

Н.З.БОР

МЕТОДЫ  
МИНИМИЗАЦИИ  
НЕДИФФЕРЕНЦИРУЕМЫХ  
ФУНКЦИЙ  
И ИХ ПРИЛОЖЕНИЯ

АКАДЕМИЯ НАУК УКРАИНСКОЙ ССР  
ОРДЕНА ЛЕНИНА ИНСТИТУТ КИБЕРНЕТИКИ

Н.З.ШОР

МЕТОДЫ  
МИНИМИЗАЦИИ  
НЕДИФФЕРЕНЦИРУЕМЫХ  
ФУНКЦИЙ  
И ИХ ПРИЛОЖЕНИЯ

КИЕВ  
«НАУКОВА ДУМКА»  
1979

УДК 007:512.8

**Методы минимизации недифференцируемых функций и их приложения/ Шор Н. З.**— Киев: Наук. думка, 1979. — с. 200.

Рассмотрены обобщенные градиентные методы минимизации негладких функций, в разработке и обосновании которых принимал участие автор. Исследованы основные модификации метода обобщенного градиентного спуска, изложены алгоритмы градиентного типа с растяжением пространства в направлении градиента и разности двух последовательных градиентов, которые обладают ускоренной сходимостью, а также применения методов минимизации недифференцируемых функций при решении разнообразных задач оптимального планирования и проектирования.

Книга рассчитана на специалистов, занятых в области математического программирования и его приложений.

Список лит.: с. 195—200 (106 назв.).

Ответственный редактор *В. С. Михалевиц*

Рецензенты *В. С. Королюк, Б. Н. Пшеничный*

Редакция физико-математической литературы

Ш  $\frac{20205-001}{M221(04)-79}$  157-79 1502000000

© Издательство «Наукова думка», 1979

# ОГЛАВЛЕНИЕ

---

Предисловие . . . . .	5
Введение . . . . .	7

## Глава 1

<b>Специальные классы недифференцируемых функций и обобщения понятия градиента . . . . .</b>	<b>13</b>
§ 1. Необходимость введения специальных классов негладких функций . . . . .	13
§ 2. Выпуклые функции. Понятие субградиента . . . . .	14
§ 3. Некоторые способы вычисления субградиента . . . . .	21
§ 4. Почти дифференцируемые функции . . . . .	26
§ 5. Полугладкие и полувывуклые функции . . . . .	30

## Глава 2

<b>Обобщенный градиентный спуск . . . . .</b>	<b>33</b>
§ 1. Проблема регулировки шага в методе обобщенного градиентного спуска . . . . .	33
§ 2. Основные теоремы о сходимости обобщенного градиентного спуска . . . . .	36
§ 3. Случай сходимости ОГС со скоростью геометрической прогрессии . . . . .	44
§ 4. Обобщенный градиентный спуск и фейеровские приближения . . . . .	52
§ 5. $\varepsilon$ -субградиентные методы . . . . .	58
§ 6. Обобщение метода ОГС на класс невыпуклых функций. Стохастические варианты ОГС и его устойчивость . . . . .	61

## Глава 3

<b>Методы градиентного типа с растяжением пространства . . . . .</b>	<b>66</b>
§ 1. Эвристика, лежащая в основе методов с растяжением пространства . . . . .	66
§ 2. Операторы растяжения пространства . . . . .	68
§ 3. Обобщенный градиентный спуск с растяжением пространства в направлении градиента . . . . .	69
§ 4. Вопросы сходимости алгоритмов ОГСРП . . . . .	72
§ 5. Применение метода ОГСРП к решению системы нелинейных уравнений . . . . .	83

§ 6. Метод минимизации, использующий операцию растяжения пространства в направлении разности двух последовательных почти-градиентов . . . . .	90
§ 7. Обоснование сходимости одного из вариантов $r$ -алгоритма . . . . .	102
§ 8. Связь между алгоритмами ОГСРП и алгоритмами последовательных отсечений . . . . .	113
§ 9. Модифицированные вычислительные схемы обобщенных градиентных методов с растяжением пространства . . . . .	120

#### Глава 4

<b>Приложения методов негладкой оптимизации к решению задач математического программирования . . . . .</b>	<b>122</b>
§ 1. Использование обобщенных градиентных методов в схемах декомпозиции . . . . .	122
§ 2. Итеративный метод решения задач линейного программирования специального вида . . . . .	132
§ 3. Решение задач распределительного типа методом обобщенного градиентного спуска . . . . .	137
§ 4. Опыт решения задач производственно-транспортного характера с использованием обобщенных градиентных методов с растяжением пространства . . . . .	143
§ 5. Использование $r$ -алгоритмов для решения нелинейных минимаксных задач . . . . .	172
§ 6. Применение методов минимизации негладких функций для решения задач интерпретации гравиметрических наблюдений . . . . .	182
§ 7. Другие области применения обобщенных градиентных методов . . . . .	187
Заключение . . . . .	192
Список литературы . . . . .	194

## ПРЕДИСЛОВИЕ

---

В настоящее время большое внимание уделяется созданию автоматизированных систем планирования, проектирования и управления в различных областях народного хозяйства. На первый план выдвигаются вопросы качества принимаемых решений, в связи с чем возрастает роль методов и алгоритмов решения оптимизационных задач в математическом обеспечении автоматизированных систем различного уровня и назначения. Все большую роль приобретают методы минимизации функций с разрывным градиентом. Многие специалисты в области численных методов математического программирования склоняются к мысли, что прогресс в области разработки алгоритмов минимизации негладких функций даст ключ к построению эффективных способов решения задач большой размерности.

Настоящая монография является в определенной мере итогом пятнадцатилетней работы автора в области развития обобщенных градиентных методов минимизации негладких функций. Эта работа была начата в отделе экономической кибернетики Института кибернетики АН УССР под руководством академика АН УССР В. С. Михалевича в связи с необходимостью решения важных практических задач оптимального планирования и проектирования.

В первой главе описываются основные классы негладких почти всюду дифференцируемых функций и анализируются различные способы определения обобщенных градиентных множеств.

Во второй главе подробно исследованы различные модификации метода обобщенного градиентного спуска, указана взаимосвязь этих методов с методами фейеровских приближений, кратко изложены основные идеи  $\epsilon$ -субградиентных методов.

В третьей главе описаны алгоритмы градиентного типа с растяжением пространства в направлении градиента (ОГСРП-алгоритмы) и в направлении разности двух последовательных градиентов ( $r$ -алгоритмы). Большое внимание уделено обоснованию сходимости и скорости сходимости этих алгоритмов, рассмотрены их предельные варианты.

Отдельно проанализированы возможности алгоритмов ОГСРП при решении систем нелинейных уравнений. Обсуждены взаимосвязи между методами отсечений и методами с растяжением пространства.

Четвертая глава посвящена приложениям обобщенных градиентных методов к реализации схем декомпозиции, для решения ряда конкретных задач текущего и перспективного планирования, для решения минимаксных задач, задач интерпретации результатов гравиметрических наблюдений и др.

Автор выражает искреннюю благодарность В. С. Михалевичу, Л. А. Галустовой, Г. И. Горбач, В. И. Гершовичу, Н. Г. Журбенко, Т. В. Марчук, А. И. Момот, В. А. Трубину, Л. П. Шабашовой, Г. Н. Юну, внесшим значительный вклад в разработку и практические применения методов негладкой оптимизации, которые нашли отражение в монографии, а также В. Н. Богданюк, Г. А. Вербе, Т. Г. Трушко, Т. И. Яценко за помощь в оформлении рукописи.

Негладкие функции, т. е. функции с разрывным градиентом, широко используются в теории и приложениях математического программирования. Что касается теории, то в настоящее время принято рассматривать основные факты выпуклого анализа для всего класса выпуклых функций, в том числе негладких, и использование субдифференциалов стало в теории таким же обычным делом, как и градиентов, особенно после обобщающих работ Рокафеллара [44], Б. Н. Пшеничного [43] и др.

Имеется несколько основных источников, порождающих задачи негладкой оптимизации. Во-первых, это задачи математического программирования высокой размерности с блочной структурой и сравнительно небольшим числом связей между блоками. Использование схем декомпозиции для решения таких задач приводит к задачам минимизации (максимизации), как правило, негладких функций от связывающих переменных или от множителей Лагранжа (двойственных оценок), соответствующих связывающим ограничениям.

Во-вторых, это задачи минимизации функции максимума, характерные для моделей игрового характера, «многокритериальных» моделей оптимального планирования и исследования операций. К такого рода задачам сводятся задачи решения систем уравнений и неравенств, определения коэффициентов нелинейной регрессии, когда в качестве критерия используется чебышевский критерий минимизации максимума невязки (модуля невязки).

В-третьих, это задачи нелинейного программирования, для решения которых используется метод негладких штрафных функций. Негладкие штрафные функции определенного

вида обладают несомненным преимуществом по сравнению с обычно применяемыми гладкими функциями штрафа: при использовании негладких штрафных функций, как правило, нет необходимости устремлять штрафные коэффициенты к  $+\infty$  и можно ограничиться решением задачи при определенных значениях штрафных коэффициентов [18, 80].

В-четвертых, это задачи оптимального управления с непрерывным и дискретным временем. Использование принципа максимума или дискретного принципа максимума во многих случаях приводит к задачам минимизации функций с разрывным градиентом. Эти задачи можно также рассматривать как специальные задачи нелинейного программирования, для решения которых можно использовать схемы декомпозиции или метод негладких штрафных функций, что опять приводит к проблемам негладкой оптимизации.

В-пятых, это задачи дискретного программирования или задачи смешанного дискретно-непрерывного типа. Многие задачи такого рода достаточно успешно могут решаться с использованием метода ветвей и границ с получением оценок путем решения двойственной задачи. Двойственная задача обычно оказывается задачей минимизации выпуклой кусочно-линейной функции с огромным числом «кусков» при простых ограничениях, т. е. задачей негладкой оптимизации.

И наконец, функции с разрывным градиентом могут непосредственно входить в модель задачи оптимального планирования, проектирования или исследования операций как результат кусочно-гладкой аппроксимации технико-экономических характеристик реальных объектов.

Нужно также отметить, что с прикладной точки зрения нет резкой границы между негладкими и гладкими функциями. С позиций прикладной математики и вычислительной практики функция с очень быстро меняющимся градиентом «близка» по своим свойствам к негладкой функции. Поэтому следует ожидать, что вычислительные методы, специально разработанные для решения задач негладкой оптимизации, могут оказаться эффективными и для оптимизации «плохих» гладких функций (например, функций овражного типа).

Таким образом, сфера применения методов негладкой оптимизации огромна. Создание эффективных методов негладкой оптимизации является ключом к решению многих

вычислительных проблем математического программирования, особенно для задач большой размерности.

Вычислительные методы негладкой оптимизации развивались в нескольких направлениях: а) исследования, связанные с решением частных классов задач минимизации негладких функций, специальная структура которых задается в явном виде; б) исследования по разработке алгоритмов решения достаточно общих классов задач, не требующих априорного задания конкретной структуры оптимизируемой функции; представляется лишь возможность вычисления функции и ее градиентов (или их аналогов в случае недифференцируемости) в произвольной точке.

К первому типу относятся многочисленные работы по созданию алгоритмов минимизации функции максимума [14]. В основном разработанные алгоритмы решения минимаксных задач примыкают к алгоритмам метода возможных направлений, являются монотонными и требуют для своей реализации решения на каждом шаге вспомогательных задач линейного или квадратичного программирования.

Ряд работ посвящен минимизации кусочно-линейных выпуклых функций [13]. Здесь чаще всего используется аппарат линейного программирования.

Для решения ряда частных задач негладкой оптимизации, в которых поверхности разрыва градиентов задаются в явном виде (например, функции от модуля аргументов), разработаны специальные способы сглаживания, позволяющие применить для приближенного решения таких задач методы минимизации гладких функций [81].

Для решения систем линейных или нелинейных выпуклых неравенств применены методы фейеровских приближений [17, 19], которые являются развитием идей релаксационных алгоритмов Агмона — Мотцкина [78]. Что касается общих методов негладкой оптимизации, то здесь, если исключить такие универсальные средства, как методы систематического перебора или случайного поиска, можно выделить два основных класса, причем в обоих используется вычисление обобщенных градиентов.

1. Метод секущих гиперплоскостей для решения задач выпуклого программирования. Наиболее известны два алгоритма:

а) метод секущих Келли [97], основанный на кусочно-линейной аппроксимации графика выпуклой функции с

использованием опорных гиперплоскостей; на каждом шаге этого метода приходится решать задачу линейного программирования, причем от шага к шагу число ограничений за- да и может возрастать;

б) метод центрированных сечений А. Ю. Левина, осно- ванный на проведении опорных гиперплоскостей к поверхно- сти уровня выпуклой функции, проходящих через центр тя- жести области локализации оптимума [31]. Последовательные процедуры отсекаения в методе центрированных сечений ока- зываются весьма эффективными, однако сложность вычис- ления центра тяжести в многомерном случае сводит на нет практическую ценность метода (по поводу этого метода см. также [76, 105]). Аналогичные процедуры, в которых вме- сто центра тяжести рассматриваются другие характерные точки выпуклых тел, например чебышевский центр, исследо- ваны М. Е. Примаком [41].

2. Обобщенные градиентные методы, которые являются основным предметом исследования в данной монографии.

Для минимизации гладких функций широко применяются различные модификации градиентных процессов, и это естественно, ведь направление антиградиента в данной точке локально является направлением наискорейшего спус- ка. (Современные методы градиентного типа минимизации гладких функций изложены в работах [32, 42].) Регулиров- ка шага в большинстве алгоритмов этого типа основана на том, чтобы обеспечить монотонное и в достаточной степени «существенное» уменьшение значения функции на каждом шаге.

Попытка обобщения этих методов на функции с разрыв- ным градиентом наталкивается по крайней мере на две труд- ности: 1) доопределение аналога градиента в тех точках, где он не существует; 2) построение новых способов выбора направления спуска и регулировки шага, так как для не- гладких функций прямое перенесение способов, развитых для минимизации гладких функций, не приводит к успеху.

Первая трудность преодолевается сравнительно просто. Дело в том, что в большинстве приложений мы можем ог- раничиться классом почти везде дифференцируемых функ- ций, для которых доопределение градиента в тех точках, где он не существует, может производиться путем предель- ного перехода. Правда, этот переход может осуществляться неоднозначно, и тогда вместо обобщенных градиентов по- лучаем градиентные множества. Для удобства применения

аппарата выпуклого анализа эти множества расширяются путем применения операции выпуклого замыкания к элементам градиентных множеств. Таким образом, получаются обобщенные градиентные множества, частным случаем которых для выпуклых функций являются субградиентные множества. В алгоритмах минимизации предполагается использование произвольных элементов обобщенных градиентных множеств. Как будет показано ниже, вычисление обобщенных градиентов в большинстве конкретных приложений ничуть не сложнее, чем вычисление градиентов, тем более что при рассмотрении функций, почти везде дифференцируемых, мера тех точек, где градиент не существует, равна нулю.

Выбрать целесообразное направление движения и величину шага при решении задач негладкой оптимизации значительно труднее. Даже в случае выпуклой функции произвольный антисубградиент не всегда задает направление уменьшения значения функции. Правда, при достаточно малом шаге уменьшается расстояние до точки минимума, однако в процессе счета у нас нет возможности вычислять это расстояние, и поэтому критерий уменьшения расстояния до точки минимума нельзя использовать при регулировке шага. Нужно было найти другой способ регулировки шага. Простейший обобщенный градиентный метод состоит в движении на каждом шаге в направлении, обратном направлению обобщенного градиента. Этот метод под названием обобщенного градиентного спуска предложен автором в 1961 г. в связи с необходимостью разработки эффективного алгоритма решения транспортных задач большой размерности для задач текущего планирования, решаемых в Институте кибернетики АН УССР совместно с Госпланом УССР. Впервые метод обобщенного градиентного спуска (ОГС) для минимизации кусочно-линейных выпуклых функций использовался при решении транспортных и транспортно-производственных задач [54]. Затем метод ОГС был распространен на класс произвольных выпуклых функций [55] и на задачи выпуклого программирования в гильбертовом пространстве [38]. Широкое развитие получили стохастические аналоги ОГС [23]. Алгоритмы решения задач математического программирования, построенные на основе ОГС, отличаются простотой и, что особенно важно для задач большой размерности, экономным использованием оперативной памяти ЭВМ.

К ограничениям метода ОГС относятся его довольно медленная сходимость, сложность контроля точности решения и то, что он применим только к классу выпуклых функций. Тем не менее к 1970 г. был накоплен серьезный опыт успешного решения большого числа важных практических задач большой размерности с использованием ОГС.

В 1969—1970 гг. автором были предложены ускоренные варианты обобщенных градиентных методов, основанные на использовании операции растяжения пространства в направлении градиента и разности двух последовательных градиентов [65]. Идея этих методов существенно отлична от той, которая используется для ускорения сходимости в случае гладких функций — идеи квадратичной аппроксимации функции в окрестности минимума, в той или иной мере определяющей формализм как методов сопряженных градиентов, так и квазиньютоновских методов [96]. В то же время предельные варианты методов с растяжением пространства при определенных условиях регулярности и гладкости обладают свойством квадратичной скорости сходимости. Таким образом, предложенные алгоритмы обладают высокой эффективностью и применительно к гладким задачам минимизации. В дальнейшем алгоритмы с растяжением пространства были обобщены на задачи нахождения локальных минимумов невыпуклых негладких функций [70].

В США и Западной Европе градиентными методами минимизации негладких функций всерьез начали заниматься примерно с 1973 г. сначала в связи с приложениями в области дискретного программирования [94], а затем в целом для решения задач большой размерности [95]. Результаты работ в этом направлении на Западе достаточно полно представлены в сборнике [100]. Особенно интенсивно развивается направление так называемой  $\varepsilon$ -субградиентной оптимизации, по идее близкое, с одной стороны, к алгоритмам В. Ф. Демьянова решения минимаксных задач, а с другой, особенно в формальном отношении, — к алгоритмам метода сопряженных градиентов (или «давидоновского» типа).

И наконец, в последнее время обнаружилось очень интересные связи между алгоритмами последовательных отсечений и алгоритмами с растяжением пространства [76, 74].

Таким образом, область обобщенных градиентных методов негладкой оптимизации не представляет нечто окончательно сформировавшееся и застывшее, а, наоборот, быстро развивается.

## СПЕЦИАЛЬНЫЕ КЛАССЫ НЕДИФФЕРЕНЦИРУЕМЫХ ФУНКЦИЙ И ОБОБЩЕНИЯ ПОНЯТИЯ ГРАДИЕНТА

### § 1. Необходимость введения специальных классов негладких функций

При решении широкого круга задач математического программирования приходится выходить за рамки класса непрерывно дифференцируемых функций. При этом мы, как правило, имеем дело с непрерывными почти везде дифференцируемыми функциями, градиент которых терпит разрыв на множестве меры нуль. Хотя в принципе возможно появление функций и с большей патологией (например, функций, которые нигде не дифференцируемы на множестве положительной меры), практически такие функции не встречаются, да и построить какой-либо нетривиальный алгоритм оптимизации таких функций не представляется возможным.

Рассмотрим такие специальные классы недифференцируемых функций, которые, с одной стороны, являются не слишком широкими обобщениями класса непрерывно дифференцируемых функций в том смысле, что для функций этих классов можно достаточно разумно определить обобщенное понятие градиента и использовать его для построения процессов градиентного типа нахождения экстремума. С другой стороны, введенные нами специальные классы функций должны охватывать функции, встречающиеся в большинстве практических приложений, где речь идет о непрерывных функциях, определенных на  $n$ -мерном евклидовом пространстве  $E_n$ .

Наиболее подробно изучены классы выпуклых и вогнутых функций, представляющие исключительный интерес с точки зрения теории и приложений математического программирования. В последние годы предложены различные расширения класса выпуклых функций [35, 43, 101]. В математическом анализе подробно рассмотрены различные классы кусочно-гладких функций. Введенный автором класс

почти дифференцируемых функций можно рассматривать как расширение класса непрерывных и кусочно-непрерывно дифференцируемых функций, удовлетворяющих локальному условию Липшица [70]. Близок к классу почти дифференцируемых функций класс локально липшицевых функций, исследованный Кларком [84]. Несколько интересных классов негладких функций предложил Мифлин [101].

Для указанных выше классов функций введены обобщенные понятия градиентов (градиентных множеств) на основе использования либо теорем отделимости (так в выпуклом случае вводятся субградиенты), либо предельного перехода.

## § 2. Выпуклые функции. Понятие субградиента

Изложим ряд основных свойств выпуклых множеств и функций, которые понадобятся нам в дальнейшем.

Подмножество  $W$  линейного пространства  $L$  называется выпуклым, если из того, что  $x \in W$  и  $y \in W$  следует, что  $(1 - \alpha)x + \alpha y \in W$  при любом  $\alpha \in [0, 1]$ .

Гиперплоскость  $P$ , проходящая через точку  $x$  границы выпуклого множества  $W \subseteq E_n$  так, что  $W$  целиком лежит в одном из замкнутых полупространств, общей границей которых является  $P$ , называется опорной к  $W$  в точке  $x$ .

**Теорема 1.1.** *Через каждую точку границы выпуклого множества  $W \subseteq E_n$  можно провести по крайней мере одну опорную к  $W$  в данной точке гиперплоскость.*

Для выпуклых множеств характерны так называемые свойства отделимости. Приведем формулировки некоторых из них.

**Теорема 1.2.** *Пусть  $W_1$  и  $W_2$  — два замкнутых непересекающихся выпуклых множества, входящих в  $E_n$ , одно из которых ограничено. Тогда найдется гиперплоскость*

$$P = \{x : (a, x) + b = 0\}$$

*такая, что*

$$(a, x) + b < 0 \quad \text{при } x \in W_1,$$

$$(a, x) + b > 0 \quad \text{при } x \in W_2.$$

**Теорема 1.3.** *Пусть  $W_1, W_2 \subseteq E_n$  — выпуклые замкнутые множества, не имеющие общих внутренних точек, причем множество внутренних точек  $W_1$  не пустое. Тогда если  $\bar{x} \in W_1 \cap W_2$ , то существует гиперплоскость  $(a, x - \bar{x}) =$*

$= 0$ ,  $a \neq 0$ , такая, что

$$(a, x - \bar{x}) \leq 0 \quad \text{для } x \in W_1,$$

$$(a, x - \bar{x}) \geq 0 \quad \text{для } x \in W_2.$$

Линейное многообразие минимальной размерности, включающее выпуклое множество  $W$ , назовем носителем этого множества и обозначим  $N(W)$ . Под размерностью  $\rho(W)$  множества  $W$  будем понимать размерность многообразия  $N(W)$ . Точку  $x \in W$  назовем  $N$ -внутренней, если существует такая окрестность этой точки  $S(x)$ , что  $S(x) \cap N(W) \subseteq W$ . В противном случае  $x \in W$  будет называться  $N$ -граничной точкой. Точка  $x \in W$  называется крайней, если она не может быть представлена в следующем виде:

$$x = \alpha x_1 + (1 - \alpha) x_2; \quad 0 < \alpha < 1; \quad x_1, x_2 \in W; \quad x_1, x_2 \neq x.$$

Любая крайняя точка является  $N$ -граничной.

**Теорема 1.4.** Произвольную точку  $x$  выпуклого замкнутого ограниченного множества  $W$  можно представить в виде выпуклой комбинации не более чем  $\rho(W) + 1$  крайних точек множества  $W$ , где  $\rho(W)$  — размерность  $W$ :

$$x = \sum_{i=1}^{\rho(W)+1} \alpha_i x_i \left( \sum_{i=1}^{\rho(W)+1} \alpha_i = 1; \quad \alpha_i \geq 0; \quad x_i — \text{крайние точки } W; \right. \\ \left. i = 1, \dots, \rho(W) + 1 \right).$$

Пусть  $f(x) = f(t_1, \dots, t_n)$  — функция, определенная на множестве  $M \subseteq E_n$ . Рассмотрим  $(n+1)$ -мерное пространство точек вида  $z = \{u, t_1, \dots, t_n\} = \{u, x\}$ . Назовем надграфиком функции  $f$  множество  $M_f = \{z: u \geq f(x); x \in M\}$ .

Функция  $f(x)$ , определенная на множестве  $M$ , называется выпуклой, если надграфик этой функции — выпуклое множество. Функция  $\varphi(x)$  называется вогнутой, если  $-\varphi(x)$  — выпуклая функция.

**Теорема 1.5.** Для того чтобы функция  $f(x)$  была выпуклой, необходимо и достаточно, чтобы область ее определения  $M$  была выпуклой и для любых  $x_1, x_2 \in M$ ,  $\alpha \in [0, 1]$  было справедливо следующее неравенство:

$$(1 - \alpha) f(x_1) + \alpha f(x_2) \geq f[(1 - \alpha)x_1 + \alpha x_2].$$

**Теорема 1.6.** Выпуклая функция непрерывна во всех внутренних точках множества  $M$ , на котором она определена.

Перейдем к определению понятия субградиента (обобщенного градиента). Как следует из теоремы 1.1, через любую точку границы выпуклого множества может быть проведена, по крайней мере, одна опорная гиперплоскость. Применим это свойство к надграфику  $M_f$  выпуклой функции  $f(x)$ . Любая точка  $z$ , лежащая на графике этой функции (т. е.  $z = \{u, x\}$ ;  $u = f(x)$ ), является граничной точкой надграфика.

Пусть  $x_0$  — внутренняя точка  $M$  и  $z_0 = \{f(x_0), x_0\}$ . Запишем уравнение опорной гиперплоскости к множеству  $M_f$  в точке  $z_0$ :

$$(a, z - z_0) = 0; \quad a \neq 0, \quad (1.1)$$

где  $a$  —  $(n + 1)$ -мерный вектор. Для удобства дальнейших рассуждений выделим в  $a$  первую компоненту  $\lambda$ :  $a = \{\lambda, b\}$ , где  $b$  —  $n$ -мерный вектор. Тогда (1.1) запишется в виде

$$\lambda(u - f(x_0)) + (b, x - x_0) = 0.$$

Пусть  $a$  — вектор внешней нормали к опорной гиперплоскости по отношению к телу  $M_f$ . Тогда для любого  $z = \{u, x\} \in M_f$  должно выполняться

$$\lambda(u - f(x_0)) + (b, x - x_0) \leq 0. \quad (1.2)$$

По определению  $M_f$ , если  $\{\bar{u}, x\} \in M_f$ , то и  $\{u, x\} \in M_f$  при  $u \geq \bar{u}$ . Таким образом, неравенство (1.2) должно сохраняться при сколь угодно больших  $u$ . Значит,  $\lambda \leq 0$ .

Если предположить  $\lambda = 0$ , то, учитывая, что  $x_0$  — внутренняя точка  $M$ ,  $(b, x - x_0) \leq 0$  для всех  $x \in S(x_0)$ , где  $S(x_0)$  — достаточно малая окрестность точки  $x_0$ , получим  $b = 0$ . Но тогда  $a = \{\lambda, b\} = 0$ , что противоречит (1.1). Таким образом,  $\lambda < 0$ .

Приняв  $g(x_0) = -\frac{b}{\lambda}$ ,  $u = f(x)$ , получим

$$f(x) - f(x_0) \geq (g(x_0), x - x_0), \quad x \in M. \quad (1.3)$$

Таким образом, для произвольной внутренней точки  $x_0$  области определения выпуклой функции существует вектор  $g(x_0)$ , удовлетворяющий неравенству (1.3).

**Определение.** Пусть задана выпуклая функция  $f(x)$  с областью определения  $M$ ,  $x_0$  — внутренняя точка  $M$ . Вектор  $g(x_0)$ , удовлетворяющий соотношению (1.3) для произвольного  $x \in M$ , называется субградиентом или обобщенным градиентом функции  $f(x)$  в точке  $x_0$ .

**Теорема 1.7.** Множество обобщенных градиентов  $G(x_0)$  выпуклой функции  $f(x)$  в точке  $x_0$  не пустое, ограниченное, выпуклое и замкнутое для любой внутренней точки области определения  $M$ .

**Доказательство.** Существование хотя бы одного обобщенного градиента следует из вышеприведенных рассуждений. Покажем ограниченность  $G(x_0)$ . Пусть  $G(x_0)$  содержит точки, отличные от точки 0. Рассмотрим

$$x(\varepsilon, g) = x_0 + \frac{\varepsilon g}{\|g\|}, \quad (1.4)$$

где

$$\varepsilon > 0; g \in G(x_0); g \neq 0.$$

При достаточно малом  $\varepsilon > 0$   $x(\varepsilon, g) \in M$ . Из непрерывности  $f(x)$  в точке  $x_0$  вытекает, что для любого  $\delta > 0$  найдется  $\varepsilon > 0$  такое, что

$$|f(x) - f(x_0)| < \delta \quad \text{при} \quad \|x - x_0\| \leq \varepsilon.$$

С другой стороны, из (1.3), (1.4)  $f(x) - f(x_0) \geq \varepsilon \|g\|$ , откуда  $\|g\| < \frac{\delta}{\varepsilon}$ , т. е.  $G(x_0)$  — ограничено.

Замкнутость множества  $G(x_0)$  следует из того, что знак  $\geq$  в неравенстве (1.3) сохраняется при предельном переходе. Легко проверяется и выпуклость. В самом деле, пусть  $g_1$  и  $g_2$  — два обобщенных градиента в точке  $x_0$ . Тогда для любого  $x \in M$

$$f(x) - f(x_0) \geq (g_1, x - x_0); \quad f(x) - f(x_0) \geq (g_2, x - x_0).$$

Пусть  $\alpha_1, \alpha_2 \geq 0, \alpha_1 + \alpha_2 = 1$ . Умножим первое неравенство на  $\alpha_1$ , второе на  $\alpha_2$  и сложим:

$$\begin{aligned} (\alpha_1 + \alpha_2) |f(x) - f(x_0)| &= f(x) - f(x_0) \geq \\ &\geq (\alpha_1 g_1 + \alpha_2 g_2, x - x_0), \quad \forall x \in M. \end{aligned}$$

Отсюда следует, что  $g_3 = \alpha_1 g_1 + \alpha_2 g_2$  является субградиентом, т. е. выпуклая комбинация субградиентов является субградиентом. Значит, множество  $G(x_0)$  — выпуклое. Теорема доказана.

**Теорема 1.8.** В любой внутренней точке  $x_0$  области определения  $M$  выпуклая функция  $f(x)$  имеет производную по любому направлению  $\eta$ . Эта производная удовлетворяет соотношению

$$f'_\eta(x_0) = \max_{g \in G(x_0)} (g, \eta). \quad (1.5)$$

Доказательство. Рассмотрим функцию

$$\varphi(t) = \frac{f(x_0 + t\eta) - f(x_0)}{t}.$$

Пусть  $\varepsilon > 0$  — достаточно малое число такое, что  $x_0 \pm \varepsilon\eta \in M$ . Тогда на отрезке  $0 < t < \varepsilon$  функция  $\varphi(t)$  непрерывна. Покажем, что она монотонно не убывает. Пусть  $0 < t_1 < t_2 < \varepsilon$ . Получаем

$$\begin{aligned} \varphi(t_2) - \varphi(t_1) &= \frac{1}{t_1 t_2} [t_1 f(x_0 + t_2\eta) - t_2 f(x_0 + t_1\eta) + \\ &\quad + (t_2 - t_1) f(x_0)] = \\ &= \frac{1}{t_1} \left\{ \left[ \frac{t_1}{t_2} f(x_0 + t_2\eta) + \left(1 - \frac{t_1}{t_2}\right) f(x_0) \right] - \right. \\ &\quad \left. - f \left[ (x_0 + t_2\eta) \frac{t_1}{t_2} + x_0 \left(1 - \frac{t_1}{t_2}\right) \right] \right\}. \end{aligned}$$

Из теоремы 1.5 следует, что  $\varphi(t_2) - \varphi(t_1) \geq 0$ . Покажем, что функция  $\varphi(t)$  ограничена снизу на отрезке  $(0, \varepsilon)$ . Для этого рассмотрим  $\varphi\left(-\frac{\varepsilon}{2}\right)$ . Легко видеть, что  $\varphi\left(-\frac{\varepsilon}{2}\right)$  является границей снизу для  $\varphi(t)$  при  $0 < t < \varepsilon$ . Таким образом, существует предел  $\lim_{t \rightarrow +0} \varphi(t) = f'_\eta(x_0)$ , причем  $f'_\eta(x_0) \leq \varphi(t)$ ;  $t \geq 0$ .

Докажем теперь формулу (1.5). Для любого  $g \in G(x_0)$  справедливо неравенство ( $t > 0$ )

$$\varphi(t) = \frac{f(x_0 + t\eta) - f(x_0)}{t} \geq \frac{(g, t\eta)}{t} = (g, \eta).$$

Таким образом,

$$f'_\eta(x_0) \geq \max_{g \in G(x_0)} (g, \eta). \quad (1.6)$$

В  $(n+1)$ -мерном пространстве рассмотрим надграфик функции  $f(x)$

$$M_f = \{(u, x) : u \geq f(x)\}$$

и луч

$$L_\eta = \{(u, x) : x = x_0 + t\eta; u = f(x_0) + f'_\eta(x_0)t; t \geq 0\}.$$

Так как  $f'_\eta(x_0) \leq \varphi(t)$  для  $t > 0$ , то этот луч не пересекает внутренние точки надграфика. В силу этого и теоремы 1.3 существует опорная гиперплоскость к надграфiku, прохо-

дящая через точку  $\{f(x_0), x_0\}$  и определяемая субградиентом  $\bar{g}(x_0)$ , который обладает следующим свойством:

$$f(x) - f(x_0) \geq (\bar{g}(x_0), x - x_0) \geq t f'_\eta(x_0), \quad x = x_0 + t\eta,$$

т. е.  $(\bar{g}(x_0), \eta) \geq f'_\eta(x_0)$ . Сопоставляя это неравенство с (1.6), получаем  $f'_\eta(x_0) = \max_{g \in G(x_0)} (g, \eta)$ , что и требовалось доказать.

**Теорема 1.9.** *Для того чтобы функция  $f(x)$  была выпуклой на  $E_n$ , необходимо и достаточно, чтобы ее производные по направлению существовали во всех точках, причем  $f'_\eta(x + t\eta)$  являлись монотонно неубывающими функциями от  $t$ .*

**Доказательство.** Необходимость следует непосредственно из теоремы 1.8. Пусть  $t_1 < t_2$ . Тогда

$$f'_\eta(x + t_1\eta) = \max_{g \in G(x+t_1\eta)} (g, \eta) = (g_1^*, \eta), \quad g_1^* \in G(x + t_1\eta);$$

$$f'_\eta(x + t_2\eta) = \max_{g \in G(x+t_2\eta)} (g, \eta) = (g_2^*, \eta), \quad g_2^* \in G(x + t_2\eta);$$

$$\begin{aligned} (t_2 - t_1) f'_\eta(x + t_2\eta) &= (t_2 - t_1) (g_2^*, \eta) \geq \\ &\geq f(x + t_2\eta) - f(x + t_1\eta) \geq (g_1^*, (t_2 - t_1)\eta) = \\ &= (t_2 - t_1) (g_1^*, \eta) = (t_2 - t_1) f'_\eta(x + t_1\eta). \end{aligned}$$

Отсюда  $f'_\eta(x + t_2\eta) \geq f'_\eta(x + t_1\eta)$ .

Для доказательства достаточности заметим, что при  $t_2 > t_1$  в силу монотонности  $f'_\eta(x + t\eta)$  по  $t$

$$f(x + t_2\eta) - f\left(x + \frac{t_1 + t_2}{2}\eta\right) = \int_{\frac{t_1+t_2}{2}}^{t_2} f'_\eta(x + t\eta) dt \geq$$

$$\begin{aligned} &\frac{t_1+t_2}{2} \\ &\geq \int_{t_1}^{\frac{t_1+t_2}{2}} f'_\eta(x + t\eta) dt = f\left(x + \frac{t_1+t_2}{2}\eta\right) - f(x + t_1\eta). \end{aligned}$$

Отсюда

$$\frac{1}{2} [f(x + t_1\eta) + f(x + t_2\eta)] \geq f\left(x + \frac{t_1+t_2}{2}\eta\right),$$

т. е.  $f(x)$  — выпуклая функция. Теорема доказана.

**Теорема 1.10.** *Если функция  $f(x)$  дважды непрерывно дифференцируема, то она является выпуклой тогда и только*

тогда, когда ее гессиан  $H(x)$  в любой точке является неотрицательно определенной матрицей.

**Доказательство.** Запишем разложение  $f(x)$  по формуле Тейлора:

$$\bullet f(x) = f(x_0) + (g(x_0), x - x_0) + \frac{1}{2} (H(x_0)(x - x_0), x - x_0) + o(\|x - x_0\|^2)$$

и вычислим производные по направлению

$$f'_\eta(x_0 + t\eta) = (g(x_0), \eta) + t(H(x_0)\eta, \eta) + o(t);$$

$$f''_\eta(x_0) = (H(x_0)\eta, \eta) \geq 0.$$

Это эквивалентно тому, что вторая производная по направлению неотрицательна и  $f'_\eta(x_0 + t\eta)$  является монотонно неубывающей функцией. Используя теорему 1.9, получаем требуемый результат.

Направление  $\eta$ ,  $\eta \neq 0$ , назовем направлением наискорейшего спуска выпуклой функции  $f(x)$  в точке  $x_0$ , являющейся внутренней точкой области определения  $M$ , если

$$\min_{\|\xi\|=1} f'_\xi(x_0) = \frac{1}{\|\eta\|} f'_\eta(x_0).$$

**Теорема 1.11.** *Направление наискорейшего спуска в точке  $x_0$  задается вектором  $\eta \in -G(x_0)$ , лежащим на кратчайшем расстоянии от начала координат.*

**Доказательство.** Пусть  $G(x_0)$  не содержит точки 0. Рассмотрим шар  $S = \{x : \|x\| \leq \|\eta\|\}$ ,  $\eta$  является общей граничной точкой  $S$  и  $-G(x_0)$ . По теореме об отделимости 1.3 через точку  $\eta$  можно провести разделяющую гиперплоскость, которая является опорной одновременно к  $S$  и  $-G(x_0)$ . Так как  $S$  — шар, единственная опорная гиперплоскость в точке  $\eta$  к  $S$  имеет вид  $(\eta, x - \eta) = 0$ , причем для  $x \in S$   $(\eta, x - \eta) \leq 0$ , а для  $x \in -G(x_0)$   $(\eta, x - \eta) \geq 0$ , т. е.  $(\eta, x) \geq \|\eta\|^2$  при  $x \in -G(x_0)$ , и

$$f'_\eta(x_0) = \max_{g \in G(x_0)} (\eta, g) = - \min_{g \in -G(x_0)} (\eta, g) = -(\eta, \eta).$$

Пусть  $\|\xi\| = 1$ , тогда

$$\begin{aligned} f'_\xi(x_0) &= \max_{g \in G(x_0)} (g, \xi) = - \min_{g \in -G(x_0)} (g, \xi) \geq - \min_{g \in -G(x_0)} \|g\| = \\ &= -\|\eta\| = \frac{f'_\eta(x_0)}{\|\eta\|}. \end{aligned}$$

Таким образом,  $\min_{\|\xi\|=1} f'_\xi(x_0) = \frac{1}{\|\eta\|} f'_\eta(x_0)$ , т. е.  $\eta$  является направлением наискорейшего спуска, что и требовалось доказать. Если в точке  $x_0$  функция непрерывно дифференцируема, то  $\eta$  совпадает с антиградиентом.

*Следствие.* Для того чтобы внутренняя точка области определения  $f(x)$ ,  $x_0 \in M$ , была точкой минимума, необходимо и достаточно, чтобы  $G(x_0)$  содержало точку 0.

В самом деле, если  $G(x_0)$  содержит 0, то для любой  $x \in M$   $f(x) - f(x_0) \geq 0$  и  $x_0$  есть точка минимума. Если  $G(x_0)$  не содержит 0, то существует направление наискорейшего спуска, производная по которому отрицательна, т. е.  $x_0$  не является точкой минимума. Если область определения  $M \subset E_n$  имеет размерность, меньшую чем  $n$ , то роль внутренних точек играют  $N$ -внутренние точки. По аналогии с субградиентным множеством можно ввести понятие  $N$ -субградиентного множества  $NG(x_0)$  как совокупности векторов  $g$ , удовлетворяющих условию  $f(x) - f(x_0) \geq (g, x - x_0)$  для всех  $x \in M$  и принадлежащих подпространству векторов, получающемуся путем параллельного сдвига носителя  $M$ . «Внутренние» субградиентные множества  $NG(x_0)$  обладают теми же свойствами, что и обычные субградиентные множества.

### § 3. Некоторые способы вычисления субградиента

Рассмотрим основные операции, по отношению к которым класс выпуклых функций является замкнутым: 1) операция получения линейных комбинаций функций с неотрицательными коэффициентами; 2) операция взятия максимума по конечному семейству функций. В приложениях сложные выпуклые функции чаще всего получаются из более простых (например, линейных, квадратичных) путем суперпозиции указанных двух операций, причем операция взятия максимума является основным источником формирования негладких выпуклых функций из гладких.

Приведем теоремы, которые описывают множество субградиентов функций, являющихся линейными комбинациями выпуклых функций с неотрицательными коэффициентами и функциями максимума.

**Теорема 1.12.** Пусть  $f(x) = \sum_{i=1}^k a_i f_i(x)$ , причем  $a_i \geq 0$ ,  $f_i(x)$  — выпуклые функции ( $i = 1, \dots, k$ ). Тогда  $f(x)$  — вы-

пуклая функция и множество обобщенных градиентов в точке  $x_0$  состоит из всевозможных векторов вида

$$g(x_0) = \sum_{i=1}^k a_i g_i(x_0); \quad g_i(x_0) \in G_{f_i}(x_0). \quad (1.7)$$

Доказательство. Пусть  $x_1, x_2 \in E_n$ ;  $\alpha_1, \alpha_2 \geq 0$ ;  $\alpha_1 + \alpha_2 = 1$ , тогда

$$\begin{aligned} f(\alpha_1 x_1 + \alpha_2 x_2) &= \sum_{i=1}^k a_i f_i(\alpha_1 x_1 + \alpha_2 x_2) \leq \\ &\leq \sum_{i=1}^k a_i [\alpha_1 f_i(x_1) + \alpha_2 f_i(x_2)] = \alpha_1 \sum_{i=1}^k a_i f_i(x_1) + \\ &+ \alpha_2 \sum_{i=1}^k a_i f_i(x_2) = \alpha_1 f(x_1) + \alpha_2 f(x_2). \end{aligned}$$

Отсюда  $f(x)$  — выпуклая функция. Из определения производной по направлению сразу следует, что

$$f'(x) = \sum_{i=1}^k a_i f'_i(x). \quad (1.8)$$

Рассмотрим произвольный вектор  $g(x_0)$  вида (1.7). Получаем

$$\begin{aligned} f(x) - f(x_0) &= \sum_{i=1}^k a_i [f_i(x) - f_i(x_0)] \geq \\ &\geq \sum_{i=1}^k (g_i(x_0), x - x_0) a_i = \left( \sum_{i=1}^k a_i g_i(x_0), x - x_0 \right). \end{aligned}$$

Отсюда следует, что вектор  $g(x_0)$  является обобщенным градиентом  $f(x)$  в точке  $x_0$ . Обратно, пусть  $\hat{g}$  — крайняя точка множества  $G_f(x_0)$ ;  $\eta$  — вектор внешней нормали к такой опорной гиперплоскости  $G_f(x_0)$  в этой точке, которая имеет с  $G_f(x_0)$  единственную общую точку  $\hat{g}$ . Тогда  $f'_\eta(x_0) = \max_{g \in G_f(x_0)} (g, \eta)$ , причем максимум достигается в единственной точке  $\hat{g}$ . Из (1.8) вытекает, что тот же результат получится для  $\bar{g} = \sum_{i=1}^k a_i \bar{g}_i(x_0)$ , где  $\bar{g}_i(x_0)$  — вектор, на котором достигается  $\max (g, \eta)$ . Таким образом, подмножество  $G_f(x_0)$

крайних точек множества  $G_f(x_0)$  имеет вид  $\sum_{i=1}^k a_i \bar{g}_i(x_0)$ , где  $\bar{g}_i(x_0) \in G_i(x_0)$ , а значит, и любой вектор из  $G_f(x_0)$  как выпуклая комбинация крайних точек имеет такой вид. Теорема 1.12 доказана.

**Теорема 1.13.** Пусть  $f_i(x)$  — выпуклые функции ( $i = 1, 2, \dots, m$ ). Тогда функция  $\varphi(x) = \max_{i \in \bar{1}^m} f_i(x)$  является выпуклой, причем множество обобщенных градиентов функции  $\varphi(x)$  в точке  $x_0$  содержит обобщенные градиенты  $g_{f_i}(x_0)$ ,  $i \in I(x_0)$ , где  $I(x_0) = \{i : \varphi(x_0) = f_i(x_0)\}$ .

**Доказательство.** Пусть  $\alpha_1, \alpha_2 \geq 0$ ;  $\alpha_1 + \alpha_2 = 1$ ;  $x_1, x_2 \in E_n$ , тогда

$$\begin{aligned} \varphi(\alpha_1 x_1 + \alpha_2 x_2) &= f_{i^*}(\alpha_1 x_1 + \alpha_2 x_2) \leq \\ &\leq \alpha_1 f_{i^*}(x_1) + \alpha_2 f_{i^*}(x_2) \leq \alpha_1 \varphi(x_1) + \alpha_2 \varphi(x_2), \\ &i^* \in I(\alpha_1 x_1 + \alpha_2 x_2). \end{aligned}$$

Отсюда следует выпуклость  $\varphi(x)$ .

Далее, пусть  $i \in I(x_0)$ , тогда

$$(g_{\bar{i}}(x_0), x - x_0) \leq \bar{f}_i(x) - \bar{f}_i(x_0) \leq \varphi(x) - \varphi(x_0).$$

Значит,  $g_{\bar{i}}(x_0)$  — обобщенный градиент  $\varphi(x)$  в точке  $x_0$ . Теорема 1.13 доказана.

Таким образом, во многих практически важных случаях вычисление обобщенного градиента практически не отличается по сложности от вычисления градиента. Отметим также, что выпуклая функция почти везде дифференцируема (это значит, что почти везде, т. е. за исключением точек множества меры нуль, обобщенный градиент совпадает с градиентом). При этом градиент выпуклой функции непрерывен на множестве точек, где он определен [5].

Свойство монотонности производной по направлению от выпуклой функции позволяет легко построить алгоритм, вычисляющий производную по направлению  $f'_\eta(x_0)$  с заданной точностью, если предположить, что имеется возможность вычисления  $f(x)$  со сколь угодно высокой точностью и в точке  $x_0$  функция  $f(x)$  дифференцируема. В самом деле, справедливо неравенство

$$\frac{1}{h} [f(x + h\eta) - f(x)] \geq f'_\eta(x) \geq \frac{1}{h} [f(x) - f(x - h\eta)],$$

$$h > 0,$$

которое дает двусторонние оценки на производную, и легко подобрать  $h$  такое, чтобы погрешность не превышала заданную. Так как вычисление градиента сводится к вычислению частных производных, то получаем алгоритм вычисления градиента с заданной точностью в той точке, где он существует.

С другой стороны, легко построить примеры, показывающие, что если обобщенный градиент в данной точке определен неоднозначно, то не существует алгоритма, который бы гарантировал для любого  $\varepsilon > 0$  построение вектора  $g_\varepsilon$  такого, что

$$\min_{g \in G_f(x)} \|g_\varepsilon - g\| < \varepsilon.$$

Однако если изменить задачу о приближенном вычислении обобщенного градиента, то можно построить соответствующий алгоритм.

Построим вектор  $g^*$ , для которого при заданных  $\varepsilon, \delta$  имеется такая точка  $x$ , лежащая в  $\delta$ -окрестности точки  $x_0$ , что

$$\min_{g \in G_f(x)} \|g^* - g\| < \varepsilon.$$

Для того чтобы обосновать такое построение, используем следующее утверждение, являющееся следствием непрерывности  $f(x)$ : пусть для некоторых  $x \in E_n, \bar{\varepsilon} > 0, h > 0$  справедливо неравенство

$$f(x + h\eta) + f(x - h\eta) - 2f(x) < \bar{\varepsilon}h,$$

тогда для любого  $\bar{\delta} > 0$  найдется такая окрестность точки  $x - S_{\bar{\delta}}(x)$ , что

$$f(y + h\eta) + f(y - h\eta) - 2f(y) < (\bar{\varepsilon} + \bar{\delta})h, \quad y \in S_{\bar{\delta}}(x).$$

Построение векторов  $g^*, \bar{x}$  состоит из  $n$  этапов, на каждом из которых получаем векторы  $g_k^*, \bar{x}_k, 1 \leq k \leq n$ , так что  $g_n^* = g^*, \bar{x}_n = \bar{x}$  — искомые векторы.

Пусть  $\{e_1, \dots, e_n\}$  — ортонормированная система векторов,  $\bar{x}_k = x_0 + \sum_{i=1}^k c_i e_i$  такое, что  $\|\bar{x}_k - x_0\| < \frac{\delta_k}{n}$ , и имеется  $h_k > 0$  такое, что

$$f(\bar{x}_k + h_k e_i) + f(\bar{x}_k - h_k e_i) - 2f(\bar{x}_k) < \frac{keh_k}{n\sqrt{n}}; \quad 1 \leq i \leq k < n.$$

Рассмотрим точки вида

$$y_r = \bar{x}_k + \frac{\delta r}{\rho n} e_{k+1}, \quad -p \leq r \leq p.$$

Последовательность  $s_r = \left\{ \frac{\rho n}{\delta} [f(y_{r+1}) - f(y_r)] \right\}_{r=-p}^{r=p-1}$  монотонно возрастает. С другой стороны, при любом  $p$

$$s_{-p} \geq f'_{e_{k+1}} \left( \bar{x}_k - \frac{\delta}{n} e_{k+1} \right); \quad s_{p-1} \leq f'_{e_{k+1}} \left( \bar{x}_k + \frac{\delta}{\sqrt{n}} e_{k+1} \right).$$

В силу этого и приведенного выше утверждения при достаточно большом  $p = \bar{p}$  найдется  $j^*$ ,  $-\bar{p} \leq j^* \leq \bar{p} - 1$ , такое, что  $s_j - s_{j-1} < \frac{(k+1)\varepsilon}{n\sqrt{n}}$  и

$$\frac{f(y_{j^* + h_k e_i}) + f(y_{j^* - h_k e_i}) - 2f(y_{j^*})}{h_k} < \frac{(k+1)\varepsilon}{n\sqrt{n}}.$$

Выберем  $h_{k+1} = \min \left\{ h_k, \frac{\delta}{\rho n} \right\}$  и примем  $y_{j^*} = \bar{x}_{k+1}$ . Тогда справедливы соотношения

$$f(\bar{x}_{k+1} + h_{k+1} e_i) + f(\bar{x}_{k+1} - h_{k+1} e_i) - 2f(\bar{x}_{k+1}) < \frac{(k+1)\varepsilon h_{k+1}}{n\sqrt{n}}, \quad i = 1, \dots, k+1;$$

$$\|\bar{x}_{k+1} - x_0\| < \frac{\delta(k+1)}{n}.$$

По индукции после  $n$  шагов получим вектор  $\bar{x}_n = \bar{x}$  такой, что для некоторого  $h_n > 0$

$$\frac{1}{h_n} [f(\bar{x} + h_n e_i) + f(\bar{x} - h_n e_i) - 2f(\bar{x})] < \frac{\varepsilon}{\sqrt{n}}, \quad i = 1, \dots, n; \quad (1.9)$$

$$\|\bar{x} - x_0\| < \delta. \quad (1.10)$$

Заметим, что справедливы неравенства

$$\begin{aligned} \frac{1}{h_n} [f(\bar{x}) - f(\bar{x} - h_n e_i)] &\leq f'_{e_i}(\bar{x}) \leq \\ &\leq \frac{1}{h_n} [f(\bar{x} + h_n e_i) - f(\bar{x})], \end{aligned} \quad (1.11)$$

$$\begin{aligned} \frac{1}{h_n} [f(\bar{x}) - f(\bar{x} - h_n e_i)] &\leq -f'_{-e_i}(\bar{x}) \leq \\ &\leq \frac{1}{h_n} [f(\bar{x} + h_n e_i) - f(\bar{x})]. \end{aligned} \quad (1.12)$$

Далее,

$$f'_{e_i}(\bar{x}) = \max_{g \in G_f(\bar{x})} (g, e_i), \quad (1.13)$$

$$-f'_{-e_i}(x) = \min_{g \in G_f(\bar{x})} (g, e_i). \quad (1.14)$$

Рассмотрим гиперпараллелепипед  $\Pi$ :

$$\Pi = \left\{ z : \frac{f(\bar{x}) - f(\bar{x} - h_n e_i)}{h_n} \leq (z, e_i) \leq \frac{f(\bar{x} + h_n e_i) - f(\bar{x})}{h_n} \right\}.$$

Из (1.11) — (1.14) вытекает, что  $G_f(\bar{x}) \subseteq \Pi$ . Примем в качестве  $g^*$  центр этого параллелепипеда:

$$g^* = \left\{ \frac{f(\bar{x} + h_n e_i) - f(\bar{x} - h_n e_i)}{2h_n} \right\}_{i=1}^n.$$

Из (1.9), (1.10) вытекает, что

$$\|g^* - g\| < \frac{\varepsilon}{2} \quad \text{при } g \in G_f(\bar{x}).$$

Таким образом мы построили искомые вектор  $g^*$  и точку  $\bar{x}$ .

Описанный выше алгоритм дает возможность построить вектор, который сколь угодно близок к обобщенному градиенту  $f(x)$ , взятому в точке, лежащей в сколь угодно малой окрестности заданной точки.

#### § 4. Почти дифференцируемые функции

В большинстве работ, посвященных методам градиентного типа, предполагается непрерывность градиента, а для обобщенных градиентных методов — выпуклость минимизируемой функции. Но часто в прикладных задачах нельзя постулировать гладкость (или выпуклость) минимизируемой функции. Например, в практике планирования и проектирования элементы функции затрат обычно задаются в виде кусочно-гладких не обязательно выпуклых функций от параметра, характеризующего производительность (пропускную способность) того или иного устройства. Отсутствие свойств выпуклости и гладкости характерно для широкого круга минимаксных задач.

Все это побуждает рассматривать задачи минимизации на классе функций, достаточно широком, чтобы охватить кусочно-гладкие функции и функции, встречающиеся в минимаксных задачах, и в то же время настолько узком, что-

бы для этих функций можно было определить некоторый естественный аналог градиента и иметь возможность использовать методы градиентного типа для нахождения локальных минимумов.

Одним из таких классов является класс почти дифференцируемых функций, определенный в работе [70].

**Определение.** Функция  $f(x)$ , определенная на  $n$ -мерном евклидовом пространстве  $E_n$ , называется почти дифференцируемой, если она удовлетворяет следующим условиям: а) в любой ограниченной области удовлетворяет условию Липшица; б) почти везде дифференцируема; в) ее градиент непрерывен на том множестве  $M$ , где он определен.

Отметим, что свойство б) не является независимым. Справедлива следующая теорема [46].

Пусть действительная функция  $f(x)$ , определенная на открытом множестве  $D \subset E_n$ , имеет конечные частные производные числа по всем направлениям, т. е.

$$\overline{\lim}_{t \rightarrow 0} \left| \frac{f(x+tv) - f(x)}{t} \right| < +\infty$$

для любой точки  $x \in D$  и любого вектора  $v \in E_n$ . Тогда  $f(x)$  дифференцируема почти всюду на  $D$ .

Ясно, что если функция удовлетворяет условию Липшица в любой открытой ограниченной области  $S \subset E_n$ , т. е.

$$|f(x) - f(y)| \leq L \|x - y\|, \quad L > 0; \quad x, y \in S,$$

то

$$\overline{\lim}_{t \rightarrow 0} \left| \frac{f(x+tv) - f(x)}{t} \right| \leq L \|v\| < \infty$$

при  $x \in S$ . Таким образом, из того, что функция удовлетворяет условию Липшица, следует, что она почти всюду дифференцируема.

**Определение.** Почти-градиентом функции  $f(x)$  в точке  $x_0$  называется вектор, являющийся предельной точкой некоторой последовательности градиентов  $g(x_1), g(x_2), \dots, g(x_k), \dots$ , где  $\{x_k\}_{k=1}^{\infty}$  — последовательность точек, сходящихся к  $x_0$ , и такая, что во всех точках этой последовательности  $f(x)$  дифференцируема.

**Теорема 1.14.** Множество  $G(x)$  почти-градиентов в любой точке  $x \in E_n$  почти дифференцируемой функции является непустым, ограниченным и замкнутым.

**Доказательство.** Из свойства а) следует ограниченность последовательности  $\{g(x_k)\}$ , а отсюда —

непустота множества  $G(x)$  и его ограниченность. Замкнутость является следствием определения почти-градиентов как предельных точек. В самом деле, пусть  $\{\hat{g}^{(k)}(x)\}$  — последовательность почти-градиентов, сходящаяся к  $g^*$ . Тогда для каждого  $k$  найдется  $x^{(k)}$  такое, что  $\|x - x^{(k)}\| \leq \frac{1}{k}$ ;  $\|\hat{g}^{(k)}(x) - g(x^{(k)})\| \leq \frac{1}{k}$ , причем в точке  $x^{(k)}$   $f(x)$  дифференцируема:

$$\begin{aligned} \|g^* - g(x^{(k)})\| &\leq \|g^* - \hat{g}^{(k)}(x)\| + \\ &+ \|\hat{g}^{(k)}(x) - g(x^{(k)})\| \leq \|g^* - \hat{g}^{(k)}(x)\| + \frac{1}{k}, \end{aligned}$$

$\lim_{k \rightarrow \infty} \|g^* - g(x^{(k)})\| = 0$ , откуда следует, что  $g^*$  является почти-градиентом  $f(x)$  в точке  $x$ .

**Теорема 1.15.** *Произвольная выпуклая функция  $f(x)$ , определенная в  $E_n$ , является почти дифференцируемой, а ее любой почти-градиент в точке  $x$  совпадает с некоторым субградиентом.*

*Доказательство.* В силу теоремы Радемейстера [5] произвольная выпуклая функция, определенная в  $E_n$ , почти везде непрерывно дифференцируема. Из общих свойств выпуклой функции известно, что ее производные по направлению в любой ограниченной области равномерно ограничены. Таким образом,  $f(x)$  — почти дифференцируемая функция.

Пусть почти-градиент  $\hat{g}_f(x_0) = \lim_{k \rightarrow \infty} g_f(x_k)$ , где  $\{g_f(x_k)\}_{k=1}^{\infty}$  — градиенты в точках  $\{x_k\}$ ,  $\lim_{k \rightarrow \infty} x_k = x_0$ . Из свойства выпуклости  $f(x)$  вытекает для всех  $x \in E_n$

$$f(x) - f(x_k) \geq (g_f(x_k), x - x_k).$$

Устремив  $k$  к бесконечности, получаем

$$f(x) - f(x_0) \geq (\hat{g}_f(x_0), x - x_0),$$

т. е.  $\hat{g}_f(x_0)$  является обобщенным градиентом функции  $f(x)$  в точке  $x_0$ . Теорема доказана.

**Теорема 1.16.** *Если функции  $f_1(x)$  и  $f_2(x)$  почти дифференцируемы, то  $f_3(x) = f_1(x) \pm f_2(x)$ ,  $f_4(x) = f_1(x) f_2(x)$  также почти дифференцируемы.*

**Доказательство.** Рассмотрим  $N_3$ -объединение множеств  $N_1$  и  $N_2$ , на которых функции  $f_1(x)$  и  $f_2(x)$  не дифференцируемы. Мера  $N_3$  равна нулю. В остальных точках  $f_3(x), f_4(x)$  дифференцируемы и их градиенты подсчитываются по правилам дифференцирования суммы и произведения двух функций. Отсюда следует утверждение теоремы.

**Теорема 1.17.** Если  $f_1(x)$  и  $f_2(x)$  почти дифференцируемы, то  $f(x) = \max(f_1(x), f_2(x))$  также почти дифференцируема.

**Доказательство.** Введем обозначение

$$\varphi^+(x) = \begin{cases} 0, & \varphi(x) \leq 0, \\ \varphi(x), & \varphi(x) > 0. \end{cases}$$

Тогда

$$f(x) = f_1(x) + [f_2(x) - f_1(x)]^+.$$

Из теоремы 1.15 следует, что  $f_2(x) - f_1(x)$  почти дифференцируема. Далее, в силу непрерывности  $f_1(x)$  и  $f_2(x)$  множество  $M_{12} = \{x : f_1(x) > f_2(x)\}$  является открытым, поэтому функция  $f(x)$  будет дифференцируема во всех точках этого множества, в которых дифференцируема функция  $f_1(x)$ . Аналогичное утверждение справедливо для множества  $M_{21} = \{x : f_1(x) < f_2(x)\}$ .

Что касается множества  $M_0 = \{x : f_1(x) = f_2(x)\}$ , то мера границы этого множества равна нулю, а для всех внутренних точек дифференцируемость вытекает из дифференцируемости в соответствующей точке  $f_1(x)$ . Это доказывает теорему.

**Определение.** Назовем обобщенным почти-градиентом функции  $f(x)$  в точке  $x$  произвольный вектор  $g$ , принадлежащий выпуклому замыканию множества почти-градиентов.

Из теоремы 1.14 вытекает, что множество обобщенных почти-градиентов выпукло, ограничено и замкнуто.

В работе Кларка [85] введено понятие обобщенных градиентов, близкое к предложенному нами. В этой работе рассматривается класс функций, определенных на  $E_n$  и удовлетворяющих локальному условию Липшица, и дается следующее определение: обобщенным градиентом  $f$  в точке  $x$ , обозначаемым  $\partial f(x)$ , называется выпуклое замыкание множества пределов последовательностей вида  $\{g_i(x + h_i)\}_{i=1}^{\infty}$  при  $\lim_{i \rightarrow \infty} h_i = 0$ .

Определение, данное Е. Кларком, отличается от нашего тем, что он не требует непрерывности градиентов на множестве, где градиент определен.

### § 5. Полугладкие и полувыпуклые функции

Изложим, следуя [101], основные результаты исследований некоторых важных классов недифференцируемых функций.

