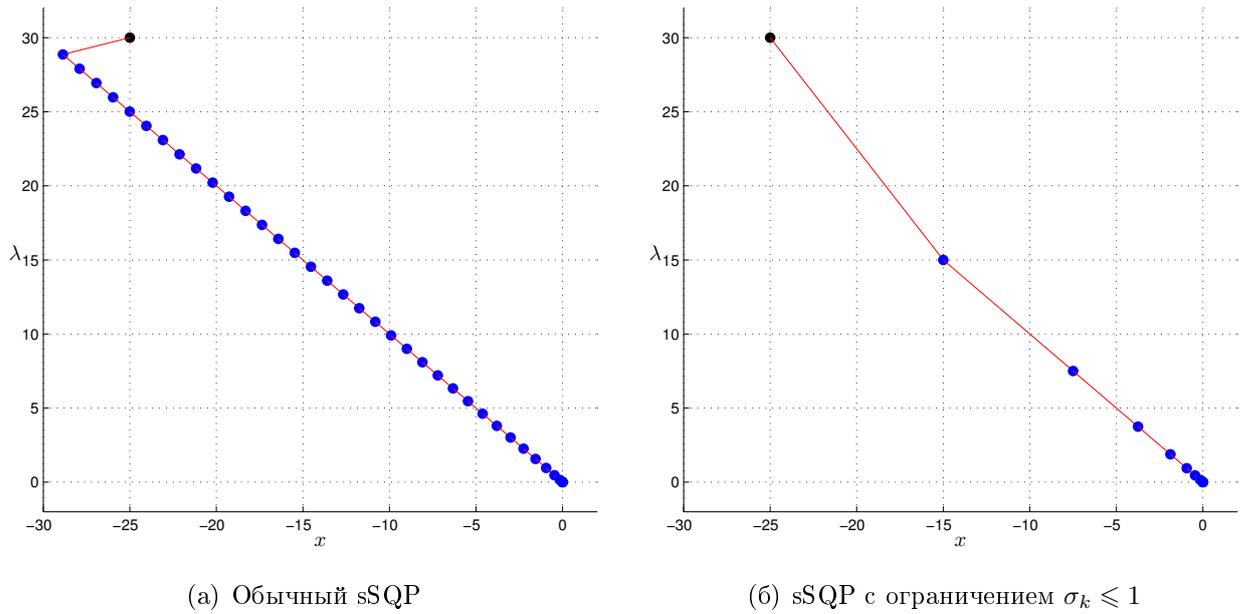


Рис. 8. Прямо-двойственные траектории для примера 1; $(x^0, \lambda^0) = (-25, 30)$.

вырождением область O_1 обычно весьма велика (во всяком случае, она содержит область, из которой выбирались начальные точки в данном эксперименте), в то время как в других задачах имеет место противоположная ситуация.

Судя по всему, если в некотором решении у матрицы Якоби ограничений есть ненулевые сингулярные числа, то у sSQP возникает эффект «отталкивания» от такого решения, который и приводит к замедлению сходимости. Рассмотрим следующую аналогию из физики. Протоны представляют собой положительно заряженные частицы в ядре атома, и поэтому они отталкиваются друг от друга вследствие электромагнитных сил. Тем не менее, ядро не распадается, поскольку на очень малых расстояниях протоны притягиваются друг к другу за счет так называемого сильного взаимодействия. Если расстояние между частицами немного увеличить, то электромагнитное взаимодействие начнет доминировать, и сильное взаимодействие не сможет больше удерживать протоны в ядре. В случае sSQP ненулевые сингулярные числа матрицы Якоби приводят к эффекту, аналогичному «электромагнитному отталкиванию», в то время как локальная сверхлинейная сходимость соответствует притяжению за счет сильного взаимодействия.

Потенциальная возможность снизить негативные эффект, описанный выше, состоит в уменьшении значения $\bar{\sigma}$ в алгоритме 3.4. Чем меньше это значение, тем ближе итерационная система sSQP к системе обычного SQP. В частности, малое верхнее ограничение на σ_k в примере 1 приводит к малым значениям S_k , а значит к быстрой сходимости. На рис. 8(б) показана траектория sSQP, в котором значение σ_k ограничено сверху константой 1. В этом случае критерий остановки был выполнен после 11 итераций (при отсутствии ограничений требуется 38 итераций; см. рис. 8(а)).

С другой стороны, чем ближе итерационная система sSQP к системе SQP, тем сильнее эффект притяжения к критическим множителям, и поэтому уменьшение $\bar{\sigma}$ может привести к уменьшению скорости сходимости. Чтобы продемонстрировать этот эффект, снова рассмотрим задачу из примера 1.3. Пусть $n = l = 1$, $f(x) = x^2$, $h(x) = x^2$. Тогда $\bar{x} = 0$ — единственное решение задачи (3.49), причем $\mathcal{M}(0) = \mathbb{R}$, и единственным критическим множителем Лагранжа является $\bar{\lambda} = -1$. Легко проверить, что из (3.81) следует

$$x^{k+1} = s_k x^k, \quad \lambda^{k+1} = \lambda^k - s_k(\lambda^k + 1),$$

где

$$s_k = \left(2 + \frac{\sigma_k(1 + \lambda^k)}{(x^k)^2} \right)^{-1}.$$

Таким образом, при малых значениях σ_k имеем

$$x^{k+1} \approx \frac{1}{2}x^k, \quad \lambda^{k+1} + 1 \approx \frac{1}{2}(\lambda^k + 1).$$

В частности, $\{(x^k, \lambda^k)\}$ линейно сходится к точке $(0, -1)$.

Согласно полученному вычислительному опыту, алгоритм 3.4 при $\bar{\sigma} \leq 1$ показывает достаточно низкую эффективность, однако его поведение может быть существенно улучшено с помощью некоторых модификаций, таких как поправки второго порядка (см., например, [39, раздел 17.3]) и немонотонный одномерный поиск [59]. Идея последнего состоит в том, чтобы заменить $\varphi_{c_1, c_2}(x^k, \lambda^k)$ в правой части (3.68) на

$$\max_{j=0, \dots, R-1} \varphi_{c_1, c_2}(x^{k-j}, \lambda^{k-j}).$$

В данном эксперименте оптимальным оказалось значение $R = 8$.

Наконец, в случае сходимости к критическим множителям существенное влияние на скорость сходимости оказывает управление параметром c_2 . В частности, следующее эвристическое правило, добавленное на шаге 5 алгоритма 3.4, значительно улучшает общее поведение: если $\alpha_k \leq 0.1$, заменить c_2 на $\min\{10c_2, 10^{10}\}$ (и перейти к шагу 6).