Кларк в работе [85] рассматривает класс локально липшицевых функций и аналог производной по направлению

$$f^0(x, d) = \limsup_{\substack{h \rightarrow 0 \\ t \rightarrow +0}} \frac{1}{t} [f(x + h + td) - f(x + h)].$$

Обобщенный градиент (обобщенное градиентное множество)  $f$  в точке  $x$  определяется следующим образом:

$$\partial f(x) = \{g \in E_n; (g, d) \leq f^0(x, d), \forall d \in E_n\}$$

и обладает такими свойствами:

- а)  $\partial f(x)$  непустое выпуклое компактное множество;
- б)  $f^0(x, d) = \max_{g \in \partial f(x)} (g, d)$ ;

в)  $f$  почти везде дифференцируемо и  $\partial f(x)$  — выпуклое замыкание всех точек  $g$  вида  $g = \lim_{k \rightarrow \infty} g_k(x_k)$ , где  $\{x_k\}_{k=1}^{\infty} \rightarrow x$  и  $f$  имеет градиент во всех точках этой последовательности;

г) если  $\{x_k\}$  сходится к  $x$  и  $g_k \in \partial f(x_k)$  для всех  $k$ , то  $\|g_k\| \leq c$  и каждая точка сгущения  $g$  последовательности  $\{g_k\}$  принадлежит  $\partial f(x)$ , т. е. точечно-множественное отображение  $x \rightarrow \partial f(x)$  ограничено и полунепрерывно сверху;

д) выполняется аналог теоремы о среднем: для двух точек  $y$  и  $z$  найдется  $\lambda \in (0, 1)$  и  $g \in \partial f(y + \lambda(z - y))$  такое, что  $f(z) - f(y) = (g, z - y)$ . Если существует производная по направлению

$$f'(x, d) = \lim_{t \rightarrow +0} \frac{1}{t} [f(x + td) - f(x)],$$

совпадающая с  $f^0(x, d)$ , то такая функция называется [43] квазидифференцируемой.

Мифлин ввел понятие полугладкой функции.

**Определение.** Функция  $f(x)$  на  $E_n$  называется полугладкой в точке  $x \in E_n$ , если:

а) функция  $f$  липшицева в некоторой окрестности  $x$ ;  
 б) для любого  $d \in E_n$  и любых положительных последовательностей  $\{t_k\}_{k=1}^{\infty}$ ,  $\{\theta_k\}_{k=1}^{\infty} \subset E_n$  и  $\{g_k\}_{k=1}^{\infty} \subset E_n$  таких, что  $t_k \rightarrow +0$ ,  $\left\{\frac{\theta_k}{t_k}\right\} \rightarrow 0 \in E_n$  и  $g_k \in \partial f(x + t_k d + \theta_k)$ , последовательность  $\{(g_k, d)\}_{k=1}^{\infty}$  имеет единственную точку сгущения.

Доказана следующая теорема: если  $f$  полугладкая в  $x$ , то для любого  $d \in E_n$   $f'(x, d)$  существует и равна  $\lim_{k \rightarrow \infty} (g_k, d)$ , где  $\{g_k\}$  — произвольная последовательность, удовлетворяющая условиям предыдущего определения.

**Определение.** Функция  $f(x)$  полувыпукла в точке  $x \in E_n$ , если она:

- а) липшицева в некоторой окрестности точки  $x$ ;
- б) квазидифференцируема в  $x$ ;
- в) из того, что  $f'(x, d) \geq 0$ , следует  $f(x + d) \geq f(x)$ .

Полувыпуклая функция, являющаяся дифференцируемой, называется псевдовыпуклой [44]. Мы говорим, что функция  $f$  полугладкая (квазидифференцируема, полувыпукла) на  $x \in E_n$ , если она полугладкая (квазидифференцируема, полувыпукла) во всех  $x \in X$ . Для полугладких функций доказана интересная теорема о суперпозиции [101].

**Теорема 1.18.** Пусть  $f_i(x)$  — полугладкие функции, определенные на  $E_n$ ,  $i = 1, \dots, m$ ;  $\Phi(y)$  — полугладкая функция, определенная на  $E_m$ ,

$$y(x) = \{f_1(x), \dots, f_m(x)\}; F(x) = \Phi(y(x));$$

$$G(x) = \text{conv} \{g \in E_n / g = G^{(m)} w; g^{(i)} \in \partial f_i(x), i = 1, \dots, m; w \in \partial \Phi(y(x))\},$$

где  $G^{(m)}$  — матрица, столбцами которой являются  $g^{(i)}$ . Тогда  $F$  — полугладкая функция и  $\partial F(x) \subseteq G(x)$ .

К сожалению, могут быть случаи, когда  $\partial F(x) \not\subseteq G(x)$ , что затрудняет вычисление обобщенного градиента в общем случае. Мифлин для полугладких функций получил интересные обобщения градиентных процессов, сходящихся к стационарным точкам, т. е. к точкам, для которых обобщенное градиентное множество содержит нуль. Эти результаты будут рассмотрены ниже.

Е. А. Нурмынский [35] ввел класс слабо выпуклых функций.

**Определение.** Непрерывная функция  $E_n \rightarrow E_1$  называется слабо выпуклой, если для любого  $x \in E_n$  существует

непустое множество  $M(x)$  векторов  $g$  таких, что для всех  $y \in E_n$

$$f(y) - f(x) \geq (g, y - x) - r(x, y),$$

где  $\frac{r(x, y)}{\|x - y\|} \rightarrow 0$  при  $y \rightarrow x$  равномерно в каждом компактном подмножестве  $E_n$ .

Слабо выпуклые функции образуют подкласс квазидифференцируемых функций. Оба эти класса обладают в определенном смысле свойством замкнутости относительно операции взятия максимума [35], что оправдывает их использование при анализе минимаксных задач.

## ОБОБЩЕННЫЙ ГРАДИЕНТНЫЙ СПУСК

### § 1. Проблема регулировки шага в методе обобщенного градиентного спуска

Пусть задана выпуклая функция  $f(x)$ , определенная на  $E_n$ . Методом обобщенного градиентного спуска будем называть процедуру построения последовательности  $\{x_k\}_{k=0}^{\infty}$  следующего вида:

$$x_{k+1} = x_k - h_{k+1}(x_k) g_i(x_k); \quad (2.1)$$

$x_0$  (нулевое приближение) задается до начала процесса спуска.

Формула (2.1) по внешнему виду ничем не отличается от формулы, задающей спуск в обычных градиентных методах, тем более что обобщенный градиент  $g_i(x)$  почти всюду совпадает с градиентом, так как выпуклая функция почти везде дифференцируема. Однако существенно различен выбор способа регулировки шага. Дело в том, что обычно применяемые для непрерывно дифференцируемых функций методы регулировки шага, например  $h_{k+1}(x_k) = h$ ,  $h$  — достаточно малая константа, в методе с постоянным шаговым множителем или выбор  $h_{k+1}(x_k)$  из условия  $\min_{h_{k+1}} f(x_k - h_{k+1} g_i(x_k))$ ; как в методе наискорейшего спуска, в случае произвольной выпуклой функции, вообще говоря, неприменимы: первый потому, что в точке минимума функция может быть недифференцируемой, и тогда  $\{g_i(x_k)\}_{k=0}^{\infty}$  не обязательно стремится к нулю, если даже  $\{x_k\}_{k=0}^{\infty}$  сходится к точке минимума; второй потому, что минимизируемая последовательность может сходиться к некоторой нестационарной точке.

**Пример [106].** Рассмотрим функцию двух переменных

$$f(x_1, x_2) = \begin{cases} 5(9x_1^2 + 16x_2^2)^{\frac{1}{2}} & \text{для } x_1 > |x_2|, \\ 9x_1 + 16|x_2| & \text{для } x_1 \leq |x_2|. \end{cases}$$

При  $x_1 > 0$   $f(x_1, x_2) = \max \left\{ 5(9x_1^2 + 16x_2^2)^{\frac{1}{2}}, 9x_1 + 16|x_2| \right\}$  и является выпуклой как результат применения операции взятия максимума к двум выпуклым функциям. Покажем, что  $f(x_1, x_2)$  в этой области непрерывно дифференцируема. Линия возможного разрыва градиента удовлетворяет условию  $x_1 = |x_2|$ . Вычислим градиенты

$$\partial \left[ 5(9x_1^2 + 16x_2^2)^{\frac{1}{2}} \right] = \left\{ \frac{5 \cdot 9x_1}{(9x_1^2 + 16x_2^2)^{\frac{1}{2}}}; \frac{5 \cdot 16x_2}{(9x_1^2 + 16x_2^2)^{\frac{1}{2}}} \right\};$$

$$\partial \left[ 5(9x_1^2 + 16x_2^2)^{\frac{1}{2}} \right] \text{ при } x_1 = |x_2| \text{ равен } \{9; 16 \operatorname{sign} x_2\};$$

$$\partial \{9x_1 + 16|x_2|\} = \{9; 16 \operatorname{sign} x_2\}.$$

Таким образом, на линии  $x_1 = |x_2|$  разрыва градиента не происходит, значит, в полуплоскости  $x_1 > 0$   $f(x_1, x_2)$  является непрерывно дифференцируемой. Если мы выберем начальное приближение в области  $x_1 > |x_2| > \left(\frac{9}{16}\right)^{\frac{1}{2}} |x_1|$ , то, как легко проверить, методу наискорейшего спуска будет соответствовать ломаная, соседние звенья которой взаимно ортогональны, а последовательность вершин сходится к точке  $\{0, 0\}$ . Как известно, для непрерывно дифференцируемых функций множество предельных точек минимизирующей последовательности, генерируемой алгоритмом наискорейшего спуска, состоит из стационарных точек. Однако в нашем примере точка  $\{0, 0\}$ , к которой сходится процесс наискорейшего спуска, не является стационарной. Это связано с тем, что в этой точке происходит разрыв градиента (множество точек разрыва градиента функции  $f(x_1, x_2)$  лежит на луче  $x_2 = 0; x_1 \leq 0; \lim_{x_1 \rightarrow -\infty} f(x_1, 0) = -\infty$ ). Этот пример показывает, что для выпуклых функций с разрывным градиентом метод наискорейшего спуска не всегда оказывается пригодным.

Таким образом, проблема регулировки шага в методе обобщенного градиентного спуска в применении к негладким функциям имеет свои специфические трудности, которые приводят к необходимости отказа от требования монотонности спуска.

Для анализа этого вопроса рассмотрим процесс обобщенного градиентного спуска с постоянным шагом, когда

$$h_{k+1}(x_k) = \frac{h}{\|g_f(x_k)\|}; \quad h > 0.$$

**Теорема 2.1.** Пусть  $f(x)$  — выпуклая функция, множество ее минимумов  $M^*$  не пустое и при использовании метода обобщенного градиентного спуска  $h_{k+1}(x_k) = \frac{h}{\|g_f(x_k)\|}$ ,  $h > 0$ . Тогда для любых  $\varepsilon > 0$  и  $x^* \in M^*$  найдутся такие  $k = k^*$  и  $\bar{x}$ , что будет выполняться следующее свойство:  $f(\bar{x}) = f(x_{k^*})$ , причем  $\|\bar{x} - x^*\| < \frac{h}{2} \times (1 + \varepsilon)$ .

**Доказательство.** Если для некоторого  $k^* g_f(x_{k^*}) = 0$ , то  $f(x_{k^*}) = f(x^*)$  и в качестве  $\bar{x}$  можно взять  $x^*$ . Если для всех  $k = 0, 1, \dots g_f(x_k) \neq 0$ , то

$$x_{k+1} = x_k - \frac{hg_f(x_k)}{\|g_f(x_k)\|}, \quad k = 0, 1, 2, \dots$$

Для любого  $x_k \in E_n$

$$\begin{aligned} \|x_{k+1} - x^*\|^2 &= \|x_k - x^* - h g_f(x_k) \|g_f(x_k)\|^{-1}\|^2 = \\ &= \|x_k - x^*\|^2 + h^2 - 2h \left( x_k - x^*, \frac{g_f(x_k)}{\|g_f(x_k)\|} \right). \end{aligned} \quad (2.2)$$

Оценим выражение  $\left( x_k - x^*, \frac{g_f(x_k)}{\|g_f(x_k)\|} \right) = a_k(x^*)$  — расстояние от  $x^*$  до опорной гиперплоскости  $L_k = \{x : (g_f(x_k), x - x_k) = 0\}$ . Обозначим  $b_k(x^*) = \rho(x^*, U_k)$ , где  $U_k = \{x : f(x) = f(x_k)\}$ . Так как множество  $U_k$  и точка  $x^*$  находятся по одну сторону от  $L_k$  и отрезок, проведенный из  $x^*$  в произвольную точку, принадлежащую  $L_k$ , пересекает  $U_k$ , то  $a_k(x^*) \geq b_k(x^*)$ . Из (2.2) получаем

$$\|x_{k+1} - x^*\|^2 \leq \|x_k - x^*\|^2 + h^2 - 2hb_k(x^*). \quad (2.3)$$

Если для всех  $k = 0, 1, 2, \dots b_k(x^*) \geq \frac{h}{2}(1 + \varepsilon)$ , то

$\|x_{k+1} - x^*\|^2 \leq \|x_k - x^*\|^2 - \varepsilon h^2 \leq \|x_0 - x^*\|^2 - \varepsilon(k+1)h^2$   
при  $k = 0, 1, 2, \dots$  Но  $\|x_{k+1} - x^*\|^2 \geq 0$ . Значит, найдется такое  $k^*$ , что

$$b_{k^*}(x^*) = \rho(x^*, U_{k^*}) \leq \frac{h}{2}(1 + \varepsilon).$$

Теорема доказана.

Рассмотрим некоторые следствия доказанной теоремы.

*Следствие 1.* При условиях теоремы 2.1 для произвольного  $\delta > 0$  можно найти такое  $h_\delta > 0$ , что при применении метода ОГС с шагом  $h_{k+1}(x_k) = \frac{h_\delta}{\|g_f(x_k)\|}$  при любом  $x_0 \in E_n$  найдется либо  $k^*$  такое, что  $x_{k^*} \in M^*$ , либо такая подпоследовательность  $k_1 < k_2 < k_3 < \dots$ , что

$$f(x_{k_i}) - f^* < \delta \quad (f^* = \min_{x \in E_n} f(x)).$$

*Следствие 2.* Если функция  $f(x)$  имеет область минимумов  $M^*$ , содержащую сферу радиуса  $r > \frac{h}{2} > 0$ , то при использовании метода ОГС с  $h_{k+1}(x_k) = \frac{h}{\|g_f(x_k)\|}$  найдется такое  $k^*$ , что  $x_{k^*} \in M^*$ .

## § 2. Основные теоремы о сходимости обобщенного градиентного спуска

Исследуем сходимость нескольких основных модификаций метода ОГС. Из доказательства теоремы 2.1 видно, что уменьшение расстояния до точек области минимумов на данном шаге гарантируется только в том случае, если мы находимся за пределами некоторой окрестности области минимумов, зависящей от шага  $h$ . Поэтому, чтобы получить обычные теоремы сходимости, нужно требовать, чтобы  $h_k \rightarrow 0$ , но не слишком быстро. В частности, если ряд  $\sum_{k=1}^{\infty} h_k$  будет

сходящимся, то последовательность  $\{x_k\}_{k=0}^{\infty}$  будет сходиться, но не обязательно к точке, принадлежащей  $M^*$ . Так мы приходим к ставшим уже классическими условиям  $h_k > 0$ ;  $h_k \rightarrow 0$ ;  $\sum_{k=1}^{\infty} h_k = +\infty$ .

Имеется несколько вариантов доказательства теоремы о сходимости метода ОГС [20, 38, 62]. Все они основаны на изучении поведения последовательности  $\{\rho_k\}_{k=0}^{\infty}$ , где  $\rho_k = \min_{y \in M^*} \|x_k - y\|$ . Наиболее общий результат (для случая выпуклых функций, определенных в гильбертовом пространстве, когда минимизация производится при наличии ограничений), получен Б. Т. Поляком [38]. Ю. М. Ермолев [20]

проводит доказательство теоремы о сходимости для конечномерного случая. Приведем принадлежащий автору вариант доказательства теоремы о сходимости ОГС.

**Теорема 2.2.** Пусть  $f(x)$  — выпуклая функция, определенная в  $E_n$ , область минимумов которой  $M^*$  ограничена, и задана последовательность положительных чисел  $\{h_k\}$ ,  $k = 1, 2, \dots$ , обладающая следующими свойствами:

$$\lim_{k \rightarrow \infty} h_k = 0; \quad \sum_{k=1}^{\infty} h_k = +\infty.$$

Тогда последовательность  $\{x_k\}$ ,  $k = 0, 1, \dots$ , образованная по формуле

$$x_{k+1} = x_k - h_{k+1} \frac{g_f(x_k)}{\|g_f(x_k)\|}, \quad (2.4)$$

при произвольном  $x_0 \in E_n$  обладает одним из следующих свойств: либо найдется такое  $k = \bar{k}$ , что  $x_{\bar{k}} \in M^*$ , либо  $\lim_{k \rightarrow \infty} \min_{y \in M^*} \|x_k - y\| = 0$  и  $\lim_{k \rightarrow \infty} f(x_k) = \min_{x \in E_n} f(x) = f^*$ .

**Доказательство.** Пусть  $x^* \in M^*$ . Как и при доказательстве теоремы 2.1, получаем при  $x_k \in M^*$

$$\|x_{k+1} - x^*\|^2 \leq \|x_k - x^*\|^2 + h_{k+1}^2 - 2h_{k+1}b_k(x^*).$$

Рассмотрим множество  $\{x : f(x) \leq f^* + a\}$  и его границу  $U_{f^*+a}$  ( $a > 0$ ). В условиях теоремы  $M^*$  — ограниченное замкнутое множество, как и  $U_{f^*+a}$ , причем  $M^* \cap U_{f^*+a} = \emptyset$ . В силу этого существует  $\min_{x \in U_{f^*+a}; y \in M^*} \|x - y\| = b(a) > 0$ . Так как  $\lim_{k \rightarrow \infty} h_k = 0$ , то найдется такое  $N_\varepsilon$ , что при  $k > N_\varepsilon$   $h_k < \varepsilon$ . Примем  $\varepsilon = b(a)$ . Для  $k > N_{b(a)}$  имеем  $h_k < b(a)$ . Если  $f(x_k) \geq f^* + a$ , то в силу (2.3)

$$\|x_k - x^*\|^2 - \|x_{k+1} - x^*\|^2 \geq b(a) h_{k+1}.$$

Так как  $\sum_k h_k = \infty$ , то найдется  $N_a^{(1)} > N_{b(a)}$  такое, что  $f(x_{N_a^{(1)}}) \leq f^* + a$ .

Введем обозначение

$$d(a) = \max_{y \in U_{f^*+a}} \min_{x \in M^*} \|y - x\|.$$

Если  $f(x_k) \leq f^* + a$ , то  $\min_{x \in M^*} \|x_k - x\| \leq d(a)$  и  $\min_{x \in M^*} \|x_{k+1} - x\| \leq d(a) + h_{k+1}$ . Если  $f(x_k) > f^* + a$  и  $k > N_{b(a)}$ , то

$$\min_{x \in M^*} \|x_{k+1} - x\| \leq \min_{x \in M^*} \|x_k - x\|.$$

Отсюда для  $k > N_a^{(1)}$

$$\min_{x \in M^*} \|x_k - x\| \leq d(a) + h(N_a^{(1)}),$$

где  $h(N_a^{(1)}) = \max_{k > N_a^{(1)}} h_k$ . Так как  $d(a) \rightarrow 0$  при  $a \rightarrow 0$ , то для произвольного  $\delta > 0$  можно выбрать такое  $a(\delta)$ , что  $d(a(\delta)) \leq \frac{\delta}{2}$ , а затем такой номер  $N_\delta$ , что  $f(x_{N_\delta}) \leq f^* + a(\delta)$  и для всех  $k > N_\delta$   $h_k \leq \frac{\delta}{2}$ . Тогда для всех  $k > N_\delta$

$$\min_{y \in M^*} \|x_k - y\| \leq \delta.$$

Это и доказывает соотношение  $\lim_{k \rightarrow \infty} \min_{y \in M^*} \|x_k - y\| = 0$ .

Из непрерывности функции  $f(x)$  следует, что

$$\lim_{k \rightarrow \infty} f(x_k) = f^*.$$

Аналогичная теорема доказана и для другого способа регулировки шага (без нормирования градиента).

**Теорема 2.3.** Если при условиях теоремы 2.2 построим последовательность  $\{x_k\}_{k=0}^\infty$  по формуле

$$x_{k+1} = x_k - h_{k+1} g_f(x), \quad k = 0, 1, \dots, \quad (2.5)$$

$x_0 \in E_n$  — некоторая начальная точка  $h_k > 0$ ,  $\lim_{k \rightarrow \infty} h_k = 0$ ,

$\sum_{k=1}^\infty h_k = +\infty$ , то возможны такие случаи:

а) последовательность  $\{g_f(x_k)\}_{k=0}^\infty$  неограничена, в этом случае сходимость отсутствует;

б) последовательность  $\{g_f(x_k)\}_{k=0}^\infty$  ограничена, тогда

$$\lim_{k \rightarrow \infty} \min_{x \in M^*} \|x_k - x\| = 0; \quad \lim_{k \rightarrow \infty} f(x_k) = f^*.$$

Доказательство. Приведем простой пример, когда реализуется случай а). Примем  $f(x) = x^4$ ;  $h_k = \frac{1}{k}$ ;

$x_0 = 1$ . Получим  $g_f(x_0) = 4$ ;  $x_1 = -3$ ;  $g_f(x_1) = -108$ ;  $x_2 = 51$  и т. д. Легко проверить, что  $\lim_{k \rightarrow \infty} |g_f(x_k)| = +\infty$ .

Рассмотрим случай б). Если для некоторого  $k^* g_f(x_{k^*}) = 0$ , то  $x_{k+1} = x_k$  для  $k \geq k^*$  и  $x_{k^*} \in M^*$ . Допустим, что  $g_f(x_k) \neq 0$ ,  $k = 0, 1, \dots$ . Рассмотрим  $\tilde{h}_k = h_k \|g_f(x_k)\|$ . Так как  $\{\|g_f(x_k)\|\}$  ограничена, то  $\lim_{k \rightarrow \infty} \tilde{h}_k = \lim_{k \rightarrow \infty} h_k = 0$ .

Если  $\sum_{k=0}^{\infty} \tilde{h}_k = +\infty$ , то, заменив (2.5) выражением  $x_{k+1} = x_k - \tilde{h}_{k+1} \frac{g_f(x_k)}{\|g_f(x_k)\|}$  и применив теорему 2.2, получаем требуемый результат.

Если  $\sum_{k=0}^{\infty} \tilde{h}_k < \infty$ , то последовательность  $\{x_k\}_{k=0}^{\infty}$  сходится к некоторой точке  $x^*$ . Покажем, что  $g_f(x^*) = 0$ . Предположим противное. Тогда  $x^* \notin M^*$ , а это значит, что, начиная с некоторого  $\bar{k}$ ,  $\|g_f(x_k)\| \geq \varepsilon > 0$ . Но тогда  $\sum_{k=\bar{k}}^{\infty} \tilde{h}_k \geq \varepsilon \sum_{k=\bar{k}}^{\infty} h_k = +\infty$  и мы пришли к противоречию. Таким образом,  $g_f(x^*) = 0$ , т. е.  $x^* \in M^*$ . Теорема доказана.

Заметим, что если  $f(x)$  — кусочно-линейная функция с конечным числом кусков, то  $\{g_f(x_k)\}_{k=0}^{\infty}$  всегда ограничена, т. е. выполняется случай б). Отметим также, что для заданного  $x_0$  всегда можно подобрать такое  $\delta > 0$ , что если  $\max_{k \geq 1} h_k \leq \delta$ , то будет выполняться случай б). В самом деле, рассмотрим для некоторого  $a > 0$   $U(a) = \{x : f(x) = f(x_0) + a\}$ ,  $S(a) = \{x : f(x) \leq f(x_0) + a\}$ . Так как  $f(x)$  определено на всем  $E_n$ ,  $M^*$  ограничено, то  $U(a)$ ,  $S(a)$  — компактные множества. Рассмотрим минимальный замкнутый шар  $S^*$  с центром в точка  $x^* \in M^*$ , содержащий  $S(a)$ , и пусть  $c = \max_{x \in S^*} \|g_f(x)\|$ ,  $\rho_1 = \rho\left(U(a), U\left(\frac{a}{2}\right)\right)$ ,  $\rho_2 = \rho\left(x^*, U\left(\frac{a}{2}\right)\right)$ .

Выберем  $\delta = \min\left\{\frac{\rho_1}{c}, \frac{\rho_2}{c}\right\}$ . Если  $h_k \leq \delta$ , то  $\tilde{h}_k = h_k \times \|g_f(x_k)\| \leq \min\{\rho_1, \rho_2\}$  при  $x_k \in S^*$ . Покажем, что  $x_k \in S^*$  при всех  $k = 0, 1, \dots$ . В самом деле,  $x_0 \in S^*$ . Если  $x_k \in S(a) \setminus S\left(\frac{a}{2}\right)$ , то из (2.3) получаем  $\rho(x_{k+1}, x^*) \leq \rho(x_k, x^*)$

и  $x_{k+1} \in S^*$ . Если  $x_k \in S\left(\frac{a}{2}\right)$ , то, так как  $\|x_{k+1} - x_k\| \leq \rho_1$ ,  $x_{k+1} \in S(a) \subseteq S^*$ . Таким образом, при всех  $k$   $x_k \in S^*$ . Но тогда  $\|g_f(x_k)\| \leq c$  и мы имеем случай б). Приведенные выше рассуждения позволяют модифицировать теорему 2.3 следующим образом.

**Теорема 2.4.** Если при условиях теоремы 2.2 образуем последовательность  $\{x_k\}_{k=0}^{\infty}$  по формуле

$$x_{k+1} = \begin{cases} x_k - h_{k+1}g_f(x_k) & \text{при } h_{k+1}\|g_f(x_k)\| \leq c, \\ x_0 & \text{в противном случае,} \end{cases}$$

где  $c > 0$  — некоторая константа,  $h_k \rightarrow 0$ ,  $\sum_{k=1}^{\infty} h_k = +\infty$ , то для любого начального приближения  $x_0$

$$\lim_{k \rightarrow \infty} \min_{x \in M^*} \|x_k - x\| = 0; \quad \lim_{k \rightarrow \infty} f(x_k) = f^*.$$

Для доказательства следует лишь отметить, что так как  $h_k \rightarrow 0$ , то, начиная с некоторого  $\bar{k}$ , условие  $h_{k+1}\|g_f(x_k)\| \leq c$  будет выполняться для всех последующих  $k$ .

**Теорема 2.5.** Если множество минимумов  $f(x)$   $M^*$  содержит сферу  $S_r$  радиуса  $r > 0$ ,  $h_k > 0$ ,  $\sum_{k=1}^{\infty} h_k = +\infty$  и  $\sup_k \operatorname{vrai} h_k < 2r$ , то для произвольного  $x_0 \in E_n$  при применении ОГС в форме (2.4) найдется  $k(x_0)$  такое, что  $x_{k(x_0)} \in M^*$ .

**Доказательство.** Пусть  $x^* \in M^*$  — центр сферы  $S_r$ . Тогда если  $x_k \notin M^*$  при  $k = 0, 1, \dots$ , то  $b_k(x^*) > r$ . Используя (2.3), получаем при  $k \geq m$

$$0 \leq \|x_{k+1} - x^*\|^2 \leq \|x_m - x^*\|^2 - \sum_{i=m}^k h_{i+1} (2b_i(x^*) - h_{i+1}), \quad (2.6)$$

Так как  $\sup_k \operatorname{vrai} h_k < 2r$ , то найдутся  $k'$  и  $\delta > 0$  такие, что  $2b_i(x^*) - h_{i+1} \geq 2r - h_{i+1} \geq \delta > 0$  при  $i > k'$ . Из (2.6) получаем

$$0 \leq \|x_{k'} - x^*\|^2 - \delta \sum_{i=k'}^{\infty} h_i.$$

Приходим к противоречию, так как по предположению  $\sum_{i=k'}^{\infty} h_i = +\infty$ . Таким образом, найдется  $x_{k(x_0)} \in M^*$ . Теорема доказана. Эта теорема имеет полезные приложения.

**Пример 1.** Решение системы выпуклых неравенств

$$f_i(x) \leq 0; \quad i = 1, \dots, m, \quad (2.7)$$

где  $f_i(x)$  — выпуклые функции, определенные на  $E_n$ .

Решение этой системы сводится к нахождению минимума функции  $\varphi(x)$ :

$$\varphi(x) = \max \left\{ \max_{1 \leq i \leq m} f_i(x), 0 \right\}.$$

Если область  $M^*$ , ограниченная системой (2.7), содержит внутренние точки, то выполняются условия применимости предыдущей теоремы. Тогда следующий итеративный алгоритм при любом начальном приближении  $x_0$  даст за конечное число шагов точку  $x_{k^*} \in M^*$ :

$$x_{k+1} = x_k - \frac{c}{k+1} \frac{g_{\varphi}(x_k)}{\|g_{\varphi}(x_k)\|},$$

где  $g_{\varphi}(x_k) = g_{j_i^*}(x_k)$ ,  $i^*$  — индекс функции, на которой достигается  $\max_{i \in \overline{1, m}} f_i(x_k)$ ,  $c > 0$ .

**Пример 2.** Задача о назначениях [95]. Найти  $x_{ij} \geq 0$ ,  $i, j = 1, \dots, n$ , удовлетворяющие следующим ограничениям:

$$\sum_{i=1}^n x_{ij} = 1; \quad \sum_{i=1}^n x_{ij} = 1,$$

при которых линейная форма  $\sum_{i=1}^n \sum_{j=1}^n c_{ij} x_{ij}$  принимает минимальное значение.

Двойственная задача состоит в следующем:

$$\max \left( \sum_{j=1}^n v_j - \sum_{i=1}^n u_i \right) \quad (2.8)$$

при ограничениях

$$v_j - u_i \leq c_{ij}, \quad i, j = 1, \dots, n. \quad (2.9)$$

Легко видеть, что в оптимальном плане задачи (2.8), (2.9)

$$v_j = \min_{j \leq i \leq n} [u_i + c_{ij}], \quad (2.10)$$

двойственная задача сводится к задаче максимизации конечно-линейной функции

$$\varphi(u) = \sum_{i=1}^n \min_{i \in \overline{1, n}} [u_i + c_{ij}] - \sum_{i=1}^n u_i.$$

Если задача о назначениях имеет единственное решение  $x_{ij(i)} = 1$  ( $i = 1, \dots, n$ );  $x_{ij} = 0$  при  $j \neq j(i)$ , где  $j(i)$  — некоторая перестановка индексов  $\{1, 2, \dots, n\}$ , то оно не изменится при достаточно малых вариациях  $\{c_{ij}\}_{i,j=1}^n$ . Это значит, что можно подобрать такие оптимальные  $u^* = \{u_i^*\}_{i=1}^n$ , что минимум в формуле (2.10) будет достигаться только при  $i = j(i)$ . Но тогда в некоторой окрестности  $u^*$   $\varphi(u)$  будет принимать постоянное значение:  $\varphi(u) = \sum_{i=1}^n c_{ij(i)}$ . Таким образом, если задача о назначениях имеет единственное решение, то множество оптимальных решений двойственной задачи содержит  $n$ -мерную сферу. Из теоремы 2.5 следует, что в этом случае метод ОГС, примененный к двойственной задаче, будет сходиться за конечное число шагов. Получаем простой алгоритм решения двойственной задачи

$$u_{k+1} = u_k + \frac{1}{k+1} g_{\varphi}(u_k),$$

где  $g_{\varphi}(u_k) = \{ |v_i| - 1 \}_{i=1}^n$ ,  $v_i = \{ j : \min_k [u_k + c_{kj}] = u_i + c_{ij} \}$ .

К сожалению, в случае неоднозначности решения задачи о назначениях, ОГС для двойственной задачи, вообще говоря, не сходится к оптимальному решению за конечное число шагов. Поэтому приведенный выше результат нельзя обобщить на задачи, двойственные к транспортной задаче, частным случаем которой является задача о назначениях. Дело в том, что произвольную линейную транспортную задачу можно свести с заданной точностью к задачам о назначениях, однако последние имеют, как правило, неоднозначное решение.

Б. Т. Поляк доказал теорему, аналогичную теореме 2.2, соответствующую методу проекции обобщенного градиента, для случая минимизации выпуклых функционалов, определенных в гильбертовом пространстве, при ограничениях довольно общего вида.

Пусть  $f(x)$  — выпуклый функционал, определенный на выпуклом множестве  $Q$  гильбертова пространства  $E$ ,  $Q = Q_1 \cap Q_2$ , где  $Q_1 = \{x : \varphi(x) \leq 0\}$ ,  $\varphi(x)$  — выпуклый на  $E$  функционал,  $Q_1$  имеет внутренние точки (множество внутренних точек обозначим  $Q_1^{(0)}$ ),  $Q_2$  — выпуклое замкнутое множество, причем  $Q_1^{(0)} \cap Q_2 \neq \emptyset$ . Обозначим через  $P_{Q_2}(x)$  проекцию точки  $x$  на  $Q_2$ , т. е.

$$\|x - P_{Q_2}(x)\| = \inf_{y \in Q_2} \|x - y\|.$$

Метод обобщенного градиентного спуска будет заключаться в построении  $\{x_k\}_{k=1}^{\infty}$  по формуле

$$x_{k+1} = P_{Q_2}(x_k + \Delta(x_k)), \quad (2.11)$$

где  $\Delta(x_k)$  — произвольный опорный функционал к множеству  $\{x : f(x) \leq f(x_k)\}$ , если  $x_k \in Q_1$ , и к множеству  $\{x : \varphi(x) \leq \varphi(x_k)\}$ , если  $x_k \in Q_1$ . (Линейный функционал называется опорным к  $R \subset E$  в точке  $x \in R$ , если  $(c, x) \leq (c, y)$  для всех  $y \in R$ .)

**Теорема 2.6.** Если  $\lim_{k \rightarrow \infty} \Delta(x_k) = 0$ ,  $\sum_{k=0}^{\infty} \|\Delta(x_k)\| = +\infty$ , то при любом  $x_0$  в методе (2.11) найдется подпоследовательность  $\{x_{n_k}\}$  такая, что  $x_{n_k} \in Q$ ,  $\lim_{k \rightarrow \infty} f(x_{n_k}) = f^* = \inf_{x \in Q} f(x)$ .

*Доказательство.* Выберем произвольные  $\alpha > f^*$  и обозначим  $S_\alpha = \{x : x \in Q_1; f(x) \leq \alpha\}$ ,  $S_\alpha^{(0)}$  — множество внутренних точек  $S_\alpha$ . Существует  $\tilde{x} \in Q_2 \cap S_\alpha^{(0)}$ . Выберем  $\rho > 0$  так, чтобы  $x \in S_\alpha$  для всех  $\|x - \tilde{x}\| \leq \rho$ . Предположим, что  $x_k \notin S_\alpha$  для всех  $k$ . Тогда если  $x_k \in Q_1$ , то  $S_\alpha \subset Q_1 \subset \{x : \varphi(x) \leq \varphi(x_k)\}$ ; если  $x_k \in Q_2$ , то  $S_\alpha \subset \{x : f(x) \leq f(x_k)\}$ , так что во всех случаях  $(\Delta(x_k), x_k) \leq \left(\Delta(x_k), \tilde{x} - \frac{\rho \Delta(x_k)}{\|\Delta(x_k)\|}\right)$ .  
Отсюда

$$\begin{aligned} \|x_{k+1} - \tilde{x}\|^2 &= \|P_{Q_2}(x_k + \Delta(x_k)) - \tilde{x}\|^2 \leq \|x_k + \Delta(x_k) - \tilde{x}\|^2 = \\ &= \|x_k - \tilde{x}\|^2 + \|\Delta(x_k)\|^2 + 2\left(\Delta(x_k), x_k - \tilde{x} + \rho \frac{\Delta(x_k)}{\|\Delta(x_k)\|}\right) - \\ &\quad - 2\rho \|\Delta(x_k)\| \leq \|x_k - \tilde{x}\|^2 + \|\Delta(x_k)\|^2 - 2\rho \|\Delta(x_k)\|. \end{aligned}$$

Выберем  $N$  такое, что  $\|\Delta(x_k)\| \leq \rho$  для всех  $k \geq N$  и просуммируем полученные выше неравенства от  $k = N$  до  $k =$

$= N + m$ . Получим

$$0 \leq \|x_{N+m+1} - \tilde{x}\|^2 \leq \|x_k - \tilde{x}\|^2 + \sum_{k=N}^{N+m} \|\Delta(x_k)\| (\|\Delta(x_k)\| - 2\rho) \leq \|x_N - \tilde{x}\|^2 - \rho \sum_{k=N}^{N+m} \|\Delta(x_k)\|, \quad m = 1, 2, \dots,$$

что противоречит расходимости  $\sum_{k=1}^{\infty} \|\Delta(x_k)\|$ . Значит, для любого  $\alpha > f^*$  найдется  $x_k \in S_\alpha$ , т. е. существует подпоследовательность  $\{x_{n_k}\}_{k=1}^{\infty}$  такая, что  $\lim_{k \rightarrow \infty} f(x_{n_k}) = f^*$ . Теорема доказана.

### § 3. Случай сходимости ОГС со скоростью геометрической прогрессии

При определенных дополнительных предположениях удастся получить варианты ОГС, сходящиеся со скоростью геометрической прогрессии.

**Теорема 2.7.** Пусть  $f(x)$  — выпуклая функция, определенная на  $E_n$ , и для всех  $x \in E_n$  при некотором  $\varphi$ ,  $0 \leq \varphi < \frac{\pi}{2}$ , выполняется неравенство

$$(g_f(x), x - x^*(x)) \geq \cos \varphi \|g_f(x)\| \|x - x^*(x)\|, \quad (2.12)$$

где  $x^*(x)$  — точка, принадлежащая множеству минимумов функции  $f(x)$  и лежащая на кратчайшем расстоянии от  $x$ . Тогда при заданном  $x_0$ , если мы выберем величину  $h_1$ , удовлетворяющую неравенству

$$h_1 \geq \left\{ \begin{array}{l} \|x^*(x_0) - x_0\| \cos \varphi, \quad \frac{\pi}{4} \leq \varphi < \frac{\pi}{2}, \end{array} \right. \quad (2.13)$$

$$\left\{ \begin{array}{l} \frac{\|x^*(x_0) - x_0\|}{2 \cos \varphi}, \quad 0 \leq \varphi \leq \frac{\pi}{4}, \end{array} \right. \quad (2.14)$$

определим  $\{h_k\}_{k=1}^{\infty}$  в соответствии с рекуррентной формулой

$$h_{k+1} = h_k r(\varphi), \quad k = 1, 2, \dots, \quad (2.15)$$

где

$$r(\varphi) = \begin{cases} \sin \varphi, & \frac{\pi}{4} \leq \varphi < \frac{\pi}{2}, \end{cases} \quad (2.16)$$

$$\begin{cases} \frac{1}{2 \cos \varphi}, & 0 \leq \varphi < \frac{\pi}{4}, \end{cases} \quad (2.17)$$

и вычислим  $\{x_k\}_{k=1}^{\infty}$  по формуле

$$x_{k+1} = x_k - h_{k+1} \frac{g_f(x_k)}{\|g_f(x_k)\|},$$

то либо при некотором  $k^*$   $q_f(x_{k^*}) = 0$  и принадлежит области минимумов, либо при всех  $k = 0, 1, 2, \dots$  будет выполняться неравенство

$$\|x_k - x^*(x_k)\| \leq \begin{cases} \frac{h_{k+1}}{\cos \varphi}, & \frac{\pi}{4} \leq \varphi < \frac{\pi}{2}, \end{cases} \quad (2.18)$$

$$\left\{ \begin{array}{l} 2 \cos \varphi \cdot h_{k+1}, \\ 0 \leq \varphi \leq \frac{\pi}{4}. \end{array} \right. \quad (2.19)$$

Доказательство. Рассмотрим сначала случай  $\frac{\pi}{4} \leq$

$\varphi < \frac{\pi}{2}$ . Для  $k = 0$  неравенство (2.18) выполняется.

Пусть оно выполняется для  $k = \rho$  и  $g_f(x_\rho) \neq 0$ . Докажем его выполнение для  $k = \rho + 1$ :

$$\begin{aligned} \|x_{\rho+1} - x^*(x_{\rho+1})\|^2 &\leq \|x_{\rho+1} - x^*(x_\rho)\|^2 = \\ &= \left\| x_\rho - h_{\rho+1} \frac{g_f(x_\rho)}{\|g_f(x_\rho)\|} - x^*(x_\rho) \right\|^2 = \\ &= \|x_\rho - x^*(x_\rho)\|^2 - \frac{2h_{\rho+1}}{\|g_f(x_\rho)\|} (x_\rho - x^*(x_\rho), g_f(x_\rho)) + h_{\rho+1}^2. \end{aligned}$$

Воспользовавшись последовательно неравенствами (2.12), (2.18) и учитывая, что при  $\varphi$ ,  $\frac{\pi}{4} \leq \varphi < \frac{\pi}{2}$ ,  $1 - 2 \cos^2 \varphi \geq 0$ , получаем

$$\begin{aligned} \|x_{\rho+1} - x^*(x_{\rho+1})\|^2 &\leq \|x_\rho - x^*(x_\rho)\|^2 (1 - 2 \cos^2 \varphi) + h_{\rho+1}^2 \leq \\ &\leq \frac{h_{\rho+1}^2}{\cos^2 \varphi} (1 - 2 \cos^2 \varphi) + h_{\rho+1}^2 = h_{\rho+1}^2 \left( \frac{\sin \varphi}{\cos \varphi} \right)^2. \end{aligned}$$

Принимая во внимание (2.15), (2.16), получаем

$$\|x_{\rho+1} - x^*(x_{\rho+1})\|^2 \leq \frac{h_{\rho+2}^2}{\cos^2 \varphi}.$$

Таким образом, для  $\frac{\pi}{4} \leq \varphi < \frac{\pi}{2}$  теорема справедлива.

Рассмотрим случай  $0 \leq \varphi \leq \frac{\pi}{4}$ . Для  $k = 0$  (2.19) выполняется. Предположим, что оно выполняется для  $k =$

$= p$ ,  $g_f(x_p) \neq 0$ , и докажем его выполнение для  $k = p + 1$ .

Воспользовавшись (2.12), (2.19), (2.17), получаем

$$\begin{aligned} \|x_{p+1} - x^*(x_{p+1})\|^2 &\leq \|x_p - x^*(x_p)\|^2 - 2h_{p+1} \cos \varphi \times \\ &\times \|x_p - x^*(x_p)\| + h_{p+1}^2 \leq \|x_p - x^*(x_p)\|^2 - \\ &- \|x_p - x^*(x_p)\|^2 + h_{p+1}^2 = h_{p+1}^2 = (2 \cos \varphi \cdot h_{p+2})^2. \end{aligned}$$

Теорема доказана.

**З а м е ч а н и е.** Из доказательства очевидно, что для справедливости теоремы достаточно выполнения условия (2.12) лишь для точек последовательности  $\{x_k\}_{k=0}^{\infty}$ .

Таким образом, если угол  $\varphi$  заранее известен, то, регулируя шаг по формулам (2.16) и (2.17), можем получить сходимость к минимуму со скоростью геометрической прогрессии со знаменателем  $q = r(\varphi)$ .

В формуле (2.12)  $\cos \varphi$  характеризует степень вытянутости поверхностей уровня функции  $f(x)$ . Если в некоторой окрестности минимума функции  $f(x)$  не существует  $\varphi < \frac{\pi}{2}$  такого, что для любого  $x$  из этой окрестности выполняется (2.12), то такую функцию будем называть существенно овражной. При минимизации существенно овражных функций приведенный в теореме способ регуляровки шаговых множителей не применим. В этом случае нужно использовать универсальный способ выбора шаговых множителей, как и в теореме 2.2.

Сформулируем теорему, аналогичную теореме 2.7, непосредственно в терминах, характеризующих степень «вытянутости» поверхностей уровня.

**Теорема 2.8.** Пусть выпуклая функция  $f(x)$  определена на  $E_n$ ,  $x^*$  — единственная точка минимума. Заданы начальное приближение  $x_0$  и числа  $\sigma$  и  $h_1$ , причем  $\sigma \geq \sqrt{2}$ ,  $h_1 \geq \frac{\|x_0 - x^*\|}{\sigma}$ . Рассмотрим множество  $Y = \{y : \|y - x^*\| \leq \sigma h_1\}$ . Если для любой пары точек  $x, z \in Y$  такой, что  $f(x) = f(z) \neq f(x^*)$ , выполняется условие

$$\|x - x^*\| \leq \sigma \|z - x^*\|, \quad (2.20)$$

то последовательность  $\{x_k\}_{k=0}^{\infty}$ , образованная с помощью рекуррентных формул

$$x_{k+1} = x_k - h_{k+1} \frac{g_f(x_k)}{\|g_f(x_k)\|},$$

причем  $h_{k+1} = h_k \frac{\sqrt{\sigma^2 - 1}}{\sigma}$ , сходится к  $x^*$  со скоростью геометрической прогрессии:

$$\|x_k - x^*\| \leq h_{k+1} \sigma, \quad (2.21)$$

за исключением случая, когда для некоторого  $k = \bar{k}$   $g_f(x_{\bar{k}}) = 0$ , т. е.  $x_{\bar{k}} = x^*$ .

**Доказательство.** При  $k = 0$  неравенство (2.21) выполняется,  $x_0 \in Y$ . Допустим, что (2.21) справедливо для  $k = p$ , и докажем его справедливость для  $k = p + 1$ . Предположим, что  $x_p \neq x^*$ , и, так как  $x_p \in Y$ , выполняется соотношение (2.20). Далее,

$$(g_f(x_k), x_k - x^*) = a_k(x^*) \|g_f(x_k)\|,$$

где  $a_k(x^*)$  — проекция  $x_k - x^*$  на  $g_f(x_k)$ . Из доказательства теоремы 2.1 следует, что  $a_k(x^*) \geq b_k(x^*)$ , где  $b_k(x^*)$  — кратчайшее расстояние от  $x^*$  до гиперповерхности уровня  $f(x) = f(x_k)$ . Таким образом,  $(g_f(x_k), x_k - x^*) \geq b_k(x^*) \times \|g_f(x_k)\|$ , а из (2.20) вытекает  $b_k(x^*) \geq \frac{\|x_k - x^*\|}{\sigma}$ . Значит,

$$(g_f(x_k), x_k - x^*) \geq \frac{1}{\sigma} \|g_f(x_k)\| \|x_k - x^*\|. \quad (2.22)$$

Введем обозначение  $\cos \varphi = 1/\sigma$ . Тогда условие (2.22) приобретает вид условия (2.12) из теоремы 2.7. Применим эту теорему. Учитывая, что при  $\sigma \geq \sqrt{2}$   $\varphi \geq \frac{\pi}{4}$ ,  $\sin \varphi = \frac{\sqrt{\sigma^2 - 1}}{\sigma}$ , получаем  $\|x_k - x^*\| \leq h_{k+1} \sigma$ , что и требовалось доказать.

**Пример.** Рассмотрим частный случай — минимизацию положительно определенной квадратичной формы  $f(x) = (Ax, x)$ . Пусть  $\lambda$  — наименьшее собственное число оператора  $A$ ,  $\mu$  — максимальное собственное число. Определим минимальное значение величины

$$\frac{(g_f(x), x)}{\|g_f(x)\| \|x\|} = \frac{(Ax, x)}{\|Ax\| \|x\|}, \quad \|x\| \neq 0.$$

Для этого зафиксируем произвольный несобственный вектор  $x_0$  и рассмотрим подпространство  $E(x_0)$  размерности два, порожденное векторами  $x_0$  и  $Ax_0$ . Обозначим через  $\tilde{A}$  матрицу квадратичной формы, определенной на  $E(x_0)$  и такой, что  $(\tilde{A}x, x) = (Ax, x)$  при  $x \in E(x_0)$ . Легко видеть, что опера-

тор  $\tilde{A}$ , определенный на подпространстве  $E(x_0)$ , является положительно определенным оператором и его собственные числа  $\tilde{\lambda}$ ,  $\tilde{\mu}$  (записанные в неубывающем порядке) удовлетворяют неравенствам  $\lambda \leq \tilde{\lambda} \leq \tilde{\mu} \leq \mu$ . Пусть  $x_1, x_2$  — ортонормированная система собственных векторов оператора  $\tilde{A}$  в  $E(x_0)$  и  $\|x_0\| = 1$ . Представим  $x_0$  в виде  $x_0 = \alpha_1 x_1 + \alpha_2 x_2$ , причем  $\alpha_1^2 + \alpha_2^2 = 1$ . Тогда

$$\frac{(Ax_0, x_0)}{\|Ax_0\| \|x_0\|} = \frac{(\tilde{A}x_0, x_0)}{\|\tilde{A}x_0\| \|x_0\|} = \frac{\alpha_1^2 \tilde{\lambda} + \alpha_2^2 \tilde{\mu}}{\sqrt{\alpha_1^2 \tilde{\lambda}^2 + \alpha_2^2 \tilde{\mu}^2}}. \quad (2.23)$$

Минимизируя правую часть (2.23) при условии  $\alpha_1^2 + \alpha_2^2 = 1$ , получаем неравенство

$$\frac{(Ax_0, x_0)}{\|Ax_0\| \|x_0\|} \geq \frac{2\tilde{\lambda}\tilde{\mu}}{\tilde{\lambda} + \tilde{\mu}},$$

которое превращается в равенство при  $\alpha_1^2 = \frac{\tilde{\mu}}{\tilde{\lambda} + \tilde{\mu}}$ ;  $\alpha_2^2 = \frac{\tilde{\lambda}}{\tilde{\lambda} + \tilde{\mu}}$ . Отсюда сразу получается, что

$$\min_{x \in E, x \neq 0} \frac{(Ax, x)}{\|Ax\| \|x\|} = \frac{2\sqrt{\tilde{\lambda}\tilde{\mu}}}{\tilde{\lambda} + \tilde{\mu}}$$

при

$$x = \sqrt{\frac{\mu}{\tilde{\lambda} + \tilde{\mu}}} s_1 + \sqrt{\frac{\tilde{\lambda}}{\tilde{\lambda} + \tilde{\mu}}} s_2,$$

где  $s_1, s_2$  — собственные нормированные векторы оператора  $A$ , соответствующие минимальному и максимальному собственным числам.

Обратимся теперь к теореме 2.7. Для нашего примера можем принять  $\cos \varphi = \frac{2\sqrt{\tilde{\lambda}\tilde{\mu}}}{\tilde{\lambda} + \tilde{\mu}}$ , при этом  $\sin \varphi = \frac{\mu - \tilde{\lambda}}{\mu + \tilde{\lambda}}$  и обобщенный градиентный спуск, описанный в условии этой теоремы, дает процесс, сходящийся к точке минимума по норме со скоростью геометрической прогрессии со знаменателем (при  $\varphi \geq \frac{\pi}{4}$ )

$$q = \frac{\mu - \tilde{\lambda}}{\mu + \tilde{\lambda}} = \frac{\rho - 1}{\rho + 1},$$

где  $\rho$  — число обусловленности  $\left(\rho = \frac{\mu}{\lambda}\right)$  [50], а по функционалу — со знаменателем  $q^2$ .

**З а м е ч а н и е.** Наши рассуждения почти полностью сохраняют силу, если  $A$  — неотрицательно определенный оператор,  $A \neq 0$ , только под  $\lambda$  нужно понимать наименьшее положительное собственное число.

Сопоставим полученный выше результат с результатом Л. В. Канторовича [28] о скорости сходимости метода наискорейшего спуска для квадратичных функционалов. Этот метод сходится по функционалу со скоростью геометрической прогрессии со знаменателем  $q^2$ . Однако схема метода наискорейшего спуска для квадратичных функционалов (или для решения систем линейных уравнений с симметричной положительно определенной матрицей) гораздо сложнее, чем схема, описанная в условиях теоремы, требует  $3n$  рабочих ячеек, т. е. в три раза больше, чем последняя схема, что имеет значение для задач большого объема (см. [50, § 70]).

Таким образом, если мы имеем возможность получить достаточно точную оценку сверху  $\hat{\rho}$  для числа обусловленности  $\rho$  и примем  $\sin \varphi = (\hat{\rho} - 1)/(\hat{\rho} + 1)$ , то алгоритм, описанный в теореме 2.7, может оказаться практически вполне конкурентоспособным по сравнению с методом наискорейшего спуска.

Заметим, что для квадратичных функционалов  $f(x) = (Ax, x)$  величина  $\sigma$ , фигурирующая в теореме 2.8, может быть взята равной  $\sqrt{\rho}$ , при этом получаем сходимость со скоростью геометрической прогрессии со знаменателем

$$\bar{q} = \frac{\sqrt{\sigma^2 - 1}}{\sigma} = \sqrt{\frac{\rho - 1}{\rho}}.$$

Рассмотрим еще один вариант метода обобщенного градиентного спуска, когда шаговый множитель остается в течение определенного числа шагов постоянным, а затем уменьшается в два раза [61].

**Теорема 2.9.** Пусть для выпуклой функции  $f(x)$  выполняются условия теоремы 2.8,  $\sigma \geq 2$ . Рассмотрим при заданном  $x_0$  следующий итеративный процесс:

$$x_{k+1} = x_k - h_{k+1} \frac{g_f(x_k)}{\|g_f(x_k)\|},$$

где  $h_{k+1} = h_0 \cdot 2^{-\left[\frac{k+1}{N}\right]}$ . При достаточно большом  $h_0$  и  $N \geq 3\sigma^2 + 1$

$$\|x_k - x^*\| \leq 2\sigma h_{k+1}, \quad k = 0, 1, 2, \dots \quad (2.24)$$

**Доказательство.** Пусть  $h_0 \geq \frac{\|x_0 - x^*\|}{2\sigma}$ . Так же как при доказательстве теоремы 2.2, получаем

$$\|x_{k+1} - x^*\|^2 \leq \|x_k - x^*\|^2 + h_{k+1}^2 - 2h_{k+1}b_k(x^*).$$

Если  $b_k(x^*) \geq \frac{h_{k+1}}{2}$ , то

$$\|x_{k+1} - x^*\| \leq \|x_k - x^*\|. \quad (2.25)$$

Если  $b_k(x^*) \geq h_{k+1}$ , то

$$\|x_{k+1} - x^*\|^2 \leq \|x_k - x^*\|^2 - h_{k+1}^2. \quad (2.26)$$

Так как  $\|x_0 - x^*\|^2 \leq 4\sigma^2 h_0^2$ , то, принимая  $h_i = h_0$  ( $i = 1, \dots, [3\sigma^2] + 1$ ), получим, если на всех шагах будет выполняться (2.26),

$$\|x_N - x^*\|^2 \leq 4\sigma^2 h_0^2 - ([3\sigma^2] + 1) h_0^2 \leq \sigma^2 h_0^2.$$

Если на некотором шаге  $\bar{k} < N$  условие (2.26) не будет выполняться, то  $b_{\bar{k}}(x^*) \leq h_0$  и  $\|x_{\bar{k}} - x^*\| \leq \sigma h_0$ .

При этом могут возникнуть две ситуации:

а) на шагах  $k = \bar{k} + 1, \dots, N$  будет выполняться условие  $b_k(x^*) \geq \frac{h_{k+1}}{2}$ ; в этом случае в силу (2.25)  $\|x_N - x^*\| \leq \sigma h_0$ ;

б) на некотором шаге  $\bar{k}$ ,  $\bar{k} \leq \bar{k} \leq N$ , выполняется условие  $b_{\bar{k}}(x^*) \leq \frac{h_0}{2}$ , тогда  $\|x_{\bar{k}} - x^*\| \leq \frac{h_0\sigma}{2}$  и  $\|x_k - x^*\| \leq \frac{h_0\sigma}{2} + h_0 \leq h_0\sigma$  при  $\sigma \geq 2$  для  $\bar{k} \leq k \leq N$ . Таким образом, во всех случаях

$$\|x_N - x^*\| \leq \sigma h_0; \quad \|x_i - x^*\| \leq 2\sigma h_0; \quad i = 0, \dots, N.$$

Принимая  $x_0^{(1)} = x_N$  и  $h_0^{(1)} = \frac{h_0}{2}$ , получим неравенство

$$h_0^{(1)} \geq \frac{\|x_0^{(1)} - x^*\|}{2\sigma}, \text{ аналогичное (2.24), откуда следует}$$

$$\|x_i - x^*\| \leq \sigma h_0, \quad i = N + 1, \dots, 2N - 1;$$

$$\|x_{2N} - x^*\| \leq \frac{\sigma h_0}{2}$$

и далее по индукции

$$\|x_{rN+1} - x^*\| \leq 2\sigma h_0 2^{-r},$$

$$i = 0, \dots, N - 1, r = 1, 2, \dots,$$

что и требовалось доказать.

Рассмотрим вопрос о выборе параметров алгоритма варианта метода ОГС, содержащегося в условиях теоремы 2.7. Для простоты будем считать, что минимум  $f(x)$  достигается в единственной точке  $x^*$ . Для реализации соответствующего алгоритма необходимо задание двух параметров:  $h$  и  $\varphi$ . Если параметр  $\varphi$  выбран так, что справедливо неравенство (2.12) для всех  $x = x_k$ ,  $k = 0, 1, \dots$ , а  $h$  удовлетворяет неравенству (2.13), то теорема 2.7 гарантирует определенную точность решения после  $k$  шагов:

$$\Delta_k = \|x_k - x^*\| \leq \frac{h \sin^k \varphi}{\cos \varphi}. \quad (2.27)$$

Так как условие (2.12) в большинстве случаев эффективно не проверяется, то нужно иметь некоторый косвенный легко проверяемый критерий «правильности» выбора  $h$  и  $\varphi$ .

Возьмем новую начальную точку  $\bar{x}_0$  и  $\bar{\varphi}$ ,  $\varphi \leq \bar{\varphi} < \frac{\pi}{2}$ .

Выберем  $\bar{h} = h + \|x_0 - \bar{x}_0\| \cos \varphi$ . Легко видеть, что если параметры  $\varphi$ ,  $h$  удовлетворяли (2.12), (2.13) для точки  $x_0$ , то  $\bar{\varphi}$ ,  $\bar{h}$  удовлетворяют этим же условиям. Пусть  $\{\bar{x}_k\}$ ,  $k = 0, 1, \dots$ , — последовательность, получаемая при применении ОГС с начальной точкой  $\bar{x}_0$  и параметрами  $\bar{\varphi}$  и  $\bar{h}$ . Тогда в силу (2.27) должно выполняться неравенство

$$\|x_k - \bar{x}_k\| \leq \frac{h \sin^k \varphi}{\cos \varphi} + \frac{\bar{h} \sin^k \bar{\varphi}}{\cos \bar{\varphi}}, \quad k = 0, 1, \dots \quad (2.28)$$

Если для некоторого  $k$  неравенство не выполняется, это свидетельствует о неправильности выбора параметров  $h$  и  $\varphi$ . С другой стороны, если хотя бы одна из последовательностей  $\{x_k\}$  не сходится к  $x^*$ , то мало вероятно при  $\bar{x}_0 \neq x_0$  и  $\bar{\varphi} \neq \varphi$ , чтобы они сходились к одной и той же точке, если же эти последовательности сходятся к разным точкам, то при достаточно большом  $k$  неравенство (2.28) не будет выполняться.

Таким образом, неравенство (2.28) можно считать удобным критерием для оценки правильности выбора параметров  $h$  и  $\varphi$  при параллельном проведении двух процессов ОГС

из разных начальных точек. Практически этот критерий можно использовать следующим образом. Допустим, что нам необходимо получить  $x_{\bar{k}}$  такое, что

$$\|x_{\bar{k}} - x^*\| \leq \delta. \quad (2.29)$$

Из неравенства (2.27) получаем, что если  $h$  и  $\varphi$  выбраны верно, то для гарантии выполнения (2.29) достаточно, чтобы

$$\bar{k} \geq \left[ \frac{\ln \frac{\delta \cos \varphi}{h}}{\ln \sin \varphi} \right]. \quad (2.30)$$

Проведем процесс из двух начальных точек  $x_0$  и  $\bar{x}_0$ , как указано выше. Если мы проделали  $\bar{k}$  шагов, где  $\bar{k}$  удовлетворяет неравенству (2.30), и при этом для всех  $k \leq \bar{k}$  выполнялось неравенство (2.28), то можно утверждать, что  $\|x_{\bar{k}} - x^*\| \leq \delta$ . Для более полной уверенности можно произвести спуск из еще одной начальной точки или более тщательно исследовать окрестность точки  $x_{\bar{k}}$ . Если же для некоторого  $k \leq \bar{k}$  неравенство (2.28) не будет выполнено, то следует изменить параметры  $h$  и  $\varphi$  (например, изменить  $\varphi$  таким образом, чтобы  $\cos \varphi$  уменьшился в два раза) и продолжить счет из наилучшей предыдущей и некоторой контрольной точек.

#### § 4. Обобщенный градиентный спуск и фейеровские приближения

Установим полезные связи между ОГС и развитым И. И. Ереминым методом фейеровских приближений.

**Определение [19].** Отображение  $\varphi \in [E_n \rightarrow E_n]$  называется  $M$ -фейеровским ( $E_n \supset M \neq \emptyset$ ), если

$$\varphi(y) = y; \quad \|\varphi(x) - y\| < \|x - y\|, \quad \forall y \in M; \quad \forall x \in M \quad (2.31)$$

Используя  $M$ -фейеровские отображения, можно строить алгоритмы фейеровских приближений, сходящиеся к  $M$ , следующего вида:  $x_{k+1} = \varphi(x_k)$ .

**Теорема 2.10 [19].** Если отображение  $\varphi(x)$  непрерывно, то последовательность  $\{x_k\}$  сходится из любого начального приближения к некоторому элементу из  $M$ .

**Доказательство.** Зафиксируем  $y \in M$ . Из (2.31) следует, что  $\|x_k - y\| \leq \|x_0 - y\|$ , т. е. последовательность  $\{x_k\}_{k=0}^{\infty}$  ограничена. Пусть  $x^*$  — предельная

точка этой последовательности. Допустим, что  $x^* \notin M$ . Тогда  $\|\varphi(x^*) - y\| \leq \|x^* - y\|$ , и из непрерывности  $\varphi$  следует, что найдется такое достаточно большое  $\bar{k}$ , что  $\|\varphi(x_{\bar{k}-1}) - y\| = \|x_{\bar{k}} - y\| < \|x^* - y\|$ . С другой стороны,  $\|x^* - y\|$  является пределом монотонно убывающей последовательности  $\{\|x_k - y\|\}$ , т. е.  $\|x^* - y\| \leq \|x_k - y\|$  для всех  $k = 0, 1, \dots$ . Получим противоречие, откуда следует, что  $x^* \in M$ . Из монотонной сходимости по расстоянию следует, что существует единственная предельная точка  $x^*$ . Теорема доказана.

Б. Т. Поляком [39] предложены варианты ОГС, являющиеся по сути алгоритмами фейеровских приближений.

Пусть  $f(x)$  — выпуклая функция, определенная на  $E_n$ ,  $f^* = \inf f(x)$ ,  $c \geq f^*$  и  $M(c) = \{x \in E_n : f(x) \leq c\}$ . Построим итеративный процесс следующего вида:

$$x_{k+1} = x_k - \frac{\gamma[f(x_k) - c]}{\|g_f(x_k)\|^2} g_f(x_k) = \varphi_c(x_k). \quad (2.32)$$

**Теорема 2.11.** При  $0 < \gamma < 2$  и любом  $x_0 \in E_n$  либо при некотором  $k^*$   $x_{k^*} \in M(c)$  или  $\lim_{k \rightarrow \infty} x_k \in M(c)$ .

**Доказательство.** Пусть  $x_k \notin M(c)$ ,  $y \in M(c)$ . Тогда

$$\begin{aligned} \|x_{k+1} - y\|^2 &= \|x_k - y\|^2 - 2\gamma \frac{f(x_k) - c}{\|g_f(x_k)\|^2} (x_k - y, g_f(x_k)) + \\ &+ \gamma^2 \frac{[f(x_k) - c]^2}{\|g_f(x_k)\|^2}. \end{aligned}$$

Заметим, что

$$0 \leq f(x_k) - c \leq f(x_k) - f(y) \leq (g_f(x_k), x_k - y),$$

откуда

$$\|x_{k+1} - y\|^2 \leq \|x_k - y\|^2 - \gamma(2 - \gamma) \frac{[f(x_k) - c]^2}{\|g_f(x_k)\|^2} < \|x_k - y\|^2. \quad (2.33)$$

Если принять  $\varphi_c(y) = y$  для  $y \in M(c)$ , получим, что отображение  $\varphi_c(x)$  является фейеровским. Правда, непосредственно теоремой 2.10 мы воспользоваться не сможем, так как  $\varphi_c(x)$  не является обязательно непрерывным отображением (из-за возможной разрывности градиента  $g_f(x)$ ).

Заметим, однако, что следствием монотонного изменения  $\|x_k - y\|$  является ограниченность  $\{x_k\}_{k=0}^\infty$  и  $\{g_f(x_k)\}$ . Если предположить, что существует предельная точка последовательности  $\{x_k\}_{k=0}^\infty$  такая, что  $f(x^*) > c$ , то найдутся

$\delta > 0$  и бесконечная подпоследовательность индексов  $k_1 < k_2 < \dots$  такие, что  $f(x_{k_i}) - c > \delta$ , т. е.

$$\|x_{k_{i+1}} - y\| \leq \|x_{k_i} - y\|^2 - \varepsilon,$$

где  $\varepsilon > 0$ ,  $i = 1, 2, \dots$ . Отсюда  $\|x_{k_{i+1}} - y\|^2 \leq \|x_0 - y\|^2 - i\varepsilon$ . Приходим к противоречию. Таким образом,  $x^* \in M(c)$ . Единственность  $x^*$  вытекает из монотонной сходимости по расстоянию. Теорема доказана.

При некоторых довольно сильных предположениях метод (2.32) сходится со скоростью геометрической прогрессии.

**Теорема 2.12.** Пусть  $f(x)$  — сильно выпуклый функционал, причем  $f(x) - f^* \geq m\|x - x^*\|^2$ ,  $g_f(x_k)$  удовлетворяет условию Липшица с константой  $L$  в области  $\|x - x^*\| \leq \|x_0 - x^*\|$ ,  $c = f^*$ . Тогда для метода (2.32) выполняется

$$\|x_k - x^*\| \leq q^k \|x_0 - x^*\|, \text{ где } q = \left(1 - \frac{\gamma(2-\gamma)m^2}{L^2}\right)^{\frac{1}{2}} < 1.$$

**Доказательство.** Из условий теоремы получаем

$$f(x_k) - f(x^*) \geq m\|x_k - x^*\|^2, \quad (2.34)$$

$$\|g_f(x_k)\| \leq L\|x_k - x^*\|. \quad (2.35)$$

Из (2.33) — (2.35) имеем

$$\begin{aligned} \|x_{k+1} - x^*\|^2 &\leq \|x_k - x^*\|^2 - \frac{m^2\gamma(2-\gamma)}{L^2} \frac{\|x_k - x^*\|^4}{\|x_k - x^*\|^2} = \\ &= \|x_k - x^*\|^2 \left[1 - \gamma(2-\gamma) \frac{m^2}{L^2}\right], \end{aligned}$$

что и требовалось доказать.

В работе [39] рассмотрим еще один случай, при котором обеспечивается сходимость (2.32) со скоростью геометрической прогрессии.

**Теорема 2.13.** Пусть  $f(x)$  — выпуклая функция,  $x^*$  — ее точка минимума и существует  $m > 0$  такое, что  $f(x) - f(x^*) \geq m\|x - x^*\|$ ,  $L$  — константа Липшица  $f(x)$  в области  $\|x - x^*\| \leq \|x_0 - x^*\|$ ,  $c = f(x^*)$ . Тогда процесс (2.32) обладает свойством

$$\|x_{k+1} - x^*\|^2 \leq q\|x_k - x^*\|^2, \quad q = 1 - \gamma(2-\gamma) \frac{m^2}{L^2}.$$

Доказательство. Из (2.33) и условий теоремы имеем

$$\begin{aligned} \|x_{k+1} - x^*\|^2 &\leq \|x_k - x^*\|^2 - \frac{\gamma(2-\gamma)m^2}{L^2} \|x_k - x^*\|^2 = \\ &= \|x_k - x^*\|^2 q, \end{aligned}$$

откуда вытекает теорема.

Рассмотрим применение процессов вида (2.32) к решению совместной системы выпуклых неравенств

$$f_i(x) \leq 0, \quad i = 1, \dots, m. \quad (2.36)$$

Решение этой системы эквивалентно задаче минимизации

$$\psi^+(x) = \begin{cases} \psi(x), & \psi(x) > 0, & \psi(x) = \max_{i \in \overline{1, m}} f_i(x); \\ 0, & \psi(x) \leq 0. \end{cases}$$

В этом случае в качестве  $c$  можно взять нуль. Получаем следующий процесс:

$$x_{k+1} = \begin{cases} x_k, & \psi^+(x) = 0, \\ x_k - \frac{\gamma g_{f_{i_k}}(x_k) \psi^+(x)}{\|g_{f_{i_k}}(x_k)\|^2}; & \psi^+(x) = f_{i_k}(x_k) > 0. \end{cases} \quad (2.37)$$

Этот процесс сходится к решению из любого начального приближения.

Рассмотрим случай, когда система (2.36) несовместна, т. е.  $\min \psi^+(x) = d > 0$ . Для анализа этого случая заметим, что процесс вида (2.32) сходится и в случае, когда  $\gamma$  зависит от  $k$ , т. е. в форме

$$x_{k+1} = x_k - \frac{\gamma_k [f(x_k) - c]}{\|g_f(x_k)\|^2} g_f(x_k)$$

при условии  $\varepsilon \leq \gamma_k \leq 2 - \varepsilon$ ,  $\varepsilon > 0$ .

Применим процесс (2.32) с некоторым  $\gamma$  к функции  $\psi(x)$  при  $c = 0$ ,  $0 < \gamma < 2$ :

$$\begin{aligned} x_{k+1} &= x_k - \frac{\gamma \psi(x_k)}{\|g_\psi(x_k)\|^2} g_\psi(x_k) = \\ &= x_k - \frac{\gamma \psi(x_k)}{\psi(x_k) - d} [\psi(x_k) - d] \frac{g_\psi(x_k)}{\|g_\psi(x_k)\|^2}. \end{aligned}$$

Если бы  $\gamma \frac{\psi(x_k)}{\psi(x_k) - d} \leq 2 - \varepsilon$  для всех  $k$ , то мы получили бы сходящийся к минимуму  $\psi(x)$  процесс. Но это невоз-

можно, так как в этом случае  $\gamma \frac{\psi(x_k)}{\psi(x_k) - d} \rightarrow \infty$ . Значит, найдется  $k^*$  такое, что

$$\gamma \frac{\psi(x_{k^*})}{\psi(x_{k^*}) - d} > 2 - \varepsilon; \quad \psi(x_{k^*}) < \frac{d(2 - \varepsilon)}{2 - \varepsilon - \gamma}.$$

Так как  $\varepsilon$  может быть взято сколь угодно малым, то справедливо следующее утверждение.

**Теорема 2.14.** Пусть  $\psi(x)$  — выпуклая функция, определенная на  $E_n$ ,  $\min_{x \in E_n} \psi(x) = d > 0$ . Рассмотрим последовательность  $\{x_k\}$ , образуемую по формуле

$$x_{k+1} = x_k - \frac{\gamma \psi(x_k)}{\|g_\psi(x_k)\|^2} g_\psi(x_k)$$

при некотором начальном приближении  $x_0$ . Тогда

$$\lim_{k \rightarrow \infty} \min_{0 \leq i \leq k} \psi(x_i) \leq \frac{2d}{2 - \gamma}.$$

Теоремы 2.11 и 2.14 указывают на принципиальную возможность построения процесса нахождения минимума выпуклой функции  $f(x)$  при неизвестном  $f^* = \min_{x \in E_n} f(x)$

путем использования оценок  $f^*$  и их последовательного уточнения на основе поведения последовательности  $\{f(x_k)\}_{k=0}^\infty$ , где

$$x_{k+1} = x_k - \frac{\gamma [f(x_k) - \bar{f}]}{\|g_f(x_k)\|^2} g_f(x_k), \quad 0 < \gamma < 2,$$

$\bar{f}$  — оценка  $f^*$ . Если  $f(x_k)$  при  $k = 0, 1, 2, \dots$  удовлетворяет соотношению  $f(x_k) - \bar{f} \geq d > 0$ , то  $\bar{f} < f^*$  и оценку  $\bar{f}$  нужно увеличивать. Однако этот признак практически использовать невозможно, так как нет алгоритмически обоснованных способов определения по конечному отрезку последовательности, сходится ли данная последовательность к некоторой константе или нет. Здесь могут быть указаны лишь эвристические правила, имеющие ограниченную сферу применения.

Группа итальянских ученых предложила модификацию ОГС, в которой используются информация о градиенте на предыдущем шаге и оценка значения функции в точке минимума [83]. Хотя в этой работе речь идет о минимизации кусочно-линейной выпуклой функции, предложенный ал-

горитм может быть применен практически без изменений для минимизации произвольной выпуклой функции.

Основная идея метода состоит в том, чтобы строить направление спуска как линейную комбинацию направления субградиента, вычисленного в данной точке, и направления движения на предыдущем шаге (по типу метода сопряженных градиентов).

Пусть  $f(x)$  — выпуклая функция,  $x \in E_n$ ;  $f^* = \min_{x \in E_n} f(x) = f(x^*)$ . Рассмотрим следующую процедуру:

$$x_{k+1} = x_k - h_k s_k; \quad h_k = \frac{[f(x_k) - f^*] \gamma_k}{\|s_k\|^2}, \quad (2.38)$$

$$s_0 = g_f(x_0); \quad s_k = g_f(x_k) + \beta_k s_{k-1}; \quad k = 1, 2, \dots$$

Приведем доказательство следующей теоремы.

**Теорема 2.15.** Если для всех  $k = 0, 1, \dots$   $0 < \gamma_k \leq 1$  и  $\beta_k \geq 0$ , то  $(x_k - x^*, s_k) \geq (x_k - x^*, g_f(x_k))$  для всех  $k$ .

**Доказательство.** Доказательство проведем методом индукции по  $k$ . Допустим, что теорема справедлива для  $k = m$ . Докажем ее для  $k = m + 1$ :

$$\begin{aligned} (x_{m+1} - x^*, s_{m+1}) &= (x_m - h_m s_m - x^*, g_f(x_{m+1}) + \beta_m s_m) \geq \\ &\geq (x_{m+1} - x^*, g_f(x_{m+1})) + \beta_m [(x_m - x^*, s_m) - (f(x_m) - f^*)]. \end{aligned}$$

Из предположения индукции следует, что

$$(x_m - x^*, s_m) \geq (x_m - x^*, g_f(x_m)) \geq f(x_m) - f^*.$$

Таким образом, учитывая, что  $\beta_m \geq 0$ , получаем

$$(x_{m+1} - x^*, s_{m+1}) \geq (x_{m+1} - x^*, g_f(x_{m+1})),$$

откуда и следует справедливость теоремы.

При специальном выборе  $\beta_k$  можно получить более сильный результат.

**Теорема 2.16.** Пусть при условиях предыдущей теоремы

$$\beta_k = \begin{cases} -\alpha_k \frac{(s_{k-1}, g_f(x_k))}{\|s_{k-1}\|^2}, & \text{если } (s_{k-1}, g_f(x_k)) < 0; \\ 0 & \text{в противном случае, } 0 \leq \alpha_k \leq 2. \end{cases}$$

Тогда

$$\frac{(x_k - x^*, s_k)}{\|s_k\|} \geq \frac{(x_k - x^*, g_f(x_k))}{\|g_f(x_k)\|}. \quad (2.39)$$

Доказательство. Если  $\beta_k = 0$ , то  $s_k = g_f(x_k)$  и (2.39) выполняется. Рассмотрим случай

$$(s_{k-1}, g_f(x_k)) < 0; \quad \beta_k = \frac{-\alpha_k (s_{k-1}, g_f(x_k))}{\|s_{k-1}\|^2}.$$

Тогда

$$\begin{aligned} \|s_k\|^2 - \|g_f(x_k)\|^2 &= \beta_k^2 \|s_{k-1}\|^2 + 2\beta_k (s_{k-1}, g_f(x_k)) = \\ &= \frac{(s_{k-1}, g_f(x_k))^2}{\|s_{k-1}\|^2} (\alpha_k^2 - 2\alpha_k) \leq 0. \end{aligned}$$

Отсюда  $\|s_k\| \leq \|g_f(x_k)\|$ . Используя теорему 2.15, получаем

$$\frac{(s_k, x_k - x^*)}{\|s_k\|} \geq \frac{(g_f(x_k), x_k - x^*)}{\|g_f(x_k)\|},$$

что и требовалось доказать.

Из доказательства теоремы видно, что направление  $s_k$ , выбираемое по формуле (2.38), образует при определенных условиях меньший угол с направлением на точку минимума по сравнению с антисубградиентом, что приводит к более быстрой сходимости процесса спуска при заданном  $f^*$ . В работе [83] рекомендуется выбирать  $\alpha_k \sim 1,5$  и приводятся результаты численных экспериментов, которые показывают преимущества алгоритма в форме (2.38) по сравнению с алгоритмом ОГС в форме (2.32).

## § 5. $\varepsilon$ -субградиентные методы

В последние годы появилось несколько работ, в которых предложены различные варианты релаксационных процессов негладкой оптимизации. Наиболее известные из них — работы Лемарешаля и Вулфа по разработке так называемого метода сопряженных субградиентов для минимизации произвольных выпуклых функций [99, 106]. Хотя формализм этого метода несколько напоминает метод сопряженных градиентов, по сути он никак не связан с квадратичной аппроксимацией и является в определенном смысле обобщением релаксационных алгоритмов В. Ф. Демьянова, разработанных для минимизации функции максимума [14]. Поиск подходящего направления спуска строится в алгоритме сопряженных субградиентов на основе информации о значениях субградиентов, взятых в точках, лежащих в некоторой окрестности данной точки. Поэтому этот метод

можно отнести к классу  $\varepsilon$ -субградиентных методов. Опишем этот метод, следуя [98].

Пусть задана выпуклая функция  $f(x)$ ,  $x \in E_n$ , и точка  $x_0 \in E_n$ . Выберем субградиент  $g(x_0) \in G_f(x_0)$  и направление  $s_0 = -g(x_0)$ , не обязательно совпадающее с направлением уменьшения функции. Если  $(f(x_0) + \rho s_0) \geq f(x_0)$  для всех  $\rho > 0$ , то  $(g(x_0 + \rho s_0), s_0) \geq 0$  для  $\rho > 0$ .

Пусть  $g_1$  — точка сгущения множества  $\{g(x_0 + \rho_i s_0), \rho_i \rightarrow +0\}_{i=1}^{\infty}$ . Тогда  $(g_1, s_0) \geq 0$  и  $\|g_0 - g_1\|^2 \geq \|g_0\|^2 + \|g_1\|^2$ . Если  $\|g_0 - g_1\|$  мала, то малы  $\|g_0\|$  и  $\|g_1\|$ , т. е. мы близки к выполнению необходимого условия минимума. В противном случае мы найдем вектор  $g_1 \in G_f(x_0)$ , достаточно далекий от  $g(x_0)$ . Далее, по индукции пусть  $g_0, \dots, g_k \in G_f(x_0)$ ,  $G_k(x_0)$  — выпуклая оболочка  $\{g_0, \dots, g_k\}$ .  $G_k(x_0)$  можно рассматривать как аппроксимацию  $G_f(x_0)$ . Если бы  $G_k(x_0) = G_f(x_0)$ , то направление наискорейшего спуска из точки  $x_0$  вычислялось бы как решение задачи

$$\min_{\|s\|=1} \max_{i \in \overline{0, k}} (s, g_i). \quad (2.40)$$

Задача (2.40) эквивалентна задаче квадратичного программирования  $\min \|s\|^2$  при условии

$$-s = \sum_{i=0}^k \lambda_i g_i; \quad \lambda_i \geq 0; \quad \sum_{i=0}^k \lambda_i = 1,$$

т. е. нахождения точки на —  $G_k(x_0)$ , лежащей на кратчайшем расстоянии от начала координат (см. теорему 1.11).

Пусть  $s_k$  — решение задачи (2.40). Тогда либо в направлении  $s_k$  возможно уменьшение функции  $f(x)$ , либо найдется  $g_{k+1} \in G_f(x_0)$  такой, что  $(g_{k+1}, s_k) \geq 0$ . Продолжая этот процесс, в итоге либо получим на некотором шаге направление спуска, либо построим сходящуюся к нулю последовательность  $\{g_k\}$ , что будет свидетельствовать о том, что точка  $x_0$  является точкой минимума. Легко представить себе алгоритм, в котором направление спуска строится выше приведенным способом, а сам спуск производится до минимума в заданном направлении. Однако этот алгоритм не будет работоспособен, так как 1) поиск минимума по направлению предполагается сколь угодно точным; 2) алгоритм, являясь аналогом наискорейшего спуска, может сойтись не к оптимальной точке.

Чтобы избавиться от этих недостатков, нужно перейти от субградиентного множества  $G_f(x)$  к  $\varepsilon$ -субградиентному

множеству  $\partial_{\varepsilon} f(x)$ , определяемому следующим образом:

$$\partial_{\varepsilon} f(x_0) = \{g \in E_n : \forall x \in E_n, f(x) - f(x_0) \geq (g, (x - x_0)) - \varepsilon\}.$$

$\varepsilon$ -субградиентное множество  $\partial_{\varepsilon} f(x)$  обладает свойствами:

1) для произвольного  $x \in E_n$  и  $\varepsilon \geq 0$  оно не пустое, выпуклое и компактное;

2)  $\partial_0 f(x) = G_f(x)$ ;

3) если  $\varepsilon > 0$ , то существует  $\delta > 0$  такое, что если  $\|y - x\| \leq \delta$  и  $y \in G_f(y)$ , то  $g \in \partial_{\varepsilon} f(x)$  [74].

Справедливо также следующее соотношение, обобщающее формулу для производной по направлению:

$$\max_{g \in \partial_{\varepsilon} f(x)} (s, g) = \inf_{\rho \geq 0} \left\{ \frac{f(x + \rho s) - f(x) + \varepsilon}{\rho} \right\}.$$

Из этой формулы следует, что если существует направление  $s$  такое, что

$$\max_{g \in \partial_{\varepsilon} f(x)} (s, g) < 0, \quad (2.41)$$

то найдется  $\rho > 0$ , при котором  $f(x + \rho s) < f(x) - \varepsilon$ ; если же такого направления не существует, то  $f(x) \leq f^* + \varepsilon$ . Таким образом, для определения  $\min f(x)$  с точностью до  $\varepsilon$  в принципе можно построить алгоритм, основанный на пошаговом определении направления  $s$ , удовлетворяющего (2.41). Трудность состоит в том, что множество  $\partial_{\varepsilon} f(x)$  в явном виде не задается. Эта трудность преодолевается в излагаемом ниже алгоритме Лемарешаля [98] путем аппроксимации  $\partial_{\varepsilon} f(x)$  множеством, являющимся выпуклым замыканием субградиентов, взятых в точках, лежащих в некоторой окрестности  $x$  и генерируемых естественным путем в процессе спуска на предыдущих итерациях.

Перейдем к описанию алгоритма (назовем его  $L$ -алгоритмом). Основу его составляет двойной цикл. Внешняя часть относится к построению минимизирующей подпоследовательности (индекс  $p$ ), внутренняя — к нахождению направления спуска (индекс  $k$ ).

Пусть  $\varepsilon > 0$ ,  $\eta > 0$ ,  $x_0 \in E_n$ ;  $g_0 \in \partial_0 f(x_0) \equiv G_f(x_0)$ ;  $p = 0$ .

1. Примем  $s_0 = -g_0$ ,  $k = 0$ .

2. Найти шаг  $\rho_k = \arg \min_{\rho_k \geq 0} f(x_p + \rho s_k)$  и  $g_{k+1} \in G_f(x_p + \rho_k s_k)$  такой, что  $(g_{k+1}, s_k) = 0$ .

3. Если  $f(x_p + \rho_k s_k) < f(x_p) - \varepsilon$ , то  $x_{p+1} = x_p + \rho_k s_k$ ,  $g_0 = g_{k+1}$  и переходим к шагу 1. Иначе переходим к шагу 4.

4. Найти  $s_{k+1}$  — проекцию нуля на выпуклое замыкание  $\{-g_0, \dots, g_{k+1}\}$ .

5. Если  $\|s_{k+1}\| \leq \eta$ , то вычисления прекращаются. Иначе переходим в шаг 6.

6. Увеличить  $k$  на единицу и перейти к шагу 2.

Легко доказывается следующая теорема.

**Теорема 2.17.** Пусть  $\lim_{\|x\| \rightarrow \infty} f(x) = +\infty$ ,  $x(\eta)$  — результат работы  $L$ -алгоритма. Если последовательность  $\{\eta_r\}_{r=0}^{\infty} \rightarrow 0$ , то  $\lim_{r \rightarrow \infty} f(x(\eta_r)) \leq \inf f(x) + \varepsilon$ . (Имеется в виду, что

после окончания работы  $L$ -алгоритма с  $\eta = \eta_m$  мы переходим к  $\eta = \eta_{m+1}$ , беря в качестве начального приближения  $x(\eta_m)$ .)

**Доказательство.** Так как  $\lim_{\|x\| \rightarrow \infty} f(x) = +\infty$ , а  $L$ -алгоритм является монотонным, то  $\{x(\eta_m)\}_{m=1}^{\infty}$  ограничена. Далее, по построению существует  $s(\eta) \in \partial_{\varepsilon} f(x(\eta))$  такое, что  $\|s(\eta)\| \leq \eta$ . Из ограниченности  $\{x(\eta_m)\}$  вытекает существование предельной точки. Обозначим ее  $x(0)$ . Из замкнутости отображения  $x \rightarrow \partial_{\varepsilon} f(x)$  вытекает, что  $0 \in \partial_{\varepsilon} f(x(0))$ . Отсюда следует, что  $\lim_{m \rightarrow \infty} f(x(\eta_m)) = f(x(0)) \leq \inf f(x) + \varepsilon$ , что и требовалось доказать.

В последние годы было предложено несколько алгоритмов, по идее очень близких к  $L$ -алгоритму, но отличающихся от него способом приближенного определения минимума по направлению, признаками переключения в п. 5, критериями остановки [99, 106, 102]. В работе Мифлина [102]  $L$ -алгоритмы перенесены на гораздо более широкий класс негладких функций и предложены модификации, дающие возможность решать задачи ограничениями\*.

## § 6. Обобщение метода ОГС на класс невыпуклых функций. Стохастические варианты ОГС и его устойчивость

При переносе немонотонных вариантов метода ОГС на класс невыпуклых функций нужно найти условие, в некотором смысле заменяющее очевидное для выпуклых функций свой-

\* В последнее время появились работы В. Ф. Демьянова, Л. В. Васильева, Е. А. Нурминского и др., посвященные исследованию  $\varepsilon$ -субградиентных процессов.

ство, что антиградиент образует острый угол с направлением из данной точки на точку минимума. Такое условие используется в теореме, доказанной Л. Г. Баженовым [1].

**Теорема 2.18.** Пусть  $f(x)$  — почти дифференцируемая функция,  $x^*$  — точка локального минимума этой функции такая, что

$$f(x^*) = \min_{x \in S_r} f(x); \quad S_r = \{x : \|x - x^*\| \leq r\}; \quad \|x_0 - x^*\| < r,$$

и для любого  $\varepsilon$ ,  $0 < \varepsilon < r$ ,

$$\inf_{x \in S_r | S_\varepsilon} (g_f(x), x - x^*) > 0.$$

Тогда последовательность  $\{x_k\}_{k=0}^\infty$ , образованная по формулам

$$x_{k+1} = \begin{cases} \bar{x}_{k+1}, & \text{если } \bar{x}_{k+1} \in S_r, \\ x_0, & \text{если } \bar{x}_{k+1} \notin S_r, \end{cases}$$

где

$$\bar{x}_{k+1} = x_k - h_k \frac{g_f(x_k)}{\|g_f(x_k)\|}; \quad h_k > 0; \quad \sum_{k=0}^{\infty} h_k = +\infty;$$

$$\lim_{k \rightarrow \infty} h_k = 0,$$

сходится к  $x^*$ .

В работе [22] введено понятие обобщенного стохастического градиента и обоснован метод случайного поиска, являющийся стохастическим аналогом ОГС. В дальнейшем этот метод получил развитие в работах Ю. М. Ермольева, Е. А. Нурминского и др. (см. [23]) и является эффективным средством решения разнообразных задач стохастического программирования. Здесь для полноты обзора кратко изложим основные идеи метода обобщенных стохастических градиентов (ОСГ). Этот метод принадлежит к разновидности методов случайного поиска, т. е. к таким итеративным процедурам, у которых направление движения на данном шаге определяется в результате реализации некоторого случайного (псевдослучайного) события в отличие от жестко определенных процедур в обычных градиентных методах (если пренебречь различного рода погрешностями при вычислениях).

При минимизации выпуклой функции  $f(x)$  метод ОСГ задается следующей формулой:

$$x_{k+1} = x_k - h_k(x_k) g_\omega(x_k), \quad k = 0, 1, \dots,$$

где  $h_k$  — шаговый множитель на  $k$ -м шаге,  $g_\omega(x_k)$  — случайный вектор, математическое ожидание которого совпадает с обобщенным градиентом функции  $f(x)$  в точке  $x_k$ . Для простоты предполагаем, что вероятностные характеристики вектора  $g_\omega(x_k)$  определяются точкой  $x_k$  и не зависят от предыстории процесса поиска (это требование не является существенным, и обоснование сходимости проходит для более общего случая).

Пусть  $x^*$  — единственная точка минимума функции  $f(x)$ . Справедлива следующая теорема [22].

**Теорема 2.19.** Пусть выполняются следующие условия:

$$1) \sum_{k=1}^{\infty} h_k(x_k) = +\infty; \quad h_k(x_k) > 0;$$

$$2) \sum_{k=1}^{\infty} h_k^2(x_k) < \infty;$$

$$3) M_\omega \|g_\omega(x_k)\|^2 < c, \quad k = 0, 1, \dots \quad (M_\omega \text{ — символ математического ожидания}).$$

Тогда с вероятностью 1

$$\lim_{k \rightarrow \infty} \|x_k - x^*\| = 0.$$

**Доказательство.** Доказательство сходимости основано на свойствах случайных последовательностей, называемых полумартингалами. Пусть  $\{y_k\}_{k=1}^{\infty}$  — последовательность случайных величин, причем выполняется свойство

$$M(y_n | y_{n-1}, \dots, y_1) \leq y_{n-1}.$$

Такая случайная последовательность называется полумартингалом. Понятие полумартингала является обобщением понятия монотонной последовательности.

Для полумартингалов доказана следующая теорема о сходимости [16]: если  $M(|y_n|) \leq c < +\infty$ , то с вероятностью 1 существует предел последовательности  $y_n$

$$\lim_{n \rightarrow \infty} y_n = y^*,$$

причем  $M(|y^*|) < +\infty$ .

Рассмотрим, как ведет себя последовательность  $\{M_\omega \|x_k - x^*\|^2\}_{k=0}^{\infty}$ :

$$\begin{aligned} \|x_{k+1} - x^*\|^2 &= \|x_k - h_k(x_k) g_\omega(x_k) - x^*\|^2 = \\ &= \|x_k - x^*\|^2 - 2h_k(x_k) (g_\omega(x_k), x_k - x^*) + h_k^2(x_k) \|g_\omega(x_k)\|^2, \end{aligned}$$

$$M_{\omega} \|x_{k+1} - x^*\|^2 = M_{\omega} \|x_k - x^*\|^2 - \\ - 2h_k(x_k) (M_{\omega} g_{\omega}(x_k), x_k - x^*) + h_k^2 M_{\omega} \|g_{\omega}(x_k)\|^2.$$

Так как  $M_{\omega} g_{\omega}(x_k) = g_f(x_k)$ ;  $(g_f(x_k), x_k - x^*) \geq 0$ , то

$$M_{\omega} (\|x_{k+1} - x^*\|^2 | x_k) \leq \|x_k - x^*\|^2 + ch_k^2(x). \quad (2.42)$$

Рассмотрим величину  $z_k = \|x_k - x^*\|^2 + c \sum_{s=k}^{\infty} h_s^2$ . Неравенство (2.42) равносильно неравенству  $M(z_{k+1} | z_k, \dots, z_1) \leq z_k$ . Следовательно,  $\{z_k\}_{k=1}^{\infty}$  — полумартингал и сходится к пределу  $z^*$  с вероятностью 1. Так как  $\lim_{k \rightarrow \infty} \sum_{s=k}^{\infty} h_s^2 = 0$ , то к этому же пределу сходится с вероятностью 1 и  $\{\|x_k - x^*\|^2\}$ . Покажем теперь, что предел равен нулю. Доказательство проведем от противного. Если  $\lim z_k$  не равен  $z^*$ , то найдутся  $\varepsilon > 0$ ,  $\delta > 0$ , что с вероятностью  $\delta$ , отличной от нуля, существует бесконечная подпоследовательность  $x_{k_1}(\omega), \dots, x_{k_i}(\omega) \dots$   $k_1 < k_2 < \dots$ , такая, что  $\|x_{k_i}(\omega) - x^*\| \geq \varepsilon$ ,

откуда  $\sum_{k=1}^{\infty} h_k(x_k) (M_{\omega} g_{\omega}(x_k), x_k - x^*) = +\infty$ , что противоречит соотношению

$$M_{\omega} \|x_{k+1} - x^*\|^2 = \|x_0 - x^*\|^2 - 2 \sum_{s=0}^k h_s(x_s) (M_{\omega} g_{\omega}(x_s), \\ x_s - x^*) + \sum_{s=0}^k h_s^2(x_s) M_{\omega} \|g_{\omega}(x_s)\|^2.$$

Теорема доказана.

Метод ОСГ находит большое число практических применений, особенно при решении многоэтапных задач стохастического программирования [22, 72].

Условие 2) в теореме более сильное, чем применяемое для обоснования метода ОСГ условие  $h_k \rightarrow 0$ , хотя его выполнение не вызывает дополнительных трудностей (например, при  $h_k = \frac{1}{k}$  условие 2) выполняется).

Аналогичное условие использовано в работе М. А. Шепилова [52] при исследовании устойчивости метода ОСГ к малым ошибкам в определении  $x_k$  и  $g_f(x_k)$ . Он доказал следующую теорему.

**Теорема 2.20.** Если функция  $f(x)$  выпукла в  $E_n$ , множество минимумов  $M^*$  не пустое, то при любой начальной точке  $x_0$  и

$$x_{k+1} = x_k - h_k \frac{gf(\bar{x}_k)}{\|gf(\bar{x}_k)\|}, \quad k = 0, 1, 2, \dots,$$

где

$$\|\bar{x}_k - x_k\| \leq \delta_k; \quad \lim_{k \rightarrow \infty} \delta_k = 0; \quad h_k > 0,$$

$$\sum_{k=1}^{\infty} h_k \delta_k < \infty; \quad \sum_{k=0}^{\infty} h_k^2 < \infty; \quad \sum_{k=1}^{\infty} h_k = \infty,$$

справедливо равенство  $\lim_{k \rightarrow \infty} x_k = x^* \in M^*$ .

Эта теорема вместе со способом приближенного вычисления обобщенного градиента, описанного нами в § 3 гл. 1, обосновывает возможность построения «универсального» алгоритма нахождения минимума выпуклой функции  $f(x)$  без вычисления субградиента (при условии, что мы имеем возможность вычислять значение функции  $f(x)$  в произвольной точке с любой точностью).

### МЕТОДЫ ГРАДИЕНТНОГО ТИПА С РАСТЯЖЕНИЕМ ПРОСТРАНСТВА

#### § 1. Эвристика, лежащая в основе методов с растяжением пространства

Исследование метода обобщенного градиентного спуска показало, что с помощью одних лишь ухищрений, связанных с регулировкой шагового множителя, двигаясь на каждом шаге строго в направлении обобщенного антиградиента, трудно добиться в достаточно общем случае значительного ускорения сходимости. В самом деле, медленная сходимость связана с тем, что антиградиент образует угол, близкий к  $\frac{\pi}{2}$ , с направлением на точку минимума. В такой ситуации расстояние до точки минимума убывает на величину, намного меньшую, чем длина шага, а значит, и скорость уменьшения шага не может быть слишком большой, если мы хотим гарантировать сходимость к минимуму.

С другой стороны, имеется простая возможность изменения углов между направлением градиента и направлением на точку минимума — использование линейных неортогональных преобразований пространства. Возникает идея построения в процессе последовательных приближений линейных операторов, изменяющих метрику пространства, и выбора направления спуска, соответствующего направлению антиградиента в пространстве с измененной метрикой. Это направление может уже значительно отличаться от направления антиградиента.

Как же строить операторы преобразования пространства? Если мы обратимся к хорошо исследованным алгоритмам квазиьютоновского типа, применяемым для минимизации дважды непрерывно дифференцируемых функций, то увидим, что основная идея построения алгоритмов с изменяемой метрикой состоит в получении тем или иным способом матрицы, близкой матрице, обратной гессиану в точке минимума (или отличающейся от нее множителем), т. е.

фактически в использовании идеи квадратической аппроксимации минимизируемой функции и имитации метода Ньютона — Рафсона без вычисления вторых производных [42]. Для минимизации негладких функций такой подход в принципе невозможен. Например, для любых кусочно-линейных функций гессиан почти везде равен нулю и матрицы, обратной гессиану, просто не существует. Поэтому для построения алгоритмов градиентного типа с преобразованием пространства, рассчитанных на достаточно широкий класс негладких функций, нужно использовать другие идеи.

Как мы указывали выше, основным фактором, замедляющим сходимость градиентных методов в условиях программной регулировки шага, является близость к нулю косинуса угла между направлением градиента и направлением на точку минимума. Если бы мы могли значительно уменьшать составляющую антиградиента, ортогональную направлению на точку минимума, оставляя почти без изменения составляющую, параллельную этому направлению, то получаемый при этом вектор давал бы направление спуска более предпочтительное, чем направление антиградиента, и можно было бы ожидать ускорения сходимости при соответствующей регулировке шага. Как получить аппроксимацию направления длительного убывания функции или ортогональных к нему направлений? Заметим, что именно направление градиента в неблагоприятном случае является почти ортогональным к направлению на точку минимума. Если исходить из гипотезы почти ортогональности градиента к направлению на точку минимума, то мы должны стараться уменьшить на последующих шагах составляющие градиентов, параллельные градиенту, полученному на данном шаге, а это можно сделать, проведя операцию «растяжения пространства в направлении градиента». Так, применяя весьма нестрогие эвристические рассуждения, приходим к идее обобщенного градиентного метода с растяжением пространства в направлении градиента. Подробное описание и обоснование этого метода дано ниже.

Другой класс методов градиентного типа с растяжением пространства, но уже в направлении разности двух последовательных градиентов связан с последовательными преобразованиями пространства, которые нацелены в общих чертах на то, чтобы превратить «узкий» конус подходящих направлений в «широкий». При этом удается получить в определенном смысле почти монотонный метод, хорошо

зарекомендовавший себя на практике для широкого круга задач. Этот метод изложен в § 6, 7 настоящей главы.

В § 8 рассматривается группа методов минимизации выпуклых функций, в основу которых положены алгоритм последовательных отсечений опорными гиперплоскостями и аппроксимация оставшейся области с помощью эллипсоидов. Преобразование пространства используется для превращения эллипсоида в шар (операция симметризации). Важный частный случай алгоритма такого типа, принадлежащего к семейству алгоритмов с растяжением пространства в направлении градиента [76], впервые был рассмотрен в работе [74]. В целом алгоритмы этого типа представляют большой теоретический и практический интерес, так как обеспечивают сходимость по функционалу со скоростью геометрической прогрессии, знаменатель которой зависит лишь от размерности пространства.

## § 2. Операторы растяжения пространства

Пусть задан вектор  $\xi \in E_n$ ,  $\|\xi\| = 1$ , и число  $\alpha \geq 0$ . Каждый вектор  $x \in E_n$  однозначно представим в следующем виде:

$$x = \gamma_{\xi}(x) \xi + d_{\xi}^{\perp}(x) \quad (3.1)$$

при условии

$$(\xi, d_{\xi}^{\perp}(x)) = 0. \quad (3.2)$$

Из соотношений (3.1), (3.2) получаем  $\gamma_{\xi}(x) = (x, \xi)$ ;  $d_{\xi}^{\perp}(x) = x - (x, \xi) \xi$ .

**Определение.** Оператором растяжения пространства  $E_n$  в направлении  $\xi$  с коэффициентом  $\alpha$  назовем оператор  $R_{\alpha}(\xi)$ , действующий следующим образом на вектор  $x$ , представленный в форме (3.1):

$$R_{\alpha}(\xi) x = \alpha \gamma_{\xi}(x) \xi + d_{\xi}^{\perp}(x).$$

Из этого определения следует:

$$1) R_{\alpha}(\xi) x = \alpha (x, \xi) \xi + [x - (x, \xi) \xi] = (\alpha - 1) (x, \xi) \xi + x; \quad (3.3)$$

2) оператор  $R_{\alpha}(\xi)$  — линейный симметричный,

$$(R_{\alpha}(\xi) x, y) = (\alpha - 1) (x, \xi) (y, \xi) + (x, y) = (x, R_{\alpha}(\xi) y);$$

3)  $R_{\alpha\beta}(\xi) = R_{\alpha}(\xi) R_{\beta}(\xi)$ ;

4) при  $\alpha > 0$   $R_{\alpha}(\xi) \cdot R_{\frac{1}{\alpha}}(\xi) = R_1(\xi) = I$ ;  $I$  — единичный

оператор;

5) оператор  $R_0(\xi)$  является оператором проектирования на подпространство, ортогональное вектору  $\xi$ :

$$R_0(\xi)x = d_\xi(x);$$

6) оператор  $R_\alpha(\xi)$  при  $n \geq 2$  имеет два собственных числа  $\lambda_1 = \alpha$ ,  $\lambda_2 = 1$ : первому из них соответствует подпространство собственных векторов, порожденное вектором  $\xi$ , второму — подпространство собственных векторов, состоящее из векторов, ортогональных  $\xi$ ;

7) пусть координаты вектора  $\xi$  в некоторой ортонормированной системе координат  $s = \{e_1, e_2, \dots, e_n\}$  равны  $\xi_1, \xi_2, \dots, \xi_n$ , тогда в силу (3.3) в этой системе координат преобразованию  $R_\alpha(\xi)$  соответствует матрица  $R_\alpha(\xi)$  с элементами  $\{r_{ij}\}$ , вычисленными по следующим формулам:

$$r_{ij} = (R_\alpha(\xi)e_i, e_j) = \begin{cases} (\alpha - 1)\xi_i\xi_j & \text{для } i \neq j, \\ (\alpha - 1)\xi_i^2 + 1 & \text{для } i = j; \end{cases}$$

8) в силу формулы (3.3) вычисление вектора  $R_\alpha(\xi)x$  требует  $(2n + 1)$  операций умножения, вычисление матриц вида  $R_\alpha(\xi)A$  или  $AR_\alpha(\xi)$  при заданных матрице  $A$  и  $\xi$ ,  $\alpha$  требует  $n(2n + 1)$  операций умножения;

9) если  $x$  — произвольный ненулевой вектор из  $E_n$ , то

$$\|R_\alpha(\xi)x\| = \sqrt{\|x\|^2 + (\alpha^2 - 1)(x, \xi)^2}. \quad (3.4)$$

Этот результат получается непосредственным применением формулы (3.3). В самом деле,

$$\|R_\alpha(\xi)x\|^2 = (R_\alpha(\xi)x, R_\alpha(\xi)x) =$$

$$= (x + (\alpha - 1)(x, \xi)\xi, x + (\alpha - 1)(x, \xi)\xi) = \|x\|^2 +$$

$$+ 2(\alpha - 1)(x, \xi)^2 + (\alpha - 1)^2(x, \xi)^2 = \|x\|^2 + (\alpha^2 - 1)(x, \xi)^2,$$

откуда и следует равенство (3.4);

10)  $R_\alpha(\xi)$  в матричной форме представим в виде

$$R_\alpha(\xi) = I + (\alpha - 1)\xi\xi^T.$$

Действительно,

$$(I + (\alpha - 1)\xi\xi^T)x = x + (\alpha - 1)\xi(x, \xi), \quad (3.5)$$

что соответствует определяющему соотношению (3.3).

### § 3. Обобщенный градиентный спуск с растяжением пространства в направлении градиента

Рассмотрим класс алгоритмов минимизации выпуклых функций, на каждом шаге которых движение в направлении обобщенного градиента будет сочетаться с операцией растяжения пространства аргументов в этом же направлении.

Алгоритмы этого класса будем называть алгоритмами обобщенного градиентного спуска с растяжением пространства в направлении градиента (ОГСРП-алгоритмами).

Предположим, что имеется алгоритм, позволяющий точно вычислить обобщенный градиент  $g_f(x)$  минимизируемой выпуклой функции  $f(x)$  в произвольной точке  $x \in E_n$ , а также алгоритмы вычисления последовательностей положительных чисел  $\{h_k\}$  и  $\{\alpha_k\}$ ,  $k = 1, 2, \dots$  (шаговых множителей и коэффициентов растяжения пространства), задано начальное приближение  $x_0$  и начальная неособая матрица  $B_0 = A_0^{-1}$  (в частности,  $B_0 = I$ ). При этих условиях определим бесконечношаговый процесс,  $(k+1)$ -й шаг которого описывается следующим образом ( $k = 0, 1, 2, \dots$ ). Вычисляем:

1)  $g_f(x_k)$  (если  $g_f(x_k) = 0$ , вычисления прекращаются, так как  $x_k$  дает точку минимума);

$$2) g_{\varphi_k}(y_k) = B_k^* g_f(x_k) = \tilde{g}_k, \quad (3.6)$$

где  $\varphi_k(y) = f(B_k y)$ ;  $y_k = A_k x_k$ ;  $A_k = B_k^{-1}$ ;  $B_k^*$  — оператор, сопряженный оператору  $B_k$ .

Формула (3.6) дает возможность вычислять обобщенный градиент от функции  $\varphi_k(y) = f(A_k^{-1}y)$ , которая получается из  $f(x)$  при применении линейного преобразования пространства  $y = A_k x$ . Эта формула является частным случаем формулы для опорного функционала к выпуклому функционалу, определенному на банаховом пространстве, полученном в результате линейного преобразования некоторого исходного банахового пространства [43]. Она может быть также элементарно выведена из определения обобщенного градиента. В тех точках, где  $g_f(x)$  определен неоднозначно, формулу (3.6) следует понимать как однозначное отображение множества  $G_f(x)$  на множество  $G_{\varphi_k}(y_k)$  (если оператор  $B_k$  неособый, то отображение взаимно однозначное);

$$3) \xi_{k+1} = \frac{g_{\varphi_k}(y_k)}{\|g_{\varphi_k}(y_k)\|} = \frac{\tilde{g}_k}{\|\tilde{g}_k\|}; \quad (3.7)$$

$$4) h_{k+1};$$

$$5) \alpha_{k+1};$$

$$6) x_{k+1} = x_k - B_k h_{k+1} \xi_{k+1}. \quad (3.8)$$

Формула (3.8) получается из формулы  $A_k x_{k+1} = y_k - h_{k+1} \xi_{k+1}$ , осуществляющей шаг обобщенного градиент-

ного спуска для функции  $\varphi_k(y)$ , с последующим применением к обеим частям оператора  $B_k$  для отображения в основное пространство  $E_n$ ;

$$7) B_{k+1} = A_{k+1}^{-1} = B_k R_{1/\alpha_{k+1}}(\xi_{k+1}). \quad (3.9)$$

Формула (3.9) дает возможность вычислить оператор  $B_{k+1}$ , обратный результирующему оператору

$$A_{k+1} = R_{\alpha_{k+1}}(\xi_{k+1}) \dots R_{\alpha_1}(\xi_1) A_0$$

преобразования пространства, который получается в результате последовательного применения оператора растяжения пространства в направлении нормированных обобщенных градиентов  $\xi_1, \xi_2, \dots, \xi_{k+1}$  с коэффициентами  $\alpha_1, \dots, \alpha_{k+1}$ , откуда

$$B_{k+1} = A_{k+1}^{-1} = A_k^{-1} R_{\alpha_{k+1}}^{-1}(\xi_{k+1}) = B_k R_{1/\alpha_{k+1}}(\xi_{k+1});$$

8) переходим к  $(k+2)$ -му шагу.

Каждый шаг ОГСРП требует ряда дополнительных операций по сравнению с методом обобщенного градиентного спуска. Наиболее трудоемкими из них являются операции 2), 6) и 7), каждая из которых требует порядка  $cn^2$  арифметических операций, из которых около  $4n^2$  операций умножения. Кроме того, требуется дополнительный массив из  $n^2$  ячеек памяти для запоминания матриц  $B_k$ .

Заметим, что если  $B_0 = I$  и  $\alpha_k = 1$ ,  $k = 1, 2, \dots$ , то метод ОГСРП становится эквивалентным обобщенному градиентному спуску. Метод ОГСРП является более гибким по сравнению с методом ОГС, так как его конкретные реализации зависят от двух последовательностей  $\{\alpha_k\}$  и  $\{h_k\}$ . Применяя различные алгоритмы построения этих последовательностей, в принципе можно на единой основе строить различные модификации и выбирать из них наиболее эффективные для данного класса задач.

Алгоритмы типа ОГСРП первоначально были разработаны для минимизации выпуклых функций. В дальнейшем оказалось, что некоторые их модификации применимы для нахождения локальных минимумов почти дифференцируемых функций (при этом, естественно, роль обобщенного градиента в алгоритме играет почти-градиент). В дальнейшем при исследовании сходимости алгоритмов ОГСРП в каждом конкретном случае будем оговаривать, идет ли речь о классе почти дифференцируемых функций или о подклассе

выпуклых функций. При этом под обобщенным градиентом читатель должен понимать соответственно почти-градиент или субградиент.

#### § 4. Вопросы сходимости алгоритмов ОГСРП

Покажем, что определенные варианты алгоритмов ОГСРП сходятся по функционалу со скоростью геометрической прогрессии, причем знаменатель этой прогрессии зависит от таких характеристик минимизируемой функции, которые инвариантны по отношению к невырожденным линейным преобразованиям пространства. Отметим, что для класса выпуклых функций удастся построить алгоритм типа ОГСРП, который сходится по функционалу со скоростью геометрической прогрессии со знаменателем, зависящим только от размерности области определения функции. Этот алгоритм описан в § 8 гл. 3.

Доказательства указанных результатов основаны на изучении поведения нормы градиента минимизируемой функции в растянутом пространстве.

**Теорема 3.1.** Пусть  $f(x)$  — почти дифференцируемая функция, для минимизации которой применяется алгоритм ОГСРП (3.6) — (3.9) и в результате получается минимизирующая последовательность  $\{x_k\}_{k=0}^{\infty}$ . Если при этом выполняются для некоторых положительных чисел  $d, \alpha^*, \delta$  следующие условия при  $k = 0, 1, \dots$ :

- 1)  $\|g_f(x_k)\| \leq d$ ;
- 2)  $1 + \delta \leq \alpha_k \leq \alpha^*$ ,

то существует такая подпоследовательность  $\{x_{k_p}\}_{p=0}^{\infty}$ ,  $k_p < k_{p+1}$  и  $c > 0$ , что

$$\|\tilde{g}_{k_p}\| \leq c \left( \prod_{j=1}^{k_p} \alpha_j \right)^{-\frac{1}{n}}; \quad p = 1, 2, \dots$$

**Доказательство.** Запишем матрицу  $A_k$  в виде произведения ортогональной матрицы  $O_k$  и симметричной положительно определенной матрицы  $S_k$  (полярное разложение [7]):

$$A_k = O_k S_k; \quad A_0 = I.$$

Так как  $A_k = R_{\alpha_k}(\xi_k) \dots R_{\alpha_1}(\xi_1)$ , а произведение собственных чисел матрицы равно ее определителю, то произведение

собственных чисел матрицы  $S_k$  равно

$$\det S_k = \det A_k = \prod_{j=1}^k \alpha_j.$$

Пусть  $\{\lambda_i^{(k)}\}_{i=1}^n$  — последовательность собственных чисел оператора  $S_k$ , записанных в порядке неубывания;  $\{e_i^{(k)}\}_{i=1}^n$  — соответствующая ортонормированная система собственных векторов;

$$S_k e_i^{(k)} = \lambda_i^{(k)} e_i^{(k)}; \quad \prod_{i=1}^n \lambda_i^{(k)} = \prod_{j=1}^k \alpha_j; \quad \lambda_n^{(k)} \geq \left( \prod_{j=1}^k \alpha_j \right)^{1/n}. \quad (3.10)$$

Кроме того,  $\lambda_i^{(k)} \geq 1$ , так как  $\alpha_j > 1$ ,  $j = 1, \dots, k$ , и поэтому

$$\|S_k e_i^{(k)}\| = \|A_k e_i^{(k)}\| \geq 1.$$

Пусть  $O_k e_i^{(k)} = \tilde{e}_i^{(k)}$ . Система векторов  $\{\tilde{e}_i^{(k)}\}_{i=1}^n$  также является ортонормированной ввиду ортогональности  $O_k$ .

Запишем  $g_f(x_k)$  в следующем виде:

$$g_f(x_k) = \sum_{i=1}^n g_i^{(k)} e_i^{(k)}.$$

Так как  $\|g_f(x_k)\| \leq d$ , то

$$|g_i^{(k)}| < d; \quad i = 1, 2, \dots, n; \quad k = 1, 2, \dots \quad (3.11)$$

Доказательство теоремы проведем от противного. Допустим, что теорема не верна. Тогда для любого сколь угодно большого  $c$  найдется  $k(c)$  такое, что при  $k > k(c)$

$$\|\tilde{g}_k\| = \|B_k^* g_f(x_k)\| = \left\| \sum_{i=1}^n \frac{g_i^{(k)}}{\lambda_i^{(k)}} \tilde{e}_i^{(k)} \right\| \geq c \left( \prod_{j=1}^k \alpha_j \right)^{-\frac{1}{n}}. \quad (3.12)$$

Это значит, что для  $k > k(c)$  существует такое  $i_k^*$ ,  $1 \leq i_k^* \leq n$ , что

$$\frac{|g_{i_k^*}^{(k)}|}{\lambda_{i_k^*}^{(k)}} \geq \frac{c \left( \prod_{j=1}^k \alpha_j \right)^{-\frac{1}{n}}}{\sqrt{n}},$$

откуда  $\lambda_{i_k}^{(k)} \leq \frac{\sqrt{n} d}{c} \left( \prod_{j=1}^k \alpha_j \right)^{\frac{1}{n}}$ . Далее,

$$\begin{aligned} \lambda_n^{(k)} &= \max_{1 \leq i \leq n} \lambda_i^{(k)} \geq \left( \frac{\prod_{j=1}^k \alpha_j}{\lambda_{i_k}^{(k)}} \right)^{\frac{1}{n-1}} \geq \left( \frac{c}{\sqrt{n} d} \right)^{\frac{1}{n-1}} \left( \prod_{j=1}^k \alpha_j \right)^{\frac{1}{n}} = \\ &= \tilde{c} \left( \prod_{j=1}^k \alpha_j \right)^{\frac{1}{n}}, \quad \tilde{c} = \left( \frac{c}{\sqrt{n} d} \right)^{\frac{1}{n-1}}. \end{aligned}$$

Пусть  $e_n^{(k+1)} = \sum_{i=1}^n d_i^{(k)} e_i^{(k)}$ ;  $\sum_{i=1}^n (d_i^{(k)})^2 = 1$ . Тогда

$$A_{k+1} e_n^{(k)} = \lambda_n^{(k+1)} e_n^{(k+1)} = R_{\alpha_{k+1}}(\xi_{k+1}) \left( \sum_{i=1}^n \lambda_i^{(k)} d_i^{(k)} e_i^{(k)} \right).$$

Введем  $r_k = \min r \left\{ \frac{\lambda_n^{(k)}}{\lambda_r^{(k)}} \leq d_{k+1} c \right\}$ ;  $1 \leq r \leq n$ . Представим

$y = \sum_{i=1}^n \lambda_i^{(k)} d_i^{(k)} e_i^{(k)}$  в виде  $y = y_1 + y_2$ , где

$$y_1 = \sum_{i=1}^{r_k-1} \lambda_i^{(k)} d_i^{(k)} e_i^{(k)}; \quad y_2 = \sum_{i=r_k}^n \lambda_i^{(k)} d_i^{(k)} e_i^{(k)}$$

Оценим  $\|R_{\alpha_{k+1}}(\xi_{k+1}) y\|$ . Для этого заметим, что

$$\xi_{k+1} = \frac{\tilde{g}_k}{\|\tilde{g}_k\|} = \sum_{i=1}^n \mu_i^{(k)} e_i^{(k)},$$

где  $\mu_i^{(k)} = \frac{g_i^{(k)}}{\lambda_i^{(k)} \|\tilde{g}_k\|}$ . Из (3.11), (3.12) получаем

$$|\mu_i^{(k)}| \leq \frac{d \left( \prod_{j=1}^k \alpha_j \right)^{\frac{1}{n}}}{\lambda_i^{(k)} c}.$$

$$\text{Для } i \geq r_k, \quad |\mu_i^{(k)}| \leq \frac{\alpha \left( \prod_{j=1}^k \alpha_j \right)^{\frac{1}{n}} \alpha_{k+1}}{\lambda_n^{(k)}} \leq \frac{d\alpha^*}{\tilde{c}}$$

$$\|R_{\alpha_{k+1}}(\xi_{k+1}) y_1\| \leq \alpha_{k+1} \|y_1\| \leq \lambda_n^{(k)} \frac{1}{c} \left\| \sum_{i=1}^{r_k-1} d_i^{(k)} e_i^{(k)} \right\| \leq \frac{\lambda_n^{(k)}}{c};$$

$$\begin{aligned} \|R_{\alpha_{k+1}}(\xi_{k+1}) y_2\| &\leq \|y_2\| + (\alpha_{k+1} - 1) |(y_2, \xi_{k+1})| \leq \\ &\leq \lambda_n^{(k)} \left[ 1 + (\alpha^* - 1) \frac{d\alpha^*}{\tilde{c}} \right] \leq \lambda_n^{(k)} \left[ 1 + (\alpha^* - 1) \frac{d\alpha^*}{\tilde{c}} \right]; \end{aligned}$$

$$\lambda_n^{(k+1)} = \|R_{\alpha_{k+1}}(\xi_{k+1}) y\| \leq \lambda_n^{(k)} \left[ 1 + \frac{1}{c} + (\alpha^* - 1) \frac{d\alpha^*}{\tilde{c}} \right]. \quad (3.13)$$

Из (3.13) следует, что для любого  $\varepsilon_0 > 0$  всегда можно подобрать столь большое  $c$ , что  $\lambda_n^{(k+1)} \leq (1 + \varepsilon_0) \lambda_n^{(k)}$ ,  $k > k(c)$ . Но если выбрать  $\varepsilon_0 < \delta$ , то при достаточно большом  $k$   $\lambda_n^{(k)} < c(1 + \delta)^{\frac{k}{n}} \leq c \left( \prod_{j=1}^k \alpha_j \right)^{\frac{1}{n}}$ , что противоречит (3.10). Теорема доказана.

Близкий результат можно получить для оценки скорости убывания последовательности  $v_k = \min_{1 \leq r \leq k} \|\tilde{g}_r\|$ . Этот результат в дальнейшем будет использоваться для получения оценок скорости сходимости «рекордов» по функционалу к минимуму при использовании ОГСРП.

**Теорема 3.2.** В предположениях теоремы 3.1 при условии  $\alpha_k = \alpha > 1$

$$v_k \leq \frac{d \sqrt{k(\alpha^2 - 1)}}{\sqrt{\frac{2k}{\alpha^n} - 1}} \quad \text{при } k = 1, 2, \dots$$

**Доказательство.** При доказательстве мы воспользуемся теми же обозначениями, что и при доказательстве теоремы 3.1. В этих обозначениях

$$\begin{aligned} \xi_{k+1} &= \frac{\tilde{g}_k}{\|\tilde{g}_k\|} = \sum_{i=1}^n \mu_i^{(k)} e_i^{(k)}, \quad k = 1, 2, \dots, \\ \mu_i^{(k)} &= \frac{g_i^{(k)}}{\lambda_i^{(k)} \|\tilde{g}_k\|}, \quad i = 1, 2, \dots, n, \quad \sum_{i=1}^n (g_i^{(k)})^2 \leq d^2. \end{aligned} \quad (3.14)$$

Пусть для некоторого  $k$

$$v_k = \min_{1 \leq r \leq k} \|\tilde{g}_k\| > \frac{d\sqrt{k(\alpha^2-1)}}{\sqrt{\alpha^{\frac{2k}{n}}-1}}.$$

Тогда для  $r = k$ ,  $1 \leq i \leq n$ , из (3.14) получаем

$$|\mu_i^{(r)}| \leq \frac{|g_i^{(r)}| \sqrt{\alpha^{\frac{2k}{n}}-1}}{\lambda_i^{(r)} d \sqrt{k(\alpha^2-1)}}. \quad (3.15)$$

Используя это соотношение, оценим при  $r$ ,  $1 \leq r \leq k-1$ ,  $(\lambda_n^{(r+1)})^2 - (\lambda_n^{(r)})^2$  ( $\lambda_n^r$  — максимальные собственные числа оператора  $S_r$ ). Рассмотрим произвольный вектор  $a$ ,  $\|a\| = 1$ :

$$a = \sum_{i=1}^n a_i e_i^{(r)}, \quad \sum_{i=1}^n a_i^2 = 1,$$

$$A_r a = \sum_{i=1}^n a_i \lambda_i^{(r)} \tilde{e}_i^{(r)}, \quad (3.16)$$

$$\lambda_n^{(r+1)} = \max_{\|a\|=1} \|A_{r+1} a\| = \max_{\|a\|=1} \|R_\alpha(\xi_{r+1}) A_r a\|. \quad (3.17)$$

Воспользуемся формулой (3.4):

$$\|R_\alpha(\xi) x\| = \sqrt{\|x\|^2 + (\alpha^2 - 1)(x, \xi)^2}.$$

Приняв  $x = A_r a$ , из (3.15) — (3.17) получаем

$$[\lambda_n^{(r+1)}]^2 = \max_{\|a\|=1} \left[ \sum_{i=1}^n a_i^2 (\lambda_i^{(r)})^2 + (\alpha^2 - 1) \left( \sum_{i=1}^n a_i \lambda_i^{(r)} \mu_i^{(r)} \right)^2 \right] \leq$$

$$\leq \max_{\|a\|=1} \left[ (\lambda_n^{(r)})^2 + (\alpha^2 - 1) \left( \sum_{i=1}^n |a_i g_i^{(r)}| \right)^2 \frac{\alpha^{\frac{2k}{n}} - 1}{k d^2 (\alpha^2 - 1)} \right] <$$

$$< [\lambda_n^{(r)}]^2 + \frac{\alpha^{\frac{2k}{n}} - 1}{k},$$

$$[\lambda_n^{(r+1)}]^2 < [\lambda_n^{(r)}]^2 + \frac{\alpha^{\frac{2k}{n}} - 1}{k}, \quad r = 0, 1, \dots, k, \lambda_n^{(0)} = 1.$$

Отсюда  $[\lambda_n^{(k)}]^2 < \alpha^{\frac{2k}{n}}$ , что противоречит тому, что  $\lambda_n^{(k)} \geq \alpha^{\frac{k}{n}}$ . Теорема доказана.

Для того чтобы перейти к оценке скорости сходимости метода ОГРСП по функционалу, нужно получить оценки расстояния до области минимумов в «растянутом» пространстве. Для некоторых вариантов ОГРСП их удается получить.

**Теорема 3.3.** Пусть  $f(x)$  — почти дифференцируемая функция, определенная в некоторой сферической окрестности  $S_d$  точки  $x^*$ , являющейся локальным минимумом  $f(x)$ ,  $S_d = \{x : \|x - x^*\| \leq d\}$ , и в области  $S_d$  почти-градиент удовлетворяет двустороннему неравенству

$$N[f(x) - f(x^*)] \leq (g(x), x - x^*) \leq M[f(x) - f(x^*)] \quad (3.18)$$

для всех  $x \in S_d$  ( $M$  и  $N$  — положительные константы,  $M > N$ ). Тогда если в алгоритме ОГРСП примем

$$1) x_0 \in S_d;$$

$$2) h_{k+1} = \frac{2MN}{M+N} \frac{f(x_k) - f(x^*)}{\|\hat{g}_k\|}; \quad (3.19)$$

$$3) 1 < \alpha_{k+1} = \alpha \leq \frac{M+N}{M-N}, \quad k = 0, 1, 2, \dots, \quad (3.20)$$

то для всех  $k = 0, 1, \dots$   $\|A_k(x_k - x^*)\| \leq d$ .