Отметим, что немонотонный одномерный поиск увеличивает общую эффективность и при больших значениях $\bar{\sigma}$, в то время как остальные указанные выше эвристики практически не оказывают влияния на эффективность в этом случае.

На рис. 9 сравнивается поведение алгоритма 3.4, в котором были реализованы описанные выше эвристики, при различных значениях $\bar{\sigma}$. Как видно из рисунка, уменьшение $\bar{\sigma}$ улучшает робастность, однако значительно ухудшает эффективность. В частности, при малых значениях $\bar{\sigma}$ данная версия алгоритма 3.4 проигрывает по эффективности квазиньютоновской версии SQP для большинства задач. Одна из причин состоит в том, что алгоритм 3.4 при малых $\bar{\sigma}$ «близок» к методу SQP, глобализованному посредством гладкой штрафной

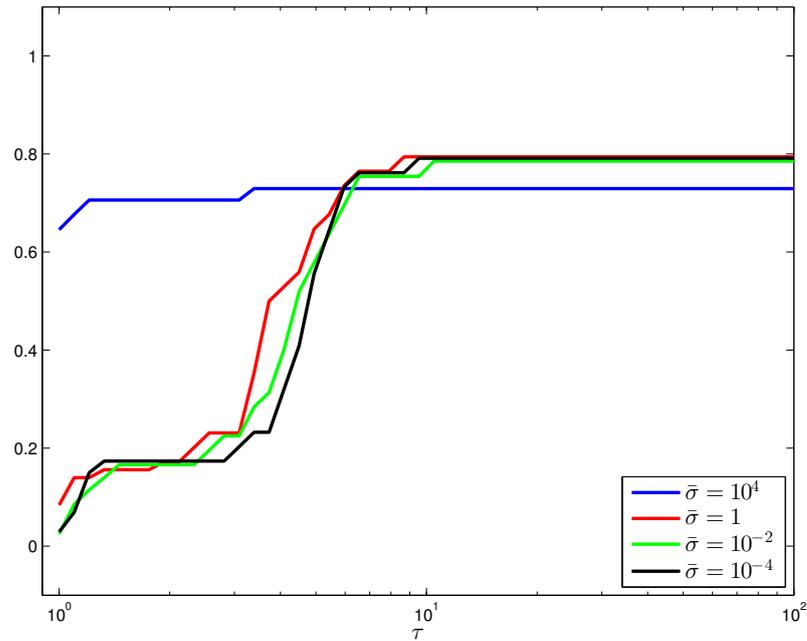


Рис. 9. Алгоритм 3.4 при различных значениях $\bar{\sigma}$.

функции, но при этом известно, что глобализации сходимости SQP с помощью негладких штрафных функций работают лучше.

Алгоритм, представленный в разделе 3.3, может быть применен к задачам со смешанными ограничениями, если переформулировать их как задачи с ограничениями-равенствами с помощью дополнительных переменных. К сожалению, вычислительный эксперимент показывает, что алгоритм 3.4 имеет достаточно низкую эффективность на задачах из DEGEN, переформулированных таким образом. Одна из причин состоит в том, что во многих случаях в результате получаются задачи с неполным вырождением, на которых для sSQP характерны длинные последовательности коротких шагов. Другая причина состоит в том, что переформулированные задачи обычно содержат «больше» критических множителей, чем исходные: введение дополнительных переменных приводит к появлению новых критических множителей (связанных с нарушением строгой дополненности).

Приложение В

Численные результаты для метода последовательного квадратичного программирования, стабилизированного вдоль подпространства

В данном приложении представлены результаты вычислительного эксперимента с методом s-sSQP и его глобализованной версией. Как было показано в приложении Б.3, метод sSQP обычно имеет медленную сходимость на задачах с неполным вырождением, поскольку в таких случаях он часто генерирует вдали от решения длинные последовательности коротких шагов. Основная цель данного эксперимента состоит в выяснении того, сможет ли стабилизация вдоль подпространства вырожденности справиться с этим негативным эффектом. Конечно, для этого необходима корректная идентификация ранга матрицы Якоби ограничений.

Будем сравнивать базовый алгоритм s-sSQP и его глобализованную версию (т. е. алгоритм 4.3) с обычными методами SQP и sSQP, а также с квазиньютоновской версией SQP, в которой матрица Гессе функции Лагранжа выбирается по правилу BFGS с модификацией Пауэлла [88, раздел 18.3], а глобализация сходимости осуществляется посредством одномерного поиска для ℓ_1 -штрафной функции (см. описание алгоритма SQP в приложении Б.1).

В алгоритме s-sSQP с исчезающей стабилизацией для всех $(x, \lambda) \in \mathbb{R}^n \times \mathbb{R}^l$ значение $\sigma(x, \lambda)$ определялось согласно (4.6) при $\beta = 1$. Далее, $P(x, \lambda)$ определялось с помощью процедуры, описанной в конце раздела 4.2, т. е. в результате работы алгоритма 4.2, применяемого к $A = h'(x)$ и $t = \tau |\sigma(x, \lambda)|^\theta$ с параметрами $\tau = 0.3$ и $\theta = 0.8$ (эти значения оказались оптимальными в эксперименте). Напомним, что если вектор (x, λ) близок к решению $(\bar{x}, \bar{\lambda})$ системы (4.2), где $\bar{\lambda}$ — некритический множитель, то данная процедура корректно определяет ранг

матрицы Якоби ограничений в точке \bar{x} . В случаях сходимости к критическим множителям величина $\sigma(x^k, \lambda^k)$ обычно становится слишком мала для больших k , и поэтому тест (4.38) не выполняется для $A = h'(x^k)$ и $t = \tau|\sigma(x^k, \lambda^k)|^\theta$ для всех $s = 0, \dots, l - 1$. Следовательно, алгоритм 4.2 возвращает значения $r = l$ и $P(x^k, \lambda^k) = 0$. В таких случаях метод s-sSQP превращается в обычный SQP, и поэтому, как обсуждалось выше, скорость сходимости является лишь линейной [72, глава 7].

В алгоритме s-sSQP с неисчезающей стабилизацией отображение P определялось так же, как описано выше, но в самом алгоритме использовалась функция $\sigma(\cdot) \equiv 1$.

В алгоритме 4.3 использовались параметры $\delta = 10$, $\varepsilon = 0.3$, $\nu = 0.5$. Начальные значения параметров штрафа были взяты из [35]: $c_1 = 100$, $c_2 = 0.01$. В тестах (4.40)–(4.43) использовались параметры $\rho = 0.1$, $q = 2$ и функции $\psi_1(t) = \psi_3(t) = 0.5t$ и $\psi_2(t) = 0.5t^2$.

Эксперименты проводились на коллекции DEGEN [42], из которой были исключены задачи 2DD01-50h, 2DD01-50v, 2DD01-500h, 2DD01-500v, 2DD01-5000h и 2DD01-5000v, а также задача 2DD01-5v (остается 104 задачи). Однако, в большинстве экспериментов использовались только задачи с ограничениями-равенствами (всего 34 задачи). Для каждой задачи проводилось по 20 запусков из случайных начальных точек $(x^0, \lambda^0, \mu^0) \in \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}_+^m$, удовлетворяющих неравенству $\|(x^0, \lambda^0, \mu^0)\|_\infty \leq 100$. Запуск считался успешным, если для некоторого номера $k \leq 500$ выполнялся критерий остановки $\sigma(x^k, \lambda^k, \mu^k) \leq 10^{-8}$.

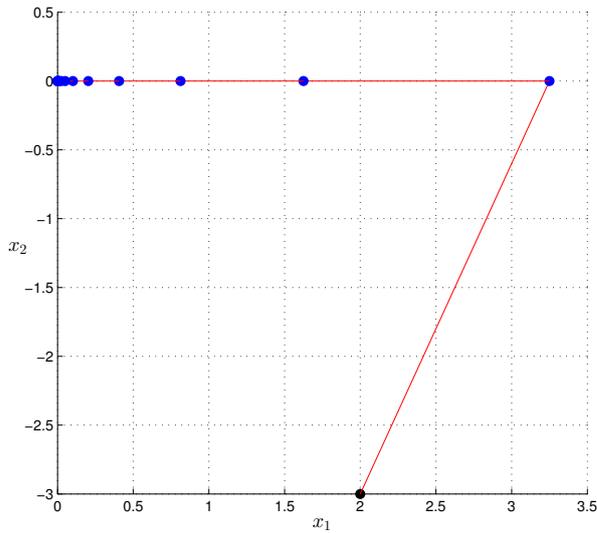
Покажем сначала, что стабилизация вдоль нужного подпространства может давать существенное повышение эффективности по сравнению с обычным sSQP.

Пример 1. Рассмотрим задачу 20204 из DEGEN. Пусть $n = l = 2$, $f(x) = (x_1^2 + x_2^2)/2$, $h(x) = ((x_1^2 + x_2^2)/2 - x_2, (x_1^2 + x_2^2)/2 + x_2)$. Тогда $\bar{x} = 0$ — единственная допустимая точка, а значит единственное решение задачи (4.1). Далее, $\text{rank } h'(\bar{x}) = 1$, и множителями Лагранжа, отвечающими \bar{x} , являются все $\lambda \in \mathbb{R}^2$, удовлетворяющие равенству $\lambda_1 = \lambda_2$. Единственным критическим множителем является $\bar{\lambda} = (-1/2, -1/2)$.

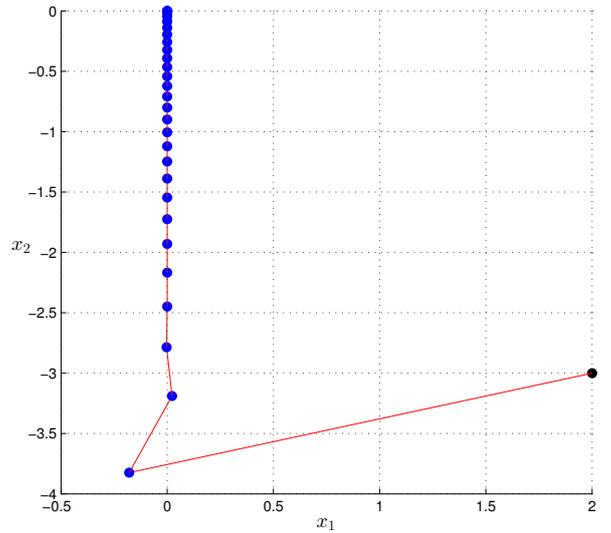
На рис. 1 показаны прямые траектории обычного SQP, стабилизированного SQP и метода s-sSQP с исчезающей и неисчезающей стабилизацией, запущенных из начальной точки $x^0 = (2, -3)$, $\lambda^0 = (-10, 15)$.

Для метода SQP критерий остановки был выполнен после 17 итераций. Из рис. 1(а) видно, что прямая траектория прижимается к $\ker h'(\bar{x})$, и скорость сходимости является линейной. Причина состоит в притяжении двойственной траектории к критическому множителю Лагранжа. Вновь подчеркнем, что наблюдаемое поведение является типичным для метода SQP на задачах с нерегулярными ограничениями.

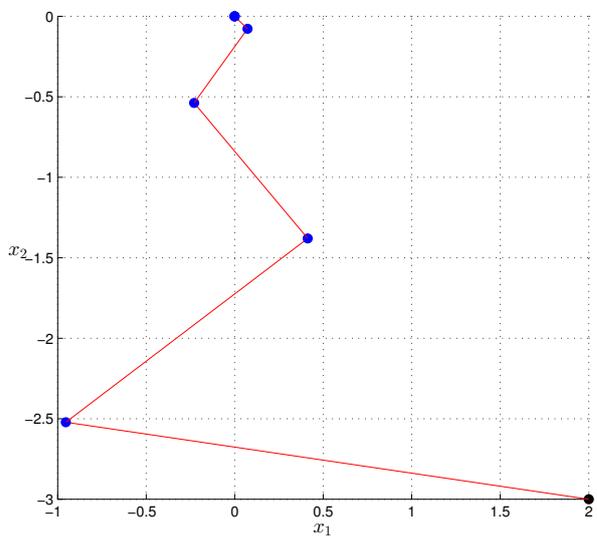
Метод sSQP демонстрирует другое поведение: проекция прямой траектории на $\ker h'(\bar{x})$ быстро сходится к нулю, однако ее проекция на $(\ker h'(\bar{x}))^\perp$ движется очень медленно. Кри-