**Доказательство.** Этот результат доказывается методом математической индукции. Пусть для некоторого  $p$

$$\|A_p(x_p - x^*)\| \leq d.$$

Оценим

$$\|A_{p+1}(x_{p+1} - x^*)\|.$$

Обозначим  $A_k(x_k - x^*) = z_k$ . Тогда

$$\begin{aligned} \|z_{p+1}\|^2 &= \|R_{\alpha_{p+1}}(\xi_{p+1})(z_p - h_{p+1}\xi_{p+1})\|^2 = \\ &= \|R_{\alpha_{p+1}}(\xi_{p+1})z_p\|^2 - 2h_{p+1}(R_{\alpha_{p+1}}(\xi_{p+1})z_p, R_{\alpha_{p+1}}(\xi_{p+1})\xi_{p+1}) + \\ &+ \alpha_{p+1}^2 h_{p+1}^2 = \left[1 + (\alpha_{p+1}^2 - 1) \frac{(z_p, \xi_{p+1})^2}{\|z_p\|^2}\right] \|z_p\|^2 - \\ &- 2h_{p+1}\alpha_{p+1}^2 (z_p, \xi_{p+1}) + \alpha_{p+1}^2 h_{p+1}^2 = \\ &= \|z_p\|^2 + \alpha_{p+1}^2 [(z_p, \xi_{p+1}) - h_{p+1}]^2 - (z_p, \xi_{p+1})^2, \\ \|z_{p+1}\|^2 &= \|z_p\|^2 + \alpha_{p+1}^2 [(z_p, \xi_{p+1}) - h_{p+1}]^2 - (z_p, \xi_{p+1})^2 \quad (3.21) \end{aligned}$$

Заметим, что

$$\begin{aligned} (z_p, \xi_{p+1}) &= (A_p(x_p - x^*), \xi_{p+1}) = \\ &= (x_p - x^*, A_p^* \xi_{p+1}) = \frac{1}{\|\hat{g}_p\|} (g_p(x_p), x_p - x^*). \end{aligned}$$

Рассмотрим два случая:

а)  $(z_p, \xi_{p+1}) \leq h_{p+1}$ , тогда из (3.18), (3.19)

$$\begin{aligned} h_{p+1} &= \frac{[f(x_p) - f(x^*)] 2MN}{\|\tilde{g}_p\| (M+N)} \leq \frac{(g_1(x_p), x_p - x^*) 2M}{\|\tilde{g}_p\| (M+N)} = \\ &= (z_p, \xi_{p+1}) \frac{2M}{M+N}, \end{aligned}$$

откуда

$$\begin{aligned} &\alpha_{p+1}^2 [(z_p, \xi_{p+1}) - h_{p+1}]^2 - (z_p, \xi_{p+1})^2 \leq \\ &\leq \alpha_{p+1}^2 \left[ (z_p, \xi_{p+1}) - (z_p, \xi_{p+1}) \frac{2M}{M+N} \right]^2 - (z_p, \xi_{p+1})^2 \leq \\ &\leq \left[ \left( \frac{M+N}{M-N} \right)^2 \left( 1 - \frac{2M}{M+N} \right)^2 - 1 \right] (z_p, \xi_{p+1})^2 = 0; \end{aligned}$$

б)  $(z_p, \xi_{p+1}) \geq h_{p+1}$  и из (3.18), (3.19) получаем

$$\begin{aligned} h_{p+1} &= \frac{[f(x_p) - f(x^*)] 2MN}{\|\tilde{g}_p\| (M+N)} \geq \frac{(g_f(x_p), x_p - x^*) 2N}{M+N} = \\ &= (z_p, \xi_{p+1}) \frac{2N}{M+N}, \end{aligned}$$

$$\begin{aligned} &\alpha_{p+1}^2 [(z_p, \xi_{p+1}) - h_{p+1}]^2 - (z_p, \xi_{p+1})^2 \leq \\ &\leq \alpha_{p+1}^2 \left[ (z_p, \xi_{p+1}) - (z_p, \xi_{p+1}) \frac{2N}{M+N} \right]^2 - (z_p, \xi_{p+1})^2 \leq \\ &\leq \left[ \left( \frac{M+N}{M-N} \right)^2 \left( 1 - \frac{2N}{M+N} \right)^2 - 1 \right] (z_p, \xi_{p+1})^2 = 0. \end{aligned}$$

Таким образом, в обоих случаях

$$\alpha_{p+1}^2 [(z_p, \xi_{p+1}) - h_{p+1}]^2 - (z_p, \xi_{p+1})^2 \leq 0.$$

Отсюда, используя равенство (3.21), получаем  $\|z_{p+1}\| \leq \|z_p\|$ . Отсюда следует, что  $\|A_k(x_k - x^*)\| \leq d$  для любого  $k = 0, 1, \dots$  Теорема доказана.

Как следствие из теорем 3.1, 3.2 и 3.3 получаем следующий результат.

**Теорема 3.4.** В условиях теоремы 3.3 найдется такая подпоследовательность индексов  $k_1, k_2, \dots$  и такое положительное число  $c$ , что

$$f(x_{k_p}) - f(x^*) \leq c \alpha^{-\frac{k_p}{n}}, \quad p = 1, 2, \dots$$

Кроме того,

$$\min_{1 \leq i \leq k} [f(x_i) - f(x^*)] \leq \frac{G \sqrt{k(\alpha^2 - 1)} d}{\sqrt{\alpha^{\frac{2k}{n}} - 1}},$$

где  $G = \max_{x \in S_d} \|g_f(x)\|$ .

**Доказательство.** Так как процесс спуска проходит в ограниченном множестве  $S_d$ , то множество почти-градиентов ограничено, значит, выполняются условия теорем 3.1 и 3.2.

Пусть  $\{k_p\}_{p=1}^{\infty}$  — подпоследовательность, получаемая в соответствии с теоремой 3.1, для которой  $\|\tilde{g}_{k_p}\| \leq \bar{c}\alpha^{-\frac{k_p}{n}}$ . Используя (3.18), получаем

$$\begin{aligned} N [f(x_{k_p}) - f(x^*)] &\leq (g_f(x_{k_p}), x_{k_p} - x^*) = \\ &= (A_{k_p}^* \tilde{g}_{k_p}, x_{k_p} - x^*) = (\tilde{g}_{k_p}, A_{k_p}(x_{k_p} - x^*)) \leq \\ &\leq \|\tilde{g}_{k_p}\| d \leq \bar{c} d \alpha^{-\frac{k_p}{n}}. \end{aligned}$$

Приняв  $\frac{\bar{c}d}{N} = c$ , получим

$$f(x_{k_p}) - f(x^*) \leq \frac{\bar{c}d}{N} \alpha^{-\frac{k_p}{n}} = c \alpha^{-\frac{k_p}{n}}.$$

Аналогичным образом получаем, используя теорему 3.2,

$$\min_{1 \leq r \leq k} [f(x_r) - f(x^*)] \leq \min_{1 \leq r \leq k} \|\tilde{g}_r\| \frac{d}{N},$$

откуда

$$\min_{1 \leq r \leq k} [f(x_r) - f(x^*)] \leq \frac{G \sqrt{k(\alpha^2 - 1)} d}{N \sqrt{\alpha^{\frac{2k}{n}} - 1}}.$$

Теорема доказана.

Теорема 3.3 обобщается на случай, когда минимум функции  $f(x)$  определен неоднозначно.

Пусть множество  $M^*$  точек минимума почти дифференцируемой функции  $f(x)$  ограничено и  $f(x)$  принимает на  $M^*$  значение  $f^*$ . Введем обозначение

$$\rho_k(x) = \min_{x^* \in M^*} \|A_k(x_k - x^*)\| = \|A_k(x - x_k^*(x))\|; \quad x_k^*(x) \in M^*.$$

**Теорема 3.5.** Пусть для всех  $x, x^*$ , удовлетворяющих условию  $\rho_0(x) \leq d$ ;  $x \in M^*$ ;  $x^* \in M^*$ , выполняются неравенства

$$N [f(x) - f^*] \leq (g_f(x), x - x^{**}) \leq M [f(x) - f^*],$$

где  $x^{**}$  — ближайшая к  $x$  точка минимума, лежащая на луче  $y = x + t(x - x^*)$ ,  $t \geq 0$ . Тогда если при применении алгоритма ОГСРП

$$1) \rho_0(x_0) \leq d;$$

$$2) h_{k+1} = \frac{2MN}{M+N} \frac{f(x_k) - f^*}{\|g_{\varphi_k}(y_k)\|};$$

$$3) 1 < \alpha_{k+1} = \alpha \leq \frac{M+N}{M-N},$$

то  $\rho_{k+1}(x_{k+1}) \leq d$ ;  $k = 0, 1, 2, \dots$

**Доказательство.** Покажем, что последовательность  $\{\rho_k(x_k)\}$  является невозрастающей:

$$\begin{aligned} \rho_{k+1}^2(x_{k+1}) &= \min_{y \in M^*} \|R_{\alpha_{k+1}}(\xi_{k+1}) A_k(x_{k+1} - y)\|^2 = \\ &= \min_{y \in M^*} \|R_{\alpha_{k+1}}(\xi_{k+1}) A_k(x_k - y) - R_{\alpha_{k+1}}(\xi_{k+1}) h_{k+1} \xi_{k+1}\|^2 \leq \\ &\leq \|R_{\alpha_{k+1}}(\xi_{k+1}) A_k(x_k - x_k^*(x_k)) - \alpha_{k+1} h_{k+1} \xi_{k+1}\|^2. \end{aligned}$$

Введем обозначение  $z_k = A_k(x_k - x_k^*(x_k))$ . После этого доказательство аналогично доказательству теоремы 3.3. В итоге получаем  $\|z_{k+1}\| \leq \|z_k\| \leq \|z_0\| \leq d$ ,  $k = 0, 1, 2, \dots$

Теорему 3.5 будем использовать при построении модификаций алгоритма ОГСРП, когда  $f^*$  неизвестно, в предположении, что функция  $f(x)$  является выпуклой. Заметим, что в этом случае  $N$  всегда можно выбирать равным единице, так как  $f(x) - f(x^*) \leq (g_f(x), x - x^*)$ ,  $x^* \in M^*$ .

**Теорема 3.6.** Пусть выпуклая функция  $f(x)$  обладает следующим свойством: существует число  $M > 1$  такое, что если

$$\varphi(\alpha) = f\{(1 - \alpha)x_1 + \alpha x_2\}, \quad 0 \leq \alpha \leq 1,$$

строго убывает по  $\alpha$ , то выполняется неравенство

$$(g_f(x_1), x_1 - x_2) \leq M [f(x_1) - f(x_2)]. \quad (3.22)$$

Кроме того,  $\lim_{\|x\| \rightarrow \infty} f(x) = +\infty$ .

Тогда если при применении алгоритма ОГСРП

$$\alpha_{k+1} = \frac{M+1}{M-1}; \quad h_{k+1} = \frac{2M}{M+1} \frac{f(x_k) - \bar{f}}{\|g_k\|},$$

причем  $\bar{f} \geq f^*$ , то последовательность  $\{h_k\}$  является ограниченной и для произвольного  $\varepsilon > 0$  найдется  $\bar{k}$  такое, что  $f(x_{\bar{k}}) < \bar{f} + \varepsilon$  (если на некотором шаге  $\bar{f}(x_k) < \bar{f}$ , то счет прекращается); если  $\bar{f} < f^*$ , то последовательность  $\{h_k\}$  является неограниченной.

Доказательство. Если  $\bar{f} \geq f^*$ , то рассмотрим функцию

$$\bar{f}(x) = \begin{cases} f(x), & f(x) \geq \bar{f}, \\ \bar{f}, & f(x) < \bar{f}. \end{cases}$$

Для этой функции справедливы условия теоремы 3.5, следовательно,  $\{\rho_k(x_k)\}$  ограничена, а значит, справедлива теорема 3.4 о сходимости по функционалу. Далее, из выпуклости  $\bar{f}(x)$  следует, что

$$\frac{f(x_k) - \bar{f}}{\|\tilde{\rho}_k\|} \leq \rho_k(x_k).$$

Таким образом,  $h_k \leq \frac{2M}{M+1} \rho_k(x_k)$ , т. е.  $\{h_k\}$  ограничена.

Если  $\bar{f}$  выбрано меньшим  $f^*$ , то возможны два случая:

а)  $\sup_k f(x_k) < \infty$ ; в этом случае выполняются условия

теоремы 3.1, из которой следует существование  $\{\|\tilde{g}_k^p\|\}_{p=1}^\infty$ , стремящейся к нулю:

$$\sup_k h_k = \sup_k \frac{2M}{M+1} \frac{f(x_k) - f}{\|\tilde{g}_k\|} \geq \sup_k \frac{f^* - f}{\|\tilde{g}_k\|} = +\infty;$$

б)  $\sup_k f(x_k) = +\infty$ .

Рассмотрим луч, исходящий из точки  $x_k$  и соединяющий ее с некоторой точкой минимума  $f(x)$ . Пусть  $x_k^*$  — первая точка минимума на этом луче. Применим неравенство (3.22):

$$\frac{f(x_k) - f^*}{\|g_f(x_k)\|} \geq \frac{1}{M} \frac{(g_f(x_k), x_k - x_k^*)}{\|g_f(x_k)\|} = \frac{R_k}{M},$$

где  $R_k$  — расстояние от точки  $x_k^*$  до опорной гиперплоскости к множеству Лебега  $\{x : f(x) \leq f(x_k)\}$ , проведенной через точку  $x_k$ . Так как область минимумов  $M^*$  ограничена,

а  $\{f(x_k)\}$  неограничена, то  $\sup_k R_k = +\infty$ . Далее,

$$\begin{aligned} \sup_k h_k &= \sup_k \frac{2M}{M+1} \frac{f(x_k) - \tilde{f}}{\|\tilde{g}_k\|} \gg \\ &\gg \sup_k \frac{2M}{M+1} \frac{f(x_k) - f^*}{\|g_f(x_k)\|} \gg \frac{2}{M+1} \sup_k R_k = +\infty. \end{aligned}$$

Теорема доказана.

Теорема 3.6 позволяет построить алгоритм минимизации функции  $f(x)$ , удовлетворяющей условию этой теоремы, при неизвестном  $f^*$ . Этот алгоритм будет состоять из последовательности этапов, на каждом из которых будет применяться ОГСРП в форме, указанной в теореме 3.6.

Вначале выбираем  $x_0^{(1)} \in E_n$ ;  $h > 0$ ;  $\Delta_0 > 0$ ;  $\tilde{f}_0 = f(x_0^{(1)})$ . Пусть выполнено  $r$  этапов алгоритма. Перед  $(r+1)$ -м этапом имеем

$$x_0^{(r+1)} \in E_n; \quad h > 0; \quad \Delta_r > 0; \quad \tilde{f}_r = f(x_0^{(r+1)}),$$

где  $x_0^{(r+1)}$  — точка, в которой получено рекордное значение функционала на  $r$  этапах,  $\tilde{f}_r$  — значение этого рекорда. Возможны два события:

- а)  $f(x_0^{(r+1)}) \leq f(x_0^{(r)}) - \frac{\Delta_r}{2}$ ; в этом случае  $\Delta_{r+1} = \Delta_r$ ;  
 б)  $f(x_0^{(r+1)}) > f(x_0^{(r)}) - \frac{\Delta_r}{2}$ ; в этом случае  $\Delta_{r+1} = \frac{\Delta_r}{2}$ .

В обоих случаях  $\tilde{f}_{r+1} = f(x_0^{(r+1)}) - \Delta_{r+1}$ . Приняв  $x_0 = x_0^{(r+1)}$  и  $\tilde{f} = \tilde{f}_{r+1}$ , применяем алгоритм ОГСРП в форме, описанной в теореме 3.6. На некотором шаге этого алгоритма  $k = \bar{k}_{r+1}$  произойдет одно из двух событий:

- а)  $f(x_{\bar{k}_{r+1}}) \leq f(x_0^{(r+1)}) - \frac{\Delta_{r+1}}{2}$ ;  
 б)  $h_{\bar{k}_{r+1}+1} > h$ .

На этом  $(r+1)$ -й этап прекращается, запоминается  $\min_{0 \leq k \leq \bar{k}_{r+1}} f(x_k^{(r+1)})$  и соответствующая точка  $x_{\bar{k}_{r+1}}^{(r+1)} = x_0^{(r+2)}$ .

Перейдем к  $(r+2)$ -му этапу. Покажем, что  $\lim_{r \rightarrow \infty} f(x_0^{(r)}) = f^*$ . Так как  $\{f(x_0^{(r)})\}$  является невозрастающей и ограни-

ченной снизу, то  $\lim_{r \rightarrow \infty} f(x_0^{(r)})$  существует. Из описания алгоритма и доказательства теоремы ясно, что при достаточно малом  $\Delta_r$  случай б) может произойти только тогда, когда  $\bar{f}_r < f^*$ . С другой стороны, случай а) может произойти после случая б) только конечное число раз подряд. Таким образом,  $\lim_{r \rightarrow \infty} \Delta_r = 0$ . Так как, начиная с достаточно малых  $\Delta_r$  (при фиксированном  $h$ ), если произошел случай б), то  $f(x_0^{(r)}) - \Delta_r$  служит оценкой снизу для  $f^*$  и

$$\lim_{r \rightarrow \infty} f(x_0^{(r)}) = f^*,$$

что и требовалось доказать.

При определенных условиях, например при решении систем нелинейных уравнений (см. § 5),  $f^*$  известно, но могут возникнуть трудности с оценкой  $M$  и  $N$ . При неправильном выборе констант  $M$  и  $N$  сходимость  $(M, N)$ -алгоритма ОГСРП, описанного в теореме 3.3, может оказаться слишком медленной или вообще отсутствовать. Поэтому важно иметь признаки, с помощью которых можно получить информацию о «ненормальном» ходе процесса минимизации. Наиболее простой из этих признаков основан на изучении поведения последовательности  $\{h_k\}$ .

**Теорема 3.7.** *Если параметры  $M, N$  в  $(M, N)$ -алгоритме выбраны правильно, то последовательность  $\{h_k\}$  ограничена. Если при неправильном выборе параметров  $M$  и  $N$  отсутствует сходимость, то либо последовательность  $\{h_k\}$  является неограниченной, либо последовательность  $\{f(x_k)\}$  неограничена.*

Эта теорема доказывается по той же схеме, что и теорема 3.6. Таким образом, если в процессе минимизации по схеме ОГСРП, описанной в теореме 3.3, при заданных  $M$  и  $N$   $h_k$  либо  $f(x_k)$  превзошли достаточно большие заданные числа  $h^{\max}$  и  $f^{\max}$ , то следует константу  $M$  увеличить,  $N$  уменьшить и начать процесс заново из начальной или наилучшей достигнутой точки.

## § 5. Применение метода ОГСРП к решению системы нелинейных уравнений

Описанный в теореме 3.3 способ выбора коэффициента растяжения пространства и шагового множителя находит непосредственное применение при решении системы нелинейных

уравнений

$$\psi_i(x) = 0, \quad i = 1, \dots, n, \quad x = \{t_1, \dots, t_n\}, \quad (3.23)$$

которую можно решать путем сведения к задаче минимизации негладкой функции  $f(x)$ , где

$$f(x) = \max_{1 \leq i \leq n} |\psi_i(x)|. \quad (3.24)$$

Если функции  $\psi_i(x)$ ,  $i = 1, \dots, n$ , почти дифференцируемы, то  $f(x)$  также почти дифференцируема (см. теорему 1.17) и для ее минимизации можно применить метод ОГСРП с указанным в теореме 3.3 способом регулировки  $\{\alpha_k\}$  и  $\{h_k\}$ , тем более что если решение системы уравнений (3.23) существует, то  $f^* = \min_{x \in E_n} f(x) = 0$ .

Обсудим вопрос о значениях параметров  $M$  и  $N$  неравенства (3.18) в случае функции  $f(x)$ , заданной выражением (3.24). Если мы ограничимся рассмотрением функций  $\psi_i(x)$ , почти всюду непрерывно дифференцируемых, то естественно считать типичным (регулярным) случаем ту ситуацию, когда в некоторой окрестности точки минимума функции  $\psi_i(x)$  непрерывно дифференцируемы, а якобиан  $\det J(x) = \det \left\{ \frac{\partial \psi_i}{\partial t_j} \right\}$  в точке минимума отличен от нуля.

**Теорема 3.8.** *В регулярном случае найдется для любого  $\delta > 0$  такая окрестность точки минимума  $x^*$ , что*

$$(1 - \delta) f(x) \leq (g_i(x), x - x^*) \leq (1 + \delta) f(x).$$

**Доказательство.** Покажем сначала, что производные по любому направлению в точке минимума положительны. Достаточно показать, что они отличны от нуля. Если бы это было не так, то нашлось бы направление  $\xi$ , ортогональное  $g_{\psi_i}(x^*)$ ,  $i = 1, \dots, n$ , что невозможно ввиду линейной независимости системы векторов  $\{g_{\psi_i}(x^*)\}_{i=1}^n$ .

Так как производные по направлению в точке  $x^*$   $f'_\xi(x^*)$  при  $\|\xi\| = 1$  являются значениями непрерывной функции, определенной на компактном множестве — единичной сфере, то множество этих значений замкнуто. Следовательно, найдется такое  $\delta' > 0$ , что  $f'_\xi(x^*) \geq \delta'$  для любого  $\xi$ ,  $\|\xi\| = 1$ .

Производные по направлению от функций  $\psi_i(x)$  равномерно непрерывны в некоторой окрестности точки  $x^*$ . Следовательно, найдется такая сферическая окрестность точки

$x^*$ , что

$$|\psi'_{i_{\xi}}(x) - \psi'_{i_{\xi}}(x^*)| \leq \frac{\delta\delta'}{3}.$$

Рассмотрим производную по направлению  $\mu(x) = \frac{x - x^*}{\|x - x^*\|}$  от функции  $f(x)$  в произвольной точке  $x$  этой окрестности. Эта производная удовлетворяет условию

$$f'_{\mu(x)}(x) - \frac{\delta\delta'}{3} \leq f'_{\mu(x)}(x^*) \leq f'_{\mu(x)}(x) + \delta\delta'. \quad (3.25)$$

В самом деле,

$$f'_{\mu(x)}(x) = \max_{i \in I(x)} [\psi'_{i_{\mu(x)}}(x) \operatorname{sign} \psi_i(x)],$$

где  $I(x)$  — множество индексов, на котором достигается равенство

$$f(x) = |\psi_i(x)|, \\ f'_{\mu(x)}(x^*) = |\psi'_{i_0\mu(x)}(x^*)| = \max_{1 \leq i \leq n} |\psi'_{i_{\mu(x)}}(x^*)|.$$

Если  $i_0 \in I(x)$ , то неравенство (3.25) следует непосредственно. Если  $i_0 \notin I(x)$ , то на отрезке  $\{\alpha x^* + (1 - \alpha)x, 0 \leq \alpha \leq 1\}$  найдется точка  $x_0$ , в которой

$$\psi'_{i_{\mu(x)}}(x_0) \operatorname{sign} \psi_{\bar{i}}(x_0) \geq \psi'_{i_0\mu(x)}(x_0) \operatorname{sign} \psi_{i_0}(x_0) \geq \\ \geq |\psi'_{i_0\mu(x)}(x^*)| - \frac{\delta\delta'}{3}, \quad \bar{i} \in I(x_0),$$

откуда

$$|\psi'_{i_{\mu(x)}}(x^*)| \geq |\psi'_{i_0\mu(x)}(x^*)| - \frac{2\delta\delta'}{3}, \\ f'_{\mu(x)}(x) \geq |\psi'_{\bar{i}\mu(x)}(x^*)| - \frac{\delta\delta'}{3} \geq f'_{\mu(x)}(x^*) - \delta\delta'.$$

Теорема доказана.

Из доказанной теоремы следует, что в регулярном случае, находясь в достаточно малой окрестности точки  $x^*$ , мы можем работать с большими  $\alpha$  и  $\gamma = \frac{2MN}{M+N}$ , близкими к единице, и получить довольно быстро сходящийся процесс. В самом деле если  $N = 1 - \delta$ ,  $M = 1 + \delta$ , то  $\alpha$  можно выбрать равным  $\frac{M+N}{M-N} = \frac{1}{\delta}$  и  $\gamma = \frac{2MN}{M+N} = 1 - \delta^2$ . При  $\delta \rightarrow 0$ ,  $\gamma \rightarrow 1$ ,  $\alpha \rightarrow \infty$ . Если же начальное приближение  $x_0$  выбрано неудачно, оценка параметров  $M$  и  $N$  затруднительна. В этом

случае полезна процедура, которая бы корректировала в процессе работы алгоритма неудачные априорные оценки  $M$  и  $N$ . Если  $M$  и  $N$  выбраны неверно, то в соответствии с результатами § 4 можем обнаружить это, исследуя поведение параметра  $\{h_k\}$ . Если  $h_k$  превзойдет заданное нами достаточно большое число (в качестве такого числа можно брать оценку сверху расстояния до точки минимума от  $x_0$ ), то нам придется изменить  $M$  и  $N$  (увеличить  $M$  и уменьшить  $N$ ), вычислить новые  $\gamma$  и  $\alpha$  и продолжить счет, взяв в качестве начальной точку, на которой  $f(x)$  принимало наименьшее значение на предыдущих итерациях.

Следует отметить, что сведение решения системы нелинейных уравнений (3.23) к задаче минимаксного типа является более эффективным по сравнению с другими способами сведения к оптимизационной задаче (например, к минимизации суммы квадратов невязок или суммы абсолютных величин невязок) при использовании алгоритма ОГСРП, так как в случае задачи минимизации функций максимума на каждой итерации нужно вычислять градиент (почти-градиент) только от одной из функций  $\psi_i(x)$ , а не от всех, как при других способах сведения. Кроме того, предельный вариант метода ОГСРП при  $\beta = 0$  можно рассматривать как метод ортогонализации градиентов, который при определенных условиях при решении систем нелинейных уравнений сходится с квадратичной скоростью. Рассмотрим этот метод подробнее. Пусть задана нелинейная система уравнений

$$\psi_i(x) = 0; \quad i = 1, 2, \dots, n; \quad x \in E_n, \quad (3.26)$$

где  $\psi_i(x) = \psi_i(t_1, t_2, \dots, t_n)$  — непрерывно дифференцируемые функции. Рассмотрим процесс нахождения решения этой системы, состоящий из последовательности этапов, каждый из которых в свою очередь состоит из  $n$  последовательных шагов. Пусть начальное для каждого  $r$ -го этапа приближение будет  $x_0^{(r)}$ .

**1-й шаг.** Вычисляем  $\psi_1(x_0^{(r)})$  и градиент  $g_{\psi_1}(x_0^{(r)})$  и находим

$$x_1^{(r)} = x_0^{(r)} - \frac{\psi_1(x_0^{(r)})}{\|g_{\psi_1}(x_0^{(r)})\|^2} g_{\psi_1}(x_0^{(r)}). \quad (3.27a)$$

**$(k+1)$ -й шаг** ( $1 \leq k < n$ ). Вычисляем  $\psi_{k+1}(x_k^{(r)})$  и  $g_{\psi_{k+1}}(x_k^{(r)})$  затем вектор  $\varphi_{k+1}^{(r)}$  — проекцию вектора  $g_{\psi_{k+1}}(x_k^{(r)})$

на подпространство, ортогональное векторам  $g_{\psi_1}(x_0^{(r)})$ , ...,  $g_{\psi_k}(x_{k-1}^{(r)})$ , и находим

$$x_{k+1}^{(r)} = x_k^{(r)} - \frac{\psi_{k+1}(x_k^{(r)})}{\|\psi_{k+1}^{(r)}\|^2} \Phi_{k+1}^{(r)}. \quad (3.276)$$

Полученное таким образом  $x_n^{(r)}$  служит начальным приближением для следующего этапа:  $x_{ii}^{(r)} = x_0^{(r+1)}$ .

Покажем, что если начальное приближение  $x_0^{(1)}$  выбрано достаточно близким к решению системы  $x^*$ , то при определенных условиях последовательность  $x_0^{(1)}$ , ...,  $x_0^{(r)}$ , ... будет сходиться с квадратической скоростью к точке  $x^*$ . Без ограничения общности предполагаем, что  $x^* = 0$ .

**Теорема 3.9.** Пусть система уравнений (3.26) имеет решение  $x^* = 0$ . Функции  $\psi_i(x)$ ,  $i = 1, 2, \dots, n$ , непрерывно дифференцируемы и их градиенты  $g_{\psi_i}(x)$  удовлетворяют условию Липшица в некоторой области  $S_\delta = \{x : \|x\| < \delta\}$

$$\|g_{\psi_i}(x') - g_{\psi_i}(x'')\| \leq L \|x' - x''\|; \quad x', x'' \in S_\delta. \quad (3.28)$$

Кроме того,  $\{g_{\psi_i}(0)\}_{i=1}^n$ ,  $i = 1, 2, \dots, n$ , образуют систему линейно независимых векторов, так что якобиан

$$\det J(x) = \det \left\{ \frac{\partial \psi_i(x)}{\partial t_j} \right\}, \quad i, j = 1, 2, \dots, n,$$

отличен от нуля в точке 0. Тогда найдется такое  $\varepsilon > 0$ , что если  $\|x_0^{(r)}\| \leq \varepsilon$ , то

$$\|x_n^{(r)}\| = \|x_0^{(r+1)}\| \leq c \|x_0^{(r)}\|^2,$$

где  $c$  — некоторое положительное число.

**Доказательство.** Рассмотрим функцию

$$D(x_0, \dots, x_{n-1}) = \det \left\{ \frac{\partial \psi_i(x_{i-1})}{\partial t_j} \right\}, \quad i, j = 1, 2, \dots, n.$$

Эта функция непрерывна, причем  $D(0, 0, \dots, 0) = \det J(0) \neq 0$ . Следовательно, найдутся такие окрестность  $S_\delta'$  точки 0 (выберем  $\delta' \leq \delta$ ) и число  $a > 0$ , что

$$|D(x_0, \dots, x_{n-1})| \geq a \text{ при } x_i \in S_\delta', \quad i = 0, 1, \dots, n-1.$$

Пусть  $x_0^{(r)}$ ,  $x_1^{(r)}$ , ...,  $x_{n-1}^{(k)} \in S_\delta'$ . Тогда найдется такое  $b > 0$ , что

$$\|\Phi_{k+1}^{(r)}\| > b, \quad k = 0, 1, \dots, n-1. \quad (3.29)$$

В самом деле, система векторов  $\{\varphi_{k+1}^{(r)}\}$ ,  $k = 0, 1, \dots, n-1$ , получается в результате процесса ортогонализации системы линейно независимых векторов  $\{g_{\psi_{k+1}}(x_k^{(r)})\}$ . Определитель Грамма системы векторов  $\{\varphi_{k+1}^{(r)}\}$  равен определителю Грамма системы векторов  $\{g_{\psi_{k+1}}(x_k^{(r)})\}$ , который в свою очередь равен

$$D^2(x_0^{(r)}, \dots, x_{n-1}^{(r)}) \geq a^2.$$

Получаем неравенство

$$\prod_{k=0}^{n-1} \|\varphi_{k+1}^{(r)}\| \geq a. \quad (3.30)$$

Так как  $\|\varphi_{k+1}^{(r)}\|$ ,  $k = 0, 1, \dots, n-1$ , ограничены сверху, то из неравенства (3.30) следует, что они ограничены снизу некоторым положительным числом, которое мы обозначим  $b$ . Далее,

$$\begin{aligned} \psi_{k+1}(x_{k+1}^{(r)}) &= \psi_{k+1}\left[x_k^{(r)} - \frac{\psi_{k+1}(x_k^{(r)})}{\|\varphi_{k+1}^{(r)}\|^2} \varphi_{k+1}^{(r)}\right] = \\ &= \psi_{k+1}(x_k^{(r)}) - \left(g_{\psi_{k+1}}^c, \frac{\varphi_{k+1}^{(r)}}{\|\varphi_{k+1}^{(r)}\|}\right) \frac{\psi_{k+1}(x_k^{(r)})}{\|\varphi_{k+1}^{(r)}\|}, \end{aligned} \quad (3.31)$$

$$k = 0, 1, \dots, n-1,$$

где  $g_{\psi_{k+1}}^c$  — значение градиента в некоторой точке, лежащей на отрезке, соединяющем  $x_k^{(r)}$  и  $x_{k+1}^{(r)}$ . Используя условия (3.28), получаем

$$g_{\psi_{k+1}}^c = g_{\psi_{k+1}}(x_k^{(r)}) + \Delta g_{k+1}, \quad (3.32)$$

где

$$\|\Delta g_{k+1}\| \leq \frac{L\psi_{k+1}(x_k^{(r)})}{\|\varphi_{k+1}^{(r)}\|}.$$

Подставляя в (3.31) правую часть (3.32) и учитывая, что

$$(g_{\psi_{k+1}}(x_k^{(r)}), \varphi_{k+1}^{(r)}) = \|\varphi_{k+1}^{(r)}\|^2,$$

получаем

$$\begin{aligned} |\psi_{k+1}(x_{k+1}^{(r)})| &\leq \frac{L\psi_{k+1}^2(x_k^{(r)})}{\|\varphi_{k+1}^{(r)}\|^2} \leq \\ &\leq \frac{L}{b^2} \psi_{k+1}^2(x_k^{(r)}), \quad k = 0, 1, \dots, n-1. \end{aligned} \quad (3.33)$$

Для дальнейшего проведения доказательства нам требуется следующая лемма.

**Лемма 3.1.** Пусть  $x_1$  и  $x_2$  — две точки,  $\xi$  — вектор такой, что  $\|\xi\| = 1$  и  $(g_{\psi_i}(x_1), \xi) = 0$  для некоторого  $i$ ,  $1 \leq i \leq n$ ,  $h$  — некоторое положительное число. Тогда если  $x_1, x_2, x_2 + h\xi \in S_\delta$  и  $\|x_1 - x_2\| = d$ , то

$$|\psi_i(x_2 + h\xi) - \psi_i(x_2)| \leq Lh(h + d).$$

**Доказательство.** Применяя теорему Лагранжа, получаем

$$|\psi_i(x_2 + h\xi) - \psi_i(x_2)| \leq h |(g_{\psi_i}(x_2 + \alpha h\xi), \xi)|,$$

где  $0 \leq \alpha \leq 1$ ,

$$\begin{aligned} |(g_{\psi_i}(x_2 + \alpha h\xi), \xi)| &= |(g_{\psi_i}(x_2 + \alpha h\xi) - g_{\psi_i}(x_1), \xi)| \leq \\ &\leq \|g_{\psi_i}(x_2 + \alpha h\xi) - g_{\psi_i}(x_1)\| \leq L(h + d). \end{aligned}$$

Отсюда

$$|\psi_i(x_2 + h\xi) - \psi_i(x_2)| \leq Lh(h + d),$$

что и требовалось доказать.

Продолжим доказательство теоремы. Предположим, что система точек  $x_0^{(r)}, \dots, x_n^{(r)}$  содержится в сфере  $S_\mu = \{x : \|x\| \leq \mu\}$  ( $\mu \leq \delta^r$ ). При этом

$$|\psi_i(x_k^{(r)})| \leq \mu \|g_{\psi_i}(0)\| + \frac{L\mu^2}{2}, \quad (3.34)$$

$$i = 1, \dots, n; \quad k = 1, \dots, n.$$

Отсюда, используя (3.33), получаем

$$|\psi_{k+1}(x_{k+1}^{(r)})| \leq \frac{L\mu^2}{b^2} \left[ \|g_{\psi_{k+1}}(0)\| + \frac{L\mu}{2} \right]^2.$$

Применяя лемму 3.1, приходим к неравенству

$$\begin{aligned} |\psi_{k+1}(x_n^{(r)})| &\leq |\psi_{k+1}(x_{k+1}^{(r)})| + \\ &+ \sum_{i=1}^{n-k-1} |\psi_{k+1}(x_{k+i+1}^{(r)}) - \psi_{k+1}(x_{k+i}^{(r)})| \leq \\ &\leq \frac{L\mu^2}{b^2} \left[ \|g_{\psi_{k+1}}(0)\| + \frac{L\mu}{2} \right]^2 + \sum_{i=1}^{n-k-1} L \cdot 2\mu \cdot 4\mu = \\ &= \bar{c}\mu^2 + o(\mu^2), \end{aligned} \quad (3.35)$$

где  $\bar{c}$  — некоторое число, выражение для которого легко выписать.

Теперь для окончания доказательства теоремы нам осталось проверить следующее утверждение: найдется такое число  $\varepsilon$  и две положительные константы  $c_1$  и  $c_2$ , что если  $\|x_0^{(r)}\| \leq \varepsilon$ , то

$$1) \max_{1 \leq k \leq n} \|x_k^{(r)}\| \leq c_1 \varepsilon; \quad (3.36)$$

$$2) f(x_n^{(r)}) = \max_{0 \leq k \leq n-1} |\psi_{k+1}(x_n^{(r)})| \geq c_2 \|x_n^{(r)}\|. \quad (3.37)$$

Первая часть этого утверждения проверяется путем непосредственной оценки норм приращений  $\|x_{k+1}^{(r)} - x_k^{(r)}\|$ ,  $k = 0, 1, \dots, n-1$ , с использованием формул (3.27) и оценок (3.29), (3.34), вторая часть является следствием первой и того, что функция  $f(x)$  в некоторой окрестности нуля имеет ограниченные снизу положительным числом производные по направлению

$$f'_{\frac{x}{\|x\|}}(x) \geq d > 0.$$

Пусть  $x_0^{(r)} = \varepsilon \leq \frac{\delta'}{c_1}$ . Тогда, воспользовавшись (3.35) — (3.37), получаем

$$\max_{0 \leq k \leq n-1} |\psi_{k+1}(x_n^{(r)})| \leq k_1 c_1^2 \varepsilon^2,$$

откуда

$$\|x_n^{(r)}\| \leq \frac{k_1 c_1^2}{c_2} \varepsilon^2 = c \|x_0^{(r)}\|^2,$$

где  $c = k_1 c_1^2 / c_2$ , что и требовалось доказать.

## § 6. Метод минимизации, использующий операцию растяжения пространства в направлении разности двух последовательных почти-градиентов

Серьезной проблемой при построении некоторых модификаций ОГСРП-алгоритмов является выбор на каждой итерации шагового множителя. Этот вопрос сравнительно просто решается в том случае, когда заранее известно значение функции в точке минимума. В более общем случае приходится усложнять алгоритм.

Изложим другой класс алгоритмов, основанный на использовании операции растяжения пространства в направлении разности двух последовательных почти-градиентов. В отличие от алгоритмов типа ОГСРП этот класс алгоритмов позволяет применить способ выбора шагового множителя, близкий к применяемому в методе наискорейшего спуска. При этом получаются алгоритмы, обеспечивающие монотонность или, в некотором смысле, почти монотонность процесса спуска. Алгоритмы этого типа будем называть  $r$ -алгоритмами. Результаты, изложенные в настоящем параграфе, основаны на работах автора и Н. Г. Журбенко [65, 67]. Применительно к минимаксным задачам  $r$ -алгоритмы подробно исследованы Л. П. Шабашовой [51, 69].

Перейдем к описанию  $r$ -алгоритмов.

Пусть  $f(x)$  — почти дифференцируемая функция, определенная в  $n$ -мерном евклидовом пространстве  $E_n$ , и выполняется условие

$$\lim_{\|x\| \rightarrow +\infty} f(x) = +\infty,$$

где  $g_f(x)$  — почти-градиент функции  $f(x)$  в точке  $x$ . Рассмотрим следующий итеративный процесс минимизации  $f(x)$ .

Первый шаг задается начальным значением  $x_0$ ; вычисляем  $g_f(x_0)$ , выбираем  $h_0 > 0$  и находим  $x_1 = x_0 - h_0 g_f(x_0)$ . Принимаем  $g_f(x_0) = \tilde{g}_1$ ,  $B = I$  (единичная матрица). Переходим ко второму шагу.

Пусть в результате вычислений после  $k$  ( $k = 1, 2, \dots$ ) шагов процесса получены определенные значения векторов  $x_k$ ,  $\tilde{g}_k \in E_n$  и матрицы  $B_k$ . Опишем  $(k+1)$ -й шаг процесса.

Вычисляем следующие величины:

1)  $g_f(x_k)$  — почти-градиент функции  $f(x)$  в точке  $x_k$ ;  
 2)  $g_k^* = B_k^* g_f(x_k)$ , где  $g_k^*$  — почти-градиент функции  $\varphi_k(y) = f(B_k y)$  в точке  $y = A_k x_k$ ,  $A_k = B_k^{-1}$ ;

$$3) r_k = g_k^* - \tilde{g}_k, \quad (3.38)$$

где  $r_k$  — разность двух градиентов от функции  $\varphi_k(y)$ , вычисленных в точках  $y_k = A_k x_k$  и  $\tilde{y}_k = A_k x_{k-1}$ ;  $r_k$  можно вычислить и другим способом при  $k = 1, 2, \dots$ :

$$r_k = B_k^* [g_f(x_k) - g_f(x_{k-1})],$$

однако использование формулы (3.38) дает некоторую экономию вычислений в целом;

$$4) \xi_{k+1} = \frac{r_k}{\|r_k\|};$$

нормировка вектора  $r_k$  нужна для подготовки очередной операции растяжения пространства в направлении, определяемом вектором  $r_k$ ;

5)  $\beta_{k+1}$  — величина, обратная коэффициенту растяжения пространства на  $(k+1)$ -м шаге;

6)  $B_{k+1} = B_k R_{\beta_{k+1}}(\xi_{k+1})$ , где  $B_{k+1} = A_{k+1}^{-1}$ ,  $A_{k+1} = R_\alpha(\xi_{k+1}) A_k$  — матрица преобразования пространства после  $(k+1)$ -го шага;  $R_\alpha(\xi)$  — оператор растяжения пространства в направлении  $\xi$  с коэффициентом  $\alpha$  ( $\alpha = \frac{1}{\beta}$ ;  $\|\xi\| = 1$ );

7)  $h_{k+1}$  — шаговый множитель в формуле спуска в заданном направлении;

$$8) \tilde{g}_{k+1} = R_{\beta_{k+1}}(\xi_{k+1}) g_k^*,$$

$$\begin{aligned} \tilde{g}_{k+1} &= R_{\beta_{k+1}}(\xi_{k+1}) g_k^* = R_{\beta_{k+1}}(\xi_{k+1}) B_k^* g_f(x_k) = \\ &= (B_k \cdot R_{\beta_{k+1}}(\xi_{k+1}))^* g_f(x_k) = B_{k+1}^* g_f(x_k), \end{aligned}$$

где  $\tilde{g}_{k+1}$  — значение градиента функции  $\varphi_{k+1}(y) = f(B_{k+1}y)$  в точке  $\tilde{y}_{k+1} = A_{k+1}x_k$ ;

$$9) x_{k+1} = x_k - h_{k+1} B_{k+1} \tilde{g}_{k+1}; \quad (3.39)$$

применив к обеим частям формулы (3.39) оператор  $A_{k+1}$ , получим

$$y_{k+1} = A_{k+1}x_{k+1} = A_{k+1}x_k - h_{k+1} \tilde{g}_{k+1} = \tilde{y}_{k+1} - h_{k+1} \tilde{g}_{k+1}.$$

Таким образом, формула (3.39) фактически реализует шаг градиентного спуска для функции  $\varphi_{k+1}(y)$ ;

10) переход к  $(k+2)$ -му шагу с запоминанием  $x_{k+1}$ ,  $\tilde{g}_{k+1}$  и  $B_{k+1}$  или окончание работы алгоритма при выполнении некоторого критерия остановки.

Приведенное описание класса алгоритмов градиентного типа с изменяемой метрикой порождает конкретные алгоритмы при уточнении способа выбора последовательностей

$\{h_{k+1}\}$  и  $\{\beta_{k+1}\}$  и критерия остановки. Такие алгоритмы будем называть  $r$ -алгоритмами.

Имеется еще одна возможность модификации  $r$ -алгоритмов — использование так называемой операции восстановления, когда после заданного числа итераций периодически происходит «восстановление» матрицы  $B_k$ , т. е. она заменяется единичной матрицей. Вопрос о сходимости алгоритмов с монотонным изменением значения минимизирующей функции и восстановлением решается сравнительно просто, так как он фактически сводится к вопросу о сходимости обычного градиентного спуска без изменения метрики, который подробно исследован в некоторых работах (см., например, [40]). В частности, если  $f(x)$  — непрерывно дифференцируема и  $h_{k+1}$  в форме (3.39) выбирается из условия минимума по направлению, то при использовании  $r$ -алгоритмов с восстановлением множество предельных точек последовательности  $\{x_k\}$  состоит из стационарных точек функции  $f(x)$  с одинаковым значением  $f(x)$ . Доказательство этого факта практически ничем не отличается от доказательства аналогичного результата для метода наискорейшего спуска. В случае, когда восстановление не предполагается, вопрос о сходимости  $r$ -алгоритма требует специальных исследований.

Рассмотрим более подробно один предельный вариант  $r$ -алгоритма с восстановлением, в котором  $\beta_k$ ,  $k = 1, 2, \dots$ , выбирается равным нулю, а  $h_{k+1}$  определяется из условия минимума выражения  $f(x_k - h_{k+1} B_{k+1} \tilde{g}_{k+1})$ . При этом, как легко заметить, либо  $g_{k^*+1} = 0$  после некоторого числа  $k^* < n$  шагов, либо  $B_{n+1} = 0$ . В самом деле, оператор  $R_0(\eta)$  означает проектирование на подпространство, ортогональное вектору  $\eta$ . Произведение операторов  $\prod_{i=1}^k R_0(\eta_i)$  не зависит от порядка сомножителей, является самосопряженным оператором и осуществляет проекцию на подпространство, являющееся ортогональным дополнением к линейной оболочке векторов  $\eta_i$ ,  $i = 1, \dots, k$ .

Запишем цепочку равенств

$$r_1 = g_f(x_1) - g_f(x_0);$$

$$\dots \dots \dots$$

$$r_k = R_0\left(\frac{r_{k-1}}{\|r_{k-1}\|}\right) R_0\left(\frac{r_{k-2}}{\|r_{k-2}\|}\right) \dots R_0\left(\frac{r_1}{\|r_1\|}\right) \Delta g_k;$$

$$\Delta g_k = g_f(x_k) - g_f(x_{k-1}), \quad k > 1.$$

Видим, что векторы  $r_1, \dots, r_k$  получаются в результате процесса последовательной ортогонализации, примененной к векторам  $g_f(x_0) - g_f(x_1), g_f(x_1) - g_f(x_2), \dots, g_f(x_{k-1}) - g_f(x_k)$ ,  $k = 1, 2, \dots$ . Если  $r_1, \dots, r_n$  отличны от нуля, то оператор  $B_{n+1} = 0$ , так как он является оператором проектирования на подпространство, ортогонально дополняющее все пространство. Если для некоторого  $k \leq n$   $r_k = 0$ , это означает, что

$$B_{k-1}^* [g_f(x_k) - g_f(x_{k-1})] = 0,$$

что возможно лишь при условии  $B_k^* g_f(x_{k-1}) = \tilde{g}_k = 0$ ,  $x_k = x_{k-1}$ .

Таким образом, в предлагаемом варианте  $r$ -алгоритма с необходимостью нужно применять восстановление при условии  $B_k^* g_f(x_{k-1}) = 0$ . Этот вариант можно рассматривать как модификацию метода сопряженных градиентов. Легко показать, что для неотрицательно определенной квадратичной формы решение получается за число шагов, не превышающее  $n$ . Для произвольной достаточно гладкой функции можно доказать теорему о квадратичной скорости сходимости.

**Теорема 3.10.** Пусть функция  $f(x)$ , определенная в  $E_n$ , дважды непрерывно дифференцируема в некоторой окрестности  $S$  точки минимума  $x^*$ , причем в этой окрестности матрица вторых производных (гессиан)  $H(x)$  удовлетворяет условию Литшица

$$\|H(x) - H(x')\| \leq L \|x - x'\|; \quad x, x' \in S. \quad (3.40)$$

Кроме того,  $H(x^*)$  — положительно определенная матрица. Тогда найдется такая окрестность точки  $x^*$   $S' \subseteq S$ , что если  $x_0 \in S'$ , то найдется такое число  $c > 0$ , что

$$\|x_n - x^*\| \leq c \|x_0 - x^*\|^2,$$

где  $x_n$  — точка, получаемая после  $n$  шагов приведенного выше алгоритма (если для некоторого  $\bar{k} < n$   $\tilde{g}_{\bar{k}} = 0$ , то примем  $x_n = x_{\bar{k}}$ ).

**Доказательство.** Без ограничения общности будем считать  $x^* = 0$  и  $f(0) = 0$ . Доказательство проведем методом индукции по размерности пространства  $E_n$ . При  $n = 1$  теорема доказывается тривиально. Пусть теорема справедлива для  $n = p$ . Докажем ее справедливость для  $n = p + 1$  ( $p \geq 1$ ). Обозначим максимальное собственное

число оператора  $H(0)$  символом  $M_0$ , минимальное — символом  $m_0$ . В силу положительной определенности  $H(0)$   $m_0 > 0$ . Вследствие непрерывности гессиана  $H(x)$  найдутся такие окрестность нуля  $S_\delta = \{x : \|x\| < \delta\}$  и положительные числа  $M$  и  $m$ , что при  $x \in S_\delta$  максимальное собственное число оператора  $H(x)$  не превышает  $M$ , а минимальное не меньше  $m$  ( $m > 0$ ). Пусть  $e$  — вектор,  $\|e\| = 1$ . Примем  $x_0 = \varepsilon e$ ,  $0 < \varepsilon < \delta$ ,

$$g_f(x_0) = \int_0^\varepsilon H(\mu e) e d\mu = \int_0^\varepsilon [H(0) + G(\mu e)] e d\mu.$$

Используя условие (3.40), получаем  $\|G(\mu e)\| \leq \mu L$ , откуда

$$g_f(x_0) = \varepsilon [H(0)e + r(x_0)], \quad (3.41)$$

причем

$$\|r(x_0)\| \leq \frac{1}{\varepsilon} \int_0^\varepsilon \mu L d\mu = \frac{L\varepsilon}{2}. \quad (3.42)$$

Определим  $x_1$  в соответствии с исследуемым алгоритмом:

$$x_1 = x_0 - h^* \frac{g_f(x_0)}{\|g_f(x_0)\|},$$

где  $h^*$  — минимальный положительный корень уравнения

$$\left( g_f(x_0), g_f\left(x_0 - h^* \frac{g_f(x_0)}{\|g_f(x_0)\|}\right) \right) = 0.$$

Легко видеть, что при достаточно малом  $\varepsilon$

$$h^* \leq \frac{\|g_f(x_0)\|}{m} \leq \frac{M\varepsilon}{m}.$$

В самом деле, пусть

$$x(h) = x_0 - h \frac{g_f(x_0)}{\|g_f(x_0)\|}, \quad \varepsilon < \frac{\delta}{\frac{M}{m} + 1}.$$

Рассмотрим функцию  $\varphi(h) = f(x(h))$ . При  $0 \leq h \leq \frac{M\varepsilon}{m}$   $x(h) \in S_\delta$ . Тогда в силу свойств гессиана  $H(x)$  в области  $S_\delta$

$$\frac{d^2\varphi(h)}{dh^2} \geq m,$$

откуда

$$\frac{d\varphi}{dh} \leq \|g_f(x_0)\| - mh \leq M\varepsilon - mh.$$

Значит,  $h^* \leq \frac{M\varepsilon}{m}$ , что и требовалось доказать.

Далее,

$$\begin{aligned} g_f(x_1) &= g_f\left(x_0 - h^* \frac{g_f(x_0)}{\|g_f(x_0)\|}\right) = \\ &= g_f(x_0) - \int_0^{h^*} H\left(x_0 - \mu \frac{g_f(x_0)}{\|g_f(x_0)\|}\right) \frac{g_f(x_0)}{\|g_f(x_0)\|} d\mu = \\ &= g_f(x_0) = \int_0^{h^*} \left[ H(x_0) + G_1\left(\mu \frac{g_f(x_0)}{\|g_f(x_0)\|}\right) \right] \frac{g_f(x_0)}{\|g_f(x_0)\|} d\mu, \end{aligned}$$

причем  $\|G_1\left(\mu \frac{g_f(x_0)}{\|g_f(x_0)\|}\right)\| \leq L\mu$ . Отсюда

$$\begin{aligned} g_f(x_1) &= g_f\left(x_0 - h^* \frac{g_f(x_0)}{\|g_f(x_0)\|}\right) = \\ &= g_f(x_0) - h^* H(x_0) \frac{g_f(x_0)}{\|g_f(x_0)\|} + h^* r_1(x_0), \end{aligned} \quad (3.43)$$

причем

$$\|r_1(x_0)\| \leq \frac{Lh^*}{2} \leq \frac{LM\varepsilon}{2m}. \quad (3.44)$$

Используя (3.41) — (3.44), получаем

$$\begin{aligned} (g_f(x_0), g_f(x_1)) &= \varepsilon^2 [(H(0)e + r(x_0), H(0)e + r(x_0)) - \\ &\quad - \frac{h}{\|g_f(x_0)\|} ((H(0) + G(x_0))(H(0)e + r(x_0)), \\ &\quad H(0)e + r(x_0))] + \varepsilon h^* (H(0)e + r(x_0), r_1(x_0)) = \\ &= \varepsilon^2 [(H(0), H(0)e) - \frac{h^*}{\|g_f(x_0)\|} (H^2(0)e, H(0)e)] + r^2(x_0), \end{aligned}$$

при этом при достаточно малом  $\varepsilon$ , как легко видеть,  $|r_2(x_0)| < d\varepsilon^3$ , где  $d$  — некоторое положительное число. Положив  $(g_f(x_0), g_f(x_1)) = 0$ , получаем

$$1 - \frac{d\varepsilon}{\|H(0)e\|^2} q \leq h^* \leq 1 + \frac{d\varepsilon}{\|H(0)e\|^2} q,$$

где

$$q = \frac{\|H(0)e\|^2 \|g_f(x_0)\|}{(H^2(0)e, H(0)e)}.$$

Теперь покажем, что расстояние от начала координат до гиперплоскости, проходящей через точку  $x_1$  и с нормалью, задаваемой вектором  $\Delta = g_f(x_0) - g_f(x_1)$ , имеет порядок  $\varepsilon^2$ . Запишем уравнение этой гиперплоскости  $(x - x_1, g_f(x_0) - g_f(x_1)) = 0$ . Расстояние от начала координат до гиперплоскости вычисляется по формуле

$$S = \frac{|(g_f(x_0) - g_f(x_1), x_1)|}{\|g_f(x_0) - g_f(x_1)\|};$$

$$x_1 = x_0 - \frac{g_f(x_0)}{\|g_f(x_0)\|} = x_0 \frac{\|H(0)e\|^2 g_f(x_0)}{(H^2(0)e, H(0)e)} - \frac{vqg_f(x_0)}{\|g_f(x_0)\|}, \quad (3.45)$$

где

$$-\frac{de}{\|H(0)e\|^2} \leq v \leq \frac{de}{\|H(0)e\|^2}. \quad (3.46)$$

Далее, используя (3.43) — (3.46), получаем

$$\begin{aligned} (g_f(x_0) - g_f(x_1), x_1) &= \frac{h^* \varepsilon^2}{\|g_f(x_0)\|} ((H(0) + G(x_0))(H(0)e + \\ &+ r(x_0)), e - \frac{\|H(0)e\|^2 (H(0)e + r(x_0))}{(H^2(0)e, H(0)e)}) - \\ &- \frac{1}{\|g_f(x_0)\|} (vqg_f(x_0), g_f(x_0) - g_f(x_1)) - (h^* r_1(x_0), x_1) = \\ &= \frac{\varepsilon^2 h^*}{\|g_f(x_0)\|} [(H^2(0)e, e) - (H(0)e, H(0)e)] + r_3(x_0) = r_3(x_0), \end{aligned}$$

где  $|r_3(x_0)| \leq c_1 \varepsilon^3$ ,  $c_1$  — положительное число, которое легко вычисляется с помощью оценок (3.42), (3.44), (3.46). Так как  $(g_f(x_0), g_f(x_1)) = 0$ , то

$$\|g_f(x_1) - g_f(x_0)\| \geq \|g_f(x_0)\| \geq m\varepsilon.$$

Отсюда

$$S \leq \frac{c_1}{m} \varepsilon^2 = c_2 \varepsilon^2, \quad (3.47)$$

где  $c_2 = c_1/m$ . Рассмотрим функцию  $\psi(x)$ , определенную в  $p$ -мерной гиперплоскости

$$L(x_0) = \{x : (g_f(x_1) - g_f(x_0), x - x_1) = 0\}$$

и равную  $f(x)$  для точек этой гиперплоскости. Покажем, что  $\min \psi(x) = \min_{x \in L(x_0)} f(x)$  достигается в точке  $x_1^*$  такой, что

$$\|x_1^*\| \leq c_3 \varepsilon^2,$$

где  $c_3$  — положительная величина, не зависящая от  $\varepsilon$  при достаточно малом  $\varepsilon$ .

Рассмотрим точку  $\bar{x}_1$ , лежащую в  $L(x_0)$  на кратчайшем расстоянии от начала координат. Из (3.47) получаем  $S = \| \bar{x}_1 \| \leq c_2 \varepsilon^2$ , откуда

$$\min_{x \in L(x_0)} f(x) \leq f(\bar{x}_2) \leq \frac{M \varepsilon^4 c_2^2}{2} = c_4 \varepsilon^4.$$

С другой стороны,

$$f(x_1^*) \geq \frac{m \| x_1^* \|^2}{2}.$$

В результате

$$\frac{m \| x_1^* \|^2}{2} \leq c_4 \varepsilon^4; \quad \| x_1^* \| \leq \sqrt{\frac{2c_4^2}{m}} \varepsilon^2 = c_3 \varepsilon^2.$$

Заметим, что  $c_3$  определяется константами  $L$ ,  $m$  и  $M$ . Так как дальнейший спуск будет проходить в гиперплоскости  $L(x_0)$  размерности  $p$ , а по предположению индукции рассматриваемая теорема справедлива для функции, определенной в  $p$ -мерном пространстве, то справедливо соотношение

$$\| x_n \| = \| x_{p+1} \| \leq \| x_1^* \| + \| x_{p+1} - x_1^* \| \leq c_3 \varepsilon^2 + c_p \| x_1 - x_1^* \|^2,$$

где  $c_p$  — число, определяемое по предположению индукции только константами  $L$ ,  $M$ ,  $m$  и  $p$  при достаточно малом значении  $\| x_1 - x_1^* \|$ . Оценим  $\| x_1 - x_1^* \|$ .

Так как  $f(x_0) > f(x_1) \geq f(x_1^*) \geq f(0) = 0$ , то

$$f(x_1) - f(x_1^*) < f(x_0) \leq \frac{M \varepsilon^2}{2}.$$

С другой стороны,

$$\frac{m \| x_1 - x_1^* \|^2}{2} \leq f(x_1) - f(x_1^*).$$

Получаем

$$\| x_1 - x_1^* \| \leq \frac{M}{m} \varepsilon^2.$$

Окончательно имеем

$$\| x_n \| \leq \left( c_3 + c_p \frac{M}{m} \right) \varepsilon^2 = c \| x_0 \|^2,$$

где  $c$  — величина, зависящая только от  $L$ ,  $m$ ,  $M$  и  $n$ . Доказательство теоремы закончено.

Приведем вариант  $r$ -алгоритма без восстановления. Хотя строгие результаты о скорости сходимости пока не получены, численные эксперименты показывают высокую скорость сходимости по сравнению с достигаемой в различных модификациях алгоритмов с изменяемой метрикой или метода сопряженных градиентов [40]. Существенная особенность, однако, состоит в том, что определенные модификации  $r$ -алгоритмов применимы для минимизации негладких функций, в то время как другие методы, обладающие ускоренной сходимостью, требуют во всяком случае непрерывности градиента.

$r$ -алгоритм с постоянным коэффициентом растяжения был реализован на БЭСМ-6. Опишем кратко программу, остановившись лишь на алгоритме одномерного поиска минимума функции  $f(x)$  в направлении  $\eta$ , с помощью которого осуществляется выбор шага  $h_{k+1}$ . Для реализации наискорейшего спуска с высокой точностью необходимо большое количество вычислений значений функции. В разработанном алгоритме поиск минимума по направлению производится таким образом, что точность поиска увеличивается в процессе счета.

Во всех приведенных ниже числовых примерах одномерный поиск минимума производится следующим образом.

Выбираем числа  $\mu \geq 1$ ,  $\gamma$ ,  $0 < \gamma < 1$ , целое число  $L$  и начальный пробный шаг  $h$ .

На  $(k + 1)$ -й итерации определены  $x_k$  — точка, полученная на  $k$ -й итерации;  $\eta$  — вектор, в направлении которого необходимо осуществить спуск из точки  $x_k$ ;  $h$  — пробный шаг.

Из точки  $x_k$  делаем шаг длиной  $h$  в направлении  $\eta$ :  $z_1 = x_k + h\eta$ . Вычисляем  $f(z_1)$ . Если  $f(z_1) < f(x_k)$ , то из точки  $z_1$  снова делаем шаг длиной  $h$  в направлении  $\eta$ :  $z_2 = z_1 + h\eta$ . Вычисляем  $f(z_2)$ . Если  $f(z_2) < f(z_1)$ , то из точки  $z_2$  опять делаем шаг по  $\eta$  и так до тех пор, пока на некотором шаге значение функции не возрастает. Эту последнюю точку и принимаем за  $x_{k+1}$ . Пусть  $l$  — число шагов при этой процедуре. Тогда если оказалось, что  $l > L$ , то за новый пробный шаг принимаем число  $\mu h/L$ . Если же функция возрастает на первом шаге, то за новый шаг принимаем число  $\gamma h$ , во всех остальных случаях пробный шаг не изменяем. Во всех приведенных задачах  $\gamma$ ,  $\mu$  и начальный шаг  $h$  принимались равными соответственно 0,1, 5/4, 0,1.

Описанная процедура не гарантирует монотонного убывания функции, однако, как показали эксперименты, возрастание функции происходит очень редко.

В программе производится контроль малости элементов  $b_{ij}$  матрицы  $B$  ( $B = A^{-1}$ , где  $A$  — матрица преобразования пространства переменных). Через каждые  $p$  итераций вычисляется величина  $d = \max_i |b_{ij}|$ , и если  $d < \varepsilon$ , то матрицу  $B$  умножаем на число  $\rho$ . При счете  $p = 10$ ,  $\varepsilon = 1$ ,  $\rho = 10$ .

Остановимся на вопросе трудоемкости программы. Основными дополнительными операциями в программе являются операции умножения матрицы на вектор, требующий  $2n^2$  операций умножения и сложения ( $n$  — размерность пространства переменных), и операция умножения матрицы  $B$  на матрицу оператора растяжения  $R_\alpha(\xi)$ . Благодаря тому что  $R_\alpha(\xi)$  — матрица специального вида, операция  $BR_\alpha(\xi)$  требует  $5n^2$  операций умножения и сложения.

Основной массив необходимой дополнительной памяти занимает матрица  $B$ , для хранения которой необходимо  $n^2$  ячеек. Поэтому, используя лишь оперативную память машины БЭСМ-6, мы можем решать задачи с числом переменных до 140.