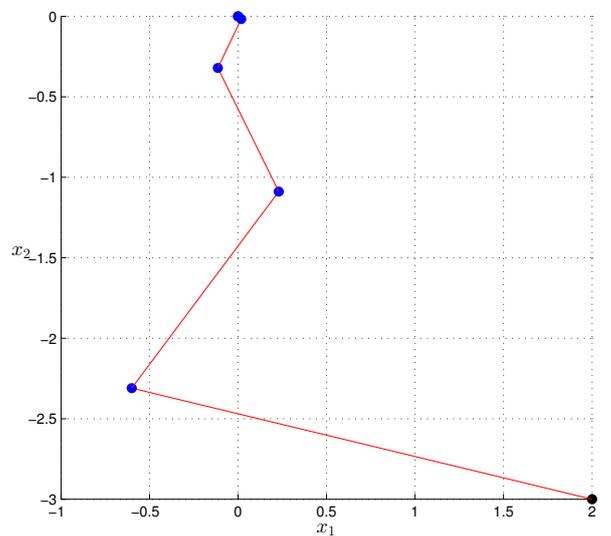
(а) Обычный SQP



(б) Стабилизированный SQP



(в) s-SQP с исчезающей стабилизацией



(г) s-SQP с неисчезающей стабилизацией

Рис. 1. Прямые траектории для примера 1; $x^0 = (2, -3)$, $\lambda^0 = (-10, 15)$.

терий остановки выполнен только после 30 итераций, несмотря на то, что предельный множитель является некритическим, и в конце имеет место сверхлинейная сходимость.

Плохое поведение SQP и sSQP можно проинтерпретировать как эффект «неправильной» двойственной стабилизации. Двойственные траектории метода SQP не стабилизируются вдоль подпространства вырожденности $\ker(h'(\bar{x}))^T$, и поэтому они сходятся к критическому множителю вдоль этого подпространства. С другой стороны, двойственные траектории sSQP стабилизируются не только вдоль подпространства вырожденности, но и вдоль его дополнения $(\ker(h'(\bar{x}))^T)^\perp = \text{im } h'(\bar{x})$, и эта «излишняя стабилизация» приводит к коротким шагам.

В то же время, метод s-sSQP корректно определяет ранг матрицы $h'(\bar{x})$ и использует значение $P(x^k, \lambda^k)$, близкое к \bar{P} . Поэтому двойственные траектории стабилизируются только вдоль подпространства вырожденности, что приводит к быстрой сходимости: критерий

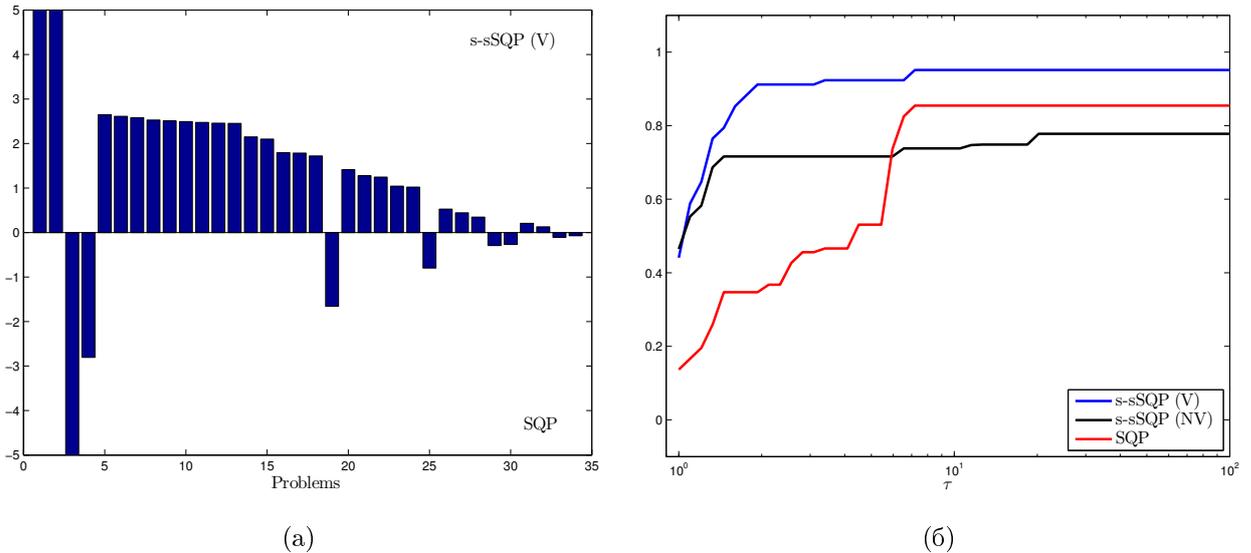


Рис. 2. Результаты для всех задач с ограничениями-равенствами из DEGEN.

остановки выполнен после 7 и 6 итераций в случае исчезающей и неисчезающей стабилизации соответственно.

Теперь сравним метод s-sSQP с глобализованным квазиньютоновским SQP на задачах с ограничениями-равенствами из DEGEN. Согласно полученному вычислительному опыту, метод s-sSQP в базовой форме (без каких-либо средств глобализации сходимости) обычно демонстрирует хорошие свойства глобальной сходимости. Процедура глобализации сходимости, реализованная в алгоритме 4.3, предоставляет гарантии отсутствия нестационарных предельных точек, однако она может несколько ухудшать эффективность и робастность метода.

Относительная эффективность алгоритмов представлена на рис. 2(a) в форме, описанной в приложении Б.3. Верхняя часть рис. 2(a) соответствует методу s-sSQP с исчезающей стабилизацией, а нижняя — квазиньютоновской версии SQP. Эти же результаты представлены на рис. 2(б) в форме «performance profiles» (см. описание в приложении А). На рис. 2(a) через (V) и (NV) обозначены методы с исчезающей и неисчезающей стабилизацией соответственно.

Рис. 2(a) показывает, что метод s-sSQP более эффективен для большинства задач: он по крайней мере в два раза эффективнее для более чем 60% задач. Только для трех задач из DEGEN квазиньютоновский SQP оказывается значительно более эффективен: 20310, 20401 и 2DD01. Однако в этих задачах решению не отвечает ни один множитель Лагранжа, и поэтому нет никаких оснований ожидать хорошего поведения от метода с двойственной стабилизацией. Тем не менее, метод s-sSQP демонстрирует хорошее поведение для многих других «плохих» задач из DEGEN, включая задачи с неизолрованными решениями, задачи, в которых все множители являются критическими, и даже некоторые задачи без множи-

телей Лагранжа. Отметим также, что метод s-sSQP с исчезающей стабилизацией также существенно эффективнее квазиньютоновского SQP, однако он несколько проигрывает по эффективности методу с исчезающей стабилизацией.

Результаты на рис. 2(б) весьма похожи: метод s-sSQP с исчезающей стабилизацией намного эффективнее и робастнее, чем квазиньютоновский SQP, а метод s-sSQP с исчезающей стабилизацией также намного более эффективен, но при этом менее робастен, чем квазиньютоновский SQP. В связи с этим, далее будем рассматривать только метод s-sSQP с исчезающей стабилизацией.

В эксперименте, результаты которого приведены выше, для каждой отдельной задачи метод s-sSQP верно идентифицировал ранг матрицы Якоби ограничений по крайней мере в 80% запусков. Средняя (по всем задачам) доля запусков, в которых ранг был верно идентифицирован, составила 95%. Неверная идентификация в большинстве случаев была связана со сходимостью к критическим множителям.

Далее сравним глобализованную версию метода s-sSQP (алгоритм 4.3) с квазиньютоновской версией SQP. Для того, чтобы компенсировать отрицательное влияние глобализации на эффективность и робастность s-sSQP, использовалась процедура немонотонного одномерного поиска (см. [59]). Идея состоит в замене выражения $\varphi_{c_1, c_2}(x^k, \lambda^k)$ в правой части (4.44) на

$$\max_{j=0, \dots, R-1} \varphi_{c_1, c_2}(x^{k-j}, \lambda^{k-j}),$$

если $k \geq R - 1$, и на

$$\max \left\{ M, \max_{j=0, \dots, k} \varphi_{c_1, c_2}(x^{k-j}, \lambda^{k-j}) \right\},$$

если $k \in \{0, 1, \dots, R - 2\}$, где R — целое, а $M > 0$ — некоторое большое число. В данном эксперименте оптимальными оказались значения $R = 8$ и $M = 10^{20}$. В этом случае полный шаг метода s-sSQP почти всегда принимается на первых 7 итерациях, что улучшает общее поведение, но в то же время использование одномерного поиска на более поздних итерациях сохраняет гарантии глобальной сходимости.

Результаты представлены на рис. 3. Сравнивая их с результатами на рис. 2, можно заключить, что глобализованная версия метода s-sSQP действительно несколько менее эффективна и робастна, чем базовая версия, но она по-прежнему существенно опережает квазиньютоновский SQP по эффективности. Подчеркнем, что для высокой эффективности крайне важна стабилизация вдоль нужного подпространства: результаты аналогичного эксперимента с обычным sSQP, приведенные в приложении Б.3, существенно хуже (см. рис. Б.6).

Наконец, приведем результаты сравнения алгоритмов на всех задачах из DEGEN, в которых ограничения-неравенства (если они присутствуют) переформулированы как равенства с использованием дополнительных переменных. Цель такой переформулировки состоит

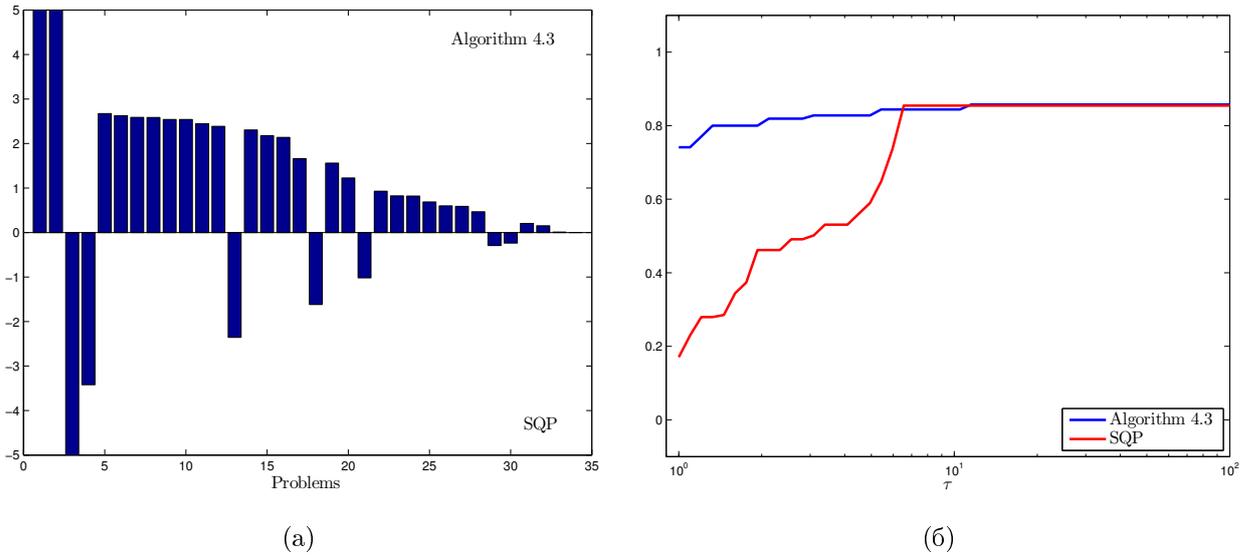


Рис. 3. Результаты для всех задач с ограничениями-равенствами из DEGEN.

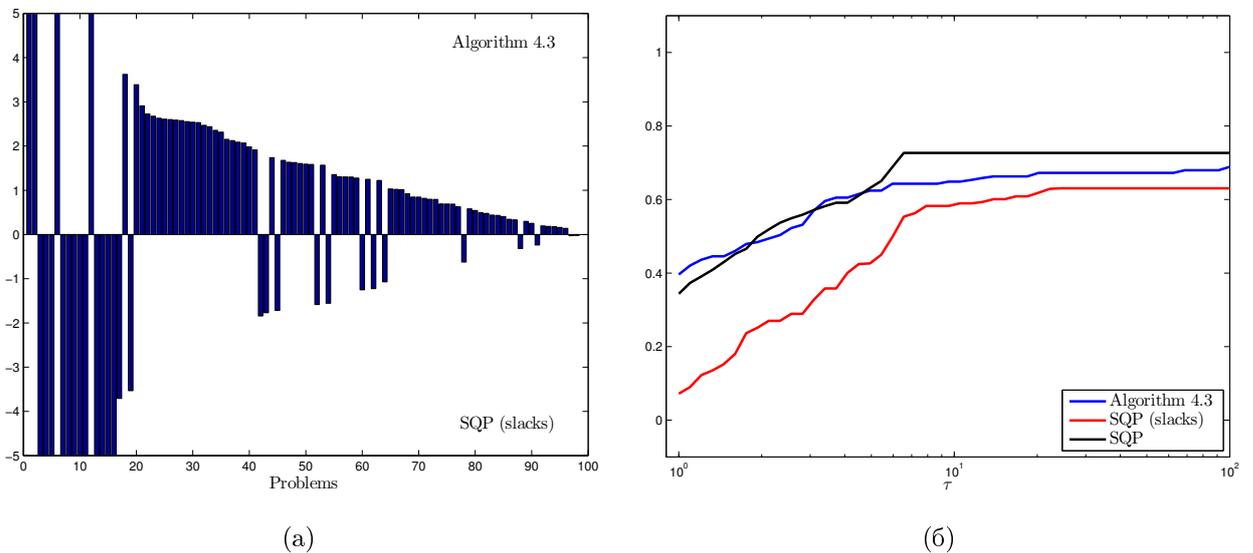


Рис. 4. Результаты для всех задач из DEGEN.