При описании результатов счета приняты следующие обозначения:  $f = f(t_1, \dots, t_n)$  — минимизируемая функция;  $x^*$  — истинная точка минимума  $f(x)$ ;  $f^*$  — истинный минимум  $f(x)$ ;  $x_0$  — начальная точка — точка на нулевой итерации;  $x_N$  — точка, полученная в результате счета;  $k$  — счетчик итерации;  $f$  — значение функции;  $\alpha$  — коэффициент растяжения;  $d$  — величина градиента в точке  $\tilde{x}$ ,  $d = \|g_f(x)\|$ ;  $\delta = \max_{i=1, n} |\tilde{t}_i - t_i^*|$ ;  $l$  — среднее количество вычислений значений функции на одной итерации. Условием прекращения счета во всех примерах являлось выполнение одного из следующих неравенств:

$\|g_f(x_N)\| \leq 10^{-6}$ ;  $\|x_N - x_{N-1}\| \leq 10^{-7}$ ;  $\|B_{g_f}(x_N)\| \leq 10^{-18}$ ;  
где  $N$  — номер итерации, при которой произошел останов.

**Пример 1** [40]. Минимизировать функцию

$$f(x) = 100(t_1^2 - t_2)^2 + (t_1 - 1)^2, \\ x^* = (1; 1); \quad f^* = 0; \quad x_0 = (-1, 2; 1)$$

при  $\alpha$ , равном 2 и 3,  $N = 63$  и 39 соответственно,  $x_N = (1,000\ 000; 1, 000\ 000)$ .

**Пример 2** [40]. Минимизировать функцию

$$f(x) = \sum_{i=1}^{10} (e^{-0,2i} + 2e^{-0,4i} - t_1 e^{-0,2t_2 i} - t_3 e^{0,4t_4 i})^2;$$

$$x^* = (1; 1; 2; 1); \quad f^* = 0; \quad x_0 = (0; 0; 0; 0);$$

$$\alpha = 2; \quad x_{33} = (0,999999; 1,000000; 2,000000; 1,000000);$$

$$\alpha = 3; \quad x_{90} = (0,999999; 0,999999; 2,000001; 1,000000).$$

**Пример 3** [40]. Минимизировать функцию

$$f(x) = 10^{-3} \cdot \sum_{i=1}^{10} (10^3 e^{-0,2i} + 2 \cdot 10^3 e^{-0,4i} - t_1 e^{-0,2t_2 i} - t_3 e^{-0,4t_4 i})^2;$$

$$x^* = (1000; 1; 2000; 2); \quad f^* = 0; \quad x_0 = (500; 0; 2500; 3);$$

$$\alpha = 2; \quad \tilde{x}_{100} = (1000,000; 1,000000; 2000,000; 2,000000);$$

$$\alpha = 3; \quad \tilde{X}_{72} = (1000,001; 1,000000; 1999,999; 2,000000).$$

**Пример 4** [67]. Минимизировать функцию

$$f(x) = 100(t_1^2 - t_2)^2 + (t_1 - 1)^2 + 90(t_3^2 - t_4)^2 + (t_3 - 1)^2 + 10,1[(t_2 - 1)^2 + (t_4 - 1)^2] + 19,8(x_2 - 1)(x_4 - 1);$$

$$x^* = (1; 1; 1; 1); \quad f^* = 0; \quad x_0 = (-3; -1; -3; -1);$$

$$\alpha = 2; \quad \tilde{x}_{99} = (0,999999; 0,999999; 1,000000; 1,000000);$$

$$\alpha = 3; \quad \tilde{x}_{76} = (1,000000; 1,000000; 1,000000; 1,000000).$$

**Пример 5** [67]. Минимизировать функцию

$$f(x) = (e^{t_1} - t_2)^4 + 100(t_2 - t_3)^6 + th^4(t_3 - t_4) + t_1^8 + (t_4 - 1)^2;$$

$$x^* = (0; 1; 1; 1); \quad f^* = 0; \quad x_0 = (1; 2; 2; 2);$$

$$\alpha = 2; \quad \tilde{x}_{36} = (0,0006; 1,0005; 1,0000; 1,0000);$$

$$\alpha = 3; \quad \tilde{x} = (-0,0004; 0,9996; 1,0000; 1,0000).$$

**Пример 6** [67]. Минимизировать функцию

$$f(x) = (t_1 + 10t_2)^2 + 5(t_3 - t_4)^3 + (t_2 - 2t_3)^2 + 10(t_1 - t_4)^4;$$

$$x^* = (0; 0; 0; 0); \quad f^* = 0; \quad x_0 = (10; 10; 10; -10);$$

$$\alpha = 2; \quad \tilde{x}_{50}^* = (4 \cdot 10^{-7}; \quad -2 \cdot 10^{-7}; \quad 6 \cdot 10^{-7}; \quad 7 \cdot 10^{-7});$$

$$\alpha = 3; \quad \tilde{x}_{46} = (3,8 \cdot 10^{-5}; \quad 0,4 \cdot 10^{-5}; \quad -4,6 \cdot 10^{-5}; \\ -4,6 \cdot 10^{-5}).$$

Более подробно результаты численных экспериментов приведены в работе [67].

Приведенные выше данные свидетельствуют о том, что  $r$ -алгоритм является эффективным средством решения ряда задач минимизации с ярко выраженными овражными особенностями. По числу итераций, требуемых для достижения определенной точности, он близок к наиболее удачным модификациям алгоритмов типа Давидона — Флетчера — Пауэлла и метода сопряженных градиентов [40]. Алгоритм устойчив по отношению к неточностям определения минимума по направлению (использован заведомо грубый способ локализации минимума по направлению). Как показал опыт вычислений, значения коэффициента растяжения пространства  $\alpha$  целесообразно выбирать равными 2 или 3 (опыты проводились и для множества других значений коэффициента растяжения), так как в большинстве случаев при дальнейшем увеличении  $\alpha$  количество итераций если и уменьшается, то незначительно, зато заметно увеличивается среднее число вычислений функции при поиске минимума по направлению (конечно, эти рекомендации относятся лишь к той конкретной модификации алгоритма, которую мы применяли).

Нами приведены также эксперименты с вычислениями по несколько измененному алгоритму. Изменение состояло в том, что минимум по направлению находился с высокой точностью. При этом наблюдалось уменьшение числа итераций на 20—30% при сохранении прежних критериев остановки.

## § 7. Обоснование сходимости одного из вариантов $r$ -алгоритма

Рассмотрим сходимость одного из «идеализированных» вариантов  $r$ -алгоритма с монотонным поведением последовательности значений функционала в точках минимизирующей последовательности [71].

Пусть заданы почти дифференцируемая функция  $f(x)$ , определенная в  $n$ -мерном евклидовом пространстве  $E_n$ ;

множество почти-градиентов  $f(x)$  в точке  $x \in G_f(x)$  (понятие почти дифференцируемой функции и почти-градиента введено в § 4 гл. 1);  $\alpha > 1$ ,  $x_0 \in E_n$ ,  $\tilde{g}_0 = 0 \in E_n$ ; неособенная матрица размерности  $n \times n$   $B_0$ ;  $\mu$ ,  $0 \leq \mu < 1$ .  $r_\mu(\alpha)$ -алгоритмом минимизации  $f(x)$  назовем итеративную процедуру следующего типа.

Пусть проделаны  $k$  итераций ( $k = 0, 1, 2, \dots$ ) и вычислены  $x_k, g_k, B_k$ . На  $(k+1)$ -й итерации вычисляем следующие величины:

$$1) g_k^* = B_k^* g_f(x_k), \text{ где } g_f(x_k) \in G_f(x_k) \text{ такой, что}$$

$$(B_k^* g_f(x_k), \tilde{g}_k) \leq \mu \|B_k^* g_f(x_k)\| \|\tilde{g}_k\|;$$

$g_k^*$  — почти-градиент функции  $\varphi_k(y) = f(B_k y)$  в точке  $y = A_k x_k$ , где  $A_k = B_k^{-1}$ ;

2)  $r_k = g_k^* - \tilde{g}_k$ , где  $r_k$  — разность двух почти-градиентов от функции  $\varphi_k(y)$ , вычисленных в точках  $y_k = A_k x_k$ ,  $\tilde{y}_k = A_k x_{k-1}$ ;

$$3) \xi_{k+1} = \frac{r_k}{\|r_k\|};$$

4)  $B_{k+1} = B_k R_\beta(\xi_{k+1})$ ;  $B_{k+1}$  — матрица, обратная  $A_{k+1}$ ;  $A_{k+1}$  — матрица преобразования пространства после  $(k+1)$ -го шага,

$$A_{k+1} = R_\alpha(\xi_{k+1}) A_k;$$

$R_\alpha(\xi)$  — оператор растяжения пространства в направлении  $\xi$ ;  $\|\xi\| = 1$ ;  $\alpha$  — коэффициент растяжения пространства  $E_n$ ;  $\beta = \frac{1}{\alpha}$  — коэффициент «сжатия» пространства градиентов;

5)  $\tilde{g}_{k+1} = R_\beta(\xi_{k+1}) g_k^* = B_{k+1}^* g_f(x_k)$ ,  $\tilde{g}_{k+1}$  — значение почти-градиента функции  $\varphi_{k+1}(y) = f(B_{k+1} y)$  в точке  $y = A_{k+1} x_k$ ;

$$6) x_{k+1} = x_{k+1}(h_{k+1}) = x_k - h_{k+1} B_{k+1} \tilde{y}_{k+1}, \quad (3.48)$$

где  $h_{k+1}$  выбирается из условий

- а) на отрезке  $[0, h_{k+1}]$   $\psi(h) = f(x_{k+1}(h))$  не возрастает;  
б)  $\exists g \in G_f(x_{k+1})$  такой, что

$$(B_{k+1}^* g, \tilde{g}_{k+1}) \leq \mu \|B_{k+1}^* g\| \|\tilde{g}_{k+1}\|. \quad (3.49)$$

Применим к обеим частям формулы (3.48) оператор  $A_{k+1}$ :

$$y_{k+1} = A_{k+1}x_{k+1} = A_{k+1}x_k - h_{k+1}\tilde{g}_{k+1} = \tilde{y}_{k+1} - h_{k+1}\tilde{g}_{k+1}.$$

Таким образом, формула (3.48) фактически реализует шаг градиентного спуска для функции  $\varphi_{k+1}(y)$ . При  $\mu = 0$  получается аналог шага наискорейшего спуска;

7) переходим к  $(k + 2)$ -й итерации с запоминанием  $x_{k+1}$ ,  $\tilde{g}_{k+1}$ ,  $B_{k+1}$  или заканчиваем работу алгоритма при выполнении некоторого критерия останова.

В отличие от других градиентных методов минимизации негладких функций (ОГС, ОГСРП)  $r_\mu(\alpha)$ -алгоритм обеспечивает монотонность спуска благодаря определенному (3.47), (3.48) способу выбора шага спуска  $f(x_0) \geq f(x_1) \geq \dots \geq f(x_k) \geq \dots$ . Однако он существенно отличается от обычных релаксационных методов тем, что если  $h_k = 0$ , то это не значит, что процесс спуска прекращается. В процессе выполнения серии итераций с нулевым шагом  $x_k$  не изменяется, но изменяются  $B_k$  и  $\tilde{g}_k$ . В это время как бы в скрытой форме осуществляется поиск подходящего направления спуска.

Чтобы избежать рассмотрения в некотором смысле патологических случаев, предположим, что функция  $f(x)$  является кусочно-гладкой.

Определим более точно класс  $K$  почти дифференцируемых кусочно-гладких функций. Пусть задано разбиение  $E_n$  на  $m$  замкнутых множеств  $\bar{D}_1, \bar{D}_2, \dots, \bar{D}_m$ , причем подмножества внутренних точек  $D_i \subset \bar{D}_i$  гомеоморфны открытой сфере или полупространству и между собой не пересекаются. На каждом из заданных открытых множеств  $D_i^+ \supset \bar{D}_i$ ,  $i = 1, 2, \dots, m$ , определена непрерывная и непрерывно дифференцируемая функция  $f_i(x)$ . Будем говорить, что  $f \in K$ , т. е.  $f$  является непрерывной кусочно-гладкой функцией с частичными функциями  $f_i(x)$  на  $\bar{D}_i$ ,  $i = 1, \dots, m$ , если

а) для  $x \in D_i$   $f(x) = f_i(x)$ ;

б) для  $x \in \bar{D}_i \cap \bar{D}_j$   $f(x) = f_i(x) = f_j(x)$ ,  $i, j = 1, \dots, m$ .

Легко увидеть, что  $f(x)$  является почти дифференцируемой функцией. При  $x \in D_i$  множество  $G_f(x)$  состоит из одного элемента  $g_{f_i}(x)$  — градиента функции  $f_i$  в точке  $x$ . Если  $x$  лежит на общей границе  $i_1, \dots, i_k$  областей гладкости (кусков),  $k \geq 2$ , то  $G_f(x) = \{g_{f_{i_1}}(x), \dots, g_{f_{i_k}}(x)\}$ . Для

$f(x) \in K$  операция вычисления почти-градиента сводится к вычислению градиента частных функций.

Предположим далее, что выполняется следующее свойство:

$$\lim_{x \rightarrow \infty} f(x) = +\infty. \quad (3.50)$$

При этом предположении пп.1,6 итерации  $r_\mu(\alpha)$ -алгоритма всегда осуществимы. В самом деле, если при выполнении п. 6 проводить вычисления до локального минимума по направлению (в силу свойства (3.50) этот минимум всегда достижим), то свойство а) выполнится тривиально; для проверки свойства б) достаточно показать, что  $\exists g \in G_f(x_{k+1})$  такой, что  $(B_{k+1}^*g, \tilde{g}_{k+1}) \leq 0$ . Но если для всех  $g \in G_f(x_{k+1})$   $0 < (B_{k+1}^*g, \tilde{g}_{k+1}) = (g, B_{k+1}\tilde{g}_{k+1})$ , то точка  $x_{k+1}$  не могла бы быть точкой локального минимума по направлению. Таким образом, свойство б) в этой точке также выполняется. Почти-градиент, существование которого предполагается в п. 6, используется на первом шаге следующей итерации.

**З а м е ч а н и е.** Введение константы  $\mu$  вместо нуля дает возможность рассматривать алгоритмы, в которых поиск минимума по направлению производится приближенно.

Для исследования поведения  $r_\mu(\alpha)$ -алгоритмов нам понадобятся некоторые вспомогательные утверждения о поведении ширины выпуклых множеств при использовании операций растяжения (сжатия) пространства.

Пусть  $W$  — выпуклое замкнутое ограниченное тело в  $E_n$ ;  $\eta$  — единичный вектор,  $\|\eta\| = 1$ . Рассмотрим гиперплоскости вида  $(\eta, x) = d = 0$ . Введем обозначения

$$\begin{aligned} d_\eta^- &= \max d \setminus (\eta, x) - d \geq 0, \quad x \in W; \\ d_\eta^+ &= \min d \setminus (\eta, x) - d \leq 0, \quad x \in W. \end{aligned} \quad (3.51)$$

Из (3.51) видно, что  $d_\eta^+ \geq d_\eta^-$ . Шириной тела  $W$  в направлении  $\eta$  называется величина  $d_\eta(W) = d_\eta^+ - d_\eta^-$ . Геометрически это расстояние между двумя параллельными гиперплоскостями, опорными к телу  $W$ , с общей нормалью  $\eta$ , между которыми находится  $W$ . Шириной  $d(W)$  тела  $W$  называется  $\min_{\|\eta\|=1} d_\eta(W)$ , диаметром  $D(W) = \max_{\|\eta\|=1} d_\eta(W)$ . Известно, что

$$D(W) = \sup_{x, y \in W} \|x - y\|.$$

Исследуем изменение ширины выпуклого тела при линейном преобразовании пространства.

**Лемма 3.2.** Пусть  $B$  — линейный оператор, имеющий полярное разложение  $B = SO$ , где  $O$  — ортогональный оператор,  $S$  — неотрицательно определенный симметричный оператор, у которого минимальное собственное число равно  $\lambda(B)$ . Тогда ширина множества  $BW$  (образа множества  $W$  при преобразовании  $B$ ) удовлетворяет неравенству

$$\lambda(B) d(W) \leq d(BW) \leq \lambda(B) D(W).$$

**Доказательство.** Пусть  $x_1$  и  $x_2$  — две точки на границе множества  $W$ , через которые можно провести две параллельные опорные гиперплоскости с нормалью  $\eta$  такие, что после преобразования  $B$  они переходят в опорные гиперплоскости к телу  $BW$ , расстояние между которыми равно ширине тела  $BW$ . Тогда

$$\|B(x_1 - x_2)\| = d(BW),$$

$$d(BW) \geq \min_{y \neq 0} \frac{\|By\|}{\|y\|} \|x_1 - x_2\| \geq \lambda(B) d_\eta(W) \geq \lambda(B) d(W).$$

Пусть теперь  $\bar{\eta}$ ,  $\|\bar{\eta}\| = 1$ , — направление, на котором достигается  $\min_{y \neq 0} \frac{\|By\|}{\|y\|}$ , т. е.  $\|B\bar{\eta}\| = \lambda(B)$ . Тогда

$$\frac{d_{\bar{B}\bar{\eta}}(BW)}{\|B\bar{\eta}\|} \leq \lambda(B) d_{\bar{\eta}}(W) \leq \lambda(B) D(W).$$

Лемма доказана.

**Лемма 3.3.** Пусть  $z_1, z_2 \in W$ ;  $\beta \leq 1$ ;  $\frac{\|z_1 - z_2\|}{d(W)} = \gamma \geq 1$ ;  $R_\beta(\xi)$  — оператор растяжения с коэффициентом  $\beta$  в направлении  $\xi$ . Тогда

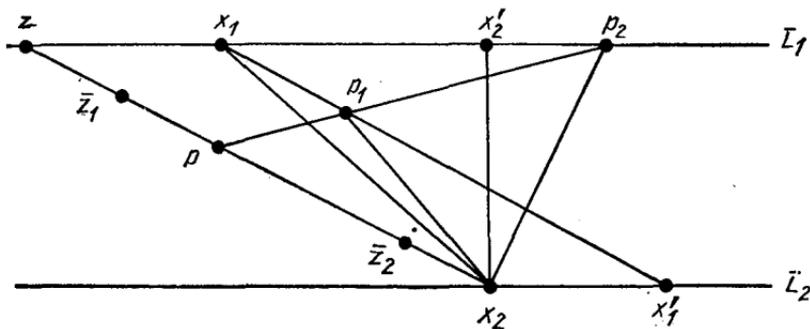
$$d\left(R_\beta\left(\frac{z_1 - z_2}{\|z_1 - z_2\|}\right)W\right) \geq \frac{d(W)}{\sqrt{1 + \frac{1 - \beta^2}{\beta^2 \gamma^2}}}.$$

**Доказательство.** Пусть  $x_1, x_2$  — точки на границе множества  $W$  такие, что через них можно провести параллельные опорные гиперплоскости  $L_1, L_2$  с нормалью  $\eta$ ,  $\|\eta\| = 1$ , которые после преобразования  $R_\beta\left(\frac{z_1 - z_2}{\|z_1 - z_2\|}\right)$  переходят в опорные гиперплоскости к телу  $R_\beta\left(\frac{z_1 - z_2}{\|z_1 - z_2\|}\right)W$ ,

определяющие ширину этого тела. Можно записать

$$d\left(R_\beta\left(\frac{z_1 - z_2}{\|z_1 - z_2\|}\right)W\right) = \left\|R_\beta\left(\frac{z_1 - z_2}{\|z_1 - z_2\|}\right)(x_1 - x_2)\right\|.$$

Проведем через точки  $x_1, x_2$  параллельно вектору  $z_1 - z_2$  плоскость;  $\bar{L}_1, \bar{L}_2$  — прямые, образуемые пересечением этой плоскости соответственно с гиперплоскостями  $L_1, L_2$ ;  $z_1^0, z_2^0$  — точки пересечения прямой, проходящей через  $z_1, z_2$ , соответственно с  $L_1, L_2$ ;  $\bar{z}_1, \bar{z}_2$  — образы точек  $z_1, z_2$ , получаемые в результате параллельного переноса отрезка  $[z_1^0, z_2^0]$  так, что  $z_2^0$  совмещается с  $x_2$ . Образ  $z_1^0$  при этом переносе обозначим буквой  $z$ . Таким образом, имеем геометри-



ческую картину, изображенную на рисунке. Пусть  $x_2'$  — проекция  $x_2$  на  $\bar{L}_1$ ,  $x_1 x_1' \parallel z x_2$ ;  $p$  — точка на отрезке  $[z, x_2]$  такая, что  $\|p - x_2\| = \beta \|z - x_2\|$ ;  $p_2$  — точка пересечения с  $\bar{L}_1$  прямой, проходящей через  $x_2$  и перпендикулярной  $[x_2, z]$ ;  $p_1$  — точка пересечения отрезков  $[x_1, x_1']$  и  $[p, p_2]$ . Из определения оператора  $R_\beta\left(\frac{z_1 - z_2}{\|z_1 - z_2\|}\right)$  очевидно, что

$$\left\|R_\beta\left(\frac{z_1 - z_2}{\|z_1 - z_2\|}\right)(x_1 - x_2)\right\| = \|p_1 - x_2\|.$$

Найдем  $\min \|p_1 - x_2\|$  при следующих ограничениях:

- 1)  $\|z_1 - x_2\| \geq \|z_1^0 - z_2^0\| = \|z_1 - z_2\| \geq \gamma d(W)$ ;
- 2)  $\|x_2 - x_2'\| \geq d_\eta(W) \geq d(W)$ .

Очевидно, что минимум достигается в том случае, когда  $\|z - x_2\| = \gamma d(W)$ ,  $\|x_2 - x_2'\| = d(W)$  и отрезок

$[x_2, p_1]$  является высотой  $\Delta r p_2 x_2$ . Непосредственным вычислением получаем, что

$$\min \|p_1 - x_2\| = \frac{d(W)}{\sqrt{1 + \frac{1 - \beta^2}{\beta^2 \gamma^2}}}.$$

Лемма доказана.

Для простоты изложения все дальнейшие рассуждения проведем для случая  $\mu = 0$ .  $r_0(\alpha)$ -алгоритм будем называть  $r(\alpha)$ -алгоритмом. Для произвольного  $\mu$ ,  $0 < \mu < 1$ , доказательство проводится по той же схеме, что и для  $\mu = 0$ , с небольшими изменениями.

Пусть  $\{x_k\}_{k=0}^{\infty}$  — последовательность точек, получаемая при минимизации  $f \in K$  с помощью  $r(\alpha)$ -алгоритма. Так как  $f(x_k) \geq f(x_{k-1})$ , то, если справедливо (3.50), существует предел

$$\lim_{k \rightarrow \infty} f(x_k) = f_{\infty} > -\infty$$

и множество  $\{x_k\}_{k=0}^{\infty}$  ограничено.

Обозначим через  $M_{\infty}$  множество предельных точек последовательности  $\{x_k\}_{k=0}^{\infty}$ . Очевидно, что если  $x \in M_{\infty}$ , то  $f(x) = f_{\infty}$ . Нас будут интересовать свойства элементов  $M_{\infty}$ . Выберем  $\delta > 0$ ,  $\varepsilon > 0$  и рассмотрим  $\delta$ -окрестность точки  $x$  — шар  $S_{\delta}(x)$ . Обозначим через  $\bar{P}_{\delta\varepsilon}(x)$  выпуклое замыкание множества

$$P_{\delta,\varepsilon}(x) = \left( \bigcup_{y \in S_{\delta}(x)} G_f(y) \right) \cup \left( \bigcup_{z \in G_f(x)} S_{\varepsilon}(z) \right).$$

Пусть  $W$  — выпуклое множество. Введем далее следующие функции:  $e_{\rho}(W) = \inf_{z \in W} |(\rho, z)|$ ;  $\rho \in E_n$ ;  $\|\rho\| = 1$ ;  $d_{\rho}(W)$  — ширина множества  $W$  в направлении  $\rho$ . Ясно, что  $d_{\rho}(\bar{P}_{\delta,\varepsilon}(x)) \geq \varepsilon$ ;

$$K_{\rho}(W) = \begin{cases} \frac{d_{\rho}(W)}{e_{\rho}(W)}, & \text{если } e_{\rho}(W) \neq 0, \\ +\infty, & \text{если } e_{\rho}(W) = 0, \end{cases}$$

$$\rho(W) = \inf_{\|\rho\|=1} K_{\rho}(W).$$

Легко показать, что  $\rho(AW) = \rho(W)$ , если  $\det A \neq 0$ . Докажем следующую теорему.

**Теорема 3.11.** Пусть  $f \in K$ ;  $\alpha > 1$ ;  $\{x_k\}_{k=0}^{\infty}$  — последовательность, полученная в результате реализации  $r(\alpha)$ -

алгоритма минимизации  $f(x)$ , удовлетворяющая условию

$$\lim_{k \rightarrow +\infty} \|x_{k+1} - x_k\| = 0. \quad (3.52)$$

Тогда для любых фиксированных  $\nu, \sqrt[n]{\beta} < \nu < 1, \varepsilon > 0, \delta > 0$  и натурального  $r$  найдется  $\bar{k} > r$  такое, что

$$\rho(\bar{P}_{\delta, \varepsilon}(x_{\bar{k}})) \geq \sqrt{\frac{\nu^2 \sqrt[n]{\alpha^2} - 1}{\alpha^2 - 1}}.$$

Доказательство. Допустим, что для всех  $k > r$

$$\rho(\bar{P}_{\delta, \varepsilon}(x_k)) < \sqrt{\frac{\nu^2 \sqrt[n]{\alpha^2} - 1}{\alpha^2 - 1}}.$$

Введем величины

$$D_{\max}^{(r)}(\delta, \varepsilon) = \sup_{k > r} D(\bar{P}_{\delta, \varepsilon}(x_k)).$$

Из условия (3.52) вытекает, что для любого натурального  $s$  существует  $N(s, \delta)$  такое, что при  $k > N(s, \delta)$   $x_t \in S_\delta(x_k)$ ,  $k \leq t \leq k + s$ . Следовательно,  $g_f(x_k), g_f(x_{k+1}), \dots, g_f(x_{k+s})$  принадлежат  $\bar{P}_{\delta, \varepsilon}(x_k)$ . Воспользуемся леммой 3.2 для получения оценки ширины множества  $B_{k+1}^* \bar{P}_{\delta, \varepsilon}(x_k)$ :

$$\varepsilon \lambda(B_{k+1}) \leq d(B_{k+1}^* \bar{P}_{\delta, \varepsilon}(x_k)) \leq \lambda(B_{k+1}) D_{\max}^{(r)}(\delta, \varepsilon); \quad k > r. \quad (3.53)$$

Из пп. 1 и 2 описания  $r$ -алгоритма получаем

$$\begin{aligned} r_{k+i} &= B_{k+1}^* [g_f(x_{k+i}) - g_f(x_{k+i-1})], \\ (B_{k+i}^* g_f(x_{k+i}), B_{k+i}^* g_f(x_{k+i+1})) &\leq 0. \end{aligned}$$

Отсюда

$$\begin{aligned} \|r_{k+i}\| &\geq \max \{ \|B_{k+i}^* g_f(x_{k+i})\|, \|B_{k+i}^* g_f(x_{k+i+1})\| \}, \\ \frac{d(B_{k+i}^* \bar{P}_{\delta, \varepsilon}(x_k))}{\|r_{k+i}\|} &\leq \rho(B_{k+i}^* \bar{P}_{\delta, \varepsilon}(x_k)), \quad i = 1, \dots, s; \\ \frac{\|r_{k+i}\|}{d(B_{k+i}^* \bar{P}_{\delta, \varepsilon}(x_k))} &\geq \frac{1}{\rho(B_{k+i}^* \bar{P}_{\delta, \varepsilon}(x_k))} \geq \sqrt{\frac{\alpha^2 - 1}{\nu^2 \sqrt[n]{\alpha^2} - 1}}. \quad (3.54) \end{aligned}$$

Воспользуемся леммой 3.3, приняв

$$\beta = \frac{1}{\alpha}, \quad \gamma \geq \sqrt{\frac{\alpha^2 - 1}{\nu^2 \sqrt[n]{\alpha^2} - 1}}.$$

Получаем

$$\begin{aligned}
 d(B_{k+i+1}^* \bar{P}_{\delta, \varepsilon}(x_k)) &= d(R_{\frac{1}{\alpha}}(\xi_{k+i+1}) B_{k+i}^* \bar{P}_{\delta, \varepsilon}(x_k)) \geq \\
 &\geq d(B_{k+i}^* \bar{P}_{\delta, \varepsilon}(x_k)) \frac{1}{\nu \sqrt[n]{\alpha}} \quad (1 > \nu > \sqrt[n]{\beta}). \quad (3.55)
 \end{aligned}$$

Выберем  $s$  настолько большим, чтобы

$$\frac{1}{\nu^s} \geq \frac{D_{\max}^{(r)}(\delta, \varepsilon)}{\varepsilon} \bar{c}, \quad (3.56)$$

где  $\bar{c} > 1$ . Тогда из (3.54) — (3.56) получим

$$\begin{aligned}
 D_{\max}^{(r)}(\delta, \varepsilon) \lambda(B_{k+s+1}) &\geq d(B_{k+s+1}^* \bar{P}_{\delta, \varepsilon}(x_k)) \geq \\
 &\geq d(B_{k+1}^* \bar{P}_{\delta, \varepsilon}(x_k)) \frac{\sqrt[n]{\beta^s}}{\nu^s} \geq \\
 &\geq d(B_{k+1}^* \bar{P}_{\delta, \varepsilon}(x_k)) \frac{D_{\max}^{(r)}(\delta, \varepsilon)}{\varepsilon} \nu \sqrt[n]{\beta^s} \bar{c} \geq \\
 &\geq \lambda(B_{k+1}) D_{\max}^{(r)}(\delta, \varepsilon) \sqrt[n]{\beta^s} \bar{c}.
 \end{aligned}$$

Отсюда

$$\frac{\lambda(B_{k+s+1})}{\lambda(B_{k+1})} \geq \bar{c} \frac{\nu \sqrt[n]{\beta^s}}{\nu^s}.$$

Приведенное выше рассуждение справедливо для любого  $k > N(s, \delta)$ . Следовательно, для любого  $q$  и  $\bar{k} > N(s, \delta)$  можно написать

$$\frac{\lambda(B_{\bar{k}+sq+1})}{\lambda(B_{\bar{k}+1})} \geq \beta \frac{qs}{n} \bar{c}^q, \quad q = 1, 2, \dots$$

Но тогда при достаточно большом  $q$   $\lambda(B_{\bar{k}+qs+1}) > \beta \frac{\bar{k}+qs+1}{n}$ ,

а это противоречит тому, что  $\lambda(B_k) \leq \beta \frac{k}{n}$ . Теорема доказана.

**Теорема 3.12.** Пусть выполняются условия теоремы 3.11 и условие (3.50). Тогда на множестве  $U = \{x : (x) = f_{\infty}\}$  найдется точка  $x^*$  такая, что множество векторов  $G_f(x^*)$  будет линейно зависимым.

**Доказательство.** Доказательство приведем от противного. Допустим, что при всех  $x \in U$  множество  $G_j(x)$  состоит из линейно независимых векторов. Заметим, что множество  $U$  компактно. Обозначим через  $\Gamma(z)$  определитель Грамма, построенный на векторах, принадлежащих  $G_j(z)$ . Покажем, что

$$\inf_{z \in U} \Gamma(z) = e_0 > 0.$$

Если бы это было не так, то нашлась бы последовательность  $\{z_k\}_{k=1}^{\infty} \subseteq U$  такая, что

$$\lim_{k \rightarrow \infty} \Gamma(z_k) = 0,$$

причем точки  $z_k$  могут быть выбраны таким образом, чтобы они принадлежали одному и тому же куску или границам одних и тех же кусков  $D_1, \dots, D_r$ . Но тогда в силу компактности  $U$  найдется  $z^* \in U$  такая, что  $\Gamma(z^*) = 0$ , а это противоречит условию независимости векторов  $G_j(z^*)$ .

Так как  $\inf_{z \in U} \Gamma(z) = e_0 > 0$ , то существует такое  $e_1 > 0$ , что определитель Грамма, построенный на любой подсистеме векторов, входящих в  $G_j(z)$ , не меньше чем  $e_1$ .

Сопоставим каждой точке  $x \in U$  сферическую окрестность  $S'_\varepsilon(x)$  такую, что: 1) она имеет общие точки только с теми кусками  $D_j$ ,  $1 \leq j \leq m$ , которые инцидентны точке  $x$ ; 2) определитель Грамма, построенный на любой подсистеме векторов, входящих в  $G_j(z)$ ,  $z \in S'_\varepsilon(x)$ , не меньше чем  $\frac{e_1}{2}$ ;  $\|g(z_1) - g(z_2)\| \leq \frac{\varepsilon}{2}$ ,  $\varepsilon > 0$ , если  $z_1, z_2 \in S'_\varepsilon(x)$  и  $z_1, z_2$  лежат в одном и том же куске. Пусть  $S''(x)$  — окрестность точки  $x$ , радиус которой равен половине радиуса  $S'_\varepsilon(x)$ ;

$$S = \bigcup_{x \in U} S''(x),$$

где  $S$  — открытое множество, содержащее  $U$ . Так как  $U$  — компакт, то можно выбрать конечную подсистему окрестностей  $S''(z_1), \dots, S''(z_q)$ , покрывающую  $U$ . Пусть  $\delta$  — минимальный радиус этих окрестностей. Начиная с некоторого достаточно большого  $N$ , каждая точка  $x_k$  минимизирующей последовательности будет находиться внутри некоторой окрестности  $S''(z_{i_k})$ ,  $1 \leq i_k \leq q$ . Тогда  $\delta$ -окрестность точки  $x_k$  будет находиться внутри  $S'_\varepsilon(z_{i_k})$ .

Расстояние от начала координат до гиперплоскости наименьшей размерности  $\Pi(z)$ , проходящей через все точки, входящие в  $G_f(x)$ , вычисляется по формуле

$$\rho[G_f(z)] = \left( \frac{\Gamma(G_f(z))}{\Gamma\{|g_2(z) - g_1(z), \dots, g_{t_z}(z) - g_1(z)\}} \right)^{1/2},$$

где  $\{g_1(z), \dots, g_{t_z}(z)\} \equiv G_f(z)$ . Так как  $\|g_2(z_{i_k}) - g_1(z_{i_k})\|, \dots, \|g_{t_z}(z_{i_k}) - g_1(z_{i_k})\|$  ограничены сверху, а  $\Gamma(G_f(z_{i_k})) \geq$

$\geq \frac{\varepsilon_1}{2}$ , то для всех  $k > N$  при достаточно малом  $\varepsilon$

$e_\rho(\bar{P}_{2\delta, \varepsilon}(z_{i_k})) = c > 0$ ;  $\rho$  — нормаль к  $\Pi(z)$ ;  $z = z_{i_k}$ ;  $d_\rho(\bar{P}_{2\delta, \varepsilon}(z_{i_k})) \leq 2\varepsilon$ . Учитывая, что  $\bar{P}_{\delta, 0}(x_k) \subseteq \bar{P}_{2\delta, 0}(z_{i_k})$ , для любого  $v > 0$  можно найти такое  $\varepsilon$ , что

$$\rho(\bar{P}_{\delta, \varepsilon}(x_k)) \leq \frac{d_\rho(\bar{P}_{2\delta, \varepsilon}(z_{i_k}))}{e_\rho(\bar{P}_{2\delta, \varepsilon}(z_{i_k}))} \leq v; \quad k > N.$$

Но это противоречит теореме 3.11. Теорема доказана.

Из теоремы 3.12 можно вывести следствие.

**Теорема 3.13.** Если выполняются условия теоремы 3.12, точка  $x^*$  — изолированная точка локального минимума,  $x_0$  — такая точка, что связная компонента множества  $\{x : f(x^*) \leq f(x) \leq f(x_0)\}$ , содержащая  $x^*$  и  $x_0$ , не содержит, кроме  $x^*$ , других точек  $z$ , у которых семейство  $G_f(z)$  линейно зависимо, то последовательность  $\{x_k\}_{k=0}^\infty$ , генерируемая  $r(\alpha)$ -алгоритмом, сходится к точке  $x^*$ .

В заключение отметим следующее. Теорема 3.13 определяет достаточные условия, при которых  $r(\alpha)$ -алгоритм сходится к локальному минимуму. Насколько «тяжелы» эти условия? Легко увидеть, что условие (3.52) выполняется, например, когда функция  $f(x)$  обладает следующим свойством: если  $x_1, x_2$  принадлежат области определения  $f(x)$  и на отрезке  $[x_1, x_2]$   $f(x)$  монотонно убывает, то для любого  $\varepsilon > 0$  можно найти такое  $\delta > 0$ , что при  $\|x_1 - x_2\| > \delta$   $f(x) - f(y) > \varepsilon$ . Это условие можно рассматривать как обобщение свойства равномерной выпуклости. К сожалению, оно не выполняется для линейных функций. Наиболее типичная ситуация, связанная с нарушением условия линейной независимости семейства  $G_f(z)$ , — это тот случай, когда точка  $z$  принадлежит более чем  $n$  кускам. Таких точек, как правило, конечное число и по отношению к  $r(\alpha)$ -алгоритму, как показывает вычислительная практика, эти точки, если

они не являются точками локального минимума, не служат «ловушками» для минимизирующей последовательности.

Численные эксперименты, проведенные нами при решении разнообразных задач минимизации кусочно-гладких функций [69, 3, 9, 25] с использованием  $r$ -алгоритма, не дали пока ни одного «патологического» случая явной расходимости процесса или сходимости к точке, не являющейся точкой локального минимума. Все же было бы желательно ослабить условия теоремы 3.13. Интересно также получить оценки скорости сходимости в достаточно общих случаях.

### § 8. Связь между алгоритмами ОГСРП и алгоритмами последовательных отсечений

Алгоритмы с растяжением пространства, близкие по своей структуре к алгоритму ОГСРП, как показано в последнее время, тесно связаны с алгоритмами последовательных отсечений и могут быть использованы для решения гораздо более общих задач, чем задачи минимизации выпуклых функций [74].

Пусть на  $E_n$  задано векторное поле  $g(x)$ , не обязательно непрерывное,  $g(x) \in E_n$ ,  $x \in E_n$ . Найдем такую точку  $x^*$ , что  $(g(x), x - x^*) \geq 0$  при всех  $x \in E_n$ . Допустим, что эта задача имеет решение, причем известно априори, что  $x^* \in S(x_0, R)$ , где  $S(x_0, R)$  — замкнутый шар с центром в точке  $x_0$  и радиусом  $R$ .

Рассмотрим следующий итеративный алгоритм решения указанной задачи при  $n > 1$ . (При описании этого алгоритма будем предполагать, что  $g(x) \neq 0$  при  $x \neq x^*$ .)

Перед началом вычислений имеем  $x_0 \in E_n$ ,  $B_0 = I_n$  — единичная матрица,  $h_0 = \frac{R}{n+1}$ . Пусть на  $k$ -м шаге мы получили  $x_k \in E_n$ ;  $B_k$  — матрица  $n \times n$ ,  $h_k > 0$ . На  $(k+1)$ -м шаге вычисляем:

1)  $g(x_k)$ ; если  $g(x_k) = 0$ , то  $x_k$  — искомая точка;

$$2) \xi_k = \frac{B_k^* g(x_k)}{\|B_k^* g(x_k)\|}; \quad (3.57)$$

$$3) x_{k+1} = x_k - h_k B_k \xi_k; \quad (3.58)$$

$$4) B_{k+1} = B_k R_\beta(\xi_k); \quad \beta = \sqrt{\frac{n-1}{n+1}}, \quad (3.59)$$

$R_\beta(\xi_k)$  — оператор растяжения пространства в направлении  $\xi_k$  с коэффициентом  $\beta$ ;

$$5) h_{k+1} = h_k r; \quad r = \frac{n}{\sqrt{n^2 - 1}}. \quad (3.60)$$

**Теорема 3.14.** Последовательность  $\{x_k\}_{k=0}^\infty$ , генерируемая алгоритмом (3.57) — (3.60), удовлетворяет неравенству

$$\|A_k(x_k - x^*)\| \leq h_k(n+1), \quad A_k = B_k^{-1}, \quad k = 0, 1, \dots \quad (3.61)$$

**Доказательство.** Доказательство проведем индукцией по  $k$ . Для  $k = 0$  неравенство (3.61) переходит в  $\|x_0 - x^*\| \leq R$  и выполняется по предположению. Предположим, что (3.61) выполняется для  $k = \bar{k}$ . Докажем его выполнение для  $k = \bar{k} + 1$ .

Пусть для краткости  $A_k(x_k - x^*) = z_k$ ,  $\alpha = \frac{1}{\beta}$ ,

$$\begin{aligned} \|z_{\bar{k}+1}\|^2 &= \|R_\alpha(\xi_{\bar{k}})(z_{\bar{k}} - h_{\bar{k}}\xi_{\bar{k}})\|^2 = \|z_{\bar{k}} - h_{\bar{k}}\xi_{\bar{k}} + \\ &+ (\alpha - 1)[(z_{\bar{k}} - h_{\bar{k}}\xi_{\bar{k}}, \xi_{\bar{k}}) \xi_{\bar{k}}]\|^2 = \|z_{\bar{k}} + [-\alpha h_{\bar{k}} + \\ &+ (\alpha - 1)(z_{\bar{k}}, \xi_{\bar{k}})] \xi_{\bar{k}}\|^2 = \|z_{\bar{k}}\|^2 - 2h_{\bar{k}}(z_{\bar{k}}, \xi_{\bar{k}})\alpha^2 + \\ &+ (\alpha^2 - 1)(z_{\bar{k}}, \xi_{\bar{k}})^2 + \alpha^2 h_{\bar{k}}^2 = \|z_{\bar{k}}\|^2 - 2h_{\bar{k}}(z_{\bar{k}}, \xi_{\bar{k}}) \frac{n+1}{n-1} + \\ &+ \frac{2}{n-1}(z_{\bar{k}}, \xi_{\bar{k}})^2 + \frac{n+1}{n-1} h_{\bar{k}}^2. \end{aligned}$$

Заметим, что

$$\begin{aligned} (z_{\bar{k}}, \xi_{\bar{k}}) &= \frac{1}{\|\tilde{g}(x_{\bar{k}})\|} (A_{\bar{k}}(x_{\bar{k}} - x^*), B_{\bar{k}}^* g(x_{\bar{k}})) = \\ &= \frac{1}{\|\tilde{g}(x_{\bar{k}})\|} (x_k - x^*, g(x_{\bar{k}})) \geq 0 \end{aligned}$$

(здесь  $\tilde{g}(x_{\bar{k}}) = B_{\bar{k}}^* g(x_{\bar{k}})$ ),  $(z_{\bar{k}}, \xi_{\bar{k}}) \leq \|z_{\bar{k}}\| \leq h_{\bar{k}}(n+1)$ .

Отсюда

$$\begin{aligned} \frac{2}{n-1}(z_{\bar{k}}, \xi_{\bar{k}})^2 - 2h_{\bar{k}} \frac{n+1}{n-1} (z_{\bar{k}}, \xi_{\bar{k}}) &\leq \\ &\leq \frac{2}{n-1} [(z_{\bar{k}}, \xi_{\bar{k}})^2 - (z_{\bar{k}}, \xi_{\bar{k}})^2] = 0. \end{aligned}$$

Таким образом,

$$\begin{aligned} \|z_{\bar{k}+1}\|^2 &\leq \|z_{\bar{k}}\|^2 + \frac{n+1}{n-1} h_k^2 \leq (n+1)^2 h_k^2 + \frac{n+1}{n-1} h_k^2 = \\ &= (n+1)^2 \frac{n^2}{n^2-1} h_k^2 = (n+1)^2 h_{\bar{k}+1}^2, \end{aligned}$$

откуда следует справедливость (3.61). Теорема доказана.  
Множество точек  $x$ , удовлетворяющих неравенству

$$\|A_k(x_k - x)\| \leq (n+1)h_k = R \left( \frac{n}{\sqrt{n^2-1}} \right)^k,$$

представляет собой эллипсоид  $\Phi_k$ , объем которого  $v(\Phi_k)$  равен  $\frac{v_0 R^n \left( \frac{n}{\sqrt{n^2-1}} \right)^{nk}}{\det A_k}$ , где  $v_0$  — объем единичного  $n$ -мерного шара. Получаем

$$\begin{aligned} \frac{v(\Phi_{k+1})}{v(\Phi_k)} &= \frac{\left( \frac{n}{\sqrt{n^2-1}} \right)^n \det A_k}{\det A_{k+1}} = \frac{\left( \frac{n}{\sqrt{n^2-1}} \right)^n \det A_k}{\det R_\alpha(\xi_k) \det A_k} = \\ &= \frac{1}{\alpha} \left( \frac{n}{\sqrt{n^2-1}} \right)^n = \sqrt{\frac{n-1}{n+1}} \left( \frac{n}{\sqrt{n^2-1}} \right)^n = q_n < 1. \end{aligned}$$

Объем эллипсоида, в котором локализуется искомая точка  $x^*$  в соответствии с неравенством (3.61), убывает со скоростью геометрической прогрессии со знаменателем  $q_n$ .

Рассмотрим возможность применения описанной процедуры для решения ряда задач математического программирования.

1. Задача минимизации выпуклой функции на шаре. Пусть на  $E_n$  определена выпуклая функция  $f(x)$ . Требуется найти минимальное значение этой функции на шаре  $S(x_0, R) = \{x : \|x - x_0\| \leq R\}$ .

Пусть  $x^* \in S(x_0, R)$  — искомый минимум. Тогда для всех  $x \in S(x_0, R)$

$$(g_f(x), x - x^*) \geq f(x) - f(x^*) \geq 0.$$

Для  $x \in S(x_0, R)$

$$\begin{aligned} (x - x_0, x - x^*) &= \|x - x_0\|^2 - (x - x_0, x^* - x_0) \geq \\ &\geq \|x - x_0\| (\|x - x_0\| - \|x^* - x_0\|) \geq 0. \end{aligned}$$

Таким образом, можем построить векторное поле  $g(x)$ , удовлетворяющее свойству  $(g(x), x - x_0) \geq 0$ ,  $x \in E_n$ .

следующим образом:

$$g(x) = \begin{cases} g_j(x), & x \in S(x_0, R); \\ \frac{x - x_0}{\|x - x_0\|}, & x \notin S(x_0, R). \end{cases} \quad (3.62)$$

Применим для решения нашей задачи алгоритм (3.57) — (3.60), определив  $g(x)$  по формуле (3.62). Тогда в соответствии с теоремой 3.14 получим, что объем области, в которой локализован минимум, убывает со скоростью геометрической прогрессии со знаменателем

$$q_n = \sqrt{\frac{n-1}{n+1}} \left( \frac{n}{\sqrt{n^2-1}} \right)^n.$$

Рассмотрим случай, когда  $\min_{x \in E_n} f(x)$  достигается в точке  $x^*$ , являющейся внутренней точкой шара  $S(x_0, R)$ . Так как алгоритм (3.57) — (3.60) применительно к задаче минимизации функции  $f(x)$  фактически реализует метод ОГСРП с небольшими изменениями, то для оценки скорости сходимости по функционалу можем применить теорему (3.2). Получаем следующую оценку:

$$\rho_k = \min_{i \in 1, k} \|B_i^* g(x_i)\| \leq \frac{d_k \sqrt{k(\alpha^2 - 1)}}{\sqrt{\alpha^{\frac{2k}{n}} - 1}}, \quad (3.63a)$$

$$d_k = \max_{i \in 1, k} \|g(x_i)\|.$$

Пусть минимум в (3.63) достигается при  $i = i(k)$ . Тогда

$$\begin{aligned} \mu_k = (x_{i(k)} - x^*, g(x_{i(k)})) &\Rightarrow (A_{i(k)}(x_{i(k)} - x^*), B_{i(k)}^* g(x_{i(k)})) \leq \\ &\leq \|A_{i(k)}(x_{i(k)} - x^*)\| \frac{d_k \sqrt{k(\alpha^2 - 1)}}{\sqrt{\alpha^{\frac{2k}{n}} - 1}} \leq h_{i(k)}(n+1) \times \\ &\times \frac{d_k \sqrt{k(\alpha^2 - 1)}}{\sqrt{\alpha^{\frac{2k}{n}} - 1}} \leq h_k(n+1) d_k \frac{\sqrt{k(\alpha^2 - 1)}}{\sqrt{\alpha^{\frac{2k}{n}} - 1}} = \\ &= \frac{R d_k \sqrt{k(\alpha^2 - 1)}}{\sqrt{1 - \beta^{\frac{2k}{n}}}} q_n. \end{aligned}$$

Заметим, что если  $x_{i(k)} \in S(x_0, R)$ , то  $g(x_{i(k)}) = \frac{x - x_0}{\|x - x_0\|}$  и, так как  $x^*$  лежит внутри  $S(x_0, R)$ ,  $\left(x_{i(k)} - x^*, \frac{x_{i(k)} - x_0}{\|x_{i(k)} - x_0\|}\right) \geq \delta > 0$ . Таким образом, из того, что  $\mu_k \rightarrow 0$ , следует, что начиная с достаточно больших  $k$   $x_{i(k)} \in S(x_0, R)$ . Но если  $x_{i(k)} \in S(x_0, R)$ , то  $g(x_{i(k)}) = g_f(x_{i(k)})$  и справедливо неравенство

$$\begin{aligned} f(x_{i(k)}) - f(x^*) &\leq (g_f(x_{i(k)}, x_{i(k)} - x^*)) \leq \\ &\leq \frac{Rd_k \sqrt{k(\alpha^2 - 1)}}{\sqrt{1 - \beta \frac{2k}{n}}} q \frac{k}{n}. \end{aligned}$$

Таким образом, если  $\min_{x \in E_n} f(x)$  достигается в точке  $x^*$ , являющейся внутренней точкой шара  $S(x_0, R)$ , то при использовании алгоритма вида (3.57) — (3.60) отклонение «рекорда» по функционалу от  $f(x^*)$  удовлетворяет неравенству

$$r_k = \min_{i \in \{1, k\}} [f(x_i) - f(x^*)] \leq \frac{Rd_k \sqrt{k(\alpha^2 - 1)}}{\sqrt{1 - \beta \frac{2k}{n}}} q \frac{k}{n}.$$

Если при всех  $k = 0, 1, \dots, x_k \in S(x_0, R)$ , можно получить более точную оценку

$$f(x_{k_i}) - f(x^*) \leq cq_n \frac{k}{n}$$

для некоторой подпоследовательности  $\{k_i\}_{i=1}^{\infty}$ , используя теорему 3.1. При больших  $n$   $q_n$  хорошо приближается следующей асимптотической формулой:

$$q_n \approx 1 - \frac{1}{2n}.$$

Отсюда следует, что, хотя  $r_k$  сходится примерно со скоростью геометрической прогрессии (если пренебречь медленно растущим множителем  $\sqrt{k}$ ), при больших  $n$  знаменатель этой прогрессии близок к единице, т. е. практически скорость сходимости может оказаться медленной.

Положительным свойством алгоритма (3.57) — (3.60) является то, что гарантированная скорость сходимости зависит лишь от размерности пространства и не требует знания специфических особенностей функции  $f(x)$ .

2. Общая задача выпуклого программирования. Найти

$$\min f_0(x) \quad (3.63б)$$

при ограничениях

$$f_i(x) \leq 0; \quad i = 1, 2, \dots, m; \quad x \in E_n; \quad n > 1, \quad (3.64)$$

где  $f_v(x)$  — выпуклые функции, определенные на  $E_n$ ;  $g_v(x)$  — соответствующие субградиенты,  $v = 0, 1, \dots, m$ , причем имеется априорная информация, что оптимальная точка  $x^*$  существует и находится в шаре  $S(x_0, R)$  (формально к системе ограничений (3.64) можно добавить ограничение  $x - x_0 \parallel \leq R$ ).

Рассмотрим поле  $g(x)$ , построенное следующим образом:

$$g(x) = \begin{cases} g_0(x), & \text{если } \max_{i \in \overline{1, m}} f_i(x) \leq 0, \\ g_{i^*}(x), & \text{если } \max_{i \in \overline{1, m}} f_i(x) = f_{i^*}(x) > 0. \end{cases} \quad (3.65)$$

Покажем, что  $(g(x), x - x^*) \geq 0$  при всех  $x \in E_n$ . Если  $\max_{i \in \overline{1, m}} f_i(x) \leq 0$ , то  $g(x) = g_0(x)$ ,

$$(g(x), x - x^*) = (g_0(x), x - x^*) \geq f_0(x) - f_0(x^*) \geq 0.$$

Если  $\max_{i \in \overline{1, m}} f_i(x) > 0$ , то  $g(x) = g_{i^*}(x)$ , причем  $f_{i^*}(x) > 0$ ,  $f_{i^*}(x^*) \leq 0$ ,  $(g(x), x - x^*) = (g_{i^*}(x), x - x^*) \geq f_{i^*}(x) - f_{i^*}(x^*) \geq 0$ . Таким образом,  $(g(x), x - x^*) \geq 0$  при всех  $x \in E_n$ . Используя это неравенство, применим для локализации  $x^*$  алгоритм (3.57) — (3.60), вычисляя  $g(x)$  по формуле (3.65).

Из теоремы 3.14 вытекает, что после  $k$  шагов алгоритма (3.57) — (3.60)  $x^*$  будет локализован в эллипсоиде объемом  $v(\Phi_k) = v_0 q_n^k$  ( $v_0$  — объем шара  $S(x_0, R)$ ) с центром в точке  $x_k$ .

Отметим, что этот результат не изменится, если во второй формуле (3.65) вместо  $g_{i^*}(x)$  брать  $g_{\bar{i}}(x)$ , где  $\bar{i}$  — произвольный индекс, для которого  $f_{\bar{i}}(x) > 0$ .

3. Задача о седловой точке. Пусть задана выпукло-вогнутая функция двухвекторных переменных  $x \in E_n$ ,  $y \in E_m$ ,  $z = (x, y) \in E_n \times E_m \equiv E_{n+m}$ ,  $z^*$  — седловая точка этой функции,  $z_0$  — заданное начальное приближение и априори известно, что  $\|z_0 - z^*\| \leq R$ .

Рассмотрим псевдоградиентное множество

$$G(z) = G_f^x(x, y) \times (-G_f^y(x, y)),$$

где  $G_f^x(x, y)$  — множество частных субградиентов функции  $f(x, y)$ , рассматриваемой как функция от  $x$  при фиксированном  $y$ ;  $-G_f^y(x, y)$  — множество субградиентов от функции  $-f(x, y)$  по  $y$  при фиксированном  $x$ .

Сформируем векторное поле  $g(z)$  следующим образом:

$$g(z) = \{g_f^x(z), -g_f^y(z)\}; \quad g_f^x(z) \in G_f^x(z); \quad g_f^y(z) \in G_f^y(z).$$

Покажем, что  $(g(z), z - z^*) \geq 0$ .

Из определения седловой точки следует, что  $f(x, y^*) \geq f(x^*, y^*) \geq f(x^*, y)$ . Далее,

$$\begin{aligned} 0 \leq f(x, y^*) - f(x^*, y) &= f(x, y^*) - f(x, y) + f(x, y) - \\ &- f(x^*, y) \leq (g_f^x(z), x - x^*) - (g_f^y(z), y - y^*) = \\ &= (g(z), z - z^*). \end{aligned}$$

Таким образом, для локализации седловой точки  $z^*$  можем применить алгоритм (3.57) — (3.60)\*.

К алгоритму типа (3.57) — (3.60) можно прийти, исходя из схемы последовательных отсечений. А. Ю. Левин в 1965 г. предложил для минимизации произвольной выпуклой функции и решения задач выпуклого программирования метод центрированных сечений [31]. Суть этого метода состоит в том, чтобы, используя информацию о локализации оптимума в некотором выпуклом замкнутом многограннике, на каждом шаге находить центр тяжести этого многогранника и вычислять субградиент в этой точке. Тогда оптимум будет находиться по определенную сторону от соответствующей опорной гиперплоскости к линии уровня, часть многогранника, лежащую в другом полупространстве, отделяемом гиперплоскостью, можем отсечь. Можно показать [31], что на каждом шаге отсекается не менее чем  $\left(\frac{n}{n+1}\right)^n$ -я часть объема многогранника, что приводит к убыванию объема области локализации со скоростью, не меньшей, чем скорость убывания геометрической прогрессии со знаменателем

$$Q_n = 1 - \left(\frac{n}{n+1}\right)^n \quad \left(\text{при больших } n \quad Q_n \approx 1 - \frac{1}{e}\right).$$

\* Обоснование других обобщенных градиентных методов нахождения седловых точек можно найти в работах [11, 33].

Однако этот алгоритм при  $n \geq 3$  практически оказывается непригодным, так как вычисление центра тяжести многомерного многогранника — весьма громоздкая процедура. Д. Б. Юдин и А. С. Немировский [76] предложили для решения задачи выпуклого программирования модифицированный метод центрированных сечений (ММЦС). Пусть оптимум  $x^*$  локализован в шаре  $S(x_0, R)$ . Проведем гиперплоскость  $(g(x_0), x - x_0) = 0$ . Область локализации  $x^*$  сужается до полусферы. Опишем вокруг этой полусферы эллипсоид минимального объема, центр которого лежит на луче  $x_0 - hg(x_0)$ ,  $h > 0$ , эллипсоид оказывается сплюснутым в направлении  $g(x_0)$ . Для того чтобы цикл вычислений замкнуть, нужно превратить эллипсоид в шар, для чего достаточно произвести растяжение пространства в направлении  $g(x_0)$  с определенным коэффициентом. Таким образом, схема отсечения естественным путем приводит к алгоритму градиентного типа с растяжением пространства в направлении градиента.

Следует отметить, что схема отсечения позволяет генерировать целое семейство алгоритмов различной сложности. Например, если производить операцию вложения области локализации  $x^*$  в эллипсоид после двух отсечений, то нам потребуется, вообще говоря, растяжение пространства в двух направлениях. Следует ожидать, что скорость сходимости при этом значительно возрастает по сравнению со скоростью сходимости алгоритма (3.57) — (3.60).

### § 9. Модифицированные вычислительные схемы обобщенных градиентных методов с растяжением пространства

В работе [45] В. А. Скоков предложил модифицированные вычислительные схемы обобщенных градиентных методов с растяжением пространства, позволяющие существенно сократить число арифметических операций, связанных с преобразованием пространства. Суть этих схем сводится к тому, что преобразование матрицы  $B_k$  заменяется преобразованием симметричной матрицы  $H_k = B_k B_k^*$ .

Как в алгоритмах ОГСРП, так и в  $r$ -алгоритмах переход от матрицы  $B_k$  к  $B_{k+1}$  и от  $x_k$  к  $x_{k+1}$  производится по формулам вида

$$B_{k+1} = B_k R_\beta(\xi_{k+1}), \quad x_{k+1} = x_k - h_{k+1} B_{k+1}^{-1} \bar{g}_k,$$

где  $\xi_{k+1}$  имеет вид  $\frac{B_k^* p_k}{\|B_k^* p_k\|}$ ,  $p_k \in E_n$ ;  $\bar{g}_k = \frac{B_{k+1}^* g_f(x_k)}{\|B_{k+1}^* g_f(x_k)\|}$ . Воспользовавшись формулой (3.5), получаем

$$\begin{aligned} H_{k+1} &= B_{k+1} B_{k+1}^* = B_k R_\beta (\xi_{k+1}) R_\beta (\xi_{k+1}) B_k^* = B_k R_\beta^2 (\xi_{k+1}) B_k^* = \\ &= B_k R_{\beta^2} (\xi_{k+1}) B_k^* = B_k (I + (\beta^2 - 1) \xi_{k+1} \xi_{k+1}^T) B_k^* = \\ &= H_k + (\beta^2 - 1) B_k \xi_{k+1} \xi_{k+1}^T B_k^* = H_k + (\beta^2 - 1) \frac{B_k B_k^* p_k p_k^T B_k B_k^*}{(B_k^* p_k, B_k^* p_k)} = \\ &= H_k + (\beta^2 - 1) \frac{H_k p_k p_k^T H_k}{(H_k p_k, p_k)}, \\ H_{k+1} &= H_k + \frac{H_k p_k p_k^T H_k}{(H_k p_k, p_k)} (\beta^2 - 1). \end{aligned}$$

Далее,

$$\begin{aligned} x_{k+1} &= x_k - h_{k+1} B_{k+1} \frac{B_{k+1}^* g_f(x_k)}{\|B_{k+1}^* g_f(x_k)\|} = \\ &= x_k - h_{k+1} \frac{H_{k+1} g_f(x_k)}{\frac{1}{2} (H_{k+1} g_f(x_k), g_f(x_k))}, \\ x_{k+1} &= x_k - h_{k+1} \frac{H_{k+1} g_f(x_k)}{\frac{1}{2} (H_{k+1} g_f(x_k), g_f(x_k))}. \end{aligned}$$

Предложенные В. А. Скоковым модификации позволяют вместо  $n^2$  ячеек, необходимых для запоминания матрицы  $B_k$ , использовать  $\frac{n(n+1)}{2}$  ячеек, необходимых для запоминания симметричной матрицы  $H_k$ . Преобразование матрицы  $H_k$  требует примерно  $2n^2$  операций умножения. Определенные дополнительные осложнения при применении описанной схемы связаны с тем, что из-за накопления погрешностей возмущенная матрица  $\tilde{H}_k$  при определенных условиях и большом числе итераций может не быть положительно определенной. При использовании первоначальной схемы вычислений вероятность того, что  $\tilde{H}_k \approx B_k B_k^*$  не является положительно определенной, намного меньше. Для повышения устойчивости схемы В. А. Скокова необходимо введение регуляризирующих поправок в матрицу  $H_k$ , однако этот вопрос пока еще не изучен.

### ПРИЛОЖЕНИЯ МЕТОДОВ НЕГЛАДКОЙ ОПТИМИЗАЦИИ К РЕШЕНИЮ ЗАДАЧ МАТЕМАТИЧЕСКОГО ПРОГРАММИРОВАНИЯ

#### § 1. Использование обобщенных градиентных методов в схемах декомпозиции

При решении задач линейного и выпуклого программирования большого объема в целях уменьшения количества обращений к внешней памяти и, в конечном итоге, экономии машинного времени применяют схемы разложения, которые сводят решения исходной задачи к решению серии задач меньшего объема (блоков). Эти схемы оказываются особенно эффективными, если структура отдельных блоков позволяет применить к ним специальные быстрые методы решения.

В математическом программировании рассматриваются в основном две двойственные между собой схемы разложения: 1) по множеству переменных; 2) по множеству ограничений. Более сложные схемы разложения обычно представляют собой в некотором смысле суперпозицию этих двух.

В последнее время большое внимание уделяется схемам так называемой параметрической декомпозиции, которая заключается в том, что в модель задачи вводятся дополнительные переменные (параметры), при фиксированных значениях которых задача приобретает простую или распадающуюся на блоки структуру.

Основу процесса решения составляет получение оптимальных значений параметров. Параметрическую декомпозицию можно рассматривать как разновидность схем разложения по переменным. Большое число работ посвящено исследованию различных схем разложения в основном для задач линейного программирования [86, 104, 77].