в получении большего множества нерегулярных задач с ограничениями-равенствами. Отметим, однако, что после переформулировки задачи имеют достаточно специальную структуру. Будем сравнивать алгоритм 4.3 (с использованием процедуры немонотонного одномерного поиска, описанной выше) с квазиньютоновской версией SQP, применяемой к тем же переформулировкам, а также к исходным задачам. Конечно, такое сравнение будет не совсем честным, поскольку введение дополнительных переменных часто негативно сказывается на свойствах задачи. В частности, результаты квазиньютоновского SQP существенно лучше, если его применять к исходной задаче.

Как показывает рис. 4, для большинства задач алгоритм 4.3 существенно более эффективен, а также более робастен, чем квазиньютоновский SQP, применяемый к переформулированным задачам. Более того, он даже несколько эффективнее, чем квазиньютоновский

SQP, применяемый к исходным задачам, но при этом менее робастен.

Полученные результаты свидетельствуют о том, что стабилизация вдоль нужного подпространства крайне важна для высокой эффективности методов с двойственной стабилизацией. Если стабилизация осуществляется вдоль меньшего подпространства (или не осуществляется вообще, как в случае SQP), двойственные траектории обычно сходятся к критическим множителям, и скорость сходимости оказывается лишь линейной. С другой стороны, стабилизация вдоль большего подпространства может приводить к длинным последовательностям коротких шагов метода. Метод s-SQP является в некотором смысле «комбинацией» обычного и стабилизированного SQP: он ведет себя как стабилизированный SQP на подпространстве вырожденности и как обычный SQP на дополнении к этому подпространству. Процедура идентификации этого подпространства, предложенная в разделе 4.2, демонстрирует высокую надежность, но при этом является очень простой и дешевой с вычислительной точки зрения (особенно по сравнению с существующими альтернативами, использующими сингулярное разложение).

Заключение

Данная работа вносит вклад в теорию численных методов решения задач оптимизации с нерегулярными ограничениями. Для таких задач исследованы причины низкой скорости сходимости существующих методов, а также разработан ряд конкурентоспособных практических алгоритмов их решения. В частности, в работе впервые предложен алгоритм, который на нерегулярных задачах существенно опережает по эффективности метод последовательного квадратичного программирования и при этом обладает теоретическими гарантиями глобальной сходимости.

Основные результаты диссертационной работы заключаются в следующем.

1. Проведено исследование эффекта притяжения двойственных траекторий к критическим множителям для метода Ньютона–Лагранжа. Для двух важных частных случаев получено полное обоснование данного эффекта.
2. Проведено исследование метода модифицированных функций Лагранжа в контексте нерегулярных задач. Получены теоретические результаты, свидетельствующие о хороших свойствах глобальной сходимости метода даже при отсутствии традиционной регулярности. В частности, для задач оптимизации с комплементарными ограничениями обоснована S -стационарность предельных точек в достаточно слабых предположениях. Также проведено исследование скорости сходимости метода и возможности ускорения его финальной фазы с помощью метода Ньютона–Лагранжа и его стабилизированной версии. Вычислительный эксперимент показал, что метод действительно обладает хорошим глобальным поведением, но несколько уступает по эффективности другим методам, даже при использовании ускорителей.
3. Предложено несколько различных подходов к глобализации сходимости стабилизированного метода последовательного квадратичного программирования, и для каждого из них получены теоретические результаты о глобальной сходимости и сверхлинейной скорости сходимости. Как показал вычислительный эксперимент, гибридный подход к глобализации сходимости и подход с использованием модифицированной функции Лагранжа не позволяет добиться высокой эффективности метода. Однако метод, гло-

бализованный с помощью точной гладкой штрафной функции, показывает высокую эффективность на некоторых классах задач.

4. 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.
5. Разработан простой и экономичный способ аппроксимации подпространства вырожденности нелинейного отображения в особом решении соответствующего уравнения. Данный способ основан на варианте метода Гаусса с выбором главного элемента. В частности, он намного эффективнее всех существующих аналогов с вычислительной точки зрения. Способ предназначен для реализации метода s-sSQP, но может быть использован в качестве ингредиента других методов, а также методов решения нерегулярных задач других классов.

Список литературы

1. Васильев Ф.П. Методы оптимизации. — М.: МЦНМО, 2011.
2. Воеводин В.В., Кузнецов Ю.А. Матрицы и вычисления. — М.: Наука, 1984.
3. Ерина М.Ю., Измаилов А.Ф. Метод Гаусса-Ньютона для отыскания особых решений систем нелинейных уравнений // Журнал вычислительной математики и математической физики. — 2007. — Т. 47, № 5. — С. 784–795.
4. Бертсекас Д. Условная оптимизация и методы множителей Лагранжа. — М.: Радио и связь, 1987.
5. Брежнева О.А., Измаилов А.Ф. О построении определяющих систем для отыскания особых решений нелинейных уравнений // Журнал вычислительной математики и математической физики. — 2002. — Т. 42, № 1. — С. 10–22.
6. Брежнева О.А., Измаилов А.Ф., Третьяков А.А., Хмура А.А. Один подход к поиску особых решений системы нелинейных уравнений общего вида // Журнал вычислительной математики и математической физики. — 2000. — Т. 40, № 3. — С. 365–377.
7. Измаилов А.Ф. Об одном классе определяющих систем для особых решений нелинейных уравнений // Вопросы моделирования и анализа в задачах принятия решений. — М.: ВЦ РАН, 2002. — С. 18–57.
8. Измаилов А.Ф. Об аналитической и вычислительной устойчивости критических множителей Лагранжа // Журнал вычислительной математики и математической физики. — 2005. — Т. 45, № 6. — С. 966–982.
9. Измаилов А.Ф. О предельных свойствах двойственных траекторий метода множителей Лагранжа // Журнал вычислительной математики и математической физики. — 2011. — Т. 51, № 1. — С. 3–23.
10. 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.

11. Измаилов А.Ф., Солодов М.В. Численные методы оптимизации. 2-е изд., перераб. и доп. — М.: Физматлит, 2008.
12. 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 имени М.В. Ломоносова, 14–23 ноября 2011 г.: Тезисы докладов. — М.: МАКС Пресс, 2011. — С. 23–24.
13. 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.
14. Измаилов А.Ф., Третьяков А.А. 2-регулярные решения нелинейных задач. Теория и численные методы. — М.: Физматлит, 1999.
15. 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.
16. 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.
17. 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.
18. 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.
19. 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.
20. 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.

21. 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.
22. Ильюшин А.А. Пластичность. Ч. 1. Упруго-пластические деформации. — Л.: Гостехтеориздат, 1948.
23. Качанов Л.М. Основы теории пластичности. — М.: Наука, 1969.
24. 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.
25. 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.
26. Усков Е.И. Численное сравнение оптимизационных алгоритмов // Теоретические и прикладные задачи нелинейного анализа. — М.: ВЦ РАН, 2012. — С. 118–131.
27. Усков Е.И. О притяжении метода Ньютона к критическим множителям Лагранжа // Журнал вычислительной математики и математической физики. — 2013. — Т. 53, № 8. — С. 1272–1286.
28. Achtziger W., Hoheisel T., Kanzow C. A smoothing-regularization approach to mathematical programs with vanishing constraints // Computational Optimization and Applications. — 2013. — Vol. 55, no. 3. — P. 733–767.
29. Achtziger W., Kanzow C. Mathematical programs with vanishing constraints: optimality conditions and constraint qualifications // Mathematical Programming. — 2008. — Vol. 114, no. 1. — P. 69–99.
30. ALGENCAN. — <http://www.ime.usp.br/~egbirgin/tango/>.
31. Andreani R., Birgin E.G., Martínez J.M., Schuverdt M.L. On augmented Lagrangian methods with general lower-level constraints // SIAM Journal on Optimization. — 2007. — Vol. 18, no. 4. — P. 1286–1309.

32. Andreani R., Birgin E.G., Martínez J.M., Schuverdt M.L. Second-order negative-curvature methods for box-constrained and general constrained optimization // *Computational Optimization and Applications*. — 2010. — Vol. 45, no. 2. — P. 209–236.
33. Andreani R., Haeser G., Schuverdt M.L., Silva P.J.S. A relaxed constant positive linear dependence constraint qualification and applications // *Mathematical Programming*. — 2012. — Vol. 135, no. 1-2. — P. 255–273.
34. Andreani R., Martínez J.M. On the solution of mathematical programming problems with equilibrium constraints // *Mathematical Methods of Operations Research*. — 2001. — Vol. 54, no. 3. — P. 345–358.
35. Bertsekas D.P. Enlarging the region of convergence of Newton's method for constrained optimization // *Journal of Optimization Theory and Applications*. — 1982. — Vol. 36, no. 2. — P. 221–252.
36. Birgin E.G., Fernández D., Martínez J.M. The boundedness of penalty parameters in an augmented Lagrangian method with constrained subproblems // *Optimization Methods and Software*. — 2012. — Vol. 27, no. 6. — P. 1001–1024.
37. Birgin E.G., Martínez J.M. Improving ultimate convergence of an augmented Lagrangian method // *Optimization Methods and Software*. — 2008. — Vol. 23, no. 2. — P. 177–195.
38. Birgin E.G., Martínez J.M. Augmented Lagrangian method with nonmonotone penalty parameters for constrained optimization // *Computational Optimization and Applications*. — 2012. — Vol. 51, no. 3. — P. 941–965.
39. Bonnans J.F., Gilbert J.Ch., Lemaréchal C., Sagastizábal C.A. *Numerical optimization: theoretical and practical aspects*. — Springer-Verlag, 2006.
40. Byrd R.H., Nocedal J., Waltz R.A. *KNITRO: an integrated package for nonlinear optimization* // *Large-Scale Nonlinear Optimization*. — Springer, 2006. — P. 35–59.
41. Conn A.R., Gould N., Sartenaer A., Toint Ph.L. Convergence properties of an augmented Lagrangian algorithm for optimization with a combination of general equality and linear constraints // *SIAM Journal on Optimization*. — 1996. — Vol. 6, no. 3. — P. 674–703.
42. DEGEN. — http://w3.impa.br/~optim/DEGEN_collection.zip.
43. Dempe S. *Foundations of bilevel programming*. — Springer, 2002.
44. Di Pillo G., Grippo L. A new class of augmented Lagrangians in nonlinear programming // *SIAM Journal on Control and Optimization*. — 1979. — Vol. 17, no. 5. — P. 618–628.

45. Dolan E.D., More J.J. Benchmarking optimization software with performance profiles // *Mathematical Programming*. — 2002. — Vol. 91, no. 2. — P. 201–213.
46. Fernández D. A quasi-Newton strategy for the sSQP method for variational inequality and optimization problems // *Mathematical Programming*. — 2013. — Vol. 137, no. 1-2. — P. 199–223.
47. Fernández D., Pilotta E.A., Torres G.A. An inexact restoration strategy for the globalization of the sSQP method // *Computational Optimization and Applications*. — 2013. — Vol. 54, no. 3. — P. 595–617.
48. Fernández D., Solodov M.V. Stabilized sequential quadratic programming for optimization and a stabilized Newton-type method for variational problems // *Mathematical Programming*. — 2010. — Vol. 125, no. 1. — P. 47–73.
49. Ferris M.C., Tin-Loi F. On the solution of a minimum weight elastoplastic problem involving displacement and complementarity constraints // *Computer Methods in Applied Mechanics and Engineering*. — 1999. — Vol. 174, no. 1. — P. 108–120.
50. Fischer A. Local behavior of an iterative framework for generalized equations with nonisolated solutions // *Mathematical Programming*. — 2002. — Vol. 94, no. 1. — P. 91–124.
51. Fletcher R., Leyffer S. User manual for filterSQP // *University of Dundee Numerical Analysis Report NA-181*. — 1998.
52. Fletcher R., Leyffer S. Solving mathematical programs with complementarity constraints as nonlinear programs // *Optimization Methods and Software*. — 2004. — Vol. 19, no. 1. — P. 15–40.
53. Fletcher R., Leyffer S., Ralph D., Scholtes S. Local convergence of SQP methods for mathematical programs with equilibrium constraints // *SIAM Journal on Optimization*. — 2006. — Vol. 17, no. 1. — P. 259–286.
54. Gill Ph.E., Murray W., Saunders M.A. SNOPT: an SQP algorithm for large-scale constrained optimization // *SIAM Journal on Optimization*. — 2002. — Vol. 12, no. 4. — P. 979–1006.
55. Gill Ph.E., Robinson D.P. A primal-dual augmented Lagrangian // *Computational Optimization and Applications*. — 2012. — Vol. 51, no. 1. — P. 1–25.
56. Gill Ph.E., Robinson D.P. A globally convergent stabilized SQP method // *SIAM Journal on Optimization*. — 2013. — Vol. 23, no. 4. — P. 1983–2010.
57. Glad S.T. Properties of updating methods for the multipliers in augmented Lagrangians // *Journal of Optimization Theory and Applications*. — 1979. — Vol. 28, no. 2. — P. 135–156.