Рассмотрим возможности применения метода обобщенного градиентного спуска для построения итеративных алгоритмов решения задач выпуклого и линейного программирования при использовании некоторых схем разложения.

#### Разложение по множеству переменных

Рассмотрим задачу выпуклого программирования, переменные которой разбиты на два подмножества. Найти

$$\min_{x,y} f_0(x, y) \quad (4.1)$$

при ограничениях

$$f_i(x, y) \leq 0, \quad i = 1, \dots, n, \quad (4.2)$$

где  $x, y$  — векторные переменные:

$$x = (x^{(1)}, \dots, x^{(l)}), \quad y = (y^{(1)}, \dots, y^{(m)});$$

$f_0(x, y), f_i(x, y)$  — выпуклые функции по совокупности переменных  $x, y; i = 1, \dots, n$ . Зафиксируем  $x = \bar{x}$  и рассмотрим следующую задачу: найти

$$\min_y f_0(\bar{x}, y) \quad (4.3)$$

при ограничениях

$$f_i(\bar{x}, y) \leq 0, \quad i = 1, \dots, n. \quad (4.4)$$

Легко видеть, что задача (4.3), (4.4) является задачей выпуклого программирования. Для тех значений  $\bar{x}$ , для которых существует оптимальное решение задачи (4.3), (4.4), определим функцию  $\Phi(\bar{x})$  следующим образом:

$$\Phi(\bar{x}) = \min_{y \in D(\bar{x})} f_0(\bar{x}, y), \quad (4.5)$$

где  $D(\bar{x})$  — множество значений  $y$ , удовлетворяющих (4.4).

**Теорема 4.1.** Если  $f_\alpha(x, y), \alpha = 0, 1, \dots, n$ , — выпуклые по совокупности переменных  $x, y$  функции, то функция  $\Phi(x)$ , определенная формулой (4.5), является выпуклой на некотором выпуклом подмножестве  $W \subseteq E_l$ . Если для некоторого  $\bar{x} \in W$  выполняется условие Слейтера, то обобщенный градиент функции  $\Phi(x)$  в точке  $x = \bar{x}$  вычисляется по формуле

$$g_\Phi(\bar{x}) = g_{L_U}^x(\bar{x}, y(\bar{x})), \quad (4.6)$$

где

$$L_U(x, y) = f_0(x, y) + \sum_{i=1}^n U_i f_i(x, y),$$

$y(\bar{x})$  — одно из оптимальных значений  $y$  в задаче (4.3), (4.4);  $U = \{U_i\}$  — множители Лагранжа задачи (4.3), (4.4), полученные в соответствии с теоремой Куна — Таккера;  $g_{LU}^x(\bar{x}, y(\bar{x}))$  — проекция такого обобщенного градиента функции  $L_U(z) = L_U(x, y)$  на подпространство  $E_1^x$ , у которого проекция на подпространство  $E_m^y$  равна нулю (обобщенный градиент берется в точке  $z = (\bar{x}, y(\bar{x}))$ ).

**Доказательство.** Рассмотрим две точки:  $\bar{x}_1$  и  $\bar{x}_2$  и их выпуклую комбинацию

$$\bar{x}_3 = \lambda_1 \bar{x}_1 + \lambda_2 \bar{x}_2 \quad (\lambda_1 + \lambda_2 = 1, \quad \lambda_1, \lambda_2 \geq 0).$$

Пусть  $\bar{y}_1 = y(\bar{x}_1)$ ,  $\bar{y}_2 = y(\bar{x}_2)$  — оптимальные значения задачи (4.3), (4.4) соответственно при  $\bar{x} = \bar{x}_1$  и  $\bar{x} = \bar{x}_2$ ;  $\bar{y}_3 = \lambda_1 \bar{y}_1 + \lambda_2 \bar{y}_2$ . Тогда в силу выпуклости функций  $f_\alpha(x, y)$  по  $x, y$  ( $\alpha = 0, 1, \dots, n$ )

$$\begin{aligned} f_0(\bar{x}_3, \bar{y}_3) &\leq \lambda_1 f_0(\bar{x}_1, \bar{y}_1) + \lambda_2 f_0(\bar{x}_2, \bar{y}_2); \\ f_i(\bar{x}_3, \bar{y}_3) &\leq \lambda_1 f_i(\bar{x}_1, \bar{y}_1) + \lambda_2 f_i(\bar{x}_2, \bar{y}_2) \leq 0. \end{aligned}$$

Таким образом,  $\bar{y}_3$  является допустимым значением задачи (4.3), (4.4) при  $x = \bar{x}_3$ . Следовательно,

$$\Phi(\bar{x}_3) = \Phi(\lambda_1 \bar{x}_1 + \lambda_2 \bar{x}_2) \leq \lambda_1 \Phi(\bar{x}_1) + \lambda_2 \Phi(\bar{x}_2),$$

что и доказывает выпуклость функции  $\Phi(\bar{x})$ .

Теперь перейдем к выводу формулы (4.6) для обобщенного градиента функции  $\Phi(x)$ . В силу теоремы Куна — Таккера, если  $D(\bar{x})$  удовлетворяет условию Слейтера, существуют множители  $U(\bar{x}) = \{U_i(\bar{x})\}$ ,  $i = 1, \dots, n$ , такие, что выполняется условие

$$\begin{aligned} \Phi(\bar{x}) &= \min_y [f_0(\bar{x}, y) + \sum_{i=1}^n U_i(\bar{x}) f_i(\bar{x}, y)] = \\ &= \max_{U \geq 0} \min_y \left[ f_0(\bar{x}, y) + \sum_{i=1}^n U_i f_i(\bar{x}, y) \right]. \end{aligned}$$

Если условие Слейтера выполняется для области  $D(\bar{x})$ , то оно будет выполняться для области  $D(x)$ , где  $x$  принадлежит некоторой окрестности  $S(x)$  точки  $\bar{x}$  (следствие непре-

ривности функций  $f_i(x, y)$  по  $x, y$ . Пусть  $x \in S(\bar{x})$ ,

$$\begin{aligned} \Phi(x) &= f_0(x, y(x)) + \sum_{i=1}^n U_i f_i(x, y(x)) \geq \\ &\geq f_0(x, y(x)) + \sum_{i=1}^n \bar{U}_i f_i(x, y(x)) \geq L_{\bar{U}}(x, y), \end{aligned}$$

где  $\bar{U} = \{\bar{U}_i\}$  — значения множителей Лагранжа задачи (4.3), (4.4) при  $x = \bar{x}$ . Выберем обобщенный градиент функции  $L_{\bar{U}}(x, y)$  в точке  $x = \bar{x}$ ,  $y = y(\bar{x})$  таким образом, чтобы его проекция  $g_{L_{\bar{U}}}^y(\bar{x}, y(\bar{x}))$  на подпространство  $F_{\bar{x}}^y$  равнялась нулю (это всегда можно сделать, так как

$$L_{\bar{U}}(\bar{x}, y(\bar{x})) = \min_y L_{\bar{U}}(\bar{x}, y),$$

а это означает, что множество обобщенных градиентов в точке  $(\bar{x}, y(\bar{x}))$  пересекает гиперплоскость  $y = 0$ ). Получаем

$$\begin{aligned} \Phi(x) - \Phi(\bar{x}) &\geq L_{\bar{U}}(x, y(x)) - L_{\bar{U}}(\bar{x}, y(\bar{x})) \geq \\ &\geq (x - \bar{x}, g_{L_{\bar{U}}}^x(\bar{x}, y(\bar{x}))) + (y(x) - y(\bar{x}), g_{L_{\bar{U}}}^y(\bar{x}, y(\bar{x}))) = \\ &= (x - \bar{x}, g_{L_{\bar{U}}}^x(\bar{x}, y(\bar{x}))), \end{aligned}$$

откуда по определению обобщенного градиента получаем формулу (4.6).

*Следствие.* Если в условиях предыдущей леммы дополнительно потребовать, чтобы  $f_\alpha(x, y)$ ,  $\alpha = 0, 1, \dots, n$ , были непрерывно дифференцируемыми по  $y$ , то формула (4.6) уточняется следующим образом:

$$g_{\Phi}(\bar{x}) = g_{f_0}^x(\bar{x}, y(\bar{x})) + \sum_{i=1}^n U_i(\bar{x}) g_{f_i}^x(\bar{x}, y(\bar{x})), \quad (4.7)$$

где  $g_{f_\alpha}^x(\bar{x}, y(\bar{x}))$  — проекции произвольных обобщенных градиентов функции  $f_\alpha$  в точке  $\bar{z} = (\bar{x}, y(\bar{x}))$  на  $E_i^x$ ,  $\alpha = 0, 1, \dots, n$ .

**Д о к а з а т е л ь с т в о.** Этот результат следует из того, что в случае непрерывной дифференцируемости функции  $f_\alpha$  по  $y$  проекция  $g_{L_{\bar{U}}}^y(\bar{x}, y(\bar{x}))$  определяется однозначно и равна нулю. Остается только применить формулу, представляющую обобщенный градиент функции  $L_{\bar{U}}$ , являющейся

линейной комбинацией с неотрицательными коэффициентами функций  $f_\alpha$ , через обобщенные градиенты составляющих функций (см. §3 гл. 1). Формула (4.7) позволяет легко строить алгоритмы решения задач выпуклого программирования (в частности, линейного) с использованием схемы разложения по координатам. Для этого мы должны предположить возможность решения за конечное число шагов задачи (4.3), (4.4) с получением множителей Лагранжа  $U = \{U_i\}$ , а также возможность получения «частных» обобщенных градиентов  $g_{i\alpha}^x$ ,  $\alpha = 0, \dots, n$ . Эти предположения реализуемы, например, когда задача (4.3), (4.4) является задачей линейного или квадратичного программирования. Осложнения могут быть связаны с тем, что задача (4.3), (4.4) не при всех  $\bar{x}$  будет иметь решения. Стандартный выход из такой ситуации заключается в использовании метода «штрафных функций», при котором задача (4.1), (4.2) заменяется следующей задачей: найти

$$\min_{x,y,v} \left[ f_0(x, y) + M \sum_{i=1}^n v_i \right] \quad (4.8)$$

при ограничениях

$$f_i(x, y) - v_i \leq 0; \quad v_i \geq 0, \quad (4.9)$$

где  $M$  — достаточно большое положительное число, а задача (4.3), (4.4) заменяется задачей: найти

$$\min_{y,v} \left[ f_0(\bar{x}, y) + M \sum_{i=1}^n v_i \right] \quad (4.10)$$

при ограничениях

$$f_i(\bar{x}, y) - v_i \leq 0; \quad v_i \geq 0. \quad (4.11)$$

Легко видеть, что задача (4.10), (4.11) имеет допустимое решение при любом  $\bar{x}$  с выполнением условия Слейтера.

Таким образом, мы можем предположить, что задача (4.1), (4.2) сведена к такому виду, что функция  $\Phi(\bar{x})$  определена для любого  $\bar{x}$  и имеется возможность вычисления обобщенного градиента  $g_\Phi(x)$  в произвольной точке  $x$ , т. е. возможность применения метода обобщенного градиентного спуска к задаче минимизации функции  $\Phi(x)$ . В результате получаем алгоритм решения задачи (4.1), (4.2),  $k$ -й шаг которого состоит в следующем:

а) для полученного  $\bar{x} = \bar{x}_k$  решаем задачу (4.3), (4.4) и получаем вектор множителей Лагранжа  $U(\bar{x}_k) = (U_i(\bar{x}_k))$ ;

б) вычисляем вектор  $g_{\Phi}(\bar{x}_k)$  в соответствии с формулой (4.6) или в благоприятном случае в соответствии с формулой (4.7);

в) находим  $\bar{x}_{k+1} = \bar{x}_k - h_{k+1} g_{\Phi}(\bar{x}_k)$ ,  $k = 0, 1, \dots$

Выбор последовательности  $\{h_k\}$  производится в соответствии с анализом и рекомендациями гл. 2.

Для частного случая, когда задача (4.1), (4.2) является задачей линейного программирования, опишем выше приведенный алгоритм более подробно. Пусть требуется найти

$$\max [(c, x) + (d, y)] \quad (4.12)$$

при ограничениях

$$Ax + By \leq e, \quad (4.13)$$

$$x \geq 0, \quad y \geq 0 \quad (4.14)$$

(здесь  $x, y$  — переменные векторы размерности соответственно  $l$  и  $m$ ;  $A$  и  $B$  — матрицы размерности  $(n \times l)$  и  $(n \times m)$  соответственно;  $c, d, e$  — постоянные векторы размерности  $l, m$  и  $n$  соответственно). Зафиксируем  $x = \bar{x}$  и рассмотрим следующую задачу: найти

$$\max_y [(c, \bar{x}) + (d, y)] \quad (4.15)$$

при ограничениях

$$By \leq e - A\bar{x}, \quad (4.16)$$

$$y \geq 0 \quad (4.17)$$

и двойственную к ней задачу: найти

$$\min_v (e - A\bar{x}, v) \quad (4.18)$$

при ограничениях

$$B^*v \geq d, \quad (4.19)$$

$$v \geq 0 \quad (4.20)$$

( $v$  — переменный вектор размерности  $n$ ;  $B^*$  — матрица, транспонированная  $B$ ).

Пусть  $\bar{v}(\bar{x})$  — оптимальное решение задачи (4.18) — (4.20),  $\alpha_j$  —  $j$ -й вектор-столбец матрицы  $A$  ( $j = 1, \dots, l$ ).

Определим  $j$ -ю компоненту  $l$ -мерного вектора  $g(\bar{x})$  следующим образом:

$$g_j(\bar{x}) = (-\alpha_j, v(\bar{x})) + c_j; \quad g(\bar{x}) = (g_1(\bar{x}), \dots, g_l(\bar{x})) \quad (4.21)$$

( $c_j$  —  $j$ -я компонента вектора  $c$ ). Тогда итерационный процесс решения задачи (4.12) — (4.14) по выше приведенной схеме разложения можно описать так: зададимся начальным значением  $x_0$ ;  $\bar{x}_k$  — значение  $\bar{x}$ , полученное после  $k$ -го шага процесса, тогда на  $(k+1)$ -м шаге:

1) решаем задачу (4.18) — (4.20) при  $\bar{x} = \bar{x}_k$  и определяем вектор  $v(\bar{x}_k)$ ;

2) определяем вектор обобщенного градиента  $g(\bar{x}_k)$  по формуле (4.21);

3) вычисляем  $\bar{x}_{k+1}$  по формуле

$$\bar{x}_{k+1} = \max(0, \bar{x}_k + h_{k+1}g(\bar{x}_k)), \quad (4.22)$$

где  $h_{k+1}$  — величина шага; операция максимизации понимается покомпонентно.

Описанный метод разложения можно применять для решения различных задач оптимизации и анализа больших транспортных сетей.

**Пример 1.** Определение оптимальных потенциалов транспортной задачи.

Рассмотрим задачу, двойственную к транспортной задаче: найти

$$\max \left[ \sum_{j=1}^m b_j y_j - \sum_{i=1}^l a_i x_i \right] \quad (4.23)$$

при ограничениях

$$y_j - x_i \leq c_{ij}, \quad i = 1, \dots, l; \quad j = 1, \dots, m, \quad (4.24)$$

и условия  $\sum_{i=1}^l a_i = \sum_{j=1}^m b_j$ .

Имеем две группы переменных:  $x = \{x_i\}_{i=1}^l$  и  $y = \{y_j\}_{j=1}^m$ . Следуя выше приведенной схеме разложения, поступаем следующим образом: зафиксируем  $x = \bar{x} = \{\bar{x}_i\}$  и рассмотрим задачу: найти

$$\max_y \left[ \sum_{j=1}^m b_j y_j - \sum_{i=1}^l a_i \bar{x}_i \right] \quad (4.25)$$

при ограничениях

$$y_j \leq c_{ij} + \bar{x}_i \quad (4.26)$$

и двойственную к ней задачу: найти

$$\min_v \sum_{i=1}^l \sum_{j=1}^m (c_{ij} + x_i) v_{ij} \quad (4.27)$$

при ограничениях

$$\sum_{i=1}^l v_{ij} = b_j, \quad (4.28)$$

$$v_{ij} \geq 0. \quad (4.29)$$

Задачи (4.25), (4.26) и (4.27) — (4.29) имеют тривиальные решения при  $b_j \geq 0$

$$y_j(\bar{x}) = \min_i (\bar{x}_i + c_{ij});$$

$$v_{ij}(\bar{x}) = \begin{cases} b_j & \text{для } i = i^*(j), \text{ для которого } \bar{x}_{i^*(j)} + c_{ij} = y_j(\bar{x}); \\ 0 & \text{при других сочетаниях индексов.} \end{cases}$$

Отсюда

$$g_i(\bar{x}) = \sum_{j \in M_i} b_j - a_i,$$

где  $M_i$  — множество индексов  $j$ , для которых  $i^*(j) = i$ .

Таким образом, формальное применение описанной выше схемы разложения приводит к известному алгоритму решения транспортной задачи методом обобщенного градиентного спуска в пространстве потенциалов [54].

**Пример 2.** Многопродуктовая транспортная задача [57]. Пусть задана транспортная сеть  $S$  с вершинами  $i$  и дугами  $d_j$  с пропускными способностями  $r_j$  и стоимостями перевозок вдоль дуги  $S_j$ . Обозначим через  $R [i_1, i_2]$  множество путей  $\{\pi_\alpha\}$  из вершины  $i_1$  в вершину  $i_2$ ; через  $D(\pi_\alpha)$  — множество дуг, из которых состоит путь  $\pi_\alpha$ ;  $K_j$  — множество путей, проходящих через дугу  $d_j$ .

Пусть имеется  $p$  видов продуктов, для каждого  $k$ -го вида продуктов заданы два множества:  $M_k$  — множество вершин, производящих  $k$ -й продукт, и  $N_k$  — множество вершин, его потребляющих ( $k = 1, \dots, p$ ). Для каждой вершины  $i \in M_k$  заданы объемы производства  $a_i^{(k)}$ , для каждой вершины  $i' \in N_k$  — объемы потребления  $b_{i'}^{(k)}$   $k$ -го продукта. Примем, что для всех  $k$

$$\sum_{i \in M_k} a_i^{(k)} = \sum_{i' \in N_k} b_{i'}^{(k)}.$$

Определим  $c(\Pi_\alpha)$  — стоимость перевозки вдоль пути  $\Pi_\alpha$  — следующим образом:

$$c(\Pi_\alpha) = \sum_{j \in D(\Pi_\alpha)} S_j. \quad (4.30)$$

Обозначим через  $x_{ii'}^{(k)}$  [ $\Pi_\alpha$ ] ( $k = 1, \dots, p$ ;  $i \in M_k$ ;  $i' \in N_k$ ;  $\Pi_\alpha \in R[i, i']$ ) количество продукта  $k$ -го вида, поставляемого из пункта  $i$  в пункт  $i'$  по пути  $\Pi_\alpha$ . Тогда многопродуктовую транспортную задачу можно сформулировать в следующем виде: найти

$$\min \sum_{k=1}^p \sum_{i \in M_k} \sum_{i' \in N_k} \sum_{\Pi_\alpha \in R[i, i']} x_{ii'}^{(k)} c(\Pi_\alpha) \quad (4.31)$$

при ограничениях

$$x_{ii'}^{(k)} [\Pi_\alpha] \geq 0, \quad (4.32)$$

$$\sum_{i' \in N_k} \sum_{\Pi_\alpha \in R[i, i']} x_{ii'}^{(k)} [\Pi_\alpha] = a_i^{(k)}, \quad (4.33)$$

$$\sum_{i \in M_k} \sum_{\Pi_\alpha \in R[i, i']} x_{ii'}^{(k)} [\Pi_\alpha] = b_{i'}^{(k)}, \quad (4.34)$$

$$\sum_{k=1}^p \sum_{i \in M_k} \sum_{i' \in N_k} \sum_{\Pi_\alpha \in K_j} x_{ii'}^{(k)} [\Pi_\alpha] \leq r_j. \quad (4.35)$$

Двойственную задачу сформулируем так:  
найти

$$\max \left[ \sum_{k=1}^p \left( \sum_{i' \in N_k} b_{i'}^{(k)} v_{i'}^{(k)} - \sum_{i \in M_k} a_i^{(k)} U_i^{(k)} - \sum_j r_j W_j \right) \right] \quad (4.36)$$

при ограничениях

$$W_j \geq 0, \quad (4.37)$$

$$v_{i'}^{(k)} - U_i^{(k)} - \sum_{i \in D(\Pi_\alpha)} W_i \leq c[\Pi_\alpha]. \quad (4.38)$$

Как и в примере 1, разобьем переменные задачи (4.36) — (4.38) на две группы: к 1-й группе отнесем переменные  $\{U_i^{(k)}\}$  и  $\{W_j\}$ , ко 2-й —  $\{v_{i'}^{(k)}\}$ . Если зафиксировать  $U_i^{(k)} = \bar{U}_i^{(k)}$ ,  $W_j = \bar{W}_j$  и рассмотреть задачу (4.36) — (4.38) относительно  $v_{i'}^{(k)}$ , то получим тривиальное решение

$$\begin{aligned} v_{i'}^{(k)} &= \min_{i, \Pi_\alpha \in R[i, i']} [\bar{U}_i^{(k)} + \sum_{i \in D(\Pi_\alpha)} \bar{W}_i + c[\Pi_\alpha]] = \\ &= \min_{i, \Pi_\alpha \in R[i, i']} [\bar{U}_i^{(k)} + \sum_{j \in D(\Pi_\alpha)} (S_i + \bar{W}_j)]. \end{aligned} \quad (4.39)$$

Также легко определить вектор обобщенного градиента  $g$  ( $\{\bar{U}_i^{(k)}\}, \{W_j\}$ ). Компонента этого вектора  $\Delta_{i^*}^{(k)}$ , относящаяся к  $U_{i^*}^{(k)}$ , равна  $(\sum_{i' \in T_{i^*}} b_{i'}^{(k)} - a_{i^*}^{(k)})$ , где множество  $T_{i^*}$

состоит из индексов  $i'$  таких, что минимум в (4.39) достигается при  $i = i^*$  (если минимум достигается при нескольких значениях  $i$ , то в качестве  $i^*$  берется одно из этих значений).

Компонента вектора  $g$ , относящаяся к  $W_j$ , вычисляется по формуле

$$\Delta_j = \sum_{i', k \in L_j} b_{i'}^{(k)} - r_j,$$

где  $L_j$  — состоит из таких пар индексов  $i', k$ , для которых минимум в (4.39) достигается при  $\Pi_\alpha \in K_j$ .

Таким образом, пользуясь общей схемой разложения, мы построили алгоритм решения задачи, двойственной к многопродуктовой транспортной задаче. Решение частного случая этой задачи, когда каждый продукт поставляет только один поставщик, подробно рассмотрено в работе [57], где применяется также метод обобщенного градиентного спуска.

#### Схема разложения по ограничениям

Схема разложения по ограничениям является в некотором смысле двойственной к схемам разложения по переменным.

Пусть рассматривается задача выпуклого программирования, множество ограничений которой некоторым образом разбито на две группы: найти

$$\min f_0(x) \quad (4.40)$$

при ограничениях

$$f_i(x) \leq 0, \quad i = 1, \dots, m, \quad (4.41)$$

$$\varphi_j(x) \leq 0, \quad j = 1, \dots, n. \quad (4.42)$$

Рассмотрим задачу следующего вида: найти

$$L(u) = \min_{x \in D} \left[ f_0(x) + \sum_{i=1}^m U_i f_i(x) \right],$$

где  $u = \{u_i\}$ ,  $i = 1, \dots, m$ , —  $m$ -мерный вектор с неотрицательными компонентами,  $D$  — множество, вырезаемое в  $E_n$  ограничениями (4.42).

Следствием теоремы Куна — Таккера является следующее утверждение: если решение задачи (4.40) — (4.42) существует

и ограничения этой задачи удовлетворяют условию Слейтера ( $D$  — ограниченное множество), то  $L(u)$  — вогнутая функция, определенная для всех  $u \geq 0$ , максимум функции  $L(u)$  существует и достигается в точке  $u^*$ , при этом  $u^*$  является  $u$ -компонентой седловой точки функции Лагранжа задачи (4.40) — (4.42)

$$\Phi(x, u, v) = f_0(x) + \sum_{i=1}^m u_i f_i(x) + \sum_{j=1}^n v_j \varphi_j(x).$$

Определение максимума вогнутой функции  $L(u)$  эквивалентно определению минимума выпуклой функции  $L_1(u) = -L(u)$ . Минимум выпуклой функции  $L_1(u)$  при ограничении  $u \geq 0$  можно находить, используя метод обобщенного градиентного спуска с проектированием на множество  $\{u : u \geq 0\}$ . Заметим, что операция проектирования вектора на это множество очень проста и сводится к тому, что положительные и нулевые координаты проектируемого вектора остаются неизменными, а отрицательные заменяются нулями. Для частного случая — задачи линейного программирования — приведенная схема разложения подробно рассмотрена в работе [56].

## § 2. Итеративный метод решения задач линейного программирования специального вида

Основная идея итеративного метода заключается в применении специальной схемы разложения и использования обобщенного градиентного спуска для увязки параметров частных задач, на которые разбивается исходная задача. В случае классической задачи составления оптимального плана производства [27] предлагаемый метод имеет простую интерпретацию. Множество технологических способов разбивается на группы  $T_1, \dots, T_l$ , и для каждой группы  $T_k$  подбирается вектор ресурсных ограничений  $b^{(k)}$ ,  $k = 1, \dots, l$ , таким образом, чтобы  $\sum_{k=1}^l b^{(k)} = b$ , где  $b$  — общий вектор ресурсных ограничений. Для каждой группы  $T_k$  решается частичная задача оптимизации; результаты этого решения используются для выбора направления изменения векторов  $b^{(k)}$ .

Обоснование метода производится для задач линейного программирования: найти

$$\max (c, x) \quad (4.43)$$

при ограничениях

$$Ax \leq b, \quad (4.44)$$

$$x \geq 0 \quad (4.45)$$

при условии, что элементы матрицы  $A$  и компоненты векторов  $b$  и  $c$  не отрицательны.

Запишем задачу (4.43) — (4.45) в следующем виде: максимизировать

$$L(x) = (c, x) = \sum_{j=1}^n c_j x_j \quad (4.46)$$

при ограничениях

$$\sum_{j=1}^n a_{ij} x_j \leq b_i, \quad i = 1, \dots, m, \quad (4.47)$$

$$x_j \geq 0, \quad j = 1, \dots, n, \quad (4.48)$$

и предположении, что

$$c_j \geq 0; \quad b_i \geq 0; \quad a_{ij} \geq 0 \quad (i = 1, \dots, m; j = 1, \dots, n). \quad (4.49)$$

Разобьем множество индексов  $N = \{1, \dots, n\}$  на  $l$  непересекающихся непустых подмножеств  $N_t$ ,  $1 \leq t \leq l$ , и рассмотрим частичные задачи: максимизировать

$$L_t(x^{(t)}) = \sum_{j \in N_t} c_j x_j \quad (4.50)$$

при ограничениях

$$\sum_{j \in N_t} a_{ij} x_j \leq b_{it}, \quad i = 1, \dots, m, \quad (4.51)$$

$$x^{(t)} \geq 0 \quad (4.52)$$

(здесь  $x^{(t)} = \{x_j\}$ , где  $j \in N_t$ ;  $t = 1, \dots, l$ ). Если  $b_{it} \geq 0$ ,  $i = 1, \dots, m$ ,  $k = 1, \dots, l$ , то каждая частичная задача (4.50) — (4.52) имеет оптимальный план (мы считаем, что каждый столбец матрицы  $A$  содержит хотя бы один ненулевой элемент, что обеспечивает ограниченность множества

решений частичных задач). Если

$$\sum_{k=1}^n b_{ik} = b_i, \quad i = 1, \dots, m, \quad (4.53)$$

$$b_{ik} \geq 0, \quad k = 1, \dots, l, \quad (4.54)$$

то вектор  $x \{b_{ik}\}$ , составленный из соответствующих компонент оптимальных планов задач (4.50) — (4.52), дает допустимое решение задачи (4.46) — (4.48), причем

$$L(x \{b_{ik}\}) = \sum_{k=1}^l L(\bar{x}^{(k)}),$$

где  $\bar{x}^{(k)}$  — оптимальный план задачи (4.50) — (4.52). Пусть  $\{u_{ik} \{b_{ik}\}\}$ ,  $i = 1, \dots, m$ , — вектор оптимального плана задачи, двойственной к задаче (4.46) — (4.48). Тогда в силу теоремы двойственности

$$\Phi \{b_{ik}\} = L(x \{b_{ik}\}) = \sum_{k=1}^l \left( \sum_{i=1}^m b_{ik} \bar{u}_{ik} \{b_{ik}\} \right). \quad (4.55)$$

Функция  $F = -\Phi \{b_{ik}\}$  является кусочно-линейной, выпуклой, и обобщенный градиент ее в точке  $\{b_{ik}\}$  вычисляется по формуле

$$g_F \{b_{ik}\} = -g_\Phi \{b_{ik}\} = \{-\bar{u}_{ik} \{b_{ik}\}\}.$$

Таким образом, задача (4.46) — (4.48) сводится к задаче минимизации выпуклой функции  $F \{b_{ik}\}$  при ограничениях (4.53), (4.54). Пусть  $D$  — множество, вырезаемое этими ограничениями. Легко видеть, что  $D$  — замкнутый выпуклый многогранник. Рассмотрим последовательность векторов  $\beta_p = \{b_{ik}\}_p$ , полученную по формулам

$$\beta_{p+1} = P_D \{\beta_p + h_{p+1} g_\Phi(\beta_p)\} = P_D \{\beta_p + h_{p+1} u_p\}, \quad (4.56)$$

где  $u_p = \{\bar{u}_{ik} \{b_{ik}\}_p\}$ ,  $\{h_p\}_{p=0}^\infty$  — последовательность положительных чисел,  $P_D$  — оператор проектирования на множество  $D$ . Легко видеть, что с помощью формулы (4.56) осуществляется шаг обобщенного градиентного спуска для минимизации функции  $F(\beta) = F \{b_{ik}\}$  с проектированием на множество ограничений. Как следует из теоремы Б. Т. Поляка [38], если  $h_p \rightarrow 0$  и  $\sum_{p=0}^\infty h_p = +\infty$ , то  $F(\beta_p)$  сходится к минимальному значению.

Рассмотрим последовательность  $x(\{\beta_{ik}\}_D)$ . Из (4.55) вытекает, что  $L(x\{b_{ik}\}_D)$  сходится к оптимальному значению  $L(x)$  задачи (4.46) — (4.48). Опишем подробнее операцию проектирования  $P_D$ , содержащуюся в (4.56). Она заключается в нахождении для произвольного вектора  $\beta$  ближайшего к нему вектора  $P_D(\beta)$ , принадлежащего выпуклому многограннику  $D$ , который можно представить как декартово произведение симплексов:

$$D = D_1 \times D_2 \times \dots \times D_m,$$

где каждый симплекс  $D_i$  определяется системой ограничений

$$\sum_{k=1}^l y_{ik} = b_i, \quad (5.57)$$

$$y_{ik} \geq 0, \quad k = 1, \dots, l. \quad (4.58)$$

Пусть  $\beta_i = \{b_{ik}\}$ ,  $k = 1, \dots, l$ . Тогда операция проектирования  $P_D$  сведется к независимому проектированию векторов  $b_i$  на симплексы  $D_i$  при всех  $i = 1, \dots, m$ . Получаем элементарную задачу квадратичного программирования: минимизировать

$$\rho^2(\{y_{ik}\}, \beta_i) = \sum_{k=1}^l (y_{ik} - b_{ik})^2 \quad (4.59)$$

при ограничениях (4.57), (4.58).

Рассмотрим совокупность задач, зависящих от параметра  $\lambda$ : найти

$$\min \left[ \sum_{k=1}^l (y_{ik} - b_{ik})^2 + 2\lambda \sum_{k=1}^l y_{ik} \right] \quad (4.60)$$

при  $y_{ik} \geq 0$ .

Эта задача имеет легко проверяемое для любого  $\lambda$  решение

$$y_{ik}(\lambda) = \max(0, b_{ik} - \lambda), \quad k = 1, \dots, l. \quad (4.61)$$

Из теоремы Куна — Таккера вытекает, что должно найтись такое  $\lambda^* = \lambda$ , при котором решение (4.61) совпадает с решением задачи (4.57) — (4.59). Поиск  $\lambda^*$  можно произвести за конечное число шагов с помощью следующей процедуры.

Примем

$$\lambda^{(1)} = \frac{\sum_{k=1}^l b_{ik} - b_i}{l}.$$

Далее, для произвольного  $q$ ,  $q = 1, 2, \dots$ , поступим следующим образом. Пусть нам известно  $\lambda^{(q)}$ . Вычислим  $y_{ik}(\lambda^{(q)})$  по формуле (4.61). Найдем

$$S_q = \sum_{k=1}^l y_{ik}(\lambda^{(q)}).$$

Если  $S_q = b_i$ , то принимаем  $\lambda^* = \lambda_q$ . Если  $S_q > b_i$ , принимаем

$$\lambda_{q+1} = \lambda_q + \frac{s_q - b_i}{n_q}, \quad (4.62)$$

где  $n_q$  — количество индексов  $k$ , для которых  $y_{ik}(\lambda_q) > 0$ , и переходим к следующему шагу. Из формул (4.61) и (4.62) легко вытекает, что случай  $s_q < b_i$  невозможен. Кроме того,  $l > n_1 > n_2 > \dots$ . Следовательно, после конечного числа шагов процесс выбора  $\lambda^*$  должен закончиться. Так как  $\lambda^*$  является единственным значением  $\lambda$ , при котором  $\{y_{ik}(\lambda)\}$  удовлетворяет ограничениям (4.57), (4.58),  $\{y_{ik}(\lambda^*)\}$  обеспечивают минимум задачи (4.57) — (4.59), т. е.  $\{y_{ik}(\lambda^*)\}$  реализует проекцию  $\beta_i$  на множество  $D_i$ . Таким образом, на каждом шаге алгоритма решения задачи (4.43) — (4.45) в соответствии с описанной схемой разложения нужно производить следующие операции:

1) при заданных  $\{b_{ik}\}_p$  решить  $l$  частных задач (4.50) — (4.52) и получить двойственные оценки  $\{u_{ik}\}$ ;

2) воспользовавшись формулой (4.56), вычислить  $\{b_{ik}\}_{p+1}$ .

Очевидно, что подобную схему разложения целесообразно применять в том случае, если каждая из частных задач решается сравнительно просто, например когда  $l$  равно  $n$ , т. е. когда мы имеем  $n$  частных задач, но каждая из них содержит одно неизвестное. В этом случае  $j$ -я компонента вектора  $x(\{b_{ik}\})$  вычисляется по формуле

$$x_j(\{b_{ik}\}) = \min_{i \in I_j^+} \frac{b_{ij}}{a_{ij}}, \quad (4.63)$$

где  $I_j^+$  — множество индексов  $i$ , для которых  $a_{ij} > 0$ .

Пусть  $i_j$  — определенный индекс, на котором достигается минимум в формуле (4.63). Тогда

$$u_{ij} = \begin{cases} \frac{c_j}{a_{ij}} & \text{для } i = i_j, \\ 0 & \text{для } i \neq i_j. \end{cases} \quad (4.64)$$

Заметим, что при этом вектор  $g_{\Phi}(\{b_{ij}\})$  имеет не более чем  $n$  ненулевых компонент.

Формулы (4.63), (4.64) дают возможность построить алгоритм, при котором вычисление обобщенного градиента  $g_{\Phi}(\{b_{ij}\})$  предельно упрощается. Используя эти формулы, легко построить итеративный процесс решения задачи линейного программирования указанного выше вида [60].

### § 3. Решение задач распределительного типа методом обобщенного градиентного спуска

К задачам распределительного типа отнесем задачи линейного программирования, которые можно сформулировать в следующем виде: найти

$$\min \sum_{i,\alpha} c_{i\alpha} x_{i\alpha} \quad (4.65)$$

при ограничениях

$$\sum_{\alpha \in M_i} x_{i\alpha} = b_i, \quad i \in I = \{1, \dots, n\}, \quad (4.66)$$

$$\sum_{i \in I; \alpha \in M_{\mu}} p_{i\alpha}^{(\mu)} x_{i\alpha} \leq a^{(\mu)}, \quad \mu \in T, \quad (4.67)$$

$$x_{i\alpha} \geq 0. \quad (4.68)$$

Множество  $I$  будем условно называть множеством заказов, множество  $T$  — множеством технологических ограничений,  $M_i, M_{\mu}$  — подмножества некоторого конечного множества  $M$ . Задачи вида (4.65) — (4.68) очень часто встречаются на практике, особенно при решении вопросов отраслевого планирования. Частными случаями являются обычные транспортные и распределительные задачи ( $\lambda$ - или  $k$ -задачи). Исследование градиентных методов решения задач распределительного типа вызвано необходимостью решения конкретных задач большой размерности, связанных с оптимальным планированием загрузки прокатных станов СССР и распределением продукции между потребителями, причем размерность задач и необходимость их решения на ЭВМ средней производительности (типа М-220 и «Минск-32») диктовали определенные требования к алгоритму, особенно в части использования оперативной памяти ЭВМ. Метод ОГС оказался пригодным для решения указанных задач.

Двойственная задача к задаче (4.65) — (4.68) формулируется следующим образом: найти

$$\max \left( \sum_{i \in I} b_i v_i - \sum_{\mu \in T} a^{(\mu)} u_\mu \right) \quad (4.69)$$

при ограничениях

$$v_i - \sum_{(\mu)} p_{i\alpha}^{(\mu)} u_\mu \leq c_{i\alpha}, \quad (4.70)$$

$$u_\mu \geq 0. \quad (4.71)$$

В силу (4.66) задача (4.65) — (4.68) имеет решение, если  $b_i \geq 0$ . С учетом этого применим к двойственной задаче (4.69) — (4.71) схему разложения по переменным, выделив в одну группу переменные  $\{v_i\}$ , в другую — переменные  $\{u_\mu\}$ . Зафиксировав  $\{u_\mu\}$ , получаем относительно переменных  $\{v_i\}$  следующую задачу: максимизировать

$$\sum_i b_i v_i \quad (4.72)$$

при ограничениях

$$v_i \leq c_{i\alpha} + \sum_{\mu} p_{i\alpha}^{(\mu)} u_\mu. \quad (4.73)$$

При  $b_i \geq 0$  задача имеет тривиальное решение

$$v_i = \min_{\alpha} \left( c_{i\alpha} + \sum_{\mu} p_{i\alpha}^{(\mu)} u_\mu \right). \quad (4.74)$$

Используя результаты § 1 гл. 4, получаем алгоритм решения двойственной задачи к задаче распределительного типа: зададимся начальными значениями  $u_\mu^0 \geq 0$  и после  $k$  шагов решения задачи получим вектор  $\{u_\mu^{(k)}\}$ . Тогда  $(k+1)$ -й шаг состоит в следующем:

1) определяем  $v_i^{(k)}$  по формуле (4.74), принимая  $u_\mu = u_\mu^k$ ;  $\alpha(k)$  — одно из значений  $\alpha$ , на котором достигается минимум в формуле (4.74);

2) принимаем

$$x_{i\alpha}^{(k)} = \begin{cases} b_i, & \alpha = \alpha(k), \\ 0, & \alpha \neq \alpha(k); \end{cases}$$

3) вычисляем «невязки»

$$\Delta_{\mu}^{(k)} = a_{\mu} - \sum_{\alpha} p_{i\alpha}^{\mu} x_{i\alpha}^{(k)}$$

4) вычисляем новые приближения для  $\{u_\mu\}$

$$u_\mu^{(k+1)} = \max(0, u_\mu - h_{k+1} \Delta_\mu^{(k)}).$$

Выбор последовательности шаговых множителей  $\{h_k\}$ , обеспечивающих сходимость процесса к оптимальному решению, если оно существует, обсуждался в § 2 гл. 2.

Таким образом, схема разложения по переменным с использованием метода обобщенного градиентного спуска дает очень простой итеративный алгоритм решения двойственной задачи (4.69) — (4.71).

Не совсем тривиальным является в данном случае вопрос восстановления оптимального решения прямой задачи (4.65) — (4.68) по полученному с определенной степенью точности оптимальному решению двойственной задачи. Теория двойственности линейного программирования подсказывает алгоритм следующего типа для выделения переменных  $x_{i\alpha}$ , входящих в оптимальный базис прямой задачи: задаться достаточно малым  $\delta > 0$  и выделить те пары индексов  $(i, \alpha)$ , для которых выполняется неравенство

$$c_{i\alpha} + \sum_{\mu} p_{i\alpha\mu}^u \bar{u}_\mu - \delta \leq \bar{v}_i, \quad (4.75)$$

где  $\{\bar{u}_\mu, \bar{v}_i\}$  — оценки двойственных переменных, полученные в результате описанного выше итеративного процесса решения задачи (4.69) — (4.71). Получив множество базисных переменных  $x_{i\alpha}$ , уже легко, решив соответствующую систему линейных алгебраических уравнений, получить оптимальное значение  $\{x_{i\alpha}\}$ .

Однако такого рода процедура не всегда является корректной. Если решение задачи (4.65) — (4.69) неоднозначно или имеется ряд базисов со значениями целевой функции, близкими к оптимальному, выбор базисных переменных в соответствии с неравенством (4.75) может дать лишние переменные и наша попытка восстановления оптимального решения прямой задачи может оказаться неудачной. Кроме того, если даже мы получим нужный базис, то для задач большого объема решение соответствующей системы линейных уравнений требует использования либо медленной внешней памяти, либо специальных процедур. Поэтому требовалось найти более простую и надежную процедуру получения решения прямой задачи. Такая процедура была разработана на основе идеи «сглаживания», которую Герстенхабер применил для построения итеративной процедуры

решения транспортной задачи [91], не приводя строгого обоснования.

Образует функцию  $S(u_\mu)$ , подставив в выражение для целевой функции (4.69) вместо  $v_i$  правую часть формулы (4.74):

$$S_\mu(\{u_\mu\}) = \sum_{i \in I} b_i \min_{\alpha} \left( c_{i\alpha} + \sum_{\mu} p_{i\alpha}^{(\mu)} u_\mu \right) - \sum_{\mu} a^{(\mu)} u_\mu.$$

Функция  $S(\{u_\mu\})$  является кусочно-линейной выпуклой. Обобщенный градиент этой функции в точке  $u = \{u_\mu\}$  можно вычислить по формуле

$$g_S(\{u_\mu\}) = \left\{ \sum_i b_i p_{i\alpha^*}^{(\mu)} - a_\mu \right\},$$

где  $\alpha^*$  — одно из значений  $\alpha$ , на котором достигается минимум в формуле (4.74).

Рассмотрим векторное поле  $G = \{g_S(u)\}$ ,  $u = \{u_\mu\}$ . Оно не является непрерывным — точки разрыва соответствуют тем векторам  $u$ , для которых минимум в формуле (4.74) определяется неоднозначно.

Выбрав  $\delta > 0$ , построим « $\delta$ -сглаженное» векторное поле  $G_\delta$ , определяемое в точке  $u$  вектором  $g^\delta(u)$  по следующим формулам:

$$g^\delta(u) = \left\{ \sum_{i,\alpha} p_{i\alpha}^{(\mu)} b_i m_{i\alpha}^{(\delta)}(u) - a_\mu \right\},$$

где

$$m_{i\alpha}^{(\delta)} = \frac{\Delta_{i\alpha}^{(\delta)}}{\sum_i \Delta_{i\alpha}^{(\delta)}},$$

а  $\Delta_{i\alpha}^{(\delta)}$  вычисляется по формулам

$$\Delta_{i\alpha}^{(\delta)} = \begin{cases} 0, & \text{если } r_{i\alpha}^{(\delta)} = \left( c_{i\alpha} + \sum_{\mu} p_{i\alpha}^{(\mu)} u_\mu \right) - \left( c_{i\alpha^*} + \sum_{\mu} p_{i\alpha^*}^{(\mu)} u_\mu \right) \leq \delta, \\ \delta - r_{i\alpha}^{(\delta)} & \text{в противном случае.} \end{cases}$$

Для учета ограничений  $u_\mu \geq 0$  рассмотрим векторное поле  $G_r^\delta$ , определенное векторами вида

$$g_r^\delta(u) = g^\delta(u) + r(u^+),$$

где  $u^+$  — вектор, который получается из вектора  $u$  путем замены его неположительных координат нулем,  $r$  — некоторое положительное число.

**Лемма 4.1.** Векторное поле  $G_r^\delta(u)$  является непрерывным при любых  $\delta > 0$  и  $r > 0$ .

**Доказательство.** Непрерывность векторного поля  $g_r^\delta(u)$  следует из непрерывности функций  $\Delta_{i\alpha}^{(\delta)}(u)$  и  $m_{i\alpha}^{(\delta)}$ , с помощью которых оно определяется. Непрерывность поля  $gu^+$  также легко проверяется.

**Лемма 4.2.** Непрерывное векторное поле  $G_r^\delta(u)$  при любых  $r > 0$ ,  $\delta > 0$  имеет по крайней мере одну неподвижную точку, если задача (4.65) — (4.68) имеет единственное оптимальное решение.

**Доказательство.** Рассмотрим прямую, проходящую через оптимальное решение задачи (4.65) — (4.68) — точку  $x^*$ . Пусть  $x_R^{(1)}$  и  $x_R^{(2)}$  — точки пересечения этой прямой со сферой с центром в точке  $x^*$  радиуса  $R$ . Легко показать, что при достаточно большом  $R$  поле  $G_r^\delta(u)$  образует в точках  $x_R^{(1)}$  и  $x_R^{(2)}$  векторы  $g_1^{(R)}$ ,  $g_2^{(R)}$  такие, что  $(g_1^{(R)}, x_R^{(1)} - x^*) > 0$  и  $(g_2^{(R)}, x_R^{(2)} - x^*) > 0$ . Отсюда следует неколлинеарность векторов  $g_1^{(R)}$  и  $g_2^{(R)}$ . Используя следствие теоремы Борсука о вращении непрерывных векторных полей [29], приходим к выводу о существовании для поля  $G_r^\delta(u)$  неподвижной точки при любом  $r > 0$  и  $\delta > 0$ . Переходя к пределу при  $r \rightarrow \infty$ , получим существование для любого  $\delta > 0$  точки  $u^{*(\delta)}$ , для которой выполняются условия

$$\begin{aligned} \sum_{i,\alpha} p_{i\alpha}^{(\mu)} b_i m_{i\alpha}^{(\delta)} \{u_\mu^{\delta*}\} - a_\mu &= 0, \quad u_\mu^{*(\delta)} > 0, \\ \sum_{i,\alpha} p_{i\alpha}^{(\mu)} b_i m_{i\alpha}^{(\delta)} \{u_\mu^{(\delta)*}\} - a_\mu &\leq 0, \quad u_\mu^{(\delta)*} = 0. \end{aligned} \quad (4.76)$$

Если принять

$$\bar{x}_{i\alpha} = B_i m_{i\alpha}^{(\delta)}(u^{(\delta)*}),$$

то условия (4.76) показывают, что  $\{\bar{x}_{i\alpha}\}$  — допустимый план задачи (4.65) — (4.68). Так как значение функции-критерия любого допустимого плана прямой задачи не меньше значения функции критерия произвольного допустимого плана двойственной задачи, можно написать неравенство

$$\sum c_{i\alpha} x_{i\alpha}^{(\delta)} \geq \sum_{i \in I} \left( b_i \min_{\alpha} \left( c_{i\alpha} + \sum_{\mu} p_{i\alpha}^{(\mu)} u_{\mu}^{(\delta)*} \right) - \sum_{\mu} a_{\mu} u_{\mu}^{(\delta)*} \right).$$

При  $\delta \rightarrow 0$  правая и левая части неравенства стремятся друг к другу, а значит, и к оптимальному решению.

Один из подходов (приемов) получения решения прямой задачи основан на следующих соображениях. Количество ненулевых  $x_{i\alpha}$  в базисном решении задачи (4.65) — (4.68) не превышает  $(n + m)$ . В задачах большого объема, встречающихся на практике (для решения именно таких задач прежде всего меняются итеративные методы),  $n \gg m$ , при этом величина  $n/m$  обычно порядка нескольких десятков, а иногда более сотни. Это означает, что в оптимальном базисном плане задачи (4.65) — (4.68) количество индексов  $i$ , для которых количество ненулевых  $x_{i\alpha}$ ,  $\alpha \in M_i$ , превышает единицу, не больше  $m$ . Заметим, что если  $\{v_i^*, u_\mu^*\}$  — оптимальный план двойственной задачи, а задача (4.65) — (4.68) имеет единственный оптимальный план, то равенство

$$v_i^* = \min_{\alpha} \left\{ c_{i\alpha} + \sum_{\mu} p_{i\alpha}^{(\mu)} u_{\mu} \right\} \quad (4.77)$$

достигается для тех  $\alpha$ , для которых  $x_{i\alpha} > 0$  в оптимальном плане прямой задачи. Поэтому если нам удалось получить с достаточной точностью оптимальные значения  $\{u_{\mu}\}$ , то, исходя из равенства (4.77) и выбрав  $\varepsilon > 0$  с учетом точности определения  $\{u_{\mu}\}$ , получаем приближенный критерий того, что  $x_{i\alpha} > 0$  в оптимальном плане задачи (4.65) — (4.68):

$$v_i - \varepsilon \leq \min_{\alpha} \left\{ c_{i\alpha} + \sum_{\mu} p_{i\alpha}^{(\mu)} u_{\mu} \right\}. \quad (4.78)$$

Пусть  $\bar{I}$  — множество индексов  $I$ , для которых неравенство (4.78) соблюдается для одного  $\alpha = \alpha(i)$ . Тогда принимаем  $x_{i\alpha(i)} = b_i$ ,  $x_{i\alpha} = 0$ , если  $\alpha \neq \alpha(i)$ . После этого естественным путем получаем задачу распределительного типа, у которой множество заказов  $I_1 = I/\bar{I}$ , а множества  $M_i$  состоят из тех индексов  $\alpha$ , для которых выполняется неравенство (4.78). Из приведенных соображений ясно, что число элементов множества  $I_1$  в практически интересных случаях в десятки раз меньше числа элементов множества  $I$ . Таким образом мы приходим к задаче следующего вида: найти

$$\min \sum c_{i\alpha} x_{i\alpha} \quad (4.79)$$

при условиях

$$\sum_{\alpha \in M_i} x_{i\alpha} = b_i, \quad i \in I_1, \quad (4.80)$$

$$\sum_{i \in I_1} \sum_{\alpha \in M_{\mu}} p_{i\alpha} x_{i\alpha} \leq a_{\mu} - \sum_{i \in \bar{I}} b_i p_{i\alpha(i)}, \quad (4.81)$$

$$x_{i\alpha} \geq 0. \quad (4.82)$$

Эту задачу, сравнительно небольшого объема, можно решать, используя, во-первых, некоторый точный алгоритм, эффективный для задач такого объема, во-вторых, схему алгоритма «со сглаживанием». Эти подходы не требуют дополнительных комментариев. Отметим только, что так как объем задачи (4.79) — (4.82) существенно меньше объема задачи (4.65) — (4.68), на ее решение практически требуется 5—10% времени решения основной задачи.

При решении большого числа практических задач распределительного типа методом обобщенного градиентного спуска в пространстве переменных двойственной задачи использовался для «восстановления» решения прямой задачи алгоритм «со сглаживанием».

По стандартной программе, разработанной Г. И. Горбач [58] для ЭВМ типа М-20, решались разнообразные задачи распределительного типа, в частности транспортные задачи большого объема, задачи распределения судов по линиям речных перевозок. Задачи распределительного типа особенно большой размерности решались при планировании загрузки прокатных станков [58]. В этих задачах количество заказов достигало 10 000, число станков превышало 50, номенклатура составляла около 1000 наименований видов продукции. Предложенные алгоритмы были использованы при планировании загрузки прокатных станков в «Союзглавметалле» Госснаба СССР со значительным экономическим эффектом.

#### **§ 4. Опыт решения задач производственно-транспортного характера с использованием обобщенных градиентных методов с растяжением пространства**

Рассмотрим несколько математических моделей задач текущего планирования, которые тем или иным способом сводятся к задачам минимизации выпуклой негладкой функции при простейших ограничениях. В настоящее время сотрудниками Института кибернетики АН УССР (Н. З. Шор, Н. Г. Журбенко, В. И. Гершович, Т. В. Белых, Л. В. Беляева и др.) накоплен определенный опыт решения задач текущего планирования с использованием  $r$ -алгоритмов, применение которых во многих случаях весьма эффективно. Рассмотрим несколько наиболее характерных примеров.

1. Динамическая транспортная задача. Имеется  $m$  поставщиков (предприятий) и  $n$  потребителей (строек). Планируемый период разбит на  $T$  интервалов. Для краткости интервал планирования будем называть кварталом. Заданы величины  $a_{it}$  — объем продукции, производимый  $i$ -м предприятием за квартал  $t$ ,  $t = \overline{1, T}$ ;  $a_i$  — объем продукции, производимый  $i$ -м предприятием за планируемый период,  $a_i = \sum_{t=1}^T a_{it}$ ;  $b_{jt}$  — объем спроса на продукцию  $j$ -го потребителя за квартал  $t$ ,  $t = \overline{1, T}$ ;  $b_j$  — потребность  $j$ -й стройки за планируемый период,  $b_j = \sum_{t=1}^T b_{jt}$ ;  $c_{ij}$  — стоимость перевозки единицы продукции от  $i$ -го предприятия к  $j$ -му потребителю. Транспортные издержки  $c_{ij}$  принимаются независимыми от времени доставки. Предполагается, что общее производство и потребление сбалансированы:

$$\sum_{i=1}^m a_i = \sum_{j=1}^n b_j,$$

однако баланс поквартально, вообще говоря, не имеет места:

$$\sum_{i=1}^m a_{it} \text{ не обязательно равна } \sum_{j=1}^n b_{jt}.$$

На интуитивном уровне задача заключается в определении такого распределения продукции между потребителями, чтобы при этом выполнялись условия:

а) для большинства предприятий или по крайней мере для предприятий с большим приоритетом удовлетворяется их потребность по кварталам;

б) объем преждевременно поставленной продукции должен быть по возможности меньшим (это условие можно интерпретировать как минимизацию затрат на хранение преждевременно произведенной продукции);

в) транспортные издержки минимальны по сравнению с другими распределениями, удовлетворяющими первым двум условиям;

г) потребители в течение планируемого периода должны по возможности прикрепляться к одному поставщику.

Введем обозначения:  $\beta_{jt}$  — объем продукции, являющейся спросом  $j$ -го потребителя за первые  $t$  кварталов,

$$\beta_{jt} = \sum_{\tau=1}^t b_{j\tau}, \quad t = \overline{1, T};$$

$\alpha_{it}$  — объем продукции, производимой  $i$ -м предприятием за первые  $t$  кварталов,

$$\alpha_{it} = \sum_{\tau=1}^t a_{i\tau}, \quad t = \overline{1, T};$$

$x_{ijt}$  — объем продукции, поставляемой  $i$ -м предприятием  $j$ -му потребителю за первые  $t$ -кварталов.

Предположим, что все  $\beta_{jt} \neq 0$ , однако отметим, что это ограничение, как будет видно из дальнейшего, не является существенным и введено лишь для простоты изложения. Окончательные результаты справедливы для случая  $\beta_{jt} \geq$

$\geq 0$ . Величину  $\frac{x_{ijt}}{\beta_{jt}}$  назовем степенью обеспеченности  $j$ -го потребителя  $i$ -м предприятием за первые  $t$  кварталов.

При построении математической модели дополним постановку задачи, не изменяя существенно ее экономической сути, следующими условиями:

1) преждевременно произведенная продукция предприятия хранится на его складах, т. е. потребители не получают излишка продукции:

$$\sum_{i=1}^m x_{ijt} \leq \beta_{jt}, \quad t = \overline{1, T-1};$$

2) степень обеспеченности потребителя не убывает со временем:

$$\frac{x_{ijt}}{\beta_{jt}} \geq \frac{x_{ijt-1}}{\beta_{jt-1}}, \quad t = \overline{2, T}. \quad (4.83)$$

Введем штрафные коэффициенты:  $R_{jt}$  — величина штрафа за недопоставку единицы продукции  $j$ -му потребителю в конце  $t$ -го квартала;  $l_{it}$  — величина штрафа за наличие единицы нераспределенной продукции на  $i$ -м предприятии в конце  $t$ -го квартала. Через  $\omega_{it}$  обозначим объем нераспределенной продукции на  $i$ -м предприятии в конце  $t$ -го квартала.

Математическая модель поставленной задачи описывается следующей задачей линейного программирования:

$$\min \left[ \sum_{j,t=1}^{n,m} c_{jt} x_{ijt} + \sum_{j,t=1}^{n,T-1} R_{jt} \left( \beta_{jt} - \sum_{i=1}^m x_{ijt} \right) + \sum_{i,t=1}^{m,T-1} l_{it} \omega_{it} \right]; \quad (4.84)$$

$$\sum_{i=1}^m x_{iT} = b_j, \quad j = \overline{1, n}, \quad (4.85)$$

$$\sum_{j=1}^n x_{iT} = a_i, \quad i = \overline{1, m}, \quad (4.86)$$

$$\beta_{jt-1} x_{ijt} \geq \beta_{jt} x_{ijt-1}, \quad i = \overline{1, m}, \quad j = \overline{1, n}, \quad t = \overline{2, T}; \quad (4.87)$$

$$\sum_{j=1}^n x_{ijt} + \omega_{it} = \alpha_{it}, \quad i = \overline{1, m}, \quad t = \overline{1, T-1}, \quad (4.88)$$

$$\sum_{i=1}^m x_{ijt} \leq \beta_{jt}, \quad j = \overline{1, n}, \quad t = \overline{1, T-1}, \quad (4.89)$$

$$x_{ijt} \geq 0, \quad t = \overline{1, T}, \quad i = \overline{1, m}, \quad j = \overline{1, n}, \quad (4.90)$$

$$\omega_{it} \geq 0, \quad t = \overline{1, T-1}, \quad i = \overline{1, m}. \quad (4.91)$$

**З а м е ч а н и я.** 1. Ограничения (4.83), имеющие смысл и для  $\beta_{jt} = 0$ , соответствуют условию (4.87).

2. Хотя естественное условие  $x_{ijt+1} \geq x_{ijt}$  не введено формально в список ограничений задачи, однако оно является следствием ограничений (4.87), (4.89) и (4.90), поскольку  $\beta_{jt+1} \geq \beta_{jt}$ .

3. Отметим, что условие (4.87) запрещает прерывать обслуживание потребителя данным поставщиком, что «способствует» выполнению условия г). Более точная формулировка требования г) приводила бы к задаче целочисленного программирования, получение удовлетворительного решения которой было бы весьма проблематичным в случае большой размерности задачи.

Опишем метод решения поставленной задачи. Решение задачи (4.84) — (4.91) будет проводиться в два этапа. На первом этапе определяется схема оптимального прикрепления потребителей к поставщикам, на втором — объемы поквартальных поставок  $x_{ijt}$ . Нахождение схемы оптимального прикрепления потребителей сводится, как и в случае классической транспортной задачи, к решению двойственной задачи.

**З а м е ч а н и е.** Под схемой оптимального прикрепления мы понимаем выделение пар индексов  $(i, j)$ , для которых  $x_{iT} > 0$ .

Прежде чем переходить к двойственной задаче запишем задачу (4.84) — (4.91) в другой эквивалентной форме:

определить

$$\min \left[ \sum_{j,i=1}^{n,m} c_{ij} x_{ijt} + \sum_{j,\tau=1}^{n,T-1} \left( \sum_{t=1}^{\tau} \beta_{jt} R_{jt} \right) \left( \sum_{i=1}^m y_{ij\tau} \right) + \sum_{i,t=1}^{m,T-1} l_{it} \omega_{it} \right], \quad (4.92)$$

$$\sum_{i=1}^m x_{ijt} = b_j, \quad j = \overline{1, n}, \quad (4.93)$$

$$\sum_{j=1}^n x_{ijt} = a_i \quad i = \overline{1, m}, \quad (4.94)$$

$$\sum_{j=1}^n \beta_j \sum_{\tau=0}^{t-1} y_{ij\tau} + \omega_{it} = \alpha_{it}, \quad i = \overline{1, m}, \quad t = \overline{1, T-1}, \quad (4.95)$$

$$-\frac{1}{b_j} x_{ijt} + \sum_{\tau=0}^{t-1} y_{ij\tau} = 0, \quad i = \overline{1, m}, \quad j = \overline{1, n}, \quad (4.96)$$

$$x_{ijt} \geq 0, \quad i = \overline{1, m}, \quad j = \overline{1, n}, \quad (4.97)$$

$$y_{ij\tau} \geq 0, \quad i = \overline{1, m}, \quad j = \overline{1, n}, \quad \tau = \overline{0, T-1}, \quad (4.98)$$

$$\omega_{it} \geq 0, \quad i = \overline{1, m}, \quad t = \overline{1, T-1}. \quad (4.99)$$

Нетрудно показать, что задача (4.92) — (4.99) эквивалентна задаче (4.84) — (4.91), причем решение  $\{x_{ijt}\}$  задачи (4.84) — (4.91) определяется формулой

$$x_{ijt} = \beta_j \sum_{\tau=0}^{t-1} y_{ij\tau}, \quad i = \overline{1, m}, \quad j = \overline{1, n}, \quad t = \overline{1, T},$$

где  $\{y_{ij\tau}\}$  — решение задачи (4.92) — (4.99).