58. Golubitsky M., Schaeffer D.G. Singularities and groups in bifurcation theory. Volume I. — Springer–Verlag, 1985.
59. Grippo L., Lampariello F., Lucidi S. A nonmonotone line search technique for Newton's method // *SIAM Journal on Numerical Analysis*. — 1986. — Vol. 23, no. 4. — P. 707–716.
60. Hestenes M.R. Multiplier and gradient methods // *Journal of Optimization Theory and Applications*. — 1969. — Vol. 4, no. 5. — P. 303–320.
61. Hock W., Schittkowski K. Test examples for nonlinear programming codes // *Journal of Optimization Theory and Applications*. — 1980. — Vol. 30, no. 1. — P. 127–129.
62. The HSL Mathematical Software Library. — <http://www.hsl.rl.ac.uk>.
63. IPOPT. — <https://projects.coin-or.org/Ipopt>.
64. Izmailov A.F., Kurennoy A.S. Abstract Newtonian frameworks and their applications // *SIAM Journal on Optimization*. — 2013. — Vol. 23, no. 4. — P. 2369–2396.
65. Izmailov A.F., Kurennoy A.S., Solodov M.V. Local convergence of the method of multipliers for variational and optimization problems under the noncriticality assumption // *Computational Optimization and Applications*. — 2014. — DOI: 10.1007/s10589-014-9658-8.
66. Izmailov A.F., Solodov M.V. Newton-type methods for optimization problems without constraint qualifications // *SIAM Journal on Optimization*. — 2004. — Vol. 15, no. 1. — P. 210–228.
67. Izmailov A.F., Solodov M.V. Examples of dual behaviour of Newton-type methods on optimization problems with degenerate constraints // *Computational Optimization and Applications*. — 2009. — Vol. 42, no. 2. — P. 231–264.
68. Izmailov A.F., Solodov M.V. On attraction of Newton-type iterates to multipliers violating second-order sufficiency conditions // *Mathematical Programming*. — 2009. — Vol. 117, no. 1-2. — P. 271–304.
69. Izmailov A.F., Solodov M.V. A truncated SQP method based on inexact interior-point solutions of subproblems // *SIAM Journal on Optimization*. — 2010. — Vol. 20, no. 5. — P. 2584–2613.
70. Izmailov A.F., Solodov M.V. On attraction of linearly constrained Lagrangian methods and of stabilized and quasi-Newton SQP methods to critical multipliers // *Mathematical Programming*. — 2011. — Vol. 126, no. 2. — P. 231–257.
71. Izmailov A.F., Solodov M.V. Stabilized SQP revisited // *Mathematical Programming*. — 2012. — Vol. 133, no. 1-2. — P. 93–120.

72. Izmailov A.F., Solodov M.V. Newtonian-type methods for optimization and variational problems. — Springer, 2014.
73. 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.
74. Izmailov A.F., Solodov M.V., Uskov E.I. Combining stabilized SQP with the augmented Lagrangian algorithm // *IMPA Preprint A754*. — 2014.
75. Izmailov A.F., Solodov M.V., Uskov E.I. Globalizing stabilized SQP by smooth primal-dual exact penalty function // *IMPA Preprint A752*. — 2014.
76. 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.
77. 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.
78. KNITRO. — <http://www.ziena.com/knitro.htm>.
79. LANCELOT. — <http://www.cse.scitech.ac.uk/nag/lancelot/lancelot.shtml>.
80. Leyffer S., López-Calva G., Nocedal J. Interior methods for mathematical programs with complementarity constraints // *SIAM Journal on Optimization*. — 2006. — Vol. 17, no. 1. — P. 52–77.
81. Li D.-H., Qi L. A stabilized SQP method via linear equations // *Applied mathematics technical report AMR00/5*, University of New South Wales. — 2000.
82. Luo Z.-Q., Pang J.-S., Ralph D. *Mathematical programs with equilibrium constraints*. — Cambridge University Press, 1996.
83. MacMPEC. — <http://wiki.mcs.anl.gov/leyffer/index.php/MacMPEC>.
84. Morales J.L. A numerical study of limited memory BFGS methods // *Applied Mathematics Letters*. — 2002. — Vol. 15, no. 4. — P. 481–487.
85. Mostafa E.M.E., Vicente L.N., Wright S.J. Numerical behavior of a stabilized SQP method for degenerate NLP problems // *Global Optimization and Constraint Satisfaction*. — Springer, 2003. — P. 123–141.

86. Murtagh B.A., Saunders M.A. MINOS 5.0 user's guide // Technical Report SOL 83.20, Stanford University. — 1983.
87. NEOS. — <http://neos-server.org/neos/>.
88. Nocedal J., Wright S.J. Numerical optimization. Second edition. — New York: Springer, 2006.
89. Outrata J., Kocvara M., Zowe J. Nonsmooth approach to optimization problems with equilibrium constraints: theory, applications and numerical results. — Springer, 1998.
90. Powell M.J.D. A method for nonlinear constraints in minimization problems // Optimization / Ed. by R. Fletcher. — Academic Press, 1969. — P. 283–298.
91. Raghunathan A.U., Biegler L.T. An interior point method for mathematical programs with complementarity constraints (MPCCs) // SIAM Journal on Optimization. — 2005. — Vol. 15, no. 3. — P. 720–750.
92. Wächter A., Biegler L.T. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming // Mathematical Programming. — 2006. — Vol. 106, no. 1. — P. 25–57.
93. Wright S.J. Superlinear convergence of a stabilized SQP method to a degenerate solution // Computational Optimization and Applications. — 1998. — Vol. 11, no. 3. — P. 253–275.
94. Wright S.J. Constraint identification and algorithm stabilization for degenerate nonlinear programs // Mathematical Programming. — 2003. — Vol. 95, no. 1. — P. 137–160.