Покажем, например, справедливость равенства

$$\sum_{\tau=1}^{T-1} \left( \sum_{t=1}^{\tau} \beta_{jt} R_{jt} \right) \left( \sum_{i=1}^m y_{ij\tau} \right) = \sum_{\tau=1}^{T-1} R_{j\tau} (\beta_{j\tau} - \tilde{x}_{j\tau}),$$

где

$$\tilde{x}_{j\tau} = \sum_{i=1}^m x_{ij\tau}.$$

Положив

$$r_{j\tau} = \sum_{i=1}^m \beta_{jt} R_{jt}, \quad \tau = \overline{1, T-1}, \quad j = \overline{1, m},$$

и учитывая (4.96), получаем

$$\begin{aligned}
 & \sum_{\tau=1}^{T-1} \left( \sum_{i=1}^{\tau} \beta_{i\tau} R_{i\tau} \right) \left( \sum_{i=1}^m y_{i\tau} \right) = \sum_{\tau=1}^{T-1} r_{j\tau} \sum_{i=1}^m \left( \frac{x_{ij\tau+1}}{\beta_{j\tau+1}} - \frac{x_{ij\tau}}{\beta_{j\tau}} \right) = \\
 & = \sum_{\tau=1}^{T-1} r_{j\tau} \left( \frac{\tilde{x}_{j\tau+1} - \beta_{j\tau+1}}{\beta_{j\tau+1}} - \frac{\tilde{x}_{j\tau} - \beta_{j\tau}}{\beta_{j\tau}} \right) = \\
 & = \sum_{\tau=1}^{T-1} r_{j\tau} \left( \frac{\beta_{j\tau} - \tilde{x}_{j\tau}}{\beta_{j\tau}} \right) - \sum_{\tau=1}^{T-1} r_{j\tau} \left( \frac{\beta_{j\tau+1} - \tilde{x}_{j\tau+1}}{\beta_{j\tau+1}} \right) = \\
 & = \sum_{\tau=1}^{T-1} r_{j\tau} \left( \frac{\beta_{j\tau} - \tilde{x}_{j\tau}}{\beta_{j\tau}} \right) - \sum_{\tau=1}^{T-1} (r_{j\tau+1} - \beta_{j\tau+1} R_{j\tau+1}) \frac{\beta_{j\tau+1} - \tilde{x}_{j\tau+1}}{\beta_{j\tau+1}} = \\
 & = r_{j1} \frac{\beta_{j1} - \tilde{x}_{j1}}{\beta_{j1}} + \sum_{\tau=2}^{T-1} R_{j\tau} (\beta_{j\tau} - \tilde{x}_{j\tau}) = \sum_{\tau=1}^{T-1} R_{j\tau} (\beta_{j\tau} - \tilde{x}_{j\tau}),
 \end{aligned}$$

что и требовалось показать.

Обозначим через  $v_j$ ,  $u_{iT}$ ,  $u_{it}$ ,  $z_{ij}$  двойственные переменные, отвечающие соответственно ограничениям (4.93) — (4.96). Тогда задача, двойственная к (4.92) — (4.99), будет состоять в следующем: найти

$$\max \left( \sum_{j=1}^n b_j v_j - \sum_{i,t=1}^{m,T} \alpha_{it} u_{it} \right), \quad (4.100)$$

$$v_j \leq c_{ij} + \frac{1}{b_j} z_{ij} + u_{iT}, \quad i = \overline{1, m}, \quad j = \overline{1, n}, \quad (4.101)$$

$$z_{ij} \leq \sum_{t=1}^{T-1} \beta_{jt} u_{it}, \quad i = \overline{1, m}, \quad j = \overline{1, n}, \quad (4.102)$$

$$z_{ij} \leq \sum_{t=1}^{\tau} \beta_{it} R_{jt} + \sum_{t=\tau+1}^{T-1} \beta_{jt} u_{it}, \quad (4.103)$$

$$i = \overline{1, m}, \quad j = \overline{1, n}, \quad \tau = \overline{1, T-2},$$

$$z_{ij} \leq \sum_{t=1}^{T-1} \beta_{it} R_{jt}, \quad i = \overline{1, m}, \quad j = \overline{1, n}, \quad (4.104)$$

$$u_{it} \geq -l_{it}, \quad i = \overline{1, m}, \quad t = \overline{1, T-1}. \quad (4.105)$$

Пусть  $v_j^*$ ,  $u_{iT}^*$ ,  $z_{ij}^*$  — решение задачи (4.100) — (4.105). Тогда, очевидно, выполняется соотношение

$$v_j^* = \min_{i=\overline{1, m}} \left[ c_{ij} + \frac{1}{b_j} z_{ij}^* + u_{iT}^* \right]. \quad (4.106)$$

Обозначим через  $I^*(j)$  множество индексов  $i$ , при которых достигается минимум в правой части формулы (4.106).

Поскольку переменные задачи (4.92) — (4.99)  $x_{ijt}$  являются двойственными оценками для системы ограничений (4.101), то  $x_{ijt} > 0$  лишь для  $i \in I^*(j)$ . Поэтому  $I^*(j)$  является множеством номеров предприятий, к которым прикрепляется  $j$ -й потребитель. Таким образом, для определения схемы оптимального прикрепления потребителей к пунктам производства достаточно решить двойственную задачу (4.100) — (4.105). Отметим, что в случае  $R_{jt} = 0$ ,  $l_{it} = 0$ ,  $i = \overline{1, m}$ ,  $j = \overline{1, n}$ ,  $t = \overline{1, T-1}$ , (4.100) — (4.105) является двойственной к классической транспортной задаче, определяемой параметрами  $c_{ij}$ ,  $a_i$ ,  $b_j$ .

Для решения задачи (4.100) — (4.105) применим схему декомпозиции по множеству переменных. Зафиксировав переменные  $u_{it}$ , приходим к задаче: определить

$$\max \sum_{j=1}^n b_j v_j, \quad (4.107)$$

$$v_i \leq c_{ij} + \frac{1}{b_j} z_{ij} + u_{iT}, \quad i = \overline{1, m}, j = \overline{1, n},$$

$$z_{ij} \leq r_{ij\tau}(u), \quad i = \overline{1, m}, j = \overline{1, n}, \tau = \overline{0, T-1},$$

где

$$r_{ij0}(u) = \sum_{t=1}^{T-1} \beta_{jt} u_{it},$$

$$r_{ij\tau}(u) = \sum_{t=1}^{\tau} \beta_{jt} R_{jt} + \sum_{t=\tau+1}^{T-1} \beta_{jt} u_{it}, \quad \tau = \overline{1, T-2},$$

$$r_{ijT-1}(u) = \sum_{t=1}^{T-1} \beta_{jt} R_{jt}.$$

Решение (4.107) определяется, очевидно, формулами

$$z_{ij}(u) = \min_{\tau=\overline{0, T-1}} r_{ij\tau}(u), \quad (4.108)$$

$$v_i(u) = \min_{i=\overline{1, m}} \left[ c_{ij} + \frac{1}{b_j} z_{ij}(u) + u_{iT} \right]. \quad (4.109)$$

Таким образом, задача (4.100) — (4.105) сводится к задаче максимизации кусочно-линейной вогнутой функции  $\tilde{F}(u)$ :

$$\max \tilde{F}(u) = \sum_{j=1}^n b_j v_j(u) - \sum_{i,t=1}^{m,T} \alpha_{it} u_{it} \quad (4.110)$$

при ограничениях

$$u_{it} \geq l_{it}, \quad i = \overline{1, m}, \quad t = \overline{1, T-1}. \quad (4.111)$$

Рассмотрим задачу безусловной максимизации также кусочно-линейной вогнутой функции  $F(u)$

$$\max_u F(u), \quad (4.112)$$

где

$$F(u) = \sum_{j=1}^n b_j v_j(u) - \sum_{i,t=1}^{m, T-1} \alpha_{it} \max\{u_{it}, -l_{it}\} - \sum_{t=1}^m \alpha_{iT} u_{iT},$$

а  $v_j(u)$  по-прежнему определяется формулами (4.108) — (4.109).

Пусть  $\tilde{u}^*$  — решение задачи (4.112). Тогда нетрудно видеть, что решение  $u^*$  задачи (4.110) — (4.111) определяется формулами

$$\begin{aligned} \tilde{u}_{it}^* &= \max\{u_{it}^*, -l_{it}\}, \quad t = \overline{1, T-1}, \\ \tilde{u}_{iT}^* &= u_{iT}^*. \end{aligned}$$

Итак, задача определения схемы прикрепления потребителей к пунктам производства сводится к задаче безусловной максимизации (4.112) кусочно-линейной вогнутой функции, для решения которой можно применить одну из модификаций метода оптимизации с растяжением пространства —  $r$ -алгоритм, опробованный на задачах минимизации кусочно-гладких функций [60, 19].

Задача определения поставок  $x_{ij}$  сводится к проблеме определения решения прямой задачи (4.92) — (4.99) по найденному решению двойственной задачи (4.100) — (4.105).

Опишем один общий метод получения решения прямой задачи, а затем покажем схему применения этого метода к нашей задаче.

Пусть задана разрешимая задача выпуклого программирования, удовлетворяющая условиям теоремы Куна — Таккера,

$$\left. \begin{aligned} \min_{x \in \Omega} f(x), \\ g_i(x) \leq 0, \quad i = \overline{1, m}. \end{aligned} \right\} \quad (4.113)$$

Запишем двойственную задачу

$$\max_{\lambda_i \geq 0} \min_{x \in \Omega} L(x, \lambda), \quad (4.114)$$

где  $L(x, \lambda) = f(x) + \sum_{i=1}^m \lambda_i g_i(x)$  — функция Лагранжа.

Легко показать, что функция  $\psi(\lambda) = \min_{x \in \Omega} L(x, \lambda)$  вогнута. Предположим, что  $\psi(\lambda)$  — собственная функция, т. е. определена для всех  $\lambda \geq 0$ . Для этого достаточно потребовать ограниченность  $\Omega$ .

Пусть  $x(\lambda)$  — решение (вообще говоря, не единственное) задачи  $\min_{x \in \Omega} L(x, \lambda)$ , т. е.  $\psi(\lambda) = L(x(\lambda), \lambda)$ . Тогда нетрудно показать, что  $g(x(\lambda)) = (g_1(x(\lambda)), \dots, g_m(x(\lambda)))$  является обобщенным градиентом  $\psi'(\lambda)$  в функции  $\psi(\lambda)$  в точке  $\lambda$ .

Пусть  $\lambda^*$  — решение двойственной задачи;  $x^*$  — решение прямой задачи, соответствующее  $\lambda^*$ , т. е. выполняются соотношения  $\lambda_i^* g_i(x^*) = 0$ ,  $i = \overline{1, m}$ . Из теоремы двойственности следует, что  $\psi(\lambda^*) = f(x^*)$ .

Определим множество  $M(\lambda^*)$ :

$$M(\lambda^*) = \{x(\lambda^*)/L(x(\lambda^*), \lambda^*) = f(x^*)\}. \quad (4.115)$$

Очевидно,  $M(\lambda^*)$  — выпуклое замкнутое множество, так как  $L(x^*, \lambda^*) = f(x^*)$ , то  $x^* \in M(\lambda^*)$ . Поэтому если  $M(\lambda^*)$  состоит из единственной точки (например,  $L(x, \lambda^*)$  строго выпукла по  $x$ ), то для определения  $x^*$  достаточно решить задачу  $\min_{x \in \Omega} L(x, \lambda^*)$ .

Рассмотрим случай, когда (4.113) является задачей линейного программирования. Предположим, что все  $\lambda_i^* > 0$  (ограничения, для которых  $\lambda_i^* = 0$ , несущественны и их можно вывести из формулировки прямой задачи). Для решения двойственной задачи  $\max_{\lambda} \psi(\lambda)$  применим метод обобщенного градиентного спуска

$$\lambda^{(k+1)} = \lambda^{(k)} + h_k \psi'(\lambda^{(k)}), \quad k \geq 1, \quad (4.116)$$

при условиях

$$h_k \geq 0; \quad h_k \rightarrow 0; \quad \sum_{k=1}^{\infty} h_k = \infty, \quad (4.117)$$

обеспечивающих сходимость  $\{\lambda^{(k)}\}$  к решению  $\lambda^*$ . Определим последовательность  $\{x_N\}$  формулой

$$x_N = \sum_{k=1}^N \mu_k^N x^{(k)}. \quad (4.118)$$

Здесь и далее  $x^{(k)} = x(\lambda^{(k)})$ ,  $\mu_k^N = \frac{h_k}{\sum_{i=1}^N h_i}$ . Так как  $\sum_{k=1}^N \mu_k^N = 1$ ,

$\mu_k^N \geq 0$ , а  $\Omega$  — выпуклое множество, то  $x_N \in \Omega$ .

Покажем, что любая предельная точка  $\tilde{x}$  последовательности  $\{x_N\}$  является решением прямой задачи (4.113).

Из замкнутости множества  $\Omega$  следует, что  $\tilde{x} \in \Omega$ . Покажем, что  $g(\tilde{x}) = 0$ .

Из линейности  $g_i(x)$  имеем

$$g(x_N) = \sum_{k=1}^N \mu_k^N g(x^{(k)}) = \sum_{k=1}^N \mu_k^N \psi'(\lambda^{(k)}).$$

Из (4.116) следует равенство

$$\sum_{k=1}^N \mu_k^N \psi'(\lambda^{(k)}) = \frac{\lambda^{(N+1)} - \lambda^{(1)}}{\sum_{i=1}^N h_i},$$

поэтому, учитывая, что  $\lambda^{(k)} \rightarrow \lambda^*$  и  $\sum_{i=1}^N h_i \rightarrow \infty$ , получаем

$\lim_{N \rightarrow \infty} g(x_N) = 0$ , откуда следует требуемый результат.

Докажем, что  $f(\tilde{x}) = f(x^*)$ . Легко видеть, что для случая задачи линейного программирования функция  $\psi(\lambda)$  является кусочно-линейной вогнутой, откуда следует существование такой окрестности  $D(\lambda^*)$  точки  $\lambda^*$ , что для  $\forall \lambda \in D(\lambda^*)$  будет  $G(\lambda) \subset G(\lambda^*)$ , где  $G(\lambda)$  — множество обобщенных градиентов функции  $\psi(\lambda)$  в точке  $\lambda$ . Так как  $\lambda^{(k)} \rightarrow \lambda^*$ , то, начиная с некоторого  $k$ , все  $\lambda^{(k)} \in D(\lambda^*)$ . Поэтому без ограничения общности можно считать, что вся последовательность  $\{\lambda^{(k)}\} \subset D(\lambda^*)$ .

Выше указано, что  $\psi'(\lambda) = g(x(\lambda))$ . Оказывается, верно в некотором смысле обратное утверждение: если  $\psi'(\lambda_0) = g(x(\lambda_1))$ , то  $x(\lambda_1) \in M(\lambda_0)$ , где  $M(\lambda_0)$  определяется аналогично  $M(\lambda^*)$  в (4.115). Отсюда, так как  $g(x^k) = \psi'(\lambda^k) \subset G(\lambda^*)$ ,  $x^k \in M(\lambda^*)$ . Но  $M(\lambda^*)$  — выпуклое

множество, поэтому  $x_N \in M(\lambda^*)$ , а из замкнутости  $M(\lambda^*)$  получаем  $\tilde{x} \in M(\lambda^*)$ . Следовательно,  $f(\tilde{x}) = f(x^*)$ ,  $\tilde{x} \in \Omega$ ,  $g(\tilde{x}) = 0$ ,  $f(\tilde{x}) = f(x^*)$ , а это означает, что  $\tilde{x}$  является решением задачи (4.113), что и требовалось доказать.

Опишем схему применения изложенного метода получения решения прямой задачи применительно к паре задач (4.92) — (4.99) и (4.100) — (4.105). Прежде всего выделим множество потребителей, для которых  $x_{ijt}$  находится непосредственно из решения  $v^*$ ,  $u^*$ ,  $z^*$  задачи (4.100) — (4.105).

Пусть  $I^*(j)$  — множество индексов  $i$ , для которых достигается минимум в формуле (4.109) при  $u = u^*$ . Через  $T_i^*(j)$  обозначим множество индексов  $t$ , для которых достигается минимум в (4.108) при  $u = u^*$ , через  $I_1$  — множество потребителей, для которых  $I^*(j)$  состоит из единственного элемента  $i^*(j)$ , через  $W$  — множество пар индексов  $(i, j)$ ,  $i \in I^*(j)$ , для которых  $T_i^*(j)$  состоит лишь из одного элемента  $\tau_{ij}$ .

Определим множество  $I_2$  формулой

$$I_2 = \{j | j \in I_1, (i^*(j), j) \in W\}.$$

Нетрудно видеть, что для  $j \in I_2$

$$x_{ijt} = \begin{cases} 0, & i \neq i^*(j), \\ 0, & t \leq \tau_{ij}^*, \\ \beta_{it}, & t > \tau_{ij}^*, i = i^*(j). \end{cases}$$

Таким образом, план поставок  $\{x_{ijt}\}$  для потребителей из  $I_2$  определяется непосредственно из решения двойственной задачи (4.100) — (4.105). Для определения плана поставок остальных потребителей необходимо решить задачу (4.92) — (4.99), в которой предварительно исключены переменные, соответствующие потребителям  $j$  из  $I_2$ , т. е. величины  $a_i$ ,  $\alpha_{it}$  заменены на  $(a_i - \sum_{j \in I_2} x_{ijt})$  и  $(\alpha_{it} - \sum_{j \in I_2} x_{ijt})$  соответственно, а индекс  $j$  принимает значения из  $\tilde{I} = [\bar{I}, \tilde{n}] / I_2$ .

Нетрудно получить следующую оценку для  $\tilde{n} = |\tilde{I}|$ :  $\tilde{n} \leq mT$ . Поэтому, определив схему оптимального прикрепления потребителей, приходим к задаче типа (4.92) — (4.99) сравнительно небольшого объема ( $n = \tilde{n}$ ), для решения которой применим описанный выше метод.

Функцию Лагранжа задачи (4.92) — (4.99) будем строить относительно ограничений (4.94), (4.95), обозначая, как и раньше, соответствующие множители через  $u_{i\tau}$  и  $u_{ii}$ . Для того чтобы множество  $\Omega$ , задаваемое системой оставшихся ограничений, было ограничено, введем дополнительное ограничение  $w_{ii} \leq \alpha_{ii}$ , являющееся следствием системы всех ограничений задачи (4.92) — (4.99). Тогда двойственная задача будет состоять в следующем:

$$\max_u \Psi(u), \quad (4.119)$$

где

$$\Psi(u) = \min_{x \in \Omega} L(x, u)$$

(здесь для краткости через  $x$  обозначено множество переменных  $\{x_{ij\tau}\}$ ,  $\{y_{ij\tau}\}$ ,  $\{w_{ii}\}$ ); множество  $\Omega$  задается системой ограничений

$$\sum_{i=1}^m x_{ij\tau} = b_j, \quad -\frac{1}{b_j} x_{ij\tau} + \sum_{\tau=0}^{T-1} y_{ij\tau} = 0, \quad w_{ii} \leq \alpha_{ii},$$

$$x_{ij\tau} \geq 0, \quad y_{ij\tau} \geq 0, \quad w_{ii} \geq 0,$$

а функция Лагранжа  $L(x, u)$  определена стандартным способом.

Нетрудно видеть, что решение  $x(u)$  задачи  $\min_{x \in \Omega} L(x, u)$  определяется следующим образом:

$$w_{ij} = \begin{cases} \alpha_{ii}, & u_{ii} + l_{ii} \leq 0, \\ 0, & u_{ii} + l_{ii} > 0; \end{cases}$$

$$y_{ij\tau} = \begin{cases} 1, & i = i^*(j), \tau = \tau^*(j), \\ 0, & i \neq i^*(j) \text{ или } \tau \neq \tau^*(j), \end{cases}$$

где  $i^*(j)$ ,  $\tau^*(j)$  — какая-нибудь пара индексов  $(i, \tau)$ , для которых достигается минимум в выражении

$$\min_{i=1, m} [c_{ij} + u_{i\tau} + \min_{\tau} r_{ij\tau}(u)],$$

$r_{ij\tau}(u)$  определено так же, как в (4.108):

$$x_{ij\tau} = \begin{cases} b_j; & i = i^*(j); \\ 0, & i \neq i^*(j). \end{cases}$$

Таким образом, вычисление значений функции  $\psi(u)$ , как и ее обобщенного градиента, не представляет трудностей.

Поэтому для решения задачи (4.119) можно применить алгоритм обобщенного градиентного спуска и по формуле (4.118) получить решение исходной задачи (4.92) — (4.99). Соответствующие программы разработаны Н. Г. Журбенко и Л. В. Беляевой.

2. Распределение коксующихся углей между коксохимическими заводами. Это задача средней размерности. При различных исходных данных она решалась нами с контролируемой точностью. Как известно, эффективность применения обобщенных градиентных методов по сравнению с конечными методами линейного программирования растет, как правило, при увеличении объема задачи. Однако даже для этой сравнительно небольшой задачи использование алгоритмов градиентного типа с растяжением пространства оказалось довольно эффективным. Соответствующие программы разработаны и практически опробованы В. И. Гершвицем [9].

Экономико-математическая модель распределения углей на коксование предложена сотрудниками Главного информационно-вычислительного центра Минчермета УССР и имеет следующий вид:

$$\min \sum_j \sum_k \sum_i c_{ij} x_{ijk}, \quad (4.120)$$

$$\sum_i x_{ijk} \leq Q_{jk}, \quad j \in J, k \in K, \quad (4.121)$$

$$\sum_j x_{ijk} \leq \alpha_i^k \theta_i, \quad i \in I, k \in K, \quad (4.122)$$

$$\underline{z}_i \theta_i \leq \sum_j \sum_k z_{jk} x_{ijk} \leq \bar{z}_i \theta_i, \quad i \in I, \quad (4.123)$$

$$\underline{S}_i \theta_i \leq \sum_j \sum_k s_{jk} x_{ijk} \leq \bar{S}_i \theta_i, \quad i \in I, \quad (4.124)$$

$$\underline{P}_i \theta_i \leq \sum_j \sum_k p_{lk} x_{ijk}, \quad i \in I, \quad (4.125)$$

$$x_{ijk} \geq 0, \quad i \in I, j \in J, k \in K. \quad (4.126)$$

Здесь индекс  $i \in I$  соответствует номеру КХЗ (коксохимического завода)  $I = \{1, 2, \dots, NI\}$ ; индекс  $j \in J$  — номеру шахты  $J = \{1, 2, \dots, Nj\}$ ;  $k \in K$  — номеру марки угля,  $k = \{1, 2, \dots, Nk\}$ . Таким образом,  $x_{ijk}$  — объем перевозки с  $j$ -й шахты на  $i$ -й КХЗ угля марки  $k$ ;  $Q_{jk}$  — производство угля марки  $k$  на шахте с номером  $j$ ;  $\theta_i$  — суммарная

(по маркам) потребность в угле  $i$ -го КХЗ;  $\alpha_i^k$  — доля  $k$ -й марки в суммарной потребности  $i$ -го КХЗ; произведение  $\alpha_i^k \theta_i$  есть потребность  $i$ -го КХЗ в  $k$ -й марке;  $z_{jk}$  ( $s_{jk}$ ) — процент золы (серы) в угле  $k$ -й марки, добываемом на  $j$ -й шахте;  $p_{jk}$  — аналогичная характеристика пластичности;  $\underline{Z}_i$  ( $\underline{S}_i$ ) — минимально допустимый процент золы (серы) во всем угле, поставляемом на  $i$ -й КХЗ;  $\bar{Z}_i$  ( $\bar{S}_i$ ) — максимально допустимый процент золы (серы) во всем угле, поставляемом на  $i$ -й КХЗ;  $\underline{P}_i$  — аналогичная характеристика пластичности.

Перепишем задачу (4.120) — (4.126) в более компактном виде, введя очевидные переобозначения:

$$\min \sum_i \sum_j c_{ij} x_{ij}, \quad (4.127)$$

$$\sum_{i=1}^{NI+1} x_{ij} = Q_j, \quad j \in J, \quad (4.128)$$

$$\sum_{j \in J_k} x_{ij} = \alpha_i^{(k)} \theta_i, \quad i \in I, \quad (4.129)$$

$$\sum_j a_j^{(l)} x_{ij} \leq b_i^l, \quad l \in L_1, i \in I, \quad (4.130)$$

$$x_{ij} \geq 0. \quad (4.131)$$

Отметим два обстоятельства, использованных нами при переходе к такому описанию.

1. На каждой шахте  $J$  добывается лишь одна марка угля. При этом множество шахт  $J$  разбивается на непересекающиеся подмножества  $J = \bigcup_k J_k$ . (Заметим, что это ограничение несущественно, так как задачу всегда можно привести к такому виду введением вместо шахты, обладающей несколькими марками, соответствующей группы шахт.)

2. От неравенства (4.121) к равенству (4.128) переходим посредством формального введения еще одного завода (его номер  $i = NI + 1$ ). Этому заводу, в отличие от других КХЗ, не соответствуют никакие технологические ограничения, и стоимость перевозок угля на этот завод из любой шахты полагаем равной нулю. Более того, замечая, что ограничения (4.129) и (4.130) имеют аналогичную структуру, и выбирая соответствующим образом коэффициенты  $a_j^i$  (фактически увеличиваем количество строк матрицы  $\{a_j^i\} = A$ ,

$l$  — строки,  $j$  — столбцы), перепишем задачу в виде

$$\min \sum_{i=1}^{Nl} \sum_{j=1}^{Nl} c_{ij} x_{ij}, \quad (4.132)$$

$$\sum_{i=1}^{Nl+1} x_{ij} = Q_j, \quad j \in J, \quad (4.133)$$

$$\sum_{j=1}^{Nl} a_j^l x_{ij} \leq b_i^l, \quad i \in I; \quad l \in L_1, \quad (4.134)$$

$$\sum_{j=1}^{Nl} a_j^l x_{ij} = b_i^l, \quad i \in I, \quad l \in L_2, \quad (4.135)$$

$$x_{ij} \geq 0, \quad (4.136)$$

$L = L_1 \cup L_2$ ,  $L_1 = \{1, 2, \dots, L\}$ ,  $L_2 = \{L+1, \dots, NL\}$ .

Рассмотрим задачу, двойственную к (4.132) — (4.136),

$$\max \left[ \sum_{j=1}^{Nl} Q_j v_j - \sum_{i=1}^{Nl} \sum_{l=1}^{NL} b_i^l u_{il} \right], \quad (4.137)$$

$$v_j - \sum_{l=1}^{NL} a_j^l u_{il} \leq c_{ij}, \quad i \in I, \quad j \in J, \quad (4.138)$$

$$u_{il} \geq 0, \quad i \in I, \quad l \in L_1. \quad (4.139)$$

Ограничениям (4.133) соответствуют двойственные переменные  $v_j$ , ограничениям (4.134) — (4.135) — двойственные переменные  $u_{il}$ .

Пусть фиксирован некоторый набор двойственных переменных  $\{\bar{u}_{il}, i \in I, l \in L\}$ . Полученная редуцированная (содержащая лишь переменные  $v_j$ ) задача имеет вид

$$\max \left[ \sum_{j=1}^{Nl} Q_j v_j - \sum_{i=1}^{Nl} \sum_{l=1}^{NL} b_i^l \bar{u}_{il} \right],$$

$$v_j \leq c_{ij} + \sum_{l=1}^{NL} a_j^l \bar{u}_{il}, \quad i \in I, \quad j \in J,$$

и тривиальное решение

$$v_j = \min_i \left\{ c_{ij} + \sum_{l=1}^{NL} a_j^l \bar{u}_{il} \right\}. \quad (4.140)$$

В частности, если фиксированный набор  $\{\bar{u}_{il}\}$  таков, что  $\bar{u}_{il} = u_{il}^*$ , где  $\{v_j^*, u_{il}^*\}$  — одно из оптимальных решений

задачи (4.137) — (4.139), то  $v_j^*$  определяются через  $u_{il}^*$  по формуле (4.140).

Для нахождения оптимального набора  $\{v_j^*, u_{il}^*\}$  воспользуемся непосредственно построениями и результатами § 1 гл. 4. Разбиваем систему ограничений задачи (4.132) — (4.136) на две группы. В первую войдут ограничения вида (4.133), во вторую — ограничения вида (4.134), (4.135). Ограничения второй группы, т. е. технологические, используем при построении функции Лагранжа, которая в этом случае имеет вид

$$L(x, u) = \sum_{i,j} c_{ij} x_{ij} + \sum_{i,l} u_l^{(i)} \left( \sum_{j=1}^{NJ} a'_{ij} x_{ij} - b_i^l \right).$$

Оптимальная для задачи (4.137) — (4.139) точка  $u^* = \{u_{il}^*\}$  есть точка максимума вогнутой функции

$$L^*(u) = \min_{x \in D} \left[ \sum_i \sum_j c_{ij} x_{ij} + \sum_i \sum_l \left( \sum_j a'_{ij} x_{ij} - b_i^l \right) \right] \quad (4.141)$$

при простейших ограничениях  $u_i^l \geq 0$ ;  $i \in I$ ;  $l \in L_i$ . Здесь область  $P$  задается ограничениями (4.133). Из выражения

$$L^*(u) = \sum_i \sum_j c_{ij} x_{ij}^*(u) + \sum_i \sum_l \left( \sum_j a'_{ij} x_{ij}^*(u) - b_i^l \right),$$

где  $x^*(u) = \{x_{ij}^*(u)\}$  — оптимальный вектор задачи (4.141), вытекает, что обобщенный градиент функции  $L^*(u)$  в точке  $u$  равен

$$g(u) = \left\{ \sum_j a'_{ij} x_{ij}^*(u) - b_i^l \right\}.$$

Указав способ вычисления обобщенного градиента  $g(u)$  в любой точке  $u$ , можем применить для максимизации  $L^*(u)$  процедуру обобщенного градиентного спуска или любую ее модификацию с ускорением. При решении данной задачи использовалась процедура обобщенного градиентного спуска с растяжением пространства вдоль разности двух последовательных градиентов [67]. Отметим также, что отыскание вектора  $x^*(u)$  — оптимального решения задачи (4.141), необходимого при вычислении обобщенного градиента  $g(u)$ , тривиально. Действительно, перепишем задачу (4.141) в виде

$$\min_{x \in D} \left[ \sum_{i=1}^{NI} \sum_{j=1}^{NJ} \left( c_{ij} + \sum_{l=1}^{NL} a_l u_{il} \right) x_{ij} - \sum_{i=1}^{NI} \sum_{l=1}^{NL} b_l u_{il} \right], \quad (4.142)$$

где область  $D$  задается равенствами (4.133), т. е.

$$\sum_{i=1}^{Nl+1} x_{ij} = Q_j. \quad (4.143)$$

Решение этой задачи

$$x_{ij}^* = \begin{cases} Q_j & \text{для одного } i^* \in I_j^*; \\ 0 & \text{при других сочетаниях индексов,} \end{cases}$$

где

$$I_j^* = \{i^* / \tilde{c}_{i^*j} = \min \tilde{c}_{ij}\},$$

$$\tilde{c}_{ij} = c_{ij} + \sum_{l=1}^{NL} a_{il}^l u_{il}.$$

**Случай несовместной системы ограничений.** Опыт проведения экспериментальных расчетов показал, что система ограничений задачи распределения коксующихся углей (4.120) — (4.126) часто оказывается несовместной. Такая несовместность возникает из-за избытка общего количества золы (либо серы) по сравнению с предельно допустимым в поставляемых углях. В этом случае будем искать приближенное решение задачи, т. е. план, оптимизирующий значение целевой функции и удовлетворяющий ограничениям (4.121), (4.122) и, возможно, нарушающий ограничения (4.123) — (4.125) не более чем на некоторую заранее указанную величину. Нам удобнее воспользоваться записью задачи в форме (4.127) — (4.131) и двойственной задачей (4.137) — (4.139).

Несовместность исходной задачи (4.127) — (4.131) по ограничениям (4.130) эквивалентна неограниченности двойственной задачи вследствие неограниченности допустимой области двойственной задачи (4.137) — (4.139) по переменным  $\{u_{il}, i \in I, l \in L_1\}$ .

Рассмотрим задачу (в дальнейшем будем называть ее квазидвойственной)

$$\max \left( \sum_{j=1}^{NJ} Q_j v_j - \sum_{i=1}^{NI} \sum_{l=1}^{NL} b_{il}^l u_{il} \right), \quad (4.144)$$

$$v_j - \sum_{l=1}^{NL} a_{il}^l u_{il} \leq c_{ij}, \quad i \in I, j \in J, \quad (4.145)$$

$$u_{il} \geq 0, \quad i \in I, l \in L_1, \quad (4.146)$$

$$u_{il} \leq d_{il}, \quad i \in I, l \in L_1. \quad (4.147)$$

Квазидвойственная задача отличается от двойственной задачи (4.137) — (4.139) лишь дополнительными ограничениями (4.147). В отличие от двойственной эта задача имеет уже конечное решение.

Наряду с квазидвойственной рассмотрим задачу (будем называть ее квазипрямой)

$$\min \left[ \sum_{i=1}^{NI} \sum_{j=1}^{NJ} c_{ij} x_{ij} + \sum_{i=1}^{NI} \sum_{l \in L_i} d_{il} y^{il} \right], \quad (4.148)$$

$$\sum_{i=1}^{NI+1} x_{ij} = Q_j, \quad j \in J, \quad (4.149)$$

$$\sum_{j=1}^{NJ} a_j^l x_{ij} - y^l \leq b_i^l, \quad i \in I, \quad l \in L_1, \quad (4.150)$$

$$\sum_{j=1}^{NJ} a_j^l x_{ij} = b_i^l, \quad i \in I, \quad l \in L_2, \quad (4.151)$$

$$x_{ij} \geq 0, \quad y^l \geq 0, \quad i \in I, \quad j \in J, \quad l \in L_1. \quad (4.152)$$

Квазипрямая (4.144) — (4.147) и квазидвойственная (4.148) — (4.152) задачи составляют пару взаимно двойственных задач, причем дополнительным переменным  $y^l$  задачи (4.148) — (4.152) соответствуют дополнительные ограничения (4.144) задачи (4.144) — (4.147).

Аналогично предыдущим построениям будем решать квазидвойственную задачу (4.144) — (4.147), сводя ее к решению задачи максимизации вогнутой функции

$$L_R^*(u) = \min_{x \in D} [L(x, u) - R(u)],$$

где

$$R(u) = \sum_{i=1}^{NI} \sum_{l \in L_i} r_{il} (u_{il} - d_{il})^+$$

есть штрафная функция, обеспечивающая учет ограничений (4.147), введенных при переходе от двойственной задачи к квазидвойственной ( $r_{il} > 0$ ).

По аналогии с предыдущим находим вектор  $u^*$  такой, что

$$L_R^*(u^*) = \max_{u^l \geq 0} L_R^*(u), \quad (4.153)$$

используя процедуру обобщенного градиентного спуска с растяжением пространства. Затем по формулам (4.140) находим оставшуюся компоненту  $v^*$  оптимального решения квазидвойственной задачи. Таким образом получаем вектор  $\{u^*, v^*\}$  — оптимальное решение задачи (4.144) — (4.147) (см. § 1 гл. 4).

Пусть  $\{x^*, y^*\}$  — оптимальное решение квазипрямой задачи (4.148) — (4.152). Если решения задач (4.144) — (4.147) и (4.153) совпадают (а это справедливо при правильном выборе штрафных коэффициентов  $r_{il}$ ), то  $y_i^l \leq r_{il}$  [80]. Но переменные  $y_i^l$  как раз и выражают меру нарушения ограничений (4.140). Это дает возможность за счет выбора штрафных коэффициентов  $r_{il}$  в известной степени задавать возможную меру нарушения этих ограничений.

**Решение прямой задачи.** Ранее основное внимание было уделено нахождению оптимального решения  $\{u^*, v^*\}$  задачи (4.137) — (4.139), двойственной к исходной задаче (4.127) — (4.131). Так как в конечном счете нас интересует оптимальное решение  $x^*$  прямой задачи, остановимся на вопросах его получения (или, что то же, восстановления) по известному  $\{u^*, v^*\}$ . Вектор  $x^*$  получаем при решении исходной задачи (4.127) — (4.131) с учетом информации об оптимальном решении  $\{u^*, v^*\}$  задачи (4.137) — (4.139). Эта информация позволяет значительно редуцировать исходную задачу. Действительно, фактически неопределенными являются лишь те компоненты  $x_{ij}^*$  вектора  $x^*$ , для которых соответствующие ограничения из (4.138) выполняются как строгие равенства при подстановке в них значений  $\{u^*, v^*\}$ . Остальные же компоненты вектора  $x^*$  заведомо равны нулю. Кроме того, те ограничения прямой задачи, которым соответствуют нулевые компоненты вектора  $\{u^*, v^*\}$ , несущественны, т. е. могут не приниматься во внимание.

Конечно, учитывая приближенность метода и вычислений на ЭВМ, выражения «выполняться как строгие равенства» и «нулевые компоненты вектора  $\{u^*, v^*\}$ » должны пониматься в несколько ином смысле. Так, неизвестные компоненты  $x_{ij}^*$  включаются в редуцированную задачу, если соответствующее ограничение из (4.138) двойственной задачи выполняется как равенство с некоторой точностью  $\varepsilon > 0$ , т. е. если

$$0 \leq c_{ij} - v_j^* + \sum_{l=1}^{NL} a_{jl}^l u_{il}^* \leq \varepsilon.$$

Аналогично некоторое ограничение прямой задачи не входит в редуцированную задачу, если соответствующая компонента оптимального двойственного вектора не превосходит некоторого малого  $\delta > 0$ .

Таким образом, восстановление прямых переменных представляет собой уже лишь нахождение конкретных величин

перевозок в соответствии со связями, установленными при решении двойственной задачи. А фактическая размерность редуцированной задачи равна количеству таких связей.

Как показывают экспериментальные расчеты, количество получаемых связей на порядок меньше количества переменных в исходной задаче. Поэтому полученная редуцированная задача имеет малую размерность и может быть легко решена любым конечным методом. В данной работе для этой цели использовался простейший вариант симплекс-метода. Время счета редуцированной задачи составляет 5—10% общих затрат машинного времени.

Наконец отметим, что в случае несовместности исходной системы ограничений все положения этого пункта остаются справедливыми, если всюду вместо прямой задачи (4.127) — (4.131) понимать квазипрямую (4.148) — (4.152), вместо двойственной (4.137) — (4.139) — квазидвойственную (4.144) — (4.147), а вместо вектора  $x^*$  — вектор  $(x^*, y^*)$ .

**Результаты экспериментальных расчетов.** Как уже указывалось выше, для решения центральной при данном подходе задачи (4.153) использовалась процедура обобщенного градиентного спуска с растяжением пространства вдоль разности двух последовательных градиентов [67]. При этом применялась схема алгоритма с симметричной матрицей, предложенная в работе [45], с некоторыми отличиями, связанными с выбором момента очередного растяжения пространства и регулировкой шагового множителя.

Кратко остановимся на этих вопросах. Допускалась возможность совершения нескольких шагов вдоль выбранного в текущем пространстве направления. Движение в данном направлении прекращалось в точке, шаг из которой в этом направлении заведомо вел к ухудшению функционала (в такой точке проекция градиента на направление движения отрицательна). В этот момент и происходило очередное растяжение пространства вдоль разности последнего полученного градиента и градиента, определившего предыдущее направление движения.

Очевидно, такую стратегию можно интерпретировать как приближенный поиск максимума по направлению. Мы не добивались монотонности по функционалу, так как требование монотонности вело к значительному увеличению количества вычислений функционала и существенно не влияло на скорость сходимости.

Величина шагового множителя регулировалась в зависимости от количества шагов вдоль выбранного направления. Через каждые  $k_1$  шагов вдоль этого направления до момента очередного растяжения пространства шаговый множитель увеличивался в  $q_1 \geq 1$  раз. Если же вдоль выбранного направления делался лишь один шаг, шаговый множитель уменьшался в  $q_2 \leq 1$  раз.

В численных расчетах применялись следующие наборы величин: 1)  $k_1 = 5$ ;  $q_1 = 1,2$ ;  $q_2 = 0,95$ ; 2)  $k_1 = 3$ ;  $q_1 = 1,2$ ;  $q_2 = 1$ .

Во втором случае уменьшение шага в исходном пространстве происходило лишь за счет уменьшения матрицы преобразования пространства. Расчеты проводились по реальным исходным данным, представленным Минчерметом УССР и Укрглавуглем. Алгоритм реализован в виде программы на языке FORTRAN-IV из ДОС ЕС.

Количество КХЗ  $N_I = 9$ , шахт  $N_J = 72$ , «технологических» ограничений на КХЗ  $N_L = 8$ . Таким образом, в исходную задачу входило  $N_I \times N_J = 648$  переменных, а градиентный спуск осуществлялся по  $N = N_I \times N_J = 648$  двойственным переменным  $u_{il}$ . Как показали экспериментальные расчеты, для определения  $u^*$ , дающего верные связи в исходной задаче, при различных наборах исходной информации требуется не более 500—600 итераций (т. е. около  $8N$ ). При этом необходимы 700—1200 вычислений максимизируемой функции, т. е. при поиске по направлению в среднем делалось не более двух шагов. Использование величины коэффициента растяжения пространства  $\alpha = 3$  предпочтительнее величины  $\alpha = 2$ .

Приведем пример изменения функционала в одном из расчетов, если начальная точка  $u = 0$ , начальный шаг  $h = 50$ , коэффициент растяжения  $\alpha = 3$ ,  $k_1 = 3$ ,  $q_1 = 1,2$ ,  $q_2 = 1$ :

Итер	$f \cdot 10^{-6}$	Итер	$f \cdot 10^{-6}$	Итер	$f \cdot 10^{-6}$
0	-0,171	200	0,386	400	0,41211
20	0,216	220	0,406	420	0,41336
40	0,309	240	0,413	440	0,41476
60	0,334	260	0,4146	460	0,41523
80	0,349	280	0,4151	480	0,41555
100	0,359	300	0,4152	500	0,41566
120	0,365	320	0,4155	520	0,41575
140	0,369	340	0,41561	540	0,41585
160	0,374	360	0,41569	570	0,41588
180	0,378	380	0,41564	600	0,41592

Потребовалось 693 вычисления функционала. Время счета на ЭВМ М-4030 55 мин.

Как показали эксперименты, время, затрачиваемое на пересчет матрицы при очередном преобразовании пространства, втрое больше времени подсчета функционала. Естественной поэтому представляется проблема сохранения информации о преобразованиях пространства в форме некоторой совокупности векторов и скалярных произведений и не в обычной матричной форме (которая, по-видимому, избыточна). Если же использовать матричную форму, то необходимо максимально экономно (с точки зрения количества операций) организовать блок пересчета матрицы при очередном преобразовании пространства.

Наконец, обратим внимание на наблюдаемое в экспериментах ухудшение сходимости в некоторый момент счета. В приведенном выше примере это происходило между 360 и 400 итерациями. Такое ухудшение свойств сходимости алгоритма связано с потерей положительной определенности матрицы преобразований из-за ограниченной точности вычислений на ЭВМ, которая начинает существенно сказываться, когда определитель матрицы становится достаточно малым. Информация о преобразовании пространства таким образом искажается и приходится восстанавливать матрицу, заменяя ее единичной. Сходимость по функционалу в этот момент ухудшается. Для борьбы с указанным эффектом незначительно корректировались диагональные элементы матрицы с целью улучшения свойств ее положительной определенности. Эффективным средством, по-видимому, является двойная точность вычислений.

3. Задача выбора состава и распределения гражданских самолетов по авиалиниям [24]. В упрощенной постановке, когда расчеты ведутся на заданный год (так называемая статическая модель), математическая модель имеет следующий вид: минимизировать суммарные годовые затраты

$$c = \sum_{i \in J} \sum_{i \in I} c_{ij} x_{ij} - \sum_{i \in I_1} c_i r_i \quad (4.154)$$

при ограничениях

$$\sum_{i \in I} p_{ij} x_{ij} \geq d_j, \quad j \in J, \quad (4.155)$$

$$\sum_{i \in I} v_{ij} x_{ij} \geq b_j, \quad j \in J_1, \quad (4.156)$$

$$\sum_{i \in J_k} \sum_{i \in I} v_{ij} x_{ij} \leq B_k, \quad k \in K, \quad (4.157)$$

$$\sum_{j \in J} x_{ij} \leq a_i, \quad i \in I_1, \quad (4.158)$$

$$\sum_{j \in J} x_{ij} \geq r_i, \quad i \in I_2, \quad (4.159)$$

$$x_{ij} \geq 0, \quad i \in I, \quad j \in J, \quad (4.160)$$

где  $x_{ij}$  — число самолетов  $i$ -го типа на  $j$ -й авиалинии;  $c_{ij}$  — суммарные годовые затраты самолета  $i$ -го типа на  $j$ -й авиалинии (тыс. руб. в год);  $p_{ij}$  — среднегодовая производительность самолета  $i$ -го типа на  $j$ -й авиалинии (пассажиров в год);  $v_{ij}$  — среднегодовая частота полетов самолета  $i$ -го типа на  $j$ -й авиалинии (рейсов в год).

Задача (4.154) — (4.160) является задачей линейного программирования большой размерности (число неизвестных порядка  $13 \cdot 1400$ , число ограничений  $2 \cdot 1400$ ). Наиболее целесообразным подходом к решению таких задач является использование схем декомпозиции в сочетании с точными или итеративными методами решения отдельных блоков [56].

Запишем задачу, двойственную к (4.154) — (4.160),

$$\max_{Z, W, U \geq 0} \psi(Z, W, U), \quad (4.161)$$

где

$$\begin{aligned} \psi(Z, W, U) = \min_{x \in R} & \left\{ \sum_{j \in J} \sum_{i \in I} c_{ij} x_{ij} + \right. \\ & + \sum_{k \in K} Z_k \left( \sum_{j \in J_k} \sum_{i \in I} v_{ij} x_{ij} - B_k \right) + \\ & \left. + \sum_{i \in I_1} W_i \left( \sum_{j \in J} x_{ij} - a_i \right) + \sum_{i \in I_2} U_i \left( r_i - \sum_{j \in J} x_{ij} \right) \right\}. \end{aligned}$$

Множество  $R$  задается системой неравенств (4.155), (4.156). Легко увидеть, что

$$\begin{aligned} \psi(Z, W, U) = \varphi(Z, W, U) - \sum_{k \in K} B_k Z_k - \\ - \sum_{i \in I_1} a_i W_i + \sum_{i \in I_2} r_i U_i, \end{aligned}$$

где

$$\varphi(Z, W, U) = \min_{x \in R} \sum_{j \in J} \sum_{i \in I} \tilde{c}_{ij}(Z, W, U) x_{ij},$$

$$\tilde{c}_{ij}(Z, W, U) = c_{ij} + \sum_{k \in K} v_{ik} Z_k \delta_{ik} + \tilde{W}_i - \tilde{U}_i,$$

$$\delta_{jk} = \begin{cases} 1, & \text{если } j \in J_k, \\ 0, & \text{если } j \notin J_k; \end{cases}$$

$$\tilde{W}_i = \begin{cases} W_i, & \text{если } i \in I_1, \\ 0, & \text{если } i \notin I_1; \end{cases}$$

$$\tilde{U}_i = \begin{cases} U_i, & \text{если } i \in I_2, \\ 0, & \text{если } i \notin I_2. \end{cases}$$

Обозначим через  $R_j$  множество, определяемое неравенствами (4.155) и (4.156) при фиксированном индексе  $j$ . Получаем

$$\varphi(Z, W, U) = \sum_{j \in J} \min_{x \in R_j} \sum_{i \in I} \tilde{c}_{ij}(Z, W, U) x_{ij},$$

т. е.

$$\varphi(Z, W, U) = \sum_j \sum_i \tilde{c}_{ij}(Z, W, U) y_{ij}^*,$$

где  $y_{ij}^*$  является решением следующей задачи: определить

$$\min_{y_{ij} \in R_j} \sum_{i \in I} \tilde{c}_{ij} y_{ij}^*, \quad j \in J. \quad (4.162)$$

Для  $j \in J/J_1$  решение задач (4.162) определяется формулой

$$y_{ij}^* = \begin{cases} 0, & \text{если } i \neq i^*, \\ \frac{d_i}{p_{ij}}, & \text{если } i = i^*, \end{cases}$$

где  $i^*$  — индекс, для которого отношение  $\frac{\tilde{c}_{ij}}{p_{ij}}$  принимает минимальное значение.

Для  $j \in J$  задачи (4.162) представляют собой задачи линейного программирования небольшого объема (10—13 переменных и два ограничения), и для их решения целесообразно применить одну из модификаций симплекс-метода.

Чтобы задачу (4.161) свести к задаче безусловной максимизации, достаточно произвести замену переменных  $Z_k = |Z'_k|$ ,  $W_i = |W'_i|$ ,  $U_i = |U'_i|$ . Хотя при такой замене переменных функций  $\tilde{\psi}(Z'_k, W'_i, U'_i) = \psi(|Z'_k|, |W'_i|, |U'_i|)$  не является вогнутой, однако все ее локальные максимумы будут соответствовать решению задачи (4.161).

Алгоритм решения задачи (4.154) — (4.160) состоит из двух этапов. На первом этапе с помощью  $r$ -алгоритма находится решение задачи (4.161)  $\{Z^*, W^*, U^*\}$ . Обобщенный градиент функции  $\psi(Z, W, U)$ , используемый при работе  $r$ -алгоритма, вычисляется по формулам

$$\psi'_{Z_k} = \sum_{j \in J_k} \sum_{i \in I} y_{ij} v_{ij} - B_k, \quad k \in K,$$

$$\psi'_{W_i} = \sum_{j \in J} y_{ij}^* - a_i, \quad i \in I_1,$$

$$\psi'_{U_i} = - \sum_{j \in J} y_{ij}^* + r_j, \quad i \in I_2.$$

В процессе анализа результатов первого этапа работы алгоритма выявляются те авиалинии, для которых  $y_{ij}^*(Z^*, W^*, U^*) = x_{ij}^*$  — решение задачи (4.154) — (4.160), т. е. авиалинии, для которых оптимальное распределение находится непосредственно из решения двойственной задачи (4.161). Эти авиалинии исключаются из задачи (4.154) — (4.160). Полученную в результате задачу (ее структура совпадает с (4.154) — (4.160)) будем обозначать через  $\overline{(4.154) - (4.160)}$ , а соответствующую двойственную —  $\overline{(4.161)}$ . Нетрудно увидеть, что число авиалиний в задаче  $\overline{(4.154) - (4.160)}$  будет равно общему числу ограничений  $\overline{(4.157) - (4.159)}$ . Таким образом, в результате выполнения первого этапа для большинства авиалиний оптимальное решение будет найдено.

На втором этапе решается задача  $\overline{(4.161)}$  методом обобщенного градиентного спуска. Для получения решения задачи  $\overline{(4.154) - (4.160)}$ , т. е. для нахождения оптимального распределения самолетов на оставшихся авиалиниях, используется формула

$$x_{ij}^N = \frac{1}{N} \sum_{k=1}^N y_{ij}^*(Z^{(k)}, W^{(k)}, U^{(k)}),$$

где  $Z^{(k)}, W^{(k)}, U^{(k)}$  — значения двойственных переменных на  $k$ -й итерации,  $N$  — общее количество итераций обобщенного градиентного спуска.

Описанный алгоритм решения задачи (4.154) — (4.160) реализован программой на языке АЛГОЛ-60 для БЭСМ-6. Поскольку разработанная программа предназначена для решения целой серии задач типа (4.154) — (4.160), в ней

предусмотрены блоки подготовки исходной информации, блоки работы с устройствами внешних носителей памяти, блоки выходной информации и т. п.

В целях экономии оперативной памяти ЭВМ целесообразно хранение лишь параметров самолетов, авиалиний и аэропортов. Константы  $c_{ij}$ ,  $p_{ij}$  и  $v_{ij}$  достаточно просто вычисляются в процессе работы алгоритма (для вычисления одной константы необходимо выполнить около 10 арифметических операций). Вследствие учета специфики задачи описанный алгоритм позволяет получить ее решение, используя лишь оперативную память БЭСМ-6 (около 32 000 слов).

Расчеты, проведенные по программе, разработанной Н. Г. Журбенко и Т. В. Белых, показали высокую эффективность применения  $r$ -алгоритма. При решении двойственной задачи требуется около 80—100 итераций для получения относительной точности по функционалу  $10^{-6}$ . Для большей части авиалиний решение принималось на первом этапе алгоритма, на втором этапе рассматривалось всего 40—50 авиалиний. Общее время счета задачи 40—50 минут.

4. Стохастическая транспортная задача [3]. Математическая модель задачи состоит в следующем:

$$\min \left[ \sum_{i,j=1}^{m,n} c_{ij} x_{ij} + \sum_{j=1}^n r_j M \left( \xi_j - \sum_{i=1}^m x_{ij} \right)^+ \right], \quad (4.163)$$

$$\sum_{j=1}^n x_{ij} \leq a_i, \quad i = \overline{1, m}, \quad (4.164)$$

$$x_{ij} \geq 0, \quad i = \overline{1, m}, \quad j = \overline{1, n}, \quad (4.165)$$

где  $m$  — число поставщиков продукции (заводов);  $n$  — число потребителей продукции (строек);  $c_{ij}$  — стоимость перевозки единицы продукции от  $i$ -го завода к  $j$ -й стройке;  $a_i$  — объем продукции, производимый  $i$ -м заводом;  $\xi_j$  — случайная величина, определяющая спрос продукции для  $j$ -й стройки;  $r_j$  — коэффициент штрафа за недопоставку (дефицит) единицы продукции  $j$ -й стройке;  $x_{ij}$  — подлежащее определению количество продукции, получаемое  $j$ -й стройкой от  $i$ -го завода;  $t^+ = \max \{0, t\}$ ;  $M$  — знак математического ожидания.

Таким образом, задача состоит в определении такого распределения продукции, при котором достигается минимум затрат на транспортировку продукции плюс математическое ожидание потерь из-за недопоставок продукции потребителям.

Эта задача является частным случаем двухэтапной задачи стохастического программирования [75]. Случайные величины спроса  $\xi_j$  предполагаются независимыми с регулярной плотностью распределения  $p_j(z)$ .

**Лемма 4.3.** *Задача (4.163) — (4.165) является задачей выпуклого программирования.*

**Доказательство.** Для доказательства достаточно показать выпуклость функции

$$f(t) = M(\xi - t)^+,$$

где  $\xi$  — случайная величина, имеющая плотность распределения  $p(z)$ . Очевидно,

$$f(t) = \int_{-\infty}^{\infty} (z - t)^+ p(z) dz = \int_t^{+\infty} (z - t) p(z) dz,$$

откуда  $f'(t) = -\int_t^{\infty} p(z) dz$ ;  $f''(t) = p(t)$ . Поскольку  $p(t) \geq 0$ , то из последнего равенства следует выпуклость функции  $f(z)$ .

Решение задачи (4.163) — (4.165) будет основано на применении одной модификации метода минимизации с растяжением пространства ( $r$ -алгоритма) к задаче, двойственной (4.163) — (4.165).

Обозначим через  $u$  двойственные переменные, соответствующие ограничениям (4.164). Тогда двойственная задача будет иметь следующий вид:

$$\max_{u \geq 0} \psi(u), \quad (4.166)$$

где

$$\psi(u) = \min_{x \geq 0} L(x, u).$$

Здесь  $x$  — совокупность переменных  $\{x_{ij}\}$ ,  $L(x, u)$  — функция Лагранжа,

$$\begin{aligned} L(x, u) = & \sum_{i,j=1}^{n,m} c_{ij} x_{ij} + \sum_{j=1}^n r_j M\left(\xi_j - \sum_{i=1}^m x_{ij}\right)^+ + \\ & + \sum_{i=1}^m u_i \left(\sum_{j=1}^n x_{ij} - a_i\right). \end{aligned}$$

Нетрудно видеть, что

$$\psi(u) = \sum_{j=1}^n \varphi_j(u) - \sum_{i=1}^m a_i u_i,$$

где

$$\varphi_j(u) = \min_{x_{ij} \geq 0} \left[ \sum_{j=1}^n (c_{ij} + u_i) x_{ij} + r_j M \left( \xi_j - \sum_{i=1}^m x_{ij} \right)^+ \right]. \quad (4.167)$$

Пусть

$$v_j = \sum_{i=1}^m x_{ij},$$

через  $i(j)$  обозначим какое-нибудь значение индекса  $i \in \{1, m\}$ , при котором достигается минимум в выражении

$$\min_{i \in \{1, m\}} (c_{ij} + u_i). \quad (4.168)$$

Тогда нетрудно показать, что  $x_{ij}$ , определяемое операцией минимизации правой части формулы (4.167), задается следующим образом:

$$x_{ij}(u) = \begin{cases} v_j(u), & i = i(j), \\ 0, & i \neq i(j). \end{cases} \quad (4.169)$$

Поэтому

$$\varphi_j(u) = \min_{u_j \geq 0} [\tilde{c}_j v_j + r_j M (\xi_j - v_j)^+]; \quad (4.170)$$

где  $\tilde{c}_j = c_{i(j)j} + u_{i(j)}$ .

**Лемма 4.4.** Величины  $v_j(u)$  в формуле (4.169) определяются следующим образом: если  $\tilde{c}_j \geq r_j$ , то  $v_j = 0$ , в противном случае  $v_j$  является решением уравнения

$$r_j (1 - F_j(v_j)) = \tilde{c}_j, \quad (4.171)$$

где  $F_j(z)$  — функция распределения случайной величины  $\xi_j$ .

**Доказательство.** Для краткости индекс  $j$  в формуле (4.170) будем опускать. Вводя обозначение  $f(x)$ , имеем

$$f(v) = \tilde{c}v + rM(\xi - v)^+ = \tilde{c}v + r \int_v^{\infty} (r - v) p(z) dz,$$

$$f'(v) = \tilde{c} - r \int_v^{\infty} p(z) dz = \tilde{c} - r(1 - F(v)).$$

Из последнего равенства следует справедливость доказываемой леммы.

Лемма 4.4 и соотношения (4.170) и (4.171) полностью определяют алгоритм вычисления функции  $\psi(u)$ . Обобщенный градиент  $\psi'(u)$  функции  $\psi(u)$  определяется формулой  $\psi'(u) = g(u)$ , где

$$g(u) = (g_1(u), \dots, g_m(u)),$$

$$g_i(u) = \sum_{j=1}^n x_{ij}(u) - a_i, \quad i = \overline{1, m}.$$

Двойственная задача (4.166) может быть решена с помощью одного из методов оптимизации, не требующего дифференцируемости функции цели (из приведенного выше определения функции  $\psi(u)$  нетрудно понять, что  $\psi(u)$  является вогнутой кусочно-линейной функцией). Одним из таких методов, обладающих ускоренной сходимостью, является метод минимизации с растяжением пространства в направлении разности двух последовательных градиентов ( $r$ -алгоритм), опробованный при решении ряда практических задач.

Чтобы применить  $r$ -алгоритмы для решения двойственной задачи (4.166), необходимо предварительно свести ее к задаче безусловной оптимизации, что легко достигается заменой переменных  $u = |u'|$ .

Пусть с помощью  $r$ -алгоритма получена точка  $\tilde{u}$ -приближения точного решения задачи (4.166)  $u^*$ . Остановимся на вопросе определения решения  $x_{ij}^*$  исходной задачи (4.163) — (4.165). Из леммы 4.4 видно, что (в силу монотонности функции распределения  $F_j(z)$ ) величины  $v_j(u^*)$  однозначно определяются значением двойственных переменных  $u^*$ , благодаря этому можно показать [3], что

$$v_j(u^*) = \sum_{i=1}^m x_{ij}^*,$$

т. е.  $v_j(u^*)$  является оптимальной величиной поставки  $j$ -му потребителю. Поскольку мы предполагаем существование (регулярной) плотности распределения  $p_j(z)$ , то  $v_j(u)$ , как нетрудно видеть из леммы 4.4, будет непрерывной функцией  $u$ . Поэтому  $v_j(\tilde{u})$  будут оценкой для  $v_j(u^*)$ , используя которую нелинейную задачу (4.163) — (4.165) легко свести к классической транспортной задаче. Для решения последней можно применить хорошо известные алгоритмы решения или (что более естественно для предлагаемого «итеративного» подхода к решению задачи) метод усреднения в процессе

решения соответствующей двойственной задачи алгоритмом обобщенного градиентного спуска в окрестности точки  $u^*$ .

Описанная математическая модель возникла в результате решения одной практической задачи, связанной с поставкой стройматериалов на строительные объекты Украины. Размерность реальной задачи  $n \sim 100$ ,  $m = 30$ . Случайные величины спроса потребителей  $\xi_j$  имеют нормальное распределение, что позволяет решение уравнений (4.171) свести к решению уравнений типа  $\operatorname{erf}(z) = d_j$ , где  $\operatorname{erf}(z)$  — функция ошибок,

$$\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt, \quad 0 \leq d_j \leq 1.$$

Поэтому для решения уравнений (4.171) достаточно затабулировать функцию  $\operatorname{erf}^{-1}(y)$ , обратную  $\operatorname{erf}(z)$ .

### § 5. Использование $r$ -алгоритмов для решения нелинейных минимаксных задач

Рассмотрим вопросы минимизации функций вида  $f(x) = \max_{1 \leq i \leq m} \varphi_i(x)$ , где  $\varphi_i(x)$  — гладкие функции,  $x \in E_n$ .

Если  $\varphi_i(x)$  — выпуклые функции,  $i = 1, \dots, m$ , то в качестве обобщенного градиента  $g_f(x)$  можно брать субградиент  $g_{\varphi_{i^*}}(x)$ , где  $i^*$  — индекс функции  $\varphi_{i^*}$ , для которой  $f(x) = \varphi_{i^*}(x)$ .

В связи с важностью минимаксных задач большое внимание в настоящее время уделяется разработке методов их решения. Ряд таких методов для определенных классов задач предложен в работах В. Ф. Демьянова, Б. Н. Пшеничного, Е. Г. Евтушенко и др. (см. [14]).

В принципе задача отыскания минимакса может быть сведена к решению следующей задачи выпуклого программирования: найти  $\min y$  при ограничениях  $f_i(x) - y \leq 0$ ,  $i = 1, 2, \dots, m$ .

Для решения этой задачи могут быть использованы итеративные методы возможных направлений [26]. Однако такой путь решения задач связан с определенными вычислительными трудностями: во-первых, нахождение минимума при данных ограничениях методами возможных направлений связано с необходимостью решения на каждой итерации ряда вспомогательных задач линейного или квадра-

тичного программирования; во-вторых, для обеспечения сходимости методов возможных направлений нужно использовать меры предосторожности против зигзагообразного движения; в-третьих, на каждом шаге метода возможных направлений нужно определять минимум по направлению — довольно трудоемкая процедура, так как приходится определять минимум негладкой функции в условиях, когда определение значения функции требует значительной вычислительной работы, и, наконец, скорость сходимости этих методов слабо изучена как теоретически, так и экспериментально.

Более простым с вычислительной точки зрения является метод обобщенного градиентного спуска, который в принципе позволяет решать выпуклые минимаксные задачи [55]. Однако скорость сходимости этого метода недостаточно высока, особенно при решении задач овражного типа, которые характерны при отыскании минимаксов. В этом случае минимум, как правило, достигается при равенстве значений нескольких функций, причем их число в регулярном случае определяет размерность оврага, равную  $(n - m_1 + 1)$ , где  $m_1$  — число функций, на которых достигается минимаксное значение.

Ускорение сходимости метода обобщенного градиентного спуска может быть достигнуто за счет растяжения пространства в направлении обобщенного градиента или в направлении разности обобщенных градиентов, вычисленных в двух последовательных точках [67]. При использовании обобщенного градиентного спуска с растяжением пространства определение минимизирующей последовательности связано с вычислением обобщенного градиента функции  $f(x)$  в каждой точке последовательности и применением оператора растяжения пространства.

Обычно при анализе эффективности вычислительных алгоритмов предполагается, что вычисление градиента функции  $n$  переменных требует примерно столько же времени, сколько  $(n + 1)$ -кратное вычисление функции. Если считать, что каждая из функций  $\varphi_i(x)$  непрерывно дифференцируема, затраты на ее вычисление не зависят от индекса  $i$ , а сложность вычисления градиента в  $(n + 1)$  раз превосходит сложность вычисления функции, то сложность вычисления функции  $f(x)$  будет составлять  $\left(\frac{m}{1+n}\right)$ -ю часть сложности вычисления ее субградиента. Таким образом,

с ростом  $m$  уменьшается относительная сложность вычисления субградиента  $g_f(x)$  по сравнению со сложностью вычисления функции  $f(x)$ . Поэтому для обеспечения высокой эффективности разрабатываемых алгоритмов решения минимаксных задач нужно особенно стремиться к упрощению процедур нахождения минимума по направлению, связанных с вычислением значений функции максимума. Исходя из этого, Л. П. Шабашова и автор предложили и экспериментально исследовали несколько модификаций  $r$ -алгоритмов решения минимаксных задач, различающихся способами регулировки шагового множителя и точностью поиска минимума по направлению [69].

Рассмотрим одну из модификаций. Движение начинаем из произвольной точки  $x_0 \in E_n$ . При этом полагаем  $\tilde{g}_0 = 0 \in E_n$  и  $B_0 = I_n$ , где  $I_n$  — единичная матрица размерности  $n \times n$ . Предположим, что выполнено  $k$  итераций ( $k = 1, 2, \dots$ ), в результате чего определены точки  $x_1, x_2, \dots, x_k$ , а также вектор  $\tilde{g}_k$  и матрица  $B_k$ . Для определения точки  $x_{k+1}$  произведем следующие действия.

1. Определяем значение функции в  $k$ -й точке последовательности

$$f(x_k) = \max_{i \in \overline{1, m}} \varphi_i(x_k) = \varphi_{i^*}(x_k).$$

2. Вычисляем обобщенный градиент функции  $f(x)$  в точке  $x_k$

$$g_f(x_k) = g_{\varphi_{i^*}}(x_k).$$

3. Находим обобщенный градиент  $\bar{g}_k$  функции  $\psi_k(y)$  в точке  $y_k$  в преобразованном пространстве по формуле

$$\bar{g}_k = g_{\psi_k}(y_k) = B_k^* g_f(x_k),$$

где

$$\psi_k(y) = f(A_k^{-1}y) = f(B_k y);$$

$A_k$  — оператор преобразования пространства;  $B_k$  — оператор, обратный оператору  $A_k$ ;  $B_k^*$  — оператор, сопряженный оператору  $B_k$ .

4. Определяем разность обобщенных градиентов  $r_k = \bar{g}_k - \tilde{g}_k$ .

5. Находим отношение норм векторов  $r_k$  и  $\bar{g}_k$

$$\beta_k = \frac{\|r_k\|}{\|\bar{g}_k\|},$$

которое характеризует изменение направления обобщенного градиента в точке  $x_{k-1}$  относительно направления обобщенного градиента в точке  $x_k$  в растянутом пространстве.

6. Сравниваем  $\beta_k$  с заданной величиной  $q_1$ : а) если  $\beta_k \leq q_1$ , то определяем следующую точку минимизирующей последовательности

$$x_{k+1} = x_k - h_k B_k \frac{\tilde{g}_k}{\|\tilde{g}_k\|},$$

соответствующую перемещению в растянутом пространстве точки  $y$  в том же направлении и с той же величиной шага  $h_k$ , что и на  $k$ -й итерации, после чего переходим к выполнению  $(k+2)$ -й итерации;

б) если  $\beta_k > q_1$ , т. е. при значительном изменении направления обобщенного градиента, переходим к выполнению последующих этапов.

7. Определяем вектор  $\xi_{k+1}$ , в направлении которого будет производиться растяжение пространства,  $\xi_{k+1} = r_k / \|r_k\|$ .

8. Находим оператор  $B_{k+1}$ , обратный оператору  $A_{k+1}$ , по формуле  $B_{k+1} = B_k R \frac{1}{\alpha} (\xi_{k+1})$ .

9. Определяем обобщенный градиент функции  $\psi_{k+1}(y)$  в точке  $y_k$ , соответствующей точке  $x_k$ , с учетом  $(k+1)$ -го оператора растяжения пространства

$$\tilde{g}_{k+1} = B_{k+1}^* g_f(x_k);$$

в направлении антиградиента осуществляется движение на данной итерации.

10. Изменяем длину шага  $h_{k+1} = h_k q_2$ , где  $0 < q_2 \leq 1$ .

11. Определяем очередную точку минимизирующей последовательности

$$x_{k+1} = x_k - h_{k+1} B_{k+1} \frac{\tilde{g}_{k+1}}{\|\tilde{g}_{k+1}\|},$$

которая соответствует точке перемещения в растянутом

пространстве в направлении  $-\frac{\tilde{g}_{k+1}}{\|\tilde{g}_{k+1}\|}$  на величину  $h_{k+1}$  из

точки  $y_k = A_k x_k$ . После этого переходим к выполнению  $(k+2)$ -й итерации ( $k = 1, 2 \dots$ ).

Для исследования эффективности описанного алгоритма были составлены стандартные автокодовые программы для ЭВМ «Минск-22» и проведены многочисленные расчеты по решению минимаксных задач. Для численных экспериментов решались задачи следующего вида: найти

$$\min_{x \in E_n} f(x),$$

где  $f(x) = \max_{i \in \overline{1, m}} \varphi_i(x)$ , при  $\varphi_i(x) = a_i \sum_{j=1}^n (x^{(j)} - a_{ij})^2$ , где  $x = \{x^{(1)}, \dots, x^{(n)}\}$ ,  $\{a_i\}$  —  $m$ -мерный вектор;  $\{a_{ij}\}$  — матрица  $m \times n$ .

В качестве примера приведем результаты расчетов конкретной задачи:  $n = 5$ ,  $m = 10$ ;  $\{a_i\} = \{1; 5; 10; 2; 4; 3; 1,7; 2,5; 6; 3,5\}$ . Величины  $\{a_{ij}\}$  задаются матрицей (в транспонированном виде)

$$\begin{pmatrix} 0 & 2 & 1 & 1 & 3 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 2 & 4 & 2 & 2 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 2 & 2 \\ 0 & 1 & 1 & 2 & 0 & 0 & 1 & 2 & 1 & 0 \\ 0 & 3 & 2 & 2 & 1 & 1 & 1 & 1 & 0 & 0 \end{pmatrix}.$$

Задача решалась при следующих значениях параметров:  $\alpha = 3$ ;  $q_1 = 0,9$ ;  $q_2 = 0,95$ ;  $h_0 = 1$ . В качестве начального приближения была взята точка  $x_0 = (0, 0, 0, 0, 1)$ , при этом  $f(x_0) = 80$ . Значения минимизирующей последовательности и функции  $f(x)$  в этих точках на соответствующих итерациях приведены в табл. 1. Точное значение  $x^*$ , при котором достигается минимум  $f(x) = \max_{i \in \overline{1, 10}} \varphi_i(x)$ , равно  $x^{(1)*} = 1,12434$ ;  $x^{(2)*} = 0,97945$ ;  $x^{(3)*} = 1,47770$ ;  $x^{(3)*} = 0,92023$ ;  $x^{(4)*} = 1,12429$ . При этом

$$f(x^*) = \min_{x \in E_n} \max_{i \in \overline{1, 10}} \varphi_i(x) = 22,60016.$$

Как видно из табл. 1, за 51 итерацию получено приближение, у которого верны шесть знаков для  $f(x)$  и 3,5 знаков для  $x^{(j)}$ . Этот результат следует признать хорошим, если учесть, что функция имела двумерный овраг.

Изучение вычислительной схемы данного алгоритма и практика вычислений показали, что данный алгоритм

Т а б л и ц а 1

Номер итерации	$x^{(1)}$	$x^{(2)}$	$x^{(3)}$	$x^{(4)}$	$x^{(5)}$	$f(x)$
Нулевое приближение	0	0	0	0	1	80
1	0,1119585	0,223917	0,1119585	0,1119585	1,111958	63,0894
6	1,097278	0,788586	0,871869	0,2815889	0,8040123	34,399
11	1,069281	0,7126785	0,8504548	0,8685811	1,102541	25,83837
16	1,063638	0,9135513	1,160926	1,101065	1,073642	24,69413
21	1,076656	0,9905223	1,366388	0,910783	1,116449	22,78248
26	1,129564	0,980844	1,46487	0,936742	1,115325	22,6504
31	1,13501	0,9733332	1,469467	0,9394724	1,11657	22,609
36	1,11975	0,9814588	1,475895	0,9139938	1,126114	22,60392
41	1,128545	0,978718	1,482843	0,9253958	1,123494	22,60168
46	1,122855	0,9793679	1,474063	0,919058	1,124103	22,60064
51	1,124632	0,9793914	1,477879	0,9206665	1,124184	22,60023

вполне пригоден для решения минимаксных задач, с точки зрения численной реализации он является простым и обеспечивает вполне достаточную скорость сходимости. Как показали эксперименты, скорость сходимости и время решения задач зависят от выбора параметров  $\alpha$  и  $q_2$ . Коэффициент растяжения пространства  $\alpha$  в принципе может быть любым числом, большим единицы. Однако если  $\alpha$  слишком близко к единице, то эффект растяжения слабо сказывается, что приводит к медленной сходимости, особенно для функций с многомерным оврагом. Наоборот, при больших  $\alpha$  происходит быстрое убывание нормы матрицы  $B$  и, как следствие, быстрое уменьшение длины шага в основном пространстве, что приводит к увеличению числа шагов без растяжения пространства, это также уменьшает скорость сходимости. Наиболее целесообразно значение  $\alpha$  выбирать в пределах от 2 до 4. Параметр  $q_2$  в принципе можно брать равным единице. Если мы выбираем  $q_2 < 1$ , то это приводит к уменьшению длины шага равномерно по всем направлениям (в то время как матрица  $B$  неравномерно уменьшает шаг по различным направлениям). Если  $q_2$  слишком отличен от единицы, то увеличивается число «холостых» шагов (без растяжения пространства), что замедляет сходимость.

С другой стороны, для многомерных задач во многих случаях целесообразно выбирать  $q_2 < 1$ , так как «средний» показатель уменьшения шага за счет уменьшения нормы матрицы  $B$  равен  $1/\sqrt[n]{\alpha}$  и при больших  $n$  может оказаться слишком близким к единице.

Описанный численный пример решался при различных  $q_2$  до получения минимального значения  $f(x) = 22,60016$  при постоянных параметрах  $\alpha = 3$ ;  $q_1 = 0,9$ . При  $q_2 = 1$  было затрачено 69 итераций, из них 52 с растяжением пространства,  $\frac{52}{69} \approx 0,75$ , при  $q_2 = 0,95$  — 57 итераций, из них 41 с растяжением пространства,  $\frac{41}{57} \approx 0,72$ , при  $q_2 = 0,9$  — 112 итераций, из них 39 с растяжением пространства,  $\frac{39}{112} \approx 0,35$ . Дальнейшее уменьшение  $q_2$  явно нецелесообразно.

Эксперименты показывают, что, по-видимому, для каждой конкретной задачи существует асимптотически «оптимальное»  $q_2$ . Можно строить алгоритмы, которые автоматически настраиваются на это значение  $q_2$ . Для этого можно анализировать отношение числа итераций с растяжением пространства к общему числу итераций, и если это отношение становится меньшим определенного числа, то необходимо увеличивать  $q_2$ , и наоборот, если оно слишком близко к единице, то уменьшать  $q_2$ . Аналогично можно регулировать и параметр  $\alpha$ . Параметр  $q_1$  рекомендуем выбирать в пределах  $0,9$ — $1$ .

Применение предложенного алгоритма целесообразно при решении задач, где требуется высокая точность. При этом в оперативной памяти ЭВМ необходимо хранить матрицу  $B_k$  размерности  $n \times n$ . Поэтому в силу ограниченности объема памяти алгоритм может применяться для нахождения минимакса функций от сравнительно небольшого числа переменных (например, оперативная память БЭСМ-6 позволяет решать задачи размерности до 150). При решении задач высокой размерности возникают трудности запоминания матрицы  $B_k$ . В таких задачах обычно минимакс достигается при равенстве значений нескольких функций, число которых  $m_1$  намного меньше  $n$  — размерности задачи. В связи с этим предлагается второй алгоритм, позволяющий избежать запоминания матрицы и сократить время счета.

Дадим описание второго алгоритма. Как и в первом алгоритме, движение начинаем из произвольной точки  $x_0 \in E_n$ , полагая  $\tilde{g}_0 = (0, 0, \dots, 0) \in E_n$ . Предположим, что после  $k$  итераций ( $k = 1, 2, \dots$ ) определены точки  $x_1, x_2, \dots, x_k$ , вектор  $\tilde{g}_k$  и последовательность векторов  $\xi_1, \xi_2, \dots, \xi_k$ , где  $\xi_v \in E_n$  ( $v = 1, 2, \dots, k$ ) — последовательность векторов единичной длины, в направлении которых производится на каждой итерации растяжение пространства  $E_n$ .

Для определения точки  $x_{k+1}$  необходимо вычислить:

$$1) f(x_k) = \max_i \varphi_i(x_k) \text{ — максимальное значение функций}$$

в точке  $x_k$ ;

$$2) g_i(x_k) = g_{\varphi_{i^*}}(x_k) \text{ — обобщенный градиент функции } f(x) \text{ в точке } x_k (f(x_k) = \varphi_{i^*}(x_k));$$

$$3) \tilde{g}_k = R_\beta(\xi_k) R_\beta(\xi_{k-1}) \dots R_\beta(\xi_1) g_i(x_k) \text{ — обобщенный градиент функции в преобразованном пространстве.}$$

В данном алгоритме вместо матрицы  $B_k$  запоминается последовательность нормированных векторов  $\xi_1, \dots, \xi_k$ . Обобщенный градиент  $\tilde{g}_k$ , вычисленный в точке  $y_k = A_k x_k$ , получается вследствие последовательного применения операторов растяжения пространства  $R_\beta(\xi_1)$  сначала к вектору  $g_i(x_k)$ , затем оператор растяжения  $R_\beta(\xi_2)$  действует на вектор  $R_\beta(\xi_1) g_i(x_k)$  и, наконец, оператор растяжения пространства с коэффициентом растяжения  $\beta = 1/\alpha$  и в направлении вектора  $\xi_k = R_\beta(\xi_k)$  действует на вектор, который получился в результате последовательного действия операторов растяжения в направлении  $\xi_1, \dots, \xi_{k-1}$ .

Выражение для определения  $\tilde{g}_k$  получается из обобщенного градиента при растяжении пространства  $E_n$  путем последовательного применения операторов растяжения в направлении векторов  $\xi_1, \xi_2, \dots, \xi_k$ :

$$y_k = R_\alpha(\xi_k) R_\alpha(\xi_{k-1}) \dots R_\alpha(\xi_1) x_k$$

и из условия  $\tilde{g}_k = g_{\psi_k}(y_k)$ , где  $\psi_k(y_k) = f(A_k^{-1} y_k)$ ,

$$\psi_k(y_k) = f[R_\beta(\xi_1) R_\beta(\xi_2) \dots R_\beta(\xi_k) y_k].$$

Далее, вычисляем:

$$4) r_k = \tilde{g}_k - \tilde{g}_k \text{ — разность обобщенных градиентов;}$$

$$5) \beta_k = \frac{\|r_k\|}{\|g_k\|};$$

б) сравниваем  $\beta_k$  с заданной величиной  $q_1$ :

а) если  $\beta_k \leq q_1$ , то

$$x_{k+1} = x_k - h_k R_\beta(\xi_1) \dots R_\beta(\xi_k) \frac{\bar{g}_k}{\|\bar{g}_k\|},$$

т. е. движение производится в том же направлении и с той же длиной шага, что на  $k$ -й итерации, после чего переходим к выполнению  $(k+2)$ -й итерации;

б) если  $\beta_k > q_1$ , то переходим к выполнению следующего действия;

7) вычисляем  $\xi_{k+1} = \frac{r_k}{\|r_k\|}$ , вектор  $\xi_{k+1}$  запоминаем;

8) вычисляем  $\bar{g}_{k+1} = R_\beta(\xi_{k+1}) R_\beta(\xi_k) \dots R_\beta(\xi_1) g_1(x_k)$ .

Учитывая выражение для получения обобщенного градиента  $\bar{g}_k$ , получаем окончательно

$$\bar{g}_{k+1} = R_\beta(\xi_{k+1}) \bar{g}_k,$$

где  $\bar{q}_{k+1}$  — обобщенный градиент функции  $\varphi_{k+1}(y)$  в точке  $y_k$ , соответствующей точке  $x_k$  с учетом на  $(k+1)$ -й итерации растяжения пространства с помощью оператора растяжения в направлении вектора  $\xi_{k+1}$  с коэффициентом растяжения  $\alpha$ ;

9) вычисляем  $h_{k+1} = h_k q_2$ , где  $0 < q_2 \leq 1$ ;

10) определяем очередную точку  $x_{k+1}$  итерационной последовательности, которая получается в результате перемещения в растянутом пространстве в направлении  $-\bar{g}_{k+1}$  из точки  $y_k = R_\alpha(\xi_k) R_\alpha(\xi_{k+1}) \dots R_\alpha(\xi_1) x_k$  в точку  $y_{k+1} = y_k - h_{k+1} \frac{\bar{g}_{k+1}}{\|\bar{g}_{k+1}\|}$ . Как указывалось выше,  $y_{k+1} = R_\alpha(\xi_{k+1}) \dots R_\alpha(\xi_1) x_{k+1}$ , откуда  $x_{k+1} = R_\beta(\xi_1) R_\beta(\xi_2) \dots R_\beta(\xi_{k+1}) y_{k+1}$ . Из этих соотношений выводится выражение для получения точки  $x_{k+1}$ :

$$x_{k+1} = x_k - h_k R_\beta(\xi_1) R_\beta(\xi_2) \dots R_\beta(\xi_{k+1}) \frac{\bar{g}_{k+1}}{\|\bar{g}_{k+1}\|}.$$

Переходим к выполнению действия на  $(k+2)$ -й итерации. Так как память для запоминания векторов  $\xi_1, \dots, \xi_k$  ограничена, приходится описанный выше процесс обрывать после определенного числа шагов  $r$  ( $r < n$ ). После

этого матрица  $B$  восстанавливается, значение  $x_r$  принимается за начальное для нового внутреннего цикла, при этом начальный шаг берется в 2—3 раза меньшим, чем в начале предыдущего цикла. Процесс продолжается до тех пор, пока не будут достигнуты приемлемые значения минимизируемой функции или пока приращение функции, полученное в результате проведения очередного цикла вычислений, не станет достаточно малым. Таким образом, второй алгоритм является промежуточным между алгоритмом обобщенного градиентного спуска и алгоритмом с растяжением пространства. Этот алгоритм применялся для решения ряда задач, связанных с проблемой построения оптимальных эталонов. Постановка задачи, исходные данные, подпрограммы вычислений функций и градиентов были любезно предоставлены нам В. А. Ковалевским и М. И. Шлезингером. Решалась сложная максиминная задача, число переменных составляло 1200, число функций, участвующих в операции взятия минимума, — 1536.

Эта задача для сравнения эффективности алгоритмов решалась также обычным методом обобщенного градиентного спуска без растяжения пространства с использованием трех различных методик регулировки шага:

- 1)  $h_{k+1} = h_k \cdot 0,95, \quad x_{k+1} = x_k + h_{k+1} \frac{g_f(x_k)}{\|g_f(x_k)\|};$
- 2)  $h_{k+1} = h_k \cdot 0,95, \quad x_{k+1} = x_k + h_{k+1} g_f(x_k);$
- 3)  $h_{k+1} = \frac{f(x^*) - f(x_k)}{\|g_f(x_k)\|}, \quad x_{k+1} = x_k + h_{k+1} \frac{g_f(x_k)}{\|g_f(x_k)\|},$

где  $f(x^*)$  — предполагаемое значение в точке максимума. Наиболее хорошие результаты решения задачи получены первым способом регулировки шага (табл. 2).

Для решения задачи с помощью описанного здесь второго алгоритма, применение которого обусловлено значительной размерностью задачи ( $n = 1200$ ), была составлена программа на языке ФОРТРАН. Решение проводилось на БЭСМ-6, при этом полагалось  $\alpha = 3$ .

Анализ результатов решения задачи, полученных обоими методами и приведенных в табл. 2, показывает, что для достижения примерно одних и тех же значений оптимизируемой функции вблизи точки максимума при использовании второго алгоритма требуются значительно меньшие число итераций и количество машинного времени, чем при использовании метода обобщенного градиентного спуска без растяжения пространства. Так, например, приемлемое

Таблица 2

Номер итерации	Значения оптимизируемой функции		Номер итерации	Значения оптимизируемой функции	
	Обобщенный градиентный метод (с помощью первой регуляровки шага)	Второй алгоритм		Обобщенный градиентный метод (с помощью первой регуляровки шага)	Второй алгоритм
1	-30,92	-30,92	40	3,350	5,010
5	-7,009	-11,26	45	3,403	5,197
10	-4,548	-8,703	50	3,826	5,277
15	-0,3739	1,143	55	3,779	5,312
20	1,456	2,785	60	4,025	5,333
25	1,911	3,965	65	3,732	5,347
30	1,911	4,551	71	3,882	
35	2,315	4,870			

значение оптимизируемой функции, равное 3,88, было получено обобщенным градиентным методом без растяжения пространства на 71-й итерации, при использовании же второй модификации алгоритма уже на 25-й итерации было получено значение  $f(x)$ , равное 3,96. Кроме того, применение второго алгоритма позволило получить значение максимизируемой функции (5,07 на 39-й итерации), с хорошей точностью совпадающее с теоретическим максимумом. При использовании обобщенного градиентного спуска такой точности достичь не удалось (для этого требовалось слишком большое число итераций и продолжительное время счета).

### § 6. Применение методов минимизации негладких функций для решения задач интерпретации гравиметрических наблюдений

Этот параграф написан на основе исследований, проведенных автором совместно с И. Г. Овруцким [37].

Широкое внедрение в практику геофизических исследований электронно-вычислительных машин позволяет применять для геологического истолкования гравитационных аномалий новые методы. При разработке новых методов интерпретации всегда следует исходить из того, что обратная задача гравитационной разведки относится к классу некор-

ректных задач математической физики и для ее решения необходимо разработать специальные математические модели и методы [10, 47].

Рассмотрим методику решения обратной задачи гравитационной разведки путем сведения ее к нелинейной минимаксной задаче с ограничениями, для решения которой используется обобщенный градиентный спуск с растяжением пространства в форме  $r$ -алгоритма.

Пусть  $x = \{t_1, \dots, t_n\}$  — вектор из пространства  $E_n$ , характеризующий параметры модели геологической среды;  $F$  — оператор решения прямой задачи гравитационной разведки, который при заданном  $x \in E_n$  дает теоретические значения аномалии силы тяжести в пунктах наблюдения;  $\Delta g_i$  — наблюдаемое значение аномалии силы тяжести в пункте  $i$ . Решение обратной задачи гравитационной разведки сводится к нахождению вектора  $x \in E_n$ , соответствующего минимуму целевой функции при заданных ограничениях, причем целевая функция характеризует отклонения теоретических значений аномалий от измеренных.

В общей постановке решение обратной задачи даже при выбранной целевой функции неоднозначно. Всегда по наблюдаемой аномалии можно найти ряд эквивалентных распределений «возмущающих масс». Тем не менее гравитационная разведка успешно используется для решения геологических задач. Это объясняется тем, что геофизик-интерпретатор использует дополнительную геологическую информацию о форме и количестве аномальных тел, о плотностной характеристике геологического разреза, данные, полученные другими геофизическими методами, вследствие чего поиск решения ведется в рамках конечномерной параметрической модели с введением априорных ограничений на значения параметров. В геофизике в качестве критерия интерпретации, как правило, используются целевые функции двух видов: оценка параметров по методу наименьших квадратов

$$Q_1 = \frac{1}{m} \sum_{i=1}^m c_i (F_i(x) - \Delta g_i)^2 \quad (4.172)$$

и оценка параметров, оптимальных в чебышевском смысле (по максимуму уклонения),

$$Q_2 = \max_i |c_i [F_i(x) - \Delta g_i]|, \quad (4.173)$$

где  $c_i$  — «весовые» коэффициенты. Выбор функции цели диктуется конкретной геологической ситуацией, а также сведениями об ошибках измерения и моделирования.

В связи с тем что измерения аномалий силы тяжести всегда содержат инструментальную случайную ошибку, а модель геологической среды является упрощенной, расхождение между наблюдаемой аномалией и теоретической, полученной в результате действия оператора  $F$  на вектор параметров  $x$ , представляет собой некоторую комбинацию инструментальной и модельной ошибок. В этих условиях оценка параметров модели по методу наименьших квадратов не имеет никаких видимых преимуществ по сравнению с оценкой по чебышевскому критерию (в большинстве реальных случаев ошибка, связанная с ограниченностью числа параметров модели, намного превосходит ошибку, вызванную неточностью измерений).

В данной работе рассмотрены алгоритмы решения обратной задачи по чебышевскому критерию. Аналогичные алгоритмы можно использовать и для метода наименьших квадратов, однако, как увидим ниже, интерпретация по чебышевскому критерию требует, так правило, меньше вычислений. Таким образом, необходимо так подобрать параметры модели  $x$ , чтобы максимальное отклонение теоретической аномалии от наблюдаемой было минимальным. Практически достаточно, чтобы величина отклонения не превышала по модулю числа  $\varepsilon$ , которое выбирает интерпретатор, исходя из точности экспериментальных данных и предполагаемых ошибок моделирования.

Обратную задачу гравитационной разведки можно рассматривать как минимаксную и математически сформулировать так: минимизировать

$$Q = \max_i |F_i(x) - \Delta g_i| \quad (4.174)$$

при ограничениях

$$f_k(x) \leq 0, \quad k = 1, \dots, p. \quad (4.175)$$

Наиболее типичные для рассматриваемых задач двусторонние ограничения типа  $x_a \leq x \leq x_b$ , где  $x_a$  и  $x_b$  — соответственно левый и правый пределы изменения параметров.

Решение задачи (4.174), (4.175) можно свести к безусловной минимаксной задаче следующего вида: минимизировать

$$Q_\lambda(x) = \max_i |F_i(x) - \Delta g_i| + \lambda \sum_{k=1}^p f_k(x) \delta_k(x), \quad (4.176)$$

где второй член в правой части (4.176) — функция штрафа,

$$\delta_k(x) = \begin{cases} 0, & \text{если } f_k(x) \leq 0, \\ 1, & \text{если } f_k(x) \geq 0; \end{cases}$$

$\lambda$  — достаточно большой штрафной множитель [18]. Функция  $Q_\lambda(x)$  негладкая. Поэтому для минимизации выражения (4.176) нужно применять специально разработанные для минимизации негладких функций обобщенные градиентные методы.

Вид оператора  $F$  целиком зависит от способа моделирования реальной геологической среды. Если модель геологической среды аппроксимировать трехмерными прямоугольными призмами, то из закона всемирного тяготения следует известная формула [48]

$$F_i(x) = -f \sum_{j=1}^m \sigma_j \left| \left| \left( \eta_j - x_i^{(2)} \right) \ln \left[ \left( \xi_j - x_i^{(1)} \right) + R_{ji} \right] + \right. \right. \\ \left. \left. + \left( \xi_j - x_i^{(1)} \right) \ln \left[ \left( -x_i^{(2)} + \eta_j \right) + R_{ji} \right] + \right. \right. \\ \left. \left. + \left( x_i^{(3)} - \zeta_j \right) \arctg \frac{\left( \xi_j - x_i^{(1)} \right) \left( \eta_j - x_i^{(2)} \right) \xi_{j2} \eta_{j2} \zeta_{j2}}{\left( \xi_j - x_i^{(3)} \right) R_{ji} \xi_{j1} \eta_{j1} \zeta_{j1}} \right|, \quad (4.177)$$

где  $f$  — гравитационная постоянная;  $x_i^{(1)}$ ,  $x_i^{(2)}$ ,  $x_i^{(3)}$  — координаты точки  $i$ ;  $\xi_{j1}$ ,  $\xi_{j2}$ ,  $\eta_{j1}$ ,  $\eta_{j2}$ ,  $\zeta_{j1}$ ,  $\zeta_{j2}$  — координаты  $j$ -й призмы по осям;  $\sigma_j$  — избыточная плотность;

$$R_{ji} = \sqrt{\left( \xi_j - x_i^{(1)} \right)^2 + \left( \eta_j - x_i^{(2)} \right)^2 + \left( \zeta_j - x_i^{(3)} \right)^2}.$$

Минимизация функции (4.176) связана с большими трудностями, так как по своей физической природе  $Q_\lambda(x)$  является негладкой овражной функцией. Это объясняется тем, что на величину  $\Delta g_i$  оказывают влияние разнородные параметры. Так, верхние кромки тел  $\zeta_{j1}$  сильнее сказываются в наблюдаемом гравитационном поле по сравнению с нижними  $\zeta_{j2}$ , поэтому в многомерном пространстве поверхности уровней функции вытянуты, а это затрудняет минимизацию.

Для решения такого класса задач предлагается использовать модифицированный обобщенный градиентный спуск с растяжением пространства в направлении разности двух последовательных градиентов. Вычисление градиента (или

его аналога) от функции максимума  $\varphi(x) = \max f_i(x)$  в точке  $x_0$  требует вычисления  $\max f_i(x)$  и определения градиента от функции  $f_i(x^*)$ , на которой достигается значение максимума при  $x = x_0$ .

Основная доля вычислений в алгоритмах градиентного типа, как правило, связана с вычислением градиента. Для облегчения программирования при сложном функционале градиент вычисляется конечноразностным методом. Сравним время вычисления градиента для функционалов типа (4.172) и (4.173). За единицу времени вычисления примем время расчета теоретического значения  $F_i(x)$  в одной точке наблюдения от одного тела в форме прямоугольной призмы. Пусть модель аппроксимируется  $n$  телами, измерения выполнены в  $m$  пунктах, вектор неизвестных параметров  $x$  имеет размерность  $t$ . Тогда затраты машинного времени составят  $mn + t$  для приближенного вычисления градиента функционала (4.173) конечноразностным методом и  $m(t +$

$\frac{t}{n} + 1$  раз больше, для аналогичного вычисления градиента функционала (4.172). Например, при  $n = 1, m = 20, t = 5$  на вычисление градиента от функционала (4.172) уходит примерно в пять раз больше времени, чем на вычисление градиента от функционала (4.173). В реальных задачах число параметров  $t$  может достигать нескольких десятков. Поэтому если градиентный метод, использованный для минимизации функционала (4.173), не уступает существенно в скорости сходимости методу, примененному для минимизации функционала (4.172), то минимизация (4.173) с определенной точностью потребует гораздо меньше времени, чем минимизация (4.172) с сопоставимой точностью. Этим можно объяснить наш интерес к построению алгоритмов интерпретации аномалий силы тяжести по чебышевскому критерию.

Для решения обратной задачи гравитационной разведки по предложенному алгоритму составлена программа на языке ФОРТРАН-IV для ЭВМ «Минск-32». Программа состоит из управляющего блока, блока вычисления функции цели  $Q_\lambda(x)$  и разностного аналога вектора обобщенного градиента и блока минимизации. Управляющий блок осуществляет ввод исходных данных, формирование констант счета, пе-

чать промежуточных и конечных результатов, организует выход из программы. В блоке вычисления функции цели  $Q_\lambda(x)$  запрограммировано по формулам (4.176), (4.177) решение прямой задачи от модели геологической среды, аппроксимируемой набором трехмерных прямоугольных призм, а также вычисление аналога вектора обобщенного градиента конечноразностным методом. В блоке минимизации реализован алгоритм обобщенного градиентного спуска с растяжением пространства в направлении разности двух антиградиентов ( $r$ -алгоритм), описанный в § 5 гл. 4. Программа опробовалась на модельных и практических примерах [37].

Опробование предложенной методики на моделях позволило оценить круг решаемых задач и точность решения. Данная методика интерпретации гравитационных аномалий может успешно применяться для определения морфологии интрузивных тел, если известна плотностная характеристика модели, для определения положения контактных гравитирующих поверхностей и вертикальных контактов тел, для оценки нижней границы распространения возмущающих масс и т. п.

## **§ 7. Другие области применения обобщенных градиентных методов**

В предыдущих параграфах речь шла о применении обобщенных градиентных методов при реализации схем декомпозиции по множеству ограничений или переменных, при некоторых типах параметрической декомпозиции, при решении отраслевых задач текущего планирования производственно-транспортного типа, при решении минимаксных задач, задач обработки результатов наблюдений. Этими примерами далеко не исчерпываются возможности применения обобщенных градиентных методов. В данном параграфе кратко изложим другие возможные применения обобщенных градиентных методов.

### **Решения задач оптимального проектирования**

Математические модели, возникающие при решении вопросов оптимального проектирования сложных динамических объектов, машин, зданий и сооружений, имеют форму задачи нелинейного программирования с ограничениями в виде равенств и неравенств, с переменными дискретного и непрерывного типа. Задача чаще всего оказывается многоэкстремальной. Обычно для решения задач оптимального про-

ектирования обобщенные градиентные методы применяются для нахождения локальных оптимумов по непрерывным переменным в сочетании с комбинаторными схемами для вариации дискретных переменных и комбинаторными или стохастическими способами поиска начального приближения. Для учета ограничений используются либо схемы исключения переменных для ограничений, заданных в форме равенств, либо метод штрафных функций с негладкой функцией штрафа [18, 80, 88] для ограничений в форме равенств и неравенств. Конкретные приложения методов минимизации негладких функций к задачам прсектирования описаны в работе [6].

Рассмотрим несколько подробнее метод негладких штрафных функций. Следуя [80], для задачи выпуклого программирования

$$\min f_0(x) \text{ при } f_i(x) \leq 0, \quad i = 1, \dots, m, \quad (4.178)$$

определим функцию

$$S(x) = f_0(x) + \sum_{i=1}^m p_i[f_i(x)], \quad (4.179)$$

где  $p_i(t)$ ,  $i = 1, \dots, m$ , удовлетворяют условиям выпуклости,  $p_i(t) = 0$  для  $t < 0$ ,  $p_i(t) \geq 0$  для  $t > 0$ . Допустим, что существует  $x^*$  — оптимальное решение задачи  $\inf_x S(x)$ .

**Теорема 4.2.** Для того чтобы  $x^*$  было оптимальным решением исходной задачи, необходимо, чтобы

$$\lim_{t \rightarrow +0} \frac{p_i(t)}{t} \geq \bar{y}_i, \quad i = 1, \dots, m,$$

где  $\bar{y} = (\bar{y}_1, \dots, \bar{y}_m)$  — некоторые множители Лагранжа задачи (4.178).

Для того чтобы (4.178) и (4.179) имели одинаковое множество оптимальных решений, достаточно, чтобы

$$\lim_{t \rightarrow +0} \frac{p_i(t)}{t} > \bar{y}_i.$$

Следует отметить, что первым серьезно исследовал негладкие функции штрафа И. И. Еремин [18]. Простейший вариант негладкой функции штрафа имеет вид

$$p(t) = \begin{cases} 0, & t \leq 0, \\ ct, & t > 0. \end{cases} \quad (4.180)$$

Из теоремы 4.2 вытекает, что для того чтобы задача минимизации  $S(x)$  была эквивалентна задаче (4.178) при

$$p_i(t) = \begin{cases} 0, & t \leq 0, \\ c_i t, & t > 0, \end{cases} \quad i = 1, \dots, m, \quad (4.181)$$

достаточно брать  $c_i > \bar{y}_i$ .

Большой интерес представляет штрафная функция в форме

$$T(x) = f_0(x) + p[\max_{i \in \overline{1, m}} f_i(x)],$$

где  $p(t)$  вычисляется по формуле (4.180). Легко видеть, что если  $c > \sum_{i=1}^m \bar{y}_i$ , то

$$p(\max_{i \in \overline{1, m}} f_i(x)) \geq \sum_{i=1}^m p_i[f_i(x)],$$

где  $p_i(t)$  вычисляется по формулам (4.181),  $c = \sum_{i=1}^m c_i$ ;  $c_i > \bar{y}_i$ . Таким образом, при  $c > \sum_{i=1}^m \bar{y}_i$  задача минимизации  $T(x)$  эквивалентна задаче (4.178).

Метод негладких штрафных функций можно с успехом использовать для получения локальных экстремумов в задачах нелинейного программирования общего вида: найти

$$\min f_0(x) \quad (4.182)$$

при ограничениях вида

$$\varphi_i(x) \leq 0, \quad i = 1, \dots, m, \quad \psi_j(x) = 0, \quad j = 1, \dots, l, \quad (4.183)$$

путем замены ее задачей минимизации функции

$$S(x) = f_0(x) + \sum_{i=1}^m c_i \varphi_i^+(x) + \sum_{j=1}^l d_j |\psi_j(x)|. \quad (4.184)$$

При достаточно больших  $c_i$  и  $d_j$  в условиях существования множителей Лагранжа в задаче (4.182) — (4.183) задача минимизации  $S(x)$  эквивалентна задаче (4.182) — (4.183).

Использование негладких штрафных функций обычно приводит к задачам минимизации функций овражного вида.

Для их решения рекомендуется использовать алгоритмы с растяжением пространства в направлении разности двух последовательных градиентов.

**Задачи дискретного программирования  
и смешанного дискретно-непрерывного типа**

Для решения задач дискретного и дискретно-непрерывного типа довольно часто используется метод «ветвей и границ». Эффективность решения задачи во многом зависит от трудоемкости получения оценок оптимума по функционалу и их точности. В последние годы большое внимание уделяется двойственному подходу к получению оценок. Суть его состоит в следующем [90].

Пусть имеется задача нелинейного программирования

$$\min_{x \in X \subset E_n} f_0(x) \quad (4.185)$$

при ограничениях

$$f_i(x) \leq 0, \quad i = 1, \dots, m; \quad (4.186)$$

$X$  — компактное множество, в частности дискретное. Рассмотрим функцию Лагранжа

$$L(x, u) = f_0(x) + \sum_{i=1}^m u_i f_i(x).$$

Пусть

$$\Phi(u) = \min_{x \in X} \left[ f_0(x) + \sum_{i=1}^m u_i f_i(x) \right], \quad (4.187)$$

$\Phi(u)$  — вогнутая функция в области, где она существует. В частности, если  $X$  — конечное дискретное множество, то  $\Phi(u)$  — кусочно-линейная вогнутая функция. Рассмотрим  $Q = \max_{u \geq 0} \Phi(u)$ .

**Теорема 4.3.** Если  $f^*$  — оптимальное значение задачи (4.185) — (4.186), то  $Q \leq f^*$ .

**Доказательство.** Пусть  $x^*$  — оптимальное решение задачи (4.185) — (4.186). Тогда  $f_i(x^*) \leq 0$ ,  $i = 1, \dots, m$ ,  $f_0(x^*) = f^*$ . При любом  $u \geq 0$

$$\Phi(u) \leq L(x^*, u) \leq f^*.$$

Отсюда  $Q = \max_{u \geq 0} \Phi(u) \leq f^*$ , что и требовалось доказать. Таким образом,  $Q$  может служить оценкой снизу для  $f^*$ .

Задача нахождения  $\max_{u \geq 0} \Phi(u)$  может решаться методом обобщенного градиентного спуска с проектированием на область  $u \geq 0$  в случае большого  $m$  или с помощью  $r$ -алгоритма, если число  $m$  не превышает нескольких сотен. Сложность вспомогательной задачи (4.187) может быть значительной, однако во многих интересных случаях она решается сравнительно просто. Так, например, рассмотрим, следуя [90], задачу линейного целочисленного программирования

$$\min(c, \omega), \quad (4.188)$$

$$A\omega = d, \quad \omega \geq 0 \text{ и целое,}$$

где  $A$  — целочисленная  $m \times (n + m)$ -матрица ранга  $m$ ;  $d$  —  $m$ -мерный целочисленный вектор;  $c$  —  $(n + m)$ -мерный вектор. Пусть  $B$  — неособая матрица размера  $m \times m$ , состоящая из столбцов матрицы  $A$ ;  $N$  — матрица оставшихся столбцов. Если выделить в  $\omega$  компоненту  $y$ , соответствующую  $B$ , и  $x$ , соответствующую  $N$ , и аналогично в  $c$  выделить компоненты  $c_B$  и  $c_N$ , то задача (4.188) сведется к виду

$$\min [c_B B^{-1} d + (c_N - c_B B^{-1} N) x]$$

при ограничениях

$$B^{-1} N x \leq B^{-1} d; \quad B^{-1} N x \equiv B^{-1} d \pmod{1}; \quad x \geq 0 \text{ и целое.}$$

Введем обозначения

$$X = \{x : B^{-1} N x \equiv B^{-1} d \pmod{1}; \quad x \geq 0 \text{ и целое}\};$$

$$f(x) = (c_N - c_B B^{-1} N) x;$$

$$g(x) = B^{-1} N x; \quad b = B^{-1} d.$$

Тогда двойственная задача имеет вид

$$\min_{x \in X} [f(x) + (u, g(x) - b)]$$

и сводится к нахождению кратчайшего пути в специальной сети, включающей  $|\det B|$  вершин. Для решения этой задачи разработан эффективный алгоритм [93].

. Хелд и Карп [94] успешно применили двойственный подход в сочетании с методом ОГС к задаче коммивояжера, Хелд, Вулф, Кроудер [95] — к многопродуктовым целочисленным потоковым задачам, Фишер [89] — к построению оптимальных расписаний.

Как видно из предыдущего изложения, наиболее простым и общим среди методов негладкой оптимизации является метод обобщенного градиентного спуска. Недостатком его является медленная и не всегда хорошо контролируемая скорость сходимости. Из алгоритмов с растяжением пространства наиболее хорошо экспериментально исследованы  $r$ -алгоритмы. Если размерность пространства, в котором происходит минимизация, не превышает 200—300, целесообразность использования  $r$ -алгоритмов вместо ОГС в большинстве случаев не вызывает сомнений. При больших размерностях значительная часть вычислительной работы идет на преобразование матрицы растяжения пространства; трудоемкость одной итерации в  $r$ -алгоритме может превысить в несколько раз трудоемкость одной итерации в методе ОГС и целесообразность применения  $r$ -алгоритма тем больше, чем меньше относительная доля времени, затраченного на преобразование матрицы, и чем большая точность требуется от результата. В каждом конкретном случае нужно оценить время решения. При этом можно пользоваться следующим эмпирическим правилом для  $r$ -алгоритмов: относительная точность решения по функционалу увеличивается на порядок через каждые  $n - 1,5n$  итераций ( $n$  — размерность минимизируемой функции). Что касается метода ОГСРП, то его, по-видимому, стоит применять лишь когда известно значение функции в точке минимума, в противном случае  $r$ -алгоритм оказывается более эффективным.

$\epsilon$ -субградиентные методы в какой-то степени являются промежуточными между методом ОГС и методами с растяжением пространства. К сожалению, они пока очень слабо исследованы экспериментально, а теоретически обоснованная

скорость сходимости и проведенные численные эксперименты пока не дают оснований для оптимизма [106].

В ближайшее время следует ожидать развития комбинированных методов минимизации выпуклых функций, основанных на схемах отсечения с растяжением пространства. На этом пути, по-видимому, можно получить алгоритмы, по эффективности не уступающие  $r$ -алгоритму, но имеющие теоретически гарантированную скорость сходимости. Большой интерес представляет проблема построения немонокотонных вариантов  $\varepsilon$ -субградиентных процессов, обладающих существенно ускоренной сходимостью по сравнению с методом ОГС и не требующих значительной дополнительной памяти.

Большое значение для решения задач большой размерности с применением двойной декомпозиции (по множеству ограничений и переменных одновременно) имеют эффективные алгоритмы нахождения седловых точек выпукло-вогнутых негладких функций. В связи с этим возникает проблема обобщения градиентных алгоритмов с преобразованием пространства на задачи нахождения седловых точек.

Решение этих проблем явится дополнительным стимулом для использования обобщенных градиентных методов минимизации негладких функций при решении сложных задач оптимального планирования, проектирования и исследования операций.

1. *Баженов Л. Г.* Об условиях сходимости методов минимизации почти дифференцируемых функций.— Кибернетика, 1972, № 4 с. 71—72.
2. *Бакаев А. А., Михалевиц В. С., Брановицкая С. В., Шор Н. З.* Методика и опыт решения сетевых транспортных задач большого объема на ЭЦВМ.— В кн.: Математические методы и проблемы производства. М., 1963, с. 247—257.
3. *Беляева Л. В., Журбенко Н. Г., Шор Н. З.* Об одной нелинейной транспортной задаче.— В кн.: Тр. семинара Науч. совета АН УССР по кибернетике «Теория оптим. решений». Киев, 1976, с. 83—89.
4. *Беляева Л. В., Журбенко Н. Г., Шор Н. З.* О методе решения одного класса динамических распределительных задач.— Экономика и мат. методы, 1978, 14, вып. 1, с. 137—146.
5. *Буземан Г.* Выпуклые поверхности.— М.: Наука, 1964,— 238 с.
6. *Вычислительные методы выбора оптимальных проектных решений/ В. С. Михалевиц, Н. З. Шор, Л. А. Галустова и др.*— Киев: Наук. думка, 1977.— 178 с.
7. *Гантмахер Ф. Р.* Теория матриц.— М.: Гостехиздат, 1953.— 492 с.
8. *Гельфанд И. М., Цетлин М. Л.* Принцип нелокального поиска в задачах автоматической оптимизации.— Докл. АН СССР, 1961, 137, № 2, с. 295—298.
9. *Гершович В. И.* Об опыте решения задачи распределения углей на коксование.— В кн.: Тр. семинара Науч. совета АН УССР по кибернетике «Теория оптим. решений». Киев, 1977, с. 22—35.
10. *Гласко В. Б., Гуцин Г. В., Старостенко В. И.* О применении метода регуляризации А. Н. Тихонова к решению нелинейных систем уравнений.— Журн. вычисл. математики и мат. физики, 1976, 16, № 2, с. 283—292.
11. *Гольштейн Е. Г.* Обобщенный градиентный метод отыскания седловых точек.— Экономика и мат. методы, 1972, 8, вып. 4, с. 569—579.
12. *Гольштейн Е. Г., Третьяков Н. В.* Градиентный метод минимизации и алгоритмы выпуклого программирования, связанные с модифицированными функциями Лагранжа.— Экономика и мат. методы, 1975, 11, вып. 4, с. 730—742.
13. *Гольштейн Е. Г., Юдин Д. Б.* Новые направления в линейном программировании.— М.: Сов. радио, 1966.— 524 с.

14. Демьянов В. Ф., Малоземов В. Н. Введение в минимакс.— М.: Наука, 1972.— 368 с.
15. Демьянов В. Ф. Методы экстремального базиса в минимаксных задачах.— Журн. вычисл. математики и мат. физики, 1977, 17, № 2, с. 512—517.
16. Дуб Дж. Л. Вероятностные процессы.— М.: Изд-во иностр. лит., 1956.— 606 с.
17. Еремин И. И. Итеративный метод для чебышевских приближений несовместных систем линейных неравенств.— Докл. АН СССР, 1962, 143, № 6, с. 1254—1256.
18. Еремин И. И. О методе «штрафов» в выпуклом программировании.— Кибернетика, 1967, № 4, с. 63—67.
19. Еремин И. И., Астафьев Н. Н. Введение в теорию линейного и выпуклого программирования.— М.: Наука, 1976.— 192 с.
20. Ермольев Ю. М. Методы решения нелинейных экстремальных задач.— Кибернетика, 1966, № 4, с. 1—17.
21. Ермольев Ю. М., Шор Н. З. О минимизации недифференцируемых функций.— Кибернетика, 1967, № 1, с. 101—102.
22. Ермольев Ю. М., Шор Н. З. Метод случайного поиска для задач двухэтапного стохастического программирования и его обобщение.— Кибернетика, 1968, № 1, с. 90—92.
23. Ермольев Ю. М. Методы стохастического программирования.— М.: Наука, 1976.— 240 с.
24. Журбенко Н. Г., Пинаев Е. Г., Шор Н. З., Юн Г. Н. Выбор состава и распределение гражданских самолетов по авиалиниям.— Кибернетика, 1976, № 4, с. 138—141.
25. Журбенко Н. Г. Исследование одного класса алгоритмов минимизации негладких функций и их применение к решению задач большой размерности: Автореф. дис. ... канд. физ.-мат. наук.— Киев, 1977.— 24 с.
26. Зойтендейк Г. Методы возможных направлений.— М.: Изд-во иностр. лит., 1963.— 175 с.
27. Канторович Л. В. Экономический расчет наилучшего использования ресурсов.— М.: Изд-во АН СССР, 1959.— 344 с.
28. Канторович Л. В. О методе наискорейшего спуска.— Докл. АН СССР, 1947, 56, вып. 3, с. 233—236.
29. Красносельский М. А. Топологические методы в теории нелинейных интегральных уравнений.— М.: Гостехиздат, 1956.— 392 с.
30. Кюнц Г. П., Крелле В. Нелинейное программирование.— М.: Сов. радио, 1965.— 301 с.
31. Левин А. Ю. Об одном алгоритме минимизации выпуклых функций.— Докл. АН СССР, 1965, 160, № 6, с. 1244—1247.
32. Любич Ю. И., Майстровский Г. Д. Общая теория релаксационных процессов для выпуклых функционалов.— Успехи мат. наук, 1970, № 1, с. 57—112.
33. Михалевич В. С., Ермольев Ю. М., Шкурба В. В., Шор Н. З. Сложные системы и решение экстремальных задач.— Кибернетика, 1967, № 5, с. 29—39.
34. Моисеев Н. Н. Численные методы в теории оптимальных систем.— М.: Наука, 1971.— 424 с.
35. Нурминский Е. А. Квазиградиентный метод решения задачи нелинейного программирования.— Кибернетика, 1973, № 1, с. 122—125.

36. *Нурминский Е. А.* О непрерывности  $\epsilon$ -субградиентных отображений.— Кибернетика, 1977, № 5, с. 148—149.
37. *Овруцкий И. Г., Шор Н. З.* Применение методов минимизации негладких функций для решения задач интерпретации гравиметрических наблюдений.— Кибернетика, 1976, № 2, с. 57—64.
- 38. *Поляк Б. Т.* Один общий метод решения экстремальных задач.— Докл. АН СССР, 1967, 174, № 1, с. 33—36.
39. *Поляк Б. Т.* Минимизация негладких функционалов.— Журн. вычисл. математики и мат. физики, 1969, 9, № 3, с. 509—521.
40. *Поляк Б. Т.* Метод сопряженных градиентов.— Тр. 11 Зим. школы по мат. программированию и смеж. вопр., 1969, вып. 1, с. 152—201.
41. *Примак М. Е.* О сходимости модифицированного метода чебышевских центров решения задачи выпуклого программирования.— Кибернетика, 1977, № 5, с. 100—102.
42. *Пшеничный Б. Н., Данилин Ю. М.* Численные методы в экстремальных задачах.— М.: Наука, 1975.— 319 с.
43. *Пшеничный Б. Н.* Необходимые условия экстремума.— М.: Наука, 1969.— 151 с.
44. *Рокафеллар Р.* Выпуклый анализ.— М.: Мир, 1973.— 470 с.
45. *Скоков В. А.* Замечание к методам минимизации, использующим операцию растяжения пространства.— Кибернетика, 1974, № 4, с. 115—117.
46. *Слободник С. Г.* Непрерывность и дифференциальные свойства функций: Автореф. дис. ... канд. физ.-мат. наук.— М., 1976.— 15 с.
47. *Тихонов А. Н., Гласко В. Б.* Применение метода регуляризации в нелинейных задачах.— Журн. вычисл. математики и мат. физики, 1965, 5, № 3, с. 463—473.
48. *Тихонов Н. А., Самарский А. А.* Уравнения математической физики.— М.: Гостехиздат, 1951.— 465 с.
49. *Фиакко А., Мак-Кормик Г.* Нелинейное программирование: Методы последоват. безусловной минимизации.— М.: Мир, 1972.— 240 с.
50. *Фаддеев Д. К., Фаддеева В. Н.* Вычислительные методы линейной алгебры.— М.: Физматгиз, 1960.— 656 с.
51. *Шабашова Л. П.* Градиентные методы решения нелинейных минимаксных задач: Автореф. дис. ... канд. физ.-мат. наук.— Днепрпетровск, 1973.— 22 с.
52. *Шпилов М. А.* О градиентных и штрафных методах в задачах математического программирования: Автореф. дис. ... канд. физ.-мат. наук.— М., 1974 — 18 с.
53. *Шпилов М. А.* О методе обобщенного градиента отыскания абсолютного минимума выпуклой функции.— Кибернетика, 1976, № 4, с. 52—57.
54. *Шор Н. З.* Применение метода градиентного спуска для решения сетевой транспортной задачи.— В кн.: Материалы науч. семинара по теорет. и прикл. вопр. кибернетики и исслед. операций: Науч. совет по кибернетике АН УССР. Киев, 1962, вып. 1, с. 9—17.
55. *Шор Н. З.* О структуре алгоритмов численного решения задач оптимального планирования и проектирования: Автореф. дис. ... канд. физ.-мат. наук.— Киев, 1964.— 10 с.
56. *Шор Н. З.* Применение обобщенного градиентного спуска в блочном программировании.— Кибернетика, 1967, № 3, с. 53—55.
57. *Шор Н. З., Росина Н. И.* Схема разложения задач линейного и выпуклого программирования и ее применение для решения задач

- планирования перевозок.— В кн.: Докл. I Всесоюз. конф. по оптимизации и моделированию транспорт. сетей. Киев, 1967, с. 225—237.
58. Шор Н. З., Горбач Г. И. Решение задач распределительного типа методом обобщенного градиентного спуска.— В кн.: Тр. семинара Науч. совета АН УССР по кибернетике «Теория оптим. решений». Киев, № 1, с. 59—71.
  59. Шор Н. З. О скорости сходимости обобщенного градиентного спуска.— Кибернетика, 1968, № 3, с. 98—99.
  60. Шор Н. З., Иванова Л. В. Об одном итеративном методе решения задач линейного программирования и матричных игр.— В кн.: Тр. семинара Науч. совета АН УССР по кибернетике «Теория оптим. решений», 1969, № 3, с. 22—30.
  61. Шор Н. З., Щепакин М. Б. Алгоритм решения двухэтапной задачи стохастического программирования.— Кибернетика, 1968, № 3, с. 56—58.
  62. Шор Н. З. Обобщенный градиентный спуск.— Тр. I Зим. школы по мат. программированию, 1969, вып. 3, с. 578—585.
  63. Шор Н. З., Билецкий В. И. Метод растяжения пространства для ускорения сходимости в задачах овражного типа.— В кн.: Тр. семинара Науч. совета АН УССР по кибернетике «Теория оптим. решений», Киев, 1969, № 2, с. 3—18.
  64. Шор Н. З. О скорости сходимости метода обобщенного градиентного спуска с растяжением пространства.— Кибернетика, 1970, № 2, с. 80—85.
  65. Шор Н. З. Методы минимизации недифференцируемых функций и их приложения: Автореф. дис. ... докт. физ.-мат. наук.— Киев, 1970.— 44 с.
  66. Шор Н. З. Использование операций растяжения пространства в задачах минимизации выпуклых функций.— Кибернетика, 1970, № 1, с. 6—12.
  67. Шор Н. З., Журбенко Н. Г. Метод минимизации, использующий операцию растяжения пространства в направлении разности двух последовательных градиентов.— Кибернетика, 1971, № 3, с. 51—59.
  68. Шор Н. З., Галушова Л. А., Момот А. И. Применение математических методов при оптимальном проектировании Единой газоснабжающей системы с учетом динамики ее развития.— Кибернетика, 1978, № 1, с. 69—74.
  69. Шор Н. З., Шабашова Л. П. О решении минимаксных задач методом обобщенного градиентного спуска с растяжением пространства.— Кибернетика, 1972, № 1, с. 82—88.
  70. Шор Н. З. О методе минимизации почти дифференцируемых функций.— Кибернетика, 1972, № 4, с. 65—70.
  71. Шор Н. З. Исследование сходимости метода градиентного типа с растяжением пространства в направлении разности двух последовательных градиентов.— Кибернетика, 1975, № 4, с. 48—53.
  72. Шор Н. З. Многоэтапное стохастическое выпуклое программирование.— В кн.: Тр. семинара Науч. совета АН УССР по кибернетике. «Теория оптим. решений», 1967, с. 48—58.
  73. Шор Н. З. Обобщенные градиентные методы минимизации негладких функций и их применение к задачам математического программирования. (Обзор).— Экономика и мат. методы, 1976, вып. 2, с. 337—356.

74. Шор Н. З. Метод отсечения с растяжением пространства для решения задач выпуклого программирования.— Кибернетика, 1977, № 1, с. 94—95.
75. Юдин Д. Б. Математические методы управления в условиях неполной информации.— М.: Сов. радио, 1974.— 400 с.
76. Юдин Д. Б., Немировский А. С. Информационная сложность и эффективные методы решения выпуклых экстремальных задач.— Экономика и мат. методы, 1976, вып. 2, с. 357—369.
77. Abadie J., Williams A. Dual and parametric methods in decomposition.— In: Recent advances in mathematical programming. New York: McGraw Hill, 1963.
78. Agmon D. The relaxation method for linear inequalities.— Can. J. Math., 1954, 6, N 2, p. 382—392.
79. Bertsekas D., Mitter S. A descent numerical method for optimization problems with nondifferentiable cost functionals.— SIAM J. Contr., 1973, 11, N 4, p. 637—652.
80. Bertsekas D. Necessary and sufficient conditions for a penalty method to be exact.— Math. Program., 9, N 1, 1975, p. 87—99.
81. Bertsekas D. Nondifferentiable Optimization via Approximation.— Math. Program., 1975, Study 3, p. 1—25.
82. Best M. A method to accelerate the rate of convergence of a class of optimization algorithms.— Math. Program., 1975, 9, N 1, p. 139—160.
83. Camerini P., Fratta L., Maffioli F. On improving relaxation methods by modified gradient techniques.— Math. Program., 1975, Study, 3, p. 26—34.
84. Clarke E. Necessary conditions for nonsmooth problems in optimal control and the calculus of variations: Diss. doctor philos. 1973.— Washington. 198 p.
85. Clarke E. Generalized gradients and applications.— Trans. Amer. Math. Soc., 1975, 205, N 2, p. 247—262.
86. Dantzig G., Wolfe P. Decomposition principles for linear programs.— Operat. res., 1960, 8, N 1, p. 101—111.
87. Elzinga J., Moore T. A central cutting plane algorithm for the convex programming problem.— Math. Program., 1975, 8, N 1, p. 134—145.
88. Evans J., Gould F., Tolle J. Exact penalty functions in nonlinear programming.— Math. Program., 1973, 4, N 1, p. 72—97.
89. Fisher M. Optimal solution of scheduling problems using Lagrange multipliers. Part. 1.— Operat. Res., 1973, 21, N 6, p. 1114—1127.
90. Fisher M., Northup W., Shapiro J. Using duality to solve discrete optimization problems: Theory and computational experience.— Math. Program., 1975, Study 3, p. 56—94.
91. Gerstenhaber M. Solution of large scale transportation problems.— In: Combinatorial analysis. New York, 1960.
92. Goldstein A. Optimization with corners.— In: Nonlinear programming. New York, 1975, p. 215—230.
93. Gorry G., Northup W., Shapiro J. Computational experience with a group integer programming algorithm. — Math. Program 1973, 4, № 2, p. 171—192.
94. Held M., Karp R. The traveling salesman problem and minimum spanning trees. Part. 2.— Math. Program., 1971, 1, N 1, p. 6—25.
95. Held M., Wolfe P., Crowder H. Validation of subgradient optimization.— Math. Program., 1974, 6, N 1, p. 62—88.

96. *Huang H.* Unified approach to quadratically convergent algorithms for function minimization.— *J. Optimizat. Theory and Appl.*, 1970, 5, N 6.
97. *Kelley J.* The 'Cutting plane' method for solving convex programs.— *SIAM Journal*, 1960, 8, N 4, p. 703—712.
98. *Lemarechal C.* An algorithm for minimizing convex functions.— In: *Proc. IFIP Cong.*— 74. Amsterdam, 1974, p. 552—556.
99. *Lemarechal C.* An extension of Davidon methods to nondifferentiable problems.— *Math. Program.*, 1975, Study 3, p. 95—100.
100. *Mathematical programming. Study 3. Nondifferentiable optimization*/Ed. M. L. Balinski and P. Wolfe.— Amsterdam: North — Holl and Publ. co. 1975.—178 p.
101. *Mifflin R.* Semismooth and semiconvex functions in constrained optimizations. RR—76—21. IIASA.— Laxenburg, Austria, 1976.— 23 p.
102. *Mifflin R.* An Algorithm for constrained optimization with semismooth functions. RR—77—3, IIASA.— Laxenburg, Austria, 1977.— 32 p.
103. *Motzkin T., Shoenberg I.* The relaxation method for linear inequalities.— *Can. J. Math.*, 1954, 6, N 2, p. 393—404.
104. *Rosen J.* Convex partition programming.— In: *Recent advances in mathematical programming.* New York: McGraw Hill, 1963.
105. *Sonnevend G.* On optimization of algorithms for function minimisation.— In: *Colloquia math. Soc. Janos Bolyai.* Eger, Hungary, 1974, p. 865—893.
106. *Wolfe P.* A method of conjugate subgradients for minimizing Nondifferentiable functions.— *Math. Program.*, 1975, Study 3, p. 145—173.

# НАУМ ЗУСЕЛЕВИЧ ШОР

## МЕТОДЫ МИНИМИЗАЦИИ НЕДИФФЕРЕНЦИРУЕМЫХ ФУНКЦИЙ И ИХ ПРИЛОЖЕНИЯ

*Печатается по постановлению ученого совета Института кибернетики АН УССР*

Редактор *Д. И. Попович*

Редактор-библиограф *Л. П. Шевченко*

Оформление художника *А. Л. Омельянюка*

Художественный редактор *И. П. Антонок*

Технический редактор *Г. Р. Боднер*

Корректоры *Л. П. Осмушниковая,*

*Л. Г. Бузишвили*

Информ. бланк № 1936.

Сдано в набор 14.06.78. Подп. в печ. 06.12.78. БФ 00384 Формат 84 × 108/32. Бумага типогр. № 1. Лит. гарн. Выс. печ. Усл. печ. л. 10,5. Уч.-изд. л. 10,15. Тираж 2200 экз. Заказ 8—1543. Цена 1 руб. 80 коп.

Издательство «Наукова думка» 252601. Киев. ГСП, Репина 3

Изготовлено Нестеровской городской типографией Львовского облполиграфиздата (г. Нестеров, ул. Горького, 8) с матриц Головного предприятия республиканского производственного объединения «Полиграфкнига» Госкомиздата УССР (г. Киев, Довженко, 3), зак. 5110